

Study of Interactive Systems Based on Brain and Computer Interfaces

Zahra Ghadiri¹, Nasim Nourafza^{2*}, Abolfazl Rasoolian³

¹Department of Management, Najafabad Branch, Islamic Azad University, Najafabad, Iran

Email: ghadiry_53@yahoo.com

^{2*} Assistant Professor, Faculty of Computer Engineering, Najafabad branch, Islamic Azad University, Najafabad, Iran

Email: noorafza@iaun.ac.ir

³Department of Management, Najafabad Branch, Islamic Azad University, Najafabad, Iran

Email: rasoolian@gmail.com

Abstract

Scientists have always been looking for ways to create an effective relationship between humans and the machine, so that this relationship is as close as possible to human relationships, since even the most sophisticated machines do not have any particular effect without human intervention. This association results from brain-generated neural responses due to motor activity or cognitive activity. Communication methods include muscle and non-muscle activities that create brain activity or brainwaves and lead to a hardware device to perform a specific task. BCI was originally designed as a communication tool for patients with neuromuscular disorders, but due to recent advances in BCI devices such as passive electrodes, wireless headset, adaptive software, and cost reduction, it has been used to link the rest of the body. The BCI is a bridge between the signals generated by thoughts in our brain and the machines. BCI has been a successful invention in the field of brain imaging, which can be used in a variety of areas, including helping motor activity, vision, hearing, and any damage that the body sustains. The BCI device records brain responses using invasive, semi-invasive and non-invasive methods including Electroencephalography (EEG), Magnetizhenophyllography (MEG), and Magnetic Resonance Imaging (MRI). Brain response using pattern recognition methods to control any translation application. In this article, a review of various techniques for extracting features and classification algorithms has been presented on brain data. A significant comparative analysis of existing BCI techniques is provided.

Keywords: Brain computer interface, Feature extraction, Signal classification, Electroencephalogram, Fuzzy inference system

1. Introduction

The human brain is the largest and most sophisticated member of human endocrinology, which consists of billions of neurons, and is considered as a multiprocessor system that receives information from different organs of our body, controlling our processing and activities. The human brain is an advantage, but the lack of a proper interface can lead to breakdown. Between man and machine is happening. If the musculoskeletal system does not respond well, BCI can be a brain-computer interface. Sometimes, it can lead to inappropriate functioning due to injuries to the brain and sometimes make the skeletal muscle system do not work well. BCI will in turn provide a platform for a wide range of applications [34]. New paradigms such as Neuroscience, Artificial Intelligence, Cognitive Science and the Brain and Computer Interface (BCI) are used to deepen the understanding of the brain. The further development of BCI to provide diagnosis for people with skeletal musculoskeletal disabilities can open the door for these people [6].

Figure 1 shows the performance of the BCI system relative to the classical performance of the brain. BCI is the discovery of progress in the field of brain mapping, which interprets the neuron language and uses this interpretation. BCI is a blessing for people who are exposed to brain damage and are limited by physical abnormalities. The brain, if used with BCI, can lead an injured person to normal life. BCI programs record brainwaves and send signals to a device that can carry out the expected work. The responsibility of BCI is to transfer electrophysiologic activities to device instructions that can lead to the continuation of normal life in people with physical disabilities. This is a straightforward brain-to-machine platform for people. BCI can, in addition to medical applications, be able to perform tasks beyond

traditional programs such as entertainment, play, experimentation and learning, translating thoughts into functions that are now as a vital issue known to be used [37].

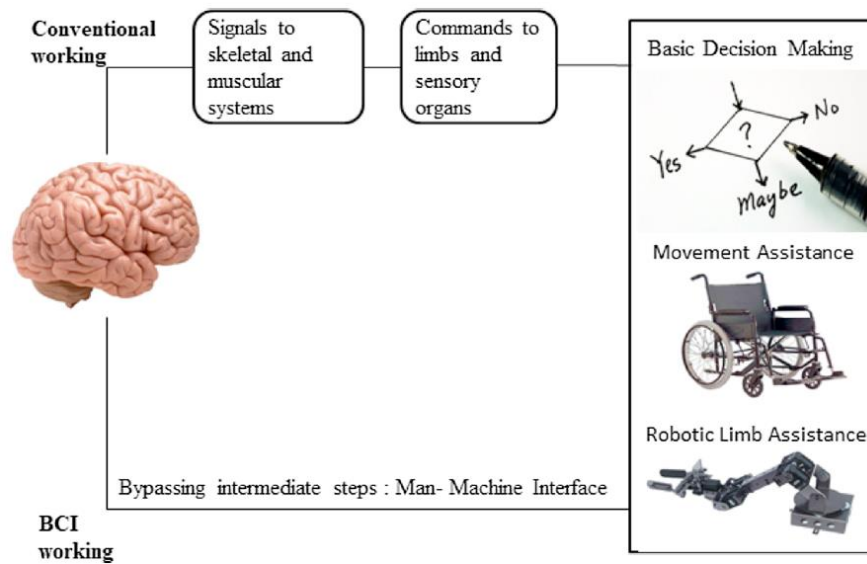


Fig. 1 execution of the command by the brain against the BCI [23].

The conventional BCI system includes a signal generation system, a signal processing technique, and an output device. Signaling can be done in three ways: aggressive, non-invasive, and semi-invasive. The invasive techniques include absorption of signals through the penetration of the micro-electrode in the brain. In semi-invasive techniques, electrodes are placed under the skin, but not in gray matter. Non-invasive techniques include inserting the electrode into the scalp without surgery. Some of the non-invasive techniques used to capture brain signals include electroencephalography (EEG) [5], magnetoencephalography (MEG) [8], and magnetic resonance imaging. Non-invasive techniques are widely used for research because these techniques are not prone to damage to brain tissues.

Brain signals have been amplified from well-known processed human signal processing devices for human use. Signal processing involves filtering, extracting features, and classifying brain potentials or brain signals. The important task of scientists and researchers is to eliminate contamination and extract useful information. Feature extraction involves the removal of noisy and artificial data for pure and non-infected data that can be used to develop BCI applications. Different extraction algorithms (also known as conversion transformations) are used to convert original data to a particular vector, such as Independent Component Analysis (ICA) [35], Common Spatial Patterns (CSPs) [27], linear filtering, Fast Fourier Transform (FFT), and Discrete Wavelet Transform (DWT). Select selected vectors by classifying algorithms such as Linear Discriminant Analysis (LDA), Support Vector Machines (SVMs) [30], Neural Networks (NNs), Fuzzy Inference Systems (FISs) [36], and many more. Classification is divided into desired classes. Finally, the signals processed by prosthetics, wheelchairs, electrical equipment, or computers are used.

The survey describes the follows: Section 2 explains various modes of signal acquisition, including techniques and devices available for recording brain signals. Section 3 reviews the various methods of feature extraction. Section 4 classifies the different types of signals. Section 5 contains our results and provides some research guidelines for the future.

2. Signal reception methods

The first step to using a brain signal to retrieve information is to obtain an appropriate signal. There are three types of signal acquisition techniques outlined below.

2.1. Invasive

This method of reading signals from within the gray matter of the brain. This necessitates the insertion of an electrode into the brain and, as a result, requires surgery, which can potentially lead to the extraction of the best quality of information, but is not used due to the resistance of the human body to any external culture.

2.2. Almost invasive

Reading the device signal outside the gray matter of the brain, which is expected to receive the signal strength clearly less powerful than the invasive methods and the possibility of being at risk of losing or damaging the brain.

Almost invasive methods are:

2.2.1. The penetration of micro-electrodes in the brain

Micro-electrodes penetrate into the gray-mattered material (the region in which the neuron is present) to produce higher-quality, high-power brain signals than non-invasive techniques. The challenge of this method is to obtain high-quality signals, such that the neural activity and the number of electrodes required to obtain a better signal and signal strength (record for a long time). Micro electrodes are used for the first time in brain signals of a monkey.

2.2.2 Electro Corticography (ECoG)

An invasive technique that requires the surgical cut of the skull to implant the device, but after implantation it can be used outside the operating room (extra ECoG). However, ECoG produces far more information than EEGs, because electrodes are placed directly on the brain surface. The level of detail of the information transmitted enables individuals to play computer games (Pong, Space Invaders) with mere mental efforts. However, ECoG only provides superficial vision in the brain. Its activity is regulated by neural networks associated with different brain structures, some of them with a greater depth of the brain, such as the thalamus sub unit and the hippocampus. Intrinsic electrodes have been created for the separation of these deep structures [1]. Kang [2] confirmed the ECoG function for two-dimensional tasks and how it functions from aggressive methods. In addition to the ECoG epidural (EECoG), a kind of ECoG provides more progress than ECoG. These methods provide rehabilitation for individuals with minor invasive states and significant accuracy. The electrode is located near the surface of the cortex (the outer layer of the nerve tissue). ECoG includes two types of electrode placement systems. The first system is a set of electrodes arranged equally on plastic strips or plastic silicone that can be changed. In order to improve spatial resolution, electrodes produce more density. The second system of electrical electrodes is separate on the spherical surface. The ECoG record has fewer effects than EEG, and this technique is less susceptible to contamination.

2.3 Non-invasive

This method is known as the most popular solution in which the electrodes are extracted from the skull to extract the signal. This method performs the quality of the signal at a low cost and ease of implementation. Non-invasive BCI methods include Electroencephalography (EEG), MEG, Magnetic Resonance imaging (MRI), functional magnetic resonance imaging (fMRI), and positron tomography (PET) release. The following figure compares some types of BCI in terms of aggression and performance range.

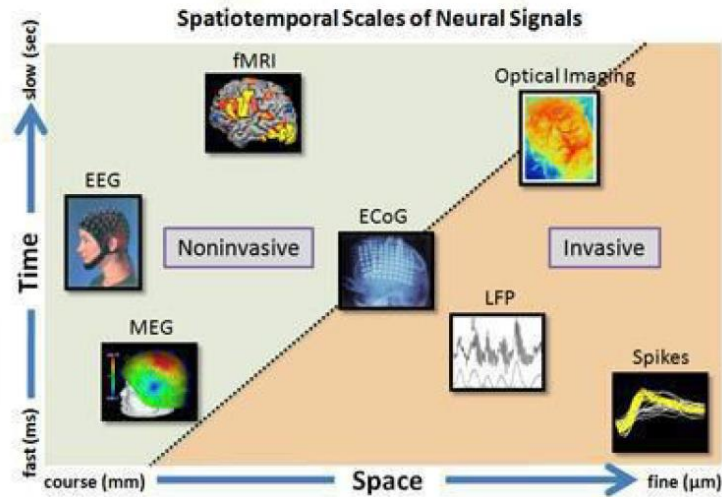


Fig. 2. Types of BCI In terms of speed [3].

Non-invasive methods are:

2.3.1 Electroencephalography

EEG is used to diagnose electrical abnormalities in the brain. Results in the output console are shown in waveforms. To get a signal from our brain, we use an electrode that attaches to the skull with the help of some sticks. This is an easy way to track the detection of separate brain regions. When performing various brain activities, there are physiological neuropsychic phenomena, some of which can be detected in the EEG signal. Some of these phenomena are dependent on an external stimulus, and some of them are created in the brain without extrinsic stimuli and during normal tasks and the ability to produce some others after a learning process [4]. EEG has many medical applications and many applications in the field of rehabilitation. In this section, EEG is used for non-invasive methods for signal reception. The advantage of this low-volume, low-cost, portable signal is that it does not have the potential risk of an offensive mode [5].

Some of the devices on the market include NeuroSky, Neuroscan, EMOTIV EPOC and Brain Products. Various studies of EEG signals by Ramadan et al. Have been investigated. [6]. At present, researchers are trying to create EEG signals that are reliable and offer high quality signals.

The brain signals are recorded by EEG devices based on frequencies called rhythmic brain activity or EEG rhythms [6]. Delta waves (1-4Hz) are commonly found in infants and during sleep. Theta waves (4-7 Hertz) are commonly found in rodents; these waves are also found during meditation, unconsciousness or drowsiness in humans. Alpha waves (8-13 milliseconds) are observed in humans in a quiet state with closed eyes. Beta waves (14 to 30 Hz) occur when a person caution, attention, or thinking. Gamma waves (> 30 Hz) are generated during voluntary movements or when some stimuli are given. Figures 3 to 6 show different EEG rhythms that record healthy conditions in an appropriate environment and record normal activities. Considering the design of an BCI system, we need to take into account some of the important features. The first feature is noisy or marginal, because the EEG signals have a weak to noise ratio. Second, always in BCI systems, feature vectors are often of great dimensions. In fact, several features are generally extracted from several channels and from several sections before joining a single-vector vector. The third feature contains timelines, in which brain activity patterns are generally related to specific EEG time variations. Finally, the BCI features are non-constant because EEG signals may vary over time, especially during sessions [7].

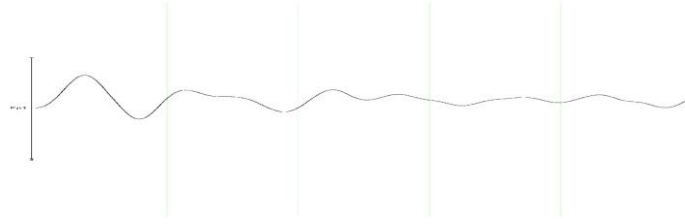


Fig. 3. Theta wave [6].

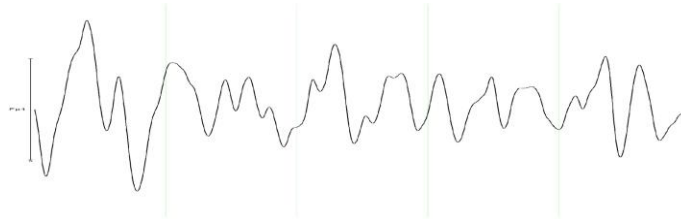


Fig. 4. Alpha wave [6].

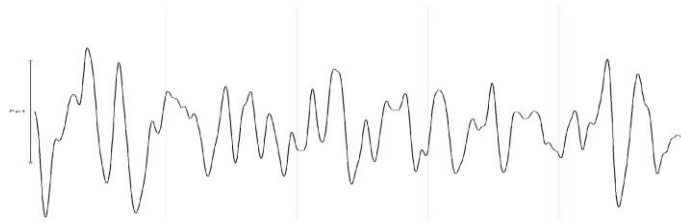


Fig. 5. Beta wave [6].

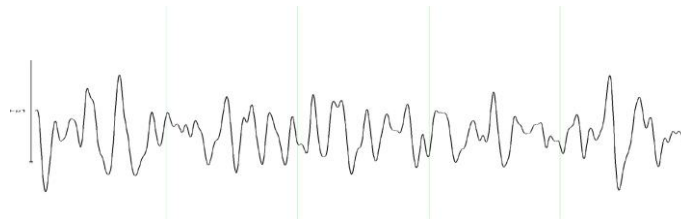


Fig. 6. Gamma wave [6].

2.3.2 Magnetic-magnetic fluography (MEG)

It was first used by David Cohen to measure the brain's electromagnetic waves from a healthy person and seizure. MEG generates magnetic circuits, which are the result of a nerve activity produced in response to an actuator, such as the EEG, and MEG, to capture the neuropsychological potentials produced by neurons, but in the form of a magnetic field. MEG has a clear spatial resolution and is not affected by artificial muscle movements in any way [8]. MEGs are based on superconducting interference devices (SQUIDS) introduced in the 1960s. SQUIDS are filled with large helium liquid units that maintain a system temperature of -269°C , with low impedance, low temperature. The SQUID device identifies and amplifies the magnetic field generated by neural activity.

Using MEG's activity, researchers have recorded differences between talking about some bilingual babies and a 11-month-old one [9]. In the recent study, a solution to the source reconstruction using MEG and EEG data has been developed by creating a Bayesian hierarchy algorithm. The data probability algorithm is used to reach the maximum speed using convergent laws, as well as auditory, visual, and facial processing data for simulation. The authors in this article considered only the location information for using the algorithm. Ford et al. [11] have performed statistical analysis of spatial-temporal data using MEG. They provided listening stimuli for their subject and found the difference

between continuous data from MEG for frequent and new stimuli. In this study, it was observed that cortical activity is more than frequent triggers for new stimuli.

2.3.3 Magnetic Resonance Imaging (MRI)

MRI provides information about the brain through full images using rays and radio waves. Along with these abnormalities, MRI helps to diagnose the causes of these disorders, and thus helps identify potential recovery methods. MRI is able to cope with the two sides of the brain, which in general can determine which side and part of the brain are damaged. MRI is capable of detecting abnormalities in the brain at an early stage, and typically provides brain images in high contrast, and thus works better than (CT).

2.3.4 Applied Magnetic Resonance Imaging (fMRI)

fMRI provides an extra feature of capturing specific brain sections, while the subject performs special tasks; therefore, we can use parts of the brain that help with changes in the blood flow during any work. Specific areas of the brain are active. fMRI shows strong evidence, since hemoglobin is a magnet in our brain, and so the magnetic field can detect activation, even with short-term simulation. Relative differences in response to hemoglobin in different parts of the body result in proper detection or inappropriate functioning of the regions.

In a study [12], fMRI was performed on preterm infants. This method was used to predict early damage to the brain and its neural growth. In another work, a new framework for improving the fMRI detection accuracy was used [13]. To increase the detection rate, the signal was extracted by converting the large volume of the brain into specific stimulus portions. The brain volume was cut for a specific point of time relative to the stimulus. Statistical analysis was applied at each time slice. Signals were used by a non-scheduled method of a non-standard time. For the analysis of fMRI data, some of the available software programs are GE's BrainWave [14], AFNI (Analysis of Functional Neuro Images), Brain Suite, and BrainVoyager.

2.3.5 Functional Near-Infrared Spectroscopy (fNIRS)

fNIRS uses light from the near-infrared spectrum (EM) to study the oxygenation and anaerobicity of the hemoglobin present in the brain. Hemoglobin Oxygenation and anaerobicity occurs in response to stimuli or during activity. The temperature measurement is done by three methods: continuous wave (CW), time interval (TR) and frequency domain (FD). fNIRS-CW is the most sophisticated system. The acquisition of fNIRS is done by placing an optode on a scalp. Optode includes a source and a tracker. The source of electromagnetic waves moves from the scalp to the brain and then transmitted to the detector. The change in intensity in the introduction of stimuli is recorded and analyzed using existing methods. Earning a signal with fNIRS is costly, but relatively cheaper than fMRI acquisition. Opri et al. Have developed a fNIRS-based BCI system in which the extraction of the characteristics and classification of data obtained after the motor vehicle tasks is performed. The fNIRS signals from the cerebral cortex have been obtained. After filtering raw data and eliminating noise, features such as mean, skewness, peak, and elongation were extracted. Later, a genetic SVM was used for data classification [15].

2.3.6 Single-Photon Emission Computed Tomography (SPECT)

SPECT is a nuclear medical method used to study the brain from gamma rays. SPECT gives a different view of how the brain functions. During the activity of the brain using SPECT, a radioactive substance is injected into the patient's body and scanned using the SPECT device. The SPECT device detects a radioactive substance that is absorbed by the brain in the patient's body. Radioactive materials allow physicians to see how the blood flow to the tissues and organs and which areas of the brain are inactive or active. SPECT produces an average brain activity within minutes. By reading these pictures, doctors can detect any reduction in brain activity.

2.3.7 Positron emission tomography (PET)

PET is another noninvasive method that measures brain function by infusing the positron material of the nuclear material. Ramadan et al. [6] used short-term radioactive drugs to monitor and diagnose any disorder in the metabolic activity of the human body. PET scan can analyze the amount of sugar consumed by the brain and analyze the information needed for both the above metabolic substances, that is, the use of a large amount of sugar or metabolism, including very small sugar, brain regions. At the same time, a positron camera is placed around the patient that provides crossover imaging of the tomography. PET generates more signal in any other way, so voltage correction is easy. It also has the benefits of achieving a higher sensitivity that is suitable for clinical admission.

The image forming process in PET with proper counting statistics is capable of performing accurate results with relatively simple algorithms. One of the main points of PET, confirmed by numerous studies, is the cost of installation and maintenance. PET is commonly used to diagnose brain disorders.

Table 1. Comparison between Various Modes of Acquisition

Non-invasive	Partially invasive	Highly invasive
Cost savings	Costly	Costly
Easily controlled	Difficult to monitor	Difficult to monitor
No medical training required	Need for medical education and assistance	Requires medical training and assistance
Poor space clarity	High spatial resolution	Higher spatial resolution
Without risk of infection	Risk of infection	Risk of infection and inflammation
Long-term recording	Short-term recording	Very shortterm recording

3. Feature extraction

The popularity of non-invasive techniques to collect signals avoids boring work and data filtering. This machine includes extraction of raw signals, including noise, and artifacts produced from the eyelid, muscle movement, hair, sweat and other factors. Figure 7 shows strips of different frequencies, their ranges and their location. Also, the results and meaning of the above, below, and normal items are specified for each detail.

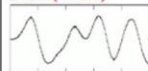
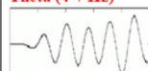
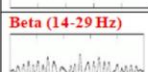
Frequency band	High	Low	Optimal	Location
Delta (<4 Hz) 	Brain injuries Learning problems Unable to think	Poor sleep Tiredness	Deep sleep Natural healing Immune system	Frontally in adults, Posteriorly in children
Theta (4-7 Hz) 	Depression Inattentiveness Hyperactivity	Anxiety Poor emotional awareness Stress	Creativity Emotional stability Intuition	Varies
Alpha (8-13 Hz) 	Daydreaming Relaxed Unable to focus	Anxiety Stress Insomnia OCD	Relaxation	Posterior regions, Occipital regions
Beta (14-29 Hz) 	Stress Anxiety Inability to relax	Day dreaming Poor cognition Depression	Conscious focus Problem solving	Typically in frontal regions
Gamma (>30 Hz) 	Anxiety Stress High Arousal	Depression Learning Disabilities	Learning Information processing Binding Senses	Somatosensory cortex

Fig. 7. Frequency band, frequency range, location, and importance [24].

3.1 Linear Filtration

Usually used to eliminate noise in the form of signals that do not belong to the frequency range of the brain signals. Linear filtering is essentially classified into low pass filters and high-speed filters. Artifacts are noise generated by the endogenous (muscle, eyes and other cardiovascular activities) or external (machine error). There are three techniques for artifact processing in gaining an EEG signal, as listed below.

- Avoiding Bug: Preventing Artifacts from Blind Artifact.
- Disapproval of flashing artifacts: abandoning contaminated experiments.
- Artifact Removal: The artifact created relies on its preprocessing methods.

The results show that the use of wet electrodes instead of dry electrodes reduces noise in the cable. The importance of massive electrical and muscular artifacts, simultaneously uncontrolled stimulus, requires online pre-processing and reduced number of electrodes. Compatibility filter and wavelet destruction are known as the appropriate time for EEG-BCI. Improving artificial purchasing techniques can help reduce the number of tests and subject training and simultaneously help to improve the extraction and classification of the feature.

3.2. Principal Component Analysis (PCA)

PCA [25, 31] is another method that maximizes the rate of data variance reduction. The PCA uses a conversion matrix that contains elements with a low variance. The transformation matrix A can be written as by the Eq (1):

$$A = \frac{1}{n} \sum (x_i - \mu) \cdot (x_i - \mu)^T, \quad (1)$$

Where x_i 's are elements of the N dimension dataset and n is the total number of elements in the original dataset can be written as according to Eq (2):

$$A \cdot Y = \Lambda \cdot Y, \quad (2)$$

where Y is the matrix containing eigenvector $y_1, y_2, \dots, y_n$ and Λ is the eigenvalue diagonal matrix with elements $\lambda_1, \lambda_2, \dots, \lambda_n$. Conventional PCA converts data to different lengths of lengths. To measure distances, functions such as the Euclid distance (ED) and dynamic time packing (DWT) are used. Li [26] applied the reduction of the EEG data dimensions and provides a novel, improved and effective PCA, which uses the covariance matrix to classify multiple time series variables (MTS) based on time - based variables.

3.3. Discrete Wavelet Transform (DWT)

Shen Sah used a spectral estimation technique in 1992 to express every general function as an infinite series of wavelets. It also allows the signal to be analyzed in various frequency bands, with different solutions. Decomposition is accomplished by using two sets of functions called scaling and wavelet functions that are related, respectively, to low pass filters and high passphrase filters. Then the coordinate space is reduced by the special vectors of the covariance matrix of the common covariance using each cluster. PCA various measurement methods like PCA, SVD, ICA and so on. The PCA uses most commonly used as the main component of the main components, the smaller it is and the more it maintains the MTS information. The use of CPCA, such as the high time complexity associated with MTS with variable-length and data mining accuracy, while the use of PCS results in a CPCA-based classification that is independent of the distance function. Hence, CCPCA has been used as the most effective method for classifying MTS theory.

3.4 Common Spatial Pattern (CSP)

CSP [27] designed a spatial or spatial filter to maximize the filtered signal disorders for classification. in order to perform CSP, the frequency and gaussian times are considered as known parameters. It converts multi - channel EEG data into a less space. This variance maximizes the classes for a two - class signal matrix. The following steps have been converted to an EEG matrix [16]. In the case of spatial filtering, the most common way to filter the common spatial pattern is the analysis of the independent components and the Laplace filter. A variety of reverse oscilloscope models allow you to recognize the actual projected in three-dimensional submarine spherical networks. The extracted features can be translated using different linear and nonlinear algorithms. First, we calculate the spatial correlation variance of the EEG naturally according to Eq (3). Then, calculate the composite spatial co-variance according to Eq (4) and Eq (5), and the projection matrix V is shown in Eq (6) and the main EEG is "X" in Finally, it transforms into Eq (7), and eventually the first and last pillar, P^1 , describes the largest one of the variance of a task and the smallest other variance according to Eq (8).

(1) Calculate the spatial correlation variance of the EEG as normal

$$C_K = \frac{X_K X_K^T}{\text{trace}(X_K X_K^T)}, \quad (3)$$

Where K represents the classes and $\text{trace}(x)$ the sum of diagonal values of x .

(2) Compute the composite spatial co-variance as

$$Cov = \sum C_K, \forall \text{classes} \quad (4)$$

$$C_K = V_K \lambda_K V_K^T, \quad (5)$$

Where V_K represents the eigenvector matrix and λ represents the diagonal matrix of the eigenvalue.

(3) The projection matrix V is denoted as

$$P = V^T U, \quad (6)$$

Where U is whitening transformation matrix $U = \sqrt{\lambda} V_0^T$. Using the projection matrix, the original EEG signal is reduced to uncorrelated components

$$W = PK \quad (7)$$

Where W is the EEG signals' source component, which includes common and specific components of different tasks.

(4) The original EEG "X" is finally transformed as

$$X = P^{-1}W. \quad (8)$$

Columns of P^{-1} are spatial patterns or can be called EEG source distribution vectors.

(5) The first and last column of P^{-1} , describes the largest variance of a task and the smallest other variance.

Herbert Ramsar et al. Proposed the above formulation to design a spatial filter for the classification of EEG data. The variance is just a few of the signals that are suitable for classification. The problem with the authors' proposed method is that if the signal is infected with an artificial signal, the filter design changes greatly. Design changes as a result of changes in covariance used to estimate spatial filters.

So there is a need for free EEG information with concepts. One of the limitations of using a CSP is that data does not provide filtered EEG signals. EEG data are not static, so we can not guarantee that the same information is extracted from the same issue at any given time. This is due to artifacts and other environmental conditions. Contaminated information affects covariance estimates and, in turn, causes additional problems.

There are various plugins for the CSP that can be easily applied to EEG data and have good performance. These are: (a) CSP permitted, (b) CSP specific gravity, and (c) fixed CSP.

3.5. Independent Component Analysis (ICA)

ICA is also a blind source separation, which, according to their statistical dependence, divides the data into various independent components. ICA provides good accuracy to remove the archive, but it is hard to get a component that contains only synthetic brain signals and also contains useful brain signals. therefore, ica improved using a combination of different methods. ICA - based algorithm was fabricated using temporal and spatial properties of

independent components (ICs) for speller P300 by Neng Xu et al. [28] the ICA follows a simple linear conversion approach. Suppose X is an initial signal matrix of dimension N , so that T decreases according to Eq (9) and Eq (10).

$$X = T.A \quad (9)$$

$$T = A^{-1}.X. \quad (10)$$

A number of studies and experiments have been conducted to remove artifacts using ICA, such as ICA in fMRI time series data, muscle removal, decay, blinking and auditory data using brain magnetic stimulation and electrosensophalography TMS -EEG. The limitation of the use of ICA is that there is no way to choose the IC automatically and the risk that the selected ICs have chosen. To improve the limitation, it is necessary to identify and remove ICA eye arthritis (OA). This method follows two steps. First, a low pass filter is applied to the EEG, and then the independently filtered components are analyzed one by one. The artifact pattern generated by the motion of the eye is analyzed by external diagnosis, and then the artifacts are identified and zeroed. Independent components are removed with artifact and EEG signals are retrieved meaningfully.

3.6 Fast Fourier Transform (FFT)

In FFT, signal characteristics are analyzed using the power spectral density estimation. The quadruple alpha, beta, gamma and theta frequencies contain major spectra for EEG. The characteristics of an EEG signal that needs to be analyzed are calculated by estimating the power spectral density (PSD) to selectively select the signal of the EEG samples.

4. Signal Classification

Lotte et al. [18] classified a generic categorization as a feature vector by selecting the most suitable classes to control a feature vector. A distinctive categorization such as a backup vector machine (SVM) is a supervised learning model that knows how to classify a feature vector.

- A static classifier, a multi-layer perceptron, can not, for instance, consider dynamic temporary information during classification. A dynamic classifier, for example: a hidden markov model, can be adapted to variable time variations.
- A stable classification, such as a linear diagnostic analyzer (LDA), is not affected by minor variations in the training data set. The unstable classification is, conversely, complex, and any change in the educational data set contributes significantly to its performance. An example is a multi-layer proptron.
- A classifier with irregularity suffers from the complexity of the training, because it does not recognize the time to stop teaching. A regular classifier is stronger on the other hand and therefore works better.

Figure 7 provides a detailed view of the type of classifiers used in BCI signal classification. Examples of each type are also included. This gives some kind of decision-making frontier, each of which can be traced. These are the following in the following section.

- Linear Classifiers: These classifiers use linear functions to binary classes. These are simple and popular. Examples of linear classifiers are the SVM and LDA.
- Nervous networks (NNs): They use artificial neurons to lead to nonlinear borders and are widely used for BCI systems, and the most popular layer of proptron is used.
- Bayesian Nonlinear Classifiers: These are not popular in real time due to the slow availability of BCI. They are productive and create non-linear decision boundaries.
- Classifiers of the nearest neighbor: The nearest neighbor classification [29] is simple categories that recognize classes with non - linear boundaries. An ordinary example is the Mahalanobis Distance (MD) [18], which is given by the Eq (11):

$$d(x) = (x - \mu_c)M^{-1}(x - \mu_c)^T \quad (11)$$

MD is a better option for classifying a sample data point, because the sample distance from the whole cluster is the same as in Eq (11) is given. The distance is another distance known as the Euclidean distance (ED), which diverges from the mean point regardless of the expansion of the data; therefore, MD is more accurate than ED.

• **Combined Classification:** A recent and popular way to improve classification, collection of classifications. Different combinational strategies reinforce each of the previous classifications, casting the simplest and most popular, and aggregate, in which the input into any meta-classification is the output of the previous classification. The most popular combination of classifications is the combination of outputs of different categories according to the application demand.

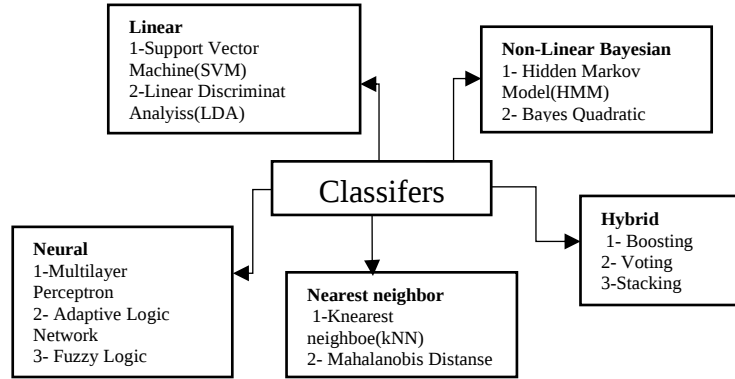


Fig. 8. Categories of signal classifiers with examples.

A brief introduction of some of the common methods of signal classification is given below.

4.1 Support Vector Machine (SVM)

SVM is an effective classifier based on the speed of training, indifference to training, resistance, and ability to overcome problems. The SVM classification is effectively the input vector X in the scalar value $f(X)$, which is indicated in Eq (12) [30].

$$f(x) = \sum_{i=1}^N a_i y_i K(X_i, X) + b \quad (12)$$

Where N is the number of support vectors, b represents bias, i is adjustable weight, y_i is a scalar in the range $\{-1, 1\}$, X_i are support vectors and $K(X_i, X)$ is the kernel.

However, high dimensional feature space can cause higher generalization errors in SVM. An improvement based on granular computing and statistical machine learning has been proposed in Guo and Wang [19]. GSVM can process fuzzy, corrupted, uncertain, incomplete and large data.

The highlighted features of GSVM are:

GSVM generalization function is better than traditional SVM because it divides the total space into the sub-sample space.

- Grain SVM has the ability to transform a linear non-separation into a linear separation problem.
- The GSVM allows parallel execution because the training data is independent of the various granules.
- GSVM. Provides better learning performance.
- It finds an approximate, but low-cost, simple solution rather than a precise, cost-effective solution.

4.2 Neural Networks

NNs are inspired by biological nervous systems that have features such as parallel computing, non-linearity, compatibility, responsiveness, and error tolerance. Inputs in NNs are called neurons that are related to (which can be positive or negative). Inputs are weighed through processing units. The processing units include a compilation section that ultimately connects to the output.

4.2.1 Convolutional Neural Network (CNN)

A CNN is another type of NN that is similar in architecture to MLP. This system divides the neurons into three dimensions: width, height, and depth. A taxonomy is according to CNN for classifying the EEG data using the P300 component of ERP has been used [32]. CNN can be created in a five-layer architecture: (1) input layer, (2) convolutional layer, (3) corrected linear unit layer, (4) aggregate layer, and (5) fully connected layer. Let L be the number of layers in the CNN architecture, x is the input vector, w is the weight vector, let M be the number of maps in the L layer, m represents the map m , J represents the number of neurons in the L layer, The number of electrodes, N_s is the number of signal values, and N_p is the number of partitions of the signal value. A number of Eq (13) to (16) apply in their work on different layers of CNNs and are used to update weights.

For Layer 1,

$$\sigma(1, M, J_1) = \sum_{i=0}^{N_e} x_{ij} w(1, M, i) + bias, \quad (13)$$

Where x_{ij} is the input vector from input layer L_0 , $0 \leq i < N_e$, and $0 \leq j < N_t$, N_t points considered for analysis according to Eq (14) and Eq (15).

For layer 2,

$$\sigma(2, M, J_2) = \sum_{i=0}^{N_s/N_p} x \left(1, M, J * \frac{N_s}{N_p} \right) w(2, M, i) + bias, \quad (14)$$

For layer 3,

$$\sigma(3, J_3) = \sum_{i=0}^{M_2} \sum_{k=0}^{N_p} x(2, i, k) w(3, i, k) + bias, \quad (15)$$

For layer 4,

$$\sigma(4, J_4) = \sum_{i=0}^{J_3} x(3, i) w(4, i) + bias. \quad (16)$$

A CNN-classifier used to classify EEG data using the P300 ERP component. CNN includes five layers and several maps. The output layer consists of a map that has two neurons. These two neurons represent two classes (class 1, P300 identified, class 2, no P300). To select the first order of CNN, the initial filters (weight) apply to the width and height of the input, then signal processing is done in the time domain. Here, the kernel is used as a vector, not as a matrix according to Eq (17) and Eq (18), and a linear sigmoidal function is applied between the hidden layer 1 and the hidden layer 2.

Hidden layer 2. Input signal convolution can be represented as [33].

$$f(\sigma) = 1.7159 \tanh\left(\frac{2}{3}\sigma\right), \quad (17)$$

Where, σ is the first deviation and classical sigmoidal function was used in between the last two hidden layers according to Eq (18).

$$f(\sigma) = \frac{1}{1 + \exp(-\sigma)}, \quad (18)$$

In the output layer, the grade score is calculated as according to Eq (19).

$$E[X] = \begin{cases} class1, & \text{if output(class1) > output(class2)} \\ class2 & \text{otherwise.} \end{cases} \quad (19)$$

CNN implements better classification than other categories because it uses hidden layers, but the number of layers that is used to obtain better classification cannot be determined. In another job, Cecotti et al. Using Eqs. (13) to (16), they classify the set of images: the human face (target) and others (non - objective) [34]. In their work, a CNN is embedded with a space filter. The filtering and classification are done on the ERP signals produced during the experiment. The learning method is based on maximizing the area under the curve (AUC). They compared CNN with SVM, BLDA (with the space filter and without). The finding suggests that CNN is better than the rest of the categories.

A CNN based on the AUC has no prior knowledge of the type of spatial filter, but requires prior information about the type of network architecture, so selecting the number of neurons and spatial filter depends on the previous experiment, which in turn affects the overall performance and performance of the network Affect. In order to achieve optimal performance for classification, CNN can be used with the legal choice of the neurons and hidden layers.

4.2.2 Perceptrons and Multilayer Perceptron.

Artificial neural networks offer a wide range of nonlinear categories, most notably MLP. Each neuron in the ANN mimics the neurons of the biological neural network, and the proper architecture can lead to effective classification; therefore, the MLP is a complex classification, and minor changes can lead to dramatic changes in the results. Experiments are not different based on having feature extraction (f), no extraction (nf), preprocessing (p), and preprocessing (np).

4.2.3 Probabilistic Neural Network (PNN)

The PNN was introduced by Spectv in 1990 and is based on the Bias rule, in which the PNN goal is a non-parametric predictor of probability density for obtaining optimal accuracy. The benefits of using the PNN are that it is easy to execute (much faster than recapture), has a parallel structure that converges optimally, has no minimum localization, and is in real time. The appropriate choice of smoothing parameter (σ) helps to correct the shape of the decision area.

The benefits of flexible neural networks include rapid training operations, internal parallel structures, ensuring the highest possible classification in the presence of adequate training data and continuing training with the use of new data without the need for re-networking. [21].

4.3. Fuzzy Inference System

In 1968, Zadeh suggests that in the real world, all classes or collections do not belong to a definite amount, such as yes or no, right or wrong, or real numbers; therefore, he introduced the concept of fuzzy sets: A fuzzy set is a definite boundless set, and the transition from a definite boundary to a flexible boundary is determined by membership function (MF). Using the benefits of a fuzzy set of flexible boundary conditions, many authors have used a fuzzy inference system in BCI applications. In another work, a fuzzy inference system was used to select the number of EEG channels for a hypothetical speech (in Spanish) [22]. It is therefore used as an artificial eye-catching and discrete WT as a feature extraction approach.

4.4 Neuro-Fuzzy Systems

Advantages of using NN and FIS. Its architecture is similar to NN and the inputs or weights (both) are fuzzy. The FNN identifies the fuzzy rules and specifies the membership function by adjusting the connection weights. Many applications use FNN-1, FNN-2 and FNN-3. A fuzzy fuzzy fuzzy system is trained by a learning algorithm derived from neural network theory. The learning method (real) acts on local information and only creates local changes in the basic fuzzy system.

A neuro-fuzzy system can be seen as a neural network of the 3-layer feeder. The first layer shows the input variables, the middle layer (hidden) represents fuzzy rules, and the third layer represents the output variables. Fuzzy sets are encoded as fusion weights. To display a fuzzy system, this does not require a learning algorithm to apply. However, this could be convenient, because it reflects the flow of data processing input and learning within the model.

Sometimes a 5-layer architecture is used, where fuzzy sets appear in units of the second and fourth layers.

A neuro-fuzzy system can always be interpreted as a system of fuzzy rules (for example, before, during and after learning). It also creates the possibility of creating a system from educational information from the beginning, because it can be analyzed using fuzzy rules with prior knowledge. All fuzzy neural models do not specify the learning methods to create fuzzy rules.

The learning method of a neuro-fuzzy system takes into account the semantic properties of the basic fuzzy system. This result in possible constraints is applicable to system parameters. All fuzzy neural methods do not have this feature.

An approximate neural fuzzy system is a subsequent (unknown) function, which is partially defined by educational data. Fuzzy rules that are coded within the system are vague examples and can be trained as examples of data. A nerve fuzzy system should not be known as a specialist system (fuzzy) and has no connection with fuzzy logic in the limited sense.

5. Conclusion

The brain and computer interface creates a way that enables users to easily control a computer through their thoughts. An interdisciplinary BCI is considered as an area for research in various aspects, such as the understanding, acquisition, and processing of brain signals. BCI research includes biology of psychology and neuroscience, engineering, computer science and applied mathematics. In this paper, we present a comprehensive review of each BCI stage. The first phase of BCI receives a brain signal. There are three types of signal acquisition systems: non-invasive, semi-invasive and aggressive. Getting an aggressive signal involves placing microelectrode and electrode chips under the scalp through surgery. Non-invasive technique eliminates brain potential by inserting a metallic electrode on the scalp (as in EEG) or recording brain activity and blood flow through specific devices (such as MEG, fMRI, etc).

These signal acquisition methods record different types of brain potentials, such as those generated by motor activity, cognitive activity, eye movement, or stimulus. Researchers prefer non-invasive techniques to aggressive methods because they are not prone to damage. The only limitation in non-invasive methods is that the resolution of the signals in the invasive methods is low. Future work can be developed to develop brain signaling devices that have lower density electrodes and more clearly. The second stage involves processing brain signals. In this paper, various extraction algorithms and classification are mentioned.

Extracting features In order to extract useful signals and remove artifacts produced by eye movement, muscle movements, features extraction techniques including Linear Filter, CSP, PCA, ICA, FFT and DWT. ICA is best suited for the removal of artifacts, and has been widely used by various researchers. CSP systems and its variants are used to filter the spatial signals of the brain. The PCA helps to convert the feature space, and DWT helps to extract time and time information from raw signals. Categorical algorithms such as LDA, SVM, NNs, and fuzzy inference systems are applied to properties acquired using feature extraction techniques.

This paper presents a summary of these classification methods and uses these techniques to provide recent information. Common BCI systems use diagnostic models for classification.

However, researchers are interested in deep learning methods such as deep belief networks, CNN, and a combination of different classification algorithms. The BCI system is a useful system that has good coordination among all of these sectors. The main goal of the BCI research is to provide a better communication approach; However, the methods used to achieve this goal can be different.

REFERENCES

- [1] Jebari, Karim. "Brain machine interface and human enhancement—an ethical review." *Neuroethics* 6, no. 3 (2013): 617-625.
- [2] Rainer Kang, J. (2017). Signal acquisition in brain-computer interface. The University of Auckland.
- [3] Naderi, p. (2016). The brain and computer interface, applications and its types, Convergent Technology Site.
- [4] Kheradpisheh, S, R. Nowzari Dalini, A. Ganjtabesh, M. (2016). A review of the brain and computer interfaces, *Journal of Computational Sciences*.
- [5] Bamdad, M., Zarshenas, H., Gerilo, H. (2013). A review of the uses of the BCI system in rehabilitation science with a rehabilitation approach to the peripheral environment, *Journal of Research in Rehabilitation Sciences*, Vol. 9, No. 6.
- [6] Rabie A. Ramadan and Athanasios V. Vasilakos. 2017. Brain computer interface: Control signals review. *Neurocomputing* 223 (2017), 26–44.
- [7] Fouad, M. M., Amin, K. M., El-Bendary, N., & Hassanien, A. E. (2015). Brain computer interface: A review. In *Brain-Computer Interfaces* (pp. 3-30). Springer, Cham.
- [8] Peter J. Uhlhaas, Peter Liddle, David Linden, Anna C. Nobre, Krish D. Singh, and Joachim Gross. 2017. Magnetoencephalography as a tool in psychiatric research: Current status and perspective. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (2017).
- [9] Naja Ferjan Ramírez, Rey R. Ramírez, Maggie Clarke, Samu Taulu, and Patricia K. Kuhl. 2017. Speech discrimination in 11-month-old bilingual and monolingual infants: a magnetoencephalography study. *Developmental Science* 20, 1 (2017), e12427.
- [10] Chang Cai, Kensuke Sekihara, and Srikantan S. Nagarajan. 2018. Hierarchical multiscale Bayesian algorithm for robust MEG/EEG source reconstruction. *NeuroImage* 183 (2018). DOI:<https://doi.org/10.1016/j.neuroimage.2018.07>.
- [11] Talitha C. Ford, Will Woods, and David P. Crewther. 2018. Magnetoencephalography reveals an increased nontarget P3a, but not target P3b, that is associated with high non-clinical psychosocial deficits. *Psychiatry Research: Neuroimaging* 271 (2018), 1–7. DOI:<https://doi.org/10.1016/j.psychres.2017.11.012>.
- [12] Charlotte Herzmans, Leire Zubiaurre-Elorza, Conor J. Wild, Annika C. Linke, Victor K. Han, David S. C. Lee, and Rhodri Cusack. 2017. Using functional magnetic resonance imaging to detect preserved function in a preterm infant with brain injury. *The Journal of Pediatrics* 189 (2017), 213–217.
- [13] Niels Janssen, Juan A. Hernández-Cabrera, and Laura Ezama Foronda. 2018. Improving the signal detection accuracy of functional magnetic resonance imaging. *NeuroImage* 176 (2018), 92–109.
- [14] GE Healthcare Systems. 2018. BrainWave. Retrieved March 2, 2018 from http://www3.gehealthcare.com/en/products/categories/advanced_visualization/applications/brainwave.
- [15] Farzan Majeed Noori, Noman Naseer, Nauman Khalid Qureshi, Hammad Nazeer, and Rayyan Azam Khan. 2017. Optimal feature selection from fNIRS signals using genetic algorithms for BCI. *Neuroscience Letters* 647 (2017), 61– 66.
- [16] Ramoser, H., Muller-Gerking, J. and Pfurtscheller, G., 2000. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE transactions on rehabilitation engineering*, 8(4), pp.441-446.
- [17] Zikuan Chen and Vince D. Calhoun. 2016. Task-evoked brain functional magnetic susceptibility mapping by independent component analysis (ICA). *Journal of Neuroscience Methods* 261 (2016), 161–171. DOI:<https://doi.org/10.1016/j.jneumeth.2016.01.007>.
- [18] Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for eeg-based brain–computer interfaces. *Journal of Neural Engineering*, 4(2), R1.
- [19] Guo, H., & Wang, W. (2017). Granular support vector machine: A review. *Artificial Intelligence Review*, 1–14.

- [20] Blankertz, B., Kawanabe, M., Tomioka, R., Hohlefeld, F., Müller, K.-R., & Nikulin, V. V. (2008). Invariant common spatial patterns: Alleviating nonstationarities in braincomputer interfacing. *Advances in Neural Information Processing Systems*, 113–120.
- [21] Designing a brain-computer communication system using Moran's emotional learning model, (2013). The first national conference on the application of intelligent systems (soft computing) in science and technology, Quchan, Islamic Azad University, Quchan Branch, https://www.civilica.com/Paper-AISST01-AISST01_120.html
- [22] Alejandro A. Torres-García, Carlos A. Reyes-García, Luis Villaseñor-Pineda, and Gregorio García-Aguilar. 2016. Implementing a fuzzy inference system in a multi-objective EEG channel selection model for imagined speech classification. *Expert Systems with Applications* 59 (2016), 1–12.
- [23] Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G. and Vaughan, T.M., 2002. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6), pp.767-791.
- [24] Wu, D., King, J.-T., Chuang, C.-H., Lin, C.-T., & Jung, T.-P. (2017). Spatial filtering for eeg-based regression problems in brain-computer interface (bci). *IEEE Transactions on Fuzzy Systems*.
- [25] Harold Hotelling. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24, 6 (1933), 417.
- [26] Hailin Li. 2016. Accurate and efficient classification based on common principal components analysis for multivariate time series. *Neurocomputing* 171 (2016), 744–753. DOI:<https://doi.org/10.1016/j.neucom.2015.07.010>.
- [27] Zoltan J. Koles, Michael S. Lazar, and Steven Z. Zhou. 1990. Spatial patterns underlying population differences in the background EEG. *Brain Topography* 2, 4 (1990), 275–284.
- [28] Xu, N., Gao, X., Hong, B., Miao, X., Gao, S. and Yang, F., 2004. BCI competition 2003-data set IIb: enhancing P300 wave detection using ICA-based subspace projections for BCI applications. *IEEE transactions on biomedical engineering*, 51(6), pp.1067-1072.
- [29] Kuncheva, L. I., & Jain, L. C. (1999). Nearest neighbor classifier: Simultaneous editing and feature selection. *Pattern Recognition Letters*, 20(11–13), 1149–1156.
- [30] Ahmadi, A., Dehzangi, O., & Jafari, R. (2012). Brain-computer interface signal processing algorithms: A computational cost vs. accuracy analysis for wearable computers. 2012 Ninth international conference on wearable and implantable body sensor networks (BSN) (pp. 40–45). IEEE.
- [31] Karl Pearson. 1901. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2, 11 (1901), 559–572.
- [32] Cecotti, H. and Graser, A., 2010. Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE transactions on pattern analysis and machine intelligence*, 33(3), pp.433-445.
- [33] Cecotti, H., Eckstein, M.P. and Giesbrecht, B., 2014. Single-trial classification of event-related potentials in rapid serial visual presentation tasks using supervised spatial filtering. *IEEE transactions on neural networks and learning systems*, 25(11), pp.2030-2042.
- [34] Sebastian Halder, Andreas Pinegger, Ivo Kathner, Selina C. Wriessnegger, Josef Faller, Joao B. Pires Antunes, Gernot R. Muller-Putz, and Andrea Kubler. 2015. Brain-controlled applications using dynamic P300 speller matrices. *Artificial Intelligence in Medicine* 63, 1 (2015), 7–17.
- [35] H. Santosa and K. S. Hong. 2013. An fNIRS-based BCI for mental arithmetic task using ICA. In 2013 International Conference on Robotics, Biomimetics, Intelligent Computational Systems. IEEE, 219–223.
- [36] Thanh Nguyen, Abbas Khosravi, Douglas Creighton, and Saeid Nahavandi. 2015. Fuzzy system with tabu search learning for classification of motor imagery data. *Biomedical Signal Processing and Control* 20 (2015), 61–70.
- [37] Douglas Creighton Thanh Nguyen, Abbas Khosravi, and Saeid Nahavandi. 2015. EEG signal classification for BCI applications by wavelets and interval type-2 fuzzy logic systems. *Expert Systems with Applications* 42, 9 (2015), 4370–4380. DOI:<https://doi.org/10.1016/j.eswa.2015.01.036>.