Finding the discriminative frequencies of motor electroencephalography signal using genetic algorithm

Shaima Miqdad Mohamed Najeeb[1], Haider Th. Salim Al Rikabi[2], Shaima Mohammed Ali[3]
[1,3]Department of Computer Technology Engineering, Northen Technical University, Iraq
[2]Department of Electrical Engineering, College of Engineering, Wasit University, Iraq

Article Info	ABSTRACT
	A crucial part of the brain-computer interface is a classification of electroencephalography (EEG) motor tasks. Artifacts such as eye and muscle movements corrupt EEG signal and reduce the classification performance. Many studies try to extract not redundant and discriminative features from EEG signals. Therefore, this study proposed a signal preprocessing and feature extraction method for EEG classification. It consists of removing the artifacts by using discrete fourier transform (DFT) as an ideal filter for specific frequencies. It also cross-correlates the EEG channels with the effective channels to emphases the EEG motor signals. Then the resultant from cross correlation are statistical calculated to extract feature for classifying a left and right finger movements using support vector machine (SVM). The genetic algorithm was applied to find the discriminative frequencies of DFT for the two EEG classes signal. The performance of the proposed method was determined by finger movement classification of 13 subjects and the experiments show that the average accuracy is above 93 percent.

Corresponding Author:

Haider Th. Salim Al rikabi
Department of Electrical Engineering
Wasit University, Iraq
Email: hdhiyab@uowasit.edu.iq

1. INTRODUCTION

The most complex organ in the human body is the human brain. The basic units of the brain cells called neurons, which is considered the center of the human nervous system and controls different organs and functions. Neurons send electrical signals to control the human body and can be measured using electroencephalography (EEG), which measures the electrical activity of the brain by recording it via electrodes placed either on the cortex or the scalp. The signal generated by this electrical activity is non-stationary and complex random signals [1, 2]. The EEG signal contains a lot of information about the human brain functions, so the EEG analysis and information extraction are very complicated. Since the EEG signal consists of the very low-frequency components, so it is corrupted with different types of artifacts (noises and power line frequencies) [3-5].

In recent years, the amount of researches and efforts have been directed towards the identification and utilization of the information from the human EEG signal. Most of the work in brain computer interface (BCI) literature on motor imagery has been towards classifying movements of the hand, foot, and tongue. These movements are large and topographically different corresponding to the brain areas.

Farid Ghani, *et al.,* classified different types of EEG data movements. They used discrete cosine transform (DCT) and independent component analysis (ICA) to reduce the number of the extracted features

and to improve the accuracy of classification [6]. There are multiple studies related to classifying EEGs into some categories like detecting normal, interictal, and epileptic signals [7]. Mohammad H. Alomari, *et al.,* obtained pretty good classification results using neural networks (NNs) and support vector machine (SVM) to discriminate between EEG right and left-hand movement after applying band pass filter (BPF) with a specific set of statistical features (mean, power and energy) [8].

R. Zarei, *et al.,* proposed a method to remove the artifacts from EEG data based on Principal component analysis (PCA) and the cross-covariance technique (CCOV) for the extraction of discriminatory mental information states from EEG signals in BCI applications [9]. Shakshi, *et al.,* removed the unwanted frequency components from the original signal by using different types of filters. Mean, skewness, standard deviation, and variance are used to extract features from the EEG signal. The information about the signal was determined with the help of different efficient DSP tools like discrete fourier transform (DFT), fast fourier transform (FFT), short-time fourier transform (STFT), and wavelet transform [7].

From the foregoing, it becomes clear that feature extraction plays an important and influential role to help the classifier for distinguishing between EEG signal classes. Therefore, the main goal of this study is to find the most related features that discriminate EEG real finger movement signal and uses the SVM classifier only as a tool to distinguish the EEG signals based on the extracted features. The genetic algorithm was employed to find the most relevant frequencies which are used as cutoff frequency of ideal filter based on DFT. Finding these frequencies improves the classification performance in terms of both accuracy and computational time. The organization of this article is: section 2 will describe the main materials used in this work. Section 3 demonstrates the proposed method. Section 4 lists and explains the classification performance. The last section will discuss and explain the effects of each stage in the proposed method.

2.   MATERIALS AND METHODOLOGY

This section covers the procedure used for solving the problem related to find the discriminative frequencies of the EEG signal. Hence, it describes the proposed method and the tools used in this article such as FFT, and cross correlation. It also describes the procedure to acquire the EEG signal.

2.1. Proposed method

This study proposes a robust scheme that consists of five stages. Figure 1 illustrates the block diagram of the proposed method. These five stages are:

− Preprocessing using FFT: this stage uses DFT as an ideal filter to filter the most discriminative EEG frequencies. The most discriminative frequencies are determined by using genetic algorithm (GA). Then, the EEG signals are reconstructed using discrete fourier transform (IDFT).

− Cross correlation of the effective channel with right/left hemisphere: The brain is divided into 2 halves, or hemispheres, that are connected by the corpus callosum. Information from both hemispheres needs to be efficiently integrated; placing electrodes (EEG channels) on the scalp are split into two groups as the right/left hemisphere. Depending on the anatomical location of the signal generated in the brain or the channels close to the motor EEG signal region, the effective channel was selected so, the right hemisphere channels are cross correlated with the F4 channel and the left hemisphere channels with F3. This is done for whole training and testing sets. Cross correlation makes a more visible magnitude difference between the two hemispheres.

− EEG feature extraction: significant and important features need to be extracted from the EEG raw data. In this study, ten statistical features are computed from the EEG data (min, max, mean, mode, median, std, range, entropy, 1st quartile, and 3rd quartile). This is done for whole training and testing sets.

− Normalization: the current study explores the application of normalized EEG data to detect and identify the patterns of information flow in the functional brain networks. It makes the EEG signal lie between 1 and -1 by dividing each channel by the maximum absolute value of the same channel.

− SVM classification: radial base Kernel function with auto kernel scale are the configuration of the SVM classifier. Ten-fold cross validation was used to evaluate the performance of the classifier.

2.2. DFT

Representation of the digital signals in the time domain describes the signal amplitude versus the sample number. Some applications, signal in the frequency domain contains more useful information than the signal in a time domain. The transformation between time-domain signal samples and frequency domain components vice versa known as the DFT and IDFT respectively. Figure 2 shows the DFT application.

In addition, the DFT is widely used in many other areas, including spectral analysis, acoustics, imaging/video, audio, instrumentation, and communications systems [10]. The DFT and IDFT equations are respectively shown below:

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-\frac{2j\pi kn}{N}} \text{ [10]} \tag{1}$$

where: $0 \leq n \leq N-1$

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k]e^{\frac{j2\pi kn}{N}} \text{ [10]} \tag{2}$$
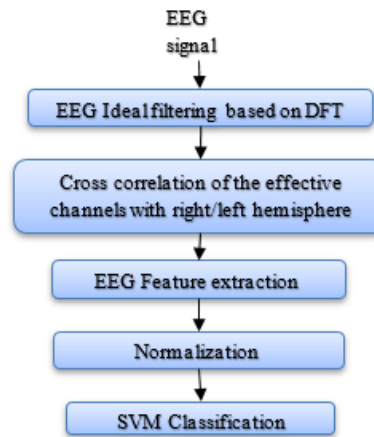

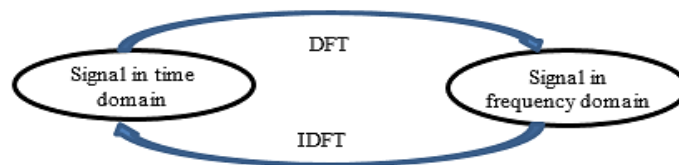Figure 1. The proposed method

Figure 2. DFT application

2.3. Cross correlation

The correlation of signals is a signal-processing technique often used for measuring the similarity between two signals and results in a cross-correlation sequence. Basic statistic parameters can be taken from the cross-correlation sequence as features of a signal and then used in classification. Correlation is also used for the detection of targets in radar or sonar signal. The sample of cross-correlation between two signals is calculated by:

$$R_{xy}[m] = \sum_{i=0}^{N-|m|-1} x[i]y[i-m] \text{ [11]} \tag{3}$$

where $R_{xy}[m]$ is the cross-correlation at m lag and $m = [-(N-1), \ldots, 0,1,2 \ldots, (N-1)]$. The samples of cross correlation for two sequences has 2*N*-1 sample length, each of the signals, *x* and *y*, consists of *N* finite number of samples [11].

2.4. EEG feature extraction

Feature extraction plays an important role in the process of classifying EEG signals. A training process will take place properly if features that describing the signal are extracted well [12, 13]. Many feature extraction algorithms are presented in the biomedical field, the simplest and most common algorithm that works to reduce the amount of data classified for EEG signal is the use of statistical approaches such as mean, median, mode, and standard deviation [14, 15].

2.5. Classification method

One of the most popular machine learning techniques is SVM. It is a statistical learning theory based on the classification method [16, 17]. SVM is applied in many applications like EEG signal classification,

cancer identification, bioinformatics, seizure prediction, face recognition, and speech disorder. The principle of SVM classification is to construct an optimal hyperplane as the decision surface to separate the training data and tries to find the nearest support vectors to that hyperplane with the minimal error of classification and maximal margin simultaneously to solve an optimization problem. The essential element in SVM is the kernel function, which maps samples in one feature space to another feature space. Radial kernel function (RBF), linear kernel function, polynomial kernel function, and gaussian function are some of the popular Kernel functions [18, 19]. The operation that takes data as input and transforms it into the required form is the function of the SVM kernel. The classification accuracy of SVM largely depends on the selection of the kernel function parameters [18].

2.6. GA

A GA is one of the heuristic methods for randomizing search and solving the optimization problems .Many different research fields used GA, genetic algorithms can be used for feature selection [20 , 21]. In GA, the chromosome is a possible solution vector, which consists of a set of genes. In the solution space, a set of chromosomes called population. The general scheme of the classic genetic algorithm as shown in Figure 3.

First, define an initial population of N chromosomes each of length L. Each chromosome in the population is then evaluated using a fitness function. Chromosomes are selected to be parents and recombine to reproduce new offspring. For a particular chromosome, a probability of selection parents should depend on the fitness function. The selection probability would be:

$$p_s = f(xi) \sum_{i=1}^{N} f(xi) \qquad (4)[23]$$

where xi represents the i-th chromosome in the population and $f(x_i)$ its fitness.

In the crossover operation, parents are selected for merging together and produced new children. Mutation consists of randomly altering genes inside chromosomes, with a very low probability. This leads the GA to escape converging towards local optima. The previous population is then replaced with a new population. Three GA operations (selection, crossover, and mutation) are iteratively applied until some stopping criterion is met or a predefined maximum number of iterations is reached. In order to obtain faster convergence towards the optimal solution and mitigate the risk of losing the best chromosome by crossover or mutation, a variation of the basic GA is introduced to improve its performance, by applying elitism, which consists of preserving the fittest chromosome in a population for the next generation [23].
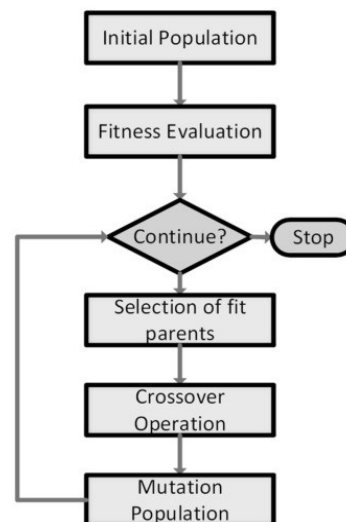

Figure 3. steps of the genetic algorithm [22]

2.7. EEG dataset acquisition

EEG raw signal from the user scalp is collected, amplified, digitized and transmitted through a Bluetooth module to the personal computer using EMOTIV EPOC headset with a sampling rate of 128 bps. EMOTIV headset measures EEG signal from 14 locations positioned at: AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, and O2 as shown in Figure 4 [24].

Thirteen subjects performed real right/left finger movements. The subjects sat in a comfortable chair wearing the headset with closed eyes. In each session, the subject was informed in advance which hand to move. Auditory stimuli were used to notify the on-action period of the subject finger movement. The duration of each movement was six seconds while the rest periods in between had different durations. This process was performed four times in each session and separated by resting periods durations. The duration of each movement was six seconds while the rest periods in between had different lengths. This process was performed four times in each session and separated by a resting period durations [25].
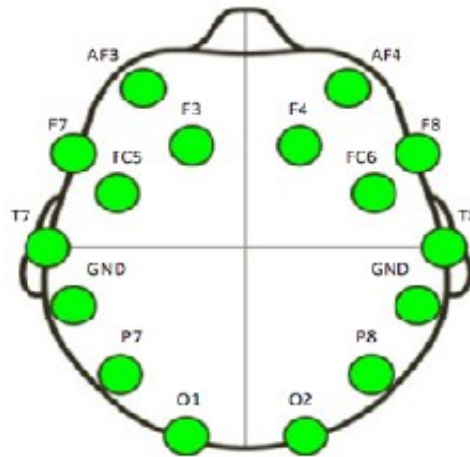

Figure 4. Emotiv EPOC electrode placement [24]

3. RESULTS AND DISCUSSIONS

Classification motor movements from the EEG signal faces a lot of difficulties, one of them is artifacts removal. Since motor signals are embedded among human body artifacts like eye movement eye blink, and internal organs signals. Motor signals are also suffered from external artifacts like bad electrode placement, environment sounds. Bad electrode placement adds a different ratio of noises to each electrode depending on the scalp connectivity with the electrode. Therefore, preprocessing is needed which tries to get rid of these artifacts and extracts the EEG motor signals. One of the most popular methods is filtering but the frequencies of the motor signals are unknown.

Since GA is used to search for these frequencies (motor discriminative frequencies). The proposed method is used as a fitness function of GA to search for the discriminative frequencies of only two subjects (subjects 2 and 6). This operation is done using the mentioned subjects in order not to fall into local optima. The population size was chosen as 20 since the diversity is ensured and to reduce the harmful effects of the mutation operator. If the size of the population is too small, this leads to the negative impact of the genetic algorithm by the mutation operator, and conversely, the latency time of the GA will increase. Therefore, the population size is chosen experimentally. The GA Twenty GA iterations were performed to explore the frequencies between 0-64HZ and it found only 27 frequencies are the most discriminative frequencies. These frequencies are 6, 7, 9-15, 18, 19, 23, 24, 27, 28, 33, 37, 39, 44-46, 48, 50, 51, 53, 59 and 64. Figure 5 shows the best and worst cost values of only two subjects during GA search.

The proposed method is applied to classify the movement of the thirteen subjects using the specified frequencies. Figure 6 illustrates the classification performance using the proposed method. The reliability of the preprocessing method (proposed method) is obviously clear and this shown with the impact performance that has an impact range of 90-100%. This is for ten subjects out of thirteen subjects. only one subject has a relatively not good impact above 70% and the rest (two subjects) have an impact range of 85-89%. The following equation is used to evaluate the classification rate:

$$Classification\ rate = \frac{number\ of\ correct\ predictions}{total\ number\ of\ predictions} x100\% \qquad (5)$$

The second stage in the proposed method (cross-correlation stage) tries to enlarge the difference between the two brain hemispheres. Figure 7 illustrates the effects of cross-correlation. The difference between the right and left finger movements is shown in Figures 7 (a) and (b) and this difference isn't obviously clear.

Meanwhile, after using cross-correlation, the difference between the two EEG signals becomes extremely clear as shown in Figures 7 (c) and (d).
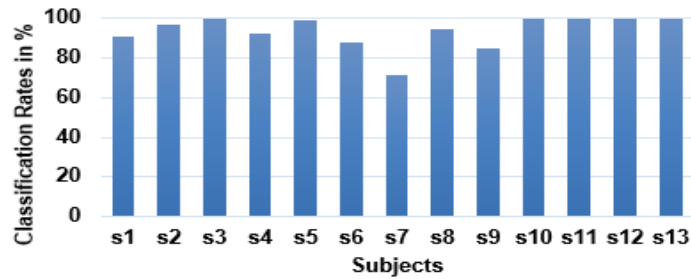

Figure 5. The best and the worst cost values of GA

Figure 6. The classification rates of 13 subject

(a)   (b)

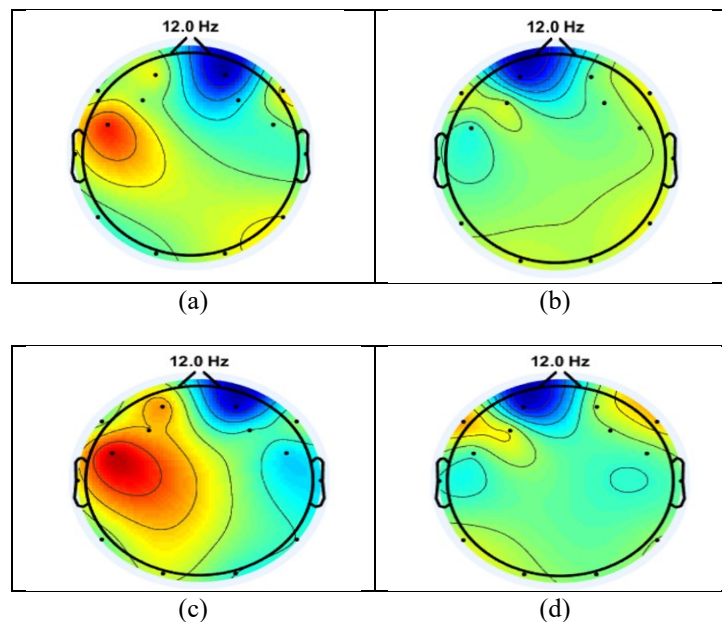(c)   (d)

Figure 7. EEG topography after and before cross-correlation effects; (a) and (b) original EEG right/left finger movements respectively, (c) and (d) same EEG signal after cross-correlation

The statistical parameter stage reduces the feature space extracted from the EEG signal which reflect to the computational time and it also filter out the unnecessary and redundant features. Table 1 illustrates the

amount of features reduction after using the ten features statistical calculation. In this table, ten statistical features produce 140 features (14 channels x 10 statistical features) and the amount of data represents the fed data before and after cross-correlation. In 5 estimate the amount of data reduction. Therefore, the amount of data fed to the classifier are reduced to 7.8% and 3.9% before and after cross correlation stage respectively.

$$Data\ reduction = \frac{No.of\ feature}{No.\ of\ input\ data}\ x100\% \tag{6}$$

Tabel 1. The data reduction percentages after and before feature extraction

No. of input data	Data Amount	No. of features	Data reduction
Original data	14 channels x 128 samples	140	7.8%
Cross correlated data	14 channels x 255 samples	140	3.9%

4. CONCLUSIONS

The paper presents the proposed method for preprocessing and extracting features from EEG real motor movements. It employs less complex tools like DFT and cross-correlation unlike using ICA or PCA mentioned in section one of some researches. The proposed method proves its effectiveness even with EEG signals acquired by gamming acquisition equipment (EMOTIV EPOC+), see Figure 6. Hence, the performance of the proposed method has good performance for thirteen subjects so that it proves that GA, which applied on two EEG subject signals, doesn't fall into local optima. The second stage of the proposed method enlarges the difference between the two EEG classes. We can clearly see that the Figure 7 meanwhile utilizing the statistic methods to reduce the amount of the processed features fed to the classifieras it is shown in the Table 1.

REFERENCES

[1] A. Ridouh, D. Boutana and M. Benidir, "Comparative Study of Time Frequency Analysis Application on Abnormal EEG Signals," *Recent Advances in Electrical Engineering and Control Applications, Springer,* pp. 355-368, 2017.

[2] H. Fauzi, A. M. Azzam, I. M. Shapiai, M. Kyoso, U. Khairuddin and T. Komura, "Energy extraction method for EEG channel selection," *TELKOMNIKA Telecommunication Computing Electronics and Control,* vol. 17, no. 5, pp. 2561-2571, 2019.

[3] Shakshi and J. Ramavtar, "Brain Wave Classification and Feature Extraction of EEG Signal by Using FFT on Lab View," *International Research Journal of Engineering and Technology (IRJET),* vol. 7, no. 3, pp. 1208-1212, 2016.

[4] S. M. Islam, M. A. El-Hajj, H. Alawieh, Z. Dawy and N. Abbas, "EEG mobility artifact removal for ambulatory epileptic seizure prediction applications," *Biomedical Signal Processing and Control, Elsevier,* vol. 55, 2020.

[5] A. Bisht, C. Kaur and P. Singh, "Recent advances in artifacts removal techniques for EEG signal processing," *Intellegent communication, control and devices, advances in intelligent and computing,* vol. 989, pp. 385-392, 2020.

[6] G. Farid, S. Hameedah, A. Dilnashin, F. Omar and U. K. Yusuf, " Classification of Wrist Movements Using EEG Signals," *Journal of Next Generation Information Technology(JNIT),* vol. 4, October 2013.

[7] V. M. M. Andrius, M. Tadas and S. Ruta, "Algorithm for automatic EEG classification according to the epilepsy type: Benign focal childhood epilepsy and structural focal epilepsy," *Biomedical Signal Processing and Control,* vol. 48, pp. 118-127, 2019.

[8] H. A. Mohammad, S. Aya and A. Khaled, "Automated Classification of L/R Hand Movement EEG Signals using Advanced Feature Extraction and Machine Learning," *International Journal of Advanced Computer Science and Applications,* vol. 4, no. 6, pp. 207-212, 2013.

[9] R. Zarei, J. He, S. Siuly and Y. Zhang, "A PCA aided cross-covariance scheme for discriminative feature extraction from EEG signals," *Computer Methods and Programs in Biomedicine, Elsevier,* vol. 146, pp. 47-57, 2017.

[10] S. Parthana, P. Tripathi, P. S. Manash and K. S. Kandarpa, "Pre-processing and Feature Extraction Techniques for EEG BCI Applications," *Journal of Engineering Technology ,* vol. 5, 2016.

[11] I. T. S. D. Hari Krishna, "Autonomuos Robot Control based on EEG and Cross-correlation," *IEEE Conferences,* pp. 1-4, 2016.

[12] H. H, M. A and S. S., "Feature Extraction ElectroEncephaloGram (EEG) using wavelet transform for cursor movement," *3rd Annual Applied Science and Engineering Conference,* vol. 434, no. 1, pp. 1-8, 2018.

[13] A. Narin, Y. Isler and M. Ozer, "Effect of Feature Selection by Genetic Algorithm on Early Prediction Performance of PAF Attack," *Innovations in Intelligent Systems and Applications Conference,* 2018.

[14] C. C. Lung, S. Rubita and S. H. Siti, "Feature extraction of eeg signal using wavelet transform for autism classificatvol," *Journal of Engineering and Applied Sciences,* vol. 10, no. 19, pp. 8533-8540, 2015.

[15] M. Kamel and R. Aicha, "Multi optimized SVM classifiers for motor imagery left and right hand movement identification," *Australas Phys Eng Sci Med,* vol. 42, no. 2, pp. 949-958, 2019.

[16] R. M. Mustaffa, W., Yee, N. Abdullah and A. N. Nasharuddin, "Colour-based building recognition using support vector machine," *TELKOMNIKA Telecommunication Computing Electronics and Control,* vol. 17, no. 1, pp. 473-480, February 2019.

[17] P. B. Harne, Y. Bobade, D. R. Dhekekar and A. Hiwale, "SVM classification of EEG signal to analyze the effect of OM mantra meditation on the brain," *2019 IEEE 16th India Council International Conference (INDICON),* 2019.

[18] B. P and S. K. J., "Support Vector Machine Technique for EEG Signals," *International Journal of Computer Applications,* vol. 63, no. 13, pp. 1-5, 2013 .

[19] E. Aydemir, T. Tuncer and S. Dogan, "A Tunable-Q wavelet transform and quadruple symmetric pattern based EEG signal classification method," *Medical Hypotheses,* vol. 134, 2020.

[20] R. Izabela and L. Krzysztof, "Genetic algorithm and forward method for feature selection in EEG feature space," *Journal of Theoretical and Applied Computer Science,* vol. 7, no. 2, pp. 72-82, 2013.

[21] aI. ,. Mohsin, S. A. Daghal and H. A. Sallomi, "A beamforming comparative study of LMS, GA and GWO algorithms for multipath smart antenna system," *TELKOMNIKA Telecommunication Computing Electronics and Control,,* vol. 18, no. 6, pp. 2911-2920, 2020.

[22] E. G. Talbi, "Metaheuristics From Design To Implementation," *John Wiley & Sons,* 2009.

[23] Y. Charles, M. Georges and R. Sandy, "A Genetic-Based Feature Selection Approach in the Identification of Left/Right Hand Motor Imagery for a Brain-Computer Interface," *Brain Sci,* vol. 7, no. 1, 2017.

[24] Emotiv, 5 May 2005. [Online]. Available: https://www.emotiv.com/.

[25] L. A. d. Mohand and O. Nalan, "EEG motor movement classification based on cross-correlation with effective channel," *Signal,Image and Video Processing,* vol. 13, pp. 567-573, 2018.

BIOGRAPHIES OF AUTHORS

Shaima Miqdad Mohamed Najeeb, she is presently one of the lecturers of technical engineering fauclty, Computer Technical engineering department, Northern Technical University, Nenwa, Iraq. She received her B.E. degree in computer Engineering in 1998 from Northern Technical University in Mosul, Iraq. Her M.Tech. degree in computer Engineering focusing on "independent component analysis of EEG signal during various mental and movement tasks" IRAQ in 2013. Her current research interests include signal and image processing, EEG and biomedical signal processing, artificial intelligent. Al Mosul city – Hay AL Moharibin, Nenwa, Iraq. The number of articles in national databases–3. The number of articles in international databases –1.

Haider Th. Salim Al Rikabi, he is presently Asst. Prof., and He is one of the faculty college of engineering, electrical engineering department, Wasit University in Al Kut, Wasit, Iraq. He received his B.Sc. degree in Electrical Engineering in 2006 from the Al Mus- tansiriya University in Baghdad, Iraq. His M.Sc. degree in Electrical Engineering focusing on Communications Systems from California state university/Fullerton, USA in 2014. His current research interests include Communications systems with mobile generation, Control systems, intelligent technologies, smart cities, and Internet of Things (IoT). Al Kut city – Hay ALRabee, Wasit, Iraq. The number of articles in national databases – 10. The number of articles in international databases –20.

Shaima Mohammed Ali, she is presently one of the lecturers of technical engineering fauclty, Computer Technical engineering department, Northern Technical University in Mosul, Nenwa, Iraq. She received her B.E. degree in computer Engineering in 2002 from Northern Technical University in Mosul, Iraq. Her M.Tech. degree in computer Engineering focusing on synthesis Systems configured on FPGA, IRAQ in 2016. Her current research interests include signal processing, EEG Signal Using Genetic Algorithm, intelligent technologies. Al Mosul city – Hay AL baker, Nenwa, Iraq. The number of articles in national databases–1. The number of articles in international databases –2.