Predicting Human Intention-Behavior Through EEG Signal Analysis Using Multi-Scale CNN

Chenxi Huang, Yutian Xiao, and Gaowei Xu

Abstract—At present, the application of Electroencephalogram (EEG) signal classification to human intention-behavior prediction has become a hot topic in the brain computer interface (BCI) research field. In recent studies, the introduction of convolutional neural networks (CNN) has contributed to substantial improvements in the EEG signal classification performance. However, there is still a key challenge with the existing CNN-based EEG signal classification methods, the accuracy of them is not very satisfying. This is because most of the existing methods only utilize the feature maps in the last layer of CNN for EEG signal classification, which might miss some local and detailed information for accurate classification. To address this challenge, this paper proposes a multi-scale CNN model-based EEG signal classification method. In this method, first, the EEG signals are preprocessed and converted to time-frequency images using the short-time Fourier Transform (STFT) technique. Then, a multi-scale CNN model is designed for EEG signal classification, which takes the converted time-frequency image as the input. Especially, in the designed multi-scale CNN model, both the local and global information is taken into consideration. The performance of the proposed method is verified on the benchmark data set 2b used in the BCI contest IV. The experimental results show that the average accuracy of the proposed method is 73.9 percent, which improves the classification accuracy of 10.4, 5.5, 16.2 percent compared with the traditional methods including artificial neural network, support vector machine, and stacked auto-encoder.

Index Terms—Electroencephalogram (EEG) signal classification, human intention-behavior, short time fourier transform (STFT), multi-scale convolutional neural networks

1 Introduction

Brain-computer interface (BCI), which is also known as brain-machine interface, is a direct connection between human's or animal's brain and external devices. It is a cutting-edge technology developed in the 1970s, which converts the human brain wave signals into control signals, enabling the human brain to interact with external processing systems or devices. To date, there have been many BCI-based techniques investigated by various researchers and engineers [1], [2], [3]. Among them, electroencephalography (EEG)-based BCI is one of the most successful and useful techniques because of its advantages, such as cheap, simple, non-invasiveness.

Currently, EEG-based BCI methods have been extensively applied in many applications, including motor imagery, sleep stage analysis, emotion recognition, and etc [4], [5], [6], [7], [8]. Motor imagery, as one of the most successful EEG-based BCI applications, aims to predict the human intention-behavior accurately and perform a task without any physical behavior. This can help and benefit the elderly and disabled people who have difficulty in moving their bodies. Traditionally, the human's intention-behavior is predicted through EEG signal analysis in the following two steps: (1) feature extraction of the raw EEG signal, and (2) the extracted features classification. Specifically, in the first step, the discriminative spatial or spectral features are often manually extracted from the raw EEG signal data. In the second step, the extracted features are classified through some machine learning models in order to predict the human intention-behavior. In the past decades, a series of feature extraction and classification methods have been proposed for motor imagery EEG-based BCI. Common spatial patterns (CSP) and independent component analysis (ICA) are two primary methods used to extract spatial features. For example, Park *et al.* [9] proposed an augmented complex CSP algorithm to process EEG data and extract signatures of motor imagery tasks. Experimental results indicate that the proposed algorithm can maximize the inter-class difference between different motor imagery tasks, verifying the effectiveness of designed CSP algorithm. Wang *et al.* [10] designed an EEG spatial filter using ICA for motor imagery task, experiments conducted on an motor imagery dataset proved that the designed method increases the practicality of motor imagery BCI systems. Apart from this, there are also many time-frequency analysis methods to extract spectral features, such as wavelet transform (WT), short time Fourier transform (STFT). Xu *et al.* [11] employed a WT method to extract time-frequency features from C3, Cz, and C4 channels for motor imagery EEG classification. Zabidi *et al.* [12] designed a STFT analysis method to extract the frequency contents varying time windows of EEG signals. In order to classify the extracted features, a number of machine learning models and algorithms, mainly including artificial neural network (ANN), and support vector

- *Chenxi Huang and Yutian Xiao are with the School of Informatics, Xiamen University, Xianmen 361005, China. E-mail: supermonkeyxi@xmu.edu.cn, 3179487641@qq.com.*
- *Gaowei Xu is with the School of Electrical Information and Engineering, Tongji University, Shanghai 201804, China. E-mail: gaowei_xu@163.com.*

Manuscript received 7 May 2020; revised 11 Nov. 2020; accepted 18 Nov. 2020. Date of publication 23 Nov. 2020; date of current version 7 Oct. 2021.
(Corresponding author: Gaowei Xu.)
Digital Object Identifier no. 10.1109/TCBB.2020.3039834

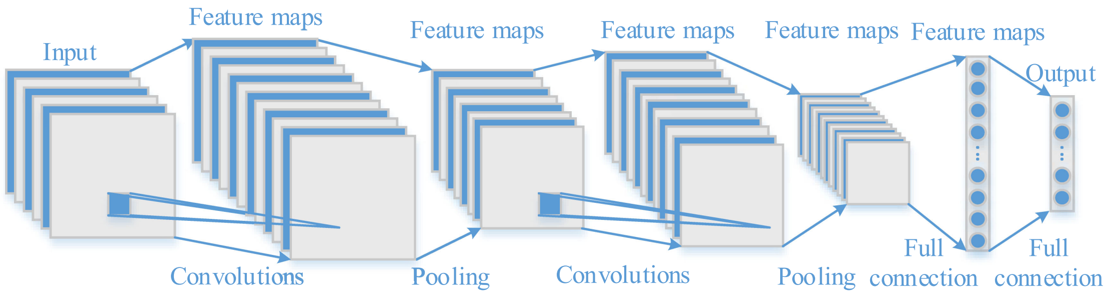


Fig. 1 The basic CNN structure.

machine (SVM), have been proposed to discriminate different motor imagery tasks and then predict human intension-behaviors. For instance, Ma *et al.* [13] proposed an optimized SVM classifier using particle swarm optimization, which obviously improves the classification accuracy of EEG-based motor imagery tasks. Maksimenko *et al.* [14] applied ANN model to recognize motor imagery tasks with the time-frequency spectral of EEG signals. Although traditional methods [15] have made considerable achievements in the field of motor imagery EEG-based BCI, there is still a critical issue need to be addressed: the feature extraction process is conducted based on the experts' knowledge and expertise with regard to EEG signal analysis, which is time-consuming and laborious.

In recent years, deep learning, as a new research hotspot of machine learning, has been widely used in various research areas, especially in image recognition and audio processing. Since deep learning has the strong abilities of automatic feature extraction and classification, it has been proposed to address the above-mentioned issue. Different deep learning models, including long short-term memory (LSTM), autoencoder (AE), especially convolutional neural network (CNN), have been previously investigated in the research area of motor imagery EEG-based BCI. Wang *et al.* [16] proposed a LSTM-based EEG classification method used in motor imagery tasks, where one-dimension aggregate approximation is used to extract signal representation and LSTM is used to classify EEG signals with the assistance of channel weighting technique. CNN has been most extensively applied to motor imagery EEG-based BCI. Dai *et al.* [17] proposed a motor imagery EEG classification framework, which is consisted of a variational AE and CNN. The proposed framework outperformed the existing methods with a classification improvement of 3 percent. Tang *et al.* [18] designed a 5-layer CNN model to classify motor imagery tasks, comparative results with three different SVM methods demonstrate the superiority of the CNN-based methods. All these proved the applicability and potential of CNN models in the research area of motor imagery BCI. Nevertheless, existing CNN-based methods mostly used high-level features in the last fully-connected for final EEG signal classification, which might loss some detailed and precise information in the low-level features. Recently, multi-scale features have been proved to be more robust and beneficial to improve the classification accuracy in many applications, such as medical image segmentation, face recognition. For example, Huang *et al.* [19] proposed a deep residual segmentation network of multi-scale feature fusion to segment IVOCT lumen contours. Experimental results showed that the proposed network can better extract the contour details, which enhance the robustness and accuracy of the lumen contour segmentation methods. Liu *et al.* [20] proposed a multi-scale CNN which contains deep and shallow neural networks for person identification, which can learn more robust and invariant features under complex scenes. Motivated by these, this paper proposes a novel motor imagery classification method using STFT and multi-scale CNN to predict the human intention-behaviors. First, STFT is introduced to convert raw EEG signals into time-frequency images. Then, a multi-scale CNN model is designed for extracting the local and global features of these images and classifying them. The performance of the proposed method is validated on the benchmark data set 2b used in the BCI contest IV. Experimental results show that the proposed method is very effective in human intention-behavior prediction, which is obviously superior to traditional methods including SVM, ANN, and stacked AE [21].

The remainder of this paper is organized as follows. The theoretical background of STFT and CNN is introduced in the Section 2. The proposed human intension-behavior prediction method is presented in the Section 3. The experimental results and analysis are given in the Section 4. Finally, we conclude this paper and discuss our future work in the Section 5.

2 Theoretical Background

2.1 STFT

STFT is one of the most popular signal processing methods proposed by Gabor in 1946 and has been widely used to analyze non-linear and non-stationary signal. It can specify the phase and magnitude of a raw signal varying with time and frequency. In general, the STFT of the raw signal is computed as follows: first, the longer raw signal is divided into several shorter signal segments; Then, the Fourier transform is separately computed on each signal segment. Finally, the Fourier spectrums obtained from all the shorter segments form a two-dimensional time-frequency spectrum of the raw signal. In mathematics, STFT is calculated as:

$$\mathrm{X}(\tau,\omega) = \int_{-\infty}^{+\infty} x(t)w(t-\tau)e^{-i\omega t}dt, \tag{1}$$

where $x(t)$ is the raw signal to be transformed, $w(t)$ is the window function which only has limited nonzero points, $\mathrm{X}(\tau,\omega)$ is the Fourier transform of $x(t)w(t-\tau)$. As τ changes, the window function shifts on the timeline. Compared with the Fourier transform, STFT has an important

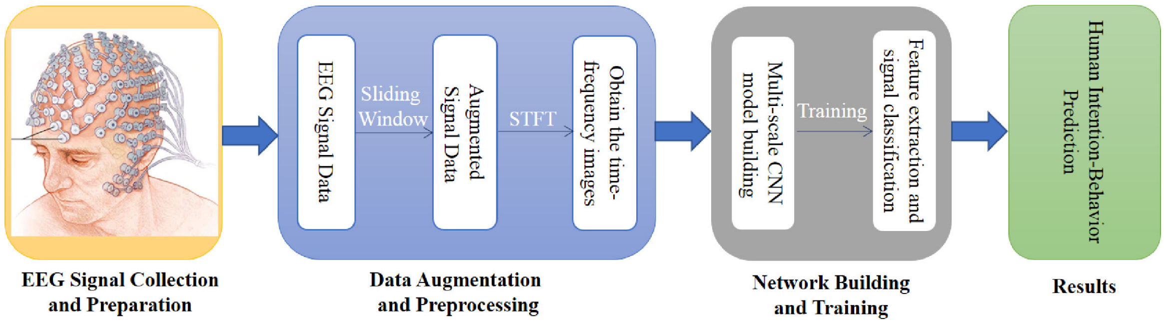


Fig. 2 The detailed procedure of the proposed method.

advantage that it can reflect the changes of signal frequency components with the change of time.

2.2 CNN

CNN is a kind of feedforward neural networks with convolution operation and has deep structure. It has three important characteristics, which are local receptive field, weight sharing, and pooling. In essence, a CNN is an input-output mapping network, which can learn the mapping relationship between input and output without any precise mathematical expression. The basic structure diagram of CNN is shown in the Fig. 1, it can be divided into three parts: input layer, output layer and hidden layer. The hidden layer can be subdivided into convolutional layer, pooling layer and full connection layer. To be specific, the input layer receives an image and propagates it to the hidden layer; The convolutional layer further extracts the feature maps of the input image with a set of convolutional filters; The pooling layer downsamples the extracted features of convolutional layer in order to reduce its dimension; The full connection layer connects its every neuron with every neuron in another layer; The output layer computes the class probability and network loss. Currently, there are many well-known and widely-used CNN architectures, such as LeNet-5, AlexNet, VGGNet, ResNet, they have achieved great success in various tasks, including image processing, speech recognition, medical image analysis and so on. In this paper, the multi-scale CNN model is designed and optimized based on the famous LeNet-5 CNN architecture.

3 Proposed Method

The procedure of the proposed method is shown in the Fig. 2, which consists of five phases, including EEG signal data collection, data preparation and augmentation, data preprocessing, multi-scale CNN model building and training, and intention-behavior prediction.

3.1 Data Collection

In this paper, the data set used to verify the proposed method is the BCI competition IV Dataset 2b [22]. The data set consisted of EEG signal data collected from nine volunteers, all the volunteers (subjects) were right-handed and none of the volunteers had any vision problems. The acquisition frequency of EEG signal data in this experimental data set is 250 Hz, and three electrodes (i.e., C3, Cz, C4) are utilized to record EEG signals [23]. A bandpass filter with the allowable pass frequency between 0.5 Hz and 100 Hz is utilized to eliminate signal noise. Each of the volunteers sat in an armchair, their eyes fixated on a screen a meter away from them. All nine volunteers were required to perform two completely different MI tasks, which are imagining left-hand and right-hand movements. Each volunteer was required to complete five sessions, all the EEG signal data is collected with feedback except the first two sessions. The detailed data collection process is illustrated in the Fig. 3. As for each collection from the first two sessions, the researchers employed a fixation cross and a warning tone to suggest the start of signal collection. Then, a visual cue is appeared for 1.25 seconds in order to indicate which motor imagery task the volunteers should perform. Finally, the volunteers imagine the corresponding motor imagery task. The timing of the remaining three sessions is similar to that of the first two sessions.

3.2 Data Preparation and Augmentation

Only the EEG signal data from the first three sessions of the BCI competition IV Dataset 2b is prepared for the method verification in this paper. To train the multi-scale model with more training data in the proposed method, we augment the original data set and obtain a much larger data set. This is because more training data will contribute to the model training and avoid the model overfitting. Here, the 3s to 7s interval of the EEG signals is considered for model training and testing. Motivated by the idea of data expansion method for time-series signals, the detailed data argumentation process is designed as follows: First, a signal segment with the length of

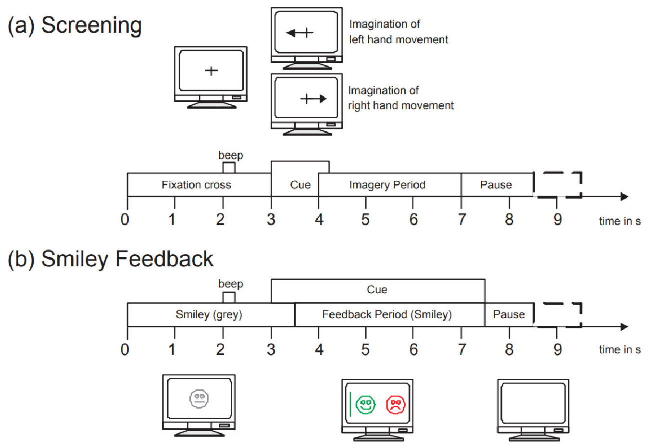


Fig. 3 The detailed data collection process: (a) The first two sessions and (b) the other three sessions.

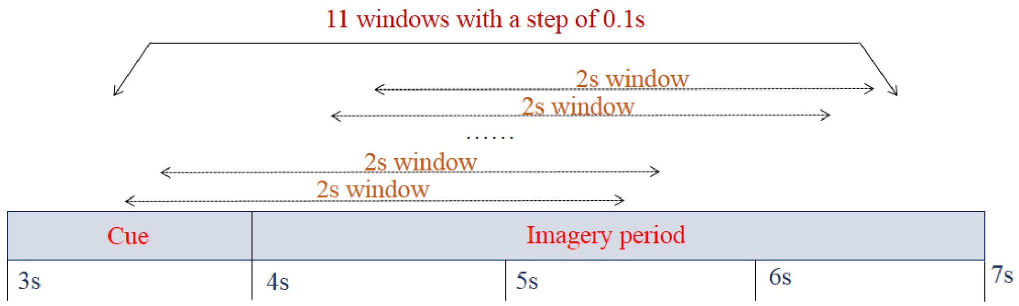


Fig. 4 The data argumentation process.

3s is randomly extracted from each raw signal. Then, the signal segment is augmented using a sliding window with step of 0.1s. Finally, 11 smaller signal segments with the length of 2s are obtained for each EEG trail. The data argumentation process is illustrated in the Fig. 4.

3.3 Data Preprocessing

In this paper, we convert raw EEG signals into time-frequency images by the STFT signal preprocessing method. It has been proven that when a person performs motor imagery tasks, event related desynchronization (ERD) and event related synchronization (ERS) [24], [25], [26], [27], [28] phenomenon will be caused. To be specific, the energy of mu band (8~13 Hz) in the motor cortex of the brain decreases, which is called ERD and represents the amplitude of the signals from the related electrodes decreases. At the same time, the energy of beta band (13–30 Hz) increases, which is called ERS [29] and represents the amplitude of the signals from the related electrodes increases. On the other hand, it should be noted that the signals at the related electrodes affected by the hand movement MI tasks are those from C3, C4, Cz electrodes. Therefore, in order to remove the intervention of other band information, only frequency bands with obvious ERD/ERS phenomena are extracted for use and analysis in this paper. Nevertheless, the frequency range of mu and beta bands adopted in this paper is not entirely consistent with the range found in [30]. We select the frequency range of 4-14 Hz for the mu band and the frequency range of 16-32 Hz for the beta band. This is because we found that the selected frequency ranges provide a better EEG signal classification performance through several comparative experiments. In this paper, the specific process of EEG signal data preprocessing is as follows: 1) Construct time-frequency image through conducting STFT on raw EEG signal: The STFT is applied to each 2s EEG signal segment with 500 signal points, the size of the sliding window is set as 64 and the overlap between two sliding windows is set as 50. Then, the size of the resulting STFT spectrogram image is 257×32; 2) Extract mu and beta band information from the converted spectrogram image: The size of the extracted mu band image is 20×32, while the size of the extracted beta band image is 33×32.To ensure that the mu band and beta band information have the same impact on the final classification result, the size of the extracted beta band image is adjusted to be the same as that of the extracted mu band image. Then, the size of the extracted beta band image is also 20×32. Ultimately, the extracted mu band image and beta band image are combined to constitute an image whose size is 40×32; 3) Form the final input image: The EEG signals obtained from C3, C4, and Cz electrodes are converted to three time-frequency images, respectively. They are formed to an image with the size of $40 \times 32 \times 3$, which is the final input to the designed multi-scale CNN model. The converted time-frequency images of nine subjects are shown in the Fig. 5. It can be found that the converted images of different motor imagery tasks of the same subject are different from each other, showing the effectiveness of the data processing method.

3.4 Multi-Scale CNN Model Building and Training

The multi-scale CNN model applied in this paper has seven layers, which are one input layer, two convolutional layers, one max-pooling layer, one multi-scale layer, one full connection layer and one softmax output layer, the structure of it is given in the Fig. 6. The input layer in the multi-scale CNN model is fed with a time-frequency image with the size of $40 \times 32 \times 3$ after EEG signals are preprocessed by STFT. The input image is first convolved with a set of convolutional kernels to extract features, and the mathematical expression of convolutional operation is defined as follows:

$$x_i^j = \psi\left(\sum\nolimits_k x_{i-1}^k * w_i^{j,\,k} + b_i^j\right), \tag{2}$$

where x_{i-1}^k and x_i^j denote the kth feature map in the $i-1$th layer and the jth feature map in the ith layer, respectively; $w_i^{j,k}$ denotes the convolutional kernel between the x_{i-1}^k feature map and the x_i^j feature map; b_i^j denotes the jth bias; $\psi()$ denotes the activation function. In this paper, the ReLU function is adopted, which is defined as follows:

$$\psi\left(x_i^j\right) = \max\left(0,\ x_i^j\right). \tag{3}$$

Then, the convolutional layer is followed by a pooling layer in order to reduce the number of CNN training parameters, and the mathematical expression of max-pooling operation used in this paper is defined as follows:

$$x_i^j(m,\ n) = \max\left(x_i^j(m',\ n')\right), \tag{4}$$

where $m \le m' < m+s$, $n \le n' < n+s$, and s is the window size of the pooling operation. (m, n) is the coordinate point of the x_i^j feature map after pooling and $(m',\ n')$ is the coordinate point of the x_i^j feature map before pooling.

After two convolutional and one pooling layers, the high-level features are extracted. Currently, most existing studies only take the high-level features in the last convolutional or

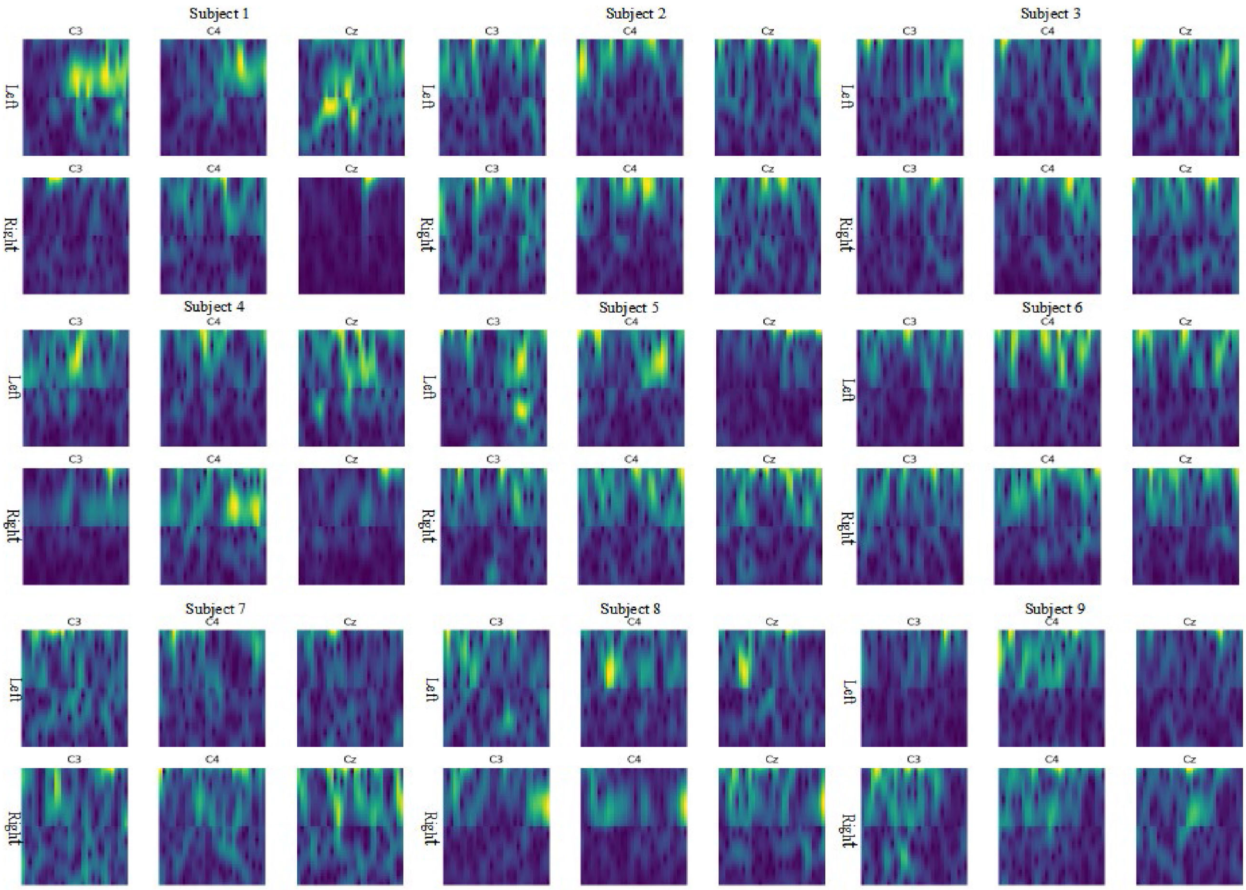


Fig. 5 The data preprocessing results for 9 subjects.

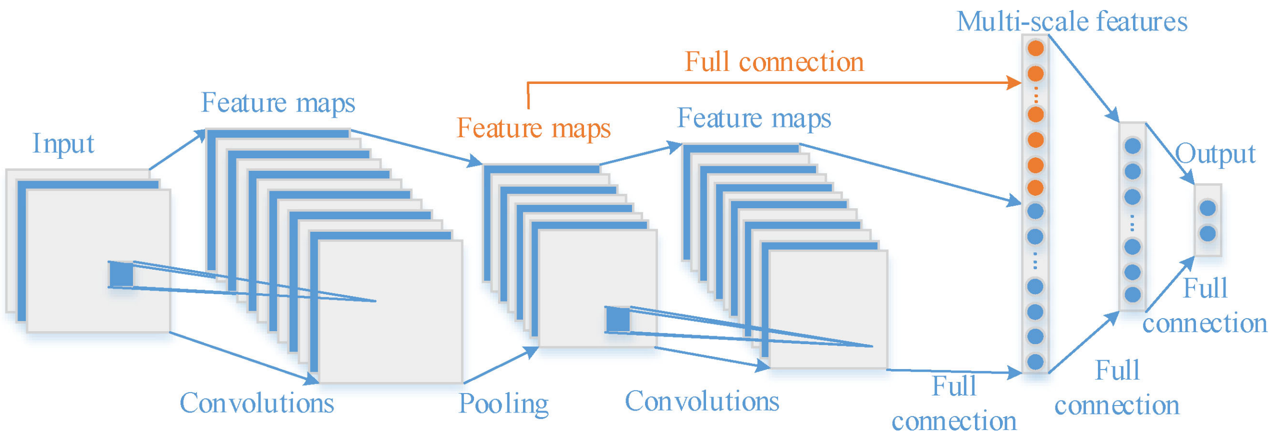


Fig. 6 The structure of the built multi-scale CNN.

pooling layer of CNN for EEG signal classification. However, it has been found that the high-level features may loss some detailed information and precision in the low-level features. Therefore, in this paper, the high-level features in the last convolutional layer and the low-level features in the last pooling layer are flattened together to form a more robust multi-scale features. The extracted multi-scale features have enough global and local information and will benefit the final classification, which are extracted as follows:

$$x_{ms} = \psi\left(\sum\nolimits_k x_c^k * w_{ms}^{j,\,k} + \sum\nolimits_k x_p^k * w_{ms}^{q,\,k} + b_{ms}\right), \qquad (5)$$

Where $x_c^k, w_{ms}^{j,\,k}$ and $x_p^k, w_{ms}^{q,\,k}$ denote the feature maps and weights in the last convolutional layer and in the last pooling layer. The multi-scale layer is then fully-connected to the softmax output layer the full connection layer passing through the full connection layer. In the output layer, the softmax function is used to calculate the probability that the input image belongs to different motor imagery classes, which is calculated as follows:

$$s_c = \frac{e^{\widetilde{s_c}}}{\sum_{c=1}^{C} e^{\widetilde{s_c}}}, \qquad (6)$$

TABLE 1
The Average Accuracy Results of the Proposed Method and the Compared Methods

	Accuracy on different subjects									
Methods \ Subject	1	2	3	4	5	6	7	8	9	Average
ANN	65.2	45.1	54.2	87.2	72.2	59.7	66.1	62.2	59.2	63.5
SVM	68.4	54.2	58.7	89.4	77.5	68.5	67.8	68.9	62.6	68.4
SAE	61.0	47.5	45.0	81.3	62.8	47.8	53.8	63.8	56.5	57.7
Multi-Scale CNN	74.3	61.8	66.5	91.5	79.5	74.8	71.4	73.3	72.0	73.9

Where $\mathrm{c} = 1, 2, \ldots, C$ represent different motor imagery classes, and C represents the number of motor imagery classes, $\widetilde{s_c}$ and s_c represent the original output and the probability output of the input image, respectively. The cross-entropy loss used for training the designed multi-scale CNN, which is calculated as follows:

$$\mathcal{L}(y,\ s) = -\sum\nolimits_{c=1}^{C} y_c log s_c, \tag{7}$$

where y_c is 1 only when the output is consistent with the true label; otherwise, it is 0. The CNN can be trained to converge on the BCI competition IV Dataset 2b through minimizing the loss error using the stochastic gradient descent method.

3.5 Intention-Behavior Prediction

After the multi-scale CNN model training, the STFT spectrograms of the raw EEG signals can be input to the trained model and the human intention-behavior can be predicted. Therefore, the proposed method can help people, especially the disabled and the elderly people, to control devices only using brain signals while without any physical movements.

4 Experimental Results

In this section, the performance of the proposed method is verified on the BCI competition IV dataset 2b. The experiments include three parts: first, the experimental results of the proposed method, both in terms of accuracy and kappa value, are given and analyzed. In addition, the proposed method is compared with traditional methods and the comparative results are also given. Finally, an additional analysis is conducted to examine the effect of different periods in the EEG signal on the classification accuracy of the proposed method. All the algorithms are written in Python with Keras deep learning framework, and all the experiments are conducted on a high-performance computer with a 256 GB SSD, Intel 12-core 3.5-GHz CPU, a GTX1080TI GPU, and 96 GB RAM.

4.1 Accuracy Results

The EEG trails from the first three sessions of the BCI competition IV Dataset 2b are selected for experimental evaluation, there are totally 400 EEG trails for each subject. After data augmentation and preprocessing of the EEG trails, the converted images of each subject are divided into two parts, i.e., the training dataset and the test dataset. The training dataset is used to train the multi-scale CNN model and the test dataset is used to verify the accuracy performance of the proposed method. The accuracy performance is separately validated on nine subjects, and all the training and testing processes for each subject are conducted ten times. The average accuracy results of nine subjects are presented in the Table 1. In addition, in order to verify the superior performance of the proposed method, the proposed method is also compared with traditional machine learning methods including support vector machine (SVM), artificial neural network (ANN), where the discriminative features are manually extracted for them, and the standard deep learning method, which is stacked auto-encoder (SAE), the comparison results are also give in Table 1. It can be found in the Table 1 that the proposed method provides a higher classification accuracy for all nine subjects compared with other methods. Specifically, it achieves an average accuracy of 73.9 percent, which obtains the accuracy improvements of 10.4, 5.5, 16.2 percent compared with the ANN, SVM and SAE, respectively. The accuracy results prove that the proposed method has excellent feature extraction and strong signal classification abilities.

4.2 Kappa Value Results

In this subsection, the kappa values of these methods are compared in the Table 2. The kappa value is a statistical measure that compares the classification accuracy with the random accuracy in order to remove the interference of random classification. It is calculated as follows:

$$kappa = \frac{accuracy - random}{1 - random}. \tag{8}$$

Here the value of random is set as 0.5 because the binary motor imagery classification task is conducted in this experiment. It can be seen from the Table 2 that the average kappa value of the proposed method is 0.48, and the kappa values of the ANN, SVM, and SAE are 0.27, 0.37 and 0.15, respectively. Therefore, it can be concluded that the proposed method obtains obvious improvement and provides more accurate EEG signal classification results.

4.3 Additional Analysis

In this subsection, in order to find the most suitable period in the EEG signals used for motor imagery task classification, an additional analysis on the effect of different time

TABLE 2
The Kappa Value of Nine Subjects

Methods \ Subjects	1	2	3	4	5	6	7	8	9	Average
	kappa value of different subjects									
ANN	0.30	-0.10	0.08	0.74	0.44	0.19	0.32	0.24	0.18	0.27
SVM	0.37	0.08	0.17	0.79	0.55	0.37	0.36	0.38	0.25	0.37
SAE	0.22	-0.05	-0.10	0.63	0.26	-0.04	0.08	0.28	0.13	0.15
Multi-Scale CNN	0.49	0.24	0.33	0.83	0.59	0.50	0.43	0.47	0.44	0.48

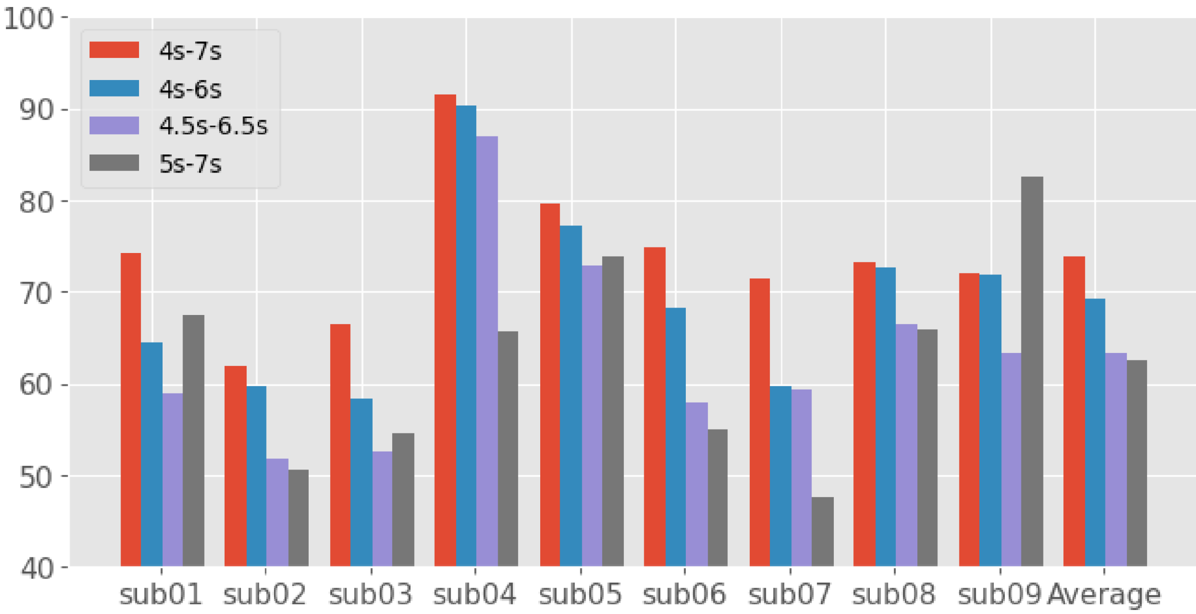


Fig. 7 The accuracy comparison results of the proposed method for different time periods.

periods on the classification accuracy is conducted. The EEG signal data from four different time periods of 4s-7s, 4s-6s, 4.5s-6.5s, 5s-7s are selected to verify the accuracy performance of the proposed method, respectively. The accuracy comparison results of the proposed method for different time periods are given in the Fig. 7. It can be found that the proposed method using the EEG signal data from the time period of 4s-7s achieve the highest classification accuracy on all the subjects except the subject 9. On the other hand, the proposed method using other periods obtains the similar accuracy. Thus it can be concluded that more experimental data obtained through the data augmentation process will benefit the model training and testing, thereby improving the accuracy of human intention-behavior prediction.

5 Conclusions and Future Work

In this paper, a human intension-behavior prediction method based on STFT and multi-scale CNN is proposed to classify motor imagery EEG signals, which is highly effective in helping the disabled and elderly people. The main contributions and innovations of this paper are as follows: 1) introducing a data argumentation method to enlarge the original motor imagery EEG dataset, which can improve the model training performance and avoid the model overfitting problem. 2) designing a STFT-based data preprocessing method based on the ERD/ERS phenomena, which can effectively extract motor imagery task-related information from EEG signals. At the same time, the EEG signals are converted to images for the multi-scale CNN model. 3) proposing a multi-scale CNN model to extract the robust features containing both global and local information, which can improve the accuracy performance of the motor imagery EEG signal classification. 4) validating the performance of the proposed method on the BCI competition IV Dataset 2b, the average accuracy of the proposed method is 73.9 percent, which is superior to traditional methods including SVM, ANN and SAE.

Although this paper has made some considerable progress, there are still two challenges need to be addressed in the future. 1) even though the multi-scale CNN model has improved the classification accuracy, the accuracy of motor imagery tasks classification is still not very satisfying. 2) the cross-subject motor imagery task has not been considered in this paper. Therefore, in future work, we will focus on the following two aspects: 1) investigating more effective and advanced models to further improve the classification accuracy; 2) focusing more attention on the cross-subject motor imagery tasks.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62002304 and the Fundamental Research Funds for the Central Universities under Grant 22120190211.

REFERENCES

[1] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R.Wolpaw, "BCI2000: A general-purpose brain-computer interface (BCI) system," in *IEEE Trans Biomed Eng*, vol. 51, no. 6, pp. 1034–1043, Jun, 2004. W.-K. Chen, Ed., in *Linear Networks and Systems*. Belmont, Calif.: Wadsworth, pp. 123–135, 1993.

[2] Y. Chae, J. Jeong, and S. Jo, "Toward brain-actuated humanoid robots: Asynchronous direct control using an EEG-based BCI," *IEEE Trans. Robot.*, vol. 28, no. 5, pp. 1131–1144, Oct. 2012.

[3] T. O. Zander *et al.*, "Enhancing human-computer interaction with input from active and passive brain-computer interfaces," in *Brain-Computer Interfaces*. London, U.K.: Springer, 2010. pp. 181–199.

[4] J. Kevric and 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, 2017.

[5] D. R. Edla *et al.*, "Classification of EEG data for human mental state analysis using random forest classifier." *Procedia Comput. Sci.*, vol. 132, pp. 1523–1532, 2018.

[6] A. Al-Nafjan *et al.*, "Review and classification of emotion recognition based on EEG brain-computer interface system research: A systematic review." *Appl. Sci.*, vol. 7, no. 12, 2017, Art. no. 1239.

[7] M. Mohammadpour, S. M. R. Hashemi, and N. Houshmand, "Classification of EEG-based emotion for BCI applications," in *Proc. IEEE Artif. Intell. Robot.*, 2017, pp. 127–131.

[8] B. Kaur, D. Singh, and P. P. Roy, "EEG based emotion classification mechanism in BCI," *Procedia Comput. Sci.*, vol. 132, pp. 752–758, 2018.

[9] C. Park, C. C. Took, and D. P. Mandic, "Augmented complex common spatial patterns for classification of noncircular EEG from motor imagery tasks," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 22, no. 1, pp. 1–10, Jan. 2014.

[10] Y. Wang, Y. T. Wang, and T. P. Jung, "Translation of EEG spatial filters from resting to motor imagery using independent component analysis," *PLoS One*, vol. 7, no. 5, 2012, Art. no. e37665.

[11] B. Xu *et al.*, "Wavelet transform time-frequency image and convolutional network-based motor imagery EEG classification," *IEEE Access*, vol. 7, pp. 6084–6093, 2019.

[12] A. Zabidi, W. Mansor, Y. K. Lee, C. W. Fadzal, "Short-time fourier transform analysis of EEG signal generated during imagined writing," in *Proc. IEEE Int. Conf. System Eng. Technol.*, 2012, pp. 1–4.

[13] M. Yuliang *et al.*, "Classification of motor imagery EEG signals with support vector machines and particle swarm optimization," *Comput. Math. Methods Med.*, 2016, pp. 1–8, 2016.

[14] V. A. Maksimenko *et al.*, "Artificial neural network classification of motor-related eeg: An increase in classification accuracy by reducing signal complexity," *Complexity*, vol. 2018, 2018, Art. no. 9385947.

[15] M. Hamedi, S. H. Salleh, and A. M. Noor., "Electroencephalographic motor imagery brain connectivity analysis for BCI: A review," *Neural Comput.*, vol. 28, no. 6, pp. 1–43, pp. 2016.

[16] P. Wang, A. Jiang, X. Liu, J. Shang, and L. Zhang, "LSTM-Based EEG classification in motor imagery tasks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2086–2095, Nov. 2018.

[17] M. Dai, D. Zheng, R. Na, S. Wang, and S. Zhang, "EEG classification of motor imagery using a novel deep learning framework," *Sensors*, vol. 19, no. 3, pp. 551, 2019.

[18] 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, 2017.

[19] C. Huang *et al.*, "A deep segmentation network of Multi-scale feature fusion based on attention mechanism for IVOCT lumen contour,"*IEEE/ACM Trans. Comput. Biol. Bioinf.*, to be published, doi: 10.1109/TCBB.2020.2973971.

[20] X. Liu *et al.*, "Multi-view multi-scale CNNs for lung nodule type classification from CT images," *Pattern Recognit.*, vol. 77, pp. 262–275, 2018.

[21] A. Narang *et al.*, "Classification of EEG signals for epileptic seizures using levenberg-marquardt algorithm based multilayer perceptron neural network," *J. Intell. Fuzzy Syst.*, vol. 34, no. 3, pp. 1669–1677, 2018.

[22] A. Kübler *et al.*, "Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface," *Neurology*, vol. 64, no. 10, pp. 1775–1777, 2005.

[23] R. Leeb, C. Brunner, and G. -MÃijller-PutzR, 2007. Accessed : Jun. 25, 2019. [Online]. Available: http://www.bbci.de/competition/iv/

[24] G. Pfurtscheller., "EEG event-related desynchronization (ERD) and synchronization (ERS)," *Electroencephalogr. Clin. Neurophysiol.*, vol. 1, no. 103, 1997, Art. no. 26.

[25] B. Graimann *et al.*, "Visualization of significant ERD/ERS patterns in multichannel EEG and ECoG data," *Clin. Neurophysiol.*, vol. 113, no. 1, pp. 43–47, 2002.

[26] M. Pesonen *et al.*, "Brain oscillatory 1–30 hz EEG ERD/ERS responses during the different stages of an auditory memory search task," *Neurosci. Lett.*, vol. 399, no. 1-2, pp. 45–50, 2006.

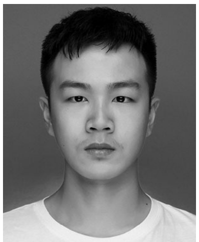
[27] A. Fink *et al.*, "EEG alpha band dissociation with increasing task demands," *Cognit. Brain Res.*, vol. 24, no. 2, pp. 252–259, 2005.

[28] E. Formaggio *et al.*, "Time–frequency modulation of ERD and EEG coherence in robot-assisted hand performance," *Brain Topogr.*, vol. 28, no. 2, pp. 352–363, 2015.

[29] C. Andrew and G. Pfurtscheller, "Lack of bilateral coherence of post-movement central beta oscillations in the human electroencephalogram," *Neuro. Lett.*, vol. 273, no. 2, pp. 89–92, 1999.

[30] G. Pfurtscheller and F. L. Da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, 1999.

ChenXi Huang is an assistant professor with the School of Informatics, Xiamen University. His research interests include image processing, image reconstruction, data fusion, 3D visualization, and machine learning.



YuTian Xiao is currently working toward the BSc degree in Xiamen University, Shanghai, China. His research interests include image processing, and machine learning.

Gaowei Xu is currently working toward the PhD degree in the College of Electronics and Information Engineering at Tongji University, Shanghai, China. His research interests include machine learning, deep learning, transfer learning and their applications in intelligent predictive maintenance, and biomedical signal processing.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/csdl.