

# Survey on Brain-Computer Interface: An Emerging Computational Intelligence Paradigm

ANNUSHREE BABLANI and DAMODAR REDDY EDLA, National Institute of Technology Goa, Goa, India

DIWAKAR TRIPATHI, Madanapalle Institute of Technology and Science, Madanapalle, Andhra Pradesh, India

RAMALINGASWAMY CHERUKU, Mahindra École Centrale, Hyderabad, India

A brain-computer interface (BCI) provides a way to develop interaction between a brain and a computer. The communication is developed as a result of neural responses generated in the brain because of motor movements or cognitive activities. The means of communication here includes muscular and non-muscular actions. These actions generate brain activities or brain waves that are directed to a hardware device to perform a specific task. BCI initially was developed as the communication device for patients suffering from neuromuscular disorders. Owing to recent advancements in BCI devices—such as passive electrodes, wireless headsets, adaptive software, and decreased costs—it is also being used for developing communication between the general public. The BCI device records brain responses using various invasive and non-invasive acquisition techniques such as electrocorticography (ECoG), electroencephalography (EEG), magnetoencephalography (MEG), and magnetic resonance imaging (MRI). In this article, a survey on these techniques has been provided. The brain response needs to be translated using machine learning and pattern recognition methods to control any application. A brief review of various existing feature extraction techniques and classification algorithms applied on data recorded from the brain has been included in this article. A significant comparative analysis of popular existing BCI techniques is presented and possible future directives are provided.

CCS Concepts: • **General and reference** → **Surveys and overviews**; • **Human-centered computing** → *Interaction design theory, concepts and paradigms*; • **Computing methodologies** → *Machine learning*; • **Applied computing** → *Bioinformatics*;

Additional Key Words and Phrases: Brain-computer interface, classification, electroencephalogram, feature extraction, fuzzy inference system

## ACM Reference format:

Annushree Bablani, Damodar Reddy Edla, Diwakar Tripathi, and Ramalingaswamy Cheruku. 2019. Survey on Brain-Computer Interface: An Emerging Computational Intelligence Paradigm. *ACM Comput. Surv.* 52, 1, Article 20 (February 2019), 32 pages.

<https://doi.org/10.1145/3297713>

Authors' addresses: A. Bablani, National Institute of Technology Goa, Goa, Farmagudi, Ponda, Goa, 403401, India; email: annubablani@nitgoa.ac.in; D. R. Edla, National Institute of Technology Goa, Goa, India; email: dr.reddy@nitgoa.ac.in; D. Tripathi, Madanapalle Institute of Technology and Science, Madanapalle, Andhra Pradesh, Madanapalle, India; email: diwakarnitgoa@gmail.com; R. Cheruku, Mahindra École Centrale, Hyderabad, India; email: rmlswamygoud@gmail.com. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2019 Association for Computing Machinery.

0360-0300/2019/02-ART20 \$15.00

<https://doi.org/10.1145/3297713>

## 1 INTRODUCTION

The human brain is largest (about 3 pound) and most complex among all the human organs, consisting of billions of neurons. It is a multiprocessing system that receives information from our peripherals, processes it and controls our actions accordingly. The human brain has a complex structure and unmatched computational capacity, including the capability of multiprocessing and learning. Therefore, it has always attracted researchers since the early ages. Various new paradigms—such as neuroscience, artificial intelligence, cognitive science, and brain-computer interface (BCI) have been developed to understand the brain in more depth. The inspiration for development of BCI come from an urge to provide social recognition to individuals who are suffering from some neuromuscular disabilities. BCI is a system that translates thoughts and provides an interface for communicating with the outside world. Recent advancements in BCI have made it possible to understand the functions and neural communication inside the brain. The study of the brain not only has helped researchers in the medical field but also in the field of engineering. In addition to recording and displaying of brain activity, BCI allows the user to control programs such as video games, computational software, spellers [55, 89], web browsers [64], and thought translation devices [19]. It is a wide area of study and requires knowledge of computer science engineering, neuroscience, psychology, signal processing, clinical rehabilitation.

A typical BCI system includes a signal acquisition system, signal processing techniques, and an output device, as shown in Figure 1. Signal acquisition can be performed in three ways: invasive, non-invasive, and semi-invasive. Invasive techniques involve signal acquisition via penetrating micro-electrodes in the dura matter of the brain. In semi-invasive approaches, electrodes are placed beneath the scalp but not in the gray matter. Non-invasive techniques involve placing of electrodes on the scalp without surgery. Some of the non-invasive techniques used to record brain signals are electroencephalography (EEG), magnetoencephalography (MEG), and magnetic resonance imaging. Non-invasive techniques are extensively used for research, as these techniques are not prone to any damage to the brain tissues.

The brain signals acquired from signal acquisition devices are processed, amplified, and converted into forms recognizable to humans using amplifiers and converters. Signal processing involves filtering, feature extraction, and classification of brain potentials or brain signals. Raw brain data is generally contaminated with motor-muscular artifacts. A major task left with scientists and researchers is to remove that contamination and extract useful data. Feature extraction involves noise and artifact removal to get pure, non-contaminated data that can be used for developing BCI applications. Various feature extraction (also known as feature transformation) algorithms are available to transform the original data into a specified feature vector, such as Independent Component Analysis (ICA) [70], Common Spatial Patterns (CSPs) [88], Principal Component Analysis (PCA) [122], and Wavelet Transform (WT) [107]. The selected feature vectors are classified into desired classes by applying classification algorithms such as Linear Discriminant Analysis (LDA) [51] Support Vector Machines (SVMs) [35], Neural Networks (NNs) [111], Fuzzy Inference Systems (FISs) [177] and many others. Finally, the processed signals are used by prosthetic devices, wheelchairs, electrical equipment, or computers.

This survey is organized as follows. Section 2 gives a brief description of the brain regions and their behavior. Section 3 explains various modes of signal acquisition, including the techniques and devices available to record brain signals. Section 4 gives an overview of the different types of signals generated from the brain and recorded via acquisition devices. Section 5 reviews the various feature extraction methods. Section 6 presents some of the important classification algorithms used in developing the BCI system. Section 7 contains our conclusions and provides some research directives for the future.

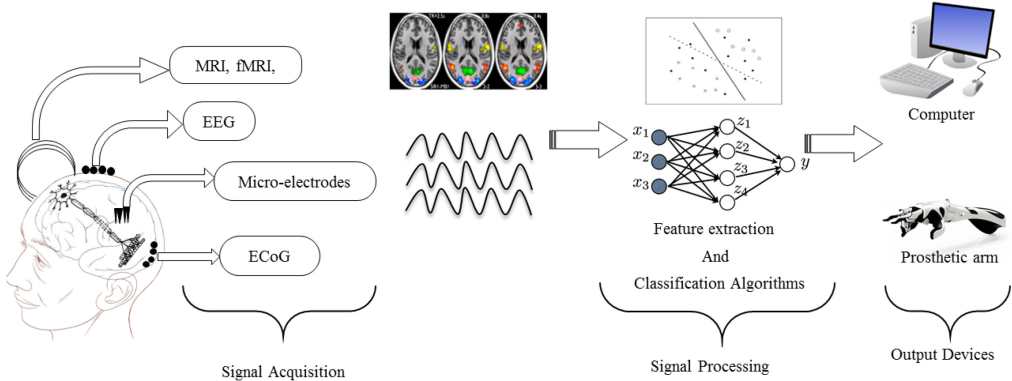


Fig. 1. Brain-computer interface system.

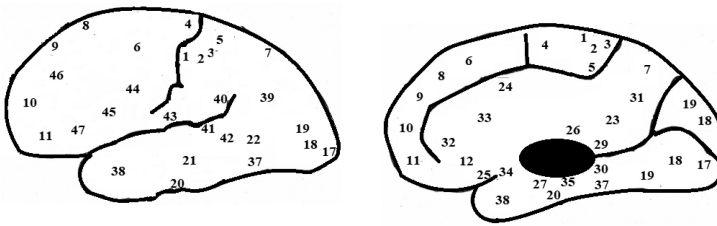


Fig. 2. Brodmann areas.

## 2 THE BRAIN AND FUNCTIONALITY OF ITS REGIONS

The brain is a multiprocessing system that receives information from the human body, processes it, and controls body actions accordingly. Different parts of the brain perform different cognitive functions, which are further discussed in this article. The brain can be macroscopically classified into two regions: the cerebral cortex and subcortical region. Another classification of the brain on the basis of cognitive functionality of different areas has been done by Brodmann [56], who divided the brain into various regions that are known as *Brodmann areas*, as shown in Figure 2 [7].

### 2.1 Cerebral Cortex

The cerebral cortex is the outer covering of gray matter, which is divided into two cerebral hemispheres: left hemisphere and right hemisphere. The functions of the left hemisphere include group coordination and communication; sensation; vision; control of the right side of body, linear thinking (step-by-step progression); verbal memory (thinking in words), including Wernicke’s speech area and Broca’s speech area (Figure 3); and goal-directed linear planning. The functions of the right hemisphere include environmental awareness, sensation, vision, control of the left side of body, non-linear thinking, visuospatial memory (thinking in pictures), mental manipulation of relationships, complex or emotional decisions, and error detection.

Each hemisphere is partitioned into four brain lobes: frontal, occipital, parietal, and temporal.

**2.1.1 Frontal Lobe.** The frontal lobe occupies the front part of the brain. Its primary functions are organizing, planning, social skills, problem solving, decision making, emotional control (Brodmann areas 10, 11, and 47), movement planning (Brodmann area 6), control of eye movement (Brodmann area 8), A-not-B task, object task [143], and comparing two items from memory (Brodmann areas 9 and 46). Brodmann areas 9 and 46 are part of frontal lobe, which together are called

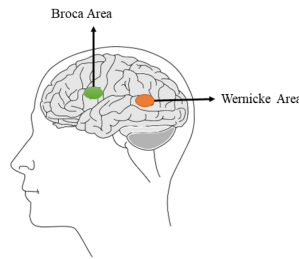


Fig. 3. Broca and Wernicke speech areas.

the Dorsolateral Prefrontal cortex (DLPFC). A person with a lesion in the DLPFC will not be able to identify objects seen a few hours earlier. The left hemisphere of the DLPFC is activated during any verbal working memory task, whereas the right hemisphere is active for visual working memory tasks [143], [115]. Shackman et al. [139] showed that the DLPFC is involved in telling the truth/lying and language processing (Broca speech area). The Broca speech area involves semantic analysis of language, that is, how to use words in a sentence. People with a lesion in the Broca area can speak but will not be able to form sentences properly (Brodmann areas 44 and 45).

**2.1.2 Occipital Lobe.** The occipital lobe is located on the back side of the skull (Brodmann areas 17, 18, and 19). The occipital lobe is involved in the processing of visual information. Electrodes are placed on the occipital lobe to record neuronal activity generated by a given visual stimulus.

**2.1.3 Parietal Lobe.** The parietal lobe is located immediately above the occipital lobe and behind the frontal lobe. It is responsible for spelling, perception, object manipulation, sense of touch (Brodmann areas 1, 2, and 3), high processing tasks (Brodmann areas 5 and 7), and visual motor coordination (Brodmann area 7).

**2.1.4 Temporal Lobe.** The temporal lobe is located behind the ears on both hemispheres of the brain. It is responsible for basic hearing function (Brodmann areas 41 and 42), recognizing faces and numbers (Brodmann area 20), memory formation and optimization during sleep (Brodmann areas 28 and 34), and understanding words (Brodmann area 22, also known as Wernicke's area) [56]. These areas are affected primarily in patients suffering from Alzheimer's disease.

## 2.2 Subcortical Region

The subcortical region is divided into two parts: the cerebrum and *cerebellum*. The cerebrum comprises the thalamus, brain stem, and limbic system. These are responsible for vital functions such as digestion, breathing, heart rate, and information transfer from the cerebrum to cerebellum. The thalamus processes information before transferring it to the cerebellum. The limbic system, commonly known as the *mini-brain*, comprises the amygdala, hippocampus, hypothalamus, cingulate gyrus, and fornix system. The hypothalamus regulates the endocrine system while the hippocampus stores memory. Functions of the limbic system include controlling emotional behavior, eating habits, anger, and sadness. (Brodmann areas 24, 32, and 33) [56].

## 3 MODES OF SIGNAL ACQUISITION

The brain's behavior can be understood by mapping its activities. Activities in the brain are generated when one neuron transfers a message to other neuron through a synapse. It does so by transferring ions, which cause a change in the electric potential inside the brain. The change in potential is generated because of the flow of ions such as sodium and potassium ions in brain cells. These ions are present inside and outside the brain cell; the difference in the concentration of ions

Table 1. Comparison between Various Modes of Acquisition

Non-invasive	Partially invasive	Highly invasive
Cost efficient	Costly	Costly
Easily monitored	Difficult to monitor	Difficult to monitor
No medical training required	Requires medical training and assistance	Requires medical training and assistance
Poor spatial resolution	High spatial resolution	Higher spatial resolution
No risk of infection	Risk of infection	Risk of infection and inflammation
Long-term recording	Short-term recording	Very shortterm recording

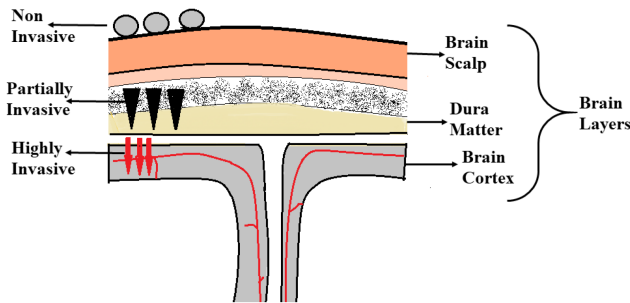


Fig. 4. Position of electrodes of different acquisition techniques on various brain layers.

causes their flow from higher concentration to lower concentration of ions present in cell. The difference in potential generated is known as resting potential ( $-80\text{mV}$ ) [91]. The flow of ions is initiated when certain actions are performed by an individual. The recording of this potential difference in brain cells caused by various activities can be done using different techniques. A few of these approaches are discussed in the following section. A comparison between various features of the acquisition techniques is provided in Table 1 and Figure 4 [6] displays how electrodes are placed on different layers of the brain.

### 3.1 Invasive Mode

The invasive mode of acquisition allows recording of brain signals by inserting electrodes surgically inside the brain.

**3.1.1 Penetrating Micro-Electrode in the Brain.** Micro-electrodes penetrate the brain gray matter (area where neurons are present) to record brain signals of higher quality and greater strength than in non-invasive approaches. The challenges for capturing good-quality signals are high penetration power so that it is easy to capture neuronal activities and number of electrodes required for better signal acquisition and durability (recording for a long interval of time). Micro-electrodes were first used to record brain signals from a monkey. Later, Jose M. Delgado [41] used this method to record human brain signals. A few variants of micro-electrodes are as follows:

- A *Micro-Wire Array* is made using wires of stainless steel or platinum or iridium alloys to improve their durability. These wires can be customized, as they are handmade. An increase in the number of electrodes increases the number of wires needed, which, in turn, increases cost and time [153].
- A *Micro-Fabricated Array* is fabricated on a single device, due to which cost and time of customization doesn't see any significant increase. It also enhances spatial resolution. Usually, micro-electrodes are fabricated with polymers [153] or with silicon [65].

Table 2. Comparison of Existing Electrode Placement Systems

Electrode placement system	Number of electrodes	Distance from nasion and inion area (%)	Distance between electrodes (%)
10-20 system	19–21	10	20
10-10 system	64–85	10	10
10-5 system	320–329	10	5

- *Polymer-Based Arrays* are less prone to chronic injury and can be easily integrated as cables or wires. Tooker et al. [153] have developed a polymer-based array that is capable of expanding, providing high-quality signals and long-term recordings.
- *Silicon-Based Arrays* have good strength to penetrate into the tissue, which make them the commonly used electrode [65].
- *Benzocyclobutene (BCB)* is a biopolymer [95] with more flexibility and higher compatibility than other polymers and also provides long-term recordings.

**3.1.2 Electrocorticography (ECoG).** ECoG is also known as Intracranial electroencephalography (iEEG). It uses electrodes that are implanted at a location close to the cortical surface (outer layer of neural tissue). Rather than penetrating into the cortical area, electrodes are placed on the brain cortical area. ECoG provides a higher spatial resolution, less noise, and higher bandwidth as compared to EEG [15]. ECoG involves two types of electrode placement systems. The first system has an array of equally spaced electrodes that are placed on strips or grids of silicone plastic, which can be altered. To improve spatial resolution, the electrodes are arranged more densely. The second system arranges individual wired electrodes over the exposed cortical surface. ECoG recordings have less artifacts than EEG (to be explained later) and this technique is less susceptible to infection.

## 3.2 Non-Invasive Mode

Non-invasive BCI techniques involve acquisition of brain signals without harming the brain tissues. Various non-invasive techniques have been adopted to acquire brain signals, some of which are explained in this article.

**3.2.1 Electroencephalography (EEG).** It is the most commonly used non-invasive technique for acquiring the electrical activity generated by brain cells (neurons or nerve cells and glial cells). The signals are recorded by placing metal electrodes on the scalp. The metal electrodes are mostly built from German silver, which is an alloy of copper, nickel, and zinc. Polytetrafluoroethylene (Teflon) is used for coating the wires and metal electrodes. An electrolyte gel or paste is applied either on the electrodes on the scalp to initiate conductivity between them. The paste is composed of lanolin and chloride ions, which helps in electrical conduction. The electrodes are positioned on the brain scalp using the standard 10-20 electrode placement system. In addition, there are 10-10 and 10-5 electrode placement systems for signal acquisition [82]. The 10-10 electrode placement protocol has been extensively used recently, as it provides more detailed brain signals. The difference between 10-20, 10-10, and 10-5 electrode systems is distance between each electrode from the nasion and inion areas, which is given in Table 2. Figure 5 shows the electrode map of the 10-20 and 10-10 electrode systems.

The 10-5 system is covered in [82]. As discussed in Section 2, each brain area works differently. In order to record a particular brain activity, the knowledge of placing an electrode on that part of brain is required. Hence, we have listed the electrodes to be placed on various Brodmann areas in Table 3.



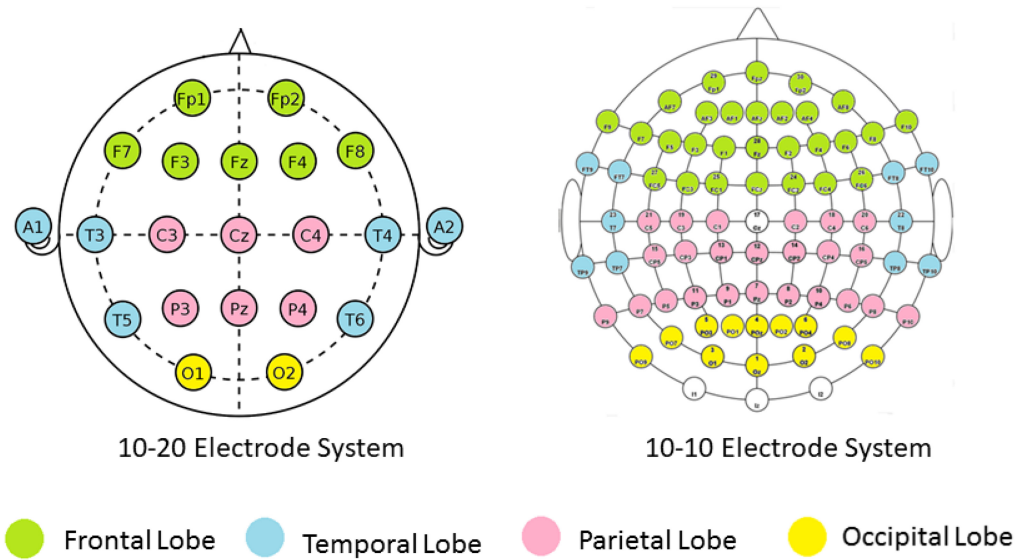


Fig. 5. The electrode systems usually followed for EEG data collection.

EEG signals are the fluctuations occurring due to the electrical potential generated as a response to the neuron activity inside the brain and recorded from the scalp [159]. Electrical impulses from a living brain of a rabbit and monkey were recorded for the first time by Richard Caton in 1875. He placed the electrodes in two positions: on the gray matter and scalp [117]. In 1890, Adolf Beck studied the brain activity of animals in response to sensory simulation [33]. In 1913, Napoleon Cybulski studied the flow of electric current in muscles using his own capacitor. He explained that the potential in the brain cells is generated because of the ions (such as sodium and potassium) that flow inside and outside of the cell. This potential is known as the resting current or *resting potential* [62]. In 1920, Hans Berger recorded EEG signals from the lesion area of the human scalp for the first time using a Siemens double-coil galvanometer and non-polarized electrodes. He observed oscillations in the galvanometer after placing two clay electrodes at a distance of 4 cm apart near a scar. However, these recordings are not clear since they contain a significant amount of artifacts. The oscillations developed have been quoted as the mirror of the brain—“the *Elektronephalogramm*” [156]. Berger analyzed waves between 8 and 13 Hz, that is, alpha waves, which are also known as Berger waves.

Currently, various devices are available to record EEG signals for medical and research purposes. Some of the commercially available devices are NeuroSky, Neuroscan, EMOTIV EPOC, and Brain Products. A review on the features of various EEG devices have been given by Ramadan et al. [126]. Currently, researchers are trying to develop EEG devices that are reliable and provide good quality signals. A low-density EEG system with 7 channels based on automated artifact removal for Alzheimer patients has been developed by Raymundo Cassani et al. by analyzing 3-minute open-eyes and awake time activity. The EEG activity of an Alzheimer patient is different from that of a healthy person. The EEG signals of an Alzheimer patient are slow because of loss of synaptic connections and decreased synchronization between cortical regions. Most of the available EEG devices are high-density devices (a large number of electrodes), such as 32-channel or 64-channel EEG systems. Because of the large number of electrodes, patients with a brain disorder may feel uneasy, drowsy, and may sweat, which results in poor signal quality with artifacts. To overcome these limitations, a low-density, portable, and easy-to-use 7-channel system has been developed

Table 3. Brodmann Area and Respective EEG Electrodes

Brodmann's areas	Electrode number (10-10 system)
6	FC1, FC3, FC2, FC4
8	F1, F2, F3, F4
9	AF4, AF3
46	AF7, AF8
10	FP1, FP2
11	NZ
47	F7, FT7, FT8
44 and 45 (Broca's areas)	F8, FC6 (left hemisphere)
17	POZ
18	O1, O2
19	PO3, PO4, PO7, PO8
1	C4
2	C3
5	C1, C2
7	P1, P2, Pz
39	P3, P4, P5, P6
37	P8, P10
21	T4
22 (Wernicke's area)	T5, T6

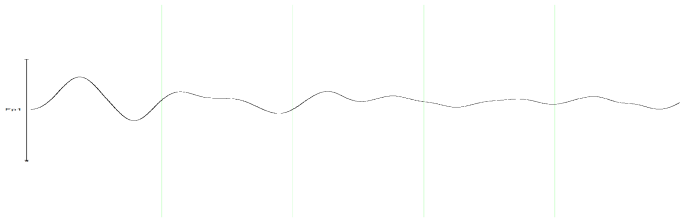


Fig. 6. Theta wave.

by the authors of [26]. In other work, a *smart wearable helmet* that monitors EEG and ECG activity has been developed by the authors of [160]. The EEG signals have been recorded by placing an electrode inside the ear canal with a wearable device. Therefore, it has been named an EarEEG. Subjects are asked to wear this smart helmet during cycling and walking to deal with artifacts generated by muscle movement.

Brain signals recorded via EEG devices are classified on the basis of frequency, called *Brain Rhythmic Activity* or *EEG Rhythms* [126]. Delta waves (1–4Hz) are usually present in infants and during deep sleep. Theta waves (4–7Hz) are generally found in rodents; they are also found in humans during meditation, the unconscious state, or drowsiness. Alpha waves (8–13Hz) are found in humans in a relaxed state with closed eyes. Mu waves lie in the alpha-wave frequency range and are observed when activity in the motor cortex area is maximum. Beta waves (14–30Hz) are present when a person is alert, attentive, or thinking. Gamma waves (>30Hz) are generated during voluntary movements or when some stimulus is given. Figures 6 to 9 show various EEG rhythms that have been recorded from a healthy subject in a suitable environment, performing normal activities.



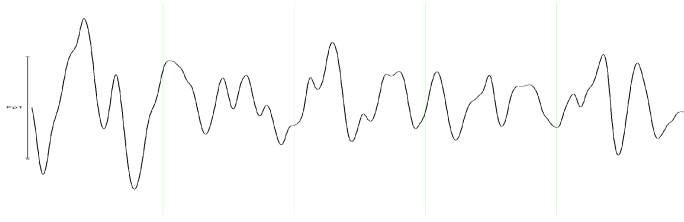


Fig. 7. Alpha wave.

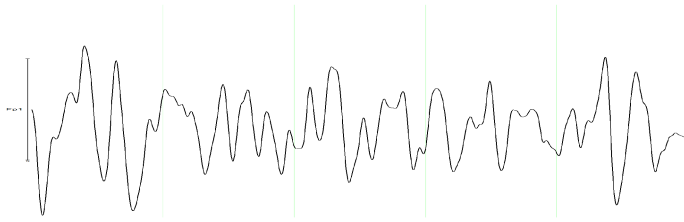


Fig. 8. Beta wave.

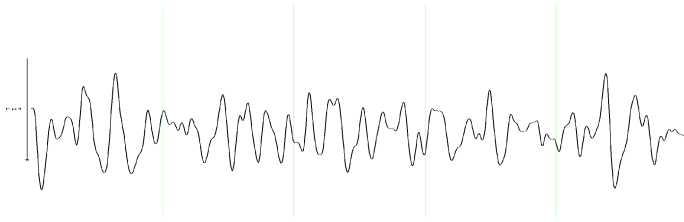


Fig. 9. Gamma wave.

**3.2.2 Magnetoencephalography (MEG).** MEG records magnetic fields produced as a result of neural activity generated in response to a stimulus. Like EEG, MEG also records postsynaptic potentials generated by neurons but in the form of magnetic fields. MEG provides good spatio-temporal resolution and is not severely affected by muscle artifacts [157]. The MEG is based on Superconducting Interference Devices (SQUIDs), which were introduced in 1960s. SQUIDs are filled with large liquid helium units that maintain the temperature of the system at approximately  $-269^{\circ}\text{C}$ . Temperature is kept low in order to achieve low impedance. The SQUID device detects and amplifies magnetic fields generated by neuronal activity. MEG was first used by David Cohen [34] to measure  $\alpha$ -brain waves from a healthy subject and an epileptic patient. The limitation of using SQUID-based devices is that it has to be maintained at a very low temperature, for which thermal isolation is required. A slow change in temperature of liquid helium will affect the system and will generate high maintenance costs. MEG does not need any referencing, as it provides better spatial resolution and is less distorted by tissue electrical activity. Hence, authors have preferred MEG over other acquisition techniques [27].

MEG has also been used recently for clinical and neurological research purposes. But, because of the cost, this device is not very popular. Hubert Cecotti et al. [27] have used the MEG technique to detect brain signals and have applied Bayesian Linear Discriminant Analysis (BLDA) on the input after spatial filtering. The authors have discussed the issues faced in the target detection system. They have proposed that single-trial detection should be used if continuous or repeated stimuli identify the target and non-target classes. The difference between functioning of an autistic

brain and a normal human brain is recorded using MEG by the authors of [148]. Also, using MEG activity, researchers have recorded differences between the speech of some 11-month-old bilingual and monolingual infants [49]. In a recent study, a solution for source reconstruction using MEG and EEG data was given by developing a hierarchical Bayesian algorithm [25]. The algorithm maximizes the likelihood of data using fast converging rules. Auditory, visual, and face-processing data have been used for simulation. The authors considered only spatial information for applying the algorithm. Ford et al. [54] have applied statistical analysis on spatio-temporal data recorded using MEG. They have provided auditory stimuli to the subject and found the difference in the continuous MEG data for repeated and novel stimuli. It has been observed that, for novel stimuli, cortical activity is more than that of repeated stimuli.

**3.2.3 Functional Magnetic Resonance Imaging (fMRI).** fMRI is also a non-invasive acquisition technique. It identifies the changes in oxygen flow of the blood or Blood Oxygen Level Dependent (BOLD) [128]. EEG and MEG both capture the rapid changes in cortical activity of the brain. This reflects the ongoing signal processing in the brain. fMRI indirectly measures neuronal activity by measuring the oxygen flow in blood. During any neuronal activity, oxygenated blood starts flowing toward the deoxygenated area and the fMRI records the difference in magnetic properties generated by oxygen flow. A typical fMRI system performs analysis of images after scanning, whereas a real-time fMRI performs simultaneous analysis. This feature of fMRI makes it useful in BCI for visualizing brain activities. The fMRI uses Echo Planar Imaging (EPI) to acquire brain activity slice by slice (slice defines time period). The real-time fMRI-BCI are influential because of their magnetic field strength, good spatial and temporal resolution, better echo time, and good magnetic field homogeneity [142]. The authors of [110] studied the functioning of the brain of individuals in different age groups. They found similarity and dissimilarity between the functionality of their brains using fMRI. An fMRI study [138] has been performed on children aged between 5 and 18 years. A task was given to the children to identify an object. They were given auditory stimuli to which they had to match line drawings (visual stimuli) of those sounds. Later, Independent Component Analysis (ICA) was applied to separate the task-related components and non-task components. In a study [71], fMRI was performed on preterm infants. It was used for early prognosis of brain injury and their neural development. In another work, a novel framework has been developed to improve the detection accuracy of fMRI [77]. For increasing the detection rate, signal extraction was performed by converting large brain volume into stimuli-specific parts. The slicing of brain volume was carried out for a specific point of time relative to the stimulus. Statistical analysis was applied on each time slice. The signals were extracted using a non-standard timepoint-by-timepoint approach. fMRI-based BCI has been used by many scientists to control various prosthetic devices [94], spelling devices [146], for playing table tennis [165] and for many other tasks. For the analysis of fMRI data, some of the available software programs are GE's BrainWave [8], AFNI (Analysis of Functional NeuroImages) [36], BrainSuite [140], and BrainVoyager [57].

**3.2.4 Functional Near-Infrared Spectroscopy (fNIRS).** fNIRS uses light from the near-infrared region of the Electromagnetic (EM) spectrum to study the oxygenation and deoxygenation of hemoglobin in the brain. The oxygenation and deoxygenation of hemoglobin occurs in response to stimuli or during activity. It works on the principle that, when EM waves are passed through a substance, the change in light intensities is visible. Based on the changes in the intensity of light, the property of a given substance can be determined. fNIRS images have high spatial resolution but low temporal resolution [126]. fNIRS measurements are done by three methods: continuous wave (CW), time-resolved (TR) and frequency domain (FD). fNIRS-CW is the most widely used acquisition system. fNIRS acquisition is performed by placing an *optode* (an optical sensor device)

on the brain scalp. The placement of an optode is done using the EEG electrode placement map (see Figure 5). An optode consists of a source and a detector. The source passes EM waves to the brain scalp, which is then transmitted and received at the detector. The source and the detector are placed at a distance of 2.5cm from each other [38]. The change in intensity on presentation of stimuli is recorded and analyzed using different available methods. Signal acquisition with fNIRS is costly but is relatively cheaper than fMRI data acquisition. The authors of [93] have developed a computer model for generating synthetic data of fNIRS that can help researchers to understand its mode of operation. Noori et al. have presented an fNIRS-based BCI system, with which they performed feature extraction and classification of data obtained after motor imagery tasks. The fNIRS signals have been acquired from the motor cortex area. Postfiltering the raw data and removing the noise, features such as mean, variance skewness, peak, and kurtosis are extracted. Later, a genetic SVM was applied on data for classification [118]. Some of the fNIRS-based BCI systems developed earlier are decoding of brain activity using fNIRS [38, 136], motor cortex activity during right-hand movement (active and passive) [96], right-hand and feet movement [1], and signal acquisition from prefrontal and primary motor cortices for mental tasks [2]. Devices available for the acquisition of fNIRS are Biopac [52], Artinis [151], and NIRx (NIRSport) [14].

**3.2.5 Positron Emission Tomography (PET).** PET is another non-invasive approach. It measures the functionality of the brain by injecting a nuclear substance-emitting positron. It records the chemical changes occurring in the brain before the symptoms of disease are visible. The dosage of radionuclide injected into the patient's body is less and does not cause any damage. Also for different regions of brain (as discussed in Section 2), different types of isotopes are used to measure brain functionality. An extensive review of the clinical application of PET has been given in [92]. PET has been mostly used for diagnosis of brain disorders. PET has been used to analyze patients with myotonic dystrophy type 1 and type 2 [3] and to study the effects of alcohol on the human brain [5]. Various methods have been applied for reconstruction of PET images [100, 109, 133].

**3.2.6 Single-Photon Emission Computed Tomography (SPECT).** SPECT is a nuclear medicine technique that uses gamma rays to study the brain. SPECT provides a look into how the brain works. While recording brain activity using SPECT, a radioactive substance is injected into the patient's body and is scanned using a SPECT machine. The SPECT machine traces the radioactive substance absorbed by the brain present in patient's body. The radioactive substance allows doctors to see how blood flows into tissues and organs. It shows the areas of the brain that are active, inactive, or overactive. SPECT averages the brain activity over a few minutes and generates an image. By reading these images, clinicians can identify any lesion in the brain or percent activity of the brain. SPECT has been widely used in healthcare to detect seizures in patients suffering from epilepsy [44, 46], Parkinson disease [66], and more.

## 4 TYPES OF BRAIN SIGNALS

### 4.1 Event-Related Potentials

Event-Related Potentials (ERPs) are the brain neuron activities stimulated by internal or external responses (cognitive, motor, or sensory) and are recorded non-invasively. ERP activity changes with time of onset of stimulus and location of the electrode on the brain scalp. The interest of researchers in the study of ERP is because of its potential in revealing the dynamics of the brain. The ERP waveform can be represented in the form of various positive and negative components. These positive and negative components are identified by their temporal occurrence. The components follow a pattern depending on the type and repetition of the stimulus. Some of the ERP component

examples are Error-Related Negativity (ERN), Contingent Negative Variation (CNV), N100, N200, P200, and P300.

Error-Related Negativity is the component of ERP that is generated when the subject responds incorrectly for any motor task [137]. It is usually found after 80 to 200ms after the erroneous response is generated [105]. ERN is the negative peak visible mostly when electrodes are placed at the frontal and central lobe of the brain [105, 137]. CNV is another ERP component that is attenuated when a single type of stimulus is repeated. CNV is visible between when the stimulus is presented for the first time and second time. The first stimulus is called a warning stimulus; the second stimulus, which is responsible for the generated response, is called an imperative stimulus. CNV usually appears after about 30 trials of the warning stimulus and imperative stimulus. The number of trials can be less if the subject easily understands the stimulus. A response to an imperative stimulus is necessary for a clear CNV elicitation [162]. CNV is also a type of cognitive ERP component.

Some of the negative and positive ERP responses generated on different types of stimuli are represented with initial “N” or “P” followed by the latency in milliseconds. N100 and P100 are the early and exogenous component of ERP having short latency [144]. These responses represent physical attributes of the presented stimuli, for example, the brightness of an image, frequency of a sound, and so on. N170 is negative peak generated after 170ms of onset of stimulus [103]. It is generated over the occipital and temporal lobe of the brain. N200 or N2 [53] is the negative peak that appears after 200ms of onset of stimulus (i.e., before P300). It usually appears on the frontocentral part of the brain. P200 is the second positive-going peak appearing after 200ms of onset of the stimulus. P200 records the cognitive activity of the brain. It is also visible in autistic patients [144]. P300 or P3 is the most widely and commonly used endogenous component of ERP for many BCI applications. It basically appears after the stimuli, which occurs rarely or is the most infrequent while presenting stimuli to the subject. It is elicited after 300 to 1,000ms of presentation of a rarely occurring stimulus [21]. It is generated by cognitive activity performed by the brain. There are two subcomponents of P3 or P300 waveform, P3a and P3b. P3a is elicited when the subject pays attention to a certain stimulus while P3b is elicited when some task-related stimuli are presented. P3a is more over the frontal lobe, whereas is P3b over the central and parietal lobes. Some of the most commonly studied P300-based BCI applications are conventional row/column spellers [45, 47], improved flash pattern for P300-based BCIs [78–81], lie detectors [9, 12, 48, 131], and 2D cursor controllers [32].

## 4.2 Evoked Brain Potential

Evoked Brain Potentials (EBPs) are generated by neurons in response to a stimulus. EBPs are a subtype of the ERP. These potentials are developed when a stimulus is given to our sensory organs, such as eyes responding to flashing of light, ears responding to sound, and the like. Some of the evoked potentials are described in this section.

**4.2.1 Visual Evoked Potential (VEP).** VEPs are observed on the occipital lobe (visual cortex) that can be captured by using various acquisition techniques such as EEG and ECoG. VEP is generated when a stimulus is presented to a subject, such as flashing a light or flashing words and pictures. If a stimulus is repetitively given in some fixed interval of time, then it is called a Steady-State Visual Evoked Potential (SSVEP) [17]. Many scientists and researchers used SSVEPbased BCI for studying brain responses, developing video games [97], controlling a prosthetic hand [113], developing a BCI speller [175], controlling 2-D cursor movement (recorded using 12 electrodes placed at the occipital lobe) [155] and much more.

**4.2.2 Auditory Evoked Potential (AEP).** When sound as a sensory stimulus is presented to a subject, the response generated is called AEP. When auditory stimulus is continuously given, it is called Auditory Steady-State Response (ASSR) [4]. A non-invasive way to record AEP using EEG is via EarEEG, in which the signal is recorded from electrodes placed within the ear canal [50, 87]. A BCI system has been purposed [4] to analyze the concentration level of subjects. To accomplish it, three types of auditory stimulus have been presented to subjects: monotone, music (violin and piano), and natural sound (cicadas singing and flowing water). Using an EEG device, four electrodes (Cz, Oz, T7, and T8) were placed on the subject's head for a BCI experiment. The subject was asked to concentrate on sounds for 20s. This system was applied to healthy subjects. In future, the same system can be used for persons with disability and on students using words and sentences to analyze their concentration power.

**4.2.3 Tactile Evoked Potential (TEP).** TEP is the response developed when a stimulus is presented to peripheral nerves, also known as the Somatosensory Evoked Potential (SEP). TEPs are usually observed at the parietal lobe. Devices [124] have been developed to provide somatosensory simulation that can be used by various acquisition techniques, such as EEG and MEG. A TEP-based BCI system has been developed in which the stimulus is presented to both index fingers. The subjects have to focus their attention on either the left or right index fingertips. This experiment has been performed on healthy subjects [114] and on visually impaired subjects [72]. Both works have used EEG for signal acquisition and applied linear discriminant analysis for classification. The authors of [114] have used only three electrodes to be placed on C3, Cz, and C4 and achieved an average accuracy of 70.42%. The authors of [72] used four types of electrode arrangements: 3, 7, 9, and 19 electrodes. They concluded that as the number of electrodes increases, accuracy improves. An average accuracy of 80% and 65% with stimuli applied on the index finger of one hand and index fingers of both hands, respectively, has been achieved.

### 4.3 Sensorimotor Rhythms (SMRs)

SMRs are voluntarily generated during muscle movement and are acquired from the motor cortex area of the brain. SMRs are initiated by users' intentional movements from bilateral limbs and change in its amplitude help in controlling physical or virtual devices. These potentials are very useful for persons with muscular disorders, as devices controlled via SMR will help them to perform motor activities. SMR provides the highest degree of freedom [176] as compared with slow cortical potential and ERP. Using SMR-based BCI, a virtual helicopter in 3D has been controlled via EEG signals [132]. Through the virtual world of the software program Blender, subjects have been asked to move a helicopter left, right, up, and down.

In future, if readers want to extend the abovementioned work, currently with hand movement, leg movement imagery can be added. This will provide more axis of movement for the virtual helicopter. Via hand motor imaginary movements and more training for the subject, flying and landing of the helicopter is also possible. Other acquisition techniques such as MEG [161, 164] and ECoG [123] were also used for capturing imaginary hand movements and the controlling cursor.

### 4.4 Slow Cortical Potential (SCP)

SCPs, also referred to as *DC potentials* or *slow oscillations* [166], are generated in the cortex area during EEG recordings. SCPs are lowest-frequency features where potential shift occurs between 0.5s to 10.0s [166]. With any increase or decrease in cortical activity, these SCPs shift toward negative or positive direction from baseline. Applications such as thought translation [112, 119] have been developed by voluntary control of SCP. This requires training the subject to produce SCPs while other potentials are involuntarily generated.

## 5 FEATURE EXTRACTION

Previous sections have explained various types of signals generated from the brain and different acquisition devices used to capture them. The acquisition device records raw signals that consist of noise or artifacts generated by eye blinks, muscle movement, hair, sweat, and other factors. To obtain useful information from raw signals, feature extraction techniques are applied to remove noise from signals, transform the signals, and reduce their dimension. In this section, a few commonly used feature extraction techniques for BCI systems are discussed.

### 5.1 Common Spatial Pattern (CSP)

The CSP [88] designs a spatial filter or spatial transform so that the filtered brain signal variance is maximized for classification. To perform CSP, gaussianity is assumed; frequency and time are considered as known parameters. It projects multichannel EEG data into lower-dimensional subspace [127]. It maximizes the variance of classes for a two-class signal matrix. The following steps are implemented in order to achieve a transformed EEG matrix.

- (1) Normalize the spatial co-variance of EEG as

$$C_K = \frac{X_K X_K^T}{\text{trace}(X_K X_K^T)}, \quad (1)$$

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

- (2) Compute the composite spatial co-variance as

$$Cov = \sum C_K, \forall \text{classes} \quad (2)$$

$$C_K = V_K \lambda_K V_K^T, \quad (3)$$

where  $V_K$  represents the eigenvector matrix and  $\lambda_K$  represents the diagonal matrix of the eigenvalue.

- (3) The projection matrix  $V$  is denoted as

$$P = V^T U, \quad (4)$$

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

$$W = P X, \quad (5)$$

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

$$X = P^{-1} W. \quad (6)$$

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

- (5) The first and last column of  $P^{-1}$  explains the largest variance of one task and smallest variance of the other.

Herbert Ramoser et al. [127] suggested the above formulas to design a spatial filter for classifying motor imagery EEG data. The variances of only a small number of signals that are most suitable for discrimination are used for the construction of the classifier. The problem with the method proposed by the authors is that if the signal is contaminated with a single artifact, the design of the filter changes severely. The design changes as a result of change in covariance, which is used to estimate the spatial filters. Thus, there is a need for artifact-free EEG data. Another limitation of using the CSP is that it does not provide the temporal data of filtered EEG signals. EEG data is



not stationary, thus we cannot guarantee extraction of the same data from the same subject every time. This is because of the artifacts and other environmental conditions. The contaminated data affects the covariance estimation and, in turn, causes overfitting problems.

There are various extensions to CSP that can be easily applied to EEG data, giving good performance. These are (a) Regularized CSP, (b) Spectrally weighted CSP, and (c) Stationary CSP. A brief review on the developments done in the past on CSP has been tabulated in Table 4.

## 5.2 Principal Component Analysis (PCA)

PCA [73, 122] is another method that performs transformation by maximizing the rate of decrease of variance of data. PCA (also known as Karhunen-Loeve transformation) uses a transformation matrix that contain elements with low variances. Transformation matrix  $A$  can be written as

$$A = \frac{1}{n} \sum (x_i - \mu) \cdot (x_i - \mu)^T, \quad (7)$$

where  $x_i$ 's are elements of the  $N$  dimension dataset and  $n$  is the total number of elements in the original dataset.

$$A \cdot Y = \Lambda \cdot Y, \quad (8)$$

where  $Y$  is the matrix containing eigenvector  $y_1, y_2, \dots, y_n$  and  $\Lambda$  is the eigenvalue diagonal matrix with elements  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Li [99] has applied dimension reduction on EEG data and proposes a new, improved and effective PCA that uses a covariance matrix to classify Multivariate Time Series (MTS) on the basis of time-based variables. PCA has been used to reduce MTS into Principal Component Sequences (PCSs) having lesser dimension than MTS. Conventional PCA (PCA) transforms data into PCSs having different lengths. To measure the distance, functions such as Euclidean Distance (ED) and Dynamic Time Wrapping (DTW) are mostly used. Li [99] has used DTW instead of ED, because ED is affected by noisy components. The Common Principal Component Analysis (CPCA) has been designed for MTS data, comprised of three steps: subspace construction, feature extraction, and classification. Subspace is constructed similar to traditional PCA, in which same-label MTS items are used to form a cluster. After constructing subspace, MTS items are projected into these subspaces. Later, these subspaces are transformed into PCS. However, traditional PCA produces groups of PCS for a single MTS. The total variance is calculated and a minima is chosen for classification task. Time consumed by CPCA has been analyzed and compared with other methods, showing that it is faster than others.

In another work [20], EEG-based BCI has been used to detect movement-related cortical potential (MRCP) by extracting variables using PCA. The authors used PCA, which acts as a temporal filter and determines a set of linear combinations of data. Coefficients of temporal filters are contained in the resulting transformation matrix. The first component produced features similar to the waveform that gives the average of various trials of MRCP. During the experiment, the authors placed EEG electrodes on the motor cortex area, which includes the frontal and parietal lobes. A generalized linear systems framework for PCA based on the Singular Value Decomposition (SVD) model for representation of spatio-temporal fMRI data was presented [11]. Time series data from brains of non-human primates for analysis was considered and PCA was employed. PCA provides both spatial and temporal characteristics present in short-term brain responses.

## 5.3 Independent Component Analysis (ICA)

ICA is also called "blind source separation" [70], which decomposes data into various independent components according to their statistical interdependence. ICA follows a simple linear transformation method. Let  $X$  be a matrix of the original signal of  $N$  dimensions, let  $T$  be the reduced

Table 4. Comparative Analysis of Various Common Spatial Pattern Methods

Extension to CSP	Method	Aim	Description	Improvement in classification accuracy from CSP
Regularized CSPs	(a) Composite CSP	To improve the performance of CSP that is deteriorated when fewer training samples are available for a subject [83]	Using records of users who have already performed the same task. Two methods are proposed: (1) de-emphasized covariance matrix and (2) emphasized covariance matrix. More advantageous when a few training samples of subjects are considered.	16.65% in mean using method 1 and 12.1% in mean using method 2 for subject with less training samples
	(b) Generic learning RCSP	To overcome problem of small number of training samples [102]	Two parameters are used to regularize covariance matrix: one increases estimation stability and the other reduces bias.	8.5% in mean
	(c) Diagonal learning RCSP	To estimate covariance [101].	Reduces covariance matrix to identity matrix. Regularized parameter is automatically identified using Ledoit and Wolf's method.	Same as CSP
	(d) A CSP regularized with selected subjects [101]		Works the same as CCSP but uses data from some selected subjects sequentially so that accuracy is maximized during training.	2.3% in mean and 3.5% in median
	(e) A weighted Tikhonov RCSP (WTRCSP)		Different large weights are applied on each channel. Performs best as it reaches at highest mean and median among all.	3.9% in mean and 9.4% in median
Spatial Filtering	(a) A spatial RCSP (SRCSP).	To obtain smooth spatial filters [101]	Applied Laplacian penalty term to get smooth filters. Penalty term has high value for non-smooth filters.	3.7% in mean and 6.6% in median
	(b) Adaptive spatial filter (ASF)	To suppress all EEG data that does not originate in the region of interest [60]	ASFs are designed to find out the maximum ratio of variance of electrical activity originated in the region of interest. A priori knowledge is required to estimate the covariance.	15% to 42%, depending on subject
	(c) Beamforming	To obtain EEG signal from predefined brain region using unsupervised spatial filtering [61]	Spatial filter is derived such that the ratio of variance is maximized for recorded EEG data produced within ROI using Rayleigh quotient.	3% in mean.
Subspace Analysis	(a) Stationary Subspace Analysis (SSA)	To identify stationary brain data [120]	SSA aims to divide EEG data into two components: non-stationary and stationary.	—
	(b) Extended SSA (Groupwise SSA)	Grouping of signals into subspaces of subjects or trials to find stationary dataset [134]	To find stationary data in each group of epochs. Distance between epochs and mean of distribution is measured in gwSSA.	0.8% to 6.7%, varying from subject to subject
	(c) Discriminant SSA (dSSA)	To extract stationary subspaces without losing information required for classification [135]	A trade-off parameter $\lambda$ has been used to project gwSSA objective function towards zero. Conjugate gradient descend has been applied to minimize objective function.	4.2% in mean and 2% in standard deviation for $\lambda = 0.75$

independent component matrix with  $M$  dimensions, and let  $A$  be the mixing matrix. Then,

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

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

ICA provides good accuracy for artifact removal, but it is difficult to get a component consisting purely of the artifact, which also contains useful brain signals. Thus, ICA has been improved by applying a combination of different methods. An ICA-based algorithm was developed employing temporal and spatial characteristics of Independent Components (ICs) for a P300 speller by Neng Xu et al. [169]. In another work [16], the authors applied Infomax ICA to retrieve P300 signals and to remove all other evoked potentials. Infomax ICA had also been applied to reduce dimensionality of 64-channel EEG data into 64 ICs [42]. A number of studies have been done and experiments have been performed for artifact removal by applying ICA, such as voluntary and involuntary eye blink detection [84], ICA on time series fMRI data [31], removal of muscles, decay, blink, and auditory data using Transcranial Magnetic Stimulation and Electroencephalography (TMS-EEG) [129]. The limitation of using ICA is that there is no procedure for automatic selection of ICs and a risk that selected ICs are desired ones. To improve on the limitation, ICA with outlier detection [181] was proposed in which Ocular Artifacts (OA) was detected and removed. The method follows two steps, initially applying a low-pass filter to EEG and then the filtered independent components are analyzed one by one. The pattern of artifact generated by eye movement is analyzed by outlier detection and then artifacts are detected and zeroed. Independent components with artifact are removed and meaningful EEG signals are retrieved.

#### 5.4 Wavelet Transformation (WT)

WT provides frequency and temporal information of the original signal. It expresses the signal in the form of a linear combination of a function. These functions are obtained by either shifting or scaling a single function known as the *mother wavelet*. After applying WT, signals are reduced into different frequency ranges. These frequency ranges are classified into approximation and detail levels. For a detailed study of WT and its various functions, refer to [40, 107, 141].

WT has been used for EEG signal transformation by Ting et al. in [152]. It reduces the original EEG signal into approximate and detail frequency coefficients. For the first level, the original EEG signal is transformed into high-frequency and low-frequency components having a length half that of the original EEG signal. This transformation can be applied till required features are obtained. As an EEG signal less than 50Hz is useful, hence, the authors of [152] have chosen the sub-band mean, where the frequency range lies between 0 to 50Hz. Another feature chosen by authors is sub-band energy, which is the square of the amplitude of the signal, lying in the range of 0 to 50Hz. The final feature vector has been selected using the Fisher distance criterion applied on the sub-band mean and sub-band energy. The feature set is finally fed to the classifier. An extension to WT (or wavelet decomposition), called wavelet packet transformation (or wavelet packet decomposition), uses multiple bases that, in turn, give variable classification output. Wavelet Packet Transformation (WPT) divides the original signal into two subspaces: Low-Frequency (LF) subspace and High-Frequency (HF) subspace. WT partitions only LF but WPT partitions both LF and HF. It uses sub-band energies obtained at final decomposition level as features. The disadvantage of using WPT is that it is a non-adaptive and non-subject-based approach. An adaptive WPT has been discussed by Yang et al. [173] to adapt the best basis function fitted for each object. After selecting the best basis, sub-band energies containing the best basis are selected as features. The number of features to be extracted depends solely on number of sub-bands containing the selected best basis. An extension to this work has used a fuzzy inference system with WT [172]. It selects the best

basis for representing EEG signals from various wavelet bases. Criteria based on the fuzzy set are defined that select the best basis.

## 6 CLASSIFICATION ALGORITHM

### 6.1 Support Vector Machines (SVMs)

SVMs [35] use the approach of projecting the input space to high-dimensional space such that non-linear data can also be easily separable. The aim of SVM is to choose an optimal separating hyperplane such that it maximizes the distance between two data points from different classes. The hyperplane ( $\bar{w}_0 \cdot x + b_0 = 0$ ) separating data with maximal width determines the direction  $\bar{w}/|w|$ , where the distance between two vectors should be maximum. For a detailed study of SVMs, see Cortes and Vapnik [35].

The SVM has been applied on various types of data to classify them into multiple classes. Similarly, the SVM has been applied on EEG signals for classifying two classes: presence and absence of P300 data [85]. P300 data has been recorded by using a P300 speller (a  $6 \times 6$  matrix having 36 alphabets arranged in rows and columns). Using Gaussian function, the value of transformation function ( $f(x)$ ) has been calculated. Several trials have been performed so that the correct value for the correct symbol is obtained. For each trial value,  $f(x)$  has been calculated and combined by summing them up with corresponding rows and column values. The target row or column is chosen having maximum summation value. Similar work using SVMs for classification of data generated using the P300 speller has been done [125]. The clustering of training data has been performed by separating these signals into homogeneous groups. A multiple classifier system has been designed, where the SVM classifier is applied on each cluster of two subjects. The function  $f(x)$  is calculated for each  $K^{th}$  partition for both subjects. For multiple classification, each single classifier is trained and assigned a real valued score  $f_K(x_{r|c})$ , where  $x_{r|c}$  is a poststimulus vector with row  $r$  and column  $c$ . It needs to maximize the score after  $J$  number of sequences with the most probable row and column:

$$S_{r|c} = \frac{1}{J} \frac{1}{K} \sum_{j=1}^J \sum_{k=1}^K f_K(x_{r|c}^{(i)}), \quad (11)$$

where,  $x_{r|c}^{(i)}$  is the poststimulus vector during the  $j^{th}$  sequence.

$$S_{r|c} = \frac{1}{K} \sum_{k=1}^K \sum_{i \in P_k} y_i \alpha_i^k \left\langle \frac{1}{J} \sum_{j=1}^J x_{r|c}^{(i)}, x_i \right\rangle + b^k \quad (12)$$

This classifier performs double averaging, the first averaging on the data space and the second being done in the classification score space. An SSVEP-based BCI has been developed in which a brain-controlled device is designed using an SVM. SSVEP data is classified into three classes: turning left, turning right, and moving forward with Radical basis function as the kernel [18].

An extension to the SVM, Dynamic Adaptation SVM (DASVM), has been proposed by Bruzzone and Marconcini [23], which used an unlabeled test sample to exploit the decision or transform function. The decision function is drawn from the target domain, which is different from the source domain of labeled training samples. The SVM uses both unlabeled and labeled data that are drawn from the sample domain. Extensive research has been conducted with SVMs and various researchers have proposed various kernel functions such as linear, Gaussian, and polynomial. However, from the literature, it is not clear which kernel function with what parameters may be more cost-effective and space-effective and sensitive to noise. To overcome the limitation of sensitivity toward noise, a Fuzzy Support Vector Machine (FSVM) has been employed by Xu et al. [170]. A fuzzy membership function is given to each training sample, which enhances the SVM by reducing the effect of outliers and noise. The FSVM gives less importance to outliers and

noise by giving them a lower fuzzy membership value. A comparison between the classical SVM and FSVM using the radial basis function as kernel and DWT as the feature extraction method is performed by calculating the error rate. The error rate is the ratio of the number of support vectors to total number of samples. To produce a better performance, the error rate ( $\delta$ ) should be less than or equal to the ratio of the number of support vectors and total training samples.

## 6.2 Neural Networks

NNs are inspired from biological neural systems having features such as parallel computing, non-linearity, adaptability, responsiveness, and fault tolerance. The inputs in NNs are called *neurons*, which are connected with *weights* (which can be positive or negative). The inputs with weights are processed through processing units. The processing units consist of a summation part, which is ultimately connected to output. The main contribution for development of the NN is by McCulloch and Pitts [111], who in 1943 purposed a neural model that takes a weighted sum as input followed by a threshold logical function.

**6.2.1 Perceptrons and Multilayer Perceptron.** Rosenblatt in 1958 [130] purposed a feed-forward NN called the Perceptron neural model. A single-layer perceptron is not able to classify non-linear data; hence, a Multilayer Perceptron (MLP) is used to resolve the problem of non-linearity. In MLP, the output of second input layer is combined to form another layer, and so on, until the problem is properly classified. These layers that are added into the network are called hidden layers. The learning law takes error received at the output layer and propagates it to hidden layers for updating weights. This learning law is called the *generalized delta rule* or *backpropagation* or *backward error propagation*.

Different MLP architectures have been applied on the EEG data in [163] after applying WT, SVM, and backpropagation on EEG data. An NN with backpropagation has achieved an accuracy of 91.4% and an SVM achieved an accuracy of 91.13%. In [13], the authors have used PCA for feature extraction and applied backpropagation on extracted data. Backpropagation has low computational complexity, but the problem is that it stuck at local minima. To solve the issue of local minima, hybrid feature selection procedures and extensions to backpropagation have been applied by researchers. In [67], the authors have compared working with MLP and Finite Impulse Response (FIR) MLP (an extension of MLP), which uses filters in place of weights in MLP architecture and performs temporal processing using filters. Keeping the architecture constant, the authors have found a number of free parameters used in both MLP and FIR-MLP. FIR-MLP reduces the number of free parameters, which improves the computational capability of the NN. To improve backpropagation learning, some improvements have been made, which are in given Table 5. Many works have used different NN architecture for classification of brain data. An MLP for MEG data analysis [22], a fully connected cascading NN, has been considered for fMRI data [43], an NN with radial basis function for classifying fMRI data [104].

**6.2.2 Convolutional Neural Network (CNN).** A CNN is another type of NN that is architecturally similar to MLP. It arranges neurons into three dimensions: width, height, and depth. A CNN can be structured in a five-layered architecture: (a) input layer, (b) convolutional layer, (c) rectified linear unit layer, (d) pooling layer, and (e) fully connected layer. Let  $L$  be the number of layers in CNN architecture,  $x$  is the input vector,  $w$  is the weight vector, let  $M$  be the number of maps in layer  $L$ ,  $m$  denotes the  $m^{th}$  map,  $J$  represents the number of neurons in layer  $L$ ,  $N_e$  is the number of electrodes,  $N_s$  is the number of signal values, and  $N_p$  is the number of partitions of signal values.

For layer 1:

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

Table 5. Variants of Backpropagation

Variants of Backpropagation	Improvements	Features
Asymptotic convergence of backpropagation [149]	Including momentum term in gradient descent equation.	Convergence rate increases for the case of sigmoidal transfer function.
Extended backpropagation [171]	Learning rate adaptation is based on correlation coefficient between local gradient and previously updated weight.	Learning rate increases and decreases exponentially, which makes algorithm more optimal in less iterations and, hence, faster.
Backpropagation with adaptive learning rate [106]	Applied Goldstein's and Armijo's [58] work to construct method that can adapt learning rate.	Automatically adapts convergence rate; faster training can handle large learning rates.
Back propagation using "Self-Determined Learning Rate" [106]	In place of learning rate, "tuning" was used to reduce large learning rate and achieve convergence.	Doesn't use learning rate and provides better generalization.
Backpropagation with learning rate different for each weight [106]	Learning rate for each weight done by estimating the local Lipschitz constant along each weight direction.	Good average, firm learning, better classification accuracy.
Hierarchical backpropagation (HBP) [174]	Divides MLP into sublayers. Each MLP has one hidden layer and input and output layer. Each sublayer is trained with backpropagation individually.	Avoids local minima, noise is not transmitted onto further layers as it has been overcome initially, and error rate is reduced.
Backpropagation based on Lyapunov stability [108]	Instead of finding global minima, it finds energy surface having single global minima; for fast errorconvergence, Lyapunov adaptive BP algorithm has been used to construct an adaptive filter.	Convergence of error to zero, adaptively updating weights, which reduced effect of input disturbances.

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.

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,

$$\sigma(3, J_3) = \sum_{i=0}^{M_2} \sum_{k=0}^{N_p} x(2, i, k) w(4, i, k) + bias. \quad (15)$$

For layer 4,

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

A classifier based on a CNN to classify EEG data having P300 component of ERP has been used [29]. The CNN consists of five layers and several maps. The output layer consists of one map, which has two neurons. These two neurons represent two classes (class 1, P300 detected; class 2, no P300). To select the order of CNN, first filters (or weights) are applied on the width and height of the input volume, then signal processing is performed in time domain. Here, authors have used kernels as vectors, not as matrix. Authors have applied a linear sigmoidal function between hidden layer 1 and hidden layer 2. Input signal convolution can be represented as

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



where,  $\sigma$  is the first deviation and classical sigmoidal function was used in between the last two hidden layers.

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

Authors have applied the set of Equations (13) to (16) in their work on various layers of CNNs and have used backpropagation for updating the weights. At the output layer, the class score was calculated as

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

CNN performs better classification than other classifiers, as it uses more hidden layers, but the number of layers to be used to achieve a better classification cannot be determined. In another work by Cecotti et al. using Equations (13) to (16), they have classified a set of images: human face (target) and others (non-target) [28]. In their work, a CNN is embedded with a spatial filter. The filtering and classification are performed on ERP signals produced during the experiment. The learning method is based on maximizing the Area Under the Curve (AUC). They had compared a CNN with SVM, BLDA (with and without spatial filter) and MLP. The finding shows that the CNN performs better than the other classifiers.

A CNN based on AUC does not require any prior knowledge about the type of spatial filter used but needs prior information about the type of architecture of the network. Hence, choice of number of neurons and spatial filter depends on the past experiment, which, in turn, affects the result and overall performance of the network. In order to achieve optimal performance for classification, a CNN can be used with legitimate choice of neuron and hidden layer.

**6.2.3 Probabilistic Neural Network (PNN).** The PNN was introduced by Specht in 1990 [147] and is based on Bayes probability rule, where the aim of the PNN is non-parametric estimation of the probability density function to obtain an optimal accuracy. The advantages of using a PNN are that it is easily operable (much faster than backpropagation), has a parallel structure, provides instantaneous training, converges optimally, has no local minima issues, and has real-time usage. An appropriate selection of the smoothing parameter ( $\sigma$ ) helps in modifying the shape of decision surface [147].

A PNN has been applied to decode the motor cortical signals generated from a rat's brain by Zhou et al. The implementation has been done on a Field Programmable Gate Array (FPGA) board. The method of acquisition is invasive; a total of 32 neurons are sorted and the number of spikes per neuron is calculated. The rate of spike firing (i.e., time vector) was considered as input to PNN. The authors have divided various neuronal activities into different classes and applied a PNN to classify in which class the current neuronal activity belongs [180]. Many other works have been done in which a PNN or hybrid PNN (PNN with other classifiers) has been applied for classification of brain data such as emotion recognition [179], PNN based on time series discriminant analysis applied with a hidden Markov model [68], a fivefold classification of motor imagery in which PNN was one of the classifiers [158], a PNN and multiclass SVM was trained on data classified using WT [63], and so on.

### 6.3 Fuzzy Inference System

Zadeh [177, 178] proposed that in the real world, all classes or sets do not belong to a crisp value such as yes or no, true or false, or a real number. Thus, he introduced the concept of fuzzy sets. "A fuzzy set is a set without crisp boundary and transition from crisp to flexible boundary is characterized through *Membership Function (MF)*" [76]. Using the advantages of a fuzzy set giving flexible boundary conditions, many authors have applied a fuzzy inference system on BCI applications.

Table 6. Types of FNN

Type of network	Input	Weight	Output
Neural network	Crisp	Crisp	Crisp
Fuzzy NN-1	Crisp	Fuzzy	Fuzzy
Fuzzy NN-2	Fuzzy	Crisp	Crisp
Fuzzy NN-3	Fuzzy	Crisp	Fuzzy
Fuzzy NN-4	Fuzzy	Fuzzy	Fuzzy

Kumar [90] used fuzzy logic for recognizing patterns of EEG data during human sleep for the first time. In another work, a fuzzy inference system was implemented to select the number of EEG channels for imagined speech (in Spanish) recognition [154]. ICA as ocular artifact removal and discrete WT as feature extracting approach has been used. The fuzzy inference method has been applied so that the system selects the combination of channels automatically. The channels are selected such that the error rate is reduced and performance is enhanced. Channel selection is an important factor, as each channel defines a different feature. Authors have used two components for channel selection: the first component searches for a non-dominated channel combination and the second component selects a single channel from a set of channel combinations. For selecting a single channel for each subject, authors have implemented the Mamdani fuzzy inference system.

In another work, classification of motor imagery tasks was done by Nguyen et al. [116], in which they applied a combination of a “fuzzy standard additive model” [59] with the tabu search learning method for classification. WT and the Wilcoxon test were used for feature extraction. A similar work for motor classification was done using a type-2 fuzzy system and applying WT for feature extraction [150]. The authors compared their results with other classifiers, such as SVM, NN, AdaBoost,  $k$ -nearest neighbors, and the “Adaptive Neuro-Fuzzy Inference System (ANFIS).” The accuracy was improved by approximately 2% to 10% with interval type-2 fuzzy systems.

#### 6.4 Neuro-Fuzzy Systems

The neuro-fuzzy system incorporates advantages of both the NN and FIS. Its architecture is similar to the NN and inputs or weights (or both) are fuzzified [24]. The FNN identifies fuzzy rules and tunes the membership function by adjusting the connection weights. Different types of neuro-fuzzy systems are listed in Table 6. Many applications have used FNN-1, FNN-2, and FNN-3. Here, a review of FNN-4 is given. In FNN-4, inputs and weights use fuzzy parameters that are mapped using various fuzzy operations such as  $\max$  ( $\vee$ ) and  $\min$  ( $\wedge$ ) [75]. An FNN has been applied on EEG recording of patients by applying *if & then rules* to the output of the NN [10]. An auditory stimulus has been given via a set of earphones to the patient, which generates AEP. AEP produced is given as input to a three-layer NN with 31 inputs, a hidden layer of 10 nodes and 5 outputs. The membership function produced by the NN is passed on to a fuzzy controller, which applies fuzzy rules. The input given to the *if and then rule* is the latency that is generated by AEP. The authors have marked a point  $Nb$  and if the amplitude goes to  $Nb$ , anesthesia is given to patient. Thus, the rule defined is *if latency is  $Nb$  then increase anesthetic dosage*. After the fuzzy controller implements the rule, defuzzification is applied using the centroid method.

A Self Organized FNN (SOFNN) is proposed by Leng et al. [98] using dynamic FNN architecture to create a self-adaptive architecture for identification of the “singleton” or “Takagi-Sugeno” fuzzy model [167]. The advantage of designing this hybrid structure is that it is more interpretable, as it makes use of the learning ability of the NN. The system has to identify the parameters of premise and consequences and the number of partitions of input and output space and fuzzy rules. The problem with SOFNN is that if non-linearity increases, the number of neurons will be increased.

Table 7. Various Latest Techniques on Brain-Computer Interface

Signal acquisition	Feature extraction	Classifier	Description
EEG (4 subjects, 5 mental tasks) [121]	Elliptic filters	Elman Neural Network (ENN)	For classifying, mental task asymmetry ratios are used. Resilient backpropagation (hyperbolic tangent function) as activation function is applied.
EEG (3 subjects) [37]	Hjorth method	SOFNN	SOFNN organizes its neurons during learning process.
EEG [145]	—	ENN	Provides a guaranteed convergence, overfitting avoidance using adaptive dead zone approach. Extended Elman backpropagation (eEBP) is applied as learning algorithm.
EEG (5 healthy subjects and 5 tetraplegia subjects, given 3 mental tasks) [30]	Hilbert-Huang transform [74]	Cross-mutated ANN with fuzzy particle swarm pptimization	Compared results of the classifier withgenetic algorithm and got higher accuracy.
EEG (9 subjects, 22 electrodes, motor imagery tasks) [39]	Robust CSP	Type-2 neuro-fuzzy classifier system	Uses a 5-layered fuzzy inference system and a self-regulatory mechanism. Gaussian membership function is used with unknown mean and known variance.
EEG, motor imagery tasks [69]	Common Bayesian network	SVM	Shows statistical relationship between activation areas and motor imagery tasks using Bayes rules. Gaussian Mixture Model is used to calculate probability density function of nodes; also gives common edges concept in CBN for feature extraction.
fMRI [86]	—	Neucube (A spiking NN)	Neucube based on evolving (eSTDM) has been proposed for modelling spatio-temporal data. It is a 5-module architecture that consists of encoding mapping, unsupervised and supervised learning, and classification.
EEG (10 subjects, motor imagery tasks) [168]	Sub-band common spatial pattern	Fuzzy integral with PSO	Fuzzy integral uses multiple decisions from different sources and collects all inferences. Sugeno integral and Choquet integral are applied and PSO is used to determine confidence of classifiers.
EEG (10 subjects) P300-based BCI [81]	Band pass filter	Bayesian Linear Discriminant Analysis (BLDA)	A new P300 paradigm has been designed with honeycomb-shaped red dots given as visual stimuli to increase 1000ms duration stimuli. This increased the classification accuracy of the system. The ERP response elicitation is the same as previous P300-based BCI paradigms.

To ensure feasibility of the network, every time a neuron is added, its impact on the performance of the system is evaluated. Checking each neuron every time as it enters the system increases the cost of the system exponentially as fuzzy MFs are expensive. Hence, a modified SOFNN [37] that improves system efficiency has been proposed. A record of neuron firing strength is kept for all clustering performed previously. The record is updated as the training progresses. This will reduce the cost and time of running the system for each training set. A lot of studies in the literature has been done in the field of BCI, a few of which we list in Table 7.

Table 8. Comparison of Various Existing Approaches

Performance Measures (in percentage)	WT				CSP			
	LDA	SVM	NN	KNN	LDA	SVM	NN	KNN
Accuracy	88.8	89.0	86.0	79.5	52.1	88.6	73.6	84.2
Sensitivity	88.7	86.0	88.7	84.1	52.6	88.6	75.6	90.6
Specificity	86.0	87.6	87.0	68.5	51.3	86.2	89.0	77.3

A comparison between existing approaches has been presented in this article. For comparing various approaches, EEG data was considered. The EEG data was recorded for a lie detection experiment. During the experiment, subjects were presented a set of stimuli for which they had to respond by either telling the truth or by lying. This stimulus generated the ERP responses, which were recorded. The ERP data was recorded for 10 subjects. The data is band pass filtered and later feature extraction and classification approaches are applied. Two feature extraction approaches were applied on 10 subject EEG data: CSP and WT. After feature extraction, the feature vector having WT coefficient was given to four classifiers: LDA, SVM, KNN, and NNs. Similarly, CSP also was applied on the EEG data and later the features were given to a similar set of classifiers. The results are tabulated in Table 8. From the results, it can be inferred that the SVM among the four classifiers performs the best for the data recorded for lie detection.

## 7 CONCLUSION AND RESEARCH DIRECTIVES

The brain-computer interface constructs a pathway that enables users to easily control a computer through their thoughts. BCI is an interdisciplinary area, since it provides scope for research in various aspects, such as understanding, acquisition, and processing of brain signals. BCI research involves biology (psychology and neurology), engineering, computer science, and applied mathematics. In this article, we have provided an extensive survey of each phase of the BCI. The first phase of the BCI is acquiring the brain signals. There are two types of acquisition systems: non-invasive and invasive. Invasive signal acquisition involves placing microelectrodes and electrode chips beneath the scalp through surgery. The non-invasive technique captures brain potentials by either placing a metal electrode on the scalp (as in EEG) or recording magnetic activity and rate of blood flow using some special apparatus (as in MEG, fMRI, etc.). These signal acquisition methods record various types of brain potentials, such as those generated by motor activity, cognitive activity, eye movement, or a stimulus. Researchers prefer non-invasive techniques over invasive approaches, as they are not prone to injury. The only limitation with non-invasive approaches is that the resolution of signals is poor in comparison to invasive approaches. Future work could be to develop brain signal acquisition devices that have low-density electrodes and provide higher resolution.

The second phase of the BCI involves processing brain signals. In this article, various feature extraction and classification algorithms have been mentioned. Feature extraction is applied on raw brain data to extract useful signals and to remove artifacts generated by eye movements, muscle movements, and the like. Various feature extraction methods have been used, including CSP, PCA, ICA, and WT. ICA works best for ocular artifact removal (eye movement) and has been widely adopted by various researchers. CSP and its variants are applied for spatial filtering of brain signals. PCA helps in transforming the feature space and WT extracts both the frequency and time information from the raw signals. Classification algorithms such as LDA, SVM, NNs, and fuzzy inference systems are applied on attributes obtained using feature extraction techniques. This article gives a brief overview of these classification approaches and presents the recent work using these techniques. Conventional BCI systems used discriminative models for classification.

However, now researchers are more interested in deep-learning approaches such as deep belief networks, CNNs, and a hybrid of various classification algorithms. An advantageous BCI system is one that has a smooth coordination between all of these parts of the BCI. The main aim of research in BCI is to provide a better way of communication; however, the methods applied to achieve this aim can differ.

## REFERENCES

- [1] Vera Kaiser, Günther Bauernfeind, Alex Kreiling, Tobias Kaufmann, Andrea Kübler, Christa Neuper, and Gernot R. Müller-Putz. 2014. Cortical effects of user training in a motor imagery based brain-computer interface measured by fNIRS and {EEG}. *NeuroImage* 85, Part 1 (2014), 432–444. DOI : <https://doi.org/10.1016/j.neuroimage.2013.04.097>
- [2] Keum-Shik Hong, Noman Naseer, and Yun-Hee Kim. 2015. Classification of prefrontal and motor cortex signals for three-class fNIRS-BCI. *Neuroscience Letters* 587 (2015), 87–92. DOI : <https://doi.org/10.1016/j.neulet.2014.12.029>
- [3] Stojan Peric, Lepsava Brajkovic, Bozidar Belanovic, Vera Ilic, Biljana Salak-Djokic, Ivana Basta, and Vidosava Rakoccevic Stojanovic. 2017. Brain positron emission tomography in patients with myotonic dystrophy type 1 and type 2. *Journal of the Neurological Sciences* 378 (2017), 187–192.
- [4] Jeong Heo, Hyun Jae Baek, Seunghyeok Hong, Min Hye Chang, Jeong Su Lee, and Kwang Suk Park. 2017. Music and natural sounds in an auditory steady-state response based brain-computer interface to increase user acceptance. *Computers in Biology and Medicine* 84 (2017), 45–52. DOI : <https://doi.org/10.1016/j.combiomed.2017.03.011>
- [5] Nora D. Volkow, Corinde E. Wiers, Ehsan Shokri-Kojori, Dardo Tomasi, Gene-Jack Wang, and Ruben Baler. 2017. Neurochemical and metabolic effects of acute and chronic alcohol in the human brain: Studies with positron emission tomography. *Neuropharmacology* 122 (2017), 175–188.
- [6] TheVisualMD.com. 2017. Brain Layer. Retrieved May 6, 2017 from <http://www.thevisualmd.com/searchimg/?idu=1083616692&q=s&p=140>.
- [7] Marie-Luise Lehmann. 2017. Brodmann Area. Retrieved May 6, 2017 from <http://www.brain-maps.com/brodmann-areas.html>.
- [8] GE Healthcare Systems. 2018. BrainWave. Retrieved March 2, 2018 from [http://www3.gehealthcare.com/en/products/categories/advanced\\_visualization/applications/brainwave](http://www3.gehealthcare.com/en/products/categories/advanced_visualization/applications/brainwave).
- [9] Vahid Abootalebi, Mohammad Hassan Moradi, and Mohammad Ali Khalilzadeh. 2009. A new approach for EEG feature extraction in P300-based lie detection. *Computer Methods and Programs in Biomedicine* 94, 1 (2009), 48–57.
- [10] Robert Allen and David Smith. 2001. Neuro-fuzzy closed-loop control of depth of anaesthesia. *Artificial Intelligence in Medicine* 21, 1 (2001), 185–191.
- [11] Anders H. Andersen, Don M. Gash, and Malcolm J. Avison. 1999. Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework. *Magnetic Resonance Imaging* 17, 6 (1999), 795–815. DOI : [https://doi.org/10.1016/S0730-725X\(99\)00028-4](https://doi.org/10.1016/S0730-725X(99)00028-4)
- [12] Annushree Bablani, Damodar Reddy Edla, Diwakar Tripathi, and Venkatanaresbhabu Kuppili. 2018. An efficient concealed information test: EEG feature extraction and ensemble classification for lie identification. *Machine Vision and Applications* (2018), 1–20. DOI : [10.1007/s00138-018-0950-y](https://doi.org/10.1007/s00138-018-0950-y)
- [13] Divya Balakrishnan and Sadasivan Puthusserypady. 2005. Multilayer perceptrons for the classification of brain computer interface data. In *Proceedings of the IEEE 31st Annual Northeast Bioengineering Conference, 2005*. IEEE, 118–119.
- [14] Randall L. Barbour, Jack Lubowsky, and Raphael Aranson. 1992. Method of imaging a random medium. US Patent 5,137,355, Filed June 8th., 1989, Issued Aug. 11th., 1992.
- [15] Jonathan Bredas. 2012. Brain-machine interfaces: Basis and advances. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42, 6 (2012), 825–836.
- [16] Anthony J. Bell and Terrence J. Sejnowski. 1995. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation* 7, 6 (1995), 1129–1159.
- [17] Fabrizio Beverina, Giorgio Palmas, Stefano Silvoni, Francesco Piccione, Silvio Giove, et al. 2003. User adaptive BCIs: SSVEP and P300 based interfaces. *PsychNology Journal* 1, 4 (2003), 331–354.
- [18] Luzheng Bi, Xin-an Fan, Ke Jie, Teng Teng, Hongsheng Ding, and Yili Liu. 2014. Using a head-up display-based steady-state visually evoked potential brain-computer interface to control a simulated vehicle. *IEEE Transactions on Intelligent Transportation Systems* 15, 3 (2014), 959–966.
- [19] N. Birbaumer, A. Kübler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, and H. Flor. 2000. The thought translation device (TTD) for completely paralyzed patients. *IEEE Transactions on Rehabilitation Engineering* 8, 2 (Jun 2000), 190–193.
- [20] Andreas Tröllund Boye, Ulrik Qvist Kristiansen, Martin Billinger, Omar Feix do Nascimento, and Dario Farina. 2008. Identification of movement-related cortical potentials with optimized spatial filtering and principal component analysis. *Biomedical Signal Processing and Control* 3, 4 (2008), 300–304. DOI : <https://doi.org/10.1016/j.bspc.2008.05.001>



- [21] Steven L. Bressler and Mingzhou Ding. 2002. Event-related potentials. *The Handbook of Brain Theory and Neural Networks* (2002), 412–415.
- [22] Matthew J. Brookes, Prejaas K. Tewarie, Benjamin A. E. Hunt, Sian E. Robson, Lauren E. Gascoyne, Elizabeth B. Liddle, Peter F. Liddle, and Peter G. Morris. 2016. A multi-layer network approach to MEG connectivity analysis. *NeuroImage* 132 (2016), 425–438.
- [23] Lorenzo Bruzzone and Mattia Marconcini. 2010. Domain adaptation problems: A DASVM classification technique and a circular validation strategy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 5 (2010), 770–787.
- [24] James J. Buckley and Yoichi Hayashi. 1994. Fuzzy neural networks: A survey. *Fuzzy Sets and Systems* 66, 1 (1994), 1–13.
- [25] 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.2018.07.056>
- [26] Raymundo Cassani, Tiago H. Falk, Francisco J. Fraga, Marco Cecchi, Dennis K. Moore, and Renato Anghinah. 2017. Towards automated electroencephalography-based Alzheimer’s disease diagnosis using portable low-density devices. *Biomedical Signal Processing and Control* 33 (2017), 261–271. DOI : <https://doi.org/10.1016/j.bspc.2016.12.009>
- [27] Hubert Cecotti. 2016. Single-trial detection with magnetoencephalography during a dual-rapid serial visual presentation task. *IEEE Transactions on Biomedical Engineering* 63, 1 (2016), 220–227.
- [28] Hubert Cecotti, Miguel P. Eckstein, and Barry Giesbrecht. 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 (2014), 2030–2042.
- [29] Hubert Cecotti and Axel Graser. 2011. Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 3 (2011), 433–445.
- [30] Rifai Chai, Sai Ho Ling, Gregory P. Hunter, Yvonne Tran, and Hung T. Nguyen. 2014. Brain-computer interface classifier for wheelchair commands using neural network with fuzzy particle swarm optimization. *IEEE Journal of Biomedical and Health Informatics* 18, 5 (2014), 1614–1624.
- [31] 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>
- [32] Luca Citi, Riccardo Poli, Caterina Cinel, and Francisco Sepulveda. 2008. P300-based BCI mouse with genetically-optimized analogue control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 16, 1 (2008), 51–61.
- [33] Anton Coenen, Edward Fine, and Oksana Zayachkivska. 2014. Adolf Beck: A forgotten pioneer in electroencephalography. *Journal of the History of the Neurosciences* 23, 3 (2014), 276–286. DOI : <https://doi.org/10.1080/0964704X.2013.867600>
- [34] David Cohen. 1972. Magnetoencephalography: detection of the brain’s electrical activity with a superconducting magnetometer. *Science* 175, 4022 (1972), 664–666.
- [35] Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine Learning* 20, 3 (1995), 273–297. DOI : <https://doi.org/10.1007/BF00994018>
- [36] Robert W. Cox. 1996. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research* 29, 3 (1996), 162–173.
- [37] Damien Coyle, Girijesh Prasad, and Thomas Martin McGinnity. 2009. Faster self-organizing fuzzy neural network training and a hyperparameter analysis for a brain-computer interface. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39, 6 (2009), 1458–1471.
- [38] Shirley M. Coyle, Tomás E. Ward, and Charles M. Markham. 2007. Brain-computer interface using a simplified functional near-infrared spectroscopy system. *Journal of Neural Engineering* 4, 3 (2007), 219.
- [39] Ankit Kumar Das, Suresh Sundaram, and Narasimhan Sundararajan. 2016. A self-regulated interval Type-2 neuro-fuzzy inference system for handling nonstationarities in EEG signals for BCI. *IEEE Transactions on Fuzzy Systems* 24, 6 (2016), 1565–1577.
- [40] Ingrid Daubechies. 1990. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory* 36, 5 (1990), 961–1005.
- [41] Reginald G. Bickford. 1969. *Physical Control of the Mind: Toward a Psychocivilized Society*. Harper & Row, New York, NY. American Association for the Advancement of Science, 666–666.
- [42] Arnaud Delorme and Scott Makeig. 2003. EEG changes accompanying learned regulation of 12-Hz EEG activity. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11, 2 (2003), 133–137.
- [43] G. Deshpande, P. Wang, D. Rangaprakash, and B. Wilamowski. 2015. Fully connected cascade artificial neural network architecture for attention deficit hyperactivity disorder classification from functional magnetic resonance imaging data. *IEEE Transactions on Cybernetics* 45, 12 (2015), 2668–2679.



- [44] Michael D. Devous, Robert F. Leroy, and Richard W. Homan. 1990. Single photon emission computed tomography in epilepsy. SPECT Imaging of the Brain: Part I. *Seminars in Nuclear Medicine* 20, 4 (1990), 325–341.
- [45] Emanuel Donchin, Kevin M. Spencer, and Ranjith Wijesinghe. 2000. The mental prosthesis: Assessing the speed of a P300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering* 8, 2 (2000), 174–179.
- [46] Roderick Duncan, James Patterson, Donald M. Hadley, P. Macpherson, M. J. Brodie, I. Bone, A. P. McGeorge, and D. J. Wyper. 1990. CT, MR and SPECT imaging in temporal lobe epilepsy. *Journal of Neurology, Neurosurgery & Psychiatry* 53, 1 (1990), 11–15.
- [47] Lawrence Ashley Farwell and Emanuel Donchin. 1988. Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology* 70, 6 (1988), 510–523.
- [48] Lawrence A. Farwell and Emanuel Donchin. 1991. The truth will out: Interrogative polygraphy (“lie detection”) with event-related brain potentials. *Psychophysiology* 28, 5 (1991), 531–547.
- [49] 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.
- [50] Russell Fischer, Joseph Ferraro, Prince Lal, and Hugh Lusted. 2007. Apparatus and method for the measurement and monitoring of bioelectric signal patterns. US Patent App. 11/549,862, Filed October 16th., 2006, Issued May 17th., 2007.
- [51] Ronald A. Fisher. 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 7, 2 (1936), 179–188.
- [52] Biopac Blood Flow Monitor. [n.d.]. Biopac Systems, Inc. ISO 9001: 2000. <https://www.biopac.com/biopac-blood-flow-measurement-devices-recognized/>.
- [53] Jonathan R. Folstein and Cyma Van Petten. 2008. Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology* 45, 1 (2008), 152–170.
- [54] Talitha C. Ford, Will Woods, and David P. Crewther. 2018. Magnetoencephalography reveals an increased non-target 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>
- [55] A. Furdea, S. Halder, D. J. Krusienski, D. Bross, F. Nijboer, N. Birbaumer, and A. Kübler. 2009. An auditory oddball (P300) spelling system for brain-computer interfaces. *Psychophysiology* 46, 3 (2009), 617–625. DOI: <https://doi.org/10.1111/j.1469-8986.2008.00783.x>
- [56] Laurence J. Garey. 1994. *Brodman’s Localisation in the Cerebral Cortex*. World Scientific.
- [57] Rainer Goebel. 2012. BrainVoyager—past, present, future. *Neuroimage* 62, 2 (2012), 748–756.
- [58] A. A. Goldstein. 1962. Cauchy’s method of minimization. *Numer. Math.* 4, 1 (1962), 146–150.
- [59] Michel Grabisch. 1997. K-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems* 92, 2 (1997), 167–189.
- [60] Moritz Grosse-Wenstrup, Klaus Gramann, and Martin Buss. 2007. Adaptive spatial filters with predefined region of interest for EEG based brain-computer-interfaces. *Advances in Neural Information Processing Systems* 19 (2007), 537.
- [61] M. Grosse-Wenstrup, C. Liefhold, K. Gramann, and M. Buss. 2009. Beamforming in noninvasive brain computer interfaces. *IEEE Transactions on Biomedical Engineering* 56, 4 (April 2009), 1209–1219. DOI : <https://doi.org/10.1109/TBME.2008.2009768>
- [62] Andrzej Grzybowski and Krzysztof Pietrzak. 2013. Napoleon Cybulski (1854–1919). *Journal of Neurology* 260, 11 (2013), 2942–2943.
- [63] Inan Guler and Elif Derya Ubeyli. 2007. Multiclass support vector machines for EEG-signals classification. *IEEE Transactions on Information Technology in Biomedicine* 11, 2 (2007), 117–126.
- [64] Sebastian Halder, Andreas Pinegger, Ivo Käthner, Selina C. Wriessnegger, Josef Faller, João B. Pires Antunes, Gernot R. Müller-Putz, and Andrea Kübler. 2015. Brain-controlled applications using dynamic P300 speller matrices. *Artificial Intelligence in Medicine* 63, 1 (2015), 7–17.
- [65] Martin Han, Panya S. Manoonkitiwongsa, Cindy X. Wang, and Douglas B. McCreery. 2012. In vivo validation of custom-designed silicon-based microelectrode arrays for long-term neural recording and stimulation. *IEEE Transactions on Biomedical Engineering* 59, 2 (2012), 346–354.
- [66] Takashi Hanakawa, Yukinori Katsumi, Hidenao Fukuyama, Manabu Honda, Takuya Hayashi, Jun Kimura, and Hiroshi Shibasaki. 1999. Mechanisms underlying gait disturbance in Parkinson’s disease—A single photon emission computed tomography study. *Brain* 122, 7 (1999), 1271–1282.
- [67] Ernst Haselsteiner and Gert Pfurtscheller. 2000. Using time-dependent neural networks for EEG classification. *IEEE Transactions on Rehabilitation Engineering* 8, 4 (2000), 457–463.
- [68] Hideaki Hayashi, Taro Shibanoki, Keisuke Shima, Yuichi Kurita, and Toshio Tsuji. 2015. A recurrent probabilistic neural network with dimensionality reduction based on time-series discriminant component analysis. *IEEE Transactions on Neural Networks and Learning Systems* 26, 12 (2015), 3021–3033.

- [69] Lianghua He, Die Hu, Meng Wan, Ying Wen, Karen M. von Deneen, and MengChu Zhou. 2016. Common Bayesian network for classification of EEG-based multiclass motor imagery BCI. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 46, 6 (2016), 843–854.
- [70] Jeanny Hérault, Christian Jutten, and John S. Denker. 1986. Space or time adaptive signal processing by neural network models. In *AIP Conference Proceedings*, Vol. 151. AIP, 206–211.
- [71] Charlotte Herzmann, 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.
- [72] Junichi Hori and Naoto Okada. 2017. Classification of tactile event-related potential elicited by Braille display for brain–computer interface. *Biocybernetics and Biomedical Engineering* 37, 1 (2017), 135–142. DOI : <https://doi.org/10.1016/j.bbe.2016.10.007>
- [73] Harold Hotelling. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24, 6 (1933), 417.
- [74] Norden Eh Huang. 2014. *Hilbert-Huang Transform and Its Applications*. Vol. 16. World Scientific.
- [75] Hisao Ishibuchi, Kitaek Kwon, and Hideo Tanaka. 1995. A learning algorithm of fuzzy neural networks with triangular fuzzy weights. *Fuzzy Sets and Systems* 71, 3 (1995), 277–293.
- [76] Jyh-Shing Roger Jang, Chuen-Tsai Sun, and Eiji Mizutani. 1997. Neuro-fuzzy and soft computing, a computational approach to learning and machine intelligence. *IEEE Transactions on automatic control* 42 (1997), 1482–1484.
- [77] 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.
- [78] Jing Jin, Brendan Z. Allison, Eric W. Sellers, Clemens Brunner, Petar Horke, Xingyu Wang, and Christa Neuper. 2011. An adaptive P300-based control system. *Journal of Neural Engineering* 8, 3 (2011), 036006.
- [79] Jing Jin, Brendan Z. Allison, Yu Zhang, Xingyu Wang, and Andrzej Cichocki. 2014. An ERP-based BCI using an oddball paradigm with different faces and reduced errors in critical functions. *International Journal of Neural Systems* 24, 08 (2014), 1450027.
- [80] Jing Jin, Eric W. Sellers, Sijie Zhou, Yu Zhang, Xingyu Wang, and Andrzej Cichocki. 2015. A P300 brain–computer interface based on a modification of the mismatch negativity paradigm. *International Journal of Neural Systems* 25, 03 (2015), 1550011.
- [81] Jing Jin, Hanhan Zhang, Ian Daly, Xingyu Wang, and Andrzej Cichocki. 2017. An improved P300 pattern in BCI to catch user’s attention. *Journal of Neural Engineering* 14, 3 (2017), 036001.
- [82] Valer Jurcak, Daisuke Tsuzuki, and Ippaita Dan. 2007. 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. *Neuroimage* 34, 4 (2007), 1600–1611.
- [83] H. Kang, Y. Nam, and S. Choi. 2009. Composite common spatial pattern for subject-to-subject transfer. *IEEE Signal Processing Letters* 16, 8 (Aug 2009), 683–686. DOI : <https://doi.org/10.1109/LSP.2009.2022557>
- [84] Suguru Kanoga, Masaki Nakanishi, and Yasue Mitsukura. 2016. Assessing the effects of voluntary and involuntary eyeblinks in independent components of electroencephalogram. *Neurocomputing* 193 (2016), 20–32. DOI : <https://doi.org/10.1016/j.neucom.2016.01.057>
- [85] Matthias Kaper, Peter Meinicke, Ulf Grossekhoefer, Thomas Lingner, and Helge Ritter. 2004. BCI competition 2003-data set IIb: Support vector machines for the P300 speller paradigm. *IEEE Transactions on Biomedical Engineering* 51, 6 (2004), 1073–1076.
- [86] Nikola K. Kasabov, Maryam Gholami Doborjeh, and Zohreh Gholami Doborjeh. 2017. Mapping, learning, visualization, classification, and understanding of fMRI Data in the NeuCube evolving spatiotemporal data machine of spiking neural networks. *IEEE Transactions on Neural Networks and Learning Systems* 28, 4 (2017), 887–899.
- [87] P. Kidmose, D. Looney, M. Ungstrup, M. L. Rank, and D. P. Mandic. 2013. A study of evoked potentials from ear-EEG. *IEEE Transactions on Biomedical Engineering* 60, 10 (Oct 2013), 2824–2830. DOI : <https://doi.org/10.1109/TBME.2013.2264956>
- [88] 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.
- [89] Andrea Kübler, Adrian Furdea, Sebastian Halder, Eva Maria Hammer, Femke Nijboer, and Boris Kotchoubey. 2009. A brain–computer interface controlled auditory event-related potential (P300) spelling system for locked-in patients. *Annals of the New York Academy of Sciences* 1157, 1 (2009), 90–100.
- [90] Anand Kumar. 1977. A real-time system for pattern recognition of human sleep stages by fuzzy system analysis. *Pattern Recognition* 9, 1 (1977), 43–46.
- [91] Satish Kumar. 2004. *Neural Networks: A Classroom Approach*. Tata McGraw-Hill Education.
- [92] Sudhir Kumar, G. Rajsekher, Subhashini Prabhakar, et al. 2005. Positron emission tomography in neurological diseases. *Neurology India* 53, 2 (2005), 149.

- [93] Darren J. Leamy, Tomás E. Ward, and Kevin Sweeney. 2011. Functional near infrared spectroscopy (fNIRS) synthetic data generation. In *Proceedings of the 33rd Annual International Conference of the IEEE EMBS*. IEEE, 6589–6592.
- [94] Jong-Hwan Lee, Jeongwon Ryu, Ferenc A. Jolesz, Zang-Hee Cho, and Seung-Schik Yoo. 2009. Brain-machine interface via real-time fMRI: Preliminary study on thought-controlled robotic arm. *Neuroscience Letters* 450, 1 (2009), 1–6.
- [95] Kee-Keun Lee, Jiping He, Amarjit Singh, and Bruce Kim. 2003. Benzocyclobutene (BCB) based intracortical neural implant. In *Proceedings of the International Conference on MEMS, NANO and Smart Systems*. IEEE, 418–422.
- [96] S. H. Lee, S. H. Jin, G. Jang, Y. J. Lee, J. An, and H. K. Shik. 2015. Cortical activation pattern for finger movement: A feasibility study towards a fNIRS based BCI. In *10th Asian Control Conference (ASCC'15)*. IEEE, 1–4.
- [97] J. Legény, R. Viciano-Abad, and A. Lécuyer. 2013. Toward contextual SSVEP-based BCI controller: Smart activation of stimuli and control weighting. *IEEE Transactions on Computational Intelligence and AI in Games* 5, 2 (June 2013), 111–116. DOI : <https://doi.org/10.1109/TCLAI.2013.2252348>
- [98] Gang Leng, Girijesh Prasad, and Thomas Martin McGinnity. 2004. An on-line algorithm for creating self-organizing fuzzy neural networks. *Neural Networks* 17, 10 (2004), 1477–1493.
- [99] 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>
- [100] B. Lipinski, Hans Herzog, E. Rota Kops, Walter Oberschelp, and H. W. Muller-Gartner. 1997. Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information. *IEEE Transactions on Medical Imaging* 16, 2 (1997), 129–136.
- [101] F. Lotte and C. Guan. 2011. Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms. *IEEE Transactions on Biomedical Engineering* 58, 2 (Feb 2011), 355–362. DOI : <https://doi.org/10.1109/TBME.2010.2082539>
- [102] Haiping Lu, Konstantinos N. Plataniotis, and Anastasios N. Venetsanopoulos. 2009. Regularized common spatial patterns with generic learning for EEG signal classification. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'09)*. IEEE, 6599–6602.
- [103] Steven J. Luck. 2012. Event-related potentials. *APA Handbook of Research Methods in Psychology* 1 (2012), 523–546.
- [104] Huairen Luo and Sadasivan Puthusserypady. 2007. Estimation of the hemodynamic response of fMRI data using RBF neural network. *IEEE Transactions on Biomedical Engineering* 54, 8 (2007), 1371–1381.
- [105] Phan Luu, Tobias Flaisch, and Don M. Tucker. 2000. Medial frontal cortex in action monitoring. *Journal of Neuroscience* 20, 1 (2000), 464–469.
- [106] George D. Magoulas, Michael N. Vrahatis, and George S. Androulakis. 1999. Improving the convergence of the backpropagation algorithm using learning rate adaptation methods. *Neural Computation* 11, 7 (1999), 1769–1796.
- [107] Stephane G. Mallat. 1989. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11, 7 (1989), 674–693.
- [108] Zhihong Man, Hong Ren Wu, Sophie Liu, and Xinghuo Yu. 2006. A new adaptive backpropagation algorithm based on Lyapunov stability theory for neural networks. *IEEE Transactions on Neural Networks* 17, 6 (2006), 1580–1591.
- [109] Berta Marti-Fuster, Oscar Esteban, Kris Thielemans, Xavier Setoain, Andres Santos, Domenec Ros, and Javier Pavia. 2014. Including anatomical and functional information in MC simulation of PET and SPECT brain studies. Brain-VISET: A voxel-based iterative method. *IEEE Transactions on Medical Imaging* 33, 10 (2014), 1931–1938.
- [110] Anna Martin, Matthias Schurz, Martin Kronbichler, and Fabio Richlan. 2015. Reading in the brain of children and adults: A meta-analysis of 40 functional magnetic resonance imaging studies. *Human Brain Mapping* 36, 5 (2015), 1963–1981.
- [111] Warren S. McCulloch and Walter Pitts. 1943. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics* 5, 4 (1943), 115–133.
- [112] B. D. Mensh, J. Werfel, and H. S. Seung. 2004. BCI competition 2003-data set Ia: Combining gamma-band power with slow cortical potentials to improve single-trial classification of electroencephalographic signals. *IEEE Transactions on Biomedical Engineering* 51, 6 (June 2004), 1052–1056. DOI : <https://doi.org/10.1109/TBME.2004.827081>
- [113] G. R. Muller-Putz and G. Pfurtscheller. 2008. Control of an electrical prosthesis with an SSVEP-based BCI. *IEEE Transactions on Biomedical Engineering* 55, 1 (Jan 2008), 361–364. DOI : <https://doi.org/10.1109/TBME.2007.897815>
- [114] G. R. Muller-Putz, R. Scherer, C. Neuper, and G. Pfurtscheller. 2006. Steady-state somatosensory evoked potentials: Suitable brain signals for brain-computer interfaces? *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 14, 1 (March 2006), 30–37.
- [115] D. G. M. Murphy, E. M. Daly, T. Van Amelsvoort, D. Robertson, A. Simmons, and H. D. Critchley. 1998. Functional neuroanatomical dissociation of verbal, visual and spatial working memory. *Schizophrenia Research* 29, 1–2 (1998), 105–106. DOI : [https://doi.org/10.1016/S0920-9964\(97\)88566-0](https://doi.org/10.1016/S0920-9964(97)88566-0) Abstracts of the Ninth Biennial Winter Workshop on Schizophrenia.
- [116] 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.

- [117] Ernst Niedermeyer and F. H. Lopes da Silva. 2005. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins.
- [118] 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.
- [119] B. Obermaier, G. R. Muller, and G. Pfurtscheller. 2003. “Virtual keyboard” controlled by spontaneous EEG activity. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 11, 4 (Dec 2003), 422–426.
- [120] Paul P. Von Bnau, Frank C. Meinecke, Simon Scholler, and Klaus-Robert Muller. 2010. Finding stationary brain sources in EEG data. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC’10)*. IEEE, 2810–2813.
- [121] Ramaswamy Palaniappan. 2006. Utilizing gamma band to improve mental task based brain-computer interface design. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 14, 3 (2006), 299–303.
- [122] 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.
- [123] Tobias Pistohl, Tonio Ball, Andreas Schulze-Bonhage, Ad Aertsen, and Carsten Mehring. 2008. Prediction of arm movement trajectories from ECoG-recordings in humans. *Journal of Neuroscience Methods* 167, 1 (2008), 105–114. DOI: <https://doi.org/10.1016/j.jneumeth.2007.10.001>
- [124] C. Pokorny, C. Breitwieser, and G. R. Müller-Putz. 2014. A tactile stimulation device for EEG measurements in clinical use. *IEEE Transactions on Biomedical Circuits and Systems* 8, 3 (June 2014), 305–312.
- [125] Alain Rakotomamonjy and Vincent Guigue. 2008. BCI competition III: Dataset II-ensemble of SVMs for BCI P300 speller. *IEEE Transactions on Biomedical Engineering* 55, 3 (2008), 1147–1154.
- [126] Rabie A. Ramadan and Athanasios V. Vasilakos. 2017. Brain computer interface: Control signals review. *Neurocomputing* 223 (2017), 26–44.
- [127] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller. 2000. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering* 8, 4 (Dec 2000), 441–446. DOI: <https://doi.org/10.1109/86.895946>
- [128] P. Rinck. 2014. Magnetic resonance: A critical peer-reviewed introduction. In *Magnetic Resonance in Medicine. The Basic Textbook of the European Magnetic Resonance forum* (8th. ed.). <http://www.magnetic-resonance.org/ch/21-01.html>.
- [129] Nigel C. Rogasch, Richard H. Thomson, Faranak Farzan, Bernadette M. Fitzgibbon, Neil W. Bailey, Julio C. Hernandez-Pavon, Zafiris J. Daskalakis, and Paul B. Fitzgerald. 2014. Removing artefacts from TMS-EEG recordings using independent component analysis: Importance for assessing prefrontal and motor cortex network properties. *NeuroImage* 101 (2014), 425–439. DOI: <https://doi.org/10.1016/j.neuroimage.2014.07.037>
- [130] Frank Rosenblatt. 1958. The perceptron A probabilistic model for information storage and organization in the brain. *Psychological Review* 65, 6 (1958), 386.
- [131] J. Peter Rosenfeld, Elena Labkovsky, Michael Winograd, Ming A. Lui, Catherine Vandenboom, and Erica Chedid. 2008. The complex trial protocol (CTP): A new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology* 45, 6 (2008), 906–919.
- [132] A. S. Royer, A. J. Doud, M. L. Rose, and B. He. 2010. EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18, 6 (Dec 2010), 581–589.
- [133] André Salomon, Benjamin Goldschmidt, René Botnar, Fabian Kiessling, and Volkmar Schulz. 2012. A self-normalization reconstruction technique for PET scans using the positron emission data. *IEEE Transactions on Medical Imaging* 31, 12 (2012), 2234–2240.
- [134] Wojciech Samek, Motoaki Kawanabe, and Carmen Vidaurre. 2011. *Group-wise Stationary Subspace Analysis a Novel Method for Studying Non-stationarities*. Citeseer.
- [135] W. Samek, K. R. Muller, M. Kawanabe, and C. Vidaurre. 2012. Brain-computer interfacing in discriminative and stationary subspaces. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2873–2876. DOI: <https://doi.org/10.1109/EMBC.2012.6346563>
- [136] 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.
- [137] Marten K. Scheffers and Michael G. H. Coles. 2000. Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance* 26, 1 (2000), 141.
- [138] Vincent J. Schmithorst, Scott K. Holland, and Elena Plante. 2007. Object identification and lexical/semantic access in children: A functional magnetic resonance imaging study of word-picture matching. *Human Brain Mapping* 28, 10 (2007), 1060–1074.



- [139] Alexander J. Shackman, Brenton W. McMenamin, Jeffrey S. Maxwell, Lawrence L. Greischar, and Richard J. Davidson. 2009. Right dorsolateral prefrontal cortical activity and behavioral inhibition. *Psychological Science* 20, 12 (2009), 1500–1506. DOI : <https://doi.org/10.1111/j.1467-9280.2009.02476.x>
- [140] David W. Shattuck and Richard M. Leahy. 2002. BrainSuite: An automated cortical surface identification tool. *Medical Image Analysis* 6, 2 (2002), 129–142.
- [141] Mark J. Shensa. 1992. The discrete wavelet transform: Wedding the a trous and Mallat algorithms. *IEEE Transactions on Signal Processing* 40, 10 (1992), 2464–2482.
- [142] Ranganatha Sitaram, Nikolaus Weiskopf, Andrea Caria, Ralf Veit, Michael Erb, and Niels Birbaumer. 2008. fMRI brain-computer interfaces. *IEEE Signal Processing Magazine* 25, 1 (2008), 95–106.
- [143] Edward E. Smith, John Jonides, and Robert A. Koeppe. 1996. Dissociating verbal and spatial working memory using PET. *Cerebral Cortex* 6, 1 (1996), 11–20.
- [144] Estate M. Sokhadze, Manuel F. Casanova, Emily L. Casanova, Eva Lamina, Desmond P. Kelly, and Irma Khachidze. 2017. Event-related potentials (ERP) in cognitive neuroscience research and applications. *NeuroRegulation* 4, 1 (2017), 14.
- [145] Qing Song. 2010. On the weight convergence of Elman networks. *IEEE Transactions on Neural Networks* 21, 3 (2010), 463–480.
- [146] Bettina Sorger, Joel Reithler, Brigitte Dahmen, and Rainer Goebel. 2012. A real-time fMRI-based spelling device immediately enabling robust motor-independent communication. *Current Biology* 22, 14 (2012), 1333–1338. DOI : <https://doi.org/10.1016/j.cub.2012.05.022>
- [147] Donald F. Specht. 1990. Probabilistic neural networks. *Neural Networks* 3, 1 (1990), 109–118.
- [148] Franca Tecchio, Francesca Benassi, Filippo Zappasodi, Leonardo Emberti Gialloreti, Mark Palermo, Stefano Seri, and Paolo Maria Rossini. 2003. Auditory sensory processing in autism: A magnetoencephalographic study. *Biological Psychiatry* 54, 6 (2003), 647–654.
- [149] Gerald Tesauro, Yu He, and Subutai Ahmad. 1989. Asymptotic convergence of backpropagation. *Neural Computation* 1, 3 (1989), 382–391.
- [150] 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>
- [151] MAO Thijssen, KR Bijkerk, and RJM Van Der Burght. 1998. Manual: Contrast-detail phantom Artinis CDRAD type 2.0. *University Medical Center Nijmegen* (1998). [https://www.radiograf.dk/fileadmin/user\\_upload/dokumenter/Bacheloropgaver/Billedkvalitet\\_ved\\_CR\\_og\\_DR\\_-\\_et\\_fantomforsoeg/CR\\_og\\_DR\\_Bilag\\_C.pdf](https://www.radiograf.dk/fileadmin/user_upload/dokumenter/Bacheloropgaver/Billedkvalitet_ved_CR_og_DR_-_et_fantomforsoeg/CR_og_DR_Bilag_C.pdf).
- [152] Wu Ting, Yan Guo-zheng, Yang Bang-hua, and Sun Hong. 2008. {EEG} feature extraction based on wavelet packet decomposition for brain computer interface. *Measurement* 41, 6 (2008), 618–625. DOI : <https://doi.org/10.1016/j.measurement.2007.07.007>
- [153] Angela Tooker, Daniel Liu, Emily B. Anderson, Sarah Felix, Kedar G. Shah, Kye Young Lee, Jason E. Chung, Satinderpall Pannu, Loren Frank, and Vanessa Tolosa. 2014. Towards a large-scale recording system: Demonstration of polymer-based penetrating array for chronic neural recording. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'14)*. IEEE, 6830–6833.
- [154] 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.
- [155] L. J. Trejo, R. Rosipal, and B. Matthews. 2006. Brain-computer interfaces for 1-D and 2-D cursor control: Designs using volitional control of the EEG spectrum or steady-state visual evoked potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 14, 2 (June 2006), 225–229. DOI : <https://doi.org/10.1109/TNSRE.2006.875578>
- [156] Mario Tudor, Lorainne Tudor, and Katarina Ivana Tudor. 2004. Hans Berger (1873-1941)—the history of electroencephalography. *Acta Medica Croatica: Casopis Hrvatske Akademije Medicinskih Znanosti* 59, 4 (2005), 307–313.
- [157] 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).
- [158] Nischal K. Verma, L. S. Vishnu Sai Rao, and Suresh K. Sharma. 2014. Motor imagery EEG signal classification on DWT and crosscorrelated signal features. In *9th International Conference on Industrial and Information Systems (ICIIS'14)*. IEEE, 1–6.
- [159] Jacques J. Vidal. 1973. Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering* 2, 1 (1973), 157–180.
- [160] Wilhelm Von Rosenberg, Theerasak Chanwimalueang, Valentin Goverdovsky, David Looney, David Sharp, and Danilo P. Mandic. 2016. Smart helmet: Wearable multichannel ECG and EEG. *IEEE Journal of Translational Engineering in Health and Medicine* 4 (2016) 1–11.

- [161] Stephan Waldert, Hubert Preissl, Evariste Demandt, Christoph Braun, Niels Birbaumer, Ad Aertsen, and Carsten Mehring. 2008. Hand movement direction decoded from MEG and EEG. *Journal of Neuroscience* 28, 4 (2008), 1000–1008.
- [162] W. Grey Walter. 1964. Contingent negative variation: An electric sign of sensori-motor association and expectancy in the human brain. *Nature* 230 (1964), 380–384.
- [163] Baojun Wang, Liu Jun, Jing Bai, Le Peng, Guang Li, and Yan Li. 2006. EEG recognition based on multiple types of information by using wavelet packet transform and neural networks. In *27th Annual International Conference of the Engineering in Medicine and Biology Society (IEEE-EMBS'05)*. IEEE, 5377–5380.
- [164] Wei Wang, Gustavo P. Sudre, Yang Xu, Robert E. Kass, Jennifer L. Collinger, Alan D. Degenhart, Anto I. Bagic, and Douglas J. Weber. 2010. Decoding and cortical source localization for intended movement direction with MEG. *Journal of Neurophysiology* 104, 5 (2010), 2451–2461.
- [165] Nikolaus Weiskopf, Klaus Mathiak, Simon W. Bock, Frank Scharnowski, Ralf Veit, Wolfgang Grodd, Rainer Goebel, and Niels Birbaumer. 2004. Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE Transactions on Biomedical Engineering* 51, 6 (2004), 966–970.
- [166] Erwin Josef Speckmann Alexander I. Roitbak (eds.) Wolfgang Haschke (auth.), Wolfgang Haschke. 1993. *Slow Potential Changes in the Brain*. Birkhäuser Basel. <http://gen.lib.rus.ec/book/index.php?md5=9b1a21cc2fbeb0ef27c59246ff19374>.
- [167] Shiqian Wu and Meng Joo Er. 2000. Dynamic fuzzy neural networks—a novel approach to function approximation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 30, 2 (2000), 358–364.
- [168] Shang-Lin Wu, Yu-Ting Liu, Tsung-Yu Hsieh, Yang-Yin Lin, Chih-Yu Chen, Chun-Hsiang Chuang, and Chin-Teng Lin. 2017. Fuzzy integral with particle swarm optimization for a motor-imagery-based brain-computer interface. *IEEE Transactions on Fuzzy Systems* 25, 1 (2017), 21–28.
- [169] Neng Xu, Xiaorong Gao, Bo Hong, Xiaobo Miao, Shangkai Gao, and Fusheng Yang. 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 (2004), 1067–1072.
- [170] Qi Xu, Hui Zhou, Yongji Wang, and Jian Huang. 2009. Fuzzy support vector machine for classification of EEG signals using wavelet-based features. *Medical Engineering & Physics* 31, 7 (2009), 858–865.
- [171] Y. F. Yam and T. W. S. Chow. 1993. Extended backpropagation algorithm. *Electronics Letters* 29, 19 (1993), 1701–1702.
- [172] Bang-hua Yang, Guo-zheng Yan, Ting Wu, and Rong-guo Yan. 2007. Subject-based feature extraction using fuzzy wavelet packet in brain-computer interfaces. *Signal Processing* 87, 7 (2007), 1569–1574.
- [173] Bang-hua Yang, Guo-zheng Yan, Rong-guo Yan, and Ting Wu. 2007. Adaptive subject based feature extraction in brain computer interfaces using wavelet packet best basis decomposition. *Medical Engineering & Physics* 29, 1 (2007), 48–53.
- [174] Sheng-Sung Yang, Chia-Lu Ho, and Chien-Min Lee. 2006. HBP: Improvement in BP algorithm for an adaptive MLP decision feedback equalizer. *IEEE Transactions on Circuits and Systems II: Express Briefs* 53, 3 (2006), 240–244.
- [175] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu. 2015. A dynamically optimized SSVEP brain computer interface (BCI) speller. *IEEE Transactions on Biomedical Engineering* 62, 6 (June 2015), 1447–1456. DOI : <https://doi.org/10.1109/TBME.2014.2320948>
- [176] H. Yuan and B. He. 2014. Brain computer interfaces using sensorimotor rhythms: Current state and future perspectives. *IEEE Transactions on Biomedical Engineering* 61, 5 (May 2014), 1425–1435. DOI : <https://doi.org/10.1109/TBME.2014.2312397>
- [177] L. A. Zadeh. 1965. Fuzzy sets. *Information and Control* 8, 3 (1965), 338–353. DOI : [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
- [178] Lotfi A. Zadeh. 1968. Fuzzy algorithms. *Information and Control* 12, 2 (1968), 94–102.
- [179] Jianhai Zhang, Ming Chen, Sanqing Hu, Yu Cao, and Robert Kozma. 2016. PNN for EEG-based emotion recognition. In *IEEE International Conference on Systems, Man, and Cybernetics (SMC'16)*. IEEE, 002319–002323.
- [180] Fan Zhou, Jun Liu, Yi Yu, Xiang Tian, Hui Liu, Yaoyao Hao, Shaomin Zhang, Weidong Chen, Jianhua Dai, and Xiaoxiang Zheng. 2010. Field-programmable gate array implementation of a probabilistic neural network for motor cortical decoding in rats. *Journal of Neuroscience Methods* 185, 2 (2010), 299–306.
- [181] Salim Çinar and Nurettin Acir. 2017. A novel system for automatic removal of ocular artefacts in EEG by using outlier detection methods and independent component analysis. *Expert Systems with Applications* 68 (2017), 36–44. DOI : <https://doi.org/10.1016/j.eswa.2016.10.009>

Received October 2017; revised August 2018; accepted November 2018