Research Article

Convolutional neural network based features for motor imagery EEG signals classification in brain–computer interface system

Samaneh Taheri[1] · Mehdi Ezoji[1] · Sayed Mahmoud Sakhaei[1]

Received: 12 October 2019 / Accepted: 29 February 2020 / Published online: 5 March 2020

Abstract
One of the essential challenges in brain–computer interface is to classify motor imagery (MI) signals. In this paper, an ensemble SVM-based voting system is proposed. In each line of this system, the EEG signal is transformed into different representations based on discrete cosine transform, Fourier transform, common spatial pattern, and empirical mode decomposition, and then these representations are combined in a triple-frame matrix. These frames are fed into a pre-trained deep convolutional neural network as a feature extractor. For each line, an SVM is employed to classify the extracted feature vectors. Finally, a decision is made based on voting between these SVMs. Performance of the proposed method is examined on the BCI Competition III dataset Iva to separate right hand and foot movement imagery. The simple proposed method achieves the average accuracy of 96.34% for all of the subjects, and 99.70% for the best situation that is an improvement in MI classification. In addition, it can be seen that right side of the brain is more effective than the other side in EEG-based MI classification.

Keywords Motor imagery · EEG signal · Convolutional neural network · Support vector machine · Voting-based classifier

1 Introduction

Human can control devices by his imaginary mental tasks through brain–computer interfaces (BCI). BCI is communication between the brain and an external device and it can help disabled people to do some muscle movements which is also called mind-machine interface (MMI), synthesis telepathy interface (STI), or direct neural interface (DNI) [1]. A BCI system can make users control computer actions without any physical activity [2]. It usually uses electroencephalography (EEG) to record brain activities, a non-invasive, and easy to apply approach [3].

Motor imagery (MI) is a mental processing state of the brain during which an individual simulates a motor action. MI has been used widely in BCI. By imagining or performing a muscle movement action, the power of mu (8–13 Hz) and beta (16–31 Hz) rhythms in the sensorimotor cortex will decrease or increase. Researchers hypothesized that an appropriate pattern of this phenomenon can be used as a suitable feature/signature in the classification. Since the EEG signals have low signal to noise ratio (SNR) and include useless information like artefacts, MI classification is known as a challenging and difficult problem [3, 4]. Therefore, MI classification attracted many researchers to work on its various issues. In this work, we make an effort to classify EEG-based MI signals of right hand and right foot presented in appropriate domains using a DCNN.

In [5], linear discriminant analysis (LDA) is used to classify left and right hand MI. LDA separates classes of objects by finding a linear combination of features. So, it cannot perform accurately at non-linear problems [6, 7]. Another classifier called support vector machine (SVM) is suggested to product non-linear decision boundary. And also, there are some defined kernel functions in case the datasets are not linearly separable [8, 9].

✉ Mehdi Ezoji, m.ezoji@nit.ac.ir | [1]Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran.

For BCI, one of the most commonly used features is the common spatial pattern (CSP) [9–15]. In [9], CSP was combined with some features such as variance, entropy, energy and logarithmic band power (LBP) and then, LDA, SVM and artificial neural network (ANN) were used for classification of hand and foot movement. Their results showed that the combination of CSP and LBP with LDA method performed better than others. Selim et al. [15] introduced a hybrid attractor metagene algorithm along with the Bat optimization algorithm to select the best subset of CSP features and also optimize the SVM parameters. Some studies have tried to use CSP as a filter to obtain a new time series. In [16], after some pre-processing method like band pass filtering (BPF) and independent component analysis (ICA) to remove artifacts, CSP was employed to design spatial filters for each class of signal and LDA was used as a classifier. Osuagwu et al. [16] achieved a maximum accuracy of $81 \pm 8\%$ for explicit MI and $83 \pm 3\%$ for implicit MI method in discrimination between left and right hand. Some other studies suggested combining some methods to achieve better results. In [17], a hybrid algorithm for two-class MI classification based on a cross-correlation method for feature extraction and a least square SVM (LS-SVM) as a classifier was proposed. The performances were evaluated with a tenfold cross-validation procedure, compared with eight other methods and achieved a 7.4% improvement.

Another most important method for feature extraction and classification is to use a convolutional neural network (CNN) [4, 12, 18]. In [4], BCI performance was improved by the combination of a long term short memory (LTSM) network and a spatial CNN. Then, a discrete wavelet transform (DWT) was used to obtain a feature vector. Results indicated a high level of accuracy.

Recently, in [12] a method was introduced for MI tasks classification based on deep convolutional neural network (DCNN). They applied short time Fourier transform and continuous wavelet transform as a time–frequency approach to transform data into images and then a network to classify. This method reached 99.35% accuracy on separating right hand and foot MI signals of BCI Competition dataset.

Empirical mode decomposition (EMD) for EEG signals is used to obtain IMF-based features for classifying MI EEG signals in BCI [19, 20]. In [21], band power (BP) on IMFs was applied to detect mu and beta rhythms. Then, new signals are reconstructed by IMFs, and BP was used on reconstructed signals. Then, hidden Markov model (HMM) was chosen as a classifier. The results showed that EMD caused suitable features to be extracted. In [20], normalized features were applied as inputs to LS-SVM classifier with different kernel functions. There is another study which tried to remove the noise and retain useful information of MI EEG signals [22]. For this purpose, Ensemble EMD (EEMD) method was designed as a new de-noising approach. The signal was decomposed by EEMD and then, an improved wavelet threshold method applied to de-noise high frequency components. At the end, IMFs were superimposed to yield a clean signal. Results implied a higher SNR and lower RMSE compared to other methods such as wavelet, and wavelet based on EMD/ EEMD. Another research used a combination of complete EEMD (CEEMD) and ICA to remove artefacts [23]. The authors of [24] proposed an optimal allocation system and employed a naïve Bayes classifier to discriminate the MI signals.

In this work, an algorithm is proposed for classification of an MI EEG signals based on a pre-trained DCNN by employing different representations of the EEG signal in transformed domains. After that, these representations are combined in a triple-frame matrix to form the input structure of the DCNN. Thus, a 4096-tuple feature vector is extracted for each MI EEG observation. Finally, each feature vector and its corresponding motor imagery task can be classified by an SVM-based voting strategy.

The rest of the paper is organized as follows: data description, other prerequisites, and CNN-based feature extraction strategy are described in Sect. 2. The proposed voting-based classification is described in Sect. 3. Experimental results are presented in Sect. 4.

2 Materials and methods

Our work is based on classification of single-trial EEG signals. According to Fig. 1, the approach includes several steps in order to process and discriminate MI EEG signals: representation of MI EEG signal by 3-dimensional structure, feature extraction using a pre-trained CNN, and finally SVM-based classification to classify the benchmark data. As shown, for each observation of MI EEG signal, a constructed 3-dimensional structure is fed to the pre-train CNN to extracting the corresponding high-level feature vector. SVM-based classifiers are used in a voting manner. Each SVM is trained using feature vectors extracted from its corresponding 3-dimentional structure. More details are provided in the Sects. 2–4.

2.1 Data description

In the proposed algorithm, we applied the EEG data from BCI Competition III dataset Iva provided by the Berlin BCI group [25, 26]. This dataset was recorded from 5 healthy subjects named as aa, al, av, aw, ay. These EEG signals were recorded from 118 electrodes according to the international 10/20 system. There were 280 observations for each subject, 140 observations for each task per subject. During each trial, subjects were required to perform 2 MI tasks i.e.

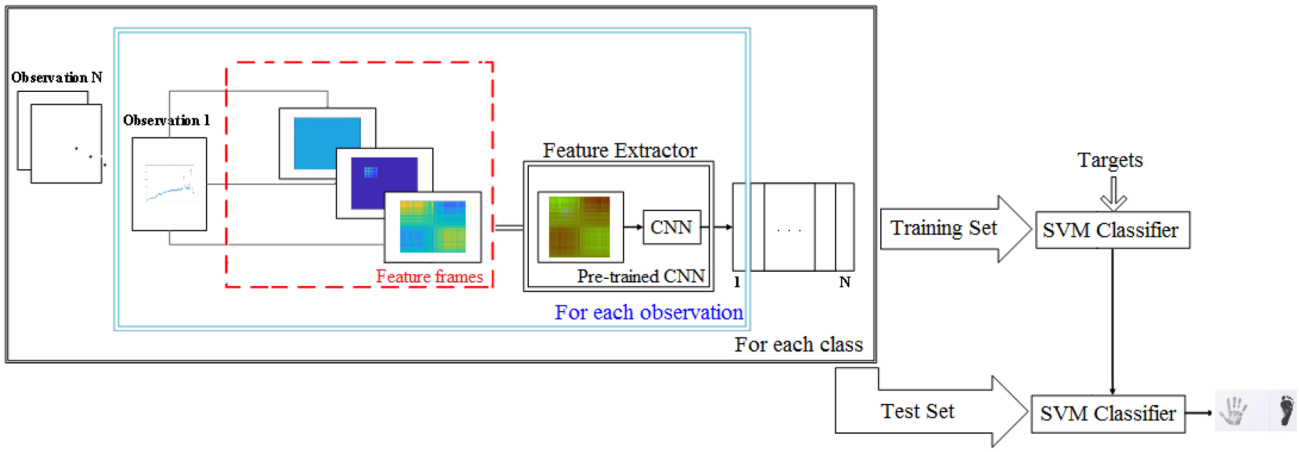

Fig. 1 Graphical abstract of the proposed CNN-based feature extraction and SVM-based classification of MI EEG-signals, including the representation of MI EEG signal by 3-dimensional structure, feature extraction using a pre-trained CNN, and SVM-based classification

right hand (class I) and right foot (class II) for 3.5 s. The sampling rate is 1000 Hz and also the version of downsampled at 100 Hz is available.

In Fig. 2, some observations from these classes are shown for subjects aa and al. It can be seen that EEG signals from the different motor imagery sometimes appear more similar than the same MI observed in different trials, i.e., between-class similarity caused in different trials sometimes exceeds within-class similarity, in MI EEG classification. Therefore, simple algorithms do not perform well under this challenges and EEG-based MI signal classification is a challenging problem.

EEMD is a good tool to be used as a time frequency transform that will be discussed by details in the next section.

2.2 Empirical mode decomposition (EMD) and Ensemble EMD (EEMD) algorithm

As we know about non-stationary property of MI EEG signals, a-time frequency method is suitable. One of the most effective methods is EMD which decomposes EEG signals into sets of time series named intrinsic mode functions (IMFs). Therefore, IMFs can be appropriate features for accurate classification of MI-EEG signals.

EMD is an adaptive time–frequency data analysis method. At the first step, we need to consider signals at the level of their local oscillations. Then, between every two local extrema (let's say between t_- and t_+), the signal can be decomposed into high-frequency part (detail) $d(t)$ and low-frequency part $m(t)$. So that we have $x(t) = m(t) + d(t)$ for $t_- < t < t_+$. Assuming that this happens for the entire signal, we get that what is referred to as an IMF and a residual consisting of all local trends. This residual can be considered as a new signal to decompose. Therefore, sequential fundamental components of a signal can be iteratively extracted [27].

The algorithm of EMD can be summarized as follows [27, 28].

(1) Identify all local extrema of signal $x(t)$.
(2) Find a lower "envelope", $e_l(t)$ that can interpolate all local minima.
(3) Find an upper "envelope", $e_u(t)$ that can interpolate all local maxima.
(4) Compute the local mean, $m(t) = (e_l(t) + e_u(t))/2$.
(5) Subtract m(t) from x(t) to extract the detail $d_i(t) = x(t) - m(t)$, (i indicates the order of IMF)
(6) Let $x(t) = d_i(t)$ and go to step 1.

The algorithm is applied to the residue iteratively, until residue becomes a uniform function from which no more IMFs can be generated. Then, the signal describes as follow:

$$x(t) = \sum_{i=1}^{M} d_i(t) + r(t) \quad (1)$$

where $d_i(t)$, $i = 1, \ldots, M$, are the IMFs and $r(t)$ is the last residue.

As we know, all data are a combination of signal and noise and it can represent undesired effects. Another popular problem that we may be faced with in EMD is mode mixing phenomenon that is caused by interference of signal and noise [29]. To solve these problems, there is an approach named Ensemble EMD (EEMD) which suggested to add white noise to each observation of data.

EEMD As [19, 30] believe, noise can be useful for data analysis in the EMD. A white noise with finite amplitude has a uniform representation in the time–frequency space.

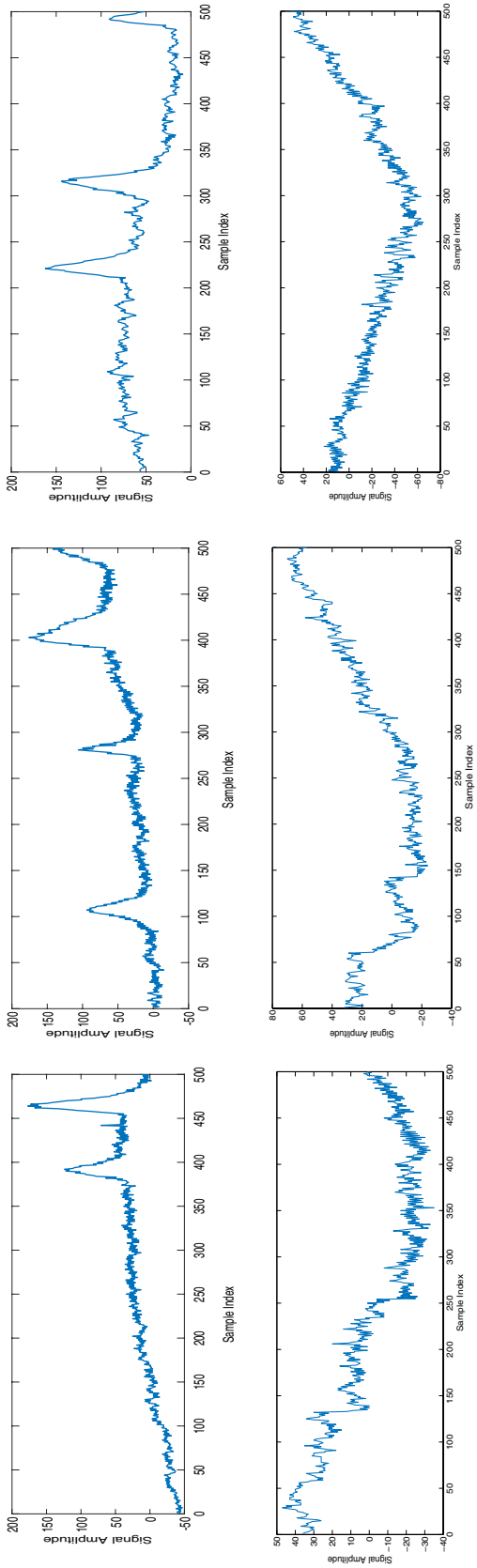

Fig. 2 Different observations of EEG signals from different subjects in 2 Classes: first row Observation 1 and 5 from class 1 and observation 1 from class 2 of subject aa, and the second row Observation 1 from class 1 and observation 1 and 5 from class 2 of subject al. *x* axis is sample index in time domain, and *y* axis is signal amplitude

Then, the signal is added to this 'uniformly distributed white background'. It is obvious that each signal trial indicates very noisy results, but since the added noises in each trial are different, so on the average of enough number of trials, noises will cancel out eventually. The mean of whole trials is used as a true IMF component.

In this work, we apply an EEMD method on data. The first 3 IMFs of one observation from each class of subject aa and al are shown in Fig. 3.

In this work, we use the 1st and 2nd IMFs along with other representations to generate a 3-dimension representation of EEG data.

2.3 Common spatial pattern (CSP)

CSPs are designed to transform signals into a new time series which is more appropriate for classification. In fact, in a new space, variances of classes are optimal for discrimination. The method is based on designing such spatial filters based on the simultaneous diagonalization of covariance matrices [31], i.e. a projection matrix is given by the CSP algorithm. By multiplying it in the main signal, we have it in optimal discriminative space.

Let's describe briefly about the CSP process. Assume a raw EEG signal named E with $N \times T$ dimensions, where N refers to the number of channels and T refers to the number of samples in each channel. At first, we should normalize the spatial covariance of each class.

$$C_f = \frac{E_f E_f'}{trace\left(E_f E_f'\right)} \quad C_h = \frac{E_h E_h'}{trace\left(E_h E_h'\right)} \tag{2}$$

E_f and E_h are data matrix of right foot and right hand movement, respectively. E' indicates transpose of matrix E, and $trace\left(EE'\right)$ is the sum of diagonal elements of EE'.

If C_c denotes the combination of averaged normalized covariance of two classes as $C_c = \overline{C}_f + \overline{C}_h$, it can be decomposed based on eigenvectors and eigenvalues like below:

$$C_c = U_c \Lambda_c U_c' \tag{3}$$

where U_c is the matrix of eigenvectors and Λ_c is the diagonal matrix of eigenvalues with descending order arrangement. Now, we need to compute the whitening transformation as $P = \sqrt{\Lambda_c^{-1}} U_c$. Then, we have:

$$S_h = P\overline{C}_h P', \quad S_f = P\overline{C}_f P' \tag{4}$$

The sum of corresponding eigenvalues of S_h and S_f should be equal to identity matrix and they have the same eigenvectors. Therefore, we can decompose S_h and S_f as $S_h = B\Lambda_h B'$ and $S_f = B\Lambda_f B'$, where $\Lambda_h + \Lambda_h = I$ and B is any orthogonal matrix which applies to the Eq. (5):

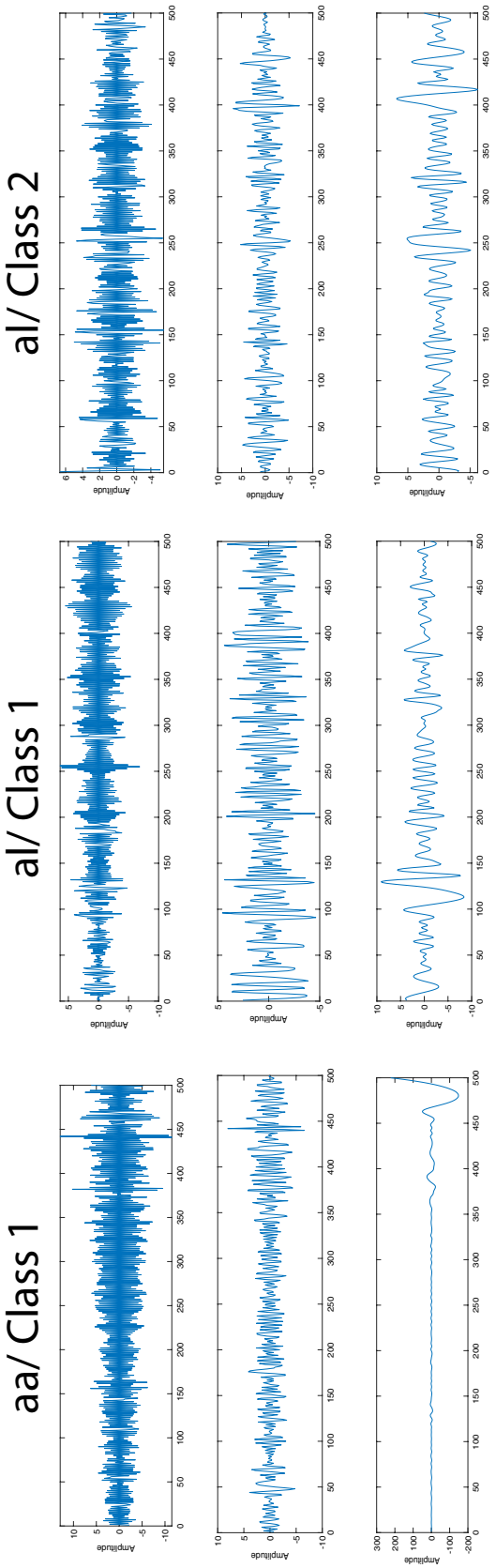

Fig. 3 Different IMFs of the first observation from subject aa and subject al: first row IMF1, second row IMF2, third row IMF3. *x* and *y* axes are entry index and amplitude of transformed signal

$$B'\left(S_h + S_f\right)B = I \tag{5}$$

Therefore, an eigenvector of the corresponding largest eigenvalue in S_h has the smallest eigenvalue in S_f and vice versa. This shows that the covariance between two classes is maximized. This property causes B as an eigenvector to be suitable for classification.

Now, we can access a set of CSP filters (projection matrixes) as follows:

$$W_{csp} = P'B = \left[w_1 w_2 \ldots w_{N-1} w_N\right] \in R^{N\times N} \tag{6}$$

W_{csp} is the CSP filter and provides maximum ratio of variance between two classes. The filtered signal is generated by multiplying CSP filters into the original signal [8].

Figure 4 illustrates the CSP-based representation of EEG signals corresponding to different MI tasks based on a W_{csp} originated from a typical training set.

2.4 Feature extraction

The goal of this work is to propose an algorithm to classify MI EEG signals. To this end, we need to extract some appropriate and high-level features. Here we decided to use a pre-trained convolutional neural network (CNN) called Alex net [32] as a feature extractor without any fine tuning. AlexNet is a CNN, which is trained on a large set of 3-dimensional images and can classify these images into 1000 categories. The architecture of Alex net is illustrated in Fig. 5 with five convolution layers, max-pooling layer, and three fully connected layers.

Through these convolution layers, the $227\times227\times3$ input image is filtered with 96, 256, 384, 384 and, 256 kernels, respectively. The first two fully connected layers include 4096 neurons with ReLU function as the activation function.

In this work, we represent the EEG signal based on Alex net's description of it from 4096-tuple FC7 layer. Then, these feature vectors are used to train an SVM-based classifier.

The input of network is 3-dimensional images in size of $227\times227\times3$, but EEG raw signals are 1-dimensional signals. Therefore, in this work, different 1-dimensional representations of input EEG signal s are considered such as DCT(s) [33], CSP(s), $\mathrm{IMF}_1(s)$, $\mathrm{IMF}_2(s)$, $\mathrm{IMF}_3(s)$, F(s) [34], and Ph(s) which are defined as follow:

- DCT(s) denotes the coefficients of Discrete Cosine Transform of s which are cropped in an appropriate window after discarding the DC component.
- CSP(s), denotes the magnitude of CSP-based representation of s.
- $\mathrm{IMF}_k(s)$ denotes k^{th} IMF of s based on EEMD.

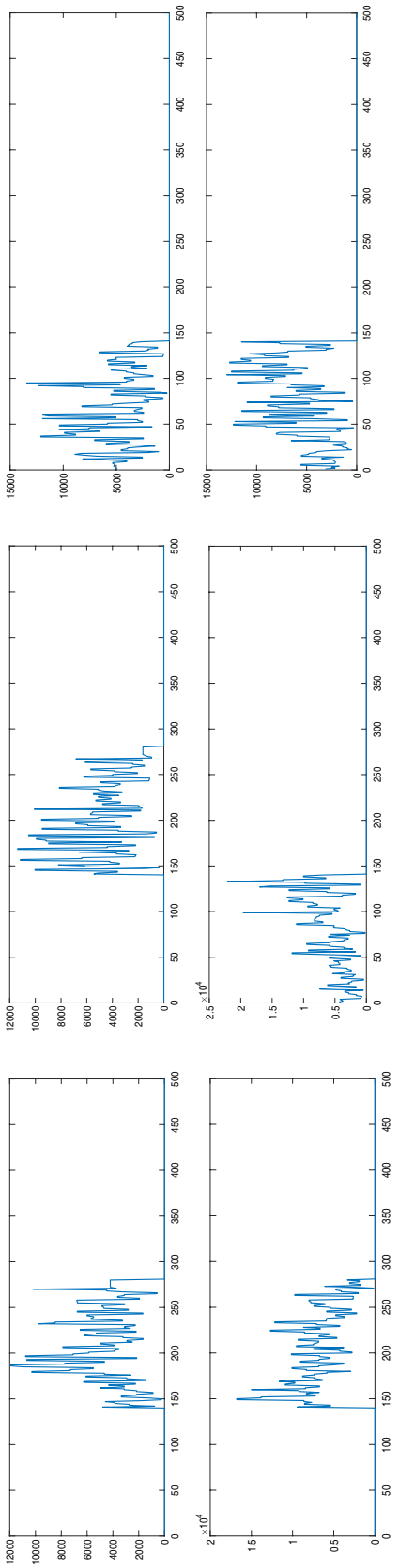

Fig. 4 CSP-based representation of EEG signal from different subjects in 2 Classes corresponding to observations illustrated in Fig. 2. *x* and *y* axes are entry index and amplitude of transformed signal

- Ph(*s*) denotes phase of Fourier transform of *s*.
- F(*s*) denotes the product of Ph(s) and the magnitude of Fourier transform of *s*.

In this work, we combine two 1-dimensional vectors x and y (from these representations) as $x * y'$ to transform them into 2-dimensional frame Z and feed three normalized frames as a 3-dimensional structure to the pre-trained DCNN. If $y = x$, Z is called to self-combination of x, and if $y \neq x$, Z is called to cross-combination of x and y. We denote this frame by $Z = f(x, y)$.

3 Classification

In this step, the feature vectors extracted from training set are used to train the SVM-based classifiers due to the good generalization ability of SVM. In the SVM framework for classification, a hyperplane is found to separate two classes such that the smallest distance between the hyperplane and any of the training samples is maximized [35]. This hyperplane is determined through a convex optimization problem with a sparse solution. In order to classify new data point x, the sign of $y(x)$ is evaluated defined as follows:

$$y(x) = \sum_{n=1}^{N} a_n t_n k\left(x, x_n\right) + b \tag{7}$$

where x_n, a_n and t_n are training samples, trained support values, and corresponding class labels, respectively. N is the number of the training d-dimensional feature vectors and b is the bias term. $k\left(x, x_n\right)$ is a kernel function such as Linear, polynomial (e.g. quadratic and cubic), Gaussian and etc.

In this work, in order to employ the ability of different combinations of features to highlight the discriminant information for each observation, a voting ensemble classifier is proposed. The block diagram of this system is illustrated in Fig. 6. In this figure, each block of feature extractor is introduced in Fig. 1 in detail.

4 Experiments and results

As mentioned before, BCI Competition III dataset Iva [26] is used to evaluate the proposed algorithm. It includes right hand and foot movement MI from 5 subjects, and for each of them there are 280 observations (140 observation for each class) as 500-tuple vectors. As explained in Sect. 2, each observation is transformed into a 3D structure by applying some time/frequency

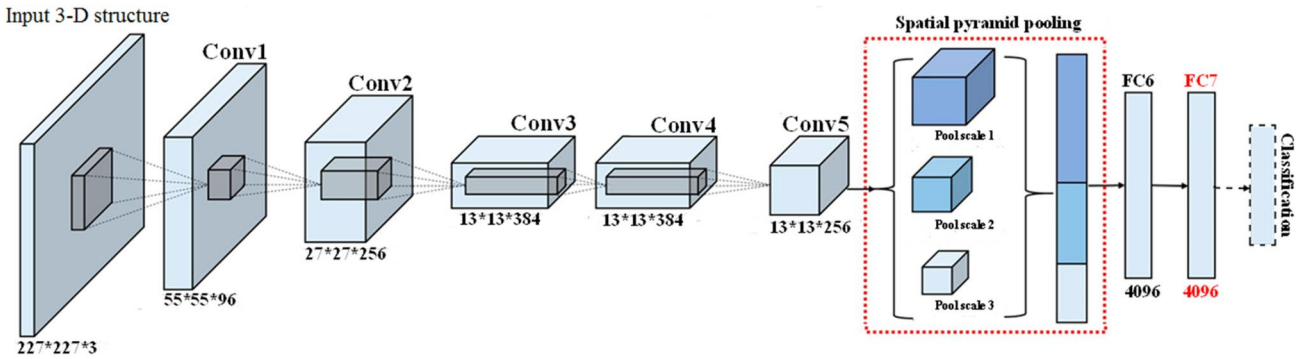

Fig. 5 The main architecture of Alex net [32] which used as the feature extractor in the proposed algorithm. The FC7 layer is selected as 4096-tuple feature vector per each input 3-dimensional structure

transformations. This 3D structure (corresponding to each observation) is fed to the DCNN and turn into a 4096-tuple feature vector. This feature vector is classified using SVM-based voting strategy.

The observations set of each subject are randomly partitioned into two disjoint and balanced subsets, 196 observations (98 observation per class) are allocated to the training/validating set for training the SVMs, and the rest to assess the performance the trained model based on recognition rate through several experiments. The accuracy or recognition rate criterion is calculated based on Eq. (8):

$$Recognition\ Rate = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

where TP, TN, FP, and FN refer to true positive, true negative, false positive and false negative, respectively.

In this work, the average recognition rate of the algorithm across multiple runs of the observation set partitioning, is reported to assess the performance of the proposed algorithm.

Experiment I In this experiment, we investigate the domain which the feature vectors are extracted from it. To this end, for each observation of subject-aa from channel-1, the Fourier phase and amplitude, DCT-based cropped coefficients, EEMD-based IMFs, and CSP-based representation, etc. are extracted. Then, we frame each of these feature vectors (denoted by x) as its self-combination $x * x'$. The average recognition rate of the proposed algorithm in multiple runs for each input structure is tabulated in Table 1.

It can be seen from Table 1, each of the proposed structure contains useful and discriminative information. We decided to discard M_3 from the proposed signal representation method due to its low generalization ability and less informative content in comparison to M_1 and M_2.

Experiment II In this work, different frames are investigated such as $Z_1 = f(s,s)$, $Z_2 = f(d,d)$, $Z_3 = f(c,c)$, $Z_4 = f(F,F)$, $Z_5 = f(M_1,M_1)$, $Z_6 = f(M_2,M_2)$, $Z_7 = f(s,c)$, $Z_8 = f(s,M_1)$, $Z_9 = f(s,M_2)$, $Z_{10} = f(s,Ph)$, $Z_{11} = f(s,F)$, $Z_{12} = f(d,s)$, $Z_{13} = f(d,c)$, $Z_{14} = f(d,M_1)$, $Z_{15} = f(d,M_2)$, $Z_{16} = f(d,F)$, $Z_{17} = f(d,Ph)$, $Z_{18} = f(c,F)$, $Z_{19} = f(c,Ph)$, $Z_{20} = f(c,M_1)$, $Z_{21} = f(c,M_2)$, $Z_{22} = f(F,M_1)$, $Z_{23} = f(F,M_2)$, $Z_{24} = f(M_1,M_2)$ for constructing a triple-frames ($Z_kZ_lZ_m$). We examined different reconstructions of the network's inputs (i.e. triple frames) to recognizing the MI movements from multiple runs for the subject-1. Results are tabulated in Table 2.

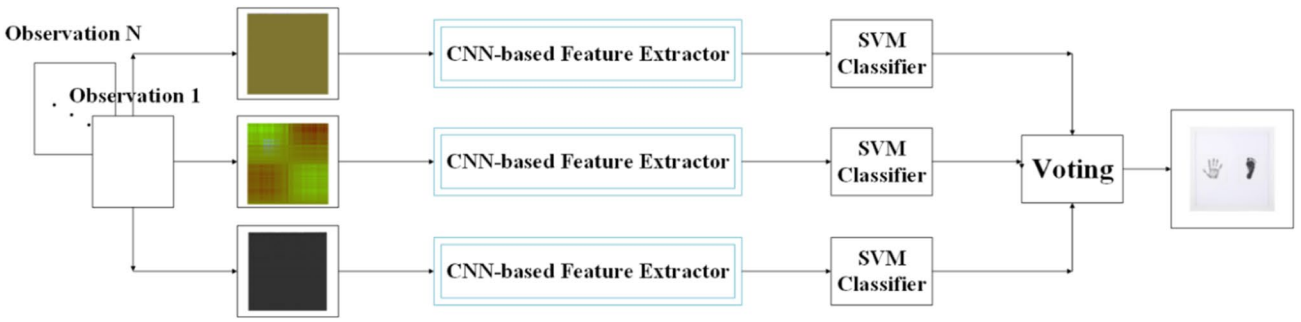

Fig.6 Block-diagram of the proposed voting ensemble classifier

Table 1 Average recognition rate (%) of the proposed algorithm versus different feature frames fed into the CNN

	Raw signal	DCT-based coefficients	CSP-based coefficients	F(s)	Ph(s)	M_1	M_2	M_3
Train set	72	70	100	85	70	70	70	75
Test set	76.5	75	76.4	88.7	76	72	75	70

Table 2 Average recognition rate (%) of the proposed algorithm based on different construction of the CNN inputs

	$Z_{12}Z_{02}Z_{13}$	$Z_{14}Z_{15}Z_{16}$	$Z_{01}Z_{07}Z_{08}$	$Z_{11}Z_{18}Z_{19}$	$Z_{03}Z_{20}Z_{18}$	$Z_{19}Z_{07}Z_{22}$	$Z_{01}Z_{10}Z_{04}$	$Z_{02}Z_{16}Z_{01}$
Train set	100	89.25	88.1	100	100	82.67	85	99.5
Test set	85.4	85.42	79.76	76.2	81	82.14	88.1	90
	$Z_{16}Z_{16}Z_{16}$	$Z_{16}Z_{17}Z_{3}$	$Z_{17}Z_{11}Z_{03}$	$Z_{05}Z_{24}Z_{06}$	$Z_{19}Z_{07}Z_{22}$	$Z_{24}Z_{10}Z_{14}$	$Z_{16}Z_{18}Z_{23}$	
Train set	90	95	95	63.0	82.67	72	95	
Test set	93.04	91.94	86	39.0	82.15	71.43	77.4	

It can be seen from Table 2, each of the proposed 3-dimensional structures extracts discriminative information from the training set. Due to high ability for test samples classification, we decide to construct the network input 3-dimensional structure based on Z_{03}, Z_{11},Z_{14}, Z_{15}, Z_{16}, Z_{17}.

Experiment III In order to take ability of the different structures in train and test phases, we proposed to ensemble these structures in a voting system. We examined various structures and concluded empirically that $Z_{17}Z_{11}Z_{03}$, $Z_{14}Z_{15}Z_{16}$ and $Z_{16}Z_{16}Z_{16}$ are the best 3-dimensional structures for the proposed SVM-based voting system.

The prepared 3-dimensional structures which provided for Alex net to extract the high-level feature vectors from different subjects and classes are illustrated in Fig. 7. The average recognition rate in multiple runs of the algorithm for different subjects is summarized in Table 3 in comparison with the other methods.

Experiment IV As we know, BCI Competition III dataset Iva recorded from 118 channels. Our entire results are shown in the above-mentioned tables are obtained from channel 1 (Fp1). In order to investigate the selected channel location in the performance of the proposed algorithm, the algorithm is applied to a variety of other channels in different locations which is illustrated in Fig. 8. The results are summarized in Table 4.

As it can be seen from Table 4,Query the proposed method for MI EEG signal classification, achieve accuracy above 93% in the worst case and obtain 99.7% in the best situation which is the best in comparison with the others.

Experiment V To evaluate the performance of the proposed method in distorted EEG data classification, white Gaussian noise with different Signal to Noise Ratio (SNRs) is added to the original EEG signals of the test data (channel F4 from subject aa). The best structures from the previous experiments, i.e., ($Z_{17}Z_{11}Z_{03}$), ($Z_{14}Z_{15}Z_{16}$), and ($Z_{16}Z_{16}Z_{16}$) are used. The results of the classification across multiple runs of the algorithm are shown in Fig. 9.

5 Discussion

Nowadays, BCI plays an important role in technology and rehabilitation processes. One of the most vital fields is MI-based BCI which draws researchers' attention. So, classifying different types of MI signals such as EEG, will be a way to improve the BCI systems. During brain–computer interface based on MI EEG signals' classification, we are facing serious challenges due to low discriminatory capacity of raw EEG signals, and also presence of the different artefacts. Thus, we investigate the different time/spatial/frequency domains for EEG signal representation, which the feature vectors are extracted from the appropriate combinations of them.

Our proposed method includes different steps such as preparing recorded signals, extracting appropriate features, which needed to design an algorithm for transforming 1-D signals into 3-D structures, and designing the SVM-based voting system to classify the feature vectors.

As we can see from Tables 1 and 2, each of the proposed and examined domains contains useful and discriminative information, and also different EEMD-based features are redundant. On the other hand, combination of the different domains is more discriminant from the viewpoint of generalization ability. Based on the generalization ability of the empirically examined structures, frequency domains such as Fourier transform, DCT, time–frequency domain such as EEMD, and spatial domain such as CSP are selected

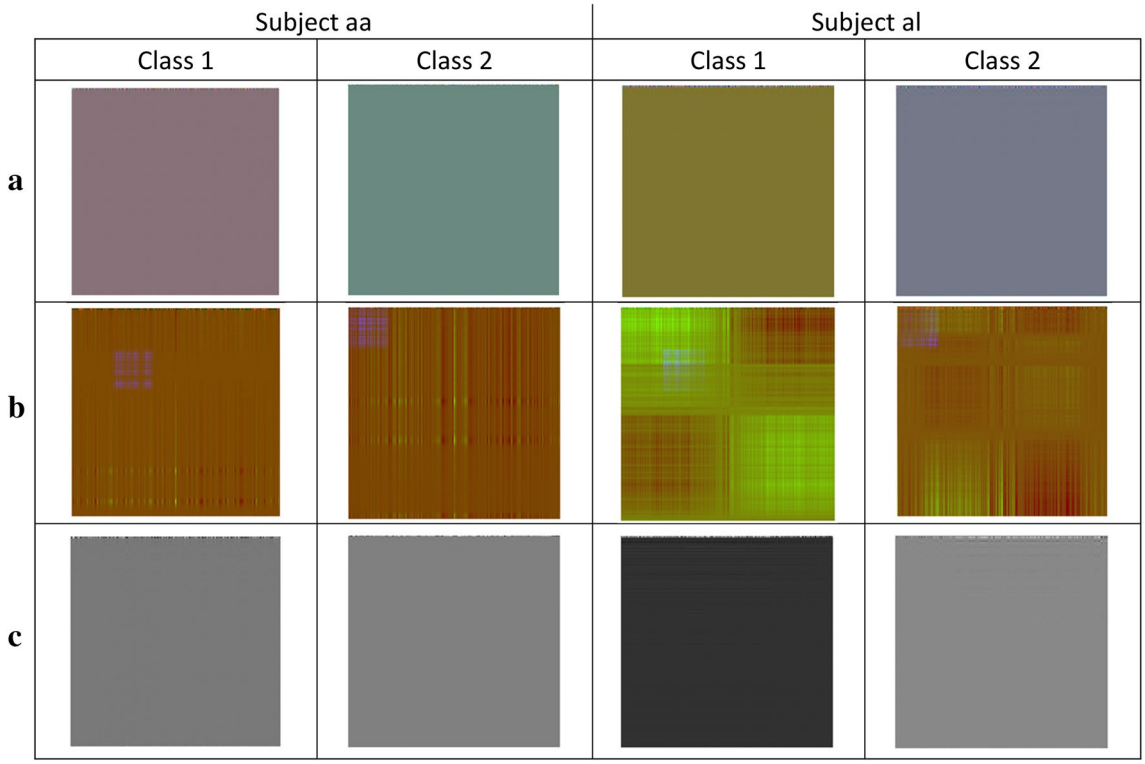

Fig. 7 The prepared 3-dimensional structure. **a** $Z_{14}Z_{15}Z_{16}$, **b**$Z_{17}Z_{11}Z_{03}$, **c** $Z_{16}Z_{16}Z_{16}$ extracted from different subjects and classes in order to feed into the Alex net

Table 3 The average recognition rate (%) and standard deviation in multiple runs of the algorithm for different subjects in comparison with the other methods

	[13]	[14]	[15]	[20]	[24]	[36]	The proposed methods
Subject 1	66.79	68.10	86.61	97.78	97.92	96	94.40
Subject 2	96.07	93.88	100	98.89	97.88	92.3	99.70
Subject 3	52.14	68.47	66.84	96.67	98.26	88.9	95.10
Subject 4	71.43	90.58	90.63	98.89	94.47	95.4	95.60
Subject 5	50	84.65	80.95	95.56	93.26	91.4	96.9
Average (std.)	67.29 (16.58)	81.14 (10.91)	85.01 (11.01)	97.56 (1.30)	96.36 (2.08)	92.8 (2.63)	96.34 (1.87)

for the EEG signals' representation and the proposed 3-D structure construction. As it can be seen from Table 1, the generalization ability of IMF3 is less than IMF1 and IMF2 and therefore IMF3 is discarded.

In experiment III and based on the Fp1 channel, the proposed method reached the classification accuracy of 99.7% in the best case and 94.4% in worst case, which are referring to subject 2 and subject 1, respectively. It can be seen from Table 3 that the works in [20, 24] are more accurate than the proposed method in the sense of the average recognition rate. Nevertheless, the proposed method achieves the best accuracy per subject (99.70%) in comparison with [20] (98.89%), and [24] (98.26%). Besides, in contrast to [20, 24], the proposed method is based on single electrode of EEG to classify MI EEG signals.

In another experiment, it was illustrated that classification accuracy changed in different EEG channels. We gained the average accuracy of 97.5% in the best channel among selected channels, which is related to channel F4.

As it can be seen from Fig. 9, although average accuracy of the algorithm fell down by decreasing the SNR, the algorithm is robust against the white Gaussian noise.

In addition, we can see that MI classification from electrodes located on the right/left side of the brain achieve the average accuracy of 96.2%/95.59%, respectively.

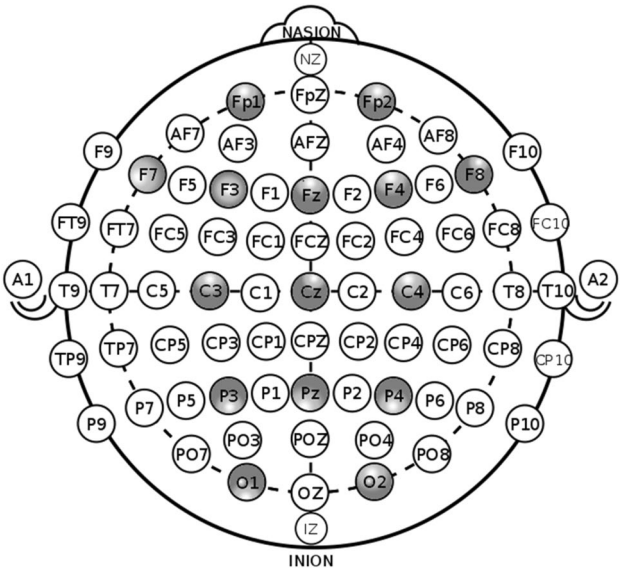

Fig. 8 An international recording 10–20 system. The grey channels are tested in the proposed algorithm

In comparison with the different methods, the proposed method achieves impressive results without any channel selection process.

About the limitations of this work, it can be noted that a two-class problem of MI EEG signal has been considered. We presented each MI EEG signal as a 3-D structure to feed a pre-trained CNN to obtain the high-level features. Extending the proposed method to the multiple class MI classification problem, and using other MI signals can be considered as the future works. And also, the effect of decision-making models on the proposed MI EEG-based representation can be examined in future work.

6 Conclusion

In this study, we proposed a method to classify MI-EEG signals into right hand and foot motor imagery. Dataset Iva from BCI Competition III is used to examine the proposed method. In this work, In order to extract the effective features, several representations of the signal in different domains such as DCT, Fourier, EEMD, and CSP are combined and are fed to the DCNN. Then, a framework based on deep convolutional neural network is employed as a feature extractor. Then, 3 SVMs are trained to classify the corresponding feature vectors. The decision is made through an ensemble voting system.

Table 4 Recognition rate (%) of the proposed algorithm in different channels

	F7	O1	P3	C3	F3	Fp1	Pz	Cz	Fz	O2	P4	C4	F4	Fp2	F8
Subject 1	97.8	97.8	97.4	98.8	97.3	94.40	97.8	98.8	97.6	97	91.3	97.7	98.0	97.6	97
Subject 2	93.8	96.4	94	96.1	97.3	99.70	96	95.2	96.1	98.2	93	92	99.28	99.1	99.4
Subject 3	99.7	97.6	98.2	93.1	94	95.10	97	92.8	96.4	92.8	93.1	92.8	97.3	94	95.2
Subject 4	95.2	94	97.3	93.2	97	95.60	95	96.4	94.9	94.3	94.8	96.8	94.6	91.6	96.4
Subject 5	97.6	94	97	93.6	96.4	96.9	96.7	98.2	96.4	95.2	94.04	96	98.8	95.2	95.2
Average	96.82	95.96	96.7	94.96	96.4	96.34	96.5	96.28	96.28	95.5	93.24	95.06	97.59	95.5	96.64

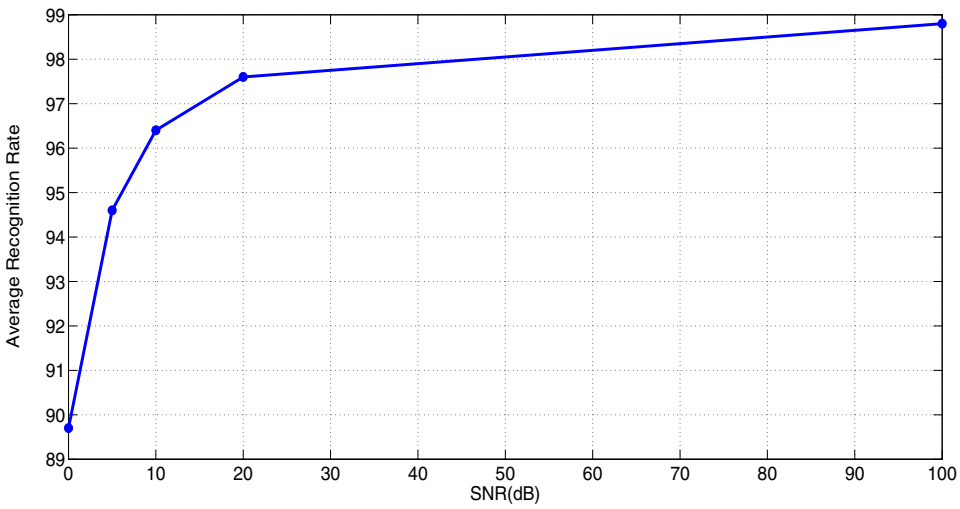

Fig.9 Average recognition rate of the proposed algorithm against white Gaussian noise

In order to evaluate the proposed algorithm, we designed different experiments and gained impressive results. The algorithm achieved classification accuracy of 99.7% for discriminating two motor imagery tasks, and the average accuracy of 96.34% for all of five subjects. Experimental results show that the proposed method is superior in comparison with other methods. In another experiment, different EEG channels were examined, and we have noticed that the proposed algorithm performed better MI classification from electrodes located on the right side of the brain than the left side, where is the source of MI signals.

In future, it would be of interest to examine the effect of decision-making or evidence accumulation models from cognitive neuroscience on the proposed MI signal-based representation, to extend the proposed method to the multiple class MI classification problem, and also to design a hardware platform for implementation of the proposed method for online applications.

Acknowledgements The authors acknowledge the funding support of Babol Noshirvani University of Technology (No. BNUT/389079/98-4).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Human and animal rights This paper does not contain any studies with human participants or animals performed by any of the authors.

References

1. Wolpaw JR et al (2000) Brain–computer interface technology: a review of the first international meeting. IEEE Trans Rehabil Eng 8(2):164–173
2. Lotte F et al (2018) A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. J Neural Eng 15(3):031005
3. Tang Z et al (2017) Single-trial EEG classification of motor imagery using deep convolutional neural networks. Optik 130:11–18
4. Yang J, Yao S, Wang JJIA (2018) Deep fusion feature learning network for mi-eeg classification. IEEE Access 6:79050–79059
5. Masoomi R, Khadem A (2015) Enhancing LDA-based discrimination of left and right hand motor imagery: outperforming the winner of BCI Competition II. In: 2015 2nd international conference on knowledge-based engineering and innovation (KBEI). IEEE
6. Vidaurre C et al (2010) Toward unsupervised adaptation of LDA for brain–computer interfaces. IEEE Trans Biomed Eng 58(3):587–597
7. Sai CY et al (2018) Automated classification and removal of EEG artifacts with SVM and wavelet-ICA. IEEE J Biomed Health Inform 22(3):664–670
8. Li Y et al (2008) A self-training semi-supervised SVM algorithm and its application in an EEG-based brain computer interface speller system. Pattern Recognit Lett 29(9):1285–1294
9. Aljalal M et al (2018) Feature extraction of EEG based motor imagery using CSP based on logarithmic band power, entropy and energy. In: 2018 1st international conference on computer applications and information security (ICCAIS). IEEE
10. Roijendijk L et al (2015) Classifying regularized sensor covariance matrices: an alternative to CSP. IEEE Trans Neural Syst Rehabil Eng 24(8):893–900
11. Blankertz B et al (2007) Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Process Mag 25(1):41–56
12. Chaudhary S et al (2019) Convolutional neural network based approach towards motor imagery tasks EEG signals classification. IEEE Sens J 19:4494–4500
13. Belwafi K et al (2018) An embedded implementation based on adaptive filter bank for brain–computer interface systems. J Neurosci Methods 305:1–16
14. Dai M et al (2018) Transfer kernel common spatial patterns for motor imagery brain-computer interface classification. Comput Math Methods Med 2018:9871603. https://doi.org/10.1155/2018/9871603
15. Selim S et al (2018) A CSP\AM-BA-SVM approach for motor imagery BCI system. IEEE Access 6:49192–49208
16. Osuagwu BA et al (2017) Is implicit motor imagery a reliable strategy for a brain–computer interface? IEEE Trans Neural Syst Rehabil Eng 25(12):2239–2248
17. Siuly S, Li Y (2012) Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain–computer interface. IEEE Trans Neural Syst Rehabil Eng 20(4):526–538
18. Schirrmeister RT et al (2017) Deep learning with convolutional neural networks for EEG decoding and visualization. Hum Brain Mapp 38(11):5391–5420
19. Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 1(01):1–41
20. Taran S et al (2018) Features based on analytic IMF for classifying motor imagery EEG signals in BCI applications. Measurement 116:68–76
21. Trad D, Al-Ani T, Jemni M (2015) A feature extraction technique of EEG based on EMD-BP for motor imagery classification in BCI. In: 2015 5th international conference on information and communication technology and accessibility (ICTA). IEEE
22. Zhang S et al (2018) Motor imagery electroencephalogram denoising method based on EEMD and improved wavelet threshold. In: 2018 Chinese control and decision conference (CCDC). IEEE
23. Pramudita BA et al (2017) Removing ocular artefacts in EEG signals by using combination of complete EEMD (CEEMD)—independent component analysis (ICA) based outlier data. In: 2017 international conference on robotics, automation and sciences (ICORAS). IEEE
24. Wang H, Zhang YJM (2016) Detection of motor imagery EEG signals employing Naïve Bayes based learning process. Measurement 86:148–158
25. Tangermann M et al (2012) Review of the BCI competition IV. Front Neurosci 6:55
26. Blankertz B et al (2006) The BCI competition III: validating alternative approaches to actual BCI problems. IEEE Trans Neural Syst Rehabil Eng 14(2):153–159
27. Flandrin P, Rilling G, Goncalves P (2004) Empirical mode decomposition as a filter bank. IEEE Signal Process Lett 11(2):112–114
28. Park C et al (2013) Classification of motor imagery BCI using multivariate empirical mode decomposition. IEEE Trans Neural Syst Rehabil Eng 21(1):10–22

29. Gao Y et al (2008) Analysis and solution to the mode mixing phenomenon in EMD. In: 2008 congress on image and signal processing. IEEE
30. Flandrin P, Gonçalves P, Rilling G (2005) EMD equivalent filter banks, from interpretation to applications. In: Huang NE, Shen SSP (eds) Hilbert–Huang transform and its applications. World Scientific, Singapore, pp 57–74
31. Ramoser H, Muller-Gerking J, Pfurtscheller G (2000) Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans Rehabil Eng 8(4):441–446
32. Han X et al (2017) Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification. Remote Sens 9(8):848
33. Gonzalez RC, Woods RE (2008) Digital image processing. Prentice Hall, New Jersey
34. Oppenheim AV, Schafer RW (1999) Digital signal processing. Prentice Hall, New Jersey
35. Bishop C (2006) Pattern recognition and machine learning. Springer, New York
36. Kevric J, Subasi A (2017) Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system. Biomed Signal Process Control 31:398–406