Rong Yuying (Orcid ID: 0000-0002-7359-0705)

Classification of Motor Imagery EEG Signals using CSCNN

Yuying Rong[2] | Xiaojun Wu[1,2] | Yumei Zhang[1,2]

[1] Key Laboratory of Modern Teaching Technology, Ministry of Education, Shaanxi Normal University, Xi'an, China
[2] School of Computer Science, Shaanxi Normal University, Xi'an, China

Correspondence

Yumei Zhang

Email: zym0910@snnu.edu.cn

ABSTRACT As an important part of brain-computer interface (BCI), the electroencephalography (EEG) technology of motor imagery has been gradually recognized for its great theoretical value and practical application. In this study, in view of the different motor imagery tasks corresponding active region of the EEG signals, we adopt a two-dimensional form including time, frequency and electrode location information, then we design a classification method containing continuous small convolution of convolutional neural network (CSCNN). This method is mainly used for feature extraction through continuous small convolutional kernels and one rectangle convolutional kernel, and the softmax classifier for classification. In the experiment, classification accuracy and kappa value are used as evaluation criteria to verify the effectiveness of the method proposed in this paper. For classification accuracy, BCI competition IV dataset 2b is used to compare with the other five classification methods (CNN, CNN-SAE, SAE, SVM and 1DCGRU). The results demonstrate that the overall accuracy of CSCNN is higher than other methods, and CSCNN obtains an average accuracy of 82.8%. For kappa value, BCI competition IV dataset 2b is used to compare with the other three methods (FBCSP, Twin SVM and CNN-SAE). The performance of CSCNN is better with an average value of 0.663. Overall, the results show that CSCNN maintains a small number of parameters and improves the classification accuracy.

KEY WORDS BCI; Motor imagery EEG signals; EEG classification; CSCNN

1 | INTRODUCTION

Brain-computer interface (BCI) is a new type of information transmission between the brain and the external device. By collecting electroencephalography (EEG) signals which then go through a series of signal analysis and process algorithms, the brain information is finally turned into the control commands that the output device can understand, which achieves the control of the human brain to the external device, such as the movement of the cursor and the grasping of the robot arm[1-3]. BCI can not only be used in the engineering and entertainment, but also play a more important role in providing communication for locked-in patients to the outside world [4].

Motor imagery (MI) EEG processing technology is an important part of the BCI, mainly involving spontaneous EEG which are characterized by spontaneous electrical activity of neurons in the sensorimotor cortex without fixed waveforms [5]. The spontaneous EEG frequency varies from 1 to 30 Hz and is mostly divided into four frequency bands: alpha (8-13Hz), mu (8-12Hz), beta (14-30Hz) and theta (4-7Hz) [6].

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ima.22405

Different MI tasks can cause changes in the specific frequency band of the sensorimotor cortex. So, we use the frequency referred to the MI tasks as control signals for MI EEG identification and classification.

Common spatial pattern (CSP) is most commonly used in signal feature extraction because it can synchronously utilize the spatial correlation of EEG signals and is suitable for processing EEG multi-dimensional signals [7]. Linear discriminant analysis (LDA) and Support vector machine (SVM) algorithms are used in different researches to classify the obtained one-dimensional signal features [8-9].

Recently, the neural networks have become popular in BCI to discover the unknown correlations. For instance, a 5-layer CNN which composes the feature extraction is applied for classifying MI tasks [10], but the key temporal information is not used in its method. Jin Zhang et al. [11] proposed a method combining LSTM and CNN, which made a great improvement than CSP. An et al. [12] studied a more powerful deep belief network (DBN) is constructed by borrowing the Ada-boost algorithm, aiming at classifying EEG signals. Phothisonothai et al. [13] proposed a method of classifying EEG by a three-layer feed-forward neural network based on a simple backpropagation algorithm. Recurrent Neural Network (RNN) model such as GRU is more suitable for processing the sequence data. However, the RNN model does not consider the spatial combination of the EEG data [14].

Convolutional neural network (CNN) is widely applied due to good spatial locality and certain degree of translation invariance [15]. However, in reference [16], the constructed CNN has too many parameters and the accuracy of classification at limited time needs to be improved. Therefore, we employ continuous small convolution of convolutional neural network (CSCNN) to classify MI EEG.

In this paper, the input form combining time, frequency and electrode location information is introduced. mu band (4-14Hz) and beta band (14-32Hz) are selected, and the short time fourier transform (STFT) method is used to obtain the frequency domain information of C3, Cz and C4 electrodes respectively. Then the frequency domain information of mu and beta bands in three electrodes are combined as the vertical axis as well as the time are plotted as the horizontal axis, which obtains the two-dimensional feature matrix.

In our approach, CSCNN is proposed in order to improve the performance of classification in MI tasks. In addition, in consideration of the computational cost, we use continuous small convolution kernels which can reduce the number of parameters. The overall EEG signals processing and classification process is shown in figure 1.

2 | MATERIALS AND DATA PREPROCESSING

2.1 | Data description

The experimental data in this paper are derived from BCI competition IV dataset 2b, which records the left/right hand MI tasks. The dataset consists of EEG from 9 subjects. The subjects were right-handed and had normal or corrected vision. All of the volunteers sat in armchairs, looking at a flat screen monitor 1m from their eyes. Each subject participated in sessions for five times. The first two training sessions contain two types of imagery and each type of imagery contains 60 trials. Each of the other three sessions includes 160 trials. EEG data were recorded by three electrodes (C3, Cz, C4) at sampling frequency of 250Hz. In each trial, 1-2s is task-free stage, and 3-7s is MI stage.

In this paper, in order to employ the change in signals, we use 3-7 seconds of EEG signals to compare with mean value of 0-2 seconds of EEG signals before the MI task started. 2-second segment in one trial is used as the sample. 50 points of data are used as sliding window to slide forward to select the piece in the MI stage (3-7s). So one trial takes 11 pieces. 2-second segment of three electrodes (C3, Cz, C4) is shown in figure 2.

2.2 | Two-dimensional form generation

When people doing unilateral limb movements (such as the right-hand movement), corresponding to the side and the left-brain cortex regions of the amplitude of mu (8-12 Hz) and beta (18-23 Hz) will significantly reduce. At the same time corresponding ipsilateral brain and right brain cortex area of the amplitude of mu and beta bands will increase. This phenomenon is known as the event-related desynchronization / synchronization (ERD/ERS) [17]. More importantly, ERD/ERS also occurs when a person relies solely on imagination for unilateral limb movement. And the phenomenon involves both EEG signals at the C4 and C3 electrodes. Cz is also affected by hand movements.

EEG signal analysis methods are similar to speech signal, the common analysis methods are mainly to extract features of frequency domain and time domain information. Then the feature vector of each electrode calculated are summarized into one feature matrix as the input of the classifier. However, these methods obviously ignore the electrode position information. The brain is a complex structure, and it is not enough to process and analyze EEG signals. Therefore, time, frequency and location information are combined into two-dimensional form as input data to analyze in this paper.

The 2-second segment of each electrode contains 500 sample points. The window size of STFT is 64 and the time lapses is 14, so the length of time axis in the feature matrix is 32. In order to obtain Nfr*32 image, where Nfr is the size of the vertical axis. STFT is used to process time series. Then the frequency bands (mu=4-14 Hz and beta=16-32 Hz) were extracted from the output spectrum. And the cubic interpolation method is adopted for beta band, and the spectrum of C3, Cz and C4 electrode is combined to obtain Nfr=132.

3 | CLASSIFICATION METHOD

3.1 | Convolutional networks analysis

CNN is a kind of feedforward neural networks, which is very suitable for the extraction of two-dimensional features and close to the real biological neural network.

The basic structure of CNN mainly includes five parts: input layer, convolutional layer, pooling layer, full connection layer and output layer [18]. The convolutional layer is the core of CNN which is responsible for most of the calculation. The convolutional layer uses the convolution kernel of a specific size to convolve with the input image of the previous layer to obtain multiple feature maps. Inspired by [19], by the proper convolutional kernel decomposition can obtain similarly expressive local representations and speed up the training. Therefore, this paper adopts continuous two-layer convolutional kernels instead of larger convolutional kernels to feature extraction. Because the larger convolution is usually computationally expensive, for example, a 5×5 convolution with n filters is 25/9=2.78 times more computationally expensive than a 3×3 convolution with the same number of filters on a grid with m filters [19]. Although the 5×5 filter has a stronger ability to capture information. In order to balance calculation cost and information loss, we use 3×3 convolution for decomposition.

In recent years, going deep greatly improves the learning/fitting ability of the whole network, but at the same time increases the model size and computing cost [20]. In this paper, considering the cost problem, we adopted fewer layers. Besides, a gradual reduction in the representation size is used to avoid mass loss of information.

3.2 | CSCNN framework

Based on the above analysis, we adopt continuous small convolutions replacing single large convolution to maintain the range of the receptive field while reducing the number of parameters. A CSCNN model is shown in figure 3. The network is described as follows:

Input layer: the input form is a 132×32 matrix which contains three EEG feature information of time, frequency and location.

Two convolutional layers: this part includes two continuous 3×3 convolution kernels in order to increase the local perception of EEG signal features. Both convolution layers are 10 filters. After two convolution operations, 10 feature maps are obtained, each of which has a size of (128×28). The k th feature map at a given layer can be described as

$$h_{ij}^{k} = f(a) = f((W^{k} * x)_{ij} + b_{k}) \tag{1}$$

where x is the two-dimensional form of MI EEG. W^{k} is the weight matrix for filter k and b_{k} is the bias value, for k =1, 2, …, (number of filter)NF. j is equal to 1 and i = 1, 2, ..., 30. The output function f is selected as rectified linear unit (ReLU) function. ReLU is defined as

$$f(a) = \mathrm{ReLU}(a) = \max(0, a)$$

(2)

The max-pooling layer: this layer aims to MI EEG feature selection and prevention of overfitting. We use max pooling in this layer and the size of sampling factor is 4×4. So the size of output maps of this layer are subsampled into 32×7.

The convolutional layer: we used 30 convolution kernels with rectangle size of 28×3. Our aim is to discover more feature information along the vertical axis.

The max-pooling layer: the subsampling layer. The size of the sampling factor is 5×5.

The fully connection and softmax: each neuron in fully connection is connected to each neuron of the output of the max-pooling layer. The softmax classifier is used as prediction of left/right hand MI EEG tasks. The softmax is defined as

$$p_{i} = \frac{e^{z_{i}}}{\sum_{i=1}^{n} e^{z_{i}}} \tag{3}$$

where n is the number of classification label, in this study, the categories mainly involve the left and right hands, so n is equal to 2. z_{i} is the value from fully connection output.

The output layer: the left/right hand classification results can be obtained.

4 | EXPERIMENTAL RESULTS AND DISCUSSION

BCI competition IV 2b is used to evaluate our method. Each subject we use has three sessions containing 400 trials. The first two sessions contain 240 trials, and the third session contains 160 trials. Electro-oculogram (EOG) and other noises existed in each trial were not removed.

In this study, the experiments are computed by MATLAB, and the operating environment is an Intel 3.7 GHz Core i7 PC with 32GB of RAM.

In the data preprocessing, we set the time interval to 50 sample points, and the EEG signals of 2s length is evenly extracted from 3-7s. Then processed by STFT, the input images containing time, frequency and pole information is constructed. In the CSCNN method, we use the network of continuous small convolution kernels and one rectangle convolutional kernel. For 500 epochs, the effect of batch size on accuracy is shown on Figure 4 with the average time. It can be seen from figure4 that the accuracy is high when the batch size equals 30 at the minimum running time. In this study, we select batch size of 30. In each session, 90% trials are selected randomly as the training set and the remaining 10% are selected as the test set.

Table 1 shows that CNN has more than twice as many CSCNN parameters as our method. The CNN compared has one convolution layer and one max-pooling layer from reference in [15]. Compared with the 132×3 convolution kernel in the one convolutional layer of CNN, CSCNN is added two 3×3 convolution kernels, which increases the receptive field, reduces the difficulty of training and avoids overfitting caused by too many parameters.

We compare the accuracy of CSCNN with CNN [16], CNN-SAE [21], SAE [21], SVM [21] and 1DCGRU [22]. In 1DCGRU network, it is a variant of LeNet-5, and the last two fully connected layers is instead of the GRU network. As can be seen from the table 2, the average accuracy result of CSCNN is higher than other four methods, with an average improvement of 82.8%. The best result was in bold front for every subject. It can be clearly found that the classification accuracy of subject 1,2,6,7,8 perform better. This indicates that CSCNN obtains the effect in reducing information loss in term of extracting feature. Compared with CNN-based models, SAE and SVM is weaker in feature extraction. So, it can be obvious seen that SAE and SVM are both lower than CNN-based models in average accuracy. Among them, CSCNN in subject 2 has obvious advantage with 22.8% higher than another best result. This suggests that CSCNN improve extraction capacity of indistinct feature. However, in subject 4, the performance of CNN-SAE achieves the best, but the accuracy of our method is only 2% lower. In subject 5 and subject 9, the accuracy of the CSCNN is very close to that of the best-performing 1DCGRU. And CSCNN has poor performance in subject 3, although the accuracy result is much higher than the SAE. Overall, it is clearer in the figure 5 that CSCNN performs better overall than the other five models from subject 1 to subject 9.

In addition, we introduce kappa value to measure the classification accuracy by removing the influence of random classification accuracy. Kappa is calculated as below:

$$\text{kappa} = \frac{a-r}{1-r} \tag{3}$$

where a is the overall classification accuracy and r is the rate of theoretical consistency.

Table 3 summarizes the kappa value for the 9 subjects. Kappa value varies from 0 to 1, and the higher kappa value indicates the better model consistency. The best kappa value was highlighted with bold front. Compared with FBCSP [7], Twin SVM [23] and CNN-SAE in table 3, it easily suggests that the proposed CSCNN method outperformed in most subjects. On the average, there is a significant improvement in the CSCNN at 11.6% over the other method. From figure 6, CSCNN has better classification accuracy.

5 | CONCLUSION

In this study, we propose a CSCNN which mainly consists of two continuous small convolutional kernels and one rectangular convolutional kernel to improve performance of classification for motor imagery based BCI. In our network structure, the existence of small convolutional kernels cuts down the number of parameters which reduces the difficulty of training and avoids overfitting caused by too many parameters. And one rectangular convolutional kernel helps to obtain more spectrum feature information along the vertical axis. In addition, two-dimensional form of MI EEG combining time, frequency and electrode location information is used as the input of CSCNN. Finally, the proposed method is evaluated by the public dataset (BCI competition IV 2b). For the accuracy, the results show that our method is superior to others. Although the accuracy of subject 3, 4, 6, 9 is poor, it is close to or even above the average of other four advanced methods. Besides, in the assessment of kappa value, CSCNN demonstrates good consistency and significantly outperform other methods with an average value of 0.663.

ACKNOWLEDGMENT

This work was supported in part by the National Key Research and Development Program of China under Grant No. 2017YFB1402102, in part by the National Natural Science Foundation of China under Grant Nos. 11872036, 11772178, Shaanxi Science and Technology Plan Project (Grant Nos. 2019ZDLSF07-01, 2019GY-217).

REFERENCES

1. D. J. Mcfarland, W. A. Sarnacki, and J. R. Wolpaw, “Brain–computer interface (BCI) operation: optimizing information transfer rates,” *Biol. Psychol.*,” vol. 63, no. 3, pp. 237-251, Mar. 2003.
2. L. Na, W. Jie, and F. Zuren, “Feature extraction by common spatial pattern in frequency domain for motor imagery tasks classification,” in *Proc. CCDC*, May 2017, pp. 5883-5888.
3. S. Zhang, Y. Zheng, D. Wang et al., “Application of a common spatial pattern-based algorithm for an fNIRS-based motor imagery brain - computer interface,” *Neuroscience Letters*, vol. 655, pp. 35-40, Jun. 2017.
4. T. Hinterberger, A. Kübler, J. Kaiser, et al., “A brain–computer interface (BCI) for the locked-in: comparison of different EEG classifications for the thought translation device,” *Clin. Neurophysiol.*, vol. 114, no. 3, pp. 416-425, Mar. 2003.
5. J. Kevric, A. Subasi, “Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system,” *Biomed. Signal Process. Control*, vol. 31, pp. 398-406, Jan. 2017.
6. G. Pfurtscheller, C. Neuper, “Motor imagery and direct brain-computer communication,” *Proc. IEEE*, vol. 89, no. 7, pp. 1123-1134, Jul. 2001.
7. A. K. Keng, C. Z. Yang, W. Chuanchu, et al., “Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b,”. *Front. Neurosci.*, vol. 6, pp. 1-9, Mar. 2012.
8. S. I. Putu, K. Shin'Ichiro, and M. Ko-Ichiro et al., “Investigation of Methods for Extracting Features Related to Motor Imagery and Resting States in EEG-Based BCI System,”. *IEEE Trans. Electr. Information & Syst.*, vol. 129, no. 10, pp. 1828-1833, Oct. 2015.
9. E. Arbabi, M. B. Shamsollahi, and R. Sameni, “Comparison between Effective Features Used for the Bayesian and the SVM Classifiers in BCI,” in *Proc. EMBS*, Sept. 2005, pp. 5365-5368.
10. Z. Tang, C. Li, and S. Sun, “Single-trial EEG classification of motor imagery using deep convolutional neural networks,” *Optik*, vol. 130, pp. 11-18, Feb. 2017.
11. Z. Wang, L. Cao, Z. Zhang, et al., “Short time Fourier transformation and deep neural networks for motor imagery brain computer interface recognition,” *Concurrency & Computation Practice & Experience*, vol.30, no.23, pp. 1-9, Jan. 2018.
12. X. An, D. Kuang, X. Guo, et al., “A Deep Learning Method for Classification of EEG Data Based on Motor Imagery,” in *Proc. ICIC*, Aug. 2014, pp. 203-210.
13. M. Phothisonothai, M. Nakagawa, “EEG-Based Classification of Motor Imagery Tasks Using Fractal Dimension and Neural Network for Brain-Computer Interface,” *IEICE Trans. Inf. Syst.*, vol. E91D, no. 1, pp. 44-53, Jan. 2008.
14. M. Seera, C. P. Lim, K. S. Tan, et al., “Classification of transcranial Doppler signals using individual and ensemble recurrent neural networks,” *Neurocomput*, vol.249, pp. 337-344, Aug. 2017.
15. H. Yang, S. Sakhavi, K. K. Ang, et al., “On the use of convolutional neural networks and augmented CSP features for multi-class motor imagery of EEG signals classification,” in *Proc EMBC*, Aug. 2015, pp. 2620-2623.
16. G. Tian, Y. Liu, “Study on classification of left-right hands motor imagery EEG signals based on CNN,” in *Proc. ICCI*CC*, Jul. 2018, pp. 324-329.
17. I. Hossain, A. Khosravi, and S. Nahavandhi, “Active transfer learning and selective instance transfer with active learning for motor imagery based BCI,” in *Proc. IJCNN*, Jul. 2016, pp. 4048-4055.
18. X. Li, W. Wang, X. Hu, et al., “Selective Kernel Networks,” in *Proc. CVPR*, Jun. 2019, pp. 510-519.
19. C. Szegedy, V. Vanhoucke, S. Ioffe, et al., “Rethinking the Inception Architecture for Computer Vision,” in *Proc. CVPR*,

Jun. 2016, pp. 2818-2826.

20. P. Judd, J. Albericio, T. Hetherington, et al., “Proteus: Exploiting Precision Variability in Deep Neural Networks,” *Parallel Comput*, vol. 73, pp. 40-51, Apr. 2017.

21. Y. R. Tabar, U. Halici, “A novel deep learning approach for classification of EEG motor imagery signals,” *J. Neural Eng.*, vol. 14, no. 1, pp. 1-11, Feb. 2017.

22. J. Cai, C. Wei, X. L. Tang, et al., “The Motor Imagination EEG Recognition Combined with Convolution Neural Network and Gated Recurrent Unit,” in *Proc. CCC*, Jul. 2018, pp. 9598-9602.

23. S. Soman, Jayadeva, “High performance EEG signal classification using classifiability and the Twin SVM,” *Appl. Soft. Comput.*, vol. 30, pp. 305-318, May. 2015.