

Electroencephalography Artifact Removal using Optimized Radial Basis Function Neural Networks

Shoorangiz Shams Shamsabad Farahani^{1*}, Mohammad Mahdi Arefi², Amir Hossein Zaeri³

1-Department of Electrical Engineering, Islamshahr Branch, Islamic Azad University, Islamshahr, Iran.

E-mail: shoorangiz_shams@yahoo.com, shams@iiu.ac.ir, (Corresponding Author)

2-Department of Power and Control Engineering, School of Electrical and Computer Engineering, Shiraz University, 71348-51154 Shiraz, Iran.

E-mail: arefi@shirazu.ac.ir

3-Department of Electrical Engineering, Shahin shahr Branch, Islamic Azad University, Shahin shahr, Iran.

Email: amzaeri@gmail.com

Received: June 2020

Revised: August 2020

Accepted: October 2020

ABSTRACT:

Electroencephalography (EEG) is a major clinical tool to diagnose, monitor and manage neurological disorders which is mostly affected by artifacts. Given the importance and the need for an automated method to remove artifacts, in this paper some intelligent automated methods are proposed which are composed of three main parts as extraction of effective input, filtering and filter optimization. Wavelet transform is utilized to extract the effective input, and the wavelet approximation coefficients are used as an effective input signal. In addition, Radial Basis Function Neural Network (RBFNN) has been used for filtering. The appropriate number of RBFs has been selected using extensive simulations, and the optimal value of spread parameter has been achieved by Bees algorithm (BA). Finally, the proposed artifact removal schemes have been evaluated on some real contaminated EEG signals in Mashad Ghaem hospital database. The results show that the proposed artifact removal schemes are able to effectively remove artifacts from EEG signals with little underlying brain signal distortion.

KEYWORDS: Artifacts, Bees Algorithm (BA), Electroencephalography, Optimization, Radial Basis Function Neural Network (RBFNN), Wavelet Transform (WT).

1. INTRODUCTION

EEG is a major tool which is used to record the electrical activity on human brain. It is recorded using 10-20 electrode placement system on the scalp of a person with low spatial and high temporal resolution [1]. EEG is mainly affected by different artifacts which reduce its clinical usefulness. There exist two types of artifacts, namely, biological and external, where the former is mostly caused by electromyogram (EMG) or muscular activities, Electrooculogram (EOG) or ocular activities and electrocardiogram (ECG) or cardiac activities, however, the latter is mostly caused due to technical factors as electrode leads and line-interference.

Most EEG artifact processing algorithms are developed to reduce biological artifacts, however, the effect of artifacts with technical origin is greatly reduced by improving technology and paying extra

attention to electrode attachments to body surface [2]. Artifact removal is the most appropriate approach introduced to identify and remove artifacts from brain signals and keep the related neurological phenomenon intact. There exist different schemes to remove EEG signal artifacts in clinical studies. Examples include linear filtering, linear combination and regression, Blind Source Separation (BSS), Principal Component Analysis (PCA), Wavelet Transform (WT), regression-based techniques, adaptive filters and neural networks.

Linear filtering is used to remove artifacts which are located in certain frequency bands and do not overlap with those of the neurological phenomena of interest [3]. In this regard, high-pass and low-pass filtering is utilized to remove EOG and EMG artifacts, respectively. In early clinical studies, linear filtering was commonly used to remove EEG signal artifacts [4]. Linear combination of EOG signal and EOG-contaminated EEG signal can be considered as one of the most common techniques to remove ocular artifacts

from EEG signals [5]. BSS techniques are used to separate the EEG signals from their components. The aforementioned techniques are used to identify the components which are attributed to artifacts. Moreover, the EEG signal is reconstructed without them [6].

Recently, great attention has been paid to Independent Component Analysis (ICA) which can be considered as one of the most widely used techniques among BSS techniques. In ICA, mixtures of independent source signals are blindly separated and the components are forced to be independent. In [7], eye-blink artifacts are automatically removed from the EEG data. In this scheme, the EEG raw data is decomposed into independent components by ICA followed by Peak Detection Algorithm of Independent Component which is suggested to identify eye-blink artifact components. In [8], ICA and Empirical Mode Decomposition (EMD) are combined where ICA is used to obtain the independent components, however, EMD is applied to remove the ocular artifacts with larger amplitude in independent components. In [9], a hybrid approach is presented which is based on ICA and Adaptive Noise Cancellation (ANC). It is noted that ICA decomposition is utilized to extract the artifact source signal.²

In PCA, data is transformed to a new coordinate system using the eigenvectors of signal covariance matrix. Afterwards, signal components are extracted by projecting the signal onto the eigenvectors. In [10], a wavelet-based threshold scheme and a PCA based adaptive threshold scheme are used to remove the ocular artifacts from EEG signals. In [11], ICA and PCA are applied to intracranial recordings, and three methods are proposed to remove the reference signal and line noise.

Wavelet Transform is a time-frequency analysis method which is suitable for non-stationary signals as EEG. It is based on different statistical characteristics of signal and noise [12]. In [13], a wavelet enhanced ICA method (WICA) is presented which uses wavelet thresholding to de-mixed independent components instead of the observed raw EEG. In [14], the wavelet-ICA and Support Vector Machine (SVM) are used to remove target artifacts. In this paper, SVM is used to identify artifactual components which are separated by wavelet-ICA. In [15], an automatic removal method is presented for Ocular Artifacts (OAs) to overcome EEG data interference. In this method, the Discrete Wavelet Transform (DWT) is applied to every recorded signal to achieve multiple scale coefficients.

There exist some regression-based methods including time and frequency domain [16-17]. These approaches involve calibration tests to determine the propagation coefficient between EOG channel and each

EEG channel. In this case, eye blink artifact is removed by subtracting each EEG channel from the separately recorded EOG. It is worth nothing that the EOG signal which is subtracted from the recorded EEG signal contains some EEG information which introduces loss in the desired information [16-17].

Recently, adaptive filters have achieved widespread applications in different areas. Because of the time varying nature of signals arising from the human body, adaptive filtering is suggested for EEG artifact removal and many new approaches are presented to design adaptive nonlinear filters for noise cancellation [18]. In [19], ocular artifacts are removed from EEGs using DWT and ANC. In [20], an adaptive filtering technique is presented to remove ocular artifacts from EEG recordings using the forgetting factor and the filter length parameters. In [21], the combination of EEG decomposition with Adaptive Filters (AFs) is addressed to improve the overall denoising process.

Neural networks are an attractive approach in adaptive signal processing [22]. In [23], ocular artifacts are removed using JADE Algorithm and neural networks. In this regard, independent components are gained using JADE, however, neural networks are used to classify the obtained independent components. Also, there exist two neural network schemes to learn classification rules from the EEG data. In [24], a combination of adaptive noise cancellation and adaptive signal enhancement is presented in a single recurrent neural network for the adaptive removal of ocular artifacts from EEG.

In this paper, an intelligent method is proposed which is based on RBFNN. It can remove artifacts from EEG signals and is composed of three main parts as extraction of effective input, filtering and filter optimization. Also, WT is used to extract the effective input, and wavelet approximation coefficients are utilized as effective input signal.

The rest of the paper is organized as follows: Section 2 briefly presents the related preliminaries, Section 3 details the proposed method based on optimized RBFNNs with wavelet approximation coefficients as input, afterwards in section 4, the performance of our approach is assessed by a simulation set up, thereafter, section 5 provides the conclusion of this paper.

2. PRELIMINARIES

In this section, concepts on RBFNNs and WT are presented.

2.1. Concepts on RBFNNs

In this subsection, the structure of a neural network is presented [25]. In total, there are three layers in RBFNNs, i.e. input layer, hidden layer and output layer.

The importance of RBFNNs is due to the following main features:

1. Universal Approximation Property
2. Learning Capability

The aforementioned features make the weights and adjustable parameters to constantly update and improve the total performance. The output of an RBFNN with n input and m neurons is as follows [25]:

$$y(t) = (W_{RBFNN}(t))^T \times H(X) \quad (1)$$

Where, $W_{RBFNN}(t)$ is the output layer weight matrix, $X = [x_1 \ x_2 \ \dots \ x_n]^T$ is the input vector and $H = [h_1 \ h_2 \ \dots \ h_m]^T$ is the vector including network activity functions which is generally defined as a Gaussian function in these networks. Assuming the predefined parameters of c_i and b_i as the average and standard deviation of the Gaussian function, respectively, the function $h_i(x)$ for each neuron is defined as

$$h_i(x) = e^{-\frac{\|x - c_i\|^2}{2b_i^2}} \quad (2)$$

In the next subsection, WT is briefly presented.

2.2. Concepts on WT

EEG signals can be appropriately analyzed by reducing the signal noise before any further analysis [26]. An efficient method for de-noising the EEG signal is WT which is used to solve the resolution problem by using multi-resolution analysis. It is a tool which cuts up data, functions or operators into different frequency components, and then studies each component with a resolution which is matched to the corresponding scale [27].

There exist two types of WT, i.e. Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). CWT is as follows [27]:

$$CWT(a,b) = \int_{-\infty}^{\infty} x(t) \varphi_{a,b}^{\nabla}(t) dt \quad (3)$$

Where, $x(t)$ is the original signal, ∇ is the complex conjugation and a and b are the scaling factors. Also, $\varphi_{a,b}^{\nabla}(t)$ is achieved by scaling the wavelet at time b and scale a . In CWT, the inner products of analyzing function and the original CWT signal is used and the similarity between these two functions is measured by integration. In CWT, it is assumed that a and b are continuously changing. The CWT drawback is that the wavelet coefficients shall be calculated for every possible scale which ends in a large amount of data. The aforementioned drawback can be overcome by

using DWT.

DWT is another type of WT which uses the mutually orthogonal set of wavelets defined by choosing the scaling and translation parameters (a and b). DWT is as follows [27]:

$$\varphi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \varphi\left(\frac{t-b}{a}\right) \quad (4)$$

Where, $\varphi_{a,b}(t)$ is the mother wavelet.

Remark1- In the current study, RBFNN is utilized because of its stability, good generalization ability, its easy design, online learning ability, and good input noise tolerance. Also, WT is used as an efficient method for de-noising the EEG signal.

3. THE PROPOSED METHOD BASED ON OPTIMIZED RBFNN WITH WAVELET APPROXIMATION COEFFICIENTS AS INPUT

In this paper, an intelligent method based on an RBFNN is presented to remove artifacts from EEG. The proposed scheme is composed of three main parts as extraction of effective input, filtering and filter optimization. WT is utilized to extract the effective input and wavelet approximation coefficients are used as an effective input signal. Afterwards, effective inputs are filtered for complete artifact removal. By applying WT, a two-level filtering system is achieved where the first level is wavelet approximation coefficients which remove some noise signals and the second level is the neural network which removes the remaining ones. RBFNNs have high potential in estimation and filtering. The number of RBFs and the spread parameter highly affect the performance of networks.

The Bees algorithm [28] is an optimization algorithm which is introduced in recent years and its ability to solve complex nonlinear problems has been proven. It is inspired by the natural foraging behavior of honey bees to find the optimal solution. It is a population-based search algorithm which mimics the food foraging behavior of swarms of honey bees. The basic version of BA performs a sort of neighborhood search which is combined with random search and can be used for both functional and combinatorial optimization. Bees algorithm shows remarkable robustness. It is able to converge to the maximum or minimum value without being trapped at local optima and it usually outperforms other methods in terms of optimization speed and accuracy [28]. In this paper, BA is used to find the optimal spread value. The number of RBFs is chosen the same as the training data, and the neural network is trained on the training data, so it may have poor performance on the test data. A solution for this problem is to construct an RBFNN with a single RBF. If an optimal solution is extracted from this

RBFNN, the network is built. Otherwise, the number of RBFs is added by one and the network is tested again. This process is continued until the final optimal solution is achieved. The Flowchart of the proposed method based on optimized RBFNN with wavelet approximation coefficients as input is given in Fig. 1.

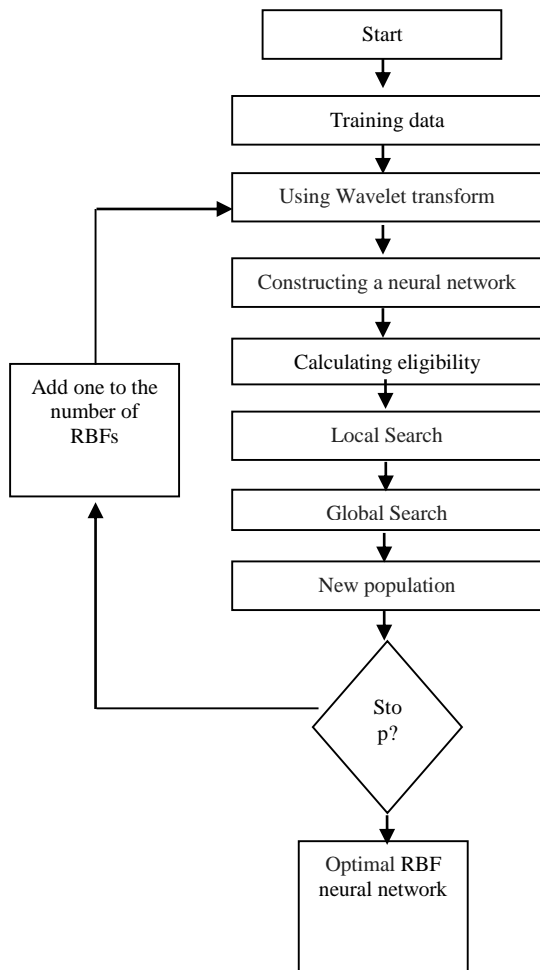


Fig. 1. Flowchart of the proposed scheme based on optimized RBFNN with wavelet approximation coefficients as input.

The algorithm of our scheme is as follows:

Step 1. Getting the EEG.

Step 2. Applying WT to EEG taken in step1.

Step 3. Constructing a neural network with an RBF.

Step 4. Using BA to find the optimal spread value for the constructed network regarding the following steps:

- **Setting BA parameters.** Setting BA parameters as the total number of bees, the number of elite bees and the number of scout bees.
- **Creating the initial population of bees.**
- **Calculating eligibility.** The mean square error (MSE) is used as eligibility criterion.

- **Local searching.** Scout bees must search around bees in elite groups one and two, until the final optimal solution is extracted.

- **Global searching.** Nonelite bees are distributed in the whole search area to prevent the optimization algorithm from being trapped in the local minimum.

- **Stop criterion.** If the stop criterion is reached, the algorithm stops and the network is built. Otherwise, go to step 4-C. In this algorithm, the stop criterion is chosen as the maximum number of iterations.

Step 5. If the network has an optimal performance, stop searching and the algorithm is finished. Otherwise, add one to the number of RBFs and go to step 4.

Continue this procedure until the number of RBFs becomes the same as the number of training data.

4. SIMULATION RESULTS

The proposed artifact removal schemes are evaluated on some real contaminated EEG signals in Mashad Ghaem hospital database with a sampling frequency of 256. There exist 3 artifact-free EEG data signal (baseline) and 355 artifact contaminating signal corresponding to 3 people. Also, three different artifacts are produced as eye blink, teeth grinding and swallowing. For each artifact, 61 time trial is recorded and the recording time is considered 10 seconds.

In our proposed methods, the RBFNN is used as a filter. In this scheme, artifact-contaminated signals and baseline signals are considered as neural network input and output, respectively, and the MSE criterion is used as a performance criterion for neural network. In the simulations, 355 artifact-contaminated signals are used. 70% of data is used to train the neural network while the remaining data is utilized for the testing purpose.

In the following, two tests without and with wavelet approximation coefficients as input are performed. In each test, our proposed schemes are evaluated based on four types of RBFNNs:

A. The proposed scheme based on precise non-optimized RBFNN

B. The proposed scheme based on precise optimized RBFNN

C. The proposed scheme based on non-optimized RBFNN

D. The proposed scheme based on optimized RBFNN

Test1. Using the original data (without wavelet approximation coefficients as input)

A. The proposed scheme based on precise non-optimized RBFNN

In this structure, the number of RBFs is chosen the same as the training data (250 RBFs are considered in this experiment). Also, the spread parameter value is obtained by trial and error. The results obtained using

this neural network are presented in Table 1.

Table 1. Results achieved using precise non-optimized RBFNN.

Spread value	Number of RBFs	MSE	Time (Sec)	Standard deviation
1	250	0.061	2	±0.01
2	250	0.056	2	±0.01
3	250	0.055	2	±0.01
4	250	0.072	2	±0.01
5	250	0.077	2	±0.01
6	250	0.087	2	±0.01
7	250	0.091	2	±0.01
8	250	0.093	2	±0.01
9	250	0.096	2	±0.01
10	250	0.096	2	±0.01

From Table 1, the impact of spread value on system performance becomes obvious and it can be concluded that it is essential to find a proper spread value. Fig. 2 shows the proposed scheme performance based on precise non-optimized RBFNN. It is clear that the performance is satisfactory.

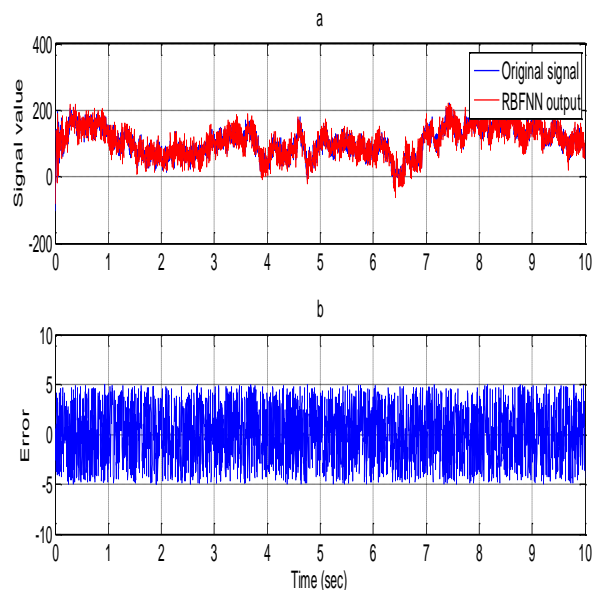


Fig. 2. Performance of the proposed scheme based on precise non-optimized RBFNN a) Comparison of the original signal and neural network prediction b) The neural network error.

B. The proposed scheme based on precise optimized RBFNN

In this structure, the number of RBFs is chosen the same as the training data (250 RBFs are considered in

this experiment). Also, the spread value is obtained using BA. The spread parameter highly impacts the network performance. Table 2 indicates the BA parameters.

Table 2. The BA parameters.

Total number of bees	20
Number of elite bees	8
Number of elite bees- group one	5
Number of elite bees- group two	3
Number of scout bees- group one	1
Number of scout bees- group two	1
The maximum number of iterations	100

The results obtained using this neural network are given in Table 3.

Table 3. Results obtained using precise optimized RBFNN.

Spread value	Number of RBFs	MSE	Time (Sec)	Standard deviation
3.11	250	0.038	2	±0.008

Comparing Tables 3 and 1, the impact of spread value can be realized. Using BA, the optimal spread 3.11 is obtained which leads to the breakdown of MSE from 0.055 to 0.038 in precise optimized RBFNN. Finally, the performance of the proposed scheme based on precise optimized RBFNN is shown in Fig. 3.

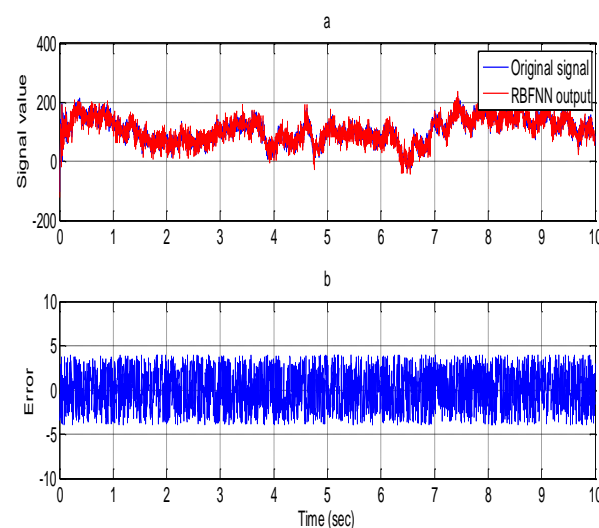


Fig. 3. Performance of the proposed scheme based on precise optimized RBFNN a) Comparison of the original signal and neural network prediction b) The neural network error.

C. The Proposed scheme based on non-optimized RBFNN

In this structure, the number of RBFs is optional and is set as an arbitrary value between 1 to 250, where 250 is considered as the total number of the training

data. The best value for the number of RBFs is chosen by testing all possible values. Furthermore, for any network with a given number of RBFs, a good spread value can lead to a satisfactory performance. So, a neural network with an RBF is constructed and the spread value which leads to the lowest MSE for the network is achieved using trial and error. If the result is not satisfactory, the number of RBFs is added by one and the network is tested again. It is recommended to increase the number of RBFs to get the best possible result. The best spread value is achieved by trial and error. The flowchart of the proposed method based on non-optimized RBFNN is presented in Fig. 4 and the results obtained using this neural network are given in Table 4.

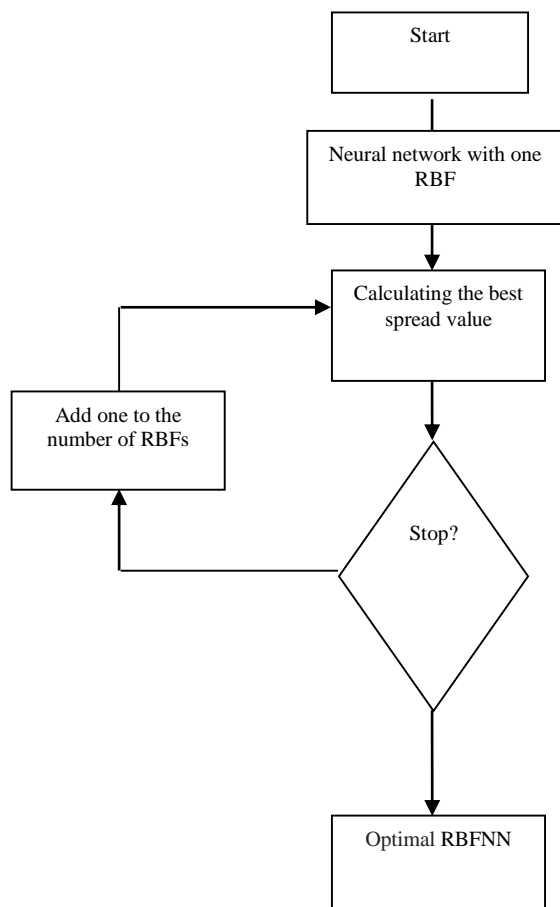


Fig. 4. Flowchart of the proposed method based on non-optimized RBFNN.

Table 4. Results obtained using non-optimized RBFNN.

Spread value	Number of RBFs	MSE	Time (Sec)	Standard deviation
3	200	0.031	1.5	±0.006

Comparing Tables 1, 3 and 4, the impact of choosing a suitable number of RBFs arbitrarily between 1 to 250 can be realized which leads to the breakdown of MSE from 0.055 and 0.038 in Tables 1 and 3, respectively, to 0.031 in Table 4. In Fig. 5, the proposed scheme performance is shown based on non-optimized RBFNN.

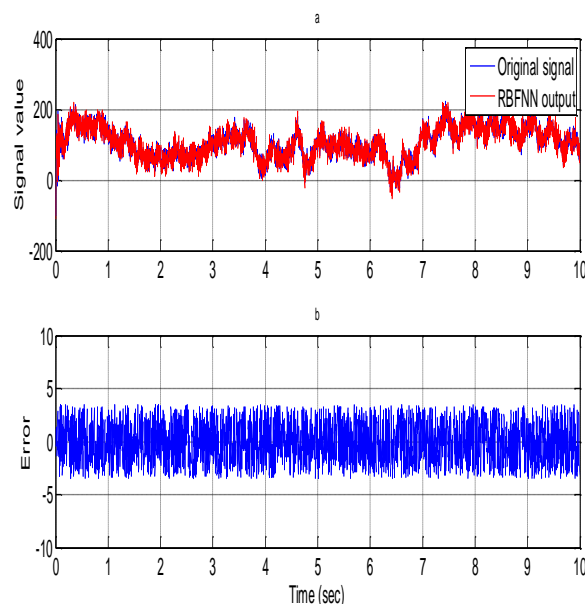


Fig. 5. Performance of the proposed scheme based on non-optimized RBFNN a) Comparison of the original signal and neural network prediction b) The neural network error.

D. The proposed scheme based on optimized RBFNN

In this structure, the number of RBFs is optional and is set as an arbitrary value between 1 to 250, where 250 is considered as the total number of the training data. Furthermore, for any network with a given number of RBFs, a good spread value can lead to a high performance. So, BA is used to find the optimal spread value in each step which enhances the neural network performance. The results obtained using this neural network are given in Table 5.

Table 5. Results obtained using optimized RBFNN.

Spread value	Number of RBFs	MSE	Time (Sec)	Standard deviation
3.42	200	0.01	1.5	±0.005

Comparing Tables 5, 4, 3 and 1, the impact of choosing a suitable number of RBFs arbitrarily between 1 to 250 and using BA to find the optimal spread value can be realized which leads to the

breakdown of MSE from 0.055, 0.038 and 0.031 in Tables 1, 3 and 4, respectively, to 0.01 in Table 5. Fig. 6 shows the proposed scheme performance based on optimized RBFNN.

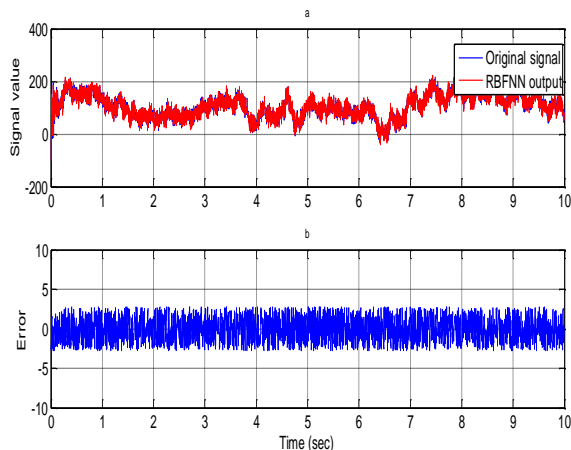


Fig. 6. Performance of the proposed scheme based on optimized RBFNN a) Comparison of the original signal and neural network prediction b) The neural network error.

Test2- Using wavelet approximation coefficients as input

In this experiment, the effect of wavelet approximation coefficients as the input of neural network is investigated. Similar to the previous test, the proposed schemes based on four types of RBFNNs are evaluated using wavelet approximation coefficients. Exhausted simulations are performed and wavelet approximation coefficients for contaminated signals with different artifacts as eye blink, teeth grinding and swallowing are achieved, however, the wavelet approximation coefficients for only two contaminated signals with eye blink and swallowing artifact are presented in Fig. 7 for simplicity. WT acts as a filter for the original signal and eliminates artifacts. In this scheme, wavelet approximation coefficients are used as an effective neural network input.

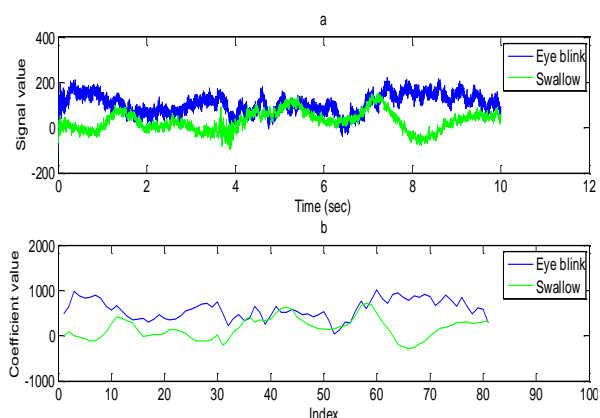


Fig. 7. a) Two contaminated signals with eye blink and swallowing artifact b) Wavelet approximation coefficients for the two signals at level five decomposition.

In Fig. 7-a, two artifact-contaminated signals are shown. It is clear that these two signals interfere with each other for many times because of artifacts. Using neural network learning algorithm, the main signal is achieved after artifact removal. However, because of the interference caused by artifacts, the incorrect training of neural network is possible which results in insufficient artifact removal. In Fig. 7-b, wavelet approximation coefficients at level five decomposition is shown for the two signals. Comparing Figs. 7-a and 7-b, it is clear that the interaction rate in Fig. 7-b is lower than Fig.7-a. It is noted that the higher level of decomposition does not mean better neural network performance. However, increasing the level of decomposition, renders more signal information to be disappeared and from a certain level, the neural network performance is degraded. Also, Haar wavelet is used in the simulations.

A. The proposed scheme based on precise non-optimized RBFNN with wavelet approximation coefficients as input

In this structure, the number of RBFs is chosen the same as the training data (250 RBFs are considered in this experiment). In addition, the spread value is obtained by trial and error. Table 6 shows the results obtained using this neural network.

Table 6. Results obtained using precise non-optimized RBFNN with wavelet approximation coefficients as input.

Decomposition level	Number of RBFs	Spread value	MSE	Time (Sec)	Standard deviation
1	250	3	0.043	1.5	±0.01
2	250	3	0.045	1.5	±0.01
3	250	2	0.048	1	±0.01
4	250	2	0.059	1	±0.01
5	250	2	0.087	1	±0.01

Table 6 shows that the best result is achieved at level one decomposition. Fig. 8 shows the proposed scheme performance based on precise non-optimized RBFNN with wavelet approximation coefficients as input.

From Table 6 and Fig. 8, it is clear that the RBFNN with wavelet approximation coefficients as input, proper number of RBFs and proper spread value at level one decomposition, renders a satisfactory

performance. Moreover, comparing Fig. 8 with Fig.2, it can be realized that error is reduced and the neural network performance is improved.

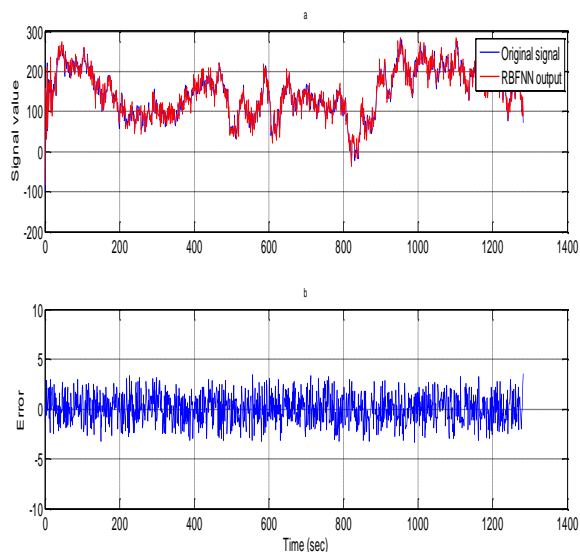


Fig. 8. Performance of the proposed scheme based on precise non-optimized RBFNN with wavelet approximation coefficients as input a) Comparison of the wavelet approximation coefficients of original signal and neural network prediction b) The neural network error.

B. The proposed scheme based on precise optimized RBFNN with wavelet approximation coefficients as input

In this structure, the number of RBFs is chosen the same as the training data (250 RBFs are considered in this experiment). In addition, the spread value is obtained using BA. Table 2 indicates the BA parameters. The results obtained using this neural network are given in Table 7.

Table 7. Results obtained using precise optimized RBFNN with wavelet approximation coefficients as input.

Decomposition level	Number of RBFs	Spread value	MSE	Time (Sec)	Standard deviation
1	250	2.93	0.029	1.5	±0.01
2	250	2.75	0.032	1.5	±0.01
3	250	2.43	0.041	1	±0.01
4	250	2.17	0.051	1	±0.01
5	250	2.13	0.065	1	±0.01

From Table 7, it is clear that the RBFNN with wavelet approximation coefficients as input, proper number of RBFs and proper spread value at level one decomposition, has better performance comparing with

other levels of decomposition. Comparing Tables 7 and 6, the impact of spread value can be realized. Using BA, the optimal spread value is 2.93 which ends to the breakdown of MSE from 0.043 to 0.029 at the first level of decomposition. Fig. 9 shows the proposed scheme performance based on precise optimized RBFNN with wavelet approximation coefficients as input.

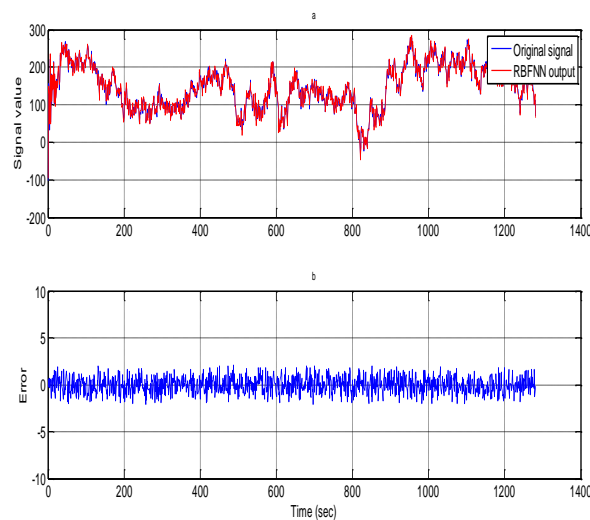


Fig. 9. Performance of the proposed method based on precise optimized RBFNN with wavelet approximation coefficients as input a) Comparison of the wavelet approximation coefficients of original signal and neural network prediction b) The neural network error.

Comparing Table 7 and Fig. 9 with Table 6 and Fig. 8, it can be realized that the former shows better performance comparing with the latter, since the spread value is achieved using BA in the former, however it is achieved by trial and error in the latter. Also, comparing Fig. 9 with Fig. 3, it can be realized that the error is reduced and the neural network performance is improved which is due to the use of wavelet approximation coefficients as input. In addition, the proposed scheme based on precise optimized RBFNN with wavelet approximation coefficients as input renders better performance comparing with the previous methods proposed in this paper.

C. The proposed scheme based on Non-optimized RBFNN with wavelet approximation coefficients as input

In this structure, the number of RBFs is optional and is set as an arbitrary value between 1 to 250, where 250 is considered as the total number of the training data. In order to find the best value for the number of RBFs, all possible values are tested and the best value is selected. Furthermore, for any network with a given number of RBFs, a good spread value ends in a

satisfactory performance. So, a neural network with an RBF is constructed and the spread value which leads to the lowest MSE for the network is achieved using trial and error. If the result is not satisfactory, the number of RBFs is added by one and the network is tested again. It is recommended to increase the number of RBFs to get the best possible result. Also, the best spread value is achieved by trial and error. Table 8 shows the results obtained using non-optimized RBFNN and using wavelet approximation coefficients as input.

Table 8. Results obtained using non-optimized RBFNN and using wavelet approximation coefficients as input.

Decomposition level	Number of RBFs	Spread value	MSE	Time (Sec)	Standard deviation
1	200	3	0.009	1.5	± 0.002
2	200	3	0.017	1.5	± 0.004
3	200	2	0.032	1	± 0.01
4	200	2	0.039	1	± 0.01
5	200	2	0.043	1	± 0.01

From Table 8, it is clear that the RBFNN with wavelet approximation coefficients as input, proper number of RBFs and proper spread value at level one decomposition has better performance comparing with other levels of decomposition. Comparing Tables 8, 7 and 6, the impact of choosing a suitable number of RBFs arbitrarily between 1 to 250 can be realized which leads to the breakdown of MSE from 0.043 and 0.029 in Tables 6 and 7, respectively, to 0.009 at the first level of decomposition in Table 8. Fig. 10 shows the performance of the proposed scheme based on non-optimized RBFNN with wavelet approximation coefficients as input.

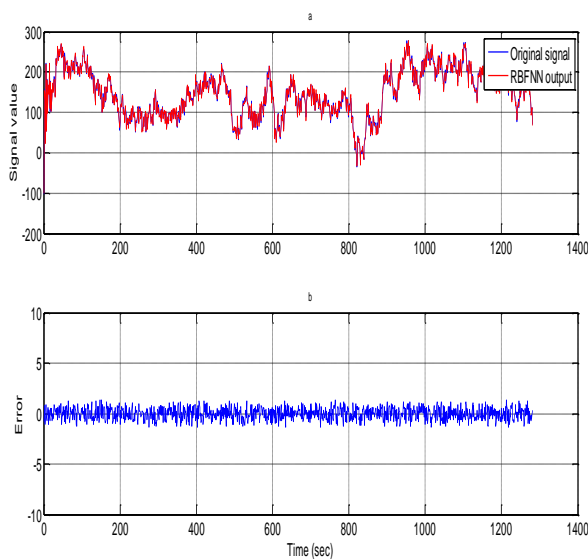


Fig. 10. Performance of the proposed scheme based on

non-optimized RBFNN with wavelet approximation coefficients as input a) Comparison of the wavelet approximation coefficients of original signal and neural network prediction b) The neural network error.

Comparing Table 8 and Fig. 10 with Table 7 and Fig. 9 and Table 6 and Fig. 8, it can be realized that the former shows better performance comparing with the latter which is due to the impact of choosing a suitable number of RBFs arbitrarily between 1 to 250. Also comparing Fig. 10 with Fig. 5, it can be realized that the error is reduced and the neural network performance is improved which is due to the use of wavelet approximation coefficients as input. Also, the proposed scheme based on non-optimized RBFNN with wavelet approximation coefficients as input renders better performance comparing with the previous methods proposed in this paper.

D. The proposed scheme based on Optimized RBFNN with wavelet approximation coefficients as input

In this structure, the number of RBFs is optional and is set as an arbitrary value between 1 to 250, where 250 is considered as the total number of the training data. Furthermore, for any network with a given number of RBFs, a good spread value can lead to high performance. So, BA is used to find the optimal spread value in each step which enhances the neural network performance. The flowchart of the proposed method based on optimized RBFNN with wavelet approximation coefficients as input is presented in Fig. 1. Table 9 shows the results obtained using the proposed method based on optimized RBFNN with wavelet approximation coefficients as input.

Table 9. Results obtained using optimized RBFNN with wavelet approximation coefficients as input.

Decomposition level	Number of RBFs	Spread value	MSE	Time (Sec)	Standard deviation
1	200	3	0.001	1.5	± 0.0009
2	200	3	0.008	1.5	± 0.001
3	200	2	0.015	1	± 0.009
4	200	2	0.024	1	± 0.009
5	200	2	0.036	1	± 0.01

From Table 9, it is clear that the RBFNN with wavelet approximation coefficients as input, proper number of RBFs and proper spread value at level one decomposition has better performance comparing with other levels of decomposition. Comparing Tables 9, 8,

7 and 6, the impact of choosing a suitable number of RBFs arbitrarily between 1 to 250 and using BA to find the optimal spread value can be realized which leads to the breakdown of MSE from 0.043, 0.029 and 0.009 in Tables 6, 7 and 8, respectively, to 0.001 at the first level of decomposition in Table 9. Fig. 11 shows the proposed method performance based on optimized RBFNN with wavelet approximation coefficients as input.

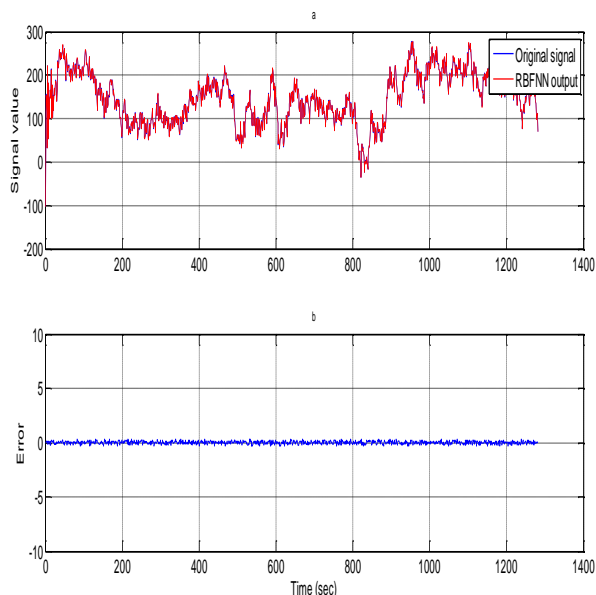


Fig.11. Performance of the proposed method based on optimized RBFNN with wavelet approximation coefficients as input a) Comparison of the wavelet approximation coefficients of original signal and neural network prediction b) The neural network error.

Comparing Table 9 and Fig.11 with Table 8 and Fig. 10, Table 7 and Fig. 9 and Table 6 and Fig. 8, it can be realized that the former shows better performance comparing with the latter which is due to the impact of choosing a suitable number of RBFs arbitrarily between 1 to 250 and using BA to find the optimal spread value. Comparing Fig. 11 with Fig. 6, it can be realized that the error is reduced and the neural network performance is improved which is due to the use of wavelet approximation coefficients as input. Also, the proposed scheme based on optimized RBFNN with wavelet approximation coefficients as input renders better performance in comparison with the previous methods proposed in this paper.

Performance evaluation of various wavelet types

In this subsection, the proposed method is evaluated using various wavelet types. The results with various wavelet types at different decomposition levels are presented in Table 10. Table 10 shows that Haar

wavelet at first level of decomposition performs better than other wavelet types.

Table 10. Performance evaluation and comparison of different wavelet types at different decomposition levels

Wavelet Type	Level1	Level2	Level3	Level4	Level5
Haar	0.001	0.008	0.015	0.024	0.036
Db2	0.002	0.009	0.018	0.029	0.041
Db3	0.002	0.01	0.017	0.032	0.043
Db4	0.002	0.009	0.019	0.028	0.045
Db5	0.003	0.008	0.02	0.027	0.044
Db6	0.002	0.008	0.016	0.031	0.051
Db7	0.003	0.008	0.017	0.034	0.055
Db8	0.004	0.009	0.018	0.034	0.046
Db9	0.003	0.018	0.023	0.026	0.041
Db10	0.004	0.01	0.019	0.029	0.043
Coif2	0.004	0.009	0.02	0.032	0.048
Coif3	0.003	0.008	0.016	0.028	0.044
Coif4	0.004	0.008	0.017	0.027	0.051
Coif5	0.003	0.008	0.018	0.031	0.051
Sym2	0.006	0.009	0.017	0.034	0.046
Sym3	0.002	0.011	0.019	0.034	0.041
Sym4	0.004	0.013	0.02	0.026	0.043
Sym5	0.002	0.009	0.016	0.028	0.045
Sym6	0.003	0.008	0.017	0.037	0.044
Sym7	0.002	0.008	0.018	0.029	0.054
Sym8	0.004	0.008	0.019	0.027	0.055
Bior1.3	0.004	0.008	0.019	0.031	0.046
Bior1.5	0.003	0.015	0.02	0.034	0.041
Bior2.2	0.004	0.01	0.018	0.034	0.043
Bior2.4	0.002	0.009	0.017	0.026	0.045
Bior2.8	0.003	0.008	0.018	0.029	0.042
Bior3.1	0.002	0.008	0.017	0.030	0.051
Bior3.3	0.003	0.008	0.019	0.028	0.055
Bior3.5	0.004	0.009	0.02	0.027	0.043
Bior3.7	0.003	0.011	0.016	0.031	0.041
Bior3.9	0.004	0.01	0.017	0.034	0.043
Bior4.4	0.002	0.009	0.017	0.034	0.045
Bior5.5	0.005	0.009	0.019	0.026	0.044
Bior6.8	0.004	0.014	0.02	0.025	0.040

5. CONCLUSION

In this study, novel artifact removal schemes based on RBFNNs are established. The proposed schemes are consist of three main parts as the neural network, the optimization and WT. In the first step, precise non-optimized RBFNN is used to remove artifacts where the number of RBFs is equal to the number of training data and the spread value is obtained by trial and error.

With this method, the neural network performance is satisfactory and the artifacts are almost removed. In the next step, BA is utilized to find the optimal spread value of precise RBFNN to enhance the performance. The simulation results show that using the optimization algorithms could significantly raise the performance of neural network. Afterwards, the number of RBFs are determined by trial and error to achieve better performance and less mean square error. Simulation results show that determining the proper number of RBFNNs and the use of BA can enhance the performance of neural networks. Next, WT is used to improve the RBFNN performance and reduce the computing time. Using wavelet approximation coefficients could significantly raise the performance of neural network. The impact of using wavelet approximation coefficients is shown by comparing the results with previous approaches where the original data is utilized as the neural network input. Finally, the optimized RBFNN with proper RBFs, optimal spread value and appropriate approximation wavelet input coefficients achieves the least mean square error among all schemes proposed in this paper.

REFERENCES

- [1] S. Sanei, J.A. Chambers, "EEG Signal processing", John Wiley & Sons, 2008.
- [2] L. Sornmo, P. Laguna, "Bioelectrical signal processing in cardiac and neurological applications", Academic Press, 2005.
- [3] J.R. Ives, D.L. Schomer, "A 6-pole filter for improving the readability of muscle contaminated EEGs," *ELECTROEN CLIN NEURO*, Vol. 69, pp. 486-490, 1988.
- [4] W. Zhou, J. Gotman, "Removing eye-movement artifacts from the EEG during the intracarotid amobarbital procedure," *Epilepsia*, Vol. 46, pp. 409-414, 2005.
- [5] R.J. Croft, R.J. Barry, "Removal of ocular artifact from the EEG: a review," *NEUROPHYSIOL CLIN*, Vol. 30, pp.5-19, 2000.
- [6] S. Choi, A. Cichocki, H.M. Park, S.Y. Lee, "Blind source separation and ICA techniques: a review," *Neural Information Proc essing-Letters and Reviews*, Vol. 6, pp. 1-57, 2005.
- [7] G. Junfeng, L. Pan, Y. Yong, W. Pei, "Automatic removal of eye-blink artifacts based on ICA and peak detection algorithm," presented at the *2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR)*, 2010.
- [8] L. Pengpai, C. Zhenxin, H. Yongmei, "A method for automatic removal of EOG artifacts from EEG based on ICA-EMD," presented at the *Chinese Automation Congress (CAC)*, 2017.
- [9] A. Jafarifarmand, M.A. Badamchizadeh, S. Khanmohammadi, M.A. Nazari, B. Mozaffari Tazehkand, "Real-time ocular artifacts removal of EEG data using a hybrid ICA-ANC approach," *Biomed. Signal Process. Control* Vol. 31, pp.199-210, 2017.
- [10] P.A. Babu, K.V. Prasad, "Removal of Ocular Artifacts from EEG Signals Using Adaptive Threshold PCA and Wavelet Transforms," presented at the *International Conference on Communication Systems and Network Technologies (IEEE 2011)*, pp. 572 – 575, 2011.
- [11] S. Hu , M. Stead , G.A. Worrell, "Removal of Scalp Reference Signal and Line Noise for Intracranial EEGs," presented at *IEEE International Conference on Networking, Sensing and Control (IEEE 2008)*, pp. 1486 – 1491, 2008.
- [12] M. Browne, T.R.H. Cutmore, "Low-probability event-detection and separation via statistical wavelet thresholding: an application to psychophysiological denoising," *CLIN NEUROPHYSIOL*, Vol. 113, pp. 1403-1411, 2002.
- [13] N.P. Castellanos, V.A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *J NEUROSCI METH*, Vol. 158, pp. 300-312, 2006.
- [14] S. Chong Yeh, M. Norrima, A. Hamzah, C. Paul, I. Masahiro, "Automated Classification and Removal of EEG Artifacts With SVM and Wavelet-ICA," presented at *IEEE Journal of Biomedical and Health Informatics*, Vol. 22, pp. 664-670, 2018.
- [15] M.Li, Y. Cui, J. Yang, "Automatic removal of ocular artifact from EEG with DWT and ICA Method," *Applied Mathematics & Information Sciences*, Vol. 7, pp. 809-816, 2013.
- [16] G.L. Wallstrom, "Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods," *RADIOTHER ONCOL*, Vol. 53, pp.105-119, 2004.
- [17] M. Fatourech, A. Bashashati, R.K. Ward, G.E. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey," *CLIN NEUROPHYSIOL*, Vol.118, pp. 480-494, 2007.
- [18] Z. Li, M.J. Er, "An online self-improved fuzzy filter and its applications," presented at the *2nd WSEAS International Conference on Electronics, Control and Signal Processing*, 2003.
- [19] H. Peng, B. Hu, Q. Shi, M. Ratcliffe, Q. Zhao, Y. Qi, G. Gao, "Removal of Ocular Artifacts in EEG-An Improved Approach Combining DWT and ANC for Portable Applications," *IEEE J. Biomed. Health. Inf.* Vol.17, pp. 600 – 607, 2013.
- [20] P. He, G. Wilson, C. Russell, M. Gerschutz, "Removal of ocular artifacts from the EEG: A comparison between time-domain regression method and adaptive filtering method using simulated data," *MED BIOL ENG COMPUT* Vol. 45, pp. 495-503, 2007.
- [21] X. Navarro, F. Porée, A. Beuchée, G. Carrault, "Denoising preterm EEG by signal decomposition and adaptive filtering: A comparative study," *MED ENG PHYS*, Vol.37, pp. 315–320, 2015.
- [22] YP. Tan, KH. Yap, L. Wang, "Intelligent Multimedia Processing with Soft Computing", Springer-Verlag Heidelberg, 2004.
- [23] V. Krishnaveni, S. Jayaraman, A. Gunasekaran, K. Ramadoss, "Automatic Removal of Ocular Artifacts using JADE Algorithm and Neural

- Network,” *Int. J. Biomed. Sci.*, Vol.2, pp. 322-333, 2009.
- [24] S. Selvan, R. Srinivasan, “**Removal of ocular artifacts from EEG using an efficient neural network based adaptive filtering technique,**” *IEEE SIGNAL PROC LET*, Vol. 6, pp. 330 – 332, 1999.
- [25] RJ. Howlett, CJ. Lakhmi, “**Radial Basis Function Networks, Recent Developments in Theory and Applications**”, *Springer*, 2001.
- [26] SS. Patil, MK. Pawar, “**Quality advancement of EEG by wavelet denoising for biomedical analysis,**” presented at the *International Conference on Communication, Information and Computing Technology, ICCICT* 2012.
- [27] A. Federico, GH. Kaufmann, “**Wavelet Transform,**” Vol. 34, pp. 2336–2338, 2009.
- [28] DT. Pham, A. Ghanbarzadeh, E. Koç, S. Otri, S. Rahim, M. Zaidi, “**The Bees Algorithm - A Novel Tool for Complex Optimisation Problems,**” presented at the *2nd Virtual International Conference on Intelligent Production Machines and Systems*, pp. 454-459, 2006.