

RESEARCH

Open Access



# Cursor movement detection in brain-computer-interface systems using the K-means clustering method and LSVM

Leila Mohammadi<sup>1</sup>, Zahra Einalou<sup>1\*</sup>, Hamidreza Hosseinzadeh<sup>1,3</sup> and Mehrdad Dadgostar<sup>2</sup>

\*Correspondence:

z.einalou@srbiau.ac.ir

<sup>1</sup> 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

In this study, we present the detection of the up-downward as well as the right-leftward motion of cursor based on feature extraction. In this algorithm, the K-means clustering method is used to recognize the available hidden patterns in each of the four modes (up, down, left, and right). The identification of these patterns can raise the accuracy of classification. The membership degree of each feature vector in the proposed new patterns is considered as a new feature vector corresponding to the previous feature vector and then, the cursor motion is detected using the linear SVM classifier. Applying the proposed method for data based on the hold-up cross validation causes the accuracy of the classifier in the up-downward and left-rightward movements in each person to increase by 2–10 %.

**Keywords:** Brain-computer-interface system, EEG signal, K-means clustering, Linear support vector machine (LSVM) classifier

## Introduction

The relationship between humans and computer is one of the research fields in which many investments have been made in recent years. Over the past two decades, various human-computer-interfaces have been designed that utilized audition, vision, somatosensation or a combination of them. Serious efforts have been made since the past decade to relate the computer to the environment based on analysis of brain electrical signals. The main purpose of these systems is to help people with corticospinal damages. Despite of having healthy brains, these people are usually not able to move or build a normal and natural relationship with their surrounding environment. Recent advances in neuroscience, physiology, signal analysis, machine learning, and hardware have made it possible to design direct brain-computer communication systems, known as brain-computer-interfaces (BCI), which enable the physically disabled patients to do their affairs without the help of others. In general, a brain-computer interface is a system that allows people with disabilities to operate electrical devices like a computer pointer, a robotic arm, or even a cellular phone with their mind. Regarding its implementation, a BCI system evaluates the specific features of the brain activity, and accordingly translates

them into the signals related to device control. The input signals may be electromagnetic [1], infrared spectroscopy [2–7], magnetic resonance imaging [8, 9], electro-echography [10] and EEG signal [11, 12]. Among these techniques, the EEG signal is used more widely due to its simplicity, low cost, non-invasiveness and ability to register and analyze the data.

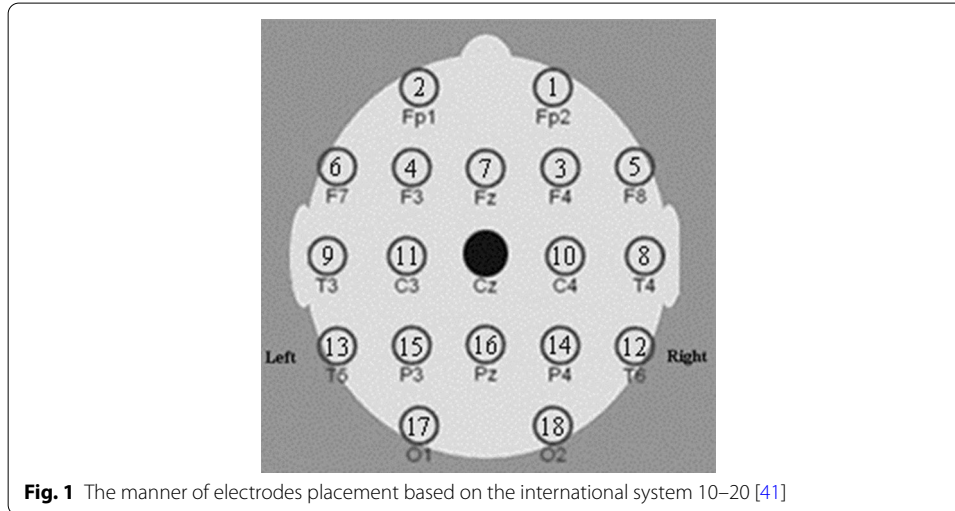
A BCI program specifies appropriate physical or mental tasks and selects electrodes related to these tasks. It extracts features of signal recording by means of these electrodes, and ultimately produces an algorithm with the highest classifiers. Then, the information is transferred to communication and control units through implementing the algorithms to translation. Developing a quick and precise classification method is a key subject for communication and control over the EEG signal. Computer cursor control is one of the most useful applications of BCI. The first studies on the cursor control based on EEG signal were focused on one-dimensional control of cursor [13, 14]. After obtaining satisfactory results, the attention of researchers was drawn to the multi-dimensional cursor control with the aim of increasing the relationship between the user and machine. Based on the studies, the multi-dimensional cursor can be controlled by implementing different brain signals, for example the P300 potential [15, 16], synchronous and asynchronous signals [17] and evoked potentials [18, 19]. They showed well that the input signals of an EEG-based brain-computer-interface system have commonly weak, non-constant and mind tasks-related noises with various artifacts such as external electromagnetic waves and electromyogram and electro-echogram waves. All of these defects were considered by researchers to improve the key features of BCI system function including the classifier velocity and precision. In addition to improving the functional features, the operational and realistic construction of the EEG-based BCI systems was another major challenge from the viewpoint of researchers. It highly relies on using a training process, which is based on a small number of training signals, classifying channel algorithms and features which can represent the studies tasks better. The proper input design, features extraction and appropriate classification of a BCI system has been undertaken over the past 20 years. Various features of the EEG signal serve as inputs for a BCI, such as the Mu band (8–12 HZ) and beta band (18–25 HZ) [20], event-related potential such as P300 [21, 22], steady-state visual-evoked response or the surface slow cortical potential (SPCs) [23, 24], and several feature extraction methods for motor imagery data as input signals, including AR parameter [25] estimation in a specific frequency band [20, 26], domain of slow cortical potential [27], common spatial pattern, event-related synchronization (ERD / ERS), wavelet correlation coefficient [28], and spatial pattern spectrum [29]. Many of the classifier algorithms, such as K-nearest Neighborhood (KNN), Linear Discriminate Analysis (LDA) [30], hidden Markov model [31], Neural Networks [32, 33], Multilayer Perceptron, and Mahalanobis Distance (MDA) [34, 35] were tested for BCI programs. The feature extraction and classifier algorithm has been investigated in BCI applications [36]. Since 2001, several BCI competitions have been organized for improved target assessment and comparison of methods for EEG data derived from mental tasks and also addressing of key issues in BCI research, and significant results have been reported for BCI data. Currently, BCI technology lacks desirable speed and precision, and these are two important variables affecting current attempts to address issues related to the creation of real BCI systems in the future. To

this aim, the classifier can discriminate between the EEG signals recorded in different sessions and days with the mental tasks belong to the BCI system. It should also be borne in mind that BCI systems have the potential to help people of all ages with severe mobility impairments. Although some studies on the BCI performance in older people didn't yield good results due to poor rate of data transfer, In the BCI competitions, [25] obtained a classifier precision of 88.7% with a linear classifier using gamma band power (Channel 4) and SCPs (Channel 1) [36]. Then, by selecting a neural network as classifier, the classifier accuracy was enhanced to 91.47%. In addition, [21, 30] by extracting features of 6 channels and using wavelet transform algorithm, they obtained a classifier accuracy of 90.8% by the neural network. [30, 37] used principal component analysis (PCA) in 6 channels and their spectral properties, reporting an accuracy of 90.44%. Trejo et al. [38] showed that people of various ages can benefit from this technology. It has also been observed in recent studies that the combination of EEG and EMG signals in BCI systems can be used for several activities in cases of mobility impairments. They believed that users would be able to understand how to interact with such interactive systems [39]. Although these methods have desirable classifier accuracy, according to complete training data, they were capable of using more than one electrode and complex vectors. Therefore, they are complex, low portable and in need of extensive training and computational time. Their results not only revealed the accuracy of diagnosis, but also significantly reduced the calculation time for training due to the reduced feature space. However, the drawback of these methods is that they used training accuracy as a criterion for the evaluation of different combinations of features. In fact, different subsets of features were trained for creating the optimal subset with the best accuracy for diagnosis, although it was ineffective. The objective of this study is to classify the EEG signal in the BCI systems using data mining technique. As we know, data mining plays a significant role in extraction of hidden information or patterns and relationships in a large amount of data. Feature extraction and selection are very important in data mining. In addition, an increase occurs in the computational time when the number of descriptive features increases. To deal with the large number of features in the present study, a method for recognizing the pattern and extracting the essential information for detection of the cursor movement has been studied. The present study aimed to find a method for the input motion efficiently and efficiently by implementing the data mining technique. This is done using the hybrid K-means clustering method and the linear support vector machine (LSVM) classifier [40] for detecting the upward, downward, leftward and rightward motion.

This study proposes a simple and straightforward algorithmic method to improve the classification of BCI data. A smaller number of features reduces the computational times. Although the speed comparison with other studies cannot be made objectively since they did not report any results on the training times or the testing times. The cursor movements pattern recognition suggests a new way on how to determine features for the the up- and downward and left- and rightward movements. The K-LSVM 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. Because the noisy information has been eliminated.

**Table 1** The total number of considered trials [41]

Subject	Number of raining trials	Number of testing trials
A	140	152
B	148	152
C	148	152


**Fig. 1** The manner of electrodes placement based on the international system 10–20 [41]

The present paper's agenda is as follows: Sect. "[Materials and methods](#)" describes the collection of data used for classification. Besides, it presents the K-means algorithm was proposed as an unsupervised algorithm to extract motion features and prevent from repetiting the training in different subsets. All of the data related to feature can be stored in a dense method for one-time training instead of performing multiple training pilots in different feature subsets since the K-means algorithm clusters the principal feature space with unsupervised learning. Therefore, a membership function was used to obtain optimal results from the K-means algorithm for training. Further, Sect. "[Results](#)" provides a summary of a combination of K-means and LSVM classifier for motion detection with high accuracy has been proposed in this study. Lastly, Sect. "[Discussion](#)" and "[Conclusion](#)" give critical insights into the performed interpretation and accuracy, and address some limitations and possibilities of the presented method.

## Materials and methods

### Data acquisitions

The research was performed by online collection of EEG signals from three normal subjects at two different sessions with a delay of one week in 2014 at the Karadeniz Technical University of Turkey [41]. The sampling frequency was 256 Hz and a notch filter was used to eliminate noise. The device consisted of 18 electrodes based on the international system 10–20 shown in Fig. 1, in which the two electrodes FP1 and FP2 were related to EOG artifacts, the two electrodes O1 and O2 were associated with EMG artifacts and the remaining 14 electrodes were used in the proposed method for extracting features.

Each person sat on a chair facing a  $19 \times 19$  screen and was asked to remain motionless during the experiment. Each trial lasted at least 10 s and was performed with a delay of 2 s. During the trial, the target appeared in one of the four possible modes on the display and the person was asked to imagine the appeared mode for 8 s; each trial ended with a beep sound. The data recording was repeated for each person one week later. Each data recording contained 4 runs and each run contained 40 trials (10 trials for each class or movement). Therefore, each data recording contained 160 trials ( $40 \times 4 = 160$ ). The length of data for each trial was 2048 samples ( $8 \times 256 = 2048$ ). The data were generally divided into two categories of train and test data as shown in Table 1. The number of considered trials varied from person to person.

### Feature extraction

The four following feature extraction parameters [41] were implemented to represent the experimental signals as shown in Table 2.

#### Continuous wavelet transform

The continuous wavelet transform (CWT) is a strong signal processing tool used as a feature in many BCI data analyses [32, 42]. CWT is a convolution of the signal  $X(t)$  with the wavelet function  $\psi_{\tau,s}(t)$  so that:

$$CWT_x^\psi(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t) \psi^*\left(\frac{t - \tau}{s}\right) dt \quad (1)$$

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t - \tau}{s}\right) \quad (2)$$

where,  $\psi(t)$  is a continuous function in time and frequency, known as the mother wavelet.  $s$  and  $\tau$  are the scaling and shifting parameters, respectively.

$$\int_{-\infty}^{\infty} \psi_{\tau,s}(t) dt = 0 \quad (3)$$

The mother wavelet applied in this method is the Morlet wavelet transform. The analysis of the statistical features in the train dataset by applying the Morlet wavelet transfer showed that the mean and standard deviation of the continuous wavelet coefficients can be used as feature. The mean and standard deviation as the first feature are calculated in accordance with the following relationships.

**Table 2** Extracted features

Feature extraction method No.	Feature extraction method
1	Mean and SD of continuous wavelet coefficients in alpha, beta & theta bands
2	Autoregressive Model
3	Skewness
4	Average derivative

$$CWT C^{Avr} = \frac{\sum |CWT C|}{L_{CWT C}} \quad (4)$$

$$CWT C^{std} = \sqrt{\frac{\sum (|CWT C| - CWT C^{Avr})^2}{L_{CWT C} - 1}} \quad (5)$$

where  $L_{CWT C}$  is the length of CWT. The mean and standard deviation of the continuous wavelet coefficients for all channels are calculated as the first feature based on the following three parameters.

1. Based on channel selection; the proposed method includes 14 channels.
2. Based on recording duration, which has been set to 8 s in the proposed method.
3. Based on EEG signal bands consisting of three bands: theta (4–8 Hz), alpha (8–13) Hz and beta (13–20 Hz).

#### **Autoregressive model**

A raw signal  $x(t)$ , as an autoregressive (AR) model, is expressed as follows:

$$x(n) = A_1 x(n-1) + A_2 x(n-2) + \dots + A_p x(n-p) + e(n) \quad (6)$$

where,  $A_1, A_2, \dots, A_p$  are the parameters of the AR model,  $p$  is the order of the model and  $n$  is an integer representing the discrete time points of the signal.  $e(n)$  is the white noise with zero variance and mean.

The suggested AR model is calculated as a second feature based on the following three parameters:

1. Based on channel selection; the proposed method includes 14 channels.
2. Based on recording duration, which has been set to 8 s in the proposed method.
3. The order of the model; the model order in the proposed method is  $p = 8$ .

#### **Skewness and average derivative**

Skewness and average derivative as third and fourth features are calculated based on the following relationships and parameters.

$$S = \frac{\frac{1}{L} \sum_{i=1}^L (x_i - \bar{x})^3}{\left( \sqrt{\frac{1}{L} \sum_{i=1}^L (x_i - \bar{x})^2} \right)^3} \quad (7)$$

$$AD = \frac{\sum_{n=1}^{L-1} x(n-1) - x(n)}{L-1} \quad (8)$$

1. Based on channel selection; the proposed method includes 14 channels.
2. Based on recording duration, which has been set to 8 s in this method and the order of the model, which has been considered  $p = 8$ .

### Description of the proposed method

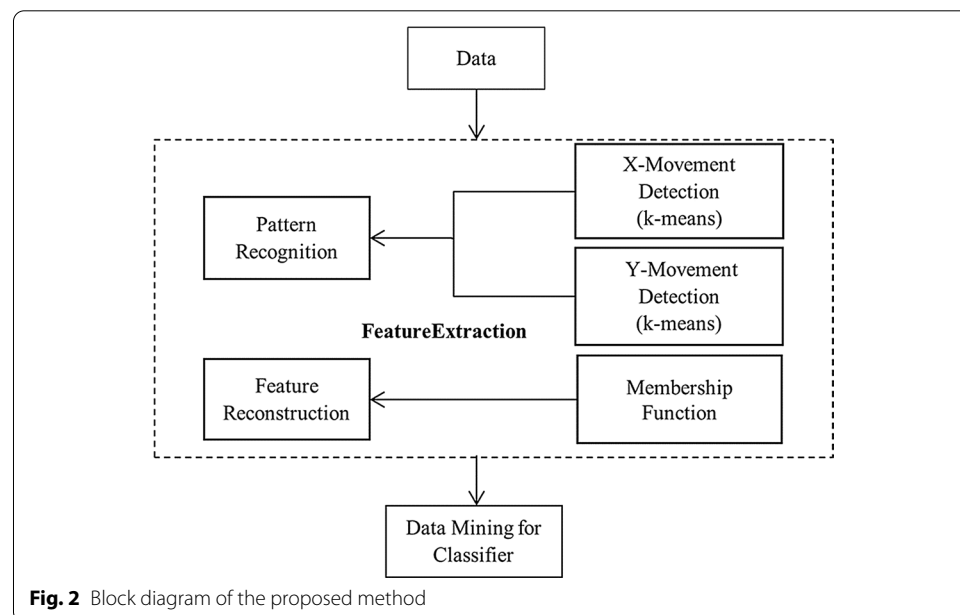
Figure 2 shows the block diagram of the proposed method, which aims to detect cursor movement in up- and downward and left- and rightward directions. The proposed method is a hybrid k-means clustering algorithm and LSVM classifier.

The feature extraction is used for patterns identification and the K-means clustering algorithm is used for clustering the movements based on the similarity of the features of the up- down, left-rightward movements considering the membership degree of data. To reduce the big dimensions of the feature space, the algorithm extracts some movement patterns before training of data for proper classification. This method considers a new pattern of k-means algorithm as a new feature of movements. The new features are different from the previous features that only included the pattern of movement features. Therefore, the dimension of feature space decreases and the new features show the similarity between the movements and patterns. Further, the boundary between up-downward and left-rightward movements are estimated by the classifier based on these new features. As the dimensions of the feature space decrease, the dataset is reconstructed with the new features. The optimal channel for each person has been calculated. In this method, once the up- downward mode and then the left-rightward mode are classified.

A K-means algorithm can play a role in applying an optimization solution in order to reduce the distance between cluster centers and cluster members, which is calculated according to the relation (9).

$$\min_{\mu_1, \mu_2, \dots, \mu_k} \sum_{k=1}^K \sum_{i \in S_k} \|X^i - \mu_k\|^2 \quad (9)$$

In the relation (9),  $k$  is the cluster index,  $S_k$  represents the  $k$ th cluster of the dataset,  $\mu_k$  indicated the center point of the cluster and  $k$  is considered as the total number of clusters, which can contribute to the normalization of the points of data for removing the



impact of the scale used for different features. The k-means algorithm adapts the location of the center repeatedly for decreasing the Euclidean distance in order to choose the cluster centers. In this regard, two methods are available for estimating the number of clusters:

1. Based on the experiences of expert.
2. Based on the similarity criterion.

The second method has been used for the similarity of movement patterns. The similarity criteria are calculated according to relations (10) and (11).

$$d_{avg} = \frac{\sum_{k=1}^K \sum_{i \in S_k} \sqrt{\sum_{j=1}^F (X_j^i - X_j^{\mu_k})^2}}{N} \quad (10)$$

$$d_{min} = \min \left[ \sqrt{\sum_{j=1}^F (X_j^{\mu_{k_1}} - X_j^{\mu_{k_2}})^2} \right] \quad \forall k_1 \neq k_2 \quad (11)$$

where  $d_{avg}$  shows the average distance from each member  $i$  to the center  $\mu_k$  in the same cluster  $S_k$ ,  $d_{min}$  indicates the smallest distance between the two clusters,  $X_j^i$  is considered as the  $j^{th}$  input element of the member  $i$ ,  $X_j^{\mu_k}$  is regarded as the  $j^{th}$  input element of the center  $\mu_k$ . Accordingly,  $N$  and  $F$  indicate the total number of data points and the dimension of the input vector, respectively. The number of optimal clusters ( $K^*$ ) is calculated by the function  $validity(\theta)$  as indicated in relation (12).

$$K^* = arg_k \min \theta = arg_k \min \frac{d_{avg}}{d_{min}} \quad (12)$$

where  $(\theta)$  represents the validity ratio for assessing the various number of clusters, upon which an acceptable number of clusters is found to recognize the hidden patterns of movement. Based on the results, the average distance from each member to the center of that cluster ( $d_{avg}$ ) decreases by computing the minimum ratio of validity. However, an increase occurs in the minimum distance between two cluster centers ( $d_{min}$ ). In other words, the determined clusters are concentrated in their own cluster and separated from others. No multiple pattern is shown with a large number of members when the value of  $k$  is close to the number of data points. In fact, the cluster density is calculated by the k-means algorithm, which is small and quite similar to the original data. Therefore, the least range of  $k$  in the proposed method ( $2 < K < 14$ ) is considered for finding the number of local minima of clusters to show the accumulated patterns of each type of motion. In this case, the magnitude of the scale of the main dataset is reduced by the available labels for each person.

After recognizing the movement patterns, some movements fall into other clusters with symbolic labels. The similarity between the untested movement and symbolic movement can pave the way for recognition. Thus, a membership function is used to assess the degree of similarity of the main data points and provided motion labels, which indicates the points' fuzzy membership of the recognized patterns. The membership function is as follows:



$$f_c(X_j^i) = \begin{cases} 1 - \frac{|X_j^{\mu_c} - X_j^i|}{\max |X_j^{\mu_c} - X_j^n|} & \text{if } \min(X_j^n) \leq X_j^i \leq \max(X_j^n), \quad \forall n \in S_c \\ 0, & \text{otherwise;} \end{cases} \quad (13)$$

$$p_c = \frac{1}{F} \sum_{j=1}^F f_c(X_j^i), \quad 1 \leq c \leq K^m + K^b \quad (14)$$

In relation (14),  $c$  is the index for the new pattern,  $X_j^i$  represents the  $j^{th}$  feature of the main input  $i$  and  $X_j^{\mu_c}$  denotes the  $j^{th}$  feature of the center  $\mu_c$  in the cluster  $S_c$ , which is obtained from the previous  $K$ -means algorithm.  $K^m$  and  $K^b$  are the number of clusters related to the class data.

Using this membership function, the similarity between movements and patterns is assessed from the viewpoint of how well-suited a movement is for pattern recognition.

## Results

The results have been obtained from three subjects by determining the optimal channel for each of them. The proposed method is the hybrid  $K$ -means clustering and LSVM classifier.

The  $K$ -means clustering has been used for recognizing the hidden patterns and the LSVM classifier for separating the up- downward and the left-rightward movements. The optimal channel for each person has been calculated according to Table 3. In this method, once the up- downward mode and then the left- rightward mode are classified. Based on the results, the LSVM classifier in the proposed method has effective results, as the applied hybrid method has led to an increase of 2–10% in the accuracy of the classifier.

Sensitivity and specificity are two statistical measures for evaluating the outcome of a binary classification test. When data can be divided into two positive and negative groups, the accuracy of the results of a test that divides the data into these two groups can be measured and described using the indices of sensitivity and specificity. Sensitivity (the true positive rate) is the proportion of positives that are correctly identified as positive by the test. Specificity (the true negative rate) is the proportion of negatives that are correctly identified as negative by the test. Mathematically, sensitivity is expressed as the ratio of true positives to combined true and false positives.

**Table 3** The results of LSVM classifier

Subject	Optimal Channel	LSVM classifier accuracy in up- downward movement hybrid	LSVM classifier accuracy in up- downward movement feature	LSVM classifier accuracy in left- rightward movement hybrid	LSVM classifier accuracy in left- rightward movement feature
1	12	69 ± 3.8	67.1 ± 4.9	63.1 ± 2.8	62.1 ± 3.6
2	14	73.3 ± 5.4	64.4 ± 4.3	68.6 ± 3.6	60.9 ± 3.2
3	12	66.3 ± 4.3	62.1 ± 3.3	74.4 ± 5.8	65.5 ± 2.8

$$\text{Sensitivity} = \frac{\text{Number of true positives}}{\text{Number of true positives} + \text{Number of false negatives}} \quad (15)$$

$$\text{Specificity} = \frac{\text{Number of true negatives}}{\text{Number of true negatives} + \text{Number of false positives}} \quad (16)$$

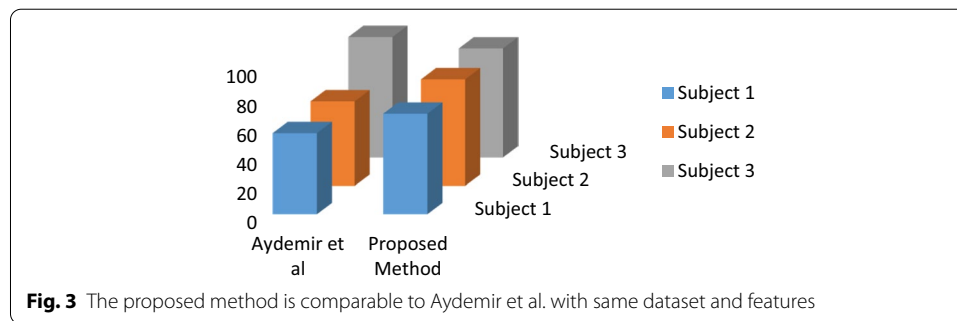
Sensitivity and specificity for up- downward and left- rightward movements in the proposed method are calculated based on the relations (15) and (16), as seen in Table 4.

## Discussion

In this study, K-LSVM method was proposed based on known special patterns. This can be comparable to conventional data mining methods in motion detection. In the feature extraction step, conventional data mining methods are not used; instead, clustering is utilized to extract symbolic motions with the aim of representing the cluster of motions. The relationship between input motion and symbolic motion as the pattern membership is evaluated to predict new motions. The proposed method is comparable to Aydemir et al. [41] with same dataset and features. They classified the EEG signal based on the decision tree structure during the cursor movement. They could separate the up- and downward and left- and rightward movements by achieving the accuracy of 55.92%, 57.92 and 80.24% for the first, second and third person, respectively. As the feature applied in their approach is according to Table 2 and the three classifiers LDA, KNN and SVM have been used for separating the movements, the applied decision tree structure is different in each person and this is one of the weak points of this approach, while the hybrid algorithm applied in the proposed method in the present study, considering the use of the same features, is similar for all subjects, and hence, the proposed method has caused an increase of 2–10% in the classifier accuracy. There are no researches on hybrid clustering and classifiers in BCI systems and recent studies using this approach have been concentrated on breast cancer diagnosis. Bennett et al. [43] tested this algorithm on a set of breast cancer data for clustering the benign and malignant tumors and suggested the detection of data patterns as one of the applications of clustering algorithm. Since the method used by the researchers did not focus on the manner of determining the quantity of clusters in datasets, the test clusters were randomly divided into two benign and malignant categories. The number of patterns should be determined to evaluate the hidden patterns of breast cancer. There is also a need for comprehensive exploration of the connection between the supervised learning and learning without

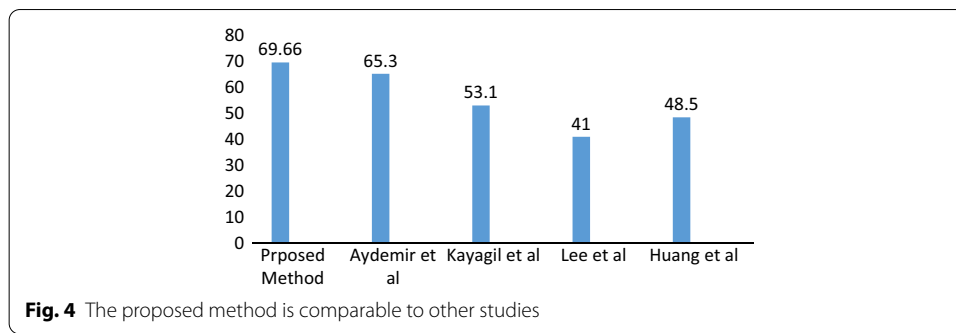
**Table 4** Sensitivity and specificity amounts

Subject	Movement	Sensitivity	Specificity
1	Up- downward	65 %	72 %
	Left- rightward	63 %	62 %
2	Up- downward	74 %	71 %
	Left- rightward	69 %	67 %
3	Up- downward	67 %	65 %
	Left- rightward	71 %	79 %



supervision to improve the breast cancer diagnosis. In order to reduce the dimensions of the dataset, some researchers began to use the technique of clustering algorithms and classifier model in the field of machine learning. Dhillon et al. [44] applied the hybrid clustering algorithm for a set of similar words to achieve a quick and accurate learning in text classification. They used C-means clustering for educational samples and SVM for Text identification. Radha et al. [45] used a breast cancer database for their researches in 2014. Their clustering was based on motor size. They applied the K-means algorithm to this dataset. Nine features were used in this method. In the first phase of the K-means clustering algorithm, the Manhattan distance criterion was used to measure the distance between nodes and the selection of the center of the primary cluster was done randomly. After implementing the algorithm at  $k=2$  on 286 data points, the average probability of the correct classification was 66.99%. In the second phase of K-means clustering, the Euclidean distance criterion was used to measure the distance between nodes and the selection of the center of the primary cluster was done randomly. After implementing the algorithm at  $k=2$  on 286 data points, the average probability of the correct classification was 74.98%. In the third phase of K-means clustering, the cluster centers were graded in order to raise the clustering accuracy and the Pearson's correlation criterion was used to measure the distance between nodes and the selection of the centers of the primary cluster was achieved through co integration. After implementing the algorithm at  $k=2$  on 286 data points, the same results were obtained from all clustering. The average probability of the correct classification was 100%. Based on the above-mentioned issues, a large number of methods have been utilized for diagnosing breast cancer based on implementing clustering and classifier. Having a large scale of data is considered as one of the main disadvantages of conventional and traditional methods. Although the time calculations in different training are decreased by applying explorations in feature selection, no comprehensive search is necessary for selecting the feature for the clustering algorithm, especially the K-means algorithm, which results in better understanding of the number of training samples without any help for extracting and selecting the related feature.

The proposed method is comparable to Aydemir et al. [41] with same dataset and features. See Fig. 3. They classified the EEG signal based on the decision tree structure during the cursor movement. The applied decision tree structure is different in each person and this is one of the weak points of this approach, while the hybrid algorithm applied in the proposed method in the present study, considering the use of the same features,



is similar for all subjects, and hence, the proposed method has caused an increase of 2–10% in the classifier accuracy.

Deep Neural Networks (DNN), however, have shown promising results in BCI research. Deep neural networks are still lagging behind in performance due to the unavailability of large training datasets [46–48]. Main disadvantage of this method was its time consuming length.

We compare our proposed method with other studies as seen as Fig. 4. Lee et al. [49] computed the Common Spatial Pattern (CSP) which is the famous feature extraction method in BCI researches. Kayagil et al. [50] and Huang et al. [51] were the most closes ones to this method because of experimental properties. Aydemir et al. have same data-set and features.

In the present study, a combination of K-means clustering algorithm and classifier were implemented to compress the feature extraction space in order to decrease the computational costs. The proposed K-LSVM method in this study emphasizes on discovering the knowledge framework, searching for more knowledge of motion detection through implementing data mining methods. In addition, descriptive features and data will be collected in the future. However, feature extraction and selection are considered as a challenging issue for researchers, although large data with lost values can cause another issue which should be taken into consideration with respect to computational time.

## Conclusions

In the present study, K-means was used as clustering algorithm in order to recognize and obtain motion patterns. These patterns are reconstructed as a new feature for the training phase. Pre-processing steps such as feature extraction and selection can play a role in creating a highly effective feature for the purpose of machine learning algorithms which is used for training the classifier. Regarding K-LSVM results, the calculation time can be significantly decreased without measuring the detection accuracy. However, using the k-means algorithm failed to decrease the size of training samples by filtering similar samples in the data, which can be regarded as a potential method for reducing the size of the training set. Based on the results, the LSVM classifier in the proposed method has effective results, as the applied hybrid method has led to an increase of 2–10% in the accuracy of the classifier. The dataset used in this study represents a complete dataset

with lost values; however, applying this method to a broad set of low-size data opens a venue of research in the future. So far, no general rule has been adopted for evaluating the number of patterns for detecting motion. In the future, a method for evaluating the number of symbolic motions can be developed, which may decrease the calculation time for extracting feature and allows experts to comprehend the motions better based on the patterns derived from the methods implemented for data mining methods.

#### Acknowledgements

Not applicable.

#### 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.

#### Funding

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

#### Availability of data and materials

Available.

#### Declarations

##### Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Karadeniz Technical university and its later amendments or comparable ethical standards.

##### Informed consent

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

##### Competing interests

The authors declare that they have no conflict of interest.

##### Author details

<sup>1</sup>Department of Biomedical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran. <sup>2</sup>Department of Biomedical Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran. <sup>3</sup>Department of Electrical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran.

Received: 26 December 2020 Accepted: 3 May 2021

Published online: 20 May 2021

#### References

- Mellinger J, Schalk G, Braun C, Preissl H, Rosenstiel W, Birbaumer N, Kübler A. An MEG-based brain-computer interface (BCI). *Neuroimage*. 2007;36(3):581–93.
- Dadgostar M, Setarehdan SK, Akin A. 2013, December. Detection of motion artifacts in fNIRS via the continuous wavelet transform. In 2013 20th Iranian Conference on Biomedical Engineering (ICBME) (pp. 243–246). IEEE.
- 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.
- Dadgostar M, Setarehdan SK, Shahzadi S, Akin A. Functional connectivity of the PFC via partial correlation. *Optik Int J Light Electron Optics*. 2016;127(11):4748–54.
- Einalou Z, Maghooli K, Setarehdan SK, Akin A. Functional near infrared spectroscopy for functional connectivity during Stroop test via mutual information. *Adv Biores*. 2015;6(1):62–7.
- Einalou Z, Maghooli K, Setarehdan SK, Akin A. Graph theoretical approach to functional connectivity in prefrontal cortex via fNIRS. *Neurophotonics*. 2017;4(4):041407.
- Dadgostar M, Setarehdan SK, Shahzadi S, Akin A. Classification of Schizophrenia using SVM VIA fNIRS. *Biomed Eng Appl Basis Commun*. 2018;30(02):1850008.
- Weiskopf N, Mathiak K, Bock SW, Scharnowski F, Veit R, Grodd W, Birbaumer N. Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE Transact Biomed Eng*. 2004;51(6):966–70.
- Faskhodi MM, Einalou Z, Dadgostar M (2018). Diagnosis of Alzheimer's disease using resting-state fMRI and graph theory. *Technology and Health Care*, 26(6), 921–931.
- Deiber MP, Ibanez V, Honda M, Sadato N, Raman R, Hallett M. (1998). Cerebral processes related to visuomotor imagery and generation of simple finger movements studied with positron emission tomography. *Neuroimage*;7(2):73–85.
- Aydemir O, Kayikcioglu T. Wavelet transform based classification of invasive brain computer interface data. *Radioengineering*. 2011;20(1):31–8.
- Nicolas-Alonso LF, Gomez-Gil J. Brain computer interfaces, a review. *Sensors*. 2012;12(2):1211–79.
- Aydemir O, Kayikcioglu T. Comparing common machine learning classifiers in low-dimensional feature vectors for brain computer interface applications. *Int J Innovative Comput Information Control*. 2013;9(3):1145–57.
- Kayikcioglu T, Aydemir O. A polynomial fitting and k-NN based approach for improving classification of motor imagery BCI data. *Pattern Recogn Lett*. 2010;31(11):1207–15.

14. Piccione F, Giorgi F, Tonin P, Priftis K, Giove S, Silvoni S, Palmas G, Beverina F. P300-based brain computer interface: reliability and performance in healthy and paralysed participants. *Clinical neurophysiology*. 2006;117(3):531–7.
15. Gao W, Guan JA, Gao J, Zhou D. Multi-ganglion ANN based feature learning with application to P300-BCI signal classification. *Biomed Signal Process Control*. 2015;18:127–37.
16. Wang M, Daly I, Allison BZ, Jin J, Zhang Y, Chen L, Wang X. A new hybrid BCI paradigm based on P300 and SSVEP. *J Neurosci Methods*. 2015;244:16–25.
17. Royer AS, Doud AJ, Rose ML, He B. EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies. *IEEE Transact Neural Syst Rehab Eng*. 2010;18(6):581–9.
18. Volosyak I, Gembler F, Stawicki P. Age-related differences in SSVEP-based BCI performance. *Neurocomputing*. 2017;250:57–64.
19. Heidari H, Einalou Z. SSVEP extraction applying wavelet transform and decision tree with Bays classification. *Int Clin Neurosci J*. 2017;4(3):91–7.
20. Guger C, Ramoser H, Pfurtscheller G. Real-time EEG analysis with subject specific spatial patterns for a brain–computer interface (BCI). *IEEE Trans Neural Syst Rehabil Eng*. 2000;8:447–56.
21. Donchin E, Spencer KM, Wijesinghe R. The mental prosthesis: assessing the speed of a P300-based brain–computer interface. *IEEE Trans Neural Syst Rehabil Eng*. 2000;8:174–9.
22. Hoffmann U, Vesin JM, Ebrahimi T, Diserens K. An efficient P300-based brain–computer interface for disabled subjects. *J Neurosci Methods*. 2008;167:115–25.
23. Middendorf M, McMillan G, Calhoun G, Jones SK. Brain–computer interfaces based on the steady-state visual-evoked response. *IEEE Trans Rehab Eng*. 2000;8(2):211–4.
24. Mensh BD, Werfel J, Seung HS. 2004. BCI competition 2003—Data set Ia: Combining gamma-band power with slow cortical potentials to improve single-trial classification of 51, 1052–1056.
25. Burke DP, Kelly SP, Chazal P, Reilly RB, Finucane C. A parametric feature extraction and classification strategy for brain–computer interfacing. *IEEE Trans Neural Syst Rehabil Eng*. 2005;13(1):12–7.
26. McFarland DJ, Sarnacki WA, Wolpaw JR. Brain–computer interface (BCI) operation: Optimizing information transfer rates. *Biol Psychol*. 2003;63:237–51.
27. Wu W, Gao X, Hong B, Gao S. Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL). *IEEE Trans Neural Syst Rehabil Eng*. 2008;55(6):1733–43.
28. Qin L, He B. A wavelet-based time–frequency analysis approach for classification of motor imagery for brain–computer interface applications. *J Neural Eng*. 2005;2:65–72.
29. Lal TN, Schröder M, Hinterberger T, Weston J, Bogdan M, Birbaumer N, Schölkopf B. Support vector channel selection in BCI. *IEEE Trans Biomed Eng*. 2004;51(6):1003–10.
30. Schlogl A, Lee F, Bischof H, Pfurtscheller G. Characterization of four-class motor imagery EEG data for the BCI-competition. *J Neural Eng*. 2005;2:L14–22.
31. Obermaier B, Guger C, Neuper C, Pfurtscheller G. Hidden Markov models for online classification of single trial EEG data. *Pattern Recognit Lett*. 2001;22:1299–309.
32. Einalou Z, Maghooli K. (2010, November). Fuzzy neural network approach for noninvasive diagnosis of digestive diseases using wavelet comparing to classification followed by fuzzy C-mean algorithm. In *Biomedical Engineering (ICBME), 2010 17th Iranian Conference of* (pp. 1–4). IEEE.
33. Pfurtscheller G, da Silva FHL. Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clin Neurophysiol*. 1999;110:1842–57.
34. Lotte F, Congedo M, Lecuyer A, Lamarche F, Arnaldi B. A review of classification algorithms for EEG-based brain–computer interfaces. *J Neural Eng*. 2007;4:R1–13.
35. Einalou Z, Najafi Z, Maghooli K, Zandi Y, Sheibeigi A. Long-term simulation of digestive sound signals by CEPSTRAL technique. *Int J Appl Math Comput Sci*. 2007;4(2):844–8.
36. Wang B, Jun L, Bai J, Peng L, Li Y, Li G. 2005. EEG recognition based on multiple types of information using wavelet packet transform and neural networks. In: *Proc. 2005 IEEE Engineering in Medicine and Biology 27th Annual Conf. Shanghai, China*, pp. 5377–5380.
37. Sun S, Zhang C. Assessing features for electroencephalographic signal categorization. *Acoust Speech Signal Process*. 2005;5:417–20.
38. Trejo LJ, Rosipal R, Matthews B. Brain-computer interfaces for 1-D and 2-D cursor control: designs using volitional control of the EEG spectrum or steady-state visual evoked potentials. *IEEE Transact Neural Syst Rehab Eng*. 2006;14(2):225–9.
39. Rouillard J, Duprès A, Cabestaing F, Leclercq S, Bekaert MH, Piau C, Vannobel JM, Lecocq C. (2015). Hybrid BCI coupling EEG and EMG for severe motor disabilities. *Procedia Manufacturing*, 3, 29–36.
40. 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 Int J Light Electron Optics*. 2016;127(6):3271–5.
41. Aydemir O, Kayikcioglu T. Decision tree structure based classification of EEG signals recorded during two dimensional cursor movement imagery. *J Neurosci Methods*. 2014;229:68–75.
42. 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*.
43. Bennett KP, Blue JA. (1998, May). A support vector machine approach to decision trees. In *Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE International Joint Conference on* (Vol. 3, pp. 2396–2401). IEEE.
44. Dhillon IS, Mallela S, Kumar R. A divisive information-theoretic feature clustering algorithm for text classification. *J Machine Learn Res*. 2003;3:1265–87.
45. Radha R, Rajendiran P. February). Using K-Means Clustering Technique to Study of Breast Cancer. In: *Computing and Communication Technologies (WCCCT), 2014 World Congress on. IEEE; 2014*. pp. 211–4.
46. Ko W, Yoon J, Kang E, Jun E, Choi JS, Suk HI. (2018, January). Deep recurrent spatio-temporal neural network for motor imagery based BCI. In *2018 6th International Conference on Brain-Computer Interface (BCI)* (pp. 1–3). IEEE.
47. Sakhavi S, Guan C, Yan S. Learning temporal information for brain-computer interface using convolutional neural networks. *IEEE Transact Neural Netw Learn Syst*. 2018;29(11):5619–29.

48. Polat H, Özerdem MS. (2020, September). Automatic Detection of Cursor Movements from the EEG Signals via Deep Learning Approach. In 2020 5th International Conference on Computer Science and Engineering (UBMK) (pp. 327–332). IEEE.
49. Lee KY, Kim S. Designing discriminative spatial filter vectors in motor imagery brain–computer interface. *Int J Imaging Syst Technol.* 2013;23(2):147–51.
50. Kayagil TA, Bai O, Henriquez CS, Lin P, Furlani SJ, Vorbach S, Hallett M. A binary method for simple and accurate two-dimensional cursor control from EEG with minimal subject training. *J Neuroeng Rehabil.* 2009;6(1):1–16.
51. Huang D, Qian K, Oxenham S, Fei DY, Bai O. (2011). Event-related desynchronization/synchronization-based brain-computer interface towards volitional cursor control in a 2D center-out paradigm. In 2011 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB) (pp. 1–8). IEEE.

## Publisher's Note

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

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---