# Designing a brain-computer interface with the aim of classifying features and enhancing the signal-to-noise ratio

Mojgan Katanforoush<sup>1</sup>, PhD Student, Reza Pourgholi<sup>1,\*</sup>, Professor

<sup>1</sup> School of Mathematics and Computer Science, Damghan University, P.O.Box 36715-364

*Abstract*: A brain-computer interface is a hardware and software communication system through which the user will be able to control computers and external devices using only their brain activities. The signal processing algorithm is the most important part of a brain-computer interface and includes the steps of data acquisition, preprocessing or signal amplification, feature extraction, and classification. This research aims to design the signal processing algorithm of a brain-computer interface and also to improve its performance using noise reduction methods. Considering the importance of feature extraction and classification steps, we must choose appropriate methods in these steps. First, the brain-computer system, signal processing algorithm, and human nervous system and brain, electroencephalogram signal have been investigated. Then, the pre-processing step and noise reduction techniques, the feature extraction step, and the classification step, and different classifiers with their applications and characteristics have been introduced. Finally, a new method based on channel selection using the placement of electrodes has been presented, which reduces noise and significantly increases the performance of the algorithm, and the use of this method increases the accuracy of the system.

Keywords: Brain-computer interface, electroencephalogram, preprocessing, classification, noise reduction

<sup>&</sup>lt;sup>1</sup> Email: katanforoush@du.ac.ir

<sup>\*</sup> Corresponding author : pourgholi@du.ac.ir

#### 1- Introduction

Brain-Computer Interfaces (BCIs) are communication devices that allow for conveying intentions solely through brain activity without involving muscles. For brain-computer communication, electrodes or sensors are placed on the patient's head and skull. These brain electrode signals are fed to the signal amplifiers of the Data Acquisition (DAQ) board, which includes an analogto-digital converter.

Once converted from analog to digital, the signals are sent to the brain-computer information processing algorithm to perform feature extraction and classification processes.

Depending on the speed and efficiency of the chosen algorithm, a specific movement can be conveyed to the user as feedback. This feedback represents the output of the BCI system and allows the user to interact with the BCI more effectively and rapidly [9].

#### **Test specifications**

The dataset used in this research was recorded by the Berlin Brain-Computer Interface Group at the Berlin Institute of Technology and the Berlin University of Medical Sciences. It corresponds to the first dataset of the 4th BCI Competition. This dataset was recorded from seven healthy individuals. During the recording sessions, participants performed motor imagery tasks without feedback. For each individual, two motor imagery tasks were selected from three classes: left hand, right hand, and foot. During the experiment, visual cues indicating left, right, or downward directions appeared on the screen, signaling the participant to imagine the respective movement. Each cue was shown for four seconds to allow the subject to perform the mental imagery. Signals were recorded using 59 electrodes, with the highest density placed over the brain's motor imagery areas.

(Refer to Fig.1 for electrode positioning on the head and corresponding channel numbers for this dataset.)



Fig. 1. Placement of electrodes on the head

## 2- Preprocessing step in signal processing algorithm

#### 2-1 Common Average Reference (CAR)

This method calculates the signal of each electrode by subtracting the average of all electrodes from the signal of the desired electrode:

$$y_i = x_i - \frac{1}{N}(x_1 + x_2 + \dots + x_N)$$
 (1)

Where i=1,2,3,..., N.

The Common Average Reference (CAR) technique is effective in reducing noise, such as 50Hz or 60Hz power line interference that is common across all electrodes. Because useful brain signals are typically concentrated in a limited number of electrodes, this method improves the signal-to-noise ratio by enhancing the brain signal relative to the average of all channels. However, CAR cannot eliminate artifacts that are not common across electrodes, such as:

Ocular artifacts (EOG), Muscle artifacts (EMG)

Eye movement signals are generally stronger near the frontal cortex, while muscle signals are more pronounced in regions close to active muscles. To address such artifacts, methods like regression or **Independent Component Analysis (ICA)** are more suitable. [4].

#### **2-2 Laplace reference**

Laplace referencing considers the spatial distribution of nearby electrodes. It subtracts a weighted sum of the potentials of neighboring electrodes from the potential of the desired electrode. The weights are determined by the inverse of the distance between electrodes:

$$y_i = x_i - \sum_{j \in s_i}^n x_i g_{ij}$$
(2)  
$$g_{ij} = \frac{1}{d_{ij}} / \sum_{j \in s_i}^n \frac{1}{d_{ij}}$$
(3)

Where:

- *s<sub>i</sub>* : set of neighboring electrodes of the i-th electrode
- $d_{ij}$ : distance between electrodes i and j

This method enhances the local signal by emphasizing activity at a specific electrode relative to its neighbors, which can improve spatial resolution.

#### 2-3 Independent Component Analysis (ICA)

ICA assumes that EEG signals are a mixture of independent sources, some of which correspond to useful brain activity and others to noise. Let:

$$x(t) = f(s(t)) + n(t)$$
(4)

- x(t): observed EEG signal
- *s*(*t*): source signals (independent components)
- *f*: unknown mixing function
- n(t): additive noise

The goal of ICA is to recover the sources s(t) from the observed mixtures x(t), assuming statistical independence of sources.

This is particularly useful for removing EOG, EMG, and other artifacts from EEG data [3].

#### **3-** Classification Step in BCI Signal Processing Algorithm

#### 3-1 k-nearest neighbors classifier (k-NNC)

This classifier is a simple yet effective method. It classifies a new sample by comparing its features with those of known training samples and assigning it to the class of the majority among its  $\mathbf{k}$  nearest neighbors.

• Distance metrics such as Euclidean distance are typically used:

$$d(x,x_i) = \sqrt{\sum (x-x_i)^2}$$
(5)

This method is intuitive and works well with small datasets, but can be computationally intensive with large datasets and sensitive to irrelevant features.

#### 3-2 Linear Discriminant Analysis (LDA)

LDA is widely used in BCI due to its low computational cost and real-time capability. It works by projecting the data onto a line (or hyperplane for multiclass problems) in a way that:

- Maximizes between-class variance
- Minimizes within-class variance

LDA finds the direction that best separates the two classes:

$$W = S_w^{-1}(m_1 - m_2)$$
 (6)

- S<sub>w</sub>: within-class scatter matrix
- m<sub>1</sub>,m<sub>2</sub>: mean vectors of each class

Although LDA performs well with linearly separable data, it may not be suitable for non-linear problems.

#### 3-3 Support vector machine (SVM)

SVM is a powerful classifier that works by finding the optimal hyperplane that maximally separates different classes [16].

• It uses kernel functions to map data to a higher-dimensional space when the data is not linearly separable.

• It maximizes the **margin** between classes, leading to better generalization.

The decision function in SVM is:

$$f(x) = \text{sign } \sum_{i=1}^{n} (\alpha_i y_i K(x_i, x) + b)$$
 (7)

 $\alpha_i$ : Lagrange multipliers  $y_i$ : class labels  $K(\boldsymbol{x}_i, \boldsymbol{x})$ : kernel function b: bias term

SVM is particularly robust to noise and overfitting, especially when combined with regularization techniques.

#### 4- Feature Extraction Step in BCI Signal Processing Algorithm

#### **Common Spatial Pattern (CSP):**

CSP is one of the most effective methods used for feature extraction in EEG-based BCI systems. It transforms multi-channel EEG data into a spatial domain where the variance of one class is maximized while the variance of the other class is minimized. This makes it easier for classifiers to distinguish between different mental states, especially in **binary classification** tasks like motor imagery (e.g., left hand vs. right hand).

#### Mathematical Basis:

Let E be the EEG signal matrix of size N×T, where:

- N: number of channels (electrodes)
- T: number of time samples

The **normalized spatial covariance matrix** C is calculated as:

$$C = \frac{EE^T}{trace(EE^T)} \tag{8}$$

Where:

 $E^T$ : transpose of the EEG matrix

*trace*(*X*): sum of the diagonal elements of matrix X.

CSP decomposes the composite covariance matrix into eigenvalues and eigenvectors. The **spatial filters** are derived from these eigenvectors.

Let:

 $C_1$ ,  $C_2$  = covariance matrices for each class

The composite spatial covariance matrix is:

$$C_c = C_1 + C_2 \tag{9}$$

Then CSP solves the **generalized eigenvalue problem:** 

$$C_1 w = \lambda C_c w \tag{10}$$

Where:

w: spatial filter,  $\lambda$ : eigenvalue

#### Feature Extraction:

After applying the spatial filters to the EEG data, the log-variance of the projected signals is computed and used as features:

$$\mathbf{f}_{i} = \log(\frac{var(w_{i}^{T}E)}{\Sigma var(w^{T}E)}) \quad (11)$$

Typically, the filters corresponding to the **largest and smallest eigenvalues** are selected, as they offer the best class discrimination. CSP is known to be **sensitive to noise and electrode selection**, which is why it is often paired with **preprocessing** and **electrode selection strategies** [18].

#### 5- Implementation of the Brain-Computer Interface Algorithm

In this section, the signal processing algorithm of the Brain-Computer Interface is implemented using a combination of preprocessing techniques, feature extraction (CSP), and classification algorithms (KNN, LDA, and SVM) [8], [10]. All signal data were filtered within the frequency band of 8 to 25 Hz, and features were extracted using Common Spatial Pattern (CSP) [7], [17].

Classification was done using 10-fold cross-validation, and the performance was evaluated based

on classification accuracy and execution time for each classifier and preprocessing configuration.

### 5-1 Classification with CSP Only (No Preprocessing)

	KNN	LDA	SVM
Person A	59.5%	50%	54%
Person B	72.5%	72%	72.5%
Person C	72.5%	71.5%	71.5%
Person D	62.5%	67%	65%
Person E	91%	93%	93%
Person F	57%	61.5%	61.5%
Person G	72.5%	70.5%	72.5%
AVERAGE	69.6429%	69.3571%	70%
Execution time (seconds)	22.8966	23.1730	22.1140

Table 1. Accuracy of Classifiers Using Only CSP



Fig. 2. Comparison of the Accuracy of Classifiers

#### 5.2 Classification with CSP + CAR

The Common Average Reference (CAR) method was used for artifact reduction before applying CSP.

Although CAR slightly changed the spatial characteristics of the signals, it did not significantly improve the classification accuracy.

- Slight improvement in subject A, D.
- Accuracy differences were not statistically significant.
- Execution time increased to around 25–30 seconds for each classifier

Table 2. Implementation with Common	Spatial Pattern and
Common Average Reference	

	KNN	LDA	SVM
Person A	59%	50.5%	53%
Person B	70%	72.5%	72.5%
Person C	71.5%	71.5%	70.5%
Person D	66%	67%	67.5%
Person E	91%	92.5%	93%
Person F	59.5%	60.5%	61.5%
Person G	71.5%	72%	73.5%
AVERAGE	69.7857%	69.2143%	70.2143%
Execution time (seconds)	25.3959	25.7285	25.3441

### **5.3** Classification with CSP + Laplace Reference

The Laplace reference method, which uses weighted averages of neighboring electrodes based on distance, was applied before CSP. While it enhanced the local spatial resolution, its impact on classification accuracy was again limited.

- No significant gains in accuracy across subjects.
- Execution time increased considerably to about 36 seconds per classifier due to the computational cost of Laplace filtering.



Table 3. Implementation with common spatial pattern and Laplace reference

#### 5.4 Classification with CSP + ICA

**Independent Component Analysis (ICA)** was applied before CSP to remove artifacts like eye and muscle movement signals.

Table 4. Implementation with Common Spatial Pattern and Analysis of Independent Components

	KNN	LDA	SVM
Person A	54.5%	54%	57.5%
Person B	73%	73%	73%
Person C	69.5%	70%	68.5%
Person D	72%	77.5%	75.5%
Person E	92.5%	93%	92%
Person F	56.5%	62%	62%
Person G	75.5%	75%	76%
AVERAGE	70.5%	72.0714%	72.0714%
Execution time (seconds)	103.5669	105.7759	105.7309



Fig. 3. Comparison of Accuracy of Classifiers Chart in Algorithm (5-4)

Despite the high computational cost, ICA yielded the best classification accuracy, especially when combined with SVM.

#### 6. Electrode Selection Algorithm

According to Fig. 4, the parts of the brain that are activated by imagining the movement of the left and right hands are identified, and in places where the intensity of the color is higher, we see more brain activity. The motor perception cortex is located in the central part and towards the ears. In the desired data set, we must choose the electrodes that are located in the central part and the motor perception cortex.



Fig. 4. The Activated Part of the Brain in Imagining Left and Right Hand Movements

To improve classification performance and reduce computation time, a two-stage electrode selection method was employed.

#### 6.1 Selected Channels (Stage 1)

The initial selection of 25 electrodes was based on their anatomical relevance to motor imagery. These include:



Fig. 5. Placement of Electrodes in the First Step

Table 5. The Accuracy of the Cl	lassifiers in the First Stage	of the
Channel Selection Algorithm		

	KNN	LDA	SVM
Person A	60.5%	55.5%	58.5%
Person B	76%	77%	77%
Person C	69%	75%	75%
Person D	50%	52.5%	51.5%
Person E	83%	86.5%	78.5%
Person F	77%	77.5%	81%
Person G	76%	76%	77.5%
AVERAGE	70.2143%	71.4286%	72.5714%
Execution time (seconds)	11.6561	11.6843	11.6369

#### 6.2 Classification Results (19 Channels)

After performance-based refinement, the number of channels was reduced to 19.



The use of fewer electrodes led to significant improvements in execution time compared to using all channels with ICA.

Table 6. The Accuracy of the Classifiers in the Second Stage of the Channel Selection Algorithm

	KNN	LDA	SVM
Person A	82.5%	66.5%	82%
Person B	70%	76%	76.5%
Person C	58.5%	63.5%	63.5%
Person D	58%	52.5%	52.5%
Person E	79.5%	79.5%	78.5%
Person F	78%	82%	81.5%
Person G	90%	87.5%	90%
AVERAGE	73.7857%	72.5%	74.7857%
Execution time (seconds)	11.2516	9.8891	9.5692

The two-stage channel selection process significantly improved both accuracy and processing speed.

The best performance was achieved using SVM with 19 selected electrodes, confirming that smart feature reduction is essential in practical BCI systems.

### 7. Electrode Selection Based on Histogram Analysis

#### 7.1 Histogram Comparison of Channels

In this section, a more advanced method for electrode selection is presented, which uses not only the spatial location of electrodes but also a statistical analysis of their activity based on **histogram charts**.

The goal is to identify the most active and informative channels by examining the distribution and strength of signals recorded during motor imagery.

For better interpretation and comparison of channel activity, histogram diagrams were generated for each selected channel.

The distribution of signal power and variance in each channel was analyzed to determine which ones contributed the most meaningful data to the classification process. As a case study, the following channels were compared:

Channel 1 (from the initial selection)

Channel 38 (not originally included in the selected channels)



Fig. 7. Histogram Chart of Trial 5, Channel 1





Channel 1 had low activity and minimal variance, indicating weak signal strength and poor discriminative potential.

Channel 38, in contrast, showed significant activity and high variance in response to motor imagery tasks.

This finding suggests that electrode selection should not be based solely on anatomical location, but also take into account actual signal quality and statistical contribution.

#### Updated Electrode Set Based on Histogram Analysis

Using the histogram results, the selected electrode set was updated to include channels with higher activity and to eliminate weak or redundant ones.

This updated set was then tested again with all three classifiers: KNN, LDA, and SVM.

### Classification Results with Histogram-Based Selection

The classification results improved across all three algorithms after the histogram-informed channel refinement.

The SVM classifier once again achieved the highest accuracy, improving from 74.8% to 79.2%, demonstrating the effectiveness of this enhanced selection method.



Fig. 9. Placement of Electrodes in Channel Selection Algorithm with Histogram Diagram

	KNN	LDA	SVM
Person A	78.5%	63.5%	80%
Person B	73.5%	71.5%	73.5%
Person C	70.5%	72%	72.5%
Person D	68%	70.5%	72%
Person E	84.5%	85.5%	85.5%
Person F	77.5%	83%	82%
Person G	88%	85%	89%
AVERAGE	77.2143%	75.8571%	79.2143%
Execution time (seconds)	6.5527	7.9322	7.4710

Table 7. Accuracy of classifiers in channel selection with Histogram Diagram

Despite the improved performance, the reduced number of optimized channels maintained a fast classification speed.

Execution time for all classifiers remained under 7.5 seconds, making this approach suitable for real-time BCI applications.



Fig. 10. Comparison of the Accuracy of Classifiers in the Channel Selection Algorithm with Histogram Chart

The histogram-based channel selection method, by incorporating statistical signal strength in addition to spatial electrode position, resulted in better classification accuracy without increasing computational time.

This demonstrates the power of data-driven channel selection in improving the performance of EEG-based Brain-Computer Interface systems.

#### 7.2 Total Data Acquisition Time

In order to better compare the introduced methods, first, the data related to the total data collection time was classified, with the difference that all the data was used, and the data related to the time when the screen was blank was not removed, nor was the data filtered.

	KNN	LDA	SVM
Person A	44.5%	50.5%	45%
Person B	52.5%	58.5%	57%
Person C	50%	46%	44%
Person D	50.5%	52%	51%
Person E	50 %	48.5%	48.5%
Person F	48%	49%	53.5%
Person G	52.5%	47.5%	45.5%
AVERAGE	49.7143%	50.2875%	49.2143%
Execution time (seconds)	31.8999	32.4852	30.4857

Table 8. Accuracy of the Classifications Related to the Total Data Collection Time

As can be seen in the above table, the average accuracy of all the classifiers and also the accuracy of the classifiers for all persons is around 50%. Due to the two classes of the data set used in this research, the probability of occurrence of each class is 50% or equal to 0.5, so if you do not remove additional data and also avoid filtering them, an accuracy of about 50% will be achieved, which In this way, it can be said that the classifiers did the classification unconsciously and by chance.



### Fig. 11. Comparing the accuracy of the classifiers in the final algorithms

#### 8- Conclusion

This study presented a comprehensive approach to designing and optimizing a Brain-Computer Interface (BCI) system by focusing on signal processing methods, particularly preprocessing, feature extraction, and classification. Key contributions of this research include:

- Evaluation of various preprocessing techniques such as CAR, Laplace, and ICA, with results showing that ICA effectively reduced noise and improved classification performance.
- Use of the Common Spatial Pattern (CSP) method for feature extraction, which enhanced class separation and feature discrimination.
- Comparison of three classifiers (KNN, LDA, SVM), with SVM consistently achieving the highest accuracy across all test scenarios.
- Development of a two-stage electrode selection strategy, starting with 25 anatomically relevant channels and refining down to 19 based on performance.
- Introduction of a histogram-based channel selection method, which further improved classification accuracy and demonstrated the importance of statistical signal quality in channel selection.

The highest classification accuracy (79.2%) was achieved using SVM combined with histogram-based channel selection, and execution time was reduced to under 7.5 seconds, confirming the method's suitability for real-time BCI applications.

In conclusion, combining CSP with proper preprocessing and intelligent channel selection significantly boosts BCI performance. This framework offers a scalable and efficient path toward accurate, low-latency brain-computer building interfaces. The excessive number of them causes redundancy and increases the execution time. It is clear that the channel selection algorithm has the best accuracy, and also the execution time of this algorithm is much less.

#### 9- Reference

[1] C. Allen and D. M. A. Mehler, "Open science challenges, benefits and tips in early career and beyond," PLoS Biol., vol. 17, no. 5, p. e3000246, 2019, doi: 10.1371/journal.pbio.3000246.

[2] M. Allen, D. Poggiali, K. Whitaker, T. R. Marshall, J. Van Langen, and R. A. Kievit, "Raincloud plots: a multi-platform tool for robust data visualization," Wellcome Open Res., vol. 4, p. 63, 2021, doi: 10.12688/wellcomeopenres 15191.2.

[3] M. Baig, N. Aslam, and H. Shum, "Filtering techniques for channel selection in motor imagery EEG applications: a survey," Artif. Intell. Rev., vol. 53, pp. 1207–1232, 2020, doi: https://doi.org/10.1007/s10462-019-09694-8

[4] M. J. Brookes et al., "Theoretical advantages of a triaxial optically pumped magnetometer magnetoencephalography system," NeuroImage, vol. 236, p. 118025, 2021, doi: 10.1016/j.neuroimage.2021.118025.

[5] B. Graimann, B. Z. Allison, and G. Pfurtscheller, Brain-computer interfaces: Revolutionizing humancomputer interaction. New York, NY, USA: Springer, 2010.

[6] C. Guger, G. Edlinger, and G. Krausz, "Hardware/software components and applications of BCIs," in Recent Advances in Brain-Computer Interface Systems, 2010, pp. 1–24.

[7] E. Hortal et al., "SVM-based Brain–Machine Interface for controlling a robot arm through four mental tasks," Neurocomputing, vol. 151, pp. 116– 121, 2015, doi: https://doi.org/10.1016/j.neucom.2014.09.078

[8] M. A. M. Joadder, S. Siuly, and E. Kabir, "A new way of channel selection in the motor imagery classification for BCI applications," in Proc. Int. Conf. Health Inf. Sci., 2018, pp. 1–10, doi: https://doi.org/10.1007/978-3-030-01078-2\_10

[9] L. Kauhanen et al., "EEG and MEG braincomputer interface for tetraplegic patients," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 14, no. 2, pp. 190–193, 2006, doi: 10.1109/TNSRE.2006.875546 [10] T. Kayikcioglu and O. Aydemir, "A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data," Pattern Recognit. Lett., vol. 31, no. 11, pp. 1207–1215, 2010, doi: https://doi.org/10.1016/j.patrec.2010.04.009

[11] F. Lotte, "A Tutorial on EEG Signal Processing Techniques for Mental State Recognition in Brain-Computer Interfaces," in Guide to Brain-Computer Music Interfacing, 2014, pp. 32–45, doi: https://doi.org/10.1007/978-1-4471-6584-2\_7

[12] E. A. Mousavi et al., "Wavelet common spatial pattern in asynchronous offline brain computer interfaces," Biomed. Signal Process. Control, vol. 6, no. 2, pp. 121–128, 2011, doi: https://doi.org/10.1016/j.bspc.2010.08.003s

[13] A. Nakamura et al., "Electromagnetic signatures of the preclinical and prodromal stages of Alzheimer's disease," Brain J. Neuron, vol. 141, no. 5, pp. 1470–1485, 2018, doi: https://doi.org/10.1093/brain/awy044

[14] S. H. F. Syam et al., "Comparing common average referencing to Laplacian referencing in detecting imagination and intention of movement for brain computer interface," in MATEC Web Conf., vol. 140, 2017, doi: https://doi.org/10.1051/matecconf/201714001028

[15] S. Vaid, P. Singh, and C. Kaur, "EEG Signal Analysis for BCI Interface: A Review," in Proc. 5th Int. Conf. Adv. Comput. Commun. Technol., 2015, doi: 10.1109/ACCT.2015.72

[16] Z. S. Mehrjard, H. Momeni, and H. A. Ardekani, "A review of machine learning algorithms to diagnose autism using the EEG signal," Soft Comput. J., 2023, doi: 10.22052/scj.2023.248522.1110.

[17] Z. Sharifi, H. Momeni, and H. A. Ardakani, "Autism Diagnosis from EEG Signals Using Machine Learning Algorithms and Convolutional Neural Networks," Soft Comput. J., 2024, doi: 10.22052/scj.2024.253566.1185.

[18] H. Veisi, H. R. Ghaedsharaf, and M. Ebrahimi, "Improving the Performance of Machine Learning Algorithms for Heart Disease Diagnosis by Optimizing Data and Features," Soft Comput. J., vol. 8, no. 1, pp. 6–12, 2019, doi: 20.1001.1.23223707.1398.8.1.6.2.