Impact of the preprocessing block on the performance of the BCI system

Soukaina Hamou[1,*], *Hafida Idrissi Azami*[1], *Said Agounad*[1], *Ousama Tarahi*[1], *and Mustapha Moufassih*[1]
[1] Laboratory of Metrology and Information Processing, Faculty of Sciences, Ibn Zohr University, Agadir , Morocco

Abstract. Electroencephalography (EEG) is considered as one of the famous and efficient used methods in the Brain Computer Interface (BCI). This is due to its simplicity for implementation, low cost and being portable. The EEG is a technique that examines the electrical activity of the brain using a non-invasive electrodes placed on the scalp. EEG-based BCI system is constituted of five blocks: signal acquisition, preprocessing, feature extraction, classification and command block. In this paper, we will study the impact of the filter type and its order on the performance of the considered BCI system. This system is composed of: bandpass (BP) filter for the preprocessing step, Common Spatial Pattern (CSP) in the feature extraction block, and for the classification block, we used Support Vector Machine (SVM). The obtained results show a good improvement of the proposed BCI system. Indeed, the accuracy of this system can achieve 88.17% and the kappa coefficient is almost 0.76.

Key words. ElectroEncephaloGram (EEG), Motor Imagery (MI), Brain Computer Interface (BCI), Band-Pass (BP) filter, Common Spatial Pattern (CSP), Support Vector Machine (SVM).

1. Introduction

The brain activities can be recorded using several invasive and non invasive methods such as ElectroEncephaloGraphy (EEG), functional Magnetic Resonance Imaging (fMRI), functional Near Infrared Spectroscopes (fNIRS), ElectroCorticoGraphy (ECoG), Magneto-EncephaloGraphy (MEG), etc. [1]. In this study, we put our interest on a non invasive EEG for many reasons, e.g., its low cost, simple to implement, portable and its high temporal resolution. EEG has many applications such as: epileptic seizure detection [2], detection of drug effects on brain activity, automatic sleep stage classification, prediction of depth of sedation and anaesthesia [3] and BCI system [4-6].
The principal objective of BCI is to record the brain activity and translate it into commands. It composed of five major blocks: signal acquisition, signal preprocessing, extraction of features, classification and command block.
In paper [7], the authors proposed a new *Wavelet-CSP* with *ICA-filter* approach that was trained and validated using SVM on BCI Competition IV datasets (BCI-IV-2a). The experimental results showed an important value of the kappa coefficient which is 0.68. Pramudita et al. [8] suggested a new method, called the *Firefly-Support vector machine (FASV*M) to enhance the classification accuracy of motor imagery. The method was applied to the features extracted by the CSP technique, and the system achieved good results with an average accuracy of 93.20%.

In this article, we suggest to study the influence of the preprocessing block on the performance of the used BCI system. This BCI system based on the triplet: BP filter for preprocessing, the CSP technique for feature extraction and the SVM technique for classification. The BCI-IV-2a dataset [9], is used to measure the influence of the chosen filter and its order by comparing the classification results of each case. Thus, we will consider only two-class problem (distinction between the imagination of two hands movement, right and left).
The following is a description of the paper's structure: Section 2 describes the different blocks of the BCI system with an overview about the used methods in our system. The methodology of our algorithm is provided in section 3. Section 4 describes the results and discussion. The conclusion is drawn in section 5.

2. BCI system

A general block diagram representing the first four blocks of the BCI system is shown in Fig. 1. These blocks will be described in the following subsections.

* Corresponding author: soukaina.hamou@edu.uiz.ac.ma

https://doi.org/10.1051/e3sconf/202235101016

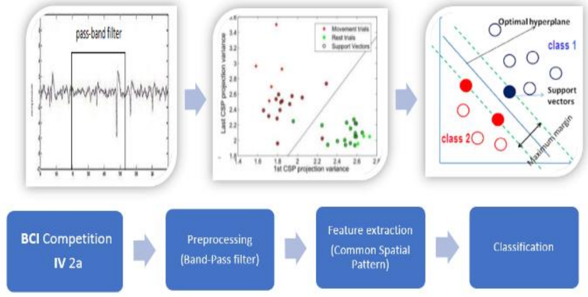

Fig. 1. Main steps of the considered BCI system.

2.1. BCI—competition IV dataset 2a

BCI-IV-2a dataset is a database that contains EEG data from nine subjects who participated in one training session and one evaluation session, and completed four various sorts of MI tasks (*right-hand*, *left-hand*, *tongue* and *both feet*). EEG signals were recorded using 22 $Ag/AgCl$ electrodes placed according to the worldwide 10-20 standard. All signals were monopolarly collected, with the right mastoid as ground and the left mastoid serving as reference. The EEG signals were sampled at 250 Hz and BP filtered between $[0.5\ and\ 100]$ Hz. The amplifier's sensitivity was set at 100 µV. To remove line noise, a supplementary 50 Hz notch filter was turned on. A training session (288 trials) and an independent evaluation session are included in this dataset. As the training and evaluation sessions were captured on two separate days [9].
The "general data format for biomedical signals" (GDF) is used to store all data sets [10].

2.2. Preprocessing

The second block of BCI system is the preprocessing of EEG signal. While, the signal recorded in the BCI-IV-2a dataset was frequently contaminated with artefacts from different noise sources either external or internal depending on their origin. The external ones, already filtered using a notch filter for line noise, and filtered by a BP for the other external noises. Internal artefacts result from physiological activities that are either the result of movements made by the subject himself or to abrupt fluctuations in bioelectric potentials, such as muscle activity (EMG), ocular activity (EOG), etc. Therefore, EEG denoising is an important preprocessing step for the majority of applications, in order to prepare the EEG data for the following steps. Furthermore, there are many methods used to remove artefacts in different domains such as spatial filtering, time-frequency filtering and adaptive filtering. In this paper, we will focus only on frequency filtering, and more precisely on infinite impulse response (IIR) filters which aim to separate frequencies by keeping only the desired bands: $mu(\mu)$ $[8 \mapsto 12\ Hz]$ and $beta\ (\beta)$ $[12 \mapsto 30\ Hz]$. The application of this filter requires the consideration of several parameters that influence the quality of the filtered EEG signal, such as bandwidth, transition bands, ripples in the bandwidth, etc.

2.3. CSP

In the feature extraction step, we used the CSP to detect and extract the features of interest from the EEG signal. This spatial method is used to improve the discrimination between two classes. The spatial filters are created to maximise the variance for the signals in the first condition while minimising it for the second condition [11-13]. To create a CSP technique that enables for the production of a time series with the best variances for discriminating between the two MI classes [14], it is necessary to start with the computation of the normalized spatial covariance of the EEG using the following formula [14]:

$$\boldsymbol{C} = \frac{\boldsymbol{\mathcal{E}.\mathcal{E}'}}{\boldsymbol{trace(\mathcal{E}.\mathcal{E}')}}$$

Where $\mathcal{E}$ represents the raw EEG data from a single trial, $\mathcal{E}'$ is its transpose and $trace(\mathcal{E}.\mathcal{E}')$ is the sum of the diagonal components of $(\mathcal{E}.\mathcal{E}')$. Averaging over the trial for each of the two distributions to be separated yields the spatial covariance. The $\mathcal{c}$omposite $\mathcal{s}$patial $\mathcal{c}$ovariance is calculated as follows: $\mathcal{C}s = \overline{\mathcal{C}}_1 + \overline{\mathcal{C}}_2$

With $\overline{\mathcal{C}}_k \in [class1, class2]$, and can be modeled by $\mathcal{C}_s = \mathcal{U}_s\, \lambda_s\, \mathcal{U}'_s$, where $\mathcal{U}_s$ is the eigenvector matrix and λ_s is the diagonal eigenvalue matrix. Note that the eigenvalues are supposed to be in descending order. The whitening transformation equalizes the variances in the space covered by $\mathcal{P} = \sqrt{\lambda_s^{-1}\, U_s^{-1}}$. If $\overline{\mathcal{C}_1}$ and $\overline{\mathcal{C}_2}$ are converted into $\mathcal{S}_1 = \mathcal{P}\, \overline{\mathcal{C}_1}\, \mathcal{P}'$ and $\mathcal{S}_2 = \mathcal{P}\, \overline{\mathcal{C}_2}\, \mathcal{P}'$ then $\mathcal{S}_1$ and $\mathcal{S}_2$ share common eigenvectors. The eigenvector with the highest eigenvalue for $\overline{\mathcal{S}_1}$ has the lowest eigenvalue for $\overline{\mathcal{S}_2}$, and vice versa, since the sum of two matching eigenvalues is always one. In the least squares sense, projecting the whitened EEG onto the first and last eigenvectors yields feature vectors that are optimal for distinguishing between two classes of EEG.
$\mathcal{W} = (\mathcal{M}'\mathcal{P})'$ is the projection matrix. Where $\mathcal{M}'$ is the mapping of a trial $\mathcal{E}$ is provided as follows
$\mathcal{Z} = \mathcal{W}\mathcal{E}$. The $\mathcal{W}^{-1}$ columns represent the common spatial patterns [14, 15].

2.4. SVM

SVM is a supervised learning method that has gained popularity in the classification areas in recent years. It was used it in the classification block. SVM is based on an algorithm that establishes a discriminant hyperplane (or a series of hyperplanes) that optimally maximizes the margins to distinguish the classes. The following expression can be used to compute the separation hyperplane in 2D feature space [16]:

$$g(x) = sgn\left(\sum_{i=1}^{P} \beta_i * y_i * ker(z_i, z) + b\right)$$

Where $ker(z_i, z)$ denotes a kernel function of the support vectors; z_i are the input vectors; y_i are the class labels y_i={1,2} (in our case we have left-hand class, right-hand class); β_i are the Lagrangian multipliers found by solving a quadratic optimization problem which varies between $0 \leq \beta_i \leq C$ (with C represents the penalty parameter.); $i = 1,2, \ldots P$ (P is the number of epochs.); b is the bias. The kernel usually applied in BCI system is the Gaussian kernel with the variable parameter γ that plays the main role in the performance of the kernel, and it has the following expression:

$$ker(z_i, z) = \exp\left(\frac{-\|z_i - z_j\|}{2\gamma^2}\right)$$

SVMs are considered as one of the most important models among the family of Kernel methods. They have gained strong popularity thanks to their good results achieved in BCI applications [17, 18]. Furthermore, the algorithm has been robust with a high-dimensional dataset and has been widely known as being the simplest algorithm used in BCI applications [19]. Moreover, the founding principle of SVMs is precisely to integrate into the estimation the control of the complexity, i.e. the number of parameters which is associated, in this case with the number of support vectors [20]. To estimate the reliability of an SVM model, the cross-validation technique was used. It is a practice of statistical partitioning of a data sample into subsets. Cross-validation is presented in three sub-methods *(𝕜-fold cross-validation, holdout and leave-one-out cross-validation*.) [21]. In our study, we used the most common one which is 𝕜 -fold cross-validation with $\Bbbk \in [4\ to\ 10]$.

3. Methodology

The EEG signals from the BCI-IV-2a dataset were processed using *Matlab.* First, a BP filter ($8 \mapsto 30\ \mathcal{H}z$) was employed to remove muscular artefacts, line-noise contamination and DC drift, we also used it to keep only the desired bands (mu and beta). Then, we employed the selection procedure to select the channel of interest. In the feature extraction step, we segmented EEG signals to epochs, after the cue appearance (our time segment = 3s). Afterwards, we calculated the spatial filters using the conventional CSP method. However, the study of a system usually requires a huge number of electrodes and an additional computing time to determine the best selection of channels. In our case, we selected ten electrodes around the sensorimotor cortex, including the channels "$C1, C2, C3, C4, C5, C6, \mathcal{F}C3, \mathcal{F}C4, CP3, CP4$" based on the international 10-20 system. This choice allows us to minimize the time needed to preprocess the dataset used. After this procedure of channel selection, we used the spatial filter that maximizes the variance of one MI class while minimizing the other class. The spatial filter is a CSP that calculated features whose variances that optimally distinguish between the two classes of MI responses. The obtained features were fed into the SVM for classification. SVM creates a hyper plane to separate the data sets. This method was applied, to improve the performance while reducing the complexity of the learning model.

To establish the impact of the filter type on the system performance, we have computed the classification accuracy (Acc) and the kappa coefficient. These parameters are defined as $kappa = \frac{p_a - p_e}{1 - p_e}$ where p_e is the expected percentage chance of agreement, p_a is the actual percentage of agreement [22].

Also, [23] $Acc = \frac{TrP + TrN}{TrP + TrN + FaP + FaN}$ where $TrP = True\_Positive$; $FaP = False\_Positive$; $TrN = True\_Negative$; $FaN = False\_Negative$.

4. Results and Discussion

Figures Fig.2, 3, 4 and 5 show the performance of the considered BCI system. The presented results (fig.2, 3, 4 and 5) were only for the first subject in the public database, BCI competition IV 2a. By conducting a test evaluation of several orders of the filter in the same database of EEG signals, we noticed that the classification accuracy fluctuates from one order to another.

The fig.2 shows that for the elliptic filter, the diagram starts from $acc = 81.25\%$, and it varies slightly with the filter order until the maximum value, $acc = 88.19\%$ which is found in $n = 7$. But this filter has also some drops in the orders $n = 5\ and\ 8$. For the Chebyshev filters, they are not steady. For different considered orders of Butterworth filter, the performance of the system increases with the increase of filter order. Fig.3 shows that almost the same evolution of the accuracy is shown for the kappa coefficient for all filters. We chose to enlarge the graphs of the elliptic filter to show the evolution of accuracy and kappa coefficient according to the filter order. As shown in the figure fig.4, the BCI system has an accuracy of $acc = 81.25$ for order $n = 1$, $acc = 86.11$ for order $n = 2$ and it takes a stable value, $acc = 86.80$ in the orders $n = 3$ and 4, then it decreases to $acc = 85.41$ for order $n = 5$, and finally the accuracy increases to a maximum value, $acc = 88.19$ for $n = 7$. In fig.5, we can see that the maximum value of the kappa coefficient is 0.7639 that can be found in order $n = 7$. Through this study, we observe that, with an increase in the filter order, the behaviour of the system becomes more accurate for the Butterworth and elliptic filters. For the Butterworth filter, we may state that the order $n = 8$ is the best performing with an accuracy of $acc = 87.5\%$. For the elliptic filter, the high accuracy was $acc = 88{,}19\%$ in $n = 7$, while the maximum accuracy for the chebyshev2 is in $n = 5$ with the value of $acc = 86{,}80\%$. From the evolution of the kappa coefficient according to filter order, the maximum value is 0.7639 which associates the $9th$ order elliptical filter. We chose the lowest order, which has the highest accuracy because the filter order is always related to the number of components that determines the cost and the complexity of the design.

The results obtained in our study demonstrate that it is a valuable and promising strategy to investigate the influence of the BP filter (preprocessing step) on the brain-computer interface system.

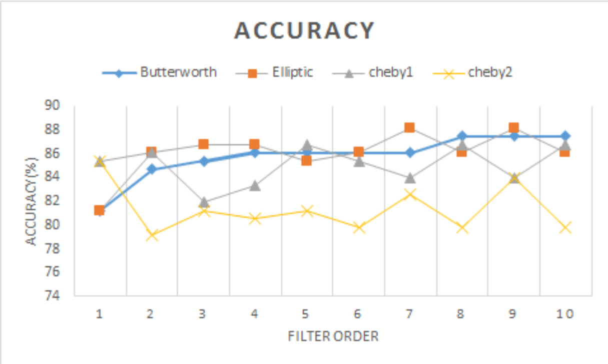

Fig. 2. Accuracy evolution of the BCI system according to the filter type and its order (for IIR filters).

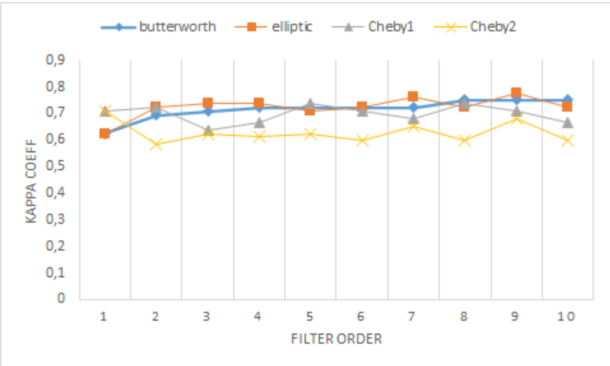

Fig. 3. Evolution of the Kappa coefficient of the BCI system according to the filter type and its order (for IIR filters).

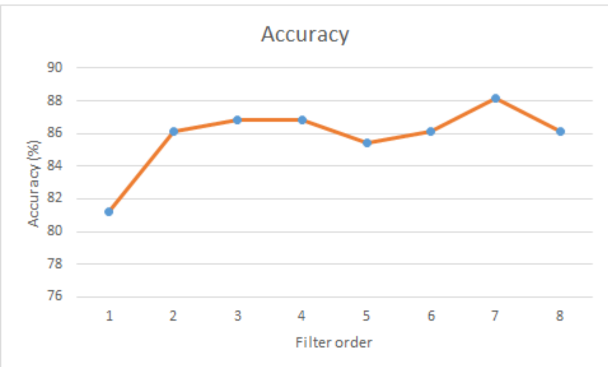

Fig. 4. Accuracy evolution of the BCI system for elliptic filter according to its order.

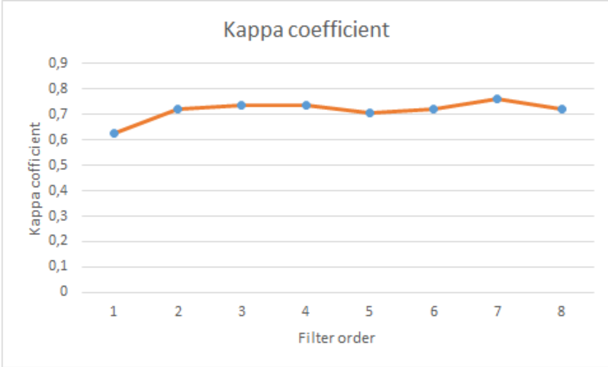

Fig. 5. Evolution of the Kappa coefficient of the BCI system for elliptic filter according to its order.

5. Conclusion

This paper is completely based on the changes made in the preprocessing step of BCI paradigm. In the accuracy term, the $7th$ order elliptic filter is the best candidate than the others with an accuracy of 88,19%. In contrast the 4th order Butterworth filter can offer the best results in several orders than the other IIR filters. The choice of filter order can affect the performance of the BCI system. Therefore, the order must be carefully chosen so as not to lose the important information, which leads to a degradation of the system performance.

References

1. M.Rashid, N.Sulaiman, A.PP Abdul Majeed, R.Muazu Musa, B.Sama Bari, S.Khatun, et al. *Frontiers in neurorobotics*, **14**, 25 (2020).
2. S.Chakrabarti, A.Swetapadma, A.Ranjan, and P.Kumar Pattnaik. *Biomedical Signal Processing and Control*, **59**, 101930 (2020).
3. F.Ebrahimi, M.Mikaeili, E.Estrada, and H.Nazeran. In 2008 30th Annual International Conference of the *IEEE Engineering in Medicine and Biology Society*, pages 1151-1154. (2008).
4. L.Cao, J.Li, H.Ji, and C.Jiang. *Journal of neuroscience methods*, **229,** pages 33-43 (2014).
5. J.Zhao, W.Li, and M.Li. *PLoS One, Public Library of Science*, **10,** e0142168 (2015).
6. J.Meng, S.Zhang, A.Bekyo, J.Olsoe, B.Baxter, and B.He. *Scientific Reports*, *Nature Publishing Group*, **6**, 38565 (2016).

7. X.Bai, X.Wang, S.Zheng, and M.Yu. *In Proceedings of the 33rd Chinese control conference*, pages 7189-7194 (2014).

8. B.Aji Pramudita, S.Wibirama, L.Iznita Izhar, N.Akhmad Setiawan, et al. In 2018 *International conference on Intelligent and Advanced System (ICIAS)*, pages 1-6. (2018).

9. C.Brunner, R.Leeb, G. Müller-Putz, A.Schlögl, and G.Pfurtscheller. *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology*, **16**, pages 1-6 (2008).

10. M.Tangermann, KR.Müller, A.Aertsen, N.Birbaumer, Ch.Braun, C.Brunner, R.Leeb, C.Mehring, K.J Miller, G.Mueller-Putz, et al. *Frontiers in neuroscience*, **6** (2012).

11. B.Blankertz, R.Tomioka, S.Lemm, M.Kawanabe, and KR.Mueller. *IEEE Signal Processing Magazine*, **25**(1), pages 41-56 (2008).

12. F.Lotte and C.Guan. *IEEE Transactions on biomedical Engineering*, **58**(2), pages 355-362, (2010).

13. GR.Müller, Ch.Neuper, and G.Pfurtscheller. *Biomedical Engineering*, **46**, pages 186-190 (2001).

14. H.Ramoser, J.Muller-Gerking, and G.Pfurtscheller. IEEE transactions on rehabilitation engineering, **8**(4), pages 441-446 (2000).

15. Y.Ning, M.Zhan, S.Yuge, and W.Xu. In 2009 Chinese Control and Decision Conference, pages 5169-5171 (2009).

16. Vl.N Vapnik. An overview of statistical learning theory. *IEEE transactions on neural networks*, **10**(5), pages 988-999 (1999).

17. D.Garrett, D.A Peterson, C.W Anderson, and M.H Thaut. *IEEE Transactions on neural systems and rehabilitation engineering*, **11**(2), pages 141-144 (2003).

18. A.Rakotomamonjy, V.Guigue, G.Mallet, and V.Alvarado. *In International conference on artificial neural networks*, *Sringer*, pages 45-50 (2005).

19. M.Kaper, P.Meinicke, U.Grossekathoefer, T.Lingner, and H.Ritter. *IEEE Transactions on biomedical Engineering*, **51**(6), pages 1073-1076 (2004).

20. V.Vapnik. The nature of statistical learning theory. *Springer science & business media* (2013).

21. D.Krstajic, L.J Buturovic, D.E Leahy, and S.Thomas. *Journal of cheminformatics*, **6**(1), pages 1-15, (2014).

22. H.C.Kraemer, *Psychometrika, Springer* **44**, Pages 461-472 (1979).

23. S.Raschka, *arXiv preprint* (2014).