Journal of Big Data

Classification of SSVEP-based BCIs using Genetic Algorithm

Hamideh Soltani[1], Zahra Einalou[2*], Mehrdad Dadgostar[1] and Keivan Maghooli[3]

*Correspondence:
z.einalou@srbiau.ac.ir
[2] Department of Biomedical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran
Full list of author information is available at the end of the article

Abstract

Brain computer interface (BCI) systems have been regarded as a new way of communication for humans. In this research, common methods such as wavelet transform are applied in order to extract features. However, genetic algorithm (GA), as an evolutionary method, is used to select features. Finally, classification was done using the two approaches support vector machine (SVM) and Bayesian method. Five features were selected and the accuracy of Bayesian classification was measured to be 80% with dimension reduction. Ultimately, the classification accuracy reached 90.4% using SVM classifier. The results of the study indicate a better feature selection and the effective dimension reduction of these features, as well as a higher percentage of classification accuracy in comparison with other studies.

Keywords: Brain signal, Visual evoked signal, Classification, Genetic algorithm, Wavelet transform

Introduction

A brain-computer interface (BCI), also called "a mind-machine interface", is a bridge between the human brain and an external device. This bridge is a new communication pathway which is expanding. This interface consists of a set of sensors and signal processing components that directly convert one's brain activity into a series of communication or control signals. In this system, brain waves should firstly be captured using a brain wave recording apparatus. In most cases, EEG is utilized to record brain waves due to its high temporal accuracy, inexpensiveness and availability. Brain-computer interfaces are pathways through which a computer either receives orders from the brain or sends signals to it [1]. In recent years, visual-evoked potential based BCIs have gained a unique place among BCI-based systems and are used for different purposes, due to various features such as high reliability diagnosis, low training requirements, high information transfer rates, and the improved voluntary function of the user. Many approaches have been proposed to perform processing and frequency identification in Steady state visually evoked potential (SSVEP)-based BCI systems [2]. Chevallier et al. [3] applied Riemannian geometry and covariance matrices to extract features from SSVEP signals. On the contrary, Cao et al. [4] used fuzzy features and fuzzy entropy and finally performed an assessment. Sözer et al. [5] presented a combined approach for extracting features from SSVEP signals. In this approach, a reference

signal is obtained by the combination of electrode signals, and multiple regression analysis is used to determine the optimal coefficients for this combination. On the contrary, Aznan et al. [6] utilized a convolutional neural network method to classify EEG signals. In this paper, seven features were extracted using wavelet transform and five features was selected using the genetic algorithm. GA chooses the best features for SSVEP signal classification to use them in BCI systems. Genetic algorithms are simple to implement. It was shown that GAs could randomly seek the optimal solution for a problem. Many feature selection methods use metaheuristic algorithms to avoid increasing computational complexity. GAs will be able to optimize the problem of feature selection with appropriate accuracy within an acceptable time [7].
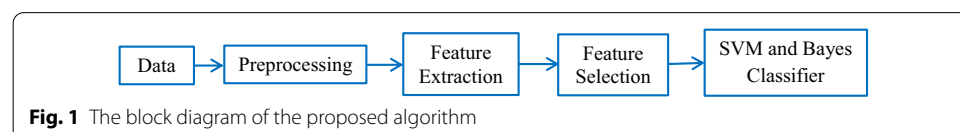
Materials and methods

This paper proposes a model for feature extraction, feature selection and, finally, classification. In the block diagram of Fig. 1, the conceptual model of the proposed approach is presented. In the first step, the data is collected. This data is processed, which includes noise elimination and data preparation for the next steps. In the second step, the features are extracted from the signal, and in the third step, the best features are selected for SSVEP signal detection by using evolutionary approaches. Finally, using several classifiers, these signals are classified to determine to which signal stimulus the subject attends. In the following, each step is investigated separately and in a more detailed manner.

Data acquisition

A device made by Biosemi Inc., Netherlands, was utilized for data recording. The data is recorded at the RIKEN Science Institute, Laboratory for Advanced Brain Signal Processing. A total of 128 active electrodes have recorded electroencephalogram signals from 4 subjects. The SSVEP stimulation was achieved using a small chessboard for 3 stimulation frequencies (8, 14 and 28 Hz). The sampling rate of signals was 256 Hz. The subjects participating in the experiment, sat at a 90-cm distance from the monitor. SSVEP begins 5 s after the start of the data and ends 20 s after it. 15 s of SSVEP are recorded from 4 subjects. For each person, five SSVEP visual stimulators are displayed to each tested subject, which are the images of personal emotions such as grief, happiness and anger [8, 11].

Pre-processing

In the data pre-processing stage, the external noise caused by mains electricity or blinking must be eliminated [25]. A signal histogram is used to display the distribution of potentials at each frequency. To remove the noise, a high-pass filter is applied to remove the excess values of each frequency. Finally, signals are denoised to regenerate the original signal according to the filters. In order to remove the effect of electrodes on the signal, a spatial filter called CAR, made up of Large and Small Laplace, is used and ultimately applied on the signal.

Fig. 1 The block diagram of the proposed algorithm

Soltani *et al. J Big Data*   (2021) 8:85

Page 3 of 11

Feature extraction

EEG signals are non-stationary and transient, the analysis of which requires the use of transform approaches such as wavelet transform [12], for the signals under study to be applied. It is presented in Eq. 1:

$$W(a,b) = \int\limits_{-\infty}^{+\infty} X(t) \frac{1}{\sqrt{a}} \psi \times \left(\frac{t-b}{a} \right) dt.$$ (1)

In the above equation, ψ is the wavelet function expressing the general condition. b is the translation parameter, indicating the value of scaling. a changes the time scale of the probe function ψ. If $a > 1$ the wavelet function is dilated, and if $a < 1$, the function would be compacted [16, 17, 19–21, 23, 24].

Since the EEG signal does not contain useful components at frequencies above 30 Hz, five levels are used for decomposition [22]. For each tested subject 128 sensors are utilized, and each signal is divided into 5 components. For each subject and the estimated characteristics from each signal, 640 components are investigated. A variety of approaches exist for extracting features from signals [28, 29]. The band Power is measured using the following equations.

$$p(n) = x^2(n),$$ (2)

$$\bar{p}(n) = \frac{1}{\omega} \sum_{k=1}^{\omega} x(n-k).$$ (3)

The Eqs. 2 and 3 indicate that the frequency band feature per n equals to the mean power, ω, of the previous samples, and the final value of the frequency band feature is obtained from the Eq. 4. The *ln* is used to improve the classification.

$$\text{BandPower} = \ln \overline{p(n)}.$$ (4)

Each signal consists of entropy and redundancy. Useful and valuable information can be acquired from variable data such as signals that have high levels of variation. This useful information is entropy, which is chaotic and completely random [9]. In a variable data, the random variable X expresses the states of the system. The values of the variable X are as shown in Eq. 5.

$$X = \{x_1, x_2, \dots, x_n\}, \quad n > 0.$$ (5)

And the probability of each is,

$$P = \{p_1, p_2, \dots, p_n\} \quad 0 \le p_i \le 1, \quad i = 1, 2, \dots, n,$$ (6)

where,

$$\sum_{i=1}^{n} p_i = 1.$$ (7)

Therefore, the entropy information of this system is obtained by the Eq. 8.

$$H = -\sum_{i=1}^{n} p_i \ln p_i. \tag{8}$$

Spectral entropy is one of these important entropies, which makes an estimation of the complexity of time series. There are several transforms to obtain the spectrum of signal. Fourier transform is one of the approaches measuring the signal spectral density. This transform expresses the spectral density distribution in the form of a function of signal frequency. Walsh transform was applied to acquire a new type of feature called the sequence entropy. The Fourier transform of the signal is obtained which yields the amount of spectrum, W_i, at the point *i*. The power spectral density is acquired.

$$\hat{P}(w_i) = \frac{1}{N}|X(w_i)|^2. \tag{9}$$

The power spectral density is normalized. The spectral density is obtained using Eq. 10.

$$P_i = \frac{\hat{P}(w_i)}{\sum_{i}^{n} \hat{P}(w_i)}. \tag{10}$$

Finally, 7 features were extracted including the features extracted by wavelet transform. According to the EEG frequency band, A5, D2, D3, D4 and D5 were chosen as shown in Table 1. The following feature extraction parameters were implemented to represent the experimental signals as shown in Table 2.

Feature selection

After extracting features, feature selection is done using the GA. In general, the process of feature selection is based on selecting the best features of all. Figure 2 presents the general process of feature selection using the GA.

In the first, different subsets of features with different dimensions are formed. These subsets are evaluated according to the fitness function written in the GA and the best subset of features is selected as the best features [18]. The GA select features by using coded decision variables, rather than features themselves. It begins with a set of answers, not one specific answer. This algorithm uses a random transfer to achieve the best answer and has few computational rules. The GA uses the information of fitness function, rather than derivatives and other information The initial population size consists of 100 generations, the crossover rate, 0.75 and the mutation rate, 0.02. The elitism

Table 1 Decomposition coefficient

Frequency range (Hz)	Decomposition Coefficient
0–2.7	A5
2.7–5.4	D5
5.4–10.8	D4
10.8–21.7	D3
21.7–43.4	D2
43.4–86.8	D1

Table 2 Extracted features

Feature No.	Features
1	Power spectrum of A5 coefficient
2	Power spectrum of D5 coefficient
3	Power spectrum of D4 coefficient
4	Power spectrum of D3 coefficient
5	Power spectrum of D2 coefficient
6	Spectral entropy
7	Approximation entropy

Fig. 2 The flowchart for feature selection using the GA
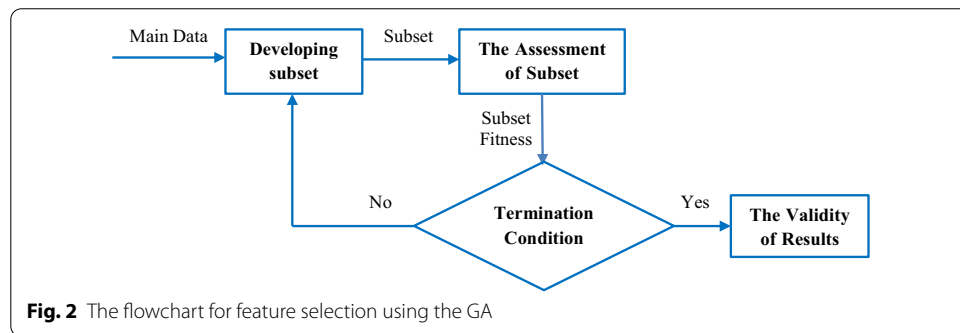
approach is applied for replacing the next generation. In order to find the best features through GA, the fitness function is designed for the selection of a set of features which would result in the best simulation (the lowest mean square error) of these features for the reconstruction of the signal.

$$Fitness = w_a \times sig - Accuracy + w_f \times \left(\sum_{i=1}^{n_f} g_f^i \right)^{-1}, \tag{11}$$

where W is the weight of each feature under study. Sig-Accuracy is the accuracy of signal obtained by the reconstruction using the selected features. In general, 7 features extracted from signals have been used for feature selection. In this feature selection approach, in order to find the solution, each of the features is considered as a gene in the chromosome and the algorithm searches for the best features to simulate the output signal. The algorithm seeks the best subset of features. One of the issues here is the number of features to select from. In order to solve this problem, 3, 4 and 5 features are selected, respectively. In other words, at first, 3 features are selected from a total of 7 features and the mean square error (MSE) is measured. In the second step, 4 features are selected and the MSE is determined. In the last step, 5 features are selected. Finally, according to the MSE, the best output is selected. It can be concluded that the MSE does not vary much among 3, 4 and 5 features.

Classification

SVM and Bayes Classifiers have been used for classification [13]. The Bayes classifier estimates the distribution of class probability and attributes, and the learning

is reduced to the level of probability estimation. As a result, there are differences in comparison with other classification approaches. This algorithm is widely used for implementation due to its high speed and simplicity. The structure of Bayesian network expresses the internal relationship between the features and can also be applied for incomplete sets of data. Experts can easily understand its structure and change it in case better predictions are necessary. Learning in this network consists of the two main tasks of graph structure learning, and then the parameter learning for structures. This algorithm originates from the main equation of Bayes theorem, which is displayed as follows:

$$p(c_i|x) = \frac{p(x|c_i) \cdot p(c_i)}{p(x)}. \tag{12}$$

In Eq. 12, c_i of the class i and x are independent variables. The condition is stated as a set of conditions. Therefore,

$$P(x) = \prod_{j=1}^{n} p(x_j), \tag{13}$$

$$P(x|c_i) = \prod_{j=1}^{n} p(x_j|C_i), \tag{14}$$

$$p(C_i|x) = \frac{\prod_{j=1}^{n} p(c_i) \times P(x_j|C_i)}{P(x)}. \tag{15}$$

Finally, the class with the highest probability is selected. The naïve Bayes network explains the conditional dependence among the features and can determine the dependence between the variables. The Bayes network is made up of nodes that are the same as variables. Nodes in Bayes network are variables selected from the set of features. The edges of each node indicate the dependence of variables on each other. SVM performs a nonlinear mapping on the data to transform them into a higher-dimension space, and searches for a separating linear hyperplane. It can be used both for classification and prediction [10, 14, 15, 18]. The goal is to find the best line, the best plane, or the best hyperplane to separate unseen data from all possible modes. SVM does this with Maximal Marginal Hyperplane (MMH). The margins should be of the same distance from either sides of the separating hyperplane and parallel to it.

The hyperplane equation is as,

$$W \cdot X + b = 0. \tag{16}$$

Where w is a weight vector and we have, $w = \{w_1, w_2 \ldots w_n\}$. In addition, n is the number of indicators and b, a scalar. For example, in a 2D feature space, the linear equation of the separator is as,

$$W_0 + W_1 \cdot X_1 + W_2 \cdot X_2 = 0. \tag{17}$$

Any point falling above this line is shown as,

Soltani *et al. J Big Data*   (2021) 8:85

Page 7 of 11

$$W_0 + W_1 \cdot X_1 + W_2 \cdot X_2 > 0. \tag{18}$$

Any point below the line is represented as,

$$W_0 + W_1 \cdot X_1 + W_2 \cdot X_2 < 0. \tag{19}$$

Result

For performance analysis, the EEG dataset is divided into training set and testing set according to 3-times cross validation scheme. Feature optimization including GA based methods are performed on the training set [26, 27]. The classification accuracy which is obtained on the testing data is employed for performance comparison. Three, four and five group-extracted feature are given to classifier respectively and MSE is measured. According to the MSE, the best output is selected as shown in Table 3.

According to results of SVM and Bayes Classifiers, we compare our proposed method with other studies as seen as Fig. 3. However, they used approaches such as decision tree and Bhattacharyya for feature selection with Bayes as a classifier. Also continuous wavelet transform (CWT) and Spectral F-Test (SFT) for feature selection with SVM [11, 13].
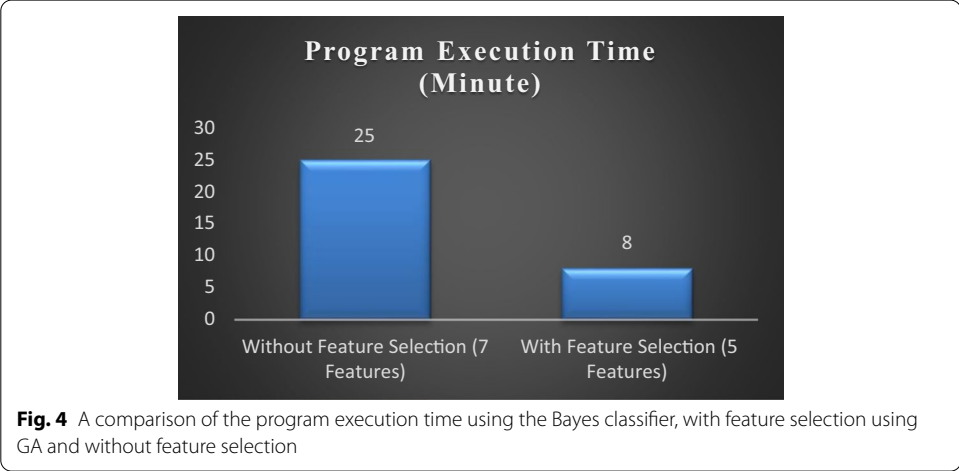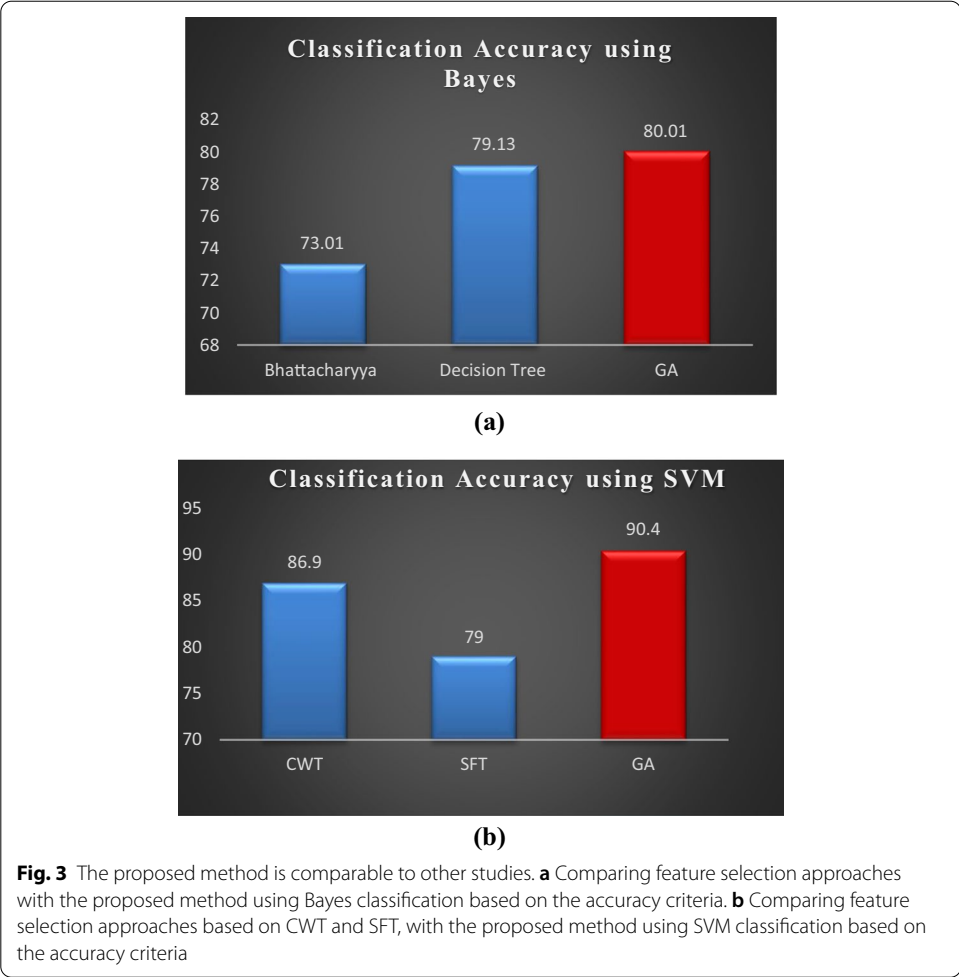
Figure 4 shows a comparison between classification approaches with and without feature extraction. This comparison is done based on the classification accuracy criterion, program execution time.

Discussion

This study proposes a simple and straightforward algorithmic method to improve the classification of BCI data. The proposed method reduces the input features space dimensions and reconstructs the format of the features to optimize the classifier. So the training time has been reduced and the accuracy is improved. Heidari et al. [11] used a same database to this study, and applied wavelet transform for feature extraction, Bayes classifiers and SVM as shown in Fig. 3. The proposed model is more accurate than the two feature selection models of decision tree and Bhattacharyya. However, the number of features selected in the proposed method using GA is lower and the selected features are more optimal. In the following, according to Anupama et al. [12], the proposed model has been compared with the SVM classification approach. The SVM classifier is very sensitive to the feature selection strategy [13]. In their study, the approaches used for feature extraction and feature selection included CWT and SFT as shown in Fig. 3. In this paper, seven features were extracted using wavelet transform and five features was selected using the genetic algorithm. Chromosomes with binary genes are considered with each chromosome length (7) equal to the number of features, while a zero state eliminates a feature and one means feature existence in classification. GA chooses the

Table3 The number of features and MSE computation in each case

The number of selected features	MSE
6, 5, 2	1.3936
3, 7, 1, 4	1.3927
2, 7, 5, 1, 3	1.3918

Soltani *et al. J Big Data*   (2021) 8:85

Page 8 of 11

Fig. 3 The proposed method is comparable to other studies. **a** Comparing feature selection approaches with the proposed method using Bayes classification based on the accuracy criteria. **b** Comparing feature selection approaches based on CWT and SFT, with the proposed method using SVM classification based on the accuracy criteria

Fig. 4 A comparison of the program execution time using the Bayes classifier, with feature selection using GA and without feature selection

best features for SSVEP signal classification to use them in BCI systems. Genetic algorithms are simple to implement. It was shown that GAs could randomly seek the optimal solution for a problem. Many feature selection methods use metaheuristic algorithms

to avoid increasing computational complexity. GAs will be able to optimize the problem of feature selection with appropriate accuracy within an acceptable time. Based on the output results, it can be concluded that the proposed model is proper for feature selection, and can be used to reduce the processing size, and also BCI costs [28, 29]. Finally, a comparison of classification is based on the classification accuracy, the program execution time and the initial database size. After feature extraction using wavelet transform, the total size of data is 950 MB at the first stage. Each tested signal from one sensor is decomposed into 5 signals, and finally, an entropy feature is extracted from the signal. After feature selection the total size of the data in the database is reduced to 240 MB. Due to the fact that the features selected using the fitness function are the best features, it is possible to reconstruct the original signal with high accuracy from these selected features. The program execution time of the classification algorithm is approximately 8 min with selected feature, and almost 25 min without feature selection. Finally, although applying the Bayes algorithm has slightly lowered the classification accuracy, the reduced execution time of the algorithm, as well as the decreased size of database, makes the implementation of feature selection based on the GA more effective for classification. However, feature extraction and selection are considered as a challenging issue for researchers.

Conclusion

This paper utilized wavelet transform, GA, and classification using Bayes and SVM to improve the performance of BCI. Based on the output results, it can be concluded that the proposed model is proper for feature selection, and can be used to reduce the processing size, and also BCI costs. Furthermore, the classification accuracy using the SVM classifier reached 90.4%, indicating better results compared with the Bayes approach. The SVM classifier is very sensitive to the feature selection strategy. The results of this study indicated better feature selection and the effective dimension reduction of these features. For future studies, it is recommended to use other databases, the empirical mode decomposition approach and other methods of signal processing, and to apply optimal feature selection approaches such as the cuckoo algorithm, colonial competitive algorithm, fuzzy logic, neuro-fuzzy network, as well as using several classifiers for classification.

Declarations

Ethics approval and consent to participate
Ethical approval was waived by the local Ethics Committee of University Islamic Azad University in view of the retrospective nature of the study and all the procedures being performed were part of the routine care.

Author details
[1]Department of Biomedical Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran. [2]Department of Biomedical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran. [3]Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

References

1. Rejer I. Wavelet transform in detection of the subject specific frequencies for SSVEP-based BCI. In: International multi-conference on advanced computer systems. Cham: Springer; 2016. p. 146–55.
2. Wang Y, Gao X, Hong B, Jia C, Gao S. Brain–computer interfaces based on visual evoked potentials. IEEE Eng Med Biol Mag. 2008;27(5):64–71.
3. Chevallier S, Kalunga E, Barthélemy Q, Yger F. Riemannian classification for SSVEP based BCI: offline versus online implementations. 2018.
4. Cao Z, Lin CT, Lai KL, Ko LW, King JT, Liao KK, Fuh JL, Wang SJ. Extraction of SSVEPs-based inherent fuzzy entropy using a wearable headband EEG in migraine patients. IEEE Trans Fuzzy Syst. 2019;28(1):14–27.
5. Sözer AT, Fidan CB. Novel spatial filter for SSVEP-based BCI: a generated reference filter approach. Comput Biol Med. 2018;96:98–105.
6. Aznan NKN, Bonner S, Connolly J, Al Moubayed N, Breckon T. On the classification of SSVEP-based dry-EEG signals via convolutional neural networks. In: 2018 IEEE international conference on systems, man, and cybernetics (SMC). IEEE; 2018. p. 3726–31.
7. Mehrdad R, Kamal B, Saman F. A novel community detection based genetic algorithm for feature selection. J Big Data. 2021;8(1):1–27.
8. Bakardjian H, Tanaka T, Cichocki A. Emotional faces boost up steady-state visual responsesforbrain–computer interface. NeuroReport. 2011;22(3):121–5.
9. Poryzala P, Materka A. Cluster analysis of CCA coefficients for robust detection of the asynchronous SSVEPs in brain–computer interfaces. Biomed Signal Process Control. 2014;10:201–8.
10. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273–97.
11. Heidari H, Einalou Z. SSVEP extraction applying wavelet transform and decision tree with bays classification. Int Clin Neurosci J. 2017;4(3):91–7.
12. Anupama HS, Jain RV, Venkatesh R, Mahadevan R, Cauvery NK, Lingaraju GM. Implementing and analyzing different feature extraction techniques using EEG-based BCI. In: Recent findings in intelligent computing techniques. Singapore: Springer; 2018. p. 377–86.
13. Carvalho SN, Costa TB, Uribe LF, Soriano DC, Yared GF, Coradine LC, Attux R. Comparative analysis of strategies for feature extraction and classification in SSVEP BCIs. Biomed Signal Process Control. 2015;21:34–42.
14. Dadgostar M, Setarehdan SK, Shahzadi S, Akin A. Classification of schizophrenia using SVM via fNIRS. Biomed Eng Appl Basis Commun. 2018;30(02):1850008.
15. Mohammadi L, Einalou Z, Hosseinzadeh H, Dadgostar M, Einalou Z. Cursor movement detection in brain–computer-interface systems using the hybrid K-means clustering method and LSVM. 2020.
16. Oskooei SG, Dadgostar M, Rad GR, Fatemizadeh E. Adaptive watermarking scheme based on ICA and RDWT. 2009.
17. Dadgostar M, Setarehdan SK, Akin A. Detection of motion artifacts in fNIRS via the continuous wavelet transform. In: 2013 20th Iranian conference on biomedical engineering (ICBME). IEEE; 2013. p. 243–6.
18. Einalou Z, Maghooli K, Setarehdan SK, Akin A. Effective channels in classification and functional connectivity pattern of prefrontal cortex by functional near infrared spectroscopy signals. Optik. 2016;127(6):3271–5.
19. Einalou Z, Maghooli K, Setarehdan SK, Akin A. Graph theoretical approach to functional connectivity in prefrontal cortex via fNIRS. Neurophotonics. 2017;4(4):041407.
20. Einalou Z, Mehran YZ, Maghuly K. Diagnosis and classification of digestive diseases by wavelet transform. In: First joint congress on fuzzy and intelligent systems (ISFS2007). 2007.
21. Einalou Z, Maghooli K. Fuzzy neural network approach for noninvasive diagnosis of digestive diseases using wavelet comparing to classification followed by fuzzy C-mean algorithm. In: 2010 17th Iranian conference of biomedical engineering (ICBME). IEEE; 2010. p. 1–4.
22. Dehghanpour P, Einalou Z. Evaluating the features of the brain waves to quantify ADHD improvement by neurofeedback. Technol Health Care. 2017;25(5):877–85.
23. Shirzadi S, Einalou Z, Dadgostar M. Investigation of functional connectivity during working memory task and hemispheric lateralization in left-and right-handers measured by fNIRS. Optik. 2020;221:165347.
24. Dadgostar M, Setarehdan SK, Shahzadi S, Akin A. Functional connectivity of the PFC via partial correlation. Optik. 2016;127(11):4748–54.
25. Faskhodi MM, Einalou Z, Dadgostar M. Diagnosis of Alzheimer's disease using resting-state fMRI and graph theory. Technol Health Care. 2018;26(6):921–31.

26. Majkowski A, Kołodziej M, Zapała D, Tarnowski P, Francuz P, Rak RJ, Oskwarek Ł. Selection of EEG signal features for ERD/ERS classification using genetic algorithms. In: 2017 18th international conference on computational problems of electrical engineering (CPEE). IEEE; 2017. p. 1–4.
27. Jukiewicz M, Buchwald M, Czyż A. Optimizing SSVEP-based brain–computer interface with CCA and genetic algorithms. In: 2019 signal processing: algorithms, architectures, arrangements, and applications (SPA). IEEE; 2019. p. 164–8.
28. Mehdizavareh MH, Hemati S, Soltanian-Zadeh H. Enhancing performance of subject-specific models via subject-independent information for SSVEP-based BCIs. PLoS ONE. 2020;15(1):e0226048.
29. Garrett D, Peterson DA, Anderson CW, Thaut MH. Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. IEEE Trans Neural Syst Rehabil Eng. 2003;11(2):141–4.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.