

Research Article

Deep Learning Algorithm for Brain-Computer Interface

Asif Mansoor,¹ Muhammad Waleed Usman,² Noreen Jamil ,² and M. Asif Naeem ²

¹National University of Sciences and Technology, Islamabad, Pakistan

²National University of Computer and Emerging Sciences, Islamabad, Pakistan

Correspondence should be addressed to Noreen Jamil; noreen.jamil@nu.edu.pk

Received 23 January 2020; Accepted 10 July 2020; Published 25 August 2020

Academic Editor: Iván García-Magariño

Copyright © 2020 Asif Mansoor et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Electroencephalography-(EEG-) based control is a noninvasive technique which employs brain signals to control electrical devices/circuits. Currently, the brain-computer interface (BCI) systems provide two types of signals, raw signals and logic state signals. The latter signals are used to turn on/off the devices. In this paper, the capabilities of BCI systems are explored, and a survey is conducted how to extend and enhance the reliability and accuracy of the BCI systems. A structured overview was provided which consists of the data acquisition, feature extraction, and classification algorithm methods used by different researchers in the past few years. Some classification algorithms for EEG-based BCI systems are adaptive classifiers, tensor classifiers, transfer learning approach, and deep learning, as well as some miscellaneous techniques. Based on our assessment, we generally concluded that, through adaptive classifiers, accurate results are acquired as compared to the static classification techniques. Deep learning techniques were developed to achieve the desired objectives and their real-time implementation as compared to other algorithms.

1. Introduction

1.1. Background. Brain-computer interface (or BCI) is basically setting up a connection between the human brain and the computer device to control or to perform certain activity using brain signals. These brain signals are translated as an action for a device. The interface thus provides a one-to-one communication pathway between the brain and the target.

The technology has advanced from mechanical devices and touch systems, and now, world is approaching towards use of neural waves as the input. Even though it is not widely applied for now, it has a promising future. Especially for the physically impaired people who face difficulties in performing physical activities and lose their brain signal to move their muscles, it is the only way to function.

A BCI system includes a device with electrodes that act as sensors and measure brain signals, an amplifier to raise the weak neural signals, and a computer which decodes the signals into controlling signals to operate devices. Mostly, the BCI device is a headset which is portable and wearable.

The BCI device has two functions. Firstly, it records the data reviewed at its electrodes, and secondly, it interprets or decodes neural signals.

Nervous system resembles an electrical system which passes nerve impulses as a message. This means neurons (brain cells) communicate by transmitting and receiving very small electrical waves, merely in range of microvolts. Now, to sense and record these signals, we require precise and advanced sensors.

Electrodes are set directly on the scalp or embedded in the brain which requires surgical procedure. The non-surgical method of electrode placement though does not damage the brain, it yields poor-quality brain signals. Those that are recorded directly from the scalp yield better results but at the risk of surgery that may induce damage in the brain. The risk of damaging brain tissues exceeds the quality obtained through the surgical method. BCI is therefore a better pathway for neurorehabilitation for paralyzed people. Apart from these, other techniques include functional MRI (fMRI) and magnetoencephalography (MEG). fMRI maps brain activity with an MRI scanner, while MEG is a brain imaging process that identifies brain activity. Electric currents flowing through the brain produce magnetic field, and these are sensed by highly sensitive magnetometers. Both fMRI and MEG techniques use large and expensive machines. Another noninvasive

methodology is near-infrared spectroscopy (NIRS). In this process, neural signals are recorded by passing NI light through the head. The quality of the brain activity measurement is not adequate for the brain computer interface.

In case of healthy people, the brain transmits signals from the central nervous system to the muscles, and thus, they can move the muscles of the body. However, in case of people suffering from stroke or neuromuscular illness, the transmission of signals between the brain and the rest of body muscles is distorted. The patient's body becomes paralyzed or loses the capability to control muscle movement, like cerebral palsy. It is observed that a patient may not be able to move a muscle, but a brain can transmit the neural signal. This means that the neural signal is transmitted from the CNS but not received by target muscles. A BCI can be designed to utilize those commands to control devices or computer programs.

Each part of the body is controlled by a particular part of the brain as shown in the figure. Using BCI techniques, it is observed which part of the brain is active and transmitting the signal. Through this, the BCI system can predict the muscle locomotion from the brain activity [1].

BCI systems can be advanced, and multiple new applications can be developed using a fact that a variety of other brain activities can also be recognized. For instance, while one performs a numeric calculation, the frontal lobe is activated, and when one comprehends a language, Wernicke's area is activated.

Currently, numerous groups are contributing to the evolution of BCIs so as to develop numerous applications, specific for each category of the consumer. Each day, scientists and engineers are improving algorithms, BCI sensor devices, and techniques for quality attainment of data and improved accuracy of systems.

The problem is which method is optimal to analyze these complex time-varying neural responses and map them accordingly to the output response desired. These signals are merely in the range of microvolts. So, these electrical signals are passed through several processes to remove noise and to gather useful signals. Next, algorithms and classification techniques are applied to the data obtained [2].

1.2. Preliminaries. To attain a better understanding of BCI systems and the processes that undergo within them, an explanation of the terminologies and the said processes is presented as follows.

1.2.1. Brain Waves. Brain waves are oscillating voltages bearing amplitudes from microvolts to some millivolts; there are 5 widely known brain waves bearing different frequency ranges exhibiting states of the brain as shown in Table 1 [3].

1.2.2. Brain Activity Recording Methods for the BCI. The neural activity of the brain can be analyzed and understood based on the recording methods used. Recording methods of the BCI can be categorized as follows:

TABLE 1: Brain waves and associated frequencies.

Frequency band	Frequency	Brain states
Gamma	>35 Hz	Concentration
Alpha	12–35 Hz	Anxiety, relaxed, external attention
Beta	8–12 Hz	Very relaxed, passive attention
Theta	4–8 Hz	Deeply relaxed, inward focus
Delta	0.5–4 Hz	Sleep

(1) *Invasive Recording Techniques.* Invasive recording methods are those in which the electrodes are inserted deep in the brain using surgical methods, and the quality of the signal generated is better as compared to its noninvasive counterpart; however, issues arise from long-term stability, and protection is required to hinder them from creating infections. One such example is electrocorticography (ECoG), which measures the brain activity from the neural cortex.

(2) *Noninvasive Recording Techniques.* Noninvasive techniques do not require any surgical treatment and thus safe from causing any sort of infections; though their signal quality is low, it is still a popular means of brain signal acquisition.

These techniques include electroencephalography (EEG) in which the electrical activity is recorded from the scalp of the brain and magnetoencephalography (MEG) in which magnetic properties exhibited due to the difference in oxygenated and deoxygenated hemoglobin are recorded.

For our project, we will opt for an EEG-based signal recording technique and explain its characteristics in the following [1].

1.3. Electroencephalography (EEG). Introduced by Hans Berger in 1929, EEG is a measurement of voltage levels that underlines the activity of the brain in response to an event or a stimulus. EEG method comprises electrodes placed on the scalp of the brain at different locations as specified in Figure 1 with temporary glue. The electric signals are generated due to the ionic content present in the brain consisting of Na⁺, Ca⁺⁺, K⁺, and Cl⁻ ions; the transportation of these ions invokes the electric potential used in EEG.

The EEG signals are of low quality because of different layers of tissues between the EEG cap and the signal source as shown in Figure 2. The potential created is in a range of tens of microvolts, and these electrodes need to have powerful amplifiers in order to acquire meaningful signals.

1.4. Need of BCI. Brain-computer interface-based technology is a developing field, and it has been under focus by many industries to innovate and make everyday life tasks easier. One of the questions which arises in the mind is why we need BCI systems? BCI system is a complex technology, no doubt, however, leading to a simpler life.

Following are the main reasons why we need to focus on this technology:

- (i) Control of devices can be made easy through just our thoughts

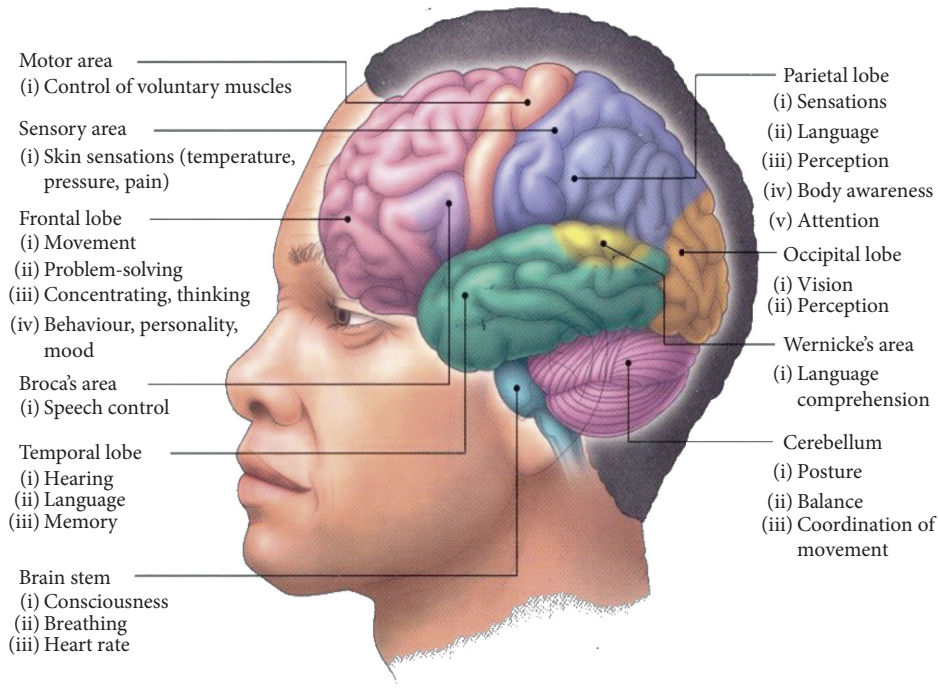
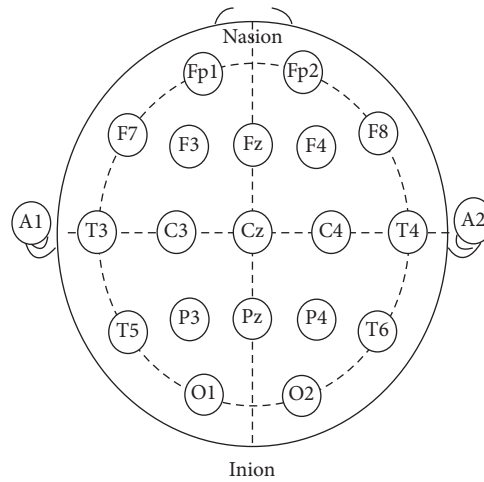


FIGURE 1: Brain anatomy.



(a)



(b)

FIGURE 2: EEG: (a) subject wearing a 32-electrode EEG cap; (b) standardized electrode placements.

- (ii) Making a decision and then performing a task takes time, while operating a device using thoughts or technically our brain waves is easier
- (iii) Re-establishing communication path from the brain to artificial limbs and assisting those affected by brain-related diseases

1.4.1. Individuals in Need of a BCI to Re-Establish Motor Control and Communication. A wide range of neurological diseases such as motor neuron diseases and spinal cord injury may lead to severe paralysis of the motor muscles, restricting the patient to control artificial devices through

only a few muscles and thereby termed as “locked-in,” while people who have completely lost their motor control are termed as “completely locked-in.”

Evident that the normal communication channel from the brain to the limbs is lost, BCI is used to re-establish the communication through an alternative route.

Even being applicable to a healthy person, BCI systems can be used to employ numerous tasks from the users using the signals generated from the brains to control applications as presented in the following [4]:

(1) *Noninvasive Brain-Computer Interface Research at the Wadsworth Center.* The research conducted at the

Wadsworth Center was to study different approaches employed in the BCI to control a computer screen cursor to analyze their advantages and disadvantages; one approach was sensory-based rhythm control in which the selected features in the frequency domain were based on the potentials created by motor imagery and linear regression was employed so that they can be converted as control signals to move the cursor.

The other procedure was the P300-based cursor control in which the user focuses attention on the desired symbol and is provided with a 6×6 matrix to produce time-varying stimuli and linear regression is utilized to allow these signals as a control input to move the cursor.

The research suggested that the BCI is an application-oriented approach and depends entirely on user training; the EEG features dictate the BCI system for speed, accuracy, bit rate, and usefulness.

Sensorimotor Rhythms (SMR) is an approach employing better results for control tasks such as controlling a screen cursor, while the P300-BCI system was slower as compared to the SMR-BCI.

(2) *The Berlin Brain-Computer Interface: Machine Learning-Based Detection of User-Specific Brain States.* The researchers for the Berlin brain-computer interface employed sensory motor rhythms, i.e., thinking of moving the left hand or right hand and used machine learning-based detection of the user specific brain states. While testing their trained model, they achieved an information transfer rate above 35 bits per minute (bpm), and overall spelling speed was 4.5 letters per minute including correcting the mistakes, using 128-channel EEG and using feedback control for untrained users in order to properly train the machine learning algorithms, thereby reducing the training user time used in the voluntary control approach [2].

1.5. *Structure of the BCI.* The steps of the brain computer interface system include the following:

- (1) Brain activity measurement/recording methods of the BCI
- (2) Preprocessing techniques
- (3) Feature extraction
- (4) Machine learning implementation/classification
- (5) Translation to control signal

1.5.1. *Preprocessing.* In BCI, preprocessing techniques consist of data acquisition of brain signals to check the signal quality without losing important information; the recorded signals are cleaned and conduct noise removal to acquire relevant information encoded in the signal. As mentioned above, the EEG signals are of poor quality; even the commercial 50 Hz frequency, due to nearby appliances, can corrupt the EEG signals, and the users are also advised not to think anything else apart from the stimuli as presented. In preprocessing, using Fourier transform or Fourier series, the signals are taken into the frequency domain and studied what

frequency content is present in the signal. The undesired 60 Hz frequency signal and undesired signal produced by performing actions other than the said stimuli are then filtered out using a notch filter as mentioned in Figure 3.

1.5.2. *Feature Extraction.* Feature extraction plays a vital role in brain-computer interface applications; the raw EEG signals are nonstationary signals that are corrupted by noise or due to artifacts present in the environment where they are being recorded, but still meaningful information can be extracted from them. The data dimensionality is also reduced to process it better, and machine learning models are applied. This method is essential to increase the classification accuracy of the BCI system.

EEG signal is a time-domain nonstationary signal, and the relevant information such as signal energy is analyzed as a function of time or frequency later; relevant statistic measures are adapted to properly explain the characteristics of the signal.

Some of the commonly used feature extraction techniques are listed as follows.

Short-time Fourier transform is a frequency-domain feature extraction technique in which the EEG signal is convolved with a window function w to extract the relevant frequency features of the brain which are broken down as sinusoids at different frequency ranges.

Mathematically, it is represented as

$$Xstft[m, n] = \sum_{k=0}^{L-1} x[k]w[n-k]e^{-j2\pi mk/L}, \quad (1)$$

where $x[k]$ is the input EEG and $w[n-k]$ is the window multiplied to extract the frequency features as shown in the figure [4].

Discrete-time and continuous-time wavelet transform is a time frequency-based feature extraction technique that allows better temporal and spatial resolution in which the EEG signals are produced in the form of wavelets at different frequency ranges of interest as shown in Figure 4.

The technique of wavelet transforms as adapted in the literature is a filtering continuous application of high-pass and low-pass filters to extract the wavelets of the signal which, when added together, constitute the original signal and downsampled by a factor of 2 as shown in Figure 5.

g and h are consecutive high-pass and low-pass filters which produce the relevant detailed and approximated coefficients of the EEG signals [5].

1.5.3. *Neural Networks.* Before starting to explain what deep learning is, it is first beneficial to explain the role of deep learning and its fundamental blocks.

Deep learning is a classification tool used in a variety of daily applications which is composed of speech recognition and computer vision to natural language processing in the context of the BCI; the input features which are different brain frequency bands are classified according to what activity the user is performing at the moment.

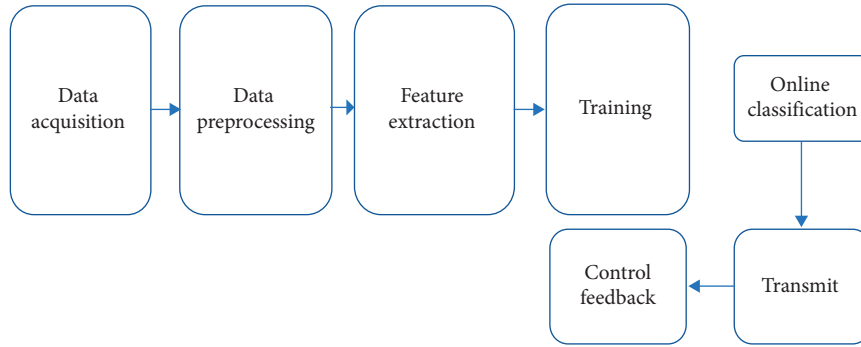


FIGURE 3: BCI block diagram.

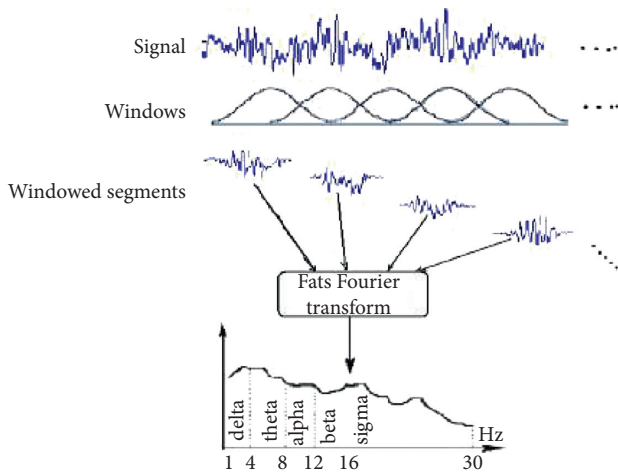


FIGURE 4: Feature extraction using short-time Fourier transform.

Neural Network. A neural network is a model similar to that of a neuron in our brain that has input nodes and output nodes; the mathematical model for a neural network is given by the following equation:

$$v = wx + b, \tag{2}$$

where v is the weighted sum of the inputs and the bias term which will be fed at the output node, b is a bias term which is mostly set to 1, and w is the random weights assigned that are multiplied with the input in order to reach closer to the desire output.

The neural network is shown in Figure 6.

These calculations are often preceded in the form of matrices; the input, the weight terms, the output, and the bias are as follows:

$$v = (w_1 \times x_1) + (w_2 \times x_2) + (w_3 \times x_3) + b. \tag{3}$$

Finally, the output node is passed to an activation function and provides the final output; the activation function calculates the characteristic of the node. The activation function acts to map the corresponding inputs to the right output y present at the output node:

$$y = \varphi(v). \tag{4}$$

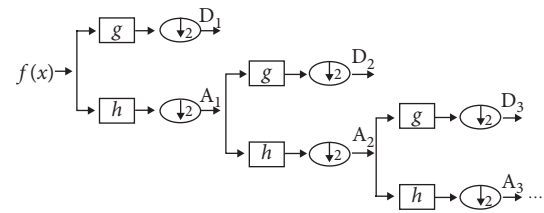


FIGURE 5: Wavelet transform.

The neural network does not get its right output at the first attempt. It needs to be trained a lot, and so a training rule is assigned to neural networks to get the right output. Many training rules are adapted, but one of the most commonly used is the delta rule, and the rule is expressed using the following equation:

$$w_{ij} \leftarrow w_{ij} + \alpha e_i x_j, \tag{5}$$

where x_j represents the number of inputs, e_i is the error generated at the output node, and α is the learning rule between $(0 < \alpha < 1)$.

The training rule is summarized as follows:

- (1) Assign adequate values to the weights.
- (2) Obtain the input from the training data and feed it into the neural network which will give an output d ; subtract the output d to obtain the correct output at the output node.

$$e_i = d_i - y_i. \tag{6}$$

- (3) Calculate the weight updates:

$$\Delta w_{ij} = \alpha e_i x_j. \tag{7}$$

- (4) Adjust the weights accordingly until the correct output or that has small tolerance is obtained:

$$w_{ij} \leftarrow w_{ij} + \Delta w_{ij}. \tag{8}$$

The above explanation was presented for a single-layer neural network; the architecture of neural networks is

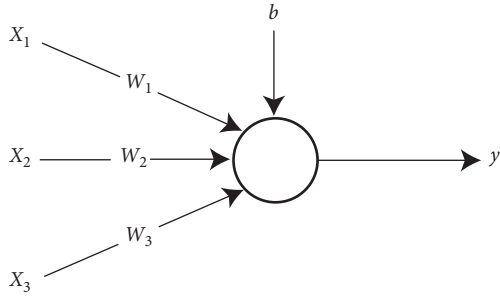


FIGURE 6: Neural network.

becoming better with the cost of greater memory, but with higher classification accuracy, we use deep neural networks which are the same as the single-layer neural network but with hidden layers added in between the input and output nodes, as shown in Figure 7.

The concepts are similar to those of a single neural network but with the adjustments of added hidden layers and a different training rule because the delta rule has a drawback of not propagating the output to the hidden layers, thereby the weights are not adjusted.

To explain how the deep neural network works, the above explained single neural network is set as basis.

In Figure 8, given a multiple-layered neural network, the weighted sum obtained at the first hidden layer is presented as

$$\begin{bmatrix} v_1^{(1)} \\ v_2^{(1)} \end{bmatrix} = \begin{bmatrix} w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (9)$$

$$\triangleq W_1 x.$$

The outputs are calculated via the sigmoid activation function:

$$\begin{bmatrix} y_1^{(1)} \\ y_2^{(1)} \end{bmatrix} = \begin{bmatrix} \varphi(v_1^{(1)}) \\ \varphi(v_2^{(1)}) \end{bmatrix}. \quad (10)$$

The process is repeated, and the outputs obtained are treated as the inputs to the other nodes, and we get the outputs as

$$\begin{bmatrix} v_1^{(1)} \\ v_2^{(1)} \end{bmatrix} = \begin{bmatrix} w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \end{bmatrix} \begin{bmatrix} y_1^{(1)} \\ x_2^{(1)} \end{bmatrix} \quad (11)$$

$$\triangleq W_1 y^{(1)}.$$

And lastly, the weighted sum is being inserted into the activation function, and we return our final output:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \varphi(v_1) \\ \varphi(v_2) \end{bmatrix}. \quad (12)$$

Deep learning training rule is given in the following.

Backpropagation algorithm is commonly used as the training instruction for the deep neural networks; the procedure is summarized as follows:

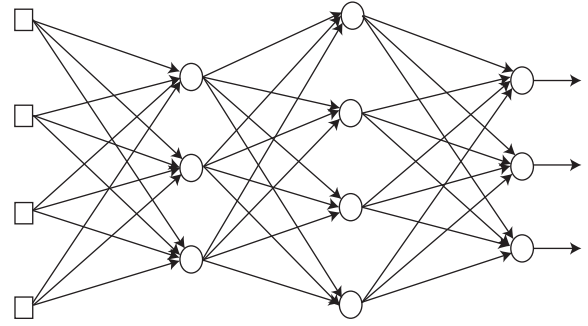


FIGURE 7: Structure of the deep neural network.

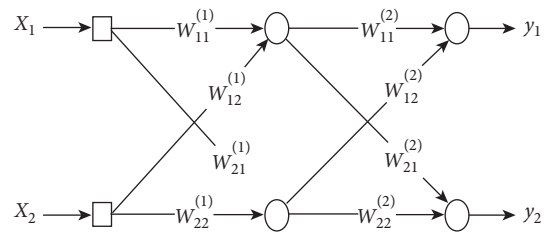


FIGURE 8: Multilayered network.

- (1) Assign adequate values to the weights.
- (2) Take the input from the training data and feed it into the neural network which will give an output d . Subtract the output d to obtain the correct output at the output node and the delta (δ) of the output nodes:

$$\begin{aligned} e &= d - y, \\ \delta &= \varphi'(v)e. \end{aligned} \quad (13)$$

- (3) Propagate the delta back towards the hidden nodes, and determine respective delta δ of nodes:

$$\begin{aligned} e^k &= W^T \delta, \\ \delta^{(k)} &= \varphi'(v^{(k)})e^{(k)}. \end{aligned} \quad (14)$$

- (4) Repeat until it reaches the input nodes.
- (5) Modify the weights according to the rule:

$$\begin{aligned} \Delta w_{ij} &= \alpha \delta_i x_j, \\ w_{ij} &\leftarrow w_{ij} + \Delta w_{ij}. \end{aligned} \quad (15)$$

- (6) These steps are repeated until the neural network is utterly trained as shown in Figure 6.

Now, the alpha-beta ranges are extracted, and consecutive energies are calculated. The features are fed into the deep learning neural network, and using the backpropagation learning rule, the model is trained, yielding an accuracy of 97.83%, as shown in Figure 9 [5].

2. Critical Review of the Related Literature

A brain-computer interface involves various stages, and development in each stage leads to an improved and efficient system. Here, the literature review of major steps including data acquisition, feature extraction, classification algorithm, and applications is presented.

- (1) Improvement of EEG signal acquisition: An electrical aspect for state of the art of front end. Computational intelligence and neuroscience: a research paper published by Ali Bulent Usakli, the NCO Academy, Turkey, presented some applicable concerns for acquiring quality EEG signals which are proven helpful for users and design engineers. One of the most important considerations is selecting suitable electrodes and headset. In the EEG-based BCI, electrodes, signal processing components including mental and environmental conditions, filtration of noise, amplification, signal translation, and data storage affect the recording process. The data acquisition is an important step in any machine learning procedure. Brain signals need to be cleaned and preprocessed so that a good result can be obtained [1].
- (2) P300 wave detection using Emotiv EPOC+ headset: effects of matrix size, flash duration, and colors: Saleh Ibrahim Alzahrani conducted research on P300 wave detection using the Emotive EPOC+ headset to study the effects of the size of the matrix, flash duration, and colors. In this study, P300 signals were obtained from five subjects with Emotive EPOC+ using all 14 channels. For this research, EEG signals obtained were communicated to software OpenViBE through a USB dongle. A sample was taken every 8 seconds at a rate of 128 samples per second. The configured sampling rate provides ample samples for the four frequency bands, containing significant ERP data. During process of signal recording, the subjects were shown a matrix of 6×6 or 3×3 on the computer screen. They were instructed to stay calm, avoid any needless movement, and set all on letter which they desire to spell. It is reported in the study that one of the advantages of using the Emotiv EPOC+ headset is that it takes merely two to three minutes for preparation as compared to other headsets that take more than ten minutes. The quality of sensors is verified through Emotiv Xavier SDK which provides feedback report of each sensor. To decrease the contact impedance, saline liquid was applied to the sensor surface. Primary objective was to assess the potential of Emotiv EPOC+ to perceive P300 signals. Finding the electrodes proficient at providing target signals helps minimize the number of channels which makes signal processing a lot easier. The results of this experiment provide evidence of the capability of Emotiv EPOC+ to detect the P300 signals from two channels, O1 and O2 [2].
- (3) Automatic seizure detection in SEEG using high frequency activities in wavelet domain. Medical engineering and physics: in this research paper, the researcher has found the method for detection of seizures using the high-frequency analysis in the wavelet domain. This method is used highly in the high-frequency domain. Because of seizure detection, the method is usually done using high frequency in the range of 80–250 Hz. Also, it can handle fast ripples in the range of 250 to 500 Hz. The biggest advantage in the seizure detection is that it can detect the seizure offset. The methodology consists of the Continuous Wavelet Transform (CWT), which was computed by convolving the SEEG signal which has to make the feature extraction, and it also includes the complex conjugate of the wavelet basis function (Ayoubian, 2013).
- (4) Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble: envelope detection is a very efficient method for detecting the impact of the signals which are based on the biological change or diagnosis. In this paper, the researchers used the Hilbert transform which has a good impact on the resultant signals so that the signals are then widespread using the DWT technique which has a unique behavior regarding the biological change; the researchers detected the changes using this method [5].
- (5) Feature selection for motor imagery EEG classification based on firefly algorithm and learning automata: in this research paper, the researchers implemented spectral linear discriminant analysis for the classification of motor imagery signals. Feature extraction method used was basically common spatial patterns; the advantage of using this feature extraction method is basically of two-class discrimination problems. This maximizes the variance of one class and decreases the variance of the other class, which is the advantage, but the disadvantage is because of the multiclass overlap structure in this method, it is not used for multiclass prediction [6].
- (6) Unsupervised adaptation of electroencephalogram signal processing based on fuzzy C-means algorithm. International Journal of Adaptive Control and Signal Processing: this research paper presented the techniques of brain mapping with emphasis on multichannel EEG and functional brain imaging techniques. During training and testing, the concentration of the subject on the target object is one of the concerns which signifies the capacity to operate a device. They have used different algorithms such as LDA and fuzzy C-means. Fuzzy C-means is an adaptive classifier, and this is probably

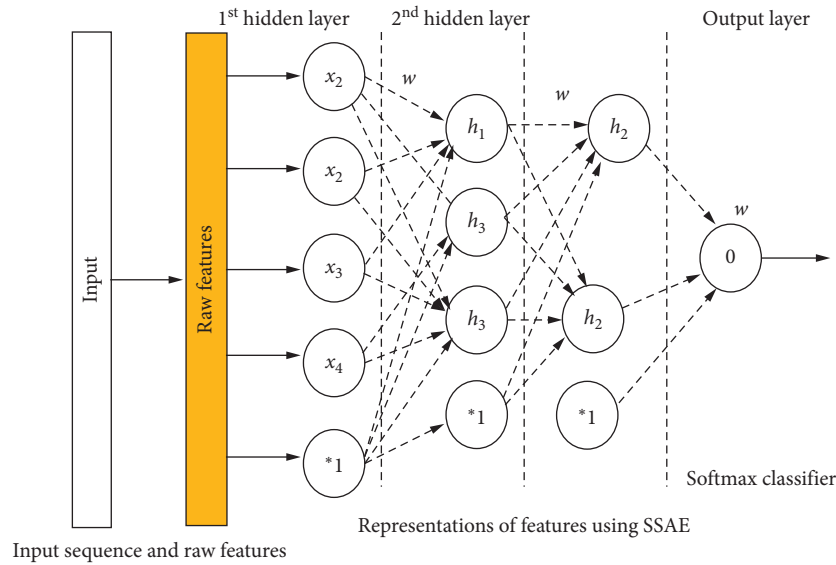


FIGURE 9: General diagram of neural networks in the BCI.

used where the device behavior is not synchronized with the classification model [7].

- (7) A review of classification algorithms for EEG-based brain-computer interfaces: a 10-year update: this paper is the latest review of the BCI classification techniques; there are some algorithms discussed in this paper taken from different papers, and their accuracies, optimization, the method of feature extraction used were compared, and every algorithm has its own advantages and disadvantages. In this paper, they have used the event-related potentials. Feature extraction method they have used is based on spatial filtering which is the most optimized filter; it can further be optimized by using the calibration [8].
- (8) Sequential non-stationary dynamic classification with sparse feedback: in this research paper, basically, they have discussed the technique for the classification of the nonstationary and nonlinear signals. As we know that the BCI signals are nonstationary in nature, sparse feedback can be used for the stability of the brain signals and classifying them using the RBF classifiers which are radial basis functions. The signals are acquired in a nonlinear manner, and we can apply linear models to them, but again, multiclass prediction is not able to be performed. This is because of the sparse feedback matrices involved [9].
- (9) Motor imagery and direct brain-computer communication: Gert Pfurtscheller and Neuper researched about the technique for the motor imaginary signals by the imagination of the left hand, right hand, and foot movements. In the neurofeedback method, real-time prediction of brain signals is difficult to achieve. We have nonlinear signals at the input of the neurofeedback

method so that we can use the Hidden Markov Method (HMM) to make predictions in real time but the accuracy is a tradeoff [10].

- (10) Toward unsupervised adaptation of LDA for brain-computer interfaces. IEEE Transactions on Biomedical Engineering: the firefly algorithm (FA) is an efficient algorithm for selecting the most appropriate subset of features and helps improving accuracy of classification. When the problem of entrapping of FA in the local optimum arises, a procedure for combining the firefly algorithm and learning automata (LA) is proposed which optimizes feature selection for motor imagery EEG. For the expected outcome of the high-dimensional feature set, a process of combining the common spatial pattern (CSP) and local characteristic-scale decomposition (LCD) algorithms is used. It is further classified by the use of the spectral regression discriminant analysis (SRDA) classifier [11].

3. Comparison of Classification Algorithms

Table 2 shows the comparison of classification algorithms.

4. Discussion

There is a large range of classifiers developed by scientists and engineers around the world. These classification algorithms can be divided into four groups.

4.1. Adaptive Classifiers. Adaptive classifiers are listed as those classifiers in which parameters are progressively recalculated and also updated with procurement of new EEG data signals which are nonstationary, and adaptive classifiers are capable to follow the change in the feature distribution.

TABLE 2: Comparison of classification algorithms.

Title	Algorithm	Input features	Efficiency	Advantages	Drawbacks/tradeoff
Vidaurre et al. [12]. Toward unsupervised adaptation of LDA for brain-computer interfaces. IEEE transactions on biomedical engineering, 587–597.	Linear discriminant analysis uses hyperplanes for different classes, assuming normal distribution, with equal covariance matrix for both classes; to solve an NC class problem, several hyperplanes are used.	Separating hyperplane is obtained by seeking the projection that maximizes the distance between two classes' means and minimizes the interclass variance.	Suitable for online BCI and provides generally good result, and fluctuations in the training data set do not affect much.	Very low computational requirement so suitable for online BCI system.	Linearity that can provide poor results on complex nonlinear EEG data (not immune to noise).
Li et al. [5]. Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble. Biomedical signal processing and control, 357–365.	Support vector machine uses a support vector hyperplane to identify classes.	Selected hyperplane is the one that maximizes margins, i.e., the distance from nearest training points.	Enables classification using linear decision boundaries, (linear SVM) has been applied, generalization capabilities, to be insensitive to overtraining and to the curse of dimensionality.	Maximizing margins and regularization are known to increase the accuracy.	These advantages are gained at the expense of a low speed of execution.
Lotte et al. [4]. A review of classification algorithms for EEG-based brain-computer interfaces: a 10-year update. Journal of neural engineering, 031005.	Neural network consists of at least three layers of nodes. Except for the input nodes, each node is a neuron that uses a nonlinear activation function which utilizes a supervised learning technique called backpropagation for training.	Neurons of the output layer determine the class of the input feature vector.	Applied to almost all BCI applications.	Because universal approximators are composed of enough neurons and hidden layers, they can approximate and classify any continuous signal.	Sensitive to overtraining, especially with such noisy and nonstationary data as EEG; therefore, careful architecture selection and regularization is required.
Usakli [1]. Improvement of EEG signal acquisition: An electrical aspect for state of the art of front end. Computational intelligence and neuroscience.	Consists of the noninvasive technique for recording brain signals which is based on the electromagnetic resonance signals compared to that of the EEG scalp signals.	Brain signals as an electrical pulse coming from the brain.	Suitable for all BCI systems. They are based on the noninvasive technique in real-time monitoring of signals.	Proven helpful for users and design engineers. One of the most important considerations is selecting suitable electrodes and headset.	Costly design because of the gold electrodes. They are costly, and everyone cannot afford that system.
Alzahrani [2]. P300 wave detection using Emotive Epoc+ headset: Effects of matrix size, flash duration,	Emotive Epoc+ 14 channel sensor was used which has 14 channels for EEG and a neutral channel as well.	Input features consist of P300 steady-state evoked potentials.	Suitable for the brain signals which are collected by the Emotive Epoc+ sensor. The signals are carried out using Emotive Epoc.	Advantage is basically being optimized, and the device is very cheap in price with affordable accuracy.	For more number of class predictions, the accuracy becomes low, and the output is affected.

TABLE 2: Continued.

Title	Algorithm	Input features	Efficiency	Advantages	Drawbacks/tradeoff
Ayoubian, L. a. (2013). Automatic seizure detection in SEEG using high frequency activities in wavelet domain. Medical engineering and physics, 35, 319–328.	Based on the continuous wavelet transform, the brain signals were computed by convolving the SEEG signal.	Brain signals were collected from the Stellate Harmonic system for EEG monitoring purpose. These signals were passed through a band-pass filter.	Suitable for the detection of the seizure. A seizure onset is added to the signals and then compared with the normal brain signal.	For automatic seizure detection, it is very useful, and it can be used for the patients who are not able to calculate when they have seizure.	The disadvantage is basically for some high-frequency seizures. The high-frequency seizures are not detected easily because of the band-pass filter.
Liu et al. [6]. Feature selection for motor imagery EEG classification based on firefly algorithm and learning automata. Sensors.	Spectral regression discriminant analysis (SRDA) is widely used in the feature classification; in this paper, they have implemented this algorithm.	Separating hyperplane is obtained by seeking the projection that maximizes distance between two classes' means and minimizes the interclass variance.	Suitable for online BCI and provides generally good result, and fluctuations in the training data set do not affect much.	Very low computational requirement so suitable for online BCI system, simple to use, and generally provides good results.	Linearity that can provide poor results on complex nonlinear EEG data (not immune to noise).
Lowne et al. [9]. Sequential non-stationary dynamic classification with sparse feedback. Pattern recognition, 897--905.	Spectral regression discriminant analysis (SRDA) is widely used in the feature classification; in this paper, they have implemented this algorithm.	Separating hyperplane is obtained by seeking the projection that maximizes distance between two classes' means and minimizes the interclass variance.	Suitable for online BCI and provides generally good result, and fluctuations in the training data set do not affect much.	Very low computational requirement so suitable for online BCI system, simple to use, and generally provides good results.	Linearity that can provide poor results on complex nonlinear EEG data (not immune to noise).
Liu et al. [6]. Unsupervised adaptation of electroencephalogram signal processing based on fuzzy C-means algorithm. International journal of adaptive control and signal processing.	Common spectral patterns were used for the feature extraction and the linear discriminant analysis, and fuzzy C-means was used for the feature classification.	The maximum distance was calculated for each fuzzy C-means, and then the mean was calculated; after that, the features were classified.	They are suitable for nonlinear EEG signals having different amplitudes for different people.	Very low computational requirement so suitable for online BCI system, simple to use, and generally provides good results.	Fuzzy behavior can be seen in the output when the frequency changes at the input.
Pfurtscheller and Neuper [10]. Motor imagery and direct brain-computer communication. Proceedings of the IEEE, 1123–1134.	Hidden Markov model was used for the classification of EEG signals as they are nonlinear in nature, so we can tune the Markov model accordingly.	The input signals were obtained from the two channels, and these signals were transformed into the HMM network.	For two channels, EEG signals, this is good, and it has a fast classification.	The method used in this paper uses low computational power, and the model functions are optimized.	Output accuracy depends on the linear behavior of the signals. When the frequency fluctuates the output, control will also change.

As mentioned in [10], a model for a motor-imagery-based self-paced BCI structure for operating a robot was proposed. They used a basic synchronous BCI, devised earlier for recording data for offline training classification before conducting the online self-paced procedure. They extracted logarithmic band power as features, and features were extracted from EEG signals. Feature selection

was manual so as to gain quality frequency bands. To extract the features, chosen frequency bands were digitally bandpass-filtered, squared, and averaged over 1 second sliding window, and natural log was applied. Utilizing the features and associated labels, two linear discriminant analysis (LDA) classifiers were trained, with one to recognize left imagery from right and the other to

isolate right imagery movement development from others.

Features related to subject's control brain signals are extracted and constantly classified by the offline LDA algorithm. These features are then used to control the robot. They used parameters of the LDA algorithm in place of accommodating threshold and dwell features. Subject's control intention and timing is the basis for adaptation for online training. The methodology proposed entailed information about the user's control needed to train and adapt which could show promise of improving the accuracy and performance. The paper used Kalman filters for online parameters to be classified using LDA. Event label assignment was introduced with a slower learning process [12].

One of the advantages of implementing adaptive classifiers is that they can employ both supervised and unsupervised. This means that even if there is no information of true labels of data being received, they can be executed. From multiple research studies, it is inferred that unsupervised learning has yielded better results than static classifiers.

Now, most of the real-world applications of BCI do not present class labels, and for this purpose, unsupervised adaption classifiers require more development [7].

4.2. Matrix and Tensor Classifiers. The approach which the researchers have used in this paper holds fewer stages for classification as compared to the classical machine learning algorithms, and they are simpler as well. Compared to other standard classifiers, Riemannian classifier does not need any parameter-tuning techniques such as cross validation to properly train and verify the accuracy of the produced model, which makes it far more robust and accurate all due to its logarithmic tendencies. Likewise, the inborn Riemannian separation for the SPD matrix is invariant both to inversion of the matrix and to any direct invertible change of the information, for example, any outside interference added to the EEG sources does not change the separations among the witnessed covariance matrices. These properties partially clarify why Riemannian classification techniques give a decent speculation ability, which empowered analysts to set up adjustment-free versatile ERP-BCIs utilizing basic subject-to-subject and session-to-session exchange learning methodologies [9].

It is shown that several approaches were implemented and gave higher performance than CSP + LDA procedures on motor imagery EEG data. The biggest advantage is quality performance. However, this is a gain at tradeoff between performance and greater number of weights because of elevated expansion in input feature dimensionality which makes suitable regularization a much-needed step [12].

EEG data can be represented in the form of tensors and are treated as analysis tools for EEG data tools for EEG data analysis which includes feature extraction, data clustering, and data classification in the BCI. EEG data are represented in more than one dimension; this includes time, frequency space, and trails and hence, these are presented by the tensor order. EEG data that have time frequency and space can be

represented by 3-dimensional tensor. Tensors have been used frequently for motor imagery-based analysis even with a large amount of data containing different categories which can be represented by the tensor and its order [10].

4.3. Transfer Learning and Deep Learning. Transfer learning is a crucial tool when it comes to session-to-session and subject-to-subject decoding performance. If transfer learning is improved enough, BCI system can be operated without calibration, and this will revolutionize the BCI systems forever.

It is observed that calibration sessions are strenuous and mentally exhausting for subjects. It is explained that accepting the input from the earliest starting point of their BCI system is promising for started subjects.

Deep learning is categorized as a ML algorithm, where features and the classifier are collectively learned straight from EEG data. There exist multiple deep learning algorithms. One of the most explored and commonly used is deep neural networks (DNNs). DNN is also performed online for slow cortical potentials (SCP) and motion-onset visual evoked potential (MVEP) which are not commonly used. The very first research conducted and paper was published on the P300-based BCI by Ciotte et al. Two convolutional layers were constructed followed by completely connected layer. One convolutional layer has a purpose to learn spatial filters and the second one was to learn temporal filters. The model was proven better than the P300 experiment, but the SVM model had more accuracy [2].

Deep extreme learning machine is used for Slow Cortical Potentials (SCP) classifications. This technique consists of multiple layers, and the last one was Kernel ELM. However, in this project, number of units, network structure, hyperparameter, and input features were not reasoned. This did not prove to be better than multiplayer ELM or standard ELM [4].

4.4. Miscellaneous Classifiers. In order to classify more than two mental tasks, two main approaches can be used to obtain a multiclass classification function. The first approach consists in directly estimating the class using multiclass techniques such as decision trees, multilayer perceptron, naive Bayes classifiers, or k -nearest neighbors. The second approach consists of decomposing the problem into several binary classification problems.

Multiclass and multilabel approaches therefore aim to recognize more than two commands. It is therefore necessary to choose carefully the mapping between mental commands and corresponding labels. However, the errors may be possible during the classification. In particular, the set of estimated labels, sometimes, may not correspond to any class, and several classes may be at equal distances, thus causing class confusion [13].

5. Methodology

Here are some methods which are discussed in the research papers for the past few years in brain-computer interface systems, as shown in Table 3.

TABLE 3: Summary of various methodologies in BCI systems.

Classification methods	Input EEG pattern	Features	References
Adaptive classifiers	Motor imaginary-based	Frequency band power, EEG time	[10, 12, 14]
Matrix and tensor classifiers	Steady-state visual evoked potentials, P300	Frequency band power, raw data	[2, 5, 7, 15]
Transfer learning and deep learning	Motor imaginary-based, P300, SSVEP	EEG time points, frequency band power, raw EEG data	[2, 4, 6, 7]
Miscellaneous classifiers	Motor imaginary-based, P300	Not specified	[1, 4, 15]

6. Conclusion

During this course of work, a question arises whether it is possible to create a brain computer interface which is affordable, with high accuracy and optimization. So, after reviewing different papers, the conclusion is that if we need an optimized model with high accuracy for the noninvasive technique of brain signals, the artificial neural network has a high accuracy and is optimal. However, there are some tradeoffs as well that are model compatibility with the brain signals. From 10 years of BCI review, we have obtained that the ANN has a high response and accuracy; after all, it optimizes the system as well. However, further research studies have been done to make it more accurate because this has to be used in health care.

Due to the fast processing that ANN allows, a form of guidance could be provided during training that enables a user to improve the classifier's performance and let it reflect his/her intents more often. This guidance was very useful for one of the three subjects.

Also, the statistical test was examined on whether it performs in a way that can be expected. Therefore, after the offline classification, it yielded an accuracy above 90 percent.

The control provided by the developed system has been sufficient to conclude that it provides enough control that a user can command an arbitrary computerized device. Also, it showed to be easily trainable.

In the future, the proposed model can provide support on multiplatforms. This can be achieved by developing applications which can help humanity and make everyday tasks easier. Furthermore, the system can be controlled with a smartphone that can override EEG headset commands. This will act as fail-safe if the BCI system experiences any malfunctioning. On the basis of this development, better EEG can be designed with higher efficiency and which is less dependent on offline classification.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] A. B. Usakli, "Improvement of EEG signal acquisition: an electrical aspect for state of the art of front end," *Computational Intelligence and Neuroscience*, vol. 2010, Article ID 630649, p. 12, 2010.
- [2] S. I. Alzahrani, "P300 Wave Detection Using Emotiv EPOC+ Headset: Effects of Matrix Size, Flash Duration, and Colors," Colorado State University, Fort Collins, CO, USA, Doctoral dissertation, 2016.
- [3] F. Yger, M. Berar, and F. Lotte, "Riemannian approaches in brain-computer interfaces: a review," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1753–1762, 2016.
- [4] F. Lotte, L. Bougrain, A. Cichocki et al., "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," *Journal of Neural Engineering*, vol. 15, no. 3, Article ID 031005, 2018.
- [5] M. Li, W. Chen, and T. Zhang, "Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble," *Biomedical Signal Processing and Control*, vol. 31, pp. 357–365, 2017.
- [6] A. Liu, K. Chen, Q. Liu, Q. Ai, Y. Xie, and A. Chen, "Feature selection for motor imagery EEG classification based on firefly algorithm and learning automata," *Sensors*, vol. 17, no. 11, p. 2576, 2017.
- [7] G. Liu, D. Zhang, J. Meng, G. Huang, and X. Zhu, "Unsupervised adaptation of electroencephalogram signal processing based on fuzzy C-means algorithm," *International Journal of Adaptive Control and Signal Processing*, vol. 26, no. 6, pp. 482–495, 2012.
- [8] A. Andreev, A. Barachant, F. Lotte, and M. Congedo, *Recreational Applications of OpenViBE: Brain Invaders and Use-The-Force*, Wiley Online Library, Hoboken, NY, USA, 2016.
- [9] D. R. Lowne, S. J. Roberts, and R. Garnett, "Sequential non-stationary dynamic classification with sparse feedback," *Pattern Recognition*, vol. 43, no. 3, pp. 897–905, 2010.
- [10] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.
- [11] A. Schlögl, C. Vidaurre, and K. R. Müller, *Adaptive Methods in BCI Research-an Introductory Tutorial*. In *Brain-Computer Interfaces*, Springer, Berlin, Heidelberg, Germany, 2009.
- [12] C. Vidaurre, M. Kawanabe, P. von Büna, B. Blankertz, and K. R. Müller, "Toward unsupervised adaptation of LDA for brain-computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 587–597, 2010.
- [13] A. B. R. Suleiman and T. A. H. Fatehi, *Features Extraction Techniques of EEG Signal for BCI Applications*, University of Mosul, Mosul, Iraq, 2007.
- [14] I. A. Fouad and F. E. Z. M. Labib, "Using emotiv EPOC neuroheadset to acquire data in brain-computer interface," *International Journal*, vol. 3, no. 11, pp. 1012–1017, 2015.
- [15] G. A.-R. Dornhege, *Toward Brain-Computer Interfacing*, MIT Press, Cambridge, MA, USA, 2007.