

Motor Imagery EEG Recognition Based on Generative and Discriminative Adversarial Learning Framework and Hybrid Scale Convolutional Neural Network

Xiaoyuan Dang, Guorui Liu, Xianlun Tang, Shifei Wang, Tianzhu Wang, and Mi Zou

Abstract—Motor imagery (MI) Electroencephalogram (EEG) signal recognition is an important part of brain computer interface (BCI). Nowadays, deep learning has been successfully applied to MI signal classification. However, due to the expensive equipment of EEG acquisition and the limitation of acquisition site that gives rise to the EEG data scarcity, and thus impacts deep learning model classification. Therefore, this paper proposes a novel EEG time-frequency transition method, which uses the continuous wavelet transform adaptive generate EEG time-frequency map. Generative and discriminative adversarial learning framework (GADALF) is proposed to enhance data. The correlation analysis of EEG time-frequency map is carried out by combining event related synchronization (ERS) with event related desynchronization (ERD). Based on the analysis results, the hybrid scale convolutional neural network (HSCNN) is proposed by improving the convolutional neural network. This network combines multi-dimensional convolution with channel attention mechanism to decode the hidden energy information of the α and β bands for EEG time-frequency map classification. The results showed that the proposed method can achieve higher recognition rates on the laboratory-collected dataset and the dataset BCI competition IV 2b compared with other methods. Finally, this paper designs an intelligent wheelchair control system based on a BCI that

verifies the effectiveness of the system by applying the proposed method to the system through online experiments.

Index Terms—Continuous wavelet transform, Generative and discriminant adversarial framework, Convolutional neural network, EEG signal recognition, Intelligent wheelchair control

I. INTRODUCTION

Brain-computer interface (BCI) is a system that realizes direct communication between the brain and external devices [1]. It collects Electroencephalography (EEG) signals through electrode devices, and then completes the control of external devices through the control system, such as smart wheelchairs [2] and robotic arms [3] are used in BCI systems.

Motor imagery (MI) signal [4] is a kind of spontaneous brain point signal. When people imagine left-hand or right-hand movements, event-related synchronization (ERS) and event-related desynchronization (ERD) respectively occur in the ipsilateral and opposite sides of the brain somatosensory cortex [5]. This phenomenon has been effectively applied to the decoding of EEG features [6].

Due to the nonlinear characteristics of EEG signals, researchers have used traditional methods such as CSP [7], EMD [8], PCA [9] and FFT [10] in the feature extraction of EEG signals. In recent years, deep learning methods such as convolutional neural networks [11] have also achieved good results in EEG recognition. However, expensive EEG acquisition equipment, the constraints of the acquisition site, and fatigue of acquisition personnel in the process of acquisition experiment all lead to the scarcity and low quality of EEG data, which affect the classification of the deep learning model. Therefore, data enhancement is needed to address the above problems, common data augmentation methods are geometric transformations such as translation, flipping, scaling, and cropping.

To address the questions posed by Y. R. Tabar et al.[12], this paper proposes an EEG time-frequency information conversion method based on continuous wavelet transform (CWT), which does not consider the window size in short-time fourier transform (STFT), and can adaptively generate EEG time-frequency maps. On this basis, an adversarial learning framework is proposed to generate a large number of artificial EEG time-frequency maps, and

Manuscript received January 28, 2022; revised September 20, 2022. This work is financially supported by the National Nature Science Foundation of China (61673079) and the Science and Technology Project of Chongqing Education Committee(KJZD-K202202401, KJZD-M202200603).

Xiaoyuan Dang is an associate professor of School of Intelligent Engineering, Chongqing College of Mobile Communication, Chongqing, 401520, China (e-mail: dxy831110@126.com).

Guorui Liu is a postgraduate student of Chongqing Key Laboratory of Complex Systems and Bionic Control, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China (corresponding author to provide phone: 152-2317-6282; e-mail: 705606221@qq.com)

Xianlun Tang is a professor of Chongqing Key Laboratory of Complex Systems and Bionic Control, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China. (e-mail: tangxl@cqupt.edu.cn)

Shifei Wang is a postgraduate student of Chongqing Key Laboratory of Complex Systems and Bionic Control, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China. (e-mail: 761736434@qq.com)

Tianzhu Wang is a postgraduate student of Chongqing Key Laboratory of Complex Systems and Bionic Control, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China. (e-mail: 2018212781@stu.cqupt.edu.cn)

Mi Zou is an associate professor of Chongqing Key Laboratory of Complex Systems and Bionic Control, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China. (e-mail: zoumi@cqupt.edu.cn)

then select high-quality data from the generated EEG time-frequency maps through five indices to solve the scarcity of EEG data. Finally, a hybrid scale convolutional neural network is proposed, which uses hybrid scale convolution combined with attention mechanism to extract the features of EEG time-frequency maps and decode the energy information of α and β bands. The proposed method is applied to the laboratory-collected dataset and the dataset BCI competition IV 2b, and the results show that the proposed method can achieve higher recognition rates than other methods.

Based on the proposed method, an intelligent prosthesis BCI system is designed to prove the feasibility of the system through online experiments, and this BCI system provides a new paradigm for the control of intelligent prosthesis.

The main highlights of this article are as follows:

- 1) The EEG time-frequency map generation method based on continuous wavelet transform is proposed in this paper, which makes the window size of the short-time Fourier transform irrelevant. The generated EEG time-frequency map can preserve various information of EEG signals in time domain, frequency domain.
- 2) An adversarial learning framework is proposed for data augmentation, and high-quality data are selected through five indices to complement the data manifold.
- 3) A network of hybrid scale convolution and attention mechanism ensemble proposed in this paper decodes the energy information of α and β bands for EEG time-frequency map classification.

II.CONSTRUCTION METHOD OF EEG TIME-FREQUENCY MAP

CWT has the properties of small-time width and concentrated frequency band, which is suitable for processing non-stationary signals such as EEG signals [13].

The continuous wavelet transform is defined as follows:

$$WT(\alpha, \tau) = \langle f(x), \varphi_{\alpha, \tau}(t) \rangle \tag{1}$$

$$= \alpha^{\frac{1}{2}} \int_{-\infty}^{+\infty} f(x) \varphi^* \left(\frac{t-\varphi}{\alpha} \right) dt \quad a > 0$$

Where $f(x)$ is the original data, and the basic wavelet $\varphi_{\alpha, \tau}(t)$ is Morl wavelet.

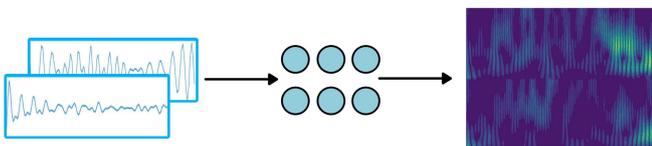


Fig. 1. Construction method of the EEG time-frequency map

CWT is used to construct the time-frequency map of EEG data. A group of EEG data is taken as an example in Fig.1. The specific steps are as follows:

- 1) Arrange the EEG data as $n \times m$ in the order of channels, where n is the number of channels, and m is the number of samples in channels.
- 2) Select channels. According to the sampling frequency of the EEG signal, the center frequency of the basic wavelet and the moving scale of the basic wavelet, the continuous wavelet transform of the EEG signal is carried out according to the sequence of channel, and

the coefficients of 8-32 parts which means α and β bands are reserved.

- 3) Construct the EEG time-frequency maps using the retention coefficient, and save them in folders according to the types of labels.

III.DATA ENHANCEMENT METHOD BASED ON GENERATIVE AND DISCRIMINATIVE ADVERSARIAL LEARNING FRAMEWORK

A.Generative network

In Fig. 2, a 5-layer generative network [15] is proposed to generate artificial EEG time-frequency maps.

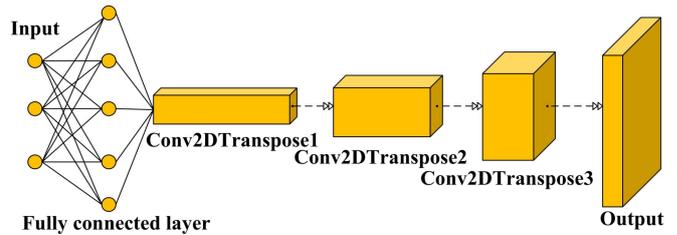


Fig. 2. Generative network

The 100-dimensional noise, which obeys a normal distribution, is provided to the input layer of the generative network. The second layer of the network consists of a fully connected layer and a reshape function to change the structure of the input noise. The last four layers of the network are transposed convolutional layers for up-sampling. The transposed convolutional layers double the size of the output tensor and reduce the number of channels until an artificial EEG time-frequency map is generated, which turns out to be consistent with the structure of the original EEG time-frequency map. The detailed information of the generative network is presented in Table I.

TABLE I
MAIN PARAMETERS OF GENERATIVE NETWORK

Layer name	Output shape	Filter size	Activation	Stride	Padding
Input	100×1×1	—	—	—	—
FC	4×4×256	—	Relu	—	—
Conv2D Transpose1	8×8	5×5(256)	Relu	2	Same
Conv2D Transpose2	16×16	5×5(128)	Relu	2	Same
Conv2D Transpose3	32×32	5×5(64)	Relu	2	Same
Conv2D Transpose4	64×64	5×5(3)	Tanh	2	Same

B.Discriminative network

In Fig.3, the discriminative network [15] is composed of a 5-layer deep convolutional neural network. The network is used to distinguish real EEG from those generated by the generative network.

In the discriminative network, RMSprop is used as the optimizer. The detailed information of the discriminative network is shown in Table II.

C.Objective function

Compared to the mutability of the Jensen-Shannon

TABLE II

MAIN PARAMETERS OF DISCRIMINATIVE NETWORK

Layer name	Output shape	Filter size	Activation	Stride	Padding
Input	64×64×3	—	—	—	—
Conv2D1	32×32	5×5(64)	2	Same	Relu
Conv2D2	16×16	5×5(128)	2	Same	Relu
Conv2D3	8×8	5×5(256)	2	Same	Relu
Conv2D4	4×4	5×5(512)	2	Same	Relu
Output	1	—	—	—	—

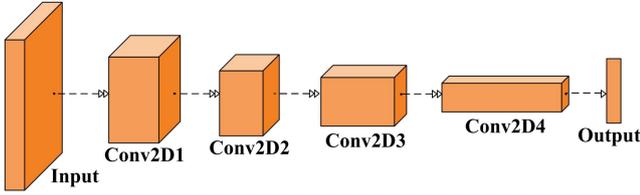


Fig. 3. Discriminative network

divergence in the original generative adversarial network (GAN) [16], the smoothness of the earth mover's distance in Wasserstein GAN [17] to effectively solve the problem of gradient disappearance. The definition of earth mover's distance is as follows:

$$W(p_{data}, p_g) = \inf_{\gamma \in \Pi(p_{data}, p_g)} E_{(x,y) \sim \gamma} [\|x - y\|] \quad (2)$$

Where p_{data} means real data distribution, p_g means fake data distribution, $\gamma \in \Pi(p_{data}, p_g)$ means the distribution of all possible joint combinations of $\gamma(x,y)$ whose marginals are respectively p_{data} and p_g , $\gamma(x,y)$ means the distance moved by distribution p_{data} to distribution p_g .

In order to reduce the complexity of solving the earth mover's distance by adopting the coefficient clipping method, which is expressed as:

$$\begin{aligned} \min_{G} \max_{D} E_{x \sim p_{data}} [D(x)] - E_{x \sim p_g} [D(G(x))] \\ s.t. w \in [-0.01, 0.01] \end{aligned} \quad (3)$$

Where $D(x)$ denotes discriminative network, $G(x)$ represents the generative network, w means the weight of the network, In order to meet the stability of the network training, this paper conducts weight initialization in the transposed convolutional layer and the convolutional layer, which follows a normal distribution with a mean of 0 and a standard deviation of 0.01. The purpose of this action is to prevent the weights of the network from dropping too fast that could result in disappearance of the model gradient when the coefficients are clipped.

D. Data enhancement algorithm for EEG time-frequency map

Algorithm: Generative and discriminative adversarial learning framework

Input: EEG time-frequency maps x_n , generative network G , the noise vector of 100 dimensions obeys the normal distribution z , discriminator network D , learning rate $lr = 0.0005$, cropping coefficient $e = 0.01$, batch size $bs = 32$, number of trainings $c_1 = 12$, $c_2 = 5$.

Output: artificial EEG time-frequency maps

1. weight initialization: $G, D \sim N(0, 0.01^2)$

2. For $t_1 = 0, \dots, n$ do

Training D

If $t_1 \leq 10$ or $(t_1 + 1) \% 100 = 0$ obtain $c = c_1$

Else $c = c_2$

For $t_1 = 0, \dots, c$ do

Sample bs batch form x_n obtain $[x_n]_{i=1}^{bs}$

Collection bs batch form z obtain $[z]_{i=1}^{bs}$

Training objective function:

$$w_D \leftarrow \nabla_D \frac{1}{bs} \left[\sum_{i=1}^{bs} D(x_i) - D(G(z_i)) \right]$$

Optimizing network parameters:

$$w_D \leftarrow w_D + lr \cdot RMSprop(w_D, D_{w_D})$$

The coefficient of cropping:

$$w_D \leftarrow clip(w_D, -e, e)$$

End For

Training G

Collection bs batch form z obtain $[z]_{i=1}^{bs}$

Training objective function:

$$w_G \leftarrow \nabla_G \frac{-1}{bs} \left[\sum_{i=1}^{bs} D(G(z_i)) \right]$$

Optimizing network parameters:

$$w_G \leftarrow w_G + lr \cdot RMSprop(w_G, G_{w_G})$$

Output: artificial EEG time-frequency maps

End while

E. Quality Evaluation and Selection

- 1) Maximum mean difference (MMD) [18], inception score (IS) [19], mode score (MS) [20], and freshet initial distance (FID) [21] are used to evaluate the correlation between the generated artificial EEG maps and the difference of original EEG maps, and the trend of scores is utilized to judge the validity of the data augmentation experiment. IS and MS reflect the quality and diversity of generated samples, and the scores of these two indices are required to be gradually higher. FID and MMD reflect the distance between generated samples and real samples, and the scores of these two indices are required to be gradually lower.
- 2) On the premise of confirming the validity of the data augmentation experiment, the EEG maps generated by generative network when the loss value of the discriminative network is close to 0 are selected as the high-quality data. When the loss value of discriminant network is close to 0, the time-frequency data distribution of artificial EEG is closer to that of the original EEG time-frequency map.

IV. HYBRID SCALE CONVOLUTIONAL NEURAL NETWORK

A. Analysis of EEG time-frequency map

The ERS/ERD phenomenon of motor imagery signals appeared in C₃ channel and C₄ channel. Therefore, the EEG time-frequency maps of C₃ channel and C₄ channel constituted by 3 groups of EEG signals are arranged in Fig.4 to analyze.

After the construction of the EEG time-frequency map using continuous wavelet transform, the energy information of α band and β band in the original time domain EEG signal is transformed into three channels of the EEG time-frequency maps, thus the color depth of the EEG time-frequency map represents the power amplitude of α band and β band, and the

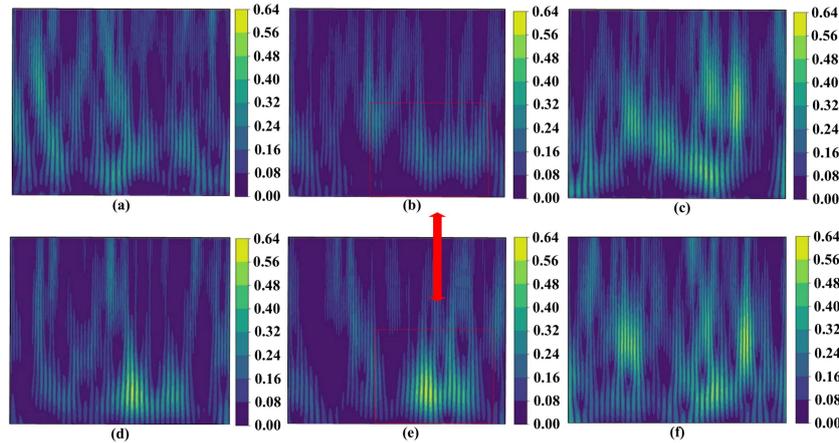


Fig. 4. (a) and (d) are EEG time-frequency maps composed of C3 channel and C4 channel of the same group of EEG data. (b) and (e) are EEG time-frequency maps composed of C3 channel and C4 channel of the same group of EEG data. (c) and (f) are EEG time-frequency maps composed of C3 channel and C4 channel of the same group of EEG data. On the left side of the EEG time-frequency map is the range of 8-32 Hz, which corresponds to the range of α band (8-13 Hz) and β band (13-32 Hz), and on the bottom of the EEG time-frequency map is the range of 0-1 second of EEG data.

two are positively correlated.

The comparison of the Fig.4(b) with Fig.4(e) suggests that the α band in Fig.4(e) is darker than Fig.4(b), indicating the power amplitude of C4 channel is higher than C3 channel, and the ERD phenomenon of EEG signals occurs in C4 channel. Therefore, speculated the EEG signals in this group are labeled right-handed motion imagery, and the inquiry confirms that the actual labels are consistent with the speculated labels.

According to the above analysis, the channel attention mechanism is perfect for this task, as it can highlight the important classification features of EEG time-frequency map, which can make the model have a higher recognition rate with many kinds of attention mechanisms, such as multi-layer attention [22], and semantic attention [23].

According to the above analysis, the channel attention mechanism is perfect for this task, as it can highlight the important classification features of EEG time-frequency maps, which can make the model have a higher recognition rate with many kinds of attention mechanisms, such as multi-layer attention [22], and semantic attention [23].

B. One-dimensional convolution layer

The one-dimensional (1D) convolution layer consists of two parallel networks. 1D convolution 1 uses the convolution kernel of $1 * e_1$ to extract the time domain information of the EEG time-frequency map, and 1D convolution 2 uses the convolution kernel of $e_2 * 1$ to extract the frequency domain information of the EEG time-frequency map. Channel attention mechanism 1 and channel attentions mechanism 2 further highlight the characteristic information of α and β bands in time domain and frequency domain. Max pooling layer 1 and max pooling layer 2 are used to reduce the number of parameters, retain effective information, and improve the training speed of the network. In order to capture the details of the EEG time-frequency map, 1D convolution 1 and 1D convolution 2 both adopt the form of small convolution kernel. In addition, with procedures above in place, feature concatenation is carried out. Specific parameters are shown in Table III.

In Fig.5, the channel attention module [24] is completed by three modules: information aggregation module, information

enhancement module, and information divergence module.

Information aggregation module: we assume that the feature F extracted by one-dimensional convolution, F contains three dimensions of length, width, and channel information, which are represented as $F \in R^{H*W*C}$. Perform information aggregation operation on the feature F , and the formula is as follows:

$$X_k = f_{IA}(F \sim C_k(i, j)) = \frac{\sum_{i=1}^H \sum_{j=1}^W (F \sim C_k(i, j))}{H \times W} \quad (4)$$

Where $F \sim C_k(i, j)$ means two-dimensional data points in the k th channel of feature F , $f_{IA}(F \sim C_k(i, j))$ means the global averaging operation on $F \sim C_k(i, j)$. X_k means the eigenvalue of the k th channel of feature F .

TABLE III
MAIN PARAMETERS OF 1D CONVOLUTIONAL LAYER

Layer name	Output shape	Kernel size	Stride	Padding
1Dconvolution1	64×64×8	1×3(8)	1	Same
1Dconvolution2	64×64×8	3×1(8)	1	Same
Channel attention 1	64×64×8	-	-	-
Channel attention 2	64×64×8	-	-	-
Max pooling1	32×32×8	2×2	1	Same
Max pooling2	32×32×8	2×2	1	Same
Concatenation	32×32×16	-	-	-

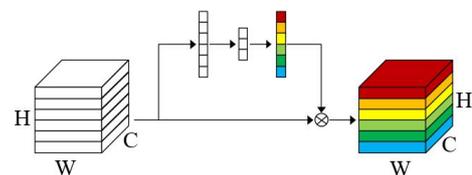


Fig. 5. The channel attention module

Information enhancement module: based on information aggregation, the three-dimensional feature becomes two-dimensional feature. The two fully connected layers are connected that change X_k to Y_k . The purpose is to reduce the degree of imitation assistance and enhance the distinguishability of features. The dimension of the first fully connected layer is b_1 , and the dimension of the second

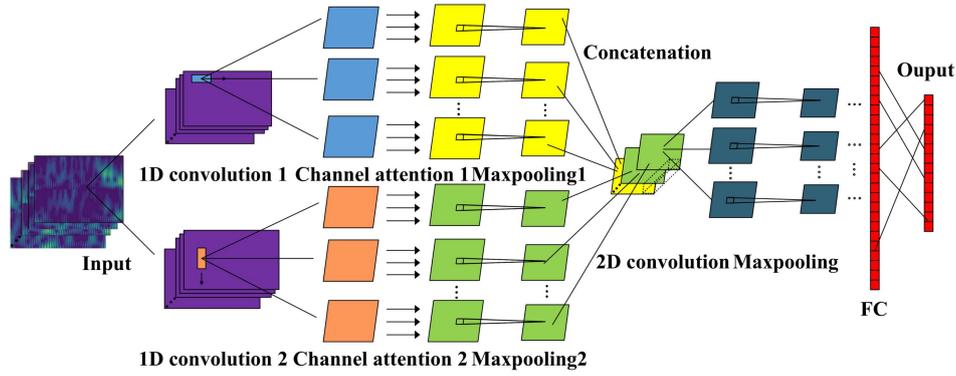


Fig. 6. Hybrid scale convolutional neural network

fully connected layer is b_2 . The formula is as follows:

$$Y_k = f_{IE}(X_k) = \sigma(W_1(\sigma(W_2 \cdot X_k))) \quad (5)$$

Where W_1, W_2 represents the parameters of the fully connected layer, and σ represents ReLU function [25].

The formula of the information divergence module is as follows:

$$f_{ID}(Y_k, F \sim C_k(i, j)) = Y_k \cdot F \sim C_k(i, j) \quad (6)$$

Where f_{ID} means the multiplication for all two-dimensional data of the k th channel of feature F by using Y_k , that is, diverging the feature value Y_k to the entire surface of the k th-channel.

C. Two-dimensional convolution layer

In this layer, a two-dimensional (2D) convolution kernel is used to extract the correlation information after feature splicing, and improve the correlation degree of feature information of α and β bands in time domain and frequency domain. The specific parameters are shown in Table IV.

TABLE IV
MAIN PARAMETERS OF 2D CONVOLUTIONAL LAYER

Layer name	Output shape	Kernel size	Stride	Padding
2Dconvolution	32×32×32	4×4	1	Same
Max pooling	16×16×32	2×2	1	Same

D. Overall Framework

The EEG time-frequency map serves as the input layer of the overall framework. Framework of the second layer to the fifth layer is 1D convolution layer, 2D convolution layer, the fully connected layer 1 and fully connected layer 2, and the full connected layer 1 contains 240 hidden neurons, the full connected layer 2 is the output layer, which contains two hidden neurons, because this paper performs two categories of tasks. In addition, the Softmax is chosen as classification function, the ReLU is chosen as the activation function, the cross entropy is chosen as the loss function and the Adam algorithm is chosen as the optimization function. Adam algorithm has faster convergence speed compared with other algorithms. The setting of initial value of learning rate directly affects the quality of final solution. If the initial value is set too high, the convergence of the loss function will oscillate. Therefore, this paper set the learning rate to 0.0003. The structure is shown in Fig.6, and the specific parameters are shown in Table V.

V. MOTOR IMAGERY SIGNAL CLASSIFICATION OFFLINE EXPERIMENT

A. Experiment introduction

This experiment uses an EPOC+ EEG acquisition instrument, as shown in Fig.7, and its sampling frequency is 128 Hz. The EPOC+ is equipped with 16 electrodes. The electrodes are placed as shown in Fig.8.

TABLE V
MAIN PARAMETERS OF OVERALL FRAMEWORK

Layer name	Output shape
1D convolutional layer	32×32×16
2D convolutional layer	16×16×32
FC	240
Output	2

Seven healthy subjects (denoted as A1, A2, ..., A7) are selected for the EEG motor imagery experiment. First, the subjects relax in a quiet experimental area, and then the experiment starts. Each experiment lasts for 6 seconds. In the first 1 second, healthy subjects are expected to see a left or right arrow as the experimental theme, at the 2nd second, the subjects would hear the beep and imagine left or right-hand movements accordingly, at the 6th second, the subjects hear the beep again and stop imagining. A total of 240 experiments have been conducted, and each type of subject experiment is repeated 120 times. A total of 120 left-hand motor imagery EEG data samples and 120 right-hand motor imagery EEG data samples are obtained. The EEG data from the 3rd second to the 4th second in the sample is recorded as the follow-up experimental object, and in order to reduce the computational complexity, only the EEG data of 6 channels (F3, F4, FC5, FC6, T7, T8) are recorded for analysis. The experimental process is shown in Fig.9.



Fig. 7. EPOC+ EEG acquisition instrument

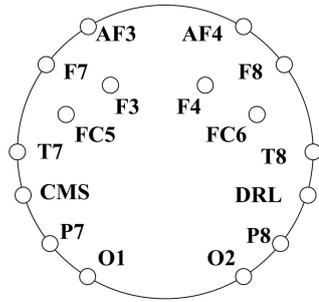


Fig.8. Electrode placement

B. Data preprocessing

EEG signals usually include a large amount of background noise such as electromyogram (EMG) and electro-oculogram (EOG). In order to reduce background noise, this paper preprocess the EEG signal in three steps. First, abnormal samples are removed. Second, 8-32 Hz band-pass [15] is adopted to filter for EEG signals. Third, the EEG signal is normalized.

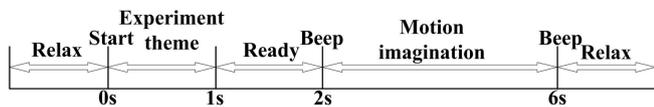


Fig. 9. The acquisition process of motor imagery signals

C. Data augmentation

1. Experiment process

According to the introduction in Section II, the EEG time-frequency map is constructed for the preprocessed EEG data. The data of a subject is chosen as an example. Two data enhancement experiments are carried out on the EEG time-frequency map of the subject according to the types of labels respectively. In the experiment, the changes of the loss value of the discriminative network are recorded, and the generative network output 32 artificial EEG time-frequency maps are generated every 100 times of discriminative network training. After the data enhancement experiments, the generated artificial EEG time-frequency maps map into the Inception-v3 network to compute four indices score according to the generated sequence with the original EEG time-frequency. the validity of the data enhancement experiments is determined according to the score, then high-quality artificial EEG time-frequency maps are selected according to the discriminative network loss value. The selected results are saved in the folder where the original EEG time-frequency maps are located at a ratio of 1:1.

2. Experimental Results and Analysis

The following procedures are described based on the EEG data of subject A7. The data enhancement experiment is trained by using the EEG time-frequency map of the subjects' left hand motor imagery, and the experiment is trained for a total of 10,000 times. Fig.10 records the scores of four indices of artificial EEG generated by 1000-8000 times of discriminative network training. IS and MS show a gradually increasing trend, indicating that the quality of the generated artificial EEG gradually is improved and became more diversified in the process of 1000-8000 training sessions. On the other hand, FID and MMD gradually decrease, indicating

that the data distribution of the generated artificial EEG is closer to that of the original EEG. In addition, it can be concluded from Fig.11 that after 8000 discriminative network training, the loss value of discriminative network gradually approaches 0. Therefore, the artificial EEG time-frequency maps generated at this time are selected as high-quality data.

The same strategy is used for training and selection of the right-hand motor imagined EEG time-frequency map of subject A7. In Fig.12 and Fig.13, the artificial EEG time-frequency maps generated after 9000 discriminative network training are selected as high-quality data.

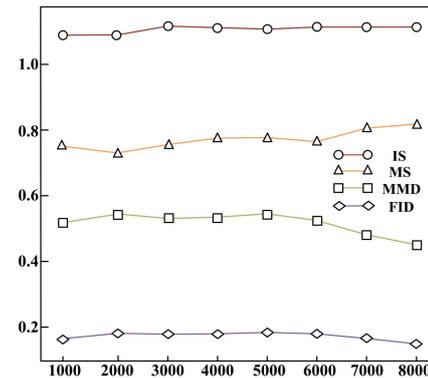


Fig.10. At the bottom of the figure is training times of discriminative network and on the left of the figure is scores of the four indices. Four different colored curves in the figure represent four different indices.

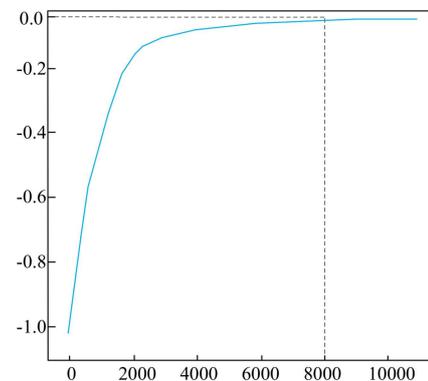


Fig.11. At the bottom of the figure is training times of discriminative network and on the left of the figure is the loss value of discriminative network. The point of contact of the two dotted lines in the figure are the number of training when the loss value is close to 0.

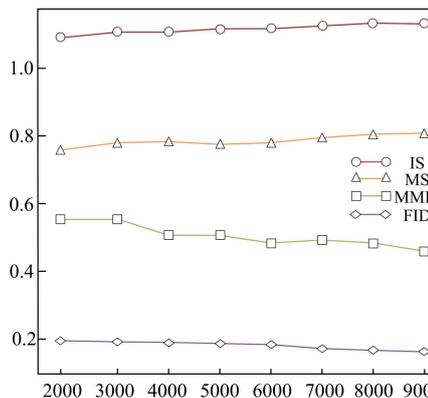


Fig.12. The bottom, left side of the figure and the curve in the figure are the same as in Fig.10

TABLE VI
THE RECOGNITION ACCURACY OF SIX ALGORITHMS FOR 7 SUBJECTS

Subject	EED	CSP-TSM	MEMD-STFT	DBN	1DCNN	HSCNN
A1	85.60	92.62	90.83	83.91	85.40	93.35
A2	76.70	72.53	71.01	92.44	71.24	95.66
A3	82.20	86.91	85.42	94.24	80.67	91.95
A4	91.64	95.39	93.70	91.03	95.35	92.40
A5	88.67	82.00	90.29	82.15	91.63	93.64
A6	70.11	71.45	76.77	84.44	81.76	89.31
A7	82.74	86.83	83.52	71.25	82.21	90.68
Average	82.52	83.97	84.51	85.64	84.04	92.43

TABLE VIII
THE RECOGNITION ACCURACY OF SEVEN ALGORITHMS FOR 9 SUBJECTS

Subject	EED	CSP	ACSP	DBN	CNN-SAE	1DCNN	Proposed
B1	57.34	67.00	64.75	67.65	74.51	77.54	84.34
B2	53.50	59.85	59.49	63.75	69.12	61.61	69.27
B3	54.69	67.32	63.45	62.00	73.34	63.02	76.51
B4	84.68	92.56	95.74	98.77	96.50	98.65	98.03
B5	72.57	79.59	79.66	81.06	83.00	84.44	86.75
B6	75.66	74.88	81.10	73.88	78.45	80.78	88.20
B7	57.45	73.48	69.20	79.84	77.54	79.47	87.12
B8	81.43	90.01	81.42	81.75	74.51	84.21	92.05
B9	72.77	84.63	80.51	87.46	72.31	80.66	87.98
Average	67.78	76.59	75.03	77.35	77.69	78.93	84.81

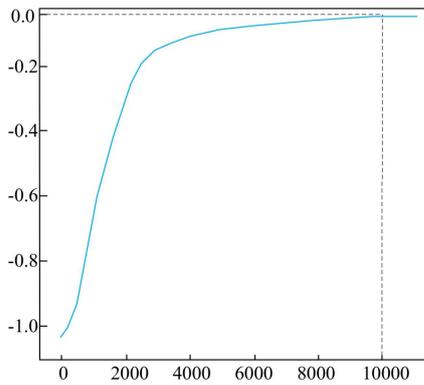


Fig.13. The bottom, left side of the figure and the curve in the figure are the same as in Fig.11.

D. Classification of EEG time-frequency maps

1. Experiment process

The tf.gfile.gfile function is used to read the EEG time-frequency map in the folder. The tf.image.decode_jpeg function is used to decode the EEG time-frequency map. The tf.image.resize_images function is used to change the decoded data format to 64×64. The EEG time-frequency map is saved in the form of pictures, the number of decoded channel is 3.

The EEG time-frequency maps before and after data enhancement are put into the HSCNN for 10 recognition experiments, then the validity of GDALF is proved according to the average recognition rate.

EED [26], CSP-TSM [27], ACSP [28], DBN [29] and 1DCNN [30] are used to conduct 10 recognition experiments on preprocessed EEG data, and the average recognition rate is compared with HSCNN to verify the validity of HSCNN.

2. Experimental Results and Analysis

In the Fig.14, the GADALF method proposed in this paper improves the recognition rate of each subject's EEG time-frequency map and effectively solves the problem of scarce EEG data, while the GADALF method does not

improve the recognition rate of subjects A2, A4 and A5 significantly, which may be caused by two reasons. First, useless information might have been included in EEG data collection. Second, the EEG time-frequency maps of some subjects are more difficult to classify by HSCNN.

In Table VI, for multiple subjects, HSCNN achieves higher EEG recognition rate compared with other methods. To draw a comparison with traditional methods such as EED, CSP-TSM and MEMD-STFT, continuous wavelet transform is used in this paper to transform temporal EEG data into EEG time-frequency maps that are more easily recognized by HSCNN. To draw a comparison with DBN and 1DCNN, HSCNN uses the combination of hybrid scale and channel attention mechanism to extract characteristic information of α band and β band in EEG time-frequency map more validity. In summary, EEG time-frequency map and HSCNN have a higher fit degree.

TABLE VII
THE P-VALUES BETWEEN THE HSCNN IN THIS PAPER AND OTHER FIVE ALGORITHMS

Method	EED	CSP-TSM	MEMD-STFT	DBN	1DCNN
P-values	0.004	0.012	0.018	0.040	0.013

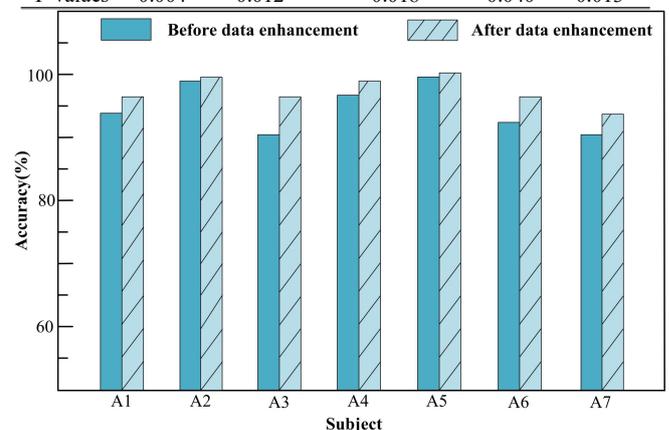


Fig. 14. The EEG time-frequency map classification rates in HSCNN before and after data supplementation

In order to verify the difference between the method proposed in this paper and other methods, the subject and method are defined as independent variables, and the classification rate is the dependent variable. The twoway analysis of variance and multiple comparisons are used to calculate the p-value between this algorithm and other algorithms. When p-value is less than 0.05, it is generally considered that there is a significant difference between the two comparison algorithms. In Table VII, the p-values are all less than 0.05.

E. Experiment on dataset BCI competition IV 2b

The dataset BCI competition IV 2b set is generated by 9 subjects performing motor imaging tasks. The subjects are all right-handed, the acquisition frequency is 250 Hz, the acquisition equipment is three bipolar recorders, and the acquisition channels are C₃, C_z, and C₄ Channel, the collection task is 5 motor imagination experiments. The first two groups of experiments adopt the non-feedback experimental paradigm collection, each group contains 120 motor imagination tasks, and the last three groups of experiments adopt the feedback experimental paradigm collection, and each group contains 160 motor imagination tasks. The first three experiments uses a band-pass filter of 0.5 -100 Hz for band-pass filtering.

Four steps are carried out on the data set of BCI competition IV 2b, including EEG data preprocessing, EEG time-frequency map construction, data enhancement and recognition. To draw a comparison with EED, ACSP, DBN, CNN-SAE [12], CSP [31] and IDCNN to verify the performance of the proposed method.

In Table VIII, for most subjects, the method proposed in this paper achieves the highest EEG recognition rate. For subject B2, excessive noise signals may be mixed in EEG data collection, resulting in low recognition rate under different methods. However, for most subjects, the GADALF method proposed in this paper can effectively improve the quality of EEG data and achieve high recognition rate in HSCNN. According to the above method, in Table IX, the p- values are all less than 0.05.

TABLE IX

THE P-VALUES BETWEEN THE METHOD PROPOSED IN THIS PAPER AND OTHER SIX ALGORITHMS

Method	EED	CSP	ACSP	DBN	CNN -SAE	IDCNN
P-values	<0.001	<0.001	<0.001	<0.001	<0.001	0.002

VI. ONLINE EXPERIMENT

A. Experiment introduction

The intelligent wheelchair system is shown in Fig.15. The system consists of the EPOC+ EEG acquisition instrument, the laptop computer, the wireless communication module, the control system and the wheelchair. The subjects took a laptop computer and wore EPOC+ EEG acquisition instrument while sitting in a wheelchair to conduct online experiments of motor imagination EEG signals. First, we grit the teeth to make the F8 channel produce obvious voltage changes, which is used to start and stop the online experiment. Second, we collect the left-hand and right-hand movement imagination EEG signal by using the F3, F4 and FC5, FC6, T7, T8 channels. And our method is employed to classify the EEG signals. The left-hand and right-hand movement imagination EEG signals are utilized to control

wheelchair to turn left and right, respectively. A total of 100 groups of online experiments are conducted, which are divided into 50 groups to control the left turn of wheelchair and 50 groups to control the right turn of wheelchair. The time interval of each group of experiments is 10 seconds for the rest of the subjects.

B. Experimental Results and Analysis

The results of the online experiment are shown in Table X. For most subjects, the online recognition rate of the method proposed in this paper is better than IDCNN. In table X and table VI, the same subjects in the online experiment, the recognition rate of the EEG data declines for two main reasons. First, subjects are vulnerable to the surrounding environment in the online experiment, second, subjects are more prone to fatigue in the online experiment, and than the quality of the EEG data collected is reduced.

TABLE X
ONLINE EXPERIMENT ACCURACY RATE OF 7 SUBJECTS

Subject	IDCNN			Proposed		
	Left	Right	Average	Left	Right	Average
A1	71.53	74.77	73.15	80.79	84.18	82.49
A2	86.91	83.87	85.39	89.64	87.28	88.46
A3	75.49	78.39	76.94	83.74	84.68	84.21
A4	70.27	73.87	72.07	82.45	87.31	84.93
A5	76.55	78.36	77.46	83.54	86.49	85.02
A6	79.26	77.43	78.35	75.64	74.05	74.85
A7	81.08	78.35	79.72	88.29	85.58	86.94

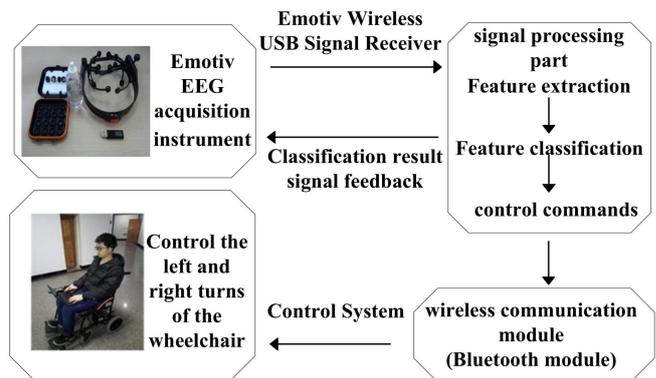


Fig. 15. The intelligent wheelchair system

VII. CONCLUSION

This paper proposes an EEG time-frequency map generation method based on continuous wavelet transform, which avoids the window selection problem of short-time Fourier transform, and proposes the generative discriminative adversarial learning framework to generate artificial EEG time-frequency maps to solve the problem of EEG data scarcity. Secondly, MMD, IS, MS, and FID are used as evaluation indicators, and the loss value of the discriminative network as the selection indicator to ensure the high quality and diversity of artificial EEG time-frequency maps. Then, by analyzing the characteristic form of α band and β band in EEG time-frequency map, the convolutional neural network is improved and a hybrid scale convolutional neural network is proposed to recognize EEG time-frequency map. The laboratory data and public dataset BCI competition IV 2b are verified the validity of the above method. Finally, an online experiment system is designed to

demonstrate the practicability of the intelligent wheelchair system.

REFERENCES

- [1] P. Dipti, and S. Dhage, "Feature Extraction Methods for Electroencephalography based Brain-Computer Interface: A Review," *IAENG International Journal of Computer Science*, vol.47, no.3, pp501-515, 2020.
- [2] Q. T. Obeidat, T. A. Campbell, and J. Kong, "Spelling with a small mobile brain-computer interface in a Moving Wheelchair," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol.25, no.11, pp2169-2179, 2017.
- [3] N. Cheng, K. S. Phua, and H. S. Lai, "Brain-computer interface-based soft robotic glove rehabilitation for stroke," *IEEE Transactions on Biomedical Engineering*, vol.67, no.12, pp3339-3351, 2020.
- [4] L. Duan, J. Li, and H. Ji, "Zero-shot learning for EEG classification in motor imagery-based BCI system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol.28, no.11, pp2411-2419, 2020.
- [5] L. Yao, N. M. Kersting, and X. Sheng, "A multi-class BCI based on somatosensory imagery," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol.26, no.8, pp1508-1515, 2018.
- [6] S. Ren, W. Wang, and Z. G. Hou, "Enhanced motor imagery based brain-computer interface via FES and VR for lower limbs," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol.28, no.8, pp1846-1855, 2020.
- [7] D. Wu, J. T. King, and C. H. Chuang C H, "Spatial filtering for EEG-based regression problems in brain-computer interface (BCI)," *IEEE Transactions on Fuzzy Systems*, vol.26, no.2, pp771-781, 2017.
- [8] X. Chen, X. Xu, and A. Liu, "The use of multivariate EMD and CCA for denoising muscle artifacts from few-channel EEG recordings," *IEEE transactions on instrumentation and measurement*, vol.67, no.3, pp359-370, 2017.
- [9] L. Farsi, S. Siuly, and E. Kabir, "Classification of alcoholic EEG signals using a deep learning method," *IEEE Sensors Journal*, vol.21, no.3, pp3552-3560, 2020.
- [10] X. Tian, Z. Deng, and W. Ying, "Deep multi-view feature learning for EEG-based epileptic seizure detection," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol.27, no.10, pp1962-1972, 2019.
- [11] C. Huang, Y. Xiao, and G. Xu, "Predicting Human Intention-Behavior through EEG signal analysis using Multi-Scale CNN," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2020.
- [12] Y. R. Tabar, and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *Journal of neural engineering*, vol.14, no.1, pp016003, 2016.
- [13] J. B. Salyers, Y. Dong, and Y. Gai, "Continuous wavelet transform for decoding finger movements from single-channel EEG," *IEEE Transactions on Biomedical Engineering*, vol.66, no.6, pp1588-1597, 2018.
- [14] P. Kaushik, A. Gupta, and P. P. Roy, "EEG-based age and gender prediction using deep BLSTM-LSTM network model," *IEEE Sensors Journal*, vol.19, no.7, pp2634-2641, 2018.
- [15] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," [Online], Available: <https://arxiv.org/abs/1511.06434>.
- [16] I. Goodfellow, J. P. Abadie, and M. Mirza, "Generative adversarial nets," *Advances in neural information processing systems*, vol.27, 2014.
- [17] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," *International conference on machine learning*, pp214-223, 2017.
- [18] A. Gretton, K. Borgwardt, and M. Rasch M, "A kernel method for the two-sample-problem," *Advances in neural information processing systems*, vol.19, pp513-520, 2006.
- [19] T. Salimans, I. Goodfellow, and W. Zaremba, "Improved techniques for training gans," *Advances in neural information processing systems*, vol.29, pp2234-2242, 2016.
- [20] T. Che, Y. Li, and A. P. Jacob, "Mode regularized generative adversarial networks," [Online], Available: <https://arxiv.org/abs/1612.02136>.
- [21] M. Heusel, H. Ramsauer, and T. Unterthiner, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," *Advances in neural information processing systems*, vol.30, 2017.
- [22] Y. Zhang, Z. Yin, and L. Nie, "Attention Based Multi-Layer Fusion of Multispectral Images for Pedestrian Detection," *IEEE Access*, vol.8, pp 165071-165084, 2020.
- [23] F. Huang, X. Zhang, and Z. Zhao, "Bi-directional spatial-semantic attention networks for image-text matching," *IEEE Transactions on Image Processing*, vol.28, no.4, pp2008-2020, 2018.
- [24] D. Li, J. Xu, and J. Wang, "A Multi-Scale Fusion Convolutional Neural Network Based on Attention Mechanism for the Visualization Analysis of EEG Signals Decoding," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol.28, no.12, pp2615-2626, 2020.
- [25] R. H. R. Hahnloser, R. Sarpeshkar, and M. A. Mahowald, "Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit," vol.405, no.6789, pp947-951, 2000.
- [26] S. Sun, "Extreme energy difference for feature extraction of EEG signals," *Expert Systems with Applications*, vol.37, no.6, pp4350-4357, 2020.
- [27] S. Kumar, K. Mamun, and A. Sharma, "CSP-TSM: Optimizing the performance of Riemannian tangent space mapping using common spatial pattern for MI-BCI," *Computers in biology and medicine*, vol.91, pp 231-242, 2017.
- [28] U. Talukdar, S. M. Hazarika, and J. Q. Gan, "Adaptation of Common Spatial Patterns based on mental fatigue for motor-imagery BCI," *Biomedical Signal Processing and Control*, vol.58, pp101829, 2020.
- [29] W. L. Zheng, and B. L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Transactions on Autonomous Mental Development*, vol.7, no.3, pp 162-175, 2015.
- [30] W. Mumtaz, and A. Qayyum, "A deep learning framework for automatic diagnosis of unipolar depression," *International journal of medical informatics*, vol.132, pp103983, 2019.
- [31] B. Chen, Y. Li, and J. Dong, "Common spatial patterns based on the quantized minimum error entropy criterion," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol.50, no.11, pp4557-4568, 2018.