# Chapter 3

# A new approach to feature extraction in MI-based BCI systems

Arefeh Nouri[a], Zahra Ghanbari[a], Mohammad Reza Aslani[b], Mohammad Hassan Moradi[a]

[a] Amirkabir University of Technology, Department of Biomedical Engineering, Tehran, Iran
[b] Shahab Danesh University, Electrical Engineering Department, Qom, Iran

an.mp@aut.ac.ir, zahraghanbari@aut.ac.ir, mr.aslani@shdu.ac.ir, mhmoradi@aut.ac.ir

## Abstract

Motor imagery-based brain-computer interface (MI-BCI) creates a path through which the brain interacts with the external environment by recording and processing EEG signals made by imagining the movement of a particular limb. In this study, with the aim of improving classification accuracy, we modify the feature extraction stage. After pre-processing and decomposing EEG signals into their frequency bands, we apply Common Spatial Pattern (CSP) algorithm to each sub-band. Then, using spatial filters and blind source separation (BSS) techniques, we identify features of brain sources, rather than only channels. Finally, by selecting the appropriate features as the input of the classification block, we discern the imagined classes. To evaluate our proposed method, we used the data set IVa of BCI competition III, which includes two classes: right hand and foot. The results of various experiments explain that the accuracy of the system is on average 98.8% and the sensitivity is 100% for all subjects and indicate that the proposed method can be used to improve the MI-based BCI.

Keywords: BCI, BSS, feature extraction, EEG, spatial filter

## 3.1 Introduction

Brain computer interface contains a set of sensors and a signal processing section that directly converts brain activities into a series of communication or controlling signals. They have three main tasks: collect the brain signals, translate them, and send output commands to a connected device. Brain signals are multiplied in a matrix of unknown coefficients and then recorded by a neural recording device as observation signals. So, it is better to obtain the real signals using BSS methods. Based on the method used to gather brain signals, BCIs can be divided into three categories of invasive, semi-invasive, and non-invasive. In invasive types of BCI, the micro-electrodes are implanted directly into the cortex, during neurosurgery. The signal quality of invasive BCI is the highest, but because neurosurgery can be a high-risk and costly procedure, it is not widely used. In section 2 and 4, we will

briefly review the main components, required steps, and applications of BCI.
To have a more practical BCI system, we propose a method in which, using CSP-based channel selection, we selected a limited number of electrodes as useful channels. Then, using an appropriate tool for Underdetermined Blind Source Separation (UBSS), we separate the brain sources and, by selecting the proper features, classify the imagined classes. In Section 3 we will explain some about BSS and its application in BCI and in Section 4 we will introduce the proposed method. Sections 6 and 7 contain the results obtained and the conclusion.

## 3.2 Types and applications

Comprehension of brain and the neurophysiological phenomena occur during various brain activities, is essential for designing BCI systems. To understand what is happening in the human brain, brain waves should be recorded using neural recording devices. Electroencephalograph (EEG) signals are usually used for this purpose due to its high time accuracy, portability, and ease of use [1]. In general, however, various brain signals recorded by different modalities can be used in BCI applications.

### 3.2.1 EEG signals

In recording brain signals using an EEG device, electrodes are placed on the surface of the scalp. The electric field produced by the activity of neurons is measured. The amplitude of the recorded signals is very small which is measured in microvolts. The main frequencies of human EEG waves are: delta, theta, alpha, beta and gamma, which can be divided into sub-bands themselves.
Several approaches are used in BCI using EEG signals. Among various EEG patterns, steady-state visual evoked potential (SSVEP)-based BCI system has been widely studied because of its fast-training rate, high signal-to-noise ratio, and high information transfer rate. SSVEP is a spontaneous evoked response induced by a visual stimulus, flickering at a constant frequency between approximately 6 to 100Hz. The response manifests as an increase in the amplitude of the stimulated frequency. SSVEP caused by an oscillating stimulus, whereas P300 is a transient neural event that can be elicited as a result of focusing at a transient stimulus. BCI systems which use P300 components often use visual tasks. However, P300-based BCI have been implemented using auditory or tactile stimuli as well [2,3].
Another prevalent method used to extract specific neural activity patterns from EEGs is motor imagery (MI), in which particular patterns are elicited in the motor cortex during the imagination of movements [4]. Luckily, everyone, including people with disabilities, can develop different mental states such as motor imagery [5]. The motion of different limbs of the body motivates different parts of the brain. In MI, the user is asked to think of a particular action, such as the clockwise rotation of the left hand [6]. These imaginations trigger some areas of the motor cortex. According to the stimulated region, BCI specifies the imagined move [7].

### 3.2.2 Other methods

As mentioned before, other neural recording devices such as Electrocorticography (ECoG), Magnetoencephalography (MEG), Functional near-infrared spectroscopy (fNIRS), etc. are also used for BCI applications [8].
MEG records the neuro-magnetic signals. The superiority of MEG over EEG is higher signal-to-noise ratio and higher spatial resolution, which is very important for BCI applications where the signal is analyzed in real-time [9].
EEG and MEG sensors gather signals on the scalp, which is actually superposition of signals generated by brain sources. Therefore, the active level of the brain cannot be accurately detected based on signals recorded by them. In addition, since measurements are strongly affected by the orientation of the neurons towards the sensors, we have lower resolutions for some areas of the brain using these sensors. So, the idea of using electrodes placed directly on or inside the cerebral cortex was proposed. ECoG is an invasive but reliable neural recording technique used to record high-frequency cortical signals in different areas of the brain [10]. Furthermore, the ECoG method is suitable for developing bilateral BCI, which provide a combination of an efferent loop which processes information from neural activity and an afferent loop which transmits information to the brain. ECoG is less risky than other invasive methods and also with the development of implantable ECoG systems; ECoG-based BCIs offer a wide range of practical solutions for many patients.
The fNIRS is a low-cost portable non-invasive method to develop BCI. The fNIRS measures changes in hemoglobin concentration in the cerebral cortex using infrared radiation [11]. The fNIRS temporal quality is not as good as the EEGs and is usually used as a complement to EEG data. It should be mentioned that a combination of two BCI systems can serve as a hybrid BCI approach. Generally, at least one BCI system combined to other systems can be used to design a hybrid BCI system [12].
No matter what method a BCI uses, but it must meet the following four criteria:

- The device should use its own signals, which are recorded directly from the brain
- There should be at least one recordable signal, which the user can intentionally modify into a targeted behavior
- Processing should be real-time
- The user should receive feedback from the system

### 3.2.3 Applications

Initially, BCI was developed for rehabilitation [13], prosthetic control [12], and neural feedback [14] application leading to the production of rehabilitation devices. In recent years, non-medical fields are widely making benefits from BCI methods. Some of these fields are including marketing, gaming, entertainment, sport, personal health, education, and military industries. Some other applications of noninvasive BCI include mood detection [15], emotions analysis [16], mind-reading, communication, controlling devices such as robotic arms [12], wheelchair [17], etc. and, personal identification systems [18].

## 3.3 BSS and its application in BCI

Blind Source Separation refers to a problem where the sources and the mixing matrix are indistinct and only observation signals are available for the separation procedure. The objective is to separate unknown and independent sources using observation signals. In the case in which the number of sources is equal to or less than the observations, it is called a determined problem and various effective algorithms are offered to solve it such as Independent Component Analysis (ICA), Nonnegative Matrix Factorization (NMF) [19] and, Principal Components Analysis (PCA) [20]. In the other case where the number of observations is less than the mixed sources, the problem is underdetermined and there is no basic solution for it. In BCI, BSS techniques are commonly used to remove noise and artifacts, such as EMG artifacts, from the recorded signals [21,22]. Also, one of the new ways to increment the dimensions of the features is to use the UBSS [23].

## 3.4 Related work

Due to the increasing applications of BCI in various industries, many efforts and studies have been done to improve the efficiency and accuracy of these systems. The main three stages of a BCI system are pre-processing, feature extraction, and classification.

### 3.4.1 Removing noise and artifacts

Removing noise and artifacts is an important part of the pre-processing stage. As some input channels have a low signal-to-noise ratio, they can be considered inefficient. In [7], to reduce the computational load and increase the accuracy of the classification stage, an attempt has been made to eliminate inefficient input signals. Figure 3.1 shows the block diagram of the proposed system in [7], which uses an ICA-based algorithm to identify and eliminate defective channels. The proposed algorithm has used a CSP-based stage to extract features related to each class. Reported results show that finding and removing inefficient input information, leads to improved classification results. As mentioned before, computational load is reduced in this work.


Figure 3.1: Block diagram of the proposed system in [7]

Moreover, the human error made by subjects also can cause problems in the learning of the system. To solve such errors, [24] provides a BSS-based algorithm that identifies and fixes cases in which the subject makes an error. Next, using a Bayesian and Linear Discriminant Analysis (LDA)-based clustering algorithm succeeds in improving classification. Figure 3.2 shows the block diagram of the proposed system in [24].

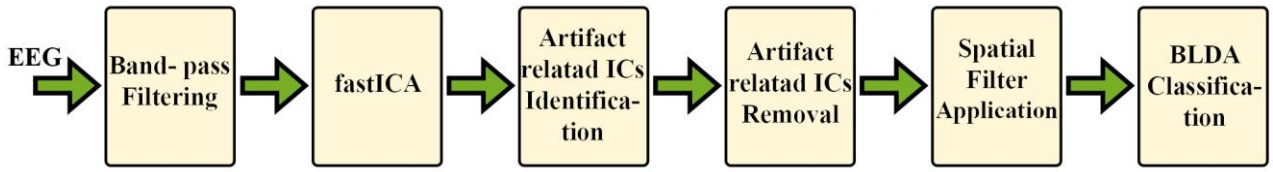


Figure 3.2: Block diagram of the proposed system in [24]

### 3.4.2 Extending input dimensions

In the study [4], an expansion of input dimension has been used to improve the BCI-based learning system. This dimensional expansion is done by input decomposition by wavelet transform. This allows input signals to propagate over time and frequency and make their analysis more efficient. The output of the wavelet transform is then divided into four spectral groups: theta (4-7 Hz), alpha (8-15 Hz), beta (16-31 Hz), and gamma (more than 32 Hz) based on its specified bandwidth, and then enter the stage including CSP algorithms in pairs and parallel. Afterwards, the statistical properties of these outputs are entered into an LDA-based clustering algorithm for classification and then a Naive Bayes-based clustering algorithm. Figure 3.3 shows these steps in the proposed system [4].

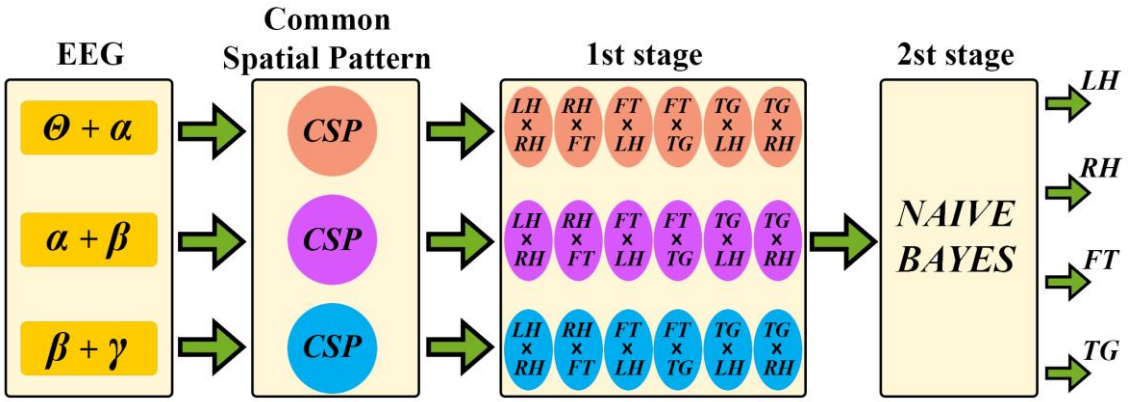


Figure 3.3: Flowchart of the proposed system in [4]

As can be seen, the proper expansion of input dimensions can be useful in creating a better learning system. A simple and effective way to increase number of features is to use a filter bank. A filter bank will be able to reduce the complexity and instability of EEG signals by preserving the harmonics in a certain bandwidth. In articles [25,26], a filter bank block is used before the main analysis begins. Supervised learning algorithms have also been used for the classification stage. According to the results of these researches, the use of a filter bank with overlapping frequency windows can improve system performance in the classification section. It is also clear that the type of clustering algorithm is of importance in the accuracy and speed of systems. Accuracy of the proposed system in [25] is higher than the system proposed in [26], but in contrast, the speed of system in [26] is higher comparing to the system speed in [25].

Numerous studies have been conducted to increase dimension of features that are more

complex than filter banks. Like the system proposed in [27], which consists of wavelet transform and Empirical Mode Decomposition (EMD) algorithm for separating the most detailed components in the EEG. In this work Support Vector Machine (SVM) algorithm is used for classification. Moreover, this approach focuses on entropy criterion analysis for classification. One of the new methods of increasing the size of inputs and increasing the number of features in the learning system is to use the concept of Underdetermined Blind Sources Separation. In [23], an experiment was performed to separate four independent sources from only two observation signals. Using such a concept in BCI can pave the way for the separation of brain sources from a limited number of channels. An example of recent research on UBSS is [22], which has shown that ICA-based algorithms are not efficient enough in isolating muscle movement artifacts from brain sources. Therefore, there is a need for using advanced methods based on UBSS.

### 3.4.3 Common spatial pattern

Most of the proposed systems have focused on improving the clustering section. Results confirm the efficiency of it to increase the accuracy of the BCI system. However, according to [28], the use of spatial filters and their improvement can also have a great impact on the quality of the final system. Fig. 3.4 illustrates the block diagram of the BCI system based on spatial filters proposed in [28].

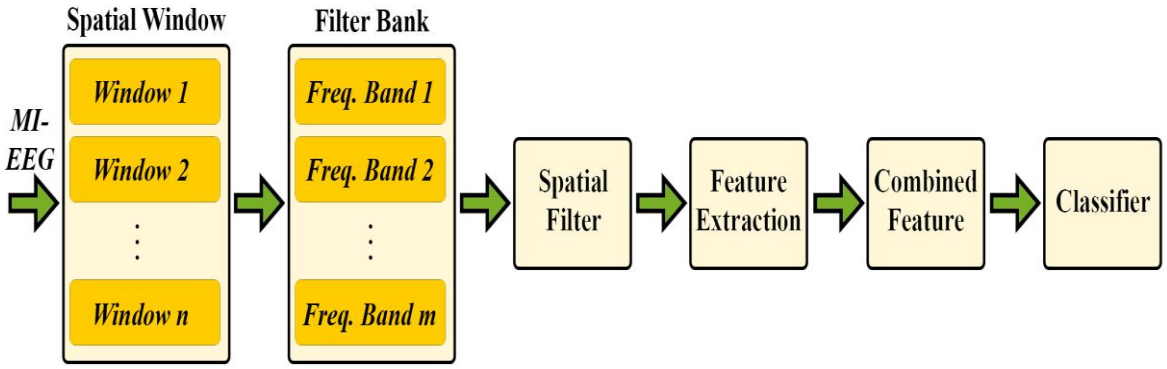


Figure 3.4: Block diagram of system [28]

The above system uses two stages of spatial filter, one to improve the input signals to reduce the effects of noise and artifacts and the second to improve and regulate the effects of brain resources in the different frequency spectrum. Therefore, to design a BCI system, it is necessary to pay special attention to the use of this type of filter in different parts. For example, in a recent study [29], spatial filters were used to eliminate the effect of inefficient channels. In other words, in the proposed system of [29], the greatest focus is on extracting the effect of optimal channels for each class. According to the results of this study, the extraction of useful features has been improved. Moreover, clustering can be done by simpler algorithms.
One way to reduce the complexity of BCI-based systems is to analyze and decide based on a smaller number of EEG samples. In [30], it is declared due to the fact that major potentials associated with motor imagery exist in the alpha and beta bands, it is possible to use a smaller number of samples for EEG analysis. On the other hand, such an approach can reduce the

variance of the target estimate in the clustering stage. Moreover, this study, by combining filter bank and CSP, has introduced a new algorithm called Component Regularized Common Spatial Patterns (CRCSP), which has results with higher accuracy compared to the CSP algorithm.

Another study has been conducted to improve the CSP algorithm and has considered this upgrade as a determining factor in the successful results of its research. In [31], to expand feature dimensions, Short-Time Fourier Transform (STFT) is used to transfer signals into time-frequency domain. The CSP algorithm is then applied to the obtained time-frequency maps. This type of algorithm is called Time-Frequency Common Spatial Pattern.

CSP is basically a two-class feature extraction algorithm and is not designed for multi-class mode. [32] is one of the studies that have generalized this tool from two-class to multi-class. The method of generalization of this research is by pairwise analysis of the useful output characteristics of the CSP algorithm by parallel layers of LDA-based clustering algorithms.

In addition to the inherent characteristics of EEG signals that are weak, unstable, and have a low signal-to-noise ratio, the motor imagernals are different not only for different subjects but also change over time for even one subject because of different physiological and psychological characteristics. In fact, the signals recorded may vary not only between subjects, but also between sessions of an individual subject for a variety of reasons. In [33], a learning transfer algorithm has been used to address the mismatch of distributions in different sessions or subjects to adapt the models.

Learning transfer, stores the knowledge gained in solving a problem and applying it to a different but related problem. By analyzing the power spectrum of data related to MI, common features in different sessions or subjects are extracted. Then, the datasets that were most relevant to the new datasets according to the Euclidean distance are selected to update the common features. Finally, common features and subject/session-specific features are shared to generate a new model. The classification accuracy of the proposed algorithm of [33] is higher than traditional machine learning algorithms.

One of the main reasons that EEG signals are none-stationary is changes in the user's cognitive status due to fatigue, frustration, low stimulation level, etc. In [34], a CSP-based scheme is proposed to overcome user’s mental fatigue. The proposed method uses LDA algorithm for CSP compliance, which uses the breaking ties criterion to select evaluation data. The training data is then updated and the output of the CSP section is calculated based on them. This algorithm is activated in case of high fatigue. The resolution of the features extracted by the proposed adaptive CSP has been compared with the standard CSP with three standard indices including: Davies Bouldin index, Fisher score, and Dunn’s index. Overall, results of this study show that the distinction of motor imagery classes is significantly improved by adaptive CSP compared to conventional CSP.

As mentioned, one of the major challenges in BCI is dealing with the none-stationarity of brain signals that can disrupt the function of BCI. The aim of [35] is identifying the most powerful type of spatial filter among common methods, and also evaluating their performance improvement using Stationary Subspace Analysis (SSA). To get closer to reality, the number of training datasets has been reduced and the test data have been recorded separately 30 minutes after the training data. A significant reduction in the accuracy of test datasets for CSP-

based methods has been observed.
However, the accuracy of the SSA preprocessing algorithm was approximately better than other algorithms. It should be noted that in this study, FBCSP and FBCSPT had higher performance compared to other methods. Also, in a similar study [36] to improve the MI-based BCI by stroke patients, three different methods of spatial filtering have been investigated. The results of this study show that FBCSP is more accurate than other methods.

### 3.4.4 Spatial filtering

Spatial filter and BSS are valuable tools in the analysis of EEG data. Spatial filtering is a method in which the amplitude of EEG in each electrode is changed according to the weight composition of the voltage values in two or more other electrodes. However, blind source separation tries to separate EEG signals into statistically different components, regardless of the underlying physiology of the component.
In [37], a new method for designing a spatial filter to improve the effects of brain sources on EEG signals is proposed using the electrical conduction model of the head and the global optimization technique. The proposed spatial filter tries to identify sources that are maximally different from each other in terms of spatial distribution. In this study, the proposed system is evaluated with both simulated EEG signals and real EEG signals. Based on the results of our proposed method, the system can decompose EEG signals into components with separate spatial distributions that are very similar to the original simulated sources.
Event Related Potentials (ERPs) are small potentials generated in response to an external event or stimulus. Since ERPs have overlaps in time and space, their separation requires using a spatial filter. Static filters such as Surface Laplacian and adaptive filters based on blind resource separation such as ICA and PCA can also be considered as types of spatial filters. The effective performance of adaptive spatial filters has led to the introduction of a new type of time adaptive spatial filter named Local Spatial Analysis (LSA) filter [38]. In the proposed system, all experiments corresponding to different times are used to obtain filter parameters and store statistical information. To ensure the proper performance of the new filter, simulated data, as well as real ERP with four stimuli including auditory, visual, touch, and pain, has been used. The results of this study show the complete superiority of LSA filters over static filters.
Using an appropriate classification model is of great importance in designing a BCI system. In a study with such an approach, the use of a hybrid algorithm consisting of machine learning with extreme kernel cores, Hybrid-KELM, based on PCA and Fisher Discriminant Analysis (FLD) is proposed to classify MI data [39]. Improving kernel cores has a considerable impact on results. Then, basic features that are generally changing over time are extracted. After that, with the aid of PCA, the dimensions of the features have been reduced. Based on the results of [39], the proposed Hybrid-KELM method results in higher accuracy compared to other methods. It also speeds up calculations and decreases error rates. Results of the above proposed system on the BCI Competition III data set show that the average accuracy of this system is 96.54%.
Recently, deep learning neural networks has been highly used in various applications. As the process of classifying EEG signals and discovering vital information must be robust,

automated, and highly accurate, the use of deep learning in the BCI field has also been considered. A review study provides the context needed for the development of BCI based on deep learning [40]. Studies published over last five years try to facilitate selecting the appropriate deep neural network structure and other super-parameters for the development of these systems. Deep neural networks create automated robust systems for classification by utilizing all the input data while learning prominent features. In particular, Convolutional Neural Networks (CNNs) and hybrid-CNNs represent the best performance compared to other networks [41].

Since the analysis of multi-channel EEG signals considerably increases the computational load, in [42] a new method for separating the two modes of imagining right-foot motion and imaging right-hand motion, using single-channel comparative analysis and EEG signal classification is proposed. The multi-cluster unsupervised learning method has also been used to select the appropriate channel. In addition, Flexible Variational Mode Decomposition (F-VMD) is used for adaptive signal analysis. It should be mentioned that F-VMD parameters are selected based on the nature of the EEG signals. Then, properties based on Hjorth parameters, entropy, and quartile criteria are extracted from the output of the F-VMD class. These features are classified using Flexible Extreme Learning Machine (F-ELM). The F-ELM algorithm works by reducing classification error, parameterization, and optimizing kernel cores. Performance of the proposed method was evaluated by measuring five indexes of accuracy (ACC), sensitivity (SEN), specificity (SPE), Math correlation coefficient (MCC), and F1 score. It should be emphasized that the proposed method can work with only one channel and two features, which include a small computational load. Results of evaluation criteria obtained in [42] indicates the high performance of the system proposed in the research.

## 3.5 Proposed method

In this study, we intend to make a low cost and practical BCI system using minimal electrodes to separate brain sources. To reach this aim, using CSP-based channel selection, we limit the number of channels. Then, with the help of a proposed technique based on UBSS, we separate the brain sources and by selecting the appropriate features, we classify the imagined classes.

In BSS, the mixing matrix coefficients determine the location of the sources. This is a physical interpretation of the mixing matrix coefficients, so that each of the coefficients is a definition of a spatial coordinate dimension. According to Fig. 3.5(a), assume that there are two sources along with two recorders (mixtures) in the environment. If source 1 and source 2 close to each other up, then $A_{11}$ get closer to $A_{21}$ and $A_{12}$ get closer to $A_{22}$. This event can prevent source separation by the independent component analysis. In other words, this leads to match the distance of hidden sources to the recorders. Whereby, sources will be embodied as a single source.

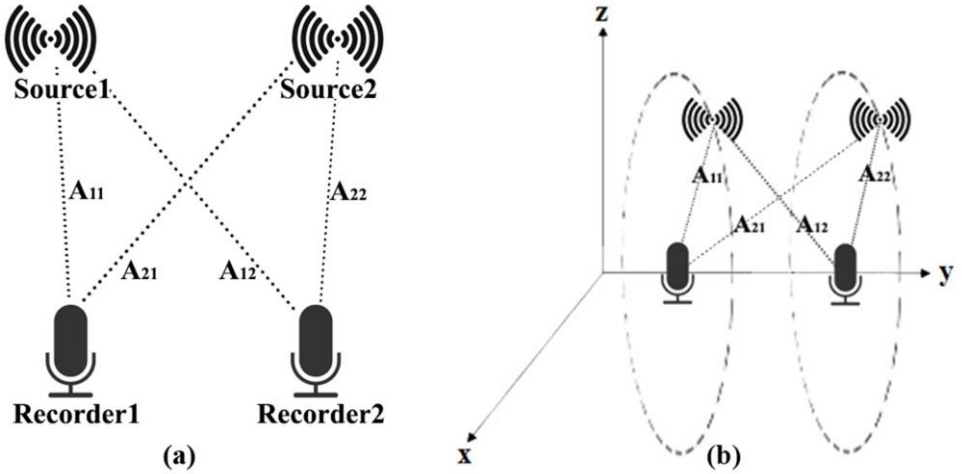


Figure 3.5: (a) Two sources with two recorders (mixtures) in the environment; $A_{11}$, $A_{12}$, $A_{21}$, and $A_{22}$ are mixing matrix coefficients. (b) The location of two different sources and their distance from the two recorders in space.

In general, if any other source is added to the circular shape in Fig. 3.5(b), all of them will be embodied as a single source. In fact, in Fig. 3.5(b) the left circle is the place of the colliding one sphere that centered by left recorder with $A_{11}$ radius and one other sphere that centered by right recorder with $A_{12}$ radius. Also, the right circle is the place of the colliding one sphere that centered by left recorder with $A_{21}$ radius and one other sphere that centered by right recorder with $A_{22}$ radius which we want two circle plate tangent to each other.
The proposed method in this chapter focuses on improving the source extraction in EEGs. In this, by analyzing the energy of observation signals derivatives, the extraction can be achieved. In this chapter, we called the proposed method Derivative Component Analysis, DCA. Now, suppose there are two observations and these signals are linear combination of brain sources and artifacts as the following:

$$X_1(t) = A_{11}S_1(t) + A_{12}S_2(t) + \dots + A_{1\times M}S_M(t) \quad (3.1)$$
$$X_2(t) = A_{21}S_1(t) + A_{22}S_2(t) + \dots + A_{2\times M}S_M(t) \quad (3.2)$$

The main concept is described as the simple equation (3.3) as follows:

$$\Delta X_i(t) = X_2(t) - i\varphi X_1(t) \quad (3.3)$$

Where 'φ' is a very small positive constant factor and 'i' represents iteration index. The number of iterations in DCA is obtained by the equation (3.4):

$$EIN = 5 \times \varphi - 1 \quad (3.4)$$

Where EIN is Effective Iteration Number. For example, the effective iteration number for an experiment with φ=0.01, EIN will be 500, this means that at least 500 iterations are required for a complete analysis of the observations by DCA. In this technique, each point of the iteration is a determiner of correlation between samples of the two observations. Because, the coefficients of the combined components in $\Delta X_i(t)$ is. DCA is looked for iteration points

with the best degree of separation between the independent components. A criterion for analyzing the degree of component separation is according to equation (3.5):

$$DCS_j = corr(\, FTSD_j(t)\, , STSD_j(t)\, ) \tag{3.5}$$

Where $DCS_j$ is degree of component separation in $\Delta X(t)$s, corr is correlation coefficient that its range is [-1, 1], $FTSD_j(t)$ is the first type of the second derivative of $\Delta X(t)$s, $STSD_j(t)$ is the second type of the second derivative of $\Delta X(t)$s. It is noted that 'j' index has a range between [1, i-2]. The $FTSD_j(t)$ and $STSD_j(t)$ are defined as the following:

$$FTSD_j(t) = \Delta X_{i+2}(t) + 2\Delta X_{i+1}(t) - \Delta X_i(t) \tag{3.6}$$

$$STSD_j(t) = \Delta X_{i+2}(t) - \Delta X_i(t) \tag{3.7}$$

In iteration points, if the DCS value is a local peak or valley, then equivalent iteration point will be selected as a new Independence Identification Point (IIP) that shown a suitable degree of separation of the independent components. It is also clear from the equations (3.1), (3.2) and (3.3) that the coefficients of the generated observation signals in iteration points are given by the following:

$$\Delta X_{i+1}(t) = (A_{21} - \mathrm{i}\varphi A_{11})S_1(t) + (A_{22} - \mathrm{i}\varphi A_{12})S_2(t) + \cdots + (A_{1\times M} - \mathrm{i}\varphi A_{1\times M})S_M(t) \tag{3.8}$$

Equation (3.8) is shown that by performing successive differences, new observations can make the suitable degrees of separation between independent components. It means that it is possible to extract some of the brain sources in the UBSS problem according to IIPs. It is indisputable that the quality of extraction does depend on the mixing matrix coefficients. This can be used against DCA technique and reduced its extraction quality. In addition, the effect of increasing the number of source components and its impact on source extraction by DCA technique is examined. In a BCI system, after executing DCA technique, the IIPs of brain sources can be obtained. With the help of these IIPs, EEG samples can be classified. The samples belonging to each class are actually attributed to independent brain sources. Now suppose we have two recorders but one of the recorders contains only one active source and we are aware of its features. In this case, if we run the DCA technic, in a specific IIP, the source-dependent samples are extracted from the recorder with unknown components. If the known source is a simple component such as a sinusoidal signal, the proposed tool becomes a powerful adaptive filter. In this chapter, we have used this tool to accurately analyze brain signals. This tool can extract all the sine signals with arbitrary frequencies. In MI, the alpha and beta bands are desired frequencies, also the frequency resolution can be adjustable. Because the proposed system uses this tool in a wide range of frequencies, we named the proposed tool "DCA-based adaptive filter bank".

## 3.5.1 Dataset used for evaluation

In this study, we used the IVa dataset of BCI Competition III. The dataset, which was recorded

from five healthy subjects using 118 channels, includes two classes of right hand and foot. 280 trials were recorded for each subject and the sampling rate was 1000 Hz. According to Table 3.1, in this dataset, most trials are labeled for 2 subjects, while less training data is provided for the other 3 subjects. The challenge is to properly categorize the two classes, even with a limited training dataset.

Table 3.1: Training and test trials

| Subject | Training trials | Test trials |
|---|---|---|
| aa | 168 | 112 |
| al | 224 | 56 |
| av | 84 | 196 |
| aw | 56 | 224 |
| ay | 28 | 252 |

## 3.5.2 Pre-processing

For the pre-processing section, we import the data related to each subject into EEGLAB. Then, using EEGLAB, we import event information and channel locations. Figure 3.6 represents the channel locations of the IVa dataset.

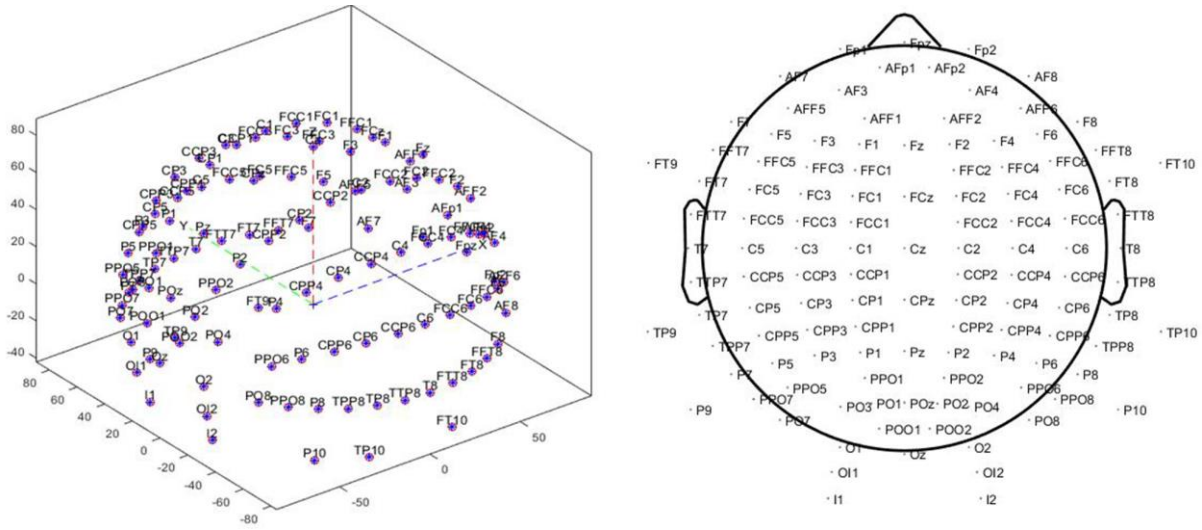

Figure 3.6: 2D and 3D plot of channel locations of the IVa dataset

To increase the signal to noise (SNR) of the channels, a spatial filter based on the Gaussian distribution with a standard deviation of 0.7 and a radius of 3 is apllied to signals, which is shown in Fig. 3.7.

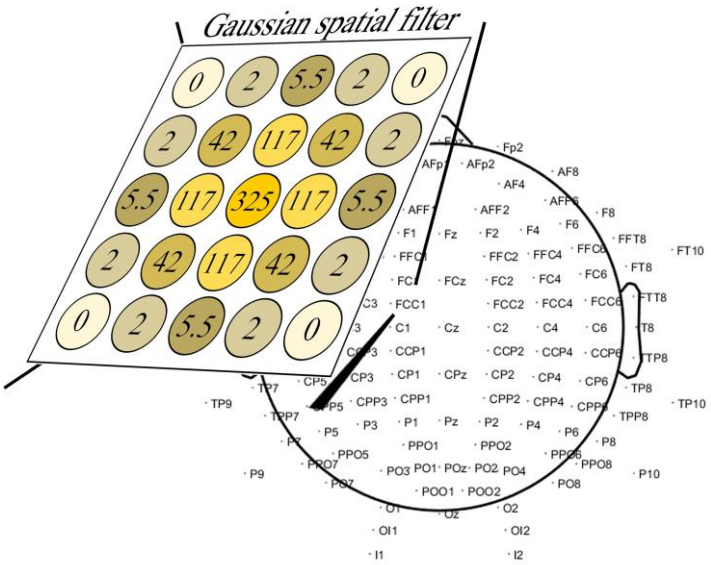


Figure 3.7: Gaussian spatial filter

To reduce the computational load, we down-sampled the sampling frequency to 100 Hz. Next, we apply a 33 order Butterworth filter with a bandwidth of 4 to 45 Hz to the data. The frequency response of the applied filter is displayed in Fig. 3.8. Thereafter, by examining the power spectrum density (PSD) of each channel, we identify and reject inappropriate channels. Next, we use the ASR algorithm to remove artifacts and clean the signals.

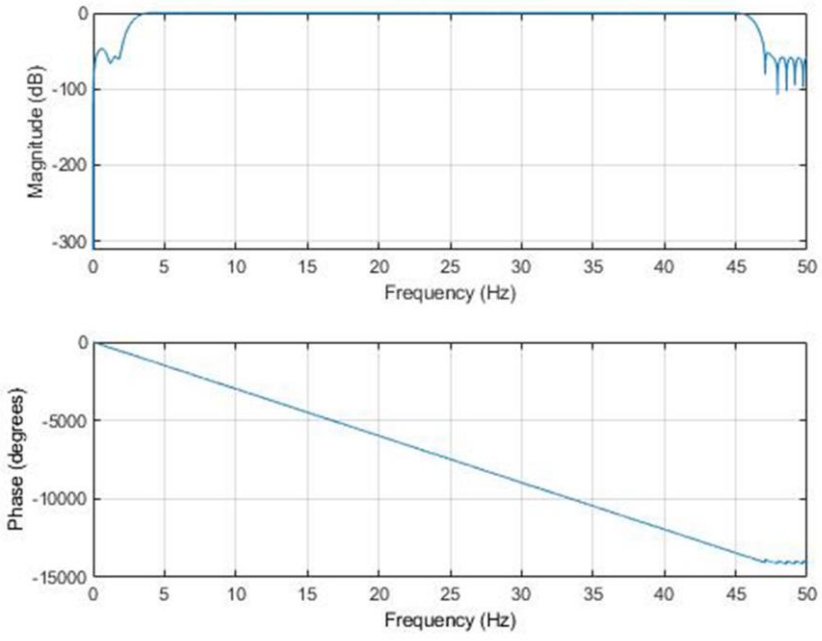


Figure 3.8: frequency response of applied filter

Afterward we run ICA and remove the inappropriate components. By inappropriate components we mean components which their properties did not match those of the brain components. The main criteria to specify if a component is artifacts or not is to check the scalp

map, the time period of the component and, the component activity power spectrum. Observations such as the smoothly decreasing EEG spectrum, a hotspot appeared in scalp map, high power at high frequencies and, absence from most time course or non-integrated presence indicate that the component may not be good and should be rejected. Figure 3.9 shows the maps of axcepted components for subject aa.

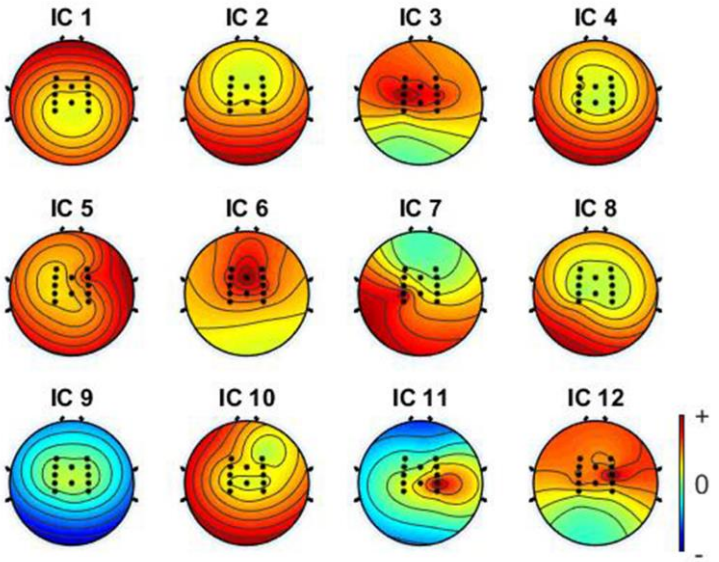


Figure 3.9: Component maps of subject aa

The steps performed for pre-processing section are summarized in Fig. 3.10.

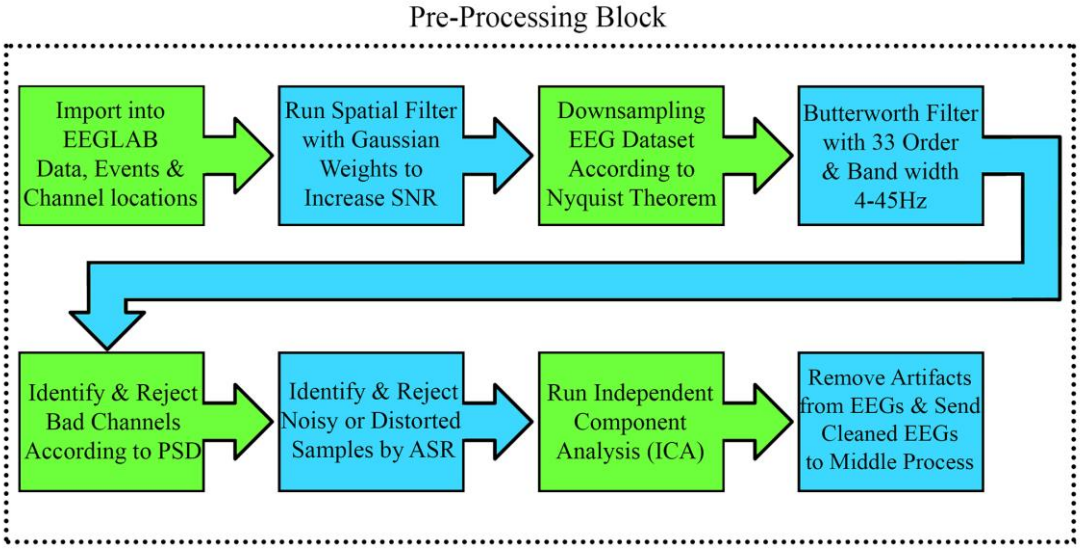


Figure 3.10: Pre-processing pipeline

### 3.5.3 Middle processing

This section is dedicated to expanding the dimensions of features through the DCA adaptive

filter bank and calculating feature parameters. DCA adaptive filter bank, which is one of our unique innovations, increases the feature dimensions to the required level of the proposed system. We used statistical parameters to analyze and categorize all sub-frequencies. Further details on the middle processing sections are given below:

### 3.5.3.1 Channel selection using CSP

CSP algorithm can differentiate the effective sources on active brain parts. We used CSP to evaluate the involvement of channels in the creation of sources. The average of the coefficients corresponding to each channel in the CSP mixing matrix is used for channel evaluation. If the average is less than the certain threshold value in the proposed system, then the channel is known as a useless channel. This way we understand which channels are more important than others. On average, for all subjects, the parietal and occipital areas were the beneficial range, so the channels located in front of the head were removed.

### 3.5.3.2 DCA adaptive filter bank

DCA adaptive filter bank, which is one of the proprietary designs of the proposed system, is able to search for different harmonics in the time string of signals and extract the correlated parts from the signal. Providing such a tool to the proposed system allows it to expand the feature dimensions along different sub-frequencies. The block diagram of this section is as shown in Fig. 3.11.


Figure 3.11: DCA adaptive filter bank block diagram in the middle processing section of the proposed system

In the figure above, the reference is a sinusoid with a specified frequency, and the signal in the proposed system is the input epochs. Figure 3.12 shows an example of the sub-frequency separation of an EEG epoch and conversion it to 22 useful features.

### 3.5.3.3 Feature extraction

After that, DCA adaptive filter bank outputs enter this stage for the computation of statistical indicators such as mean, median, variance, skewness, entropy, and power. The information is prepared in matrix blocks to be transferred to the final processing section.

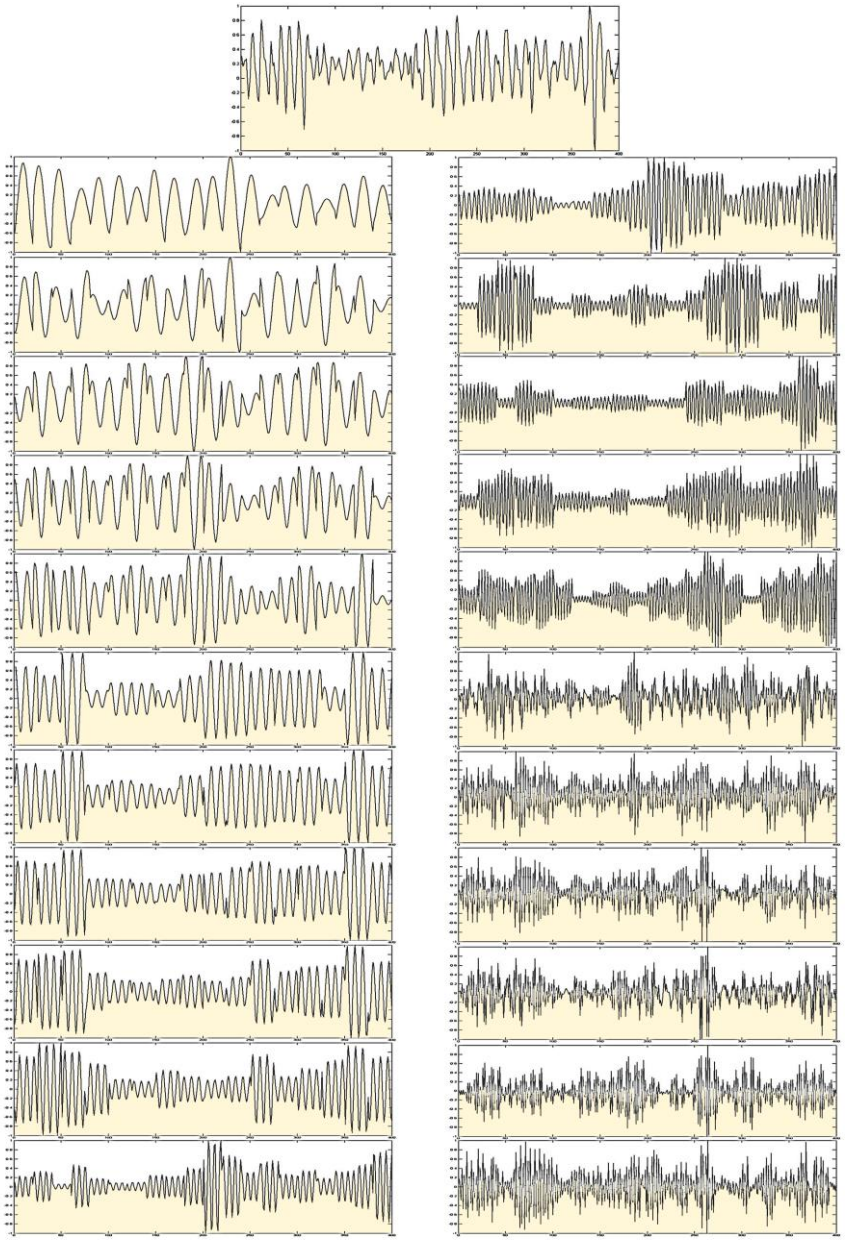

Figure 3.12: Sub-frequency separation of an EEG epoch

### 3.5.4 Final processing

The data labels attached to the data of the previous step are entered as matrix blocks in this section. The data is divided into three categories of train, test and, cross validation. A single layer perseptron neural network containing a neuron is used. The results of previous evaluations show that training by this neural network can bring system learning to an appropriate level.

## 3.6 Computer simulation and result

In this section, we execute the proposed system for two experiments with the large and small

train datasets in MATLAB 2018a. Also, the computer hardware has a core i7 processor and 8 GB of RAM.

### 3.6.1 Experiment 1

In the first experiment, we intend to enter one of the large datasets into the proposed system and measure its accuracy based on this dataset. According to Table 3.1, we selected subject al as the subject containing the large train dataset. After pre-processing according to Fig. 3.10, the proposed system proceeds to select the optimal channels and reduce the computational load. So, the output of pre-processing is entered into the CSP block. Figure 3.13 shows the channels obtained from the CSP evaluation of the channel, and these highlighted channels have been selected as useful channels.

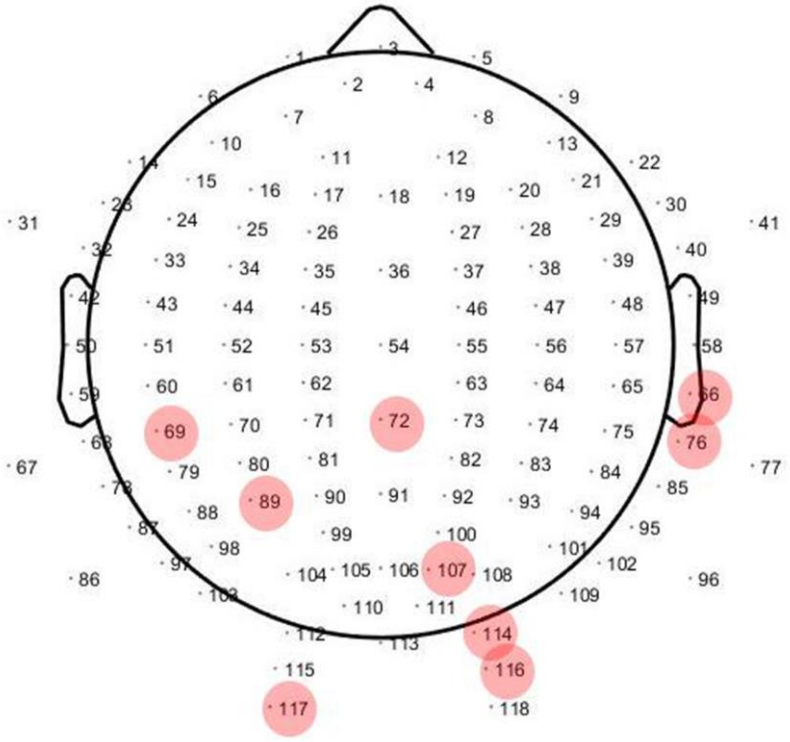


Figure 3.13: Selected channels by CSP for subject al

Also, the table below shows the number of errors in the testing on ten percent of the total subject al dataset.

Table 3.2: Error measurement in the four times training the proposed system for subject al

| Error | TP | FP | TN | FN |
|---|---|---|---|---|
| 0 | 11 | 0 | 11 | 0 |
| 0 | 11 | 0 | 11 | 0 |
| 0 | 11 | 0 | 11 | 0 |
| 1 | 10 | 0 | 11 | 1 |
| 0.25 (avg) | 10.75 | 0 | 11 | 0.25 |

Based on the DCA adaptive filter bank mechanism mentioned in sub-section 3.5.3.2, since

the pertinent information for MI is available at frequencies in the range of 8 to 30 Hz [43], we insert the sinuses from 8 to 30 Hz at a resolution of 0.5 Hz into the DCA adaptive filter bank. Then we calculate the statistical features mentioned in sub-section 3.5.3.3 for all sub-frequencies generated by the DCA adaptive filter bank. In the next level, since we had 23 sub-bands, 9 channels, and 6 features, so the neural network has 1242 inputs. The desired neural network is shown in Fig. 3.14.

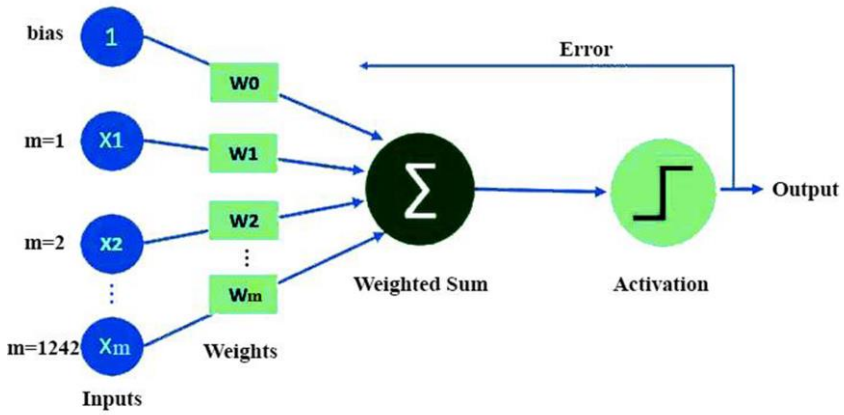


Figure 3.14: Neural network used in final processes in proposed system

## 3.6.2 Experiment 2

In this experiment, we intend to enter one of the small datasets into the proposed system. According to Table 3.1, we selected subject aw as a subject containing the small train dataset. It is noted that all of the processes are the same as in the previous experiment. Figure 3.15 shows the channels obtained from the CSP evaluation of the channel, and these highlighted channels have been selected as useful channels.

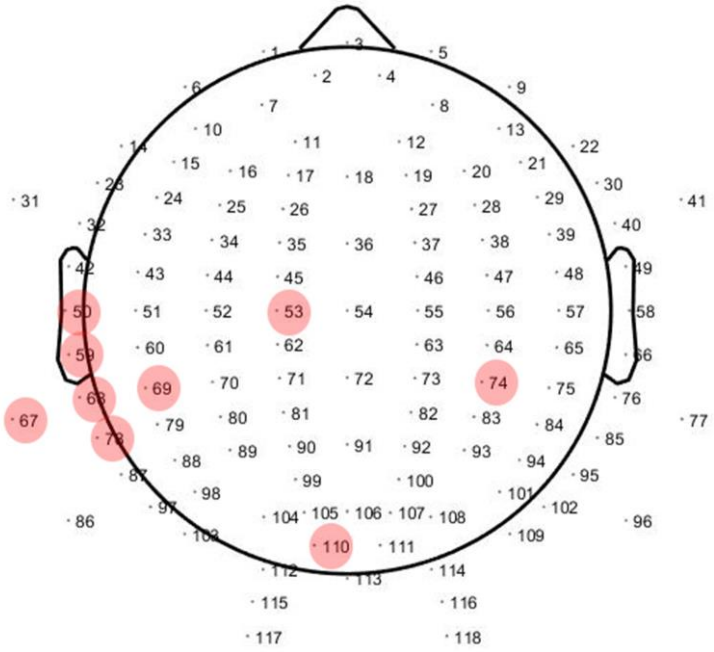

Figure 3.15: Selected channels by CSP for subject aw

Where TP is the number of cases correctly identified, FP is the number of cases incorrectly identified, TN is the number of cases correctly identified and, FN is the number of cases incorrectly identified. Likewise, precision is the TP to TP + FP ratio and recall is the TP to TP + FN ratio. Also, the table below shows the number of errors in the testing on ten percent of the total subject al dataset.

Table 3.3: Error measurement in the four times training the proposed system for subject aw

| Error | TP | FP | TN | FN |
|---|---|---|---|---|
| 0 | 3 | 0 | 3 | 0 |
| 0 | 3 | 0 | 3 | 0 |
| 0 | 3 | 0 | 3 | 0 |
| 0 | 3 | 0 | 3 | 0 |
| 0 (avg) | 3 | 0 | 3 | 0 |

## 3.7 Discussion

In this section, the results obtained from the proposed system for 5 subjects with five criteria, accuracy (ACC), sensitivity (SEN), specificity (SPE), Mathew's correlation coefficient (MCC), and F1 score, are measured. These criteria are defined as follows:

$$ACC = \frac{TP+TN}{TP+FP+TN+FN} \tag{3.9}$$

$$SEN = \frac{TP}{TP+FN} \tag{3.10}$$

$$SPE = \frac{TN}{TN+FP} \tag{3.11}$$

$$MCC = \frac{(TP \times TN)-(FP \times FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{3.12}$$

$$F1 = \frac{2 \times precision \times recall}{precision+recall} \tag{3.13}$$

Table 3.4 shows the obtained parameters for the proposed system.

Table 3.4: Obtained parameters for the proposed system

| Subject | TP | FP | TN | FN |
|---|---|---|---|---|
| aa | 9 | 1 | 7 | 0 |
| al | 11 | 0 | 11 | 0 |
| av | 4 | 0 | 4 | 0 |
| aw | 3 | 0 | 3 | 0 |
| ay | 1 | 0 | 1 | 0 |

Also, the results of the mentioned criteria are according to Table 3.5.

Table 3.5: The result of popular criteria

| Subject | ACC | SEN | SPE | MCC | F1 |
|---|---|---|---|---|---|
| aa | 94.11 | 100 | 87.5 | 88.74 | 0.94 |
| al | 100 | 100 | 100 | 100 | 1 |
| av | 100 | 100 | 100 | 100 | 1 |
| aw | 100 | 100 | 100 | 100 | 1 |
| ay | 100 | 100 | 100 | 100 | 1 |

## 3.8 Conclusion

Based on Equation 3.8, we show that during the consecutive difference of two observations, a number of IPs are generated at the local minimum and maximum position. At these points, the DCA technique is able to distinguish correlated samples from other samples and, by clustering the samples, obtain an estimate of the independent components. Using this technique and placing a known source instead of one of the two input observations, the system became a powerful adaptive filter. Then by placing a large number of sine signals as a known source, we built an adaptive filter bank and used it in the proposed system. According to the results of experiments 1 and 2, it was shown that the system contains a very small number of errors in various training. This confirms the robustness of the system against various initial conditions. The results of the system show the high efficiency of the proposed system and the improvement of the MI-based BCI systems by DCA technique and tools made based on it. Then, popular criteria were used to measure the performance of the system. According to the results, the average accuracy, sensitivity, specificity, MCC, and F1 score for all subjects were 98.8%, 100%, 97.5%, 97.75%, and 0.994 respectively. All these results show the importance of using the proposed method in BCI-based systems.

# References


[1] S. Chaudhary, S. Taran, V. Bajaj, S. Siuly, A flexible analytic wavelet transform based approach for motor-imagery tasks classification in BCI applications, Computer Methods and Programs in Biomedicine. 187 (2020). https://doi.org/10.1016/j.cmpb.2020.105325.

[2] Y. Zhang, S.Q. Xie, H. Wang, Z. Zhang, Data Analytics in Steady-State Visual Evoked Potential-Based Brain-Computer Interface: A Review, IEEE Sensors Journal. 21 (2021) 1124–1138. https://doi.org/10.1109/JSEN.2020.3017491.

[3] B.Z. Allison, A. Kübler, J. Jin, 30+ years of P300 brain–computer interfaces, Psychophysiology. 57 (2020). https://doi.org/10.1111/psyp.13569.

[4] E.M. dos Santos, R. Cassani, T.H. Falk, F.J. Fraga, Improved motor imagery brain-computer interface performance via adaptive modulation filtering and two-stage

classification, Biomedical Signal Processing and Control. 57 (2020). https://doi.org/10.1016/j.bspc.2019.101812.
[5] S. Taran, V. Bajaj, D. Sharma, S. Siuly, A. Sengur, Features based on analytic IMF for classifying motor imagery EEG signals in BCI applications, Measurement: Journal of the International Measurement Confederation. 116 (2018). https://doi.org/10.1016/j.measurement.2017.10.067.
[6] S. Zhang, S. Wang, D. Zheng, K. Zhu, M. Dai, A novel pattern with high-level commands for encoding motor imagery-based brain computer interface, Pattern Recognition Letters. 125 (2019). https://doi.org/10.1016/j.patrec.2019.03.017.
[7] M.A. Amirabadi, M.H. Kahaei, A New Fast Approach for an EEG-based Motor Imagery BCI Classification, IETE Journal of Research. (2020). https://doi.org/10.1080/03772063.2020.1816221.
[8] S. Taran, V. Bajaj, Motor imagery tasks-based EEG signals classification using tunable-Q wavelet transform, Neural Computing and Applications. 31 (2019). https://doi.org/10.1007/s00521-018-3531-0.
[9] H.-L. Halme, Development of an MEG-based brain-computer interface, n.d. www.aalto.fi.
[10] K.J. Miller, D. Hermes, N.P. Staff, The current state of electrocorticography-based brain-computer interfaces, Neurosurgical Focus. 49 (2020). https://doi.org/10.3171/2020.4.FOCUS20185.
[11] IEEE Engineering in Medicine and Biology Society. Annual International Conference (42nd : 2020 : Online), IEEE Engineering in Medicine and Biology Society, Institute of Electrical and Electronics Engineers, Canadian Medical and Biological Engineering Society. Annual Conference (43rd : 2020 : Online), 42nd Annual International Conferences of the IEEE Engineering in Medicine and Biology Society : "Enabling Innovative Technologies for Global Healthcare" : 20-24 July 2020, Montreal, Canada., n.d.
[12] Y. Zhu, Y. Li, J. Lu, P. Li, A Hybrid BCI Based on SSVEP and EOG for Robotic Arm Control, Frontiers in Neurorobotics. 14 (2020). https://doi.org/10.3389/fnbot.2020.583641.
[13] M. de Castro-Cros, M. Sebastian-Romagosa, J. Rodríguez-Serrano, E. Opisso, M. Ochoa, R. Ortner, C. Guger, D. Tost, Effects of Gamification in BCI Functional Rehabilitation, Frontiers in Neuroscience. 14 (2020). https://doi.org/10.3389/fnins.2020.00882.
[14] S. Bhattacharyya, S. Das, A. Das, R. Dey, R.S. Dhar, Neuro-feedback system for real-time BCI decision prediction, Microsystem Technologies. (2021). https://doi.org/10.1007/s00542-020-05146-4.
[15] M. Wu, H. Qi, Using passive BCI to online control the air conditioner for obtaining the individual specific thermal comfort, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2019. https://doi.org/10.1109/EMBC.2019.8856497.
[16] Z. He, Z. Li, F. Yang, L. Wang, J. Li, C. Zhou, J. Pan, Advances in multimodal emotion recognition based on brain–computer interfaces, Brain Sciences. 10 (2020). https://doi.org/10.3390/brainsci10100687.
[17] M. Eidel, A. Kübler, Wheelchair Control in a Virtual Environment by Healthy Participants Using a P300-BCI Based on Tactile Stimulation: Training Effects and Usability, Frontiers in Human Neuroscience. 14 (2020). https://doi.org/10.3389/fnhum.2020.00265.
[18] W. Kong, B. Jiang, Q. Fan, L. Zhu, X. Wei, Personal identification based on brain networks of EEG signals, International Journal of Applied Mathematics and Computer

Science. 28 (2018). https://doi.org/10.2478/amcs-2018-0057.
[19] H. Sawada, N. Ono, H. Kameoka, D. Kitamura, H. Saruwatari, A review of blind source separation methods: Two converging routes to ILRMA originating from ICA and NMF, APSIPA Transactions on Signal and Information Processing. 8 (2019). https://doi.org/10.1017/ATSIP.2019.5.
[20] A. Cimmino, A. Ciaramella, G. Dezio, P.J. Salma, Non-linear PCA Neural Network for EEG Noise Reduction in Brain-Computer Interface, in: Smart Innovation, Systems and Technologies, 2021. https://doi.org/10.1007/978-981-15-5093-5_36.
[21] Y.J. Song, F. Sepulveda, A Novel Technique for Selecting EMG-Contaminated EEG Channels in Self-Paced Brain-Computer Interface Task Onset, IEEE Transactions on Neural Systems and Rehabilitation Engineering. 26 (2018). https://doi.org/10.1109/TNSRE.2018.2847316.
[22] L. Zou, X. Chen, G. Dang, Y. Guo, Z.J. Wang, Removing Muscle Artifacts from EEG Data via Underdetermined Joint Blind Source Separation: A Simulation Study, IEEE Transactions on Circuits and Systems II: Express Briefs. 67 (2020) 187–191. https://doi.org/10.1109/TCSII.2019.2903648.
[23] T. Sgouros, N. Mitianoudis, A novel Directional Framework for Source Counting and Source Separation in Instantaneous Underdetermined Audio Mixtures, IEEE/ACM Transactions on Audio Speech and Language Processing. 28 (2020) 2025–2035. https://doi.org/10.1109/TASLP.2020.3003855.
[24] F. Ferracuti, V. Casadei, I. Marcantoni, S. Iarlori, L. Burattini, A. Monteriù, C. Porcaro, A functional source separation algorithm to enhance error-related potentials monitoring in noninvasive brain-computer interface, Computer Methods and Programs in Biomedicine. 191 (2020). https://doi.org/10.1016/j.cmpb.2020.105419.
[25] J. Khan, M.H. Bhatti, U.G. Khan, R. Iqbal, Multiclass EEG motor-imagery classification with sub-band common spatial patterns, Eurasip Journal on Wireless Communications and Networking. 2019 (2019). https://doi.org/10.1186/s13638-019-1497-y.
[26] S.Z. Zahid, M. Aqil, M. Tufail, M.S. Nazir, Online Classification of Multiple Motor Imagery Tasks Using Filter Bank Based Maximum-a-Posteriori Common Spatial Pattern Filters, IRBM. 41 (2020). https://doi.org/10.1016/j.irbm.2019.11.002.
[27] N. Ji, L. Ma, H. Dong, X. Zhang, EEG signals feature extraction based on DWT and EMD combined with approximate entropy, Brain Sciences. 9 (2019). https://doi.org/10.3390/brainsci9080201.
[28] M.K.M. Rahman, M.A.M. Joadder, A space-frequency localized approach of spatial filtering for motor imagery classification, Health Information Science and Systems. 8 (2020). https://doi.org/10.1007/s13755-020-00106-8.
[29] Y. Park, W. Chung, Optimal Channel Selection Using Correlation Coefficient for CSP Based EEG Classification, IEEE Access. 8 (2020). https://doi.org/10.1109/ACCESS.2020.3003056.
[30] Y. Guo, Y. Zhang, Z. Chen, Y. Liu, W. Chen, EEG classification by filter band component regularized common spatial pattern for motor imagery, Biomedical Signal Processing and Control. 59 (2020). https://doi.org/10.1016/j.bspc.2020.101917.
[31] A Computationally Efficient Multiclass Time-Frequency Common Spatial Pattern Analysis on EEG Motor Imagery, 2020. https://doi.org/10.0/Linux-x86_64.
[32] S. Razi, M.R. Karami Mollaei, J. Ghasemi, A novel method for classification of BCI multi-class motor imagery task based on Dempster–Shafer theory, Information Sciences. 484 (2019) 14–26. https://doi.org/10.1016/j.ins.2019.01.053.
[33] M. Zheng, B. Yang, Y. Xie, EEG classification across sessions and across subjects through transfer learning in motor imagery-based brain-machine interface system,

Medical and Biological Engineering and Computing. 58 (2020) 1515–1528. https://doi.org/10.1007/s11517-020-02176-y.
[34] U. Talukdar, S.M. Hazarika, J.Q. Gan, Adaptation of Common Spatial Patterns based on mental fatigue for motor-imagery BCI, Biomedical Signal Processing and Control. 58 (2020). https://doi.org/10.1016/j.bspc.2019.101829.
[35] A. Miladinović, M. Ajčević, J. Jarmolowska, U. Marusic, M. Colussi, G. Silveri, P.P. Battaglini, A. Accardo, Effect of power feature covariance shift on BCI spatial-filtering techniques: A comparative study, Computer Methods and Programs in Biomedicine. 198 (2021). https://doi.org/10.1016/j.cmpb.2020.105808.
[36] A. Miladinovic, M. Ajcevic, J. Jarmolowska, U. Marusic, G. Silveri, P.P. Battaglini, A. Accardo, Performance of EEG Motor-Imagery based spatial filtering methods: A BCI study on Stroke patients, in: Procedia Computer Science, 2020. https://doi.org/10.1016/j.procs.2020.09.270.
[37] I. Daly, Neural component analysis: A spatial filter for electroencephalogram analysis, Journal of Neuroscience Methods. 348 (2021). https://doi.org/10.1016/j.jneumeth.2020.108987.
[38] R.J. Bufacchi, C. Magri, G. Novembre, G.D. Iannetti, Local spatial analysis: An easy-to-use adaptive spatial EEG filter, Journal of Neurophysiology. 125 (2021). https://doi.org/10.1152/jn.00560.2019.
[39] V. K, D. A, M. J, S. M, A. A, S.A. Iraj, A novel method of motor imagery classification using eeg signal, Artificial Intelligence in Medicine. 103 (2020) 101787. https://doi.org/10.1016/j.artmed.2019.101787.
[40] A. Al-Saegh, S.A. Dawwd, J.M. Abdul-Jabbar, Deep learning for motor imagery EEG-based classification: A review, Biomedical Signal Processing and Control. 63 (2021). https://doi.org/10.1016/j.bspc.2020.102172.
[41] S. Chaudhary, S. Taran, V. Bajaj, A. Sengur, Convolutional Neural Network Based Approach Towards Motor Imagery Tasks EEG Signals Classification, IEEE Sensors Journal. 19 (2019). https://doi.org/10.1109/JSEN.2019.2899645.
[42] S.K. Khare, V. Bajaj, A facile and flexible motor imagery classification using electroencephalogram signals, Computer Methods and Programs in Biomedicine. 197 (2020). https://doi.org/10.1016/j.cmpb.2020.105722.
[43] M.A.M. Joadder, S. Siuly, E. Kabir, H. Wang, Y. Zhang, A New Design of Mental State Classification for Subject Independent BCI Systems, IRBM. 40 (2019) 297–305. https://doi.org/10.1016/j.irbm.2019.05.004.