

Classification of SSVEP-based BCIs using Genetic Algorithm

Hamideh Soltani

Islamic Azad University Central Tehran Branch

Zahra Einalou (✉ z.einalou@srbiau.ac.ir)

Islamic Azad University Tehran North Branch

Keivan Maghooli

Islamic Azad University Science and Research Branch

Research

Keywords: brain signal, visual evoked signal, classification, genetic algorithm, wavelet transform

Posted Date: December 2nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-119561/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Classification of SSVEP-based BCIs using Genetic Algorithm

Hamideh Soltani¹, Zahra Einalou^{2*}, Keivan Maghooli³

¹Department of Biomedical Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran.

^{2*}Department of Biomedical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran.

³Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

Corresponding author

Zahra Einalou

Department of Biomedical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran

E-mail: z.einalou@srbiau.ac.ir

Phone: +98 21 77318896

Fax: +98 21 77318896

Abstract

In recent years, brain-computer communication systems have been regarded as a new way of communication for humans. One of the applications of brain-computer communication is the development of systems which facilitates communication. To this end, it is necessary to extract the visually evoked signals from the EEG signal and classify it. In this research, common methods such as wavelet transform are applied in order to extract features. However, genetic algorithm, as an evolutionary method, is used to select features. Finally, after selecting features, the classification was done using the two approaches support vector machine and Bayesian method. Five features were selected and the accuracy of Bayesian classification was measured to be 80% with dimension reduction, and 78% without 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

1. 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 SSVEP-based BCI systems. [2] Chevallier et al. (2019) [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 preformed an assessment. In a 2018 research, 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. (2018) [6] utilized a convolutional neural network method to classify EEG signals. In this paper, the features were extracted through wavelet transform and feature selection was done using the genetic

algorithm. Genetic algorithm enables us to choose the best features for SSVEP signal classification among the features of the extracted signal and to use them in BCI systems.

2. Materials and Methods

This paper proposes a model for feature extraction, feature selection and, finally, classification. In the block diagram of Fig1, the conceptual model of the proposed approach is presented. In the first step, the research data is collected. The 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.

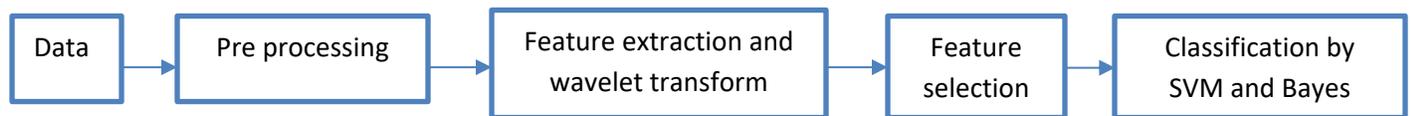


Fig1. The block diagram of the proposed algorithm

2.1. Data Acquisition

The used data are collected by Bakardjian, available on the reference link [7]. 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 captured 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 seconds after the start of the data and ends 20 seconds after it. Fifteen seconds of SSVEP are captured 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].

2.2. Pre-Processing

In the data pre-processing stage, the external noise caused by mains electricity or blinking must be eliminated. A signal histogram is used to display the distribution of potentials at each frequency. To remove the noise, all wavelet signals are extracted. A high-pass filter is applied to remove the excess values of each frequency. Finally, wavelet 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.

2.3. Feature Extraction

EEG signals are non-stationary and transient, the analysis of which requires the use of transform approaches such as wavelet transform, for the signals under study to be applied. The conversion used in this study is presented in Equation 1:

$$w(a,b) = \int_{-\infty}^{+\infty} X(t) \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) dt \quad (1)$$

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.

Since the EEG signal does not contain useful components at frequencies above 30 Hz, five levels are used for decomposition. 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. The band Power is measured using the following equation.

$$p[t] = x^2[t] \quad (2)$$

$$\bar{p}[n] = \frac{1}{\omega} \sum_{k=0}^{\omega} [n-k] \quad (3)$$

The equations 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 equation 4. The \ln is used to improve the classification.

$$\text{Band Power} = \ln(\bar{p})[n] \quad (4)$$

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 Equation 5.

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

And the probability of each is,

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

Where,

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

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

$$H = - \sum_{i=1}^n p_i \ln p_i \quad (8)$$

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 (wavelet) of the signal is obtained which yields the amount of spectrum, W_i , at the point i .

- The spectral density is acquired.

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

The spectral density is normalized.

$$P_i = \frac{\hat{P}(w_i)}{\sum_i^n \hat{P}(w_i)} \quad (10)$$

-The spectral density is obtained using Equation 10.

Finally, 8 features were extracted including the five signals extracted by wavelet transform, one frequency band power, one spectral entropy, and one approximation entropy.

2.4. Feature Selection

After extracting signal features, feature selection is done using the genetic algorithm. 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 genetic algorithm.

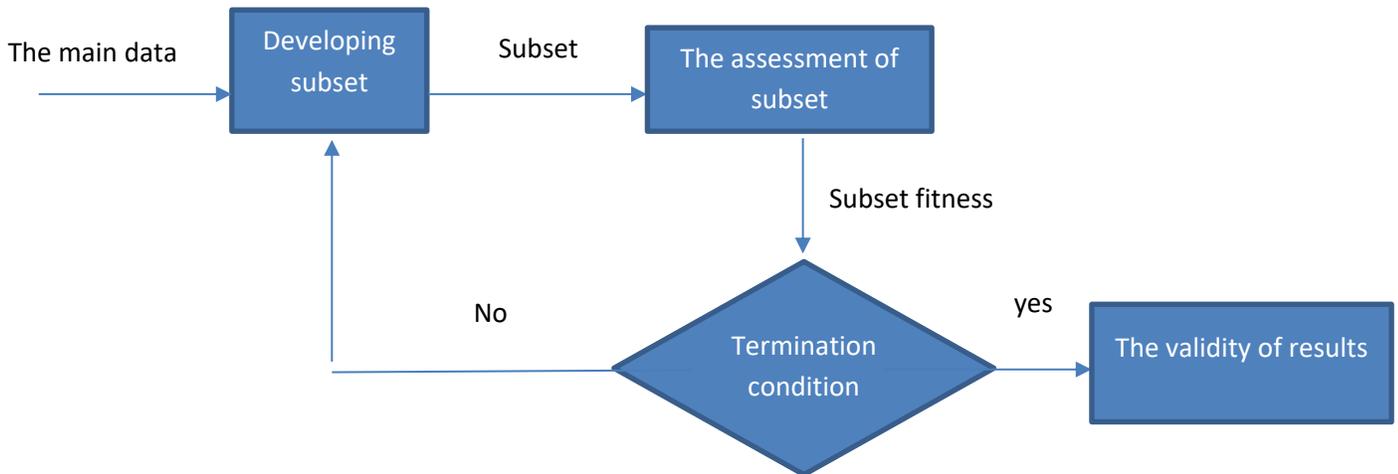


Figure 2: The flowchart for feature selection using the genetic algorithm

In the first stage, different subsets of features with different dimensions are formed. These subsets are evaluated according to the goal function written in the genetic algorithm and the best subset of features is selected as the best features. The advantages of using the genetic algorithm for feature selection are as follows:

- The genetic algorithms 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 genetic algorithm uses the information of fitness function, rather than derivatives and other information.

Due to the above advantages, the genetic algorithm is used for feature selection. Special settings have been applied for the genetic algorithm. The initial population size consists of 100 generations, the crossover rate, 0.75 and the mutation rate, 0.02. The elitism approach is applied for replacing the next generation. There are conditions set for terminating the genetic algorithm:

- If the goal function rate does not improve in 50 generations
- After the completion of 500 generations

In order to find the best features through genetic algorithm, the goal 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.

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

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, 8 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 8 features and the mean loss (the mean square error) is measured. In the second step, 4 features are selected and the mean loss is determined. In the last step, 5 features are selected. Finally according to the mean loss and the amount of features, the best output is selected. It can be concluded that the mean loss does not vary much among 3, 4 and 5 features. Therefore, 5 features have been used as a reference for the selected features.

2.5. Classification

SVM and Bayes Classifiers have been used for classification. 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:

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

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

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

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

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

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]. 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,

$$W \cdot X + b = 0 \quad (16)$$

Where w is a weight vector and we have, $w = \{w_1, w_2 \dots w_n\}$. 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,

$$w_0 + w_1 x_1 + w_2 x_2 = 0 \quad (17)$$

Any point falling above this line is shown as,

$$w_0 + w_1 x_1 + w_2 x_2 > 0 \quad (18)$$

Any point below the line is represented as,

$$w_0 + w_1 x_1 + w_2 x_2 < 0 \quad (19)$$

Finally, the confusion matrix was obtained at the frequencies 14 Hz, 8 Hz and 24 Hz. In total, according to the feature extraction and feature selection, an accuracy of 78.8 was determined for the data.

3. Result

Table1 displays the range of decomposed frequencies. The decomposed signal at various frequencies is presented in Table3.

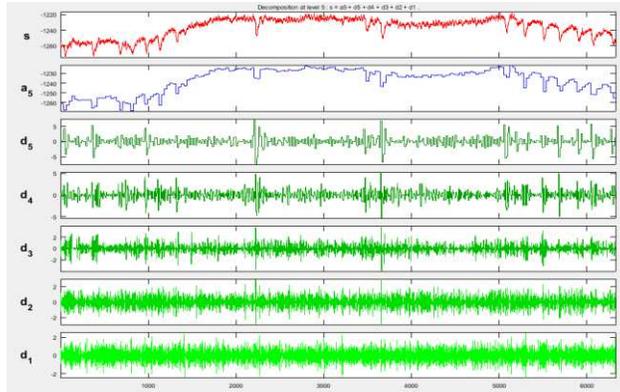


Figure3. The decomposition of the original signal into 5 levels at different frequency ranges using the wavelet transform

Table1. Decomposed signals

Frequency Range	The Decomposed Signal
0-2.7 HZ	A5
2.7-5.4 HZ	D5
5.4-10.8 HZ	D4
10.8-21.7 HZ	D3
21.7-43.4 HZ	D2
43.4-86.8 HZ	D1

In Table2 the mean square error is shown for features 4, 5 and 6. It can be concluded that the mean loss does not vary much among 3, 4 or 5 features. Thus, the 5 features are selected as a reference for the selected features. In Table 3, the accuracy is displayed using a Bayes classifier.

Table2. The number of features and MSE computation in each case

The Number of Features	(MSE)	Feature No.
[6,5,2]	1.3936	3
[3,7,1,4]	1.3927	4
[2,7,5,1,3]	1.3918	5

Table3. The classification accuracy of Bayes classifier in two modes, 1-with feature selection through genetic algorithm and 2-without feature selection

The Bayes Classifier	With Feature Selection	Without Selection	Feature
Classification accuracy	78.8%	80%	

Tables 4, 5 and 6 indicate the confusion matrix at the frequencies 14 Hz, 8 Hz and 24 Hz, and the accuracy and validity of each. Table 7 and 8 compare the accuracy criteria of the proposed method with other studies.

Table4. The confusion matrix at 14 Hz

Confusion Matrix	SSVEP	Normal	
SSVEP	26	5	
Normal	19	462	
			78.8%

Table5. The confusion matrix at 8 Hz

Confusion Matrix	SSVEP	Normal	
SSVEP	35	5	
Normal	10	462	
			78.8%

Table6. The confusion matrix at 24 Hz

Confusion Matrix	SSVEP	Normal	
SSVEP	33	5	
Normal	12	462	
			78.8%

Table7. Comparing feature selection approaches with the proposed method using Bayes classification based on the accuracy criteria

Accuracy Criteria	Classification Approach	Feature Selection Approach
83.32%	Bayes	T-test
73.01%	Bayes	Bhattacharyya
79.13%	Bayes	Decision tree
78.80%	Bayes	The proposed method

Table8. Comparing feature selection approaches based on SWT, CCA and SFT, with the proposed method using SVM classification based on the accuracy criteria

Accuracy Criteria	Classification Approach	Feature Selection Approach
86.9%	SVM	SWT
91.4%	SVM	CCA
79%	SVM	SFA
90.4%	SVM	The proposed method

Figures 4, 5, and 6 display a comparison between classification approaches with and without feature extraction. This comparison is done based on the classification accuracy criterion, program execution time, and initial database size.

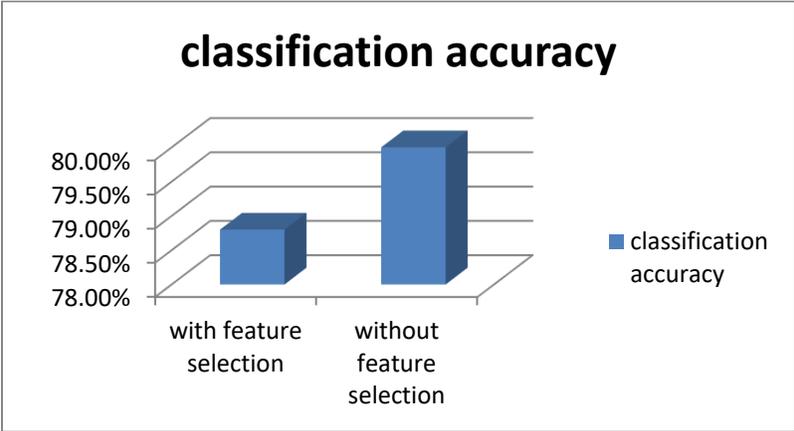


Figure4. A comparison of the accuracy of Bayes classifier, with feature selection through genetic algorithm and without feature selection

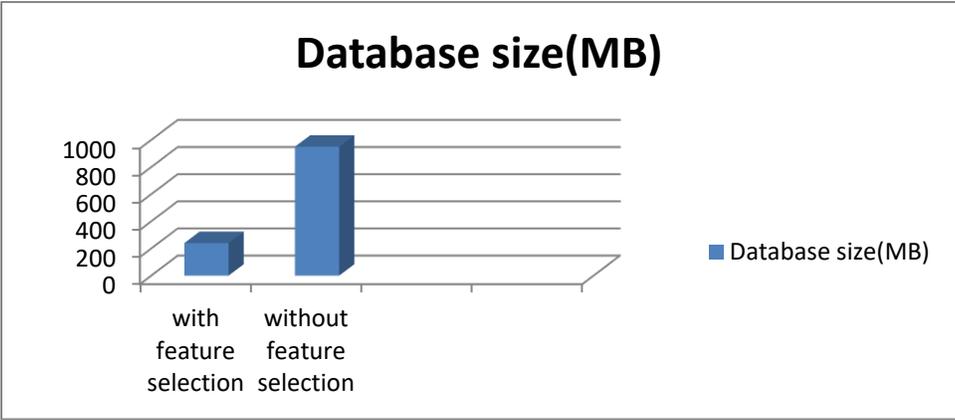


Figure5. A comparison of the Bayes classifier database size, with feature selection through genetic algorithm and without feature selection

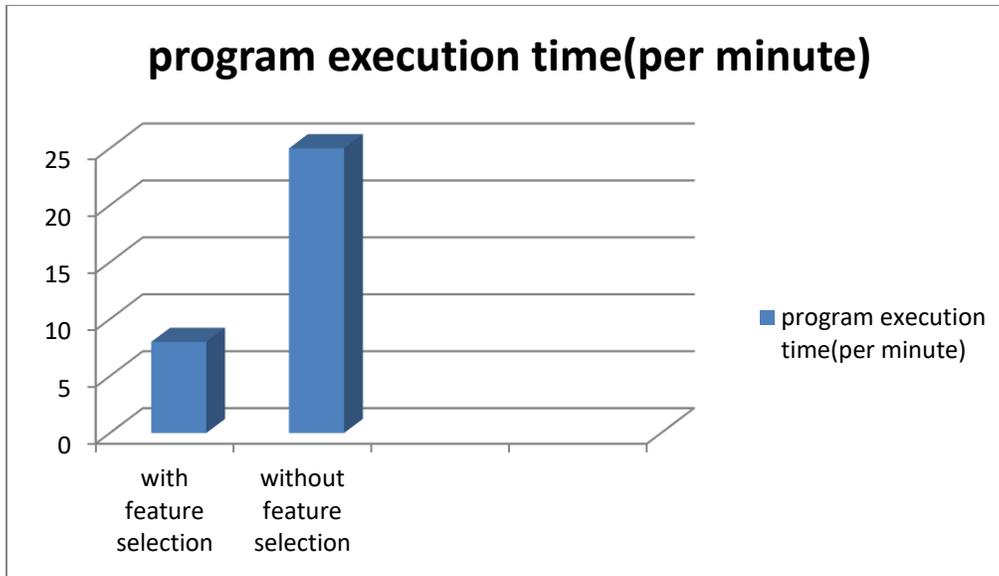


Figure6. A comparison of the program execution time using the Bayes classifier, with feature selection through genetic algorithm and without feature selection

4. Discussion

The main purpose of this paper is the classification of signals depending on whether or not it is SSVEP. The wavelet transform approach and genetic algorithm were applied for the extraction and optimization of features, in order to classify the visual evoked signal (three stimulation frequencies of 8 Hz, 14 Hz and 28 Hz). Heidari and Einalou [11] used a data bank similar to this study, and applied wavelet transform for feature extraction, Bayes classifiers and SVM. However, they used approaches such as T-test, decision tree and Bhattacharyya for feature selection. Using the Bayes classifier, the accuracy criterion in the approaches T-test, Bhattacharyya, and decision tree, and the proposed method was measured to be 83.32%, 73.01%, 79.13% and 78.80%, respectively. As it can be seen, the proposed model is more accurate than the two feature selection models of decision tree and Bhattacharyya, and less accurate than t-test method for feature selection. However, the number of features selected in the proposed method using genetic algorithm 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. In their study, the approaches used for feature extraction and feature selection included SWT, CCA, and SFT, which utilized the methods K-NN and DA for classification. The accuracy of the SVM classifier and the feature selection approaches based on SWT, CCA, SFT and the proposed method in this study were measured to be 86.9%, 91.4%, 79% and 90.4%, respectively. Finally, a comparison of classification is made with and without feature extraction, based on the classification accuracy, the program execution time and the initial database volume. Bayesian classification accuracy was measured to be 80% without dimension reduction, and 78.8% with

dimension reduction. Classification accuracy is reduced with feature selection. One of the most important reasons for reduced classification accuracy is the removal of many features. The classification with all the data obtained from feature extraction using the wavelet transform results in a higher accuracy, than when a part of features are included into the classification algorithm as a database. In other words, the classification would certainly be more accurate, when all the details of data are applied for classification. 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. In the feature selection stage, two thirds of all features are not selected, and the total size of the data in the database is reduced to 240 MB. Due to the fact that the features selected using the goal function are the best available 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 minutes with feature selection, and almost 25 minutes without feature selection. Finally, although applying the naive 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 genetic algorithm more effective for classification.

5. Conclusion

This paper utilized wavelet transform, genetic algorithm, and classification using naïve Bayes and SVM to improve the performance of brain-computer interface (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 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 including the MLP neural network for classification.

Declarations

Availability of data and materials

Available.

Competing interests

The authors declare that they have no conflict of interest.

Funding information

This work did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions

All Authors contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

Compliance with ethical standards

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.

Informed consent was obtained from all individual participants included in the study.

References

[1] Rejer I. Wavelet Transform in Detection of the Subject Specific Frequencies for SSVEP-Based BCI. 2016; In International Multi-Conference on Advanced Computer Systems (pp. 146-155). Springer International Publishing.

[2] M Z, A Maryam, S Sajid, S Awais, M Kamboh. Seizure detection using EEG. 2017; A survey of different techniques. IEEE

[3] Chevallier S, Kalunga E, Barthélemy Q, Yger F. Riemannian classification for SSVEP based BCI. 2018 offline versus online implementations.

[4] Cao, Z, Lin, C T, Lai, K L, Ko, L W, King, J T, Liao, K K, Wang, S J. Extraction of SSVEPs-based inherent fuzzy entropy using a wearable headband EEG in migraine patients. 2019; IEEE Transactions on Fuzzy Systems.

[5] Sözer, A T, Fidan, C B. Novel spatial filter for SSVEP-based BCI. 2018; A generated reference filter approach. Computers in biology and medicine, 96, 98-105.

- [6] Aznan, N K N, Bonner, S, Connolly, J, Al Moubayed, N, Breckon, T. On the Classification of SSVEP-Based Dry-EEG Signals via Convolutional Neural Networks.2018; IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 3726-3731).
- [7]<http://www.bakardjian.com/work/SSVEP/Data Bakardjian.html>
- [8] Bakardjian, H, Tanaka, T, Cichocki, A. Emotional faces boost up steady-state visual responses for brain-computer interface.2011 ;Neuroreport, 22(3), 121-125.
- [9].Poryzala, P, Materka, A. Cluster analysis of CCA coefficients for robust detection of the asynchronous SSVEPs in brain-computer interfaces.2014; Biomedical Signal Processing and Control, 10, 201-208.
- [10] Cortes, C Vapnik, V. Support-vector networks. Machine Learning.1995; 20 (3): 273.
- [11] Heidari H, Einalou Z. SSVEP Extraction Applying Wavelet Transform and Decision Tree With Bays Classification.2017; International Clinical Neuroscience Journal, 4(3), 91-97.
- [12] Anupama H S, Raj V Jain, Revanur Venkatesh, Rupa Mahadevan, N K Cauvery, and G M Lingaraju. "Implementing and Analyzing Different Feature Extraction Techniques Using EEG-Based BCI.2018; In *Recent Findings in Intelligent Computing Techniques*, pp. 377-386. Springer, Singapore.

Figures



Figure 1

The block diagram of the proposed algorithm

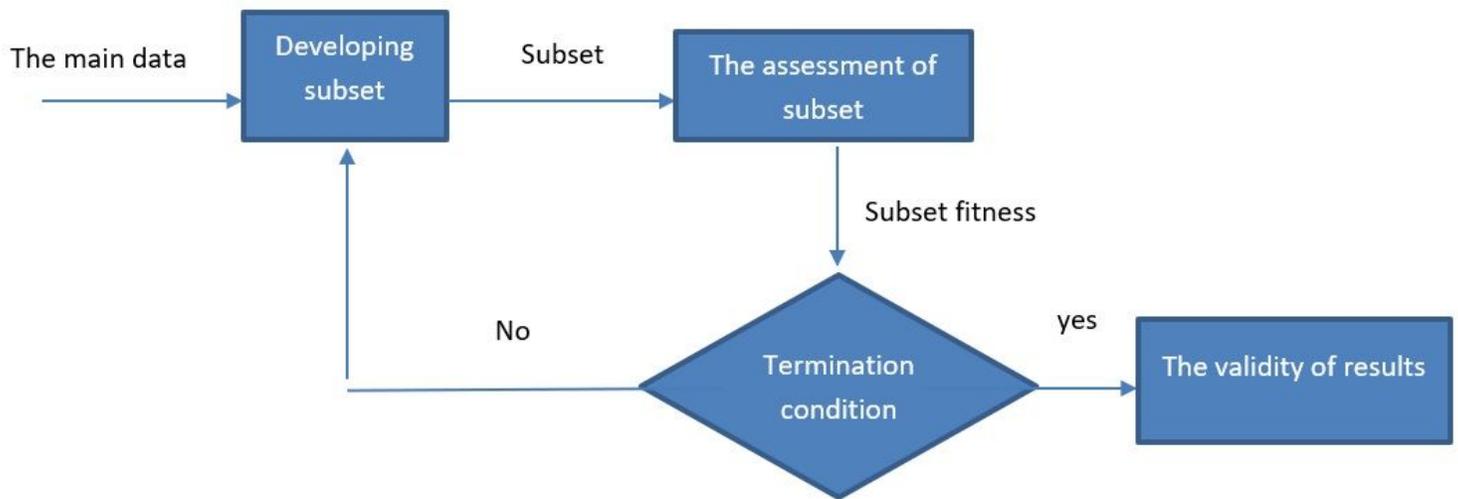


Figure 2

The flowchart for feature selection using the genetic algorithm

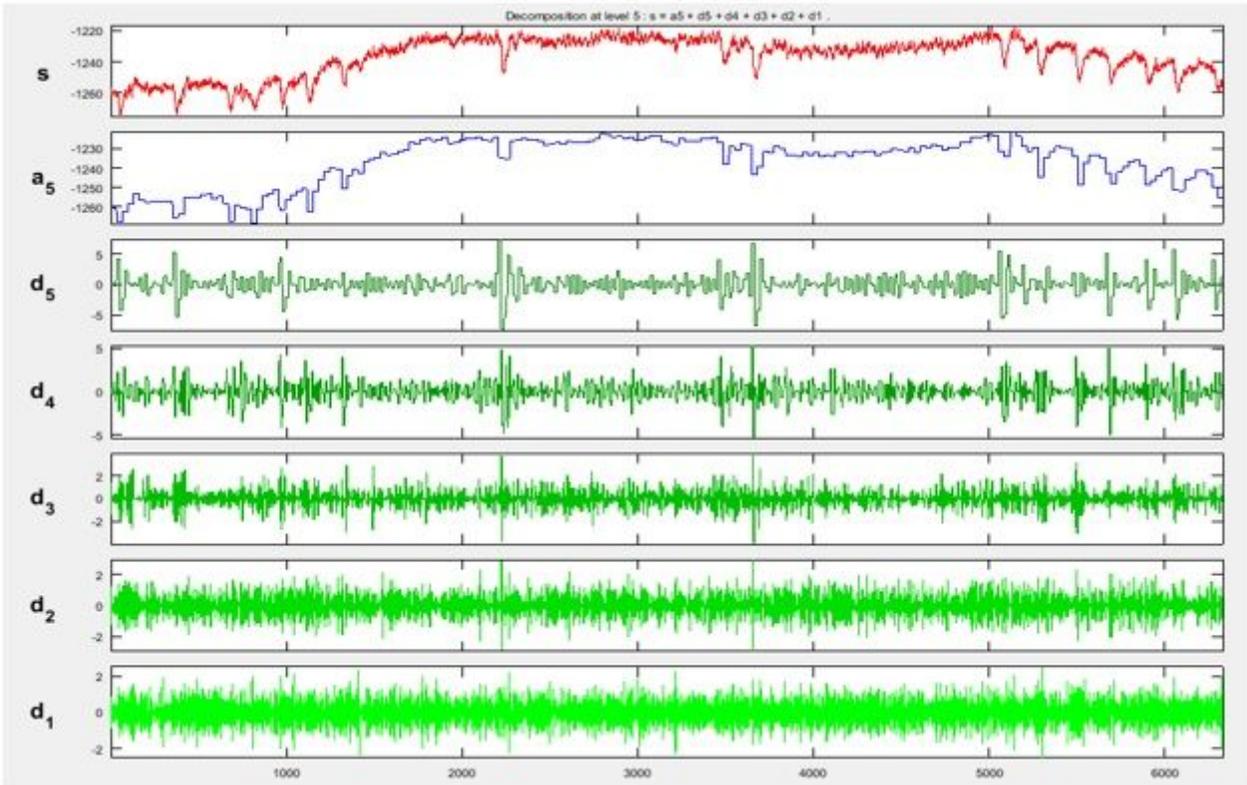


Figure 3

The decomposition of the original signal into 5 levels at different frequency ranges using the wavelet transform

classification accuracy

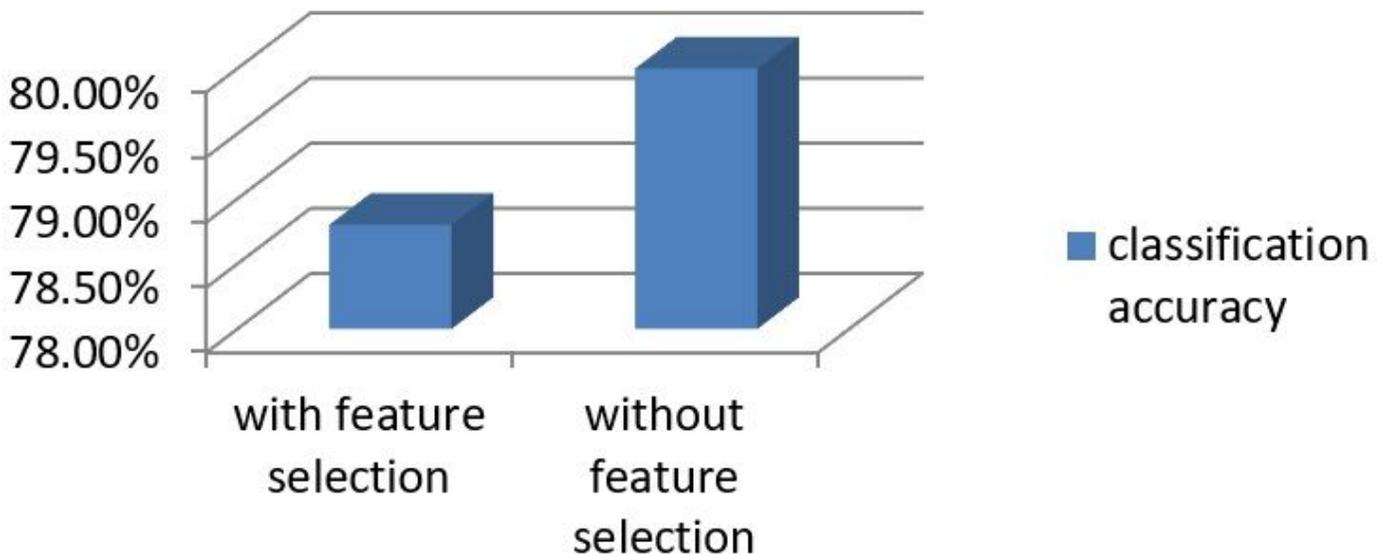


Figure 4

A comparison of the accuracy of Bayes classifier, with feature selection through genetic algorithm and without feature selection

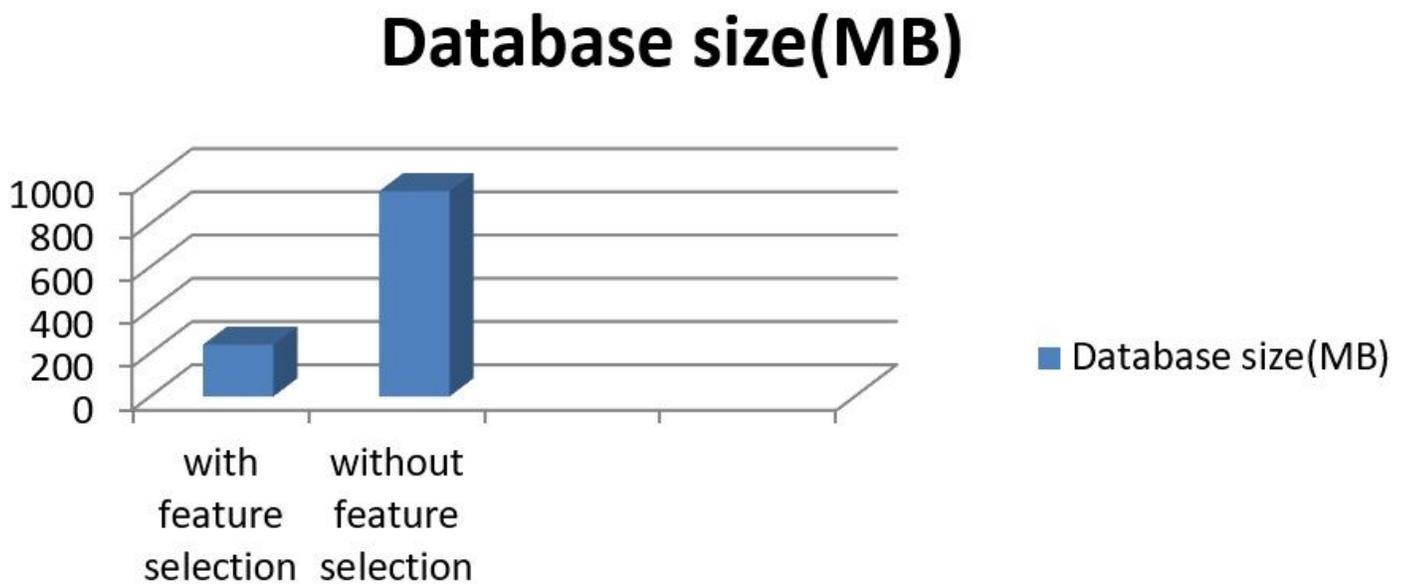


Figure 5

A comparison of the Bayes classifier database size, with feature selection through genetic algorithm and without feature selection

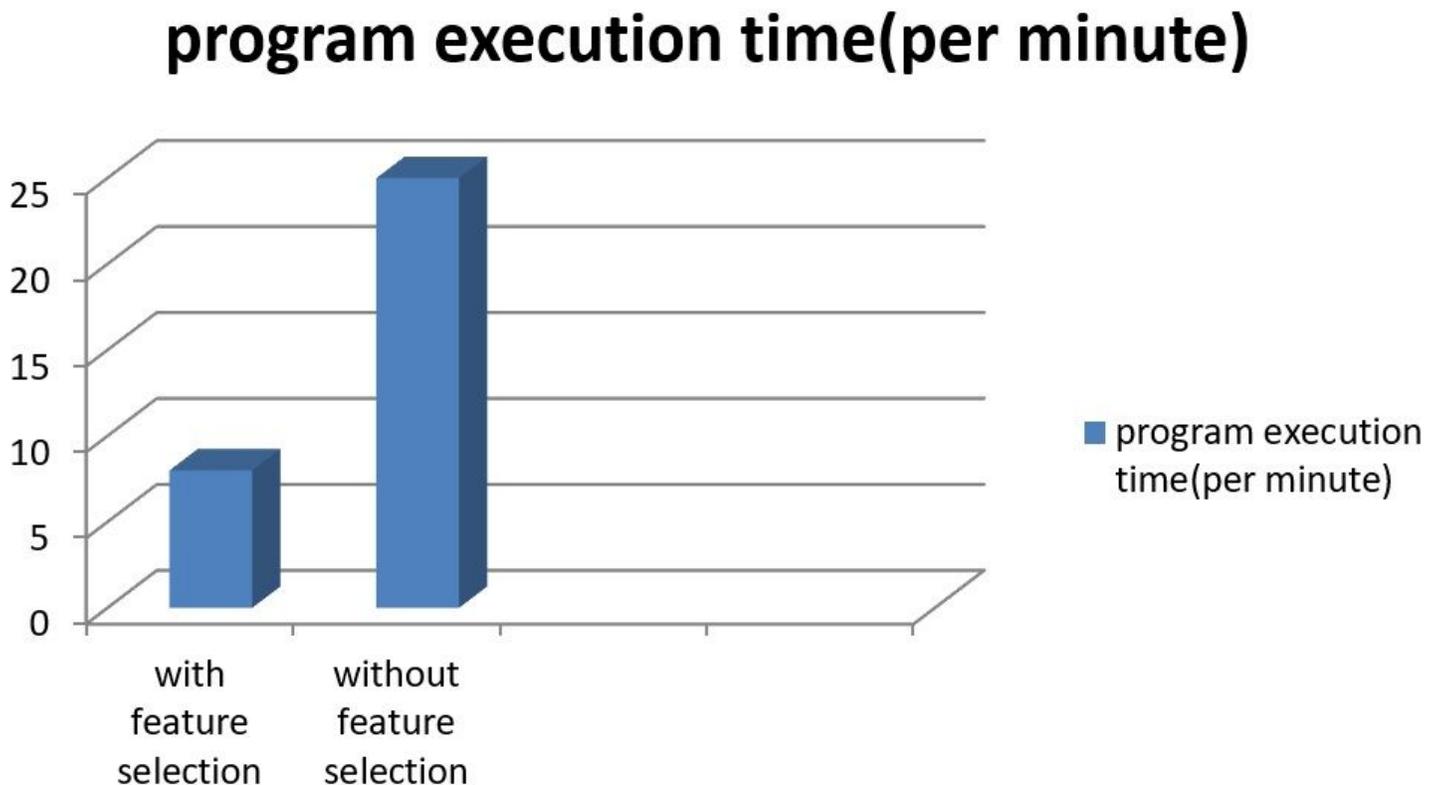


Figure 6

A comparison of the program execution time using the Bayes classifier, with feature selection through genetic algorithm and without feature selection