Task-State EEG Signal Classification for Spatial Cognitive Evaluation Based on Multiscale High-Density Convolutional Neural Network

Dong Wen, *Member, IEEE*, Rou Li, Hao Tang, Yijun Liu, Xianglong Wan, *Member, IEEE*, Xianling Dong, M. Iqbal Saripan, *Senior Member, IEEE*, Xifa Lan, Haiqing Song, and Yanhong Zhou

***Abstract*— In this study, a multi-scale high-density convolutional neural network (MHCNN) classification method for spatial cognitive ability assessment was proposed, aiming at achieving the binary classification of task-state EEG signals before and after spatial cognitive training. Besides, the multi-dimensional conditional mutual information method was used to extract the frequency band features of the EEG data. And the coupling features under the combination of multi-frequency bands were transformed into multispectral images. At the same time, the idea of Densenet was introduced to improve the multi-scale convolutional neural network. Firstly, according to the discreteness of multispectral EEG image features, two-scale convolution kernels were used to calculate and learn useful channel and frequency band feature information in multispectral image data. Secondly, to enhance feature propagation and reduce the number of parameters, the dense network was connected after the multi-scale convolutional network, and the learning rate change function of the stochastic gradient descent algorithm was optimized to objectively evaluate the training effect. The experimental results showed that compared with the classical convolution neural network (CNN) and multi-scale convolution neural network, the proposed MHCNN had better classification performance in the six frequency band combinations with the highest accuracy of 98%: Theta-Alpha2-Gamma, Alpha2-Beta2-Gamma, Beta1-Beta2-Gamma, Theta-Beta2-Gamma, Theta-Alpha1-Gamma, and Alpha1-Alpha2-Gamma. By comparing the classification results of six frequency band combinations, it was found that the combination of the Theta-Beta2-Gamma band had the best classification effect. The MHCNN classification method proposed in this research could be used as an effective biological indicator of spatial cognitive training effect and could be extended to other brain function evaluations.**

***Index Terms*— Multi-scale high-density convolutional neural network, spatial cognition evaluation, task-state EEG signals.**

Manuscript received November 12, 2021; revised March 28, 2022; accepted April 2, 2022. Date of publication April 11, 2022; date of current version April 25, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61876165 and Grant 61503326 and in part by the National Key Research and Development Program of China under Grant 2021YFF1200603. *(Corresponding author: Yanhong Zhou.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of First Hospital of Qinhuangdao in Hebei Province, China, under Approval No. 2018B006.

Dong Wen and Xianglong Wan are with the Institute of Artificial Intelligence, University of Science and Technology Beijing, Haidian, Beijing 100083, China.

Rou Li and Hao Tang are with The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, School of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066004, China.

Yijun Liu is with the School of Science, Yanshan University, Qinhuangdao, Hebei 066004, China.

Xianling Dong is with the Hebei Key Laboratory of Nerve Injury and Repair, Chengde Medical University, Chengde 067000, China.

M. Iqbal Saripan is with the Faculty of Engineering, Universiti Putra Malaysia (UPM), Serdang, Selangor Darul Ehsan 43400, Malaysia.

Xifa Lan is with the Department of Neurology, First Hospital of Qinhuangdao, Qinhuangdao 066099, China.

Haiqing Song is with the Department of Neurology, Xuanwu Hospital of Capital Medical University, Beijing 100053, China.

Yanhong Zhou is with the School of Mathematics and Information Science and Technology, Hebei Normal University of Science and Technology, Qinhuangdao, Hebei 066004, China (e-mail: yhzhou168@163.com).

Digital Object Identifier 10.1109/TNSRE.2022.3166224

I. Introduction

SPATIAL cognition refers to the information processing process of size, shape, direction, and distance of three-dimensional objects in physical space or mental space [1]. Spatial cognitive problems such as cognitive degradation of visuospatial structure and spatiotemporal disorientation are common in some brain diseases. For example, the course of Alzheimer's disease (AD) is irreversible and its etiology is not clear [2], [3]. Therefore, it is of great importance to explore spatial cognitive training and assessment methods for the prevention of spatial cognitive decline in healthy subjects. It is a problem to be solved in the field of brain science and cognitive science to analyze and study spatial cognitive ability and find a quick and effective method to evaluate spatial cognitive ability.

Due to the characteristics of the large volume and the high dimension of EEG data [4], the task-state EEG recognition method for spatial cognitive evaluation needs to input appropriate features into the classification method to evaluate the spatial cognitive ability. In traditional machine learning algorithms, linear discriminant analysis [5], K-means clustering [6], and SVM [7] were often used to classify EEG signals. However, these methods have problems such as slow computing speed, large storage space demand, and difficulty in solving multi-classification problems [8], which result in

limitations in the field of nervous system research. Deep learning, as a high-performance method in the field of machine learning, has been widely used in the field of EEG signal classification. Some studies have proved the feasibility of deep learning in the field of EEG research [9]–[13], Akrofi *et al.* used multiple discriminant analysis and K-means clustering to analyze EEG to evaluate and detect Alzheimer's disease [9], and Li *et al.* studied the classification of stroke patients and healthy controls by evolutionary multinucleate support vector machine based on genetic programming [12].

Compared with the results of other classical machine learning methods in EEG classification, the CNN classification model has not significantly improved. The main reason is that these researches on deep learning of EEG signals do not give full play to the advantages of deep learning, especially CNN for spatial information of data. The idea of multispectral image transformation proposed by Bashivan [14] made up for this defect. According to the research of Bashivan *et al.*, it was found that the classification effect of CNN with multispectral image converted EEG signal as feature input was better than other classifiers. At the same time, considering that the optimization design of CNN was crucial to its final classification performance, they combined the multi-scale convolutional neural network with the high-density network. The multi-scale convolutional neural network was proposed based on Googlenet's Inception [15]–[17], which could accelerate the convergence speed of the network and effectively improve the results of classification. To reduce the computational load of the model and avoid the inability to update the parameters, we used a two-scale convolution kernel to calculate and learn useful channel and frequency band and other feature information in the multispectral image data. The results did not achieve the desired effect. Later, by referring to the Densenet model structure proposed by Huang *et al.* [18], this structure had good performance in CNN image classification, which could greatly reduce the number of parameters. At the same time, relevant studies have applied CNN improved model to the field of cognitive evaluation. For example, Li *et al.* applied CNN to improve the cognitive training effect of network recognition [19] and many researchers also applied CNN's improved model to classify the emotion of EEG [20], [21].

Therefore, we improved the multi-scale convolutional neural network based on the feature extraction method of multi-spectral image transformation, proposed the new classifier named multi-scale high-density convolutional neural network (MHCNN), and applied it to the EEG signal classification task of spatial cognition. In this experiment, 16 task-state EEG signals collected by 7 volunteers before and after spatial cognitive training were classified as data. In addition, multi-scale high-density convolutional neural networks, multi-scale convolutional neural networks, and traditional convolutional neural network classifier models were used for comparative analysis.

II. Materials and Methods

A. Data Set

16 channels of EEG signals studied in this paper came from Yanshan University. The experimental subjects were 7 male volunteers without any mental system or neurological diseases, aged between 21 and 26. All subjects signed informed consent before the experiment [22], [23], and the experimental study was approved by the Ethics Committee of First Hospital of Qinhuangdao in Hebei Province, China (The approval number is 2018B006 in 2018). The data set included two kinds of EEG signals from 7 male volunteers participating in the game of Virtual City Walking before and after VR spatial cognitive training task (Virtual Morris Water Maze).

In the task of this study, subjects participated in a virtual pool environment by wearing a VR head-mounted display and used remote cues to navigate themselves to the target platform. First, the subjects were made to understand the task requirements and learn the location of the platform according to the remote cues. Secondly, the subjects searched for the hidden platform location from different perspectives and directions according to the distal cues from different starting positions and trained their spatial cognitive navigation ability. Finally, the platform was removed. And the spatial memory storage ability was evaluated by navigating to the location of the platform in memory according to the remote cues.

In the process of data recording, subjects need to stay awake and relaxed, wear an electrode cap in a quiet room for about 10 minutes, and conduct human-computer interaction with the virtual game through an external controller. Meanwhile, the room temperature was kept at about 23°C, with an error of no more than 2°C [22], [23]. The subjects' spatial cognition was assessed after performing a task without any training. Then the game and spatial cognition were trained for 20 days. And finally, the task was performed again on the 20th day, and the spatial cognition was evaluated after training [22], [23].

In the study, the EEGLAB toolbox in MATLAB was used to conduct preprocessing of EEG Data, including EEG preview, filters, raw data inspection, EEG frequency division, data extraction (epoch), baseline correction. First, obvious waveform drift data were removed and power frequency interference was filtered. The electromyographic effects of blinking or eye movements were then corrected by ICA and the artifact signals due to device or subject movements were removed. Then, EEG signals were divided into frequency segments. The time window for extracting EEG data was set as 1 second before and after the event, and the baseline correction was carried out based on the 1 second before the event. EEG signals were divided into 7 frequency bands: Delta(1-4Hz), Theta(4-8Hz), Alpha1(8-10.5Hz), Alpha2(10.5-13Hz), Beta1 (13-20Hz), Beta2(20-30Hz), and Gamma(30-40Hz).

The distribution of EEG electrodes is shown in Fig. 1. EEG signals from the following 16 channels were selected: Fp1, Fp2, F3, Fz, F4, of the frontal lobe, F7, P7 of the left temporal lobe, FCz, C3, C4, Cz, Pz of the parietal lobe, F8, P8, of the right temporal lobe, O1, and O2 of the occipital bone, and the bilateral auricle as the reference electrode. The channels are numbered from left to right and from top to bottom according to the scalp electrodes. As shown in reference [22].

An OpenBCI device with 16 wet electrodes was used in the experiment, and the electrodes were placed according to the international standard lead 10-20 systems. The sampling frequency of EEG signals was 125Hz, and the sliding window

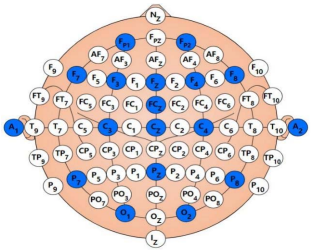

Fig. 1. Distribution of EEG electrodes.

technology (the sliding window was 2 seconds, with an overlap rate of 50%) was adopted to perform the segmentation of EEG signals. Due to the different duration of the task before and after the training, EEG signals of each frequency band before and after the training were segmented into 424 and 289 samples respectively. The corresponding relationship between the time domain and frequency domain was obtained by a fast Fourier transform of each channel in a single sample.

B. Feature Extraction Methods

In this paper, we adopted the multidimensional conditional mutual information (MCMI) method [24]. As we know, MCMI is an improved method of the conditional mutual information (CMI). While CMI only analyze the coupling dynamic relationship between two different brain regions (or channels), MCMI takes the influence of other channels on the coupling characteristics of the two selected channels into consideration. Briefly, 16 scalp electrodes were divided into 3 groups according to the relevant position. The time series of two channels are firstly selected and investigated to check whether they belong to the same group. If they are in the same group, an additional consideration for the influence of other channels should be given. Otherwise, the two channels are far away enough, and the influence of other channels is neglected. Finally, we calculate the coupling strength matrix regarding all the combinations of two channels.

In the experiment, the MCMI method was used to extract features from task-state EEG before and after spatial cognitive training. Then, the following three feature classification methods were used to classify and compare the results to verify the feasibility, effectiveness, and accuracy of the proposed multi-scale high-density convolutional neural network algorithm.

C. Existing Classifier Methods

1) Convolutional Neural Networks: The convolutional neural network consists of five basic structures: the Input Layer, the Convolutional Layer, the Pooling Layer, the Fully Connected Layer, and the Softmax Layer. The role of the convolution and pooling layer is to extract features from the input image, while the role of the full connection and classification layer is to classify. The convolutional neural network has been widely used as a deep learning model, and the most basic structure of CNN can complete feature extraction and classification. CNN has also made many research achievements in EEG analysis [14], [25]–[28]. This study took CNN as a classic classification model for experimental comparison. The CNN structure in the experiment is two groups of convolution pooling and two layers of full connection. Relu is selected as the activation function and Adam as the optimization function.

2) Multi-Scale Convolutional Neural Networks: The multi-scale convolutional neural network was proposed based on Googlenet's Inception [15]–[17]. For example, the multi-scale convolutional neural network used in the literature [15] learned the original image through three convolutional cores of different sizes. The similarity between the multi-scale convolutional neural network and the classical convolutional neural network lies in that they are stacked with the convolutional layer, the pooling layer, the full connection layer, and the final classification layer. The multi-scale convolutional neural network uses the multi-scale convolutional kernel, which enriches the image features, retains the original information of the image, and has better robustness. Some studies have improved the structure of the multi-scale CNN to analyze EEG signals [29], [30], verifying the feasibility of applying this network to analyze EEG. In this paper, the multi-scale part was added based on classical CNN for experimental comparison [31].

D. Multi-Scale High-Density Convolutional Neural Network

1) Design of Multi-Scale High-Density Convolutional Neural Networks: EEG signals can be processed by multi-dimensional conditional mutual information (MCMI) [24] method to obtain multispectral image data including channel location and frequency band characteristic information. The distribution of characteristic information in multispectral image data is global and discrete. If the convolutional layer in the constructed CNN only uses a single-size convolutional kernel, the model will not be able to effectively learn the feature information contained in the multi-spectral image data.

The number of convolutional layers is also an important factor that determines the performance of CNN. Deeper networks allow the model to learn more complex features.However, when the training sample size was relatively small, if CNN used too many convolutional layers for feature extraction of data, the gradient would disappear and the parameter update could not be transmitted to all the neural networks in the process of backpropagation [32].

Therefore, we improved the classifier from the perspective of the multi-scale convolution kernel and dense connection. Among them, Huang *et al.* proposed the MSDnet method based on this perspective, which used multi-scale network architecture to train multiple classifiers with different resource requirements [18]. To maximize the reuse of the calculation between classifiers, they were incorporated into a single deep convolutional neural network as early exits and connected through dense connections. MSDnet sets up multiple classifications exits in a network. For simple images, the results can be obtained directly from one of the previous classification exits. For images that are difficult to be classified, reliable results can only be obtained by going to a certain layer at the back of the network. This kind of network structure will affect

the accuracy of the late classifier and make the calculation of the late classifier very demanding. Therefore, MSDNet is more suitable to solve the problem of multiple classifications of images with different degrees of difficulty under the condition of limited computing resources.

The purpose of spatial cognitive EEG signal recognition in this study is to binary classify multispectral images that are difficult to classify under the condition of unlimited computational resources and obtain more accurate classification results as far as possible. Therefore, MSDNet is not suitable for the analysis of EEG signals in this study. Given the above analysis, we proposed a multi-scale high-density convolutional neural network based on deep learning knowledge and improved CNN in the following aspects.

Adopt multiple-size convolution kernels. Multi-scale features can effectively improve the results of image retrieval, image classification, and object detection tasks. If deep convolution undergoes multiple downsampling, the characteristic information of the large target may be lost. So we only used a multi-size convolution kernel once after the input of image data. At the same time, if the size of the convolution kernel is too large, the parameters and training time of the whole model will increase. To reduce the calculation amount of the model and avoid the parameter that can not be updated, we used (3×3) and (5×5) convolution kernels to calculate and learn the useful channel, frequency band, and other characteristic information in multispectral image data. Using the Leaky Relu activation function can effectively avoid some useful but negative feature information being ignored, change the data distribution through the activation function, and reduce the gradient disappearance of the deep neural network in the training process.

Adopt a highly dense network of neural connections. By feature reuse, Densenet can alleviate the gradient disappearance problem, and the characteristics of its model structure can make the network narrower, reduce the amount of computation, and avoid the occurrence of overfitting. In Dense Block, the feature graphs of each layer are of the same size and can be connected to the channel dimension. The nonlinear combination function in the Dense Block uses the structure of BN+ReLU+3×3 Conv. After the convolution of each layer in all Dense blocks, k feature graphs are output. That is, the number of channels of the obtained feature graphs is k, or k convolution cores are used. In general, a smaller k (such as 12) is used, and better performance can be obtained. Because the input of the following layer will be very large, the Dense Block can use the bottleneck layer to reduce the computing amount. Increase 1×1 Conv in the original structure (BN+ReLU+1×1 Conv+BN+ReLU+3×3 Conv) and reduce the number of features to improve the computing efficiency. The middle layer of the Dense Blocks is called the transition layer. The transition layer compiled the model with the same size by carrying out the following sampling operation (including a 1×1 convolution and 2×2 AVGpooling), and the structure was BN+ReLU+1×1 CONV+2×2 AVGpooling. Assumed that the number of channels in the feature graph obtained from the Dense Block on the Transition layer was m, and the Transition layer could generate θ*m features (through the convolutional layer). In this research, $\theta = 0.5$ was used.

TABLE I
THE NETWORK STRUCTURE ADOPTED BY THE HIGH-DENSITY CONVOLUTIONAL NEURAL NETWORK

Layers	Highly dense convolutional neural networks
Dense Block(1)	$\begin{bmatrix} 1*1 & conv \\ 3*3 & conv \end{bmatrix} * 12$
Transition Layer(1)	$1*1 \quad conv$ $2*2 \quad average\ pool$
Dense Block(2)	$\begin{bmatrix} 1*1 & conv \\ 3*3 & conv \end{bmatrix} * 12$
Classification Layer	$2*2 \quad average\ pool$ $Fully-Connected, softmax$

Before entering the first Dense Block, the high-density network used was convolved once by the MSDnet method. Because the data set used in this study was relatively simple, the deeper model was easy to overfit. It contained two Dense blocks in total, and each Dense Block contained 12 layers. Final Dense Block was followed by an Average Pooling layer, which was then fed into a Softmax classifier. Note that in Densenet, all 3×3 convolution used padding = 1 to keep the size of the feature graph unchanged.

According to the above content, the detailed structure of the multi-scale high-density convolutional neural network constructed is shown in Fig. 2. In the training process, the adaptive stochastic gradient descent method was adopted instead of the stochastic gradient descent method to update the weights in the model iteratively.

2) Parameters Setting of Adaptive Stochastic Gradient Descent Method: In Densenet, the stochastic gradient descent method is used for training. Due to problems such as selecting the value of the learning rate, unstable model, and failure to reach the global optimal solution, we optimized the performance of the adaptive stochastic gradient descent method [33].

First, the impulse is added to the stochastic gradient descent method. Impulse could prevent the model from continuing training when it reaches the global optimum. At the same time, impulse could effectively speed up the learning process of the stochastic gradient descent method, especially for the gradient with noise and high curvature, which had a good inhibition effect [33]. The core idea of the impulse is to introduce a variable v to accumulate gradients generated before the current iteration. When v is roughly in the same direction as the current gradient, gradient descent will be enhanced; otherwise, it will be reduced. Formula (1) and Formula (2) are related to impulse:

$$v_{t+1} = \mu v_t - \alpha \nabla_\theta h(\theta_t)(1) \tag{1}$$

$$\theta_{t+1} = \theta_t + v_{t+1} \tag{2}$$

Where $\mu \in [0,1]$ is a random number, where α is the learning rate.

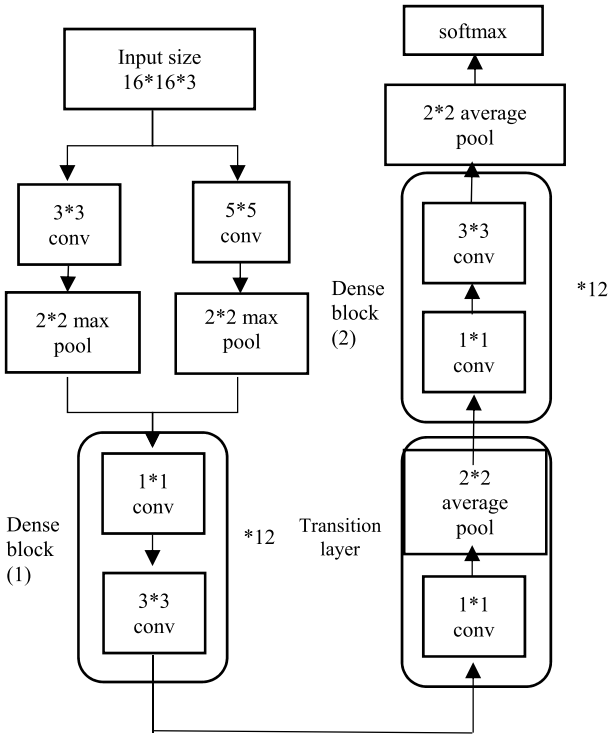

Fig. 2. Multi-scale high-density convolutional neural network.

Secondly, the learning rate is determined by setting the threshold value k, which decreases linearly in the first k*Epochs iteration, and decreases exponentially in the [k*Epochs, Epochs] iteration. The relevant formula of this method is shown in Formula (3):

$$\alpha(t) = \begin{cases} \alpha - \omega * epoch_t \\ \quad epoch_t \leqslant k * Epochs \\ (\alpha - \omega * k * Epochs)\, e^{(-\beta * (epoch_t - k * Epochs))} \\ \quad epoch_t > k * Epochs \end{cases} \tag{3}$$

where α is the initial learning rate, Epochs in the total number of iterations, the epoch is the t-th iteration value, ω is the linear rate of decline, β is the exponential rate of decline. The initial learning rate α is set to 0.01 and $\mu = 0.9$, which is the most commonly used value of impulse in stochastic gradient descent. The total number of iterations of Epochs is set to 120, the drop rate Ω is set to 1e-5, the drop rate β is set to 0.1, and the threshold K is set to 2/3.

E. Statistical Analysis

In this experiment, TensorFlow deep learning framework was adopted to build the neural network, and 5-fold cross-validation was adopted to evaluate the model. Four data were trained, one data was validated, and the average value of evaluation indexes was obtained after 5 validations as the evaluation value of model performance. Precision, F1-score, Recall, AUC, average verification accuracy curve, and average verification loss curve were used to evaluate the classification performance of the model.

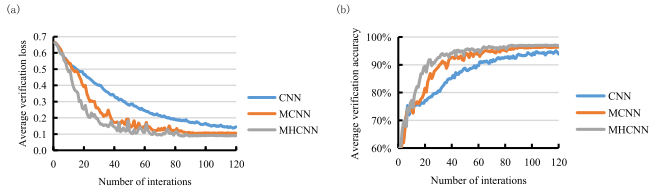

Fig. 3. Average verification loss curve (a) and accuracy curve (b) under the combination of Theta-Alpha2-Gamma band.

The verification accuracy curve is the variation curve of the model on the Precision of the verification data set with the change of the number of iterations during the training iteration of the deep learning model. The verification loss curve is the change curve of the loss value of the model to the verification data set with the change of the number of iterations during the training iteration of the deep learning model.

Since the verification accuracy curve and verification loss curve can be used to evaluate the data fitting ability of the neural network model, the average verification accuracy curve and the average verification loss curve obtained from 5-fold cross-validation are presented in this paper. At the same time, Precision, Recall, F1-score, and AUC can be obtained in each fold cross-validation process of the model. To facilitate understanding, the average values of the 5 fold cross-validation of these four indexes are presented.

III. Results

A. Comparison of Experimental Results With Different Classifiers

In this research, classical CNN, multi-scale CNN (MCNN), and multi-scale high-density CNN (MHCNN) were studied and compared. The multispectral image dataset was obtained by processing spatial cognitive EEG data with MCMI. Combined with the above three classifiers, the experimental results are presented and compared for analysis.

1) Theta-Alpha2-Gamma Band Combination: The average verification loss of multispectral image data constructed based on the MCMI method in the three CNN models under the combination of the Theta-Alpha2-Gamma band is shown in Fig. 3(a). According to Fig. 3(a), the following results can be drawn: the curve of CNN is the highest and tends to be stable between 0.15 and 0.16 when the iteration reaches 110 times. The curve of MCNN is the next, which tends to be stable between 0.10 and 0.11 when the iteration reaches 85 times. MHCNN has the lowest curve, which tends to be stable between 0.07 and 0.08 when the iteration reaches 84 times, and the curve is relatively smoothest after stabilization.

The average verification accuracy of multispectral image data constructed based on the MCMI method in the three CNN models under the condition of the combination of the Theta-Alpha2-Gamma band is shown in Fig. 3(b). According to Fig. 3(b), the following conclusions can be drawn: the curve of CNN is the lowest and tends to be stable between 0.94 and 0.95 when the iteration reaches 100 times. The curve of MCNN is the next, which tends to be stable between 0.96 and 0.97 when the iteration reaches 90 times. MHCNN

TABLE II
EVALUATION VALUES OF MODEL UNDER THE COMBINATION OF THETA-ALPHA2-GAMMA BAND

Classifier	Precision	Recall	F1-score	AUC
CNN	96.2%	95.4%	95.4%	99.0%
MCNN	96.5%	96.7%	96.6%	99.0%
MHCNN	96.9%	97.2%	97.2%	99.1%

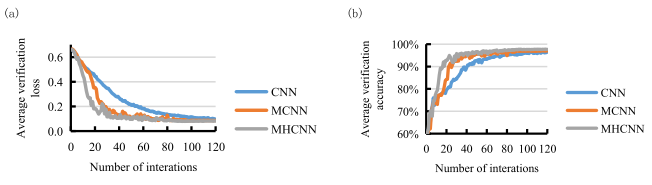

Fig. 4. Average verification loss curve (a) and accuracy curve (b) under the combination of Alpha2-Beta2-Gamma band.

TABLE III
EVALUATION VALUES OF MODEL UNDER ALPHA2-BETA2-GAMMA BAND COMBINATION

Classifier	Precision	Recall	F1-score	AUC
CNN	96.6%	97.4%	97.4%	99.2%
MCNN	96.9%	97.7%	97.6%	99.2%
MHCNN	97.6%	97.9%	97.9%	99.3%

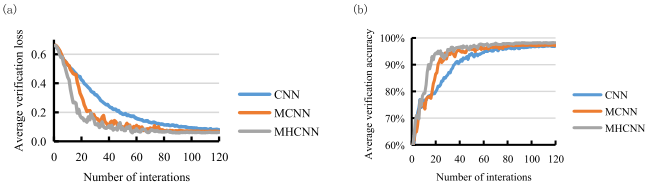

Fig. 5. Average verification loss curve (a) and accuracy curve (b) under the combination of Beta1-Beta2-Gamma band.

has the highest curve, which tends to be stable between 0.97 and 0.98 when the iteration reaches 83 times, and the curve is relatively smoothest after stabilization.

Table II shows the average Precision, F1-score, Recall, and AUC evaluation values of the multispectral image data constructed based on the MCMI method in the three CNN models under the condition of the combination of the Theta-Alpha2-Gamma band. It can be seen from the table that the four evaluation values of the multispectral image data in the MHCNN model are significantly better than those of CNN and MCNN.

2) Alpha2-Beta2-Gamma Band Combination: The average verification loss of the multispectral image data constructed based on the MCMI method in the three CNN models under the combination of Alpha2-Beta2-Gamma band is shown in Fig. 4(a). According to Fig. 4(a), the following results can be drawn: the curve of CNN is the highest and tends to be stable between 0.10 and 0.11 when the iteration reaches 105 times. The curve of MCNN is the next, which tends to be stable between 0.09 and 0.10 when the iteration reaches 95 times. MHCNN has the lowest curve, which tends to be stable between 0.07 and 0.08 when the iteration reaches 84 times, and the curve is relatively smoothest after stabilization.

The average verification accuracy of the multispectral image data constructed based on the MCMI method in the three CNN models under the combination of Alpha2-Beta2-Gamma band is shown in Fig. 4(b). According to Fig. 4(b), the following conclusions can be drawn: the curve of CNN is the lowest and tends to be stable between 0.96 and 0.97 when the iteration reaches 103 times. The curve of MCNN is the next, which tends to be stable between 0.96 and 0.97 when the iteration reaches 90 times. MHCNN has the highest curve, which tends to be stable between 0.97 and 0.98 when the iteration reaches 83 times, and the curve is relatively smoothest after stabilization. Fig. 4. Average verification loss curve (a) and accuracy curve (b) under the combination of Alpha2-Beta2-Gamma band.

Table III shows the average Precision, F1-score, Recall, and AUC evaluation values of the multispectral image data constructed based on the MCMI method in the three CNN models under the combination of Alpha2-Beta2-Gamma band. It can be seen from the table that the multispectral image data in the MHCNN model are significantly better than the 4 evaluation values of CNN and MCNN.

3) Beta1-Beta2-Gamma Band Combination: The average verification loss of multispectral image data constructed based on the MCMI method in the three CNN models under the condition of the combination of the Theta-Alpha2-Gamma band is shown in Fig. 5(a). According to Fig. 5(a), the following results can be drawn: the curve of CNN is the highest and tends to be stable between 0.07 and 0.08 when the iteration reaches 110 times. The curve of MCNN is the next, which tends to be stable between 0.06 and 0.07 when the iteration reaches 95 times. MHCNN has the lowest curve, which tends to be stable between 0.05 and 0.06 when the iteration reaches 84 times, and the curve is relatively smoothest after stabilization.

The average verification accuracy of the multispectral image data constructed based on the MCMI method in the three CNN models under the combination of Beta1-Beta2-Gamma bands is shown in Fig. 5(b). According to Fig. 5(b), the following conclusions can be drawn: the curve of CNN is the lowest and tends to be stable between 0.96 and 0.97 when the iteration reaches 100 times. The curve of MCNN is the next, which tends to be stable between 0.96 and 0.97 when the iteration reaches 90 times. MHCNN has the highest curve, which tends to be stable between 0.97 and 0.98 when the iteration reaches 75 times, and the curve is relatively smoothest after the stability.

Table IV shows the average Precision, F1-score, Recall, and AUC evaluation values of the multispectral image data constructed based on the MCMI method in the three CNN models under the combination of Beta1-Beta2-Gamma bands. It can be seen from the table that the multispectral image data in the MHCNN model are significantly better than the 4 evaluation values of CNN and MCNN.

4) Theta-Beta2-Gamma Band Combination: The average verification loss of multispectral image data constructed based on the MCMI method in the three CNN models under the condition of the combination of the Theta-Beta2-Gamma band

TABLE IV
EVALUATION VALUES OF MODEL UNDER THE COMBINATION OF BETA1-BETA2-GAMMA BAND

Classifier	Precision	Recall	F1-score	AUC
CNN	97.6%	97.9%	97.9%	99.3%
MCNN	97.6%	97.8%	97.9%	99.3%
MHCNN	97.7%	98.1%	98.0%	99.4%

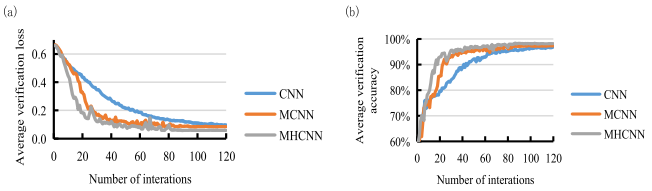

Fig. 6. Average verification loss curve (a) and accuracy curve (b) under the combination of Theta-Beta2-Gamma band.

TABLE V
EVALUATION VALUES OF MODEL UNDER THE COMBINATION OF THETA-BETA2-GAMMA BAND

Classifier	Precision	Recall	F1-score	AUC
CNN	96.6%	96.8%	96.7%	99.1%
MCNN	97.4%	97.8%	97.8%	99.2%
MHCNN	98.0%	98.0%	98.0%	99.4%

is shown in Fig. 6(a). According to Fig. 6(a), the following results can be drawn: the curve of CNN is the highest and tends to be stable between 0.10 and 0.11 when the iteration reaches 105 times. The curve of MCNN is the next, which tends to be stable between 0.08 and 0.09 when the iteration reaches 95 times. MHCNN has the lowest curve, which tends to be stable between 0.05 and 0.06 when the iteration reaches 80 times, and the curve is relatively smoothest after the stability.

Under the condition of the combination of the Theta-Beta2-Gamma band, the multispectral image data based on the MCMI method is constructed in three kinds. The average verification accuracy in the CNN model is shown in Fig. 6(b). Fig. 6(b) displayed that the curve of CNN is the lowest and tends to be stable between 0.96 and 0.97 when the iteration reaches 103 times. The curve of MCNN is the next, which tends to be stable between 0.96 and 0.97 when the iteration reaches 90 times. MHCNN has the highest curve, which tends to be stable between 0.97 and 0.98 when the iteration reaches 75 times, and the curve is relatively smoothest after the stability.

Table V shows the average Precision, F1-score, Recall, and AUC evaluation values of the multispectral image data constructed based on the MCMI method under the condition of the combination of Theta-Beta2-Gamma band in the three CNN models. It can be seen from the table that the multispectral image data in the MHCNN model are significantly better than the 4 evaluation values of CNN and MCNN.

5) Theta-Alpha1-Gamma Band Combination: The average verification loss of multispectral image data constructed based on the MCMI method in the three CNN models under the condition of the combination of Theta-Alpha1-Gamma band

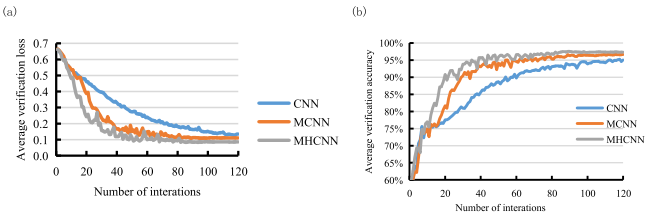

Fig. 7. Average verification loss curve (a) and accuracy curve (b) under the combination of Theta-Alpha1-Gamma band.

TABLE VI
EVALUATION VALUES OF MODEL UNDER THE COMBINATION OF THETA-ALPHA1-GAMMA BAND

Classifier	Precision	Recall	F1-score	AUC
CNN	96.8%	97.8%	97.7%	99.1%
MCNN	96.6%	96.8%	96.8%	99.2%
MHCNN	97.4%	97.4%	97.5%	99.3%

is shown in Fig. 7(a). According to Fig. 7(a), the following results can be drawn: the curve of CNN is the highest and tends to be stable between 0.10 and 0.11 when the iteration reaches 105 times. The curve of MCNN is the next, which tends to be stable between 0.08 and 0.09 when the iteration reaches 95 times. MHCNN has the lowest curve, which tends to be stable between 0.05 and 0.06 when the iteration reaches 80 times, and the curve is relatively smoothest after the stability.

The average verification accuracy of the multispectral image data constructed based on the MCMI method in the three CNN models under the condition of the combination of the Theta-Alpha1-Gamma band is shown in Fig. 7(b). According to Fig. 7(b), the following conclusions can be drawn: the curve of CNN is the lowest and tends to be stable between 0.96 and 0.97 when the iteration reaches 103 times. The curve of MCNN is the next, which tends to be stable between 0.96 and 0.97 when the iteration reaches 90 times. MHCNN has the highest curve, which tends to be stable between 0.97 and 0.98 when the iteration reaches 75 times, and the curve is relatively smoothest after the stability.

Table VI shows the average Precision, F1-score, Recall, and AUC evaluation values of the multispectral image data constructed based on the MCMI method in the three CNN models under the condition of the combination of Theta-alpha1-gamma band. It can be seen from the table that the multispectral image data in the MHCNN model are significantly better than the 4 evaluation values of CNN and MCNN.

6) Alpha1-Alpha2-Gamma Band Combination: The average verification loss of the multispectral image data constructed based on the MCMI method in the three CNN models under the condition of Alpha1-Alpha2-Gamma band combination is shown in Fig. 8(a). According to Fig. 8(a), the following results can be drawn: the curve of CNN is the highest and tends to be stable between 0.12 and 0.13 when the iteration reaches 105 times. The curve of MCNN is the next, which tends to be stable between 0.11 and 0.12 when the iteration reaches 95 times. MHCNN has the lowest curve, which tends to be stable between 0.08 and 0.09 when the iteration reaches

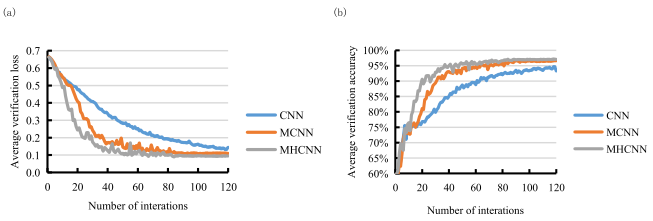

Fig. 8. Average verification loss curve (a) and accuracy curve (b) under the combination of Alpha1-Alpha2-Gamma band.

TABLE VII
EVALUATION VALUE OF MODEL UNDER THE COMBINATION OF ALPHA1-ALPHA2-GAMMA BAND

Classifier	Precision	Recall	F1-score	AUC
CNN	96.5%	96.3%	96.2%	99.1%
MCNN	96.4%	96.6%	96.7%	99.0%
MHCNN	97.2%	97.0%	97.0%	99.2%

80 times, and the curve is relatively smoothest after the stability.

The average verification accuracy of the multispectral image data constructed based on the MCMI method in the three CNN models under the condition of Alpha1-Alpha2-Gamma band combination is shown in Fig. 8(b). According to Fig. 8(b), the following conclusions can be drawn: the curve of CNN is the lowest and tends to be stable between 0.96 and 0.97 when the iteration reaches 103 times. The curve of MCNN is the next, which tends to be stable between 0.96 and 0.97 when the iteration reaches 90 times. MHCNN has the highest curve, which tends to be stable between 0.97 and 0.98 when the iteration reaches 75 times, and the curve is relatively smoothest after the stability.

Table VII shows the average Precision, F1-score, Recall, and AUC evaluation values of the multispectral image data constructed based on the MCMI method under the condition of the combination of Theta-Beta2-Gamma band in the three CNN models. It can be seen from the table that the multispectral image data in the MHCNN model are significantly better than the 4 evaluation values of CNN and MCNN.

B. Classification Results Under Different Frequency Combinations Based on MHCNN

Under the combination of Theta-Alpha2-Gamma, Alpha2-Beta2 Gamma, Beta1 Beta2 Gamma, Theta-Beta2 Gamma, Theta-Beta2 Gamma, Theta-Alpha1-gamma, Alpha1-Alpha2-Gamma band, MCMI processed the data to obtain the corresponding six kinds of multispectral image data sets. The six kinds of different feature sets are classified by the multi-scale high-density convolutional neural network, and the obtained results are presented and compared. Since the average verification accuracy and average verification loss of these six features sets all reach saturation when the iterations reach 80∼90 times, only the average verification loss and average verification accuracy between 80 and 120 iterations are shown for the convenience of observation.

According to Fig. 9, the verification loss rates of Beta1-Beta2-Gamma and Theta-Beta2-Gamma are the lowest, which

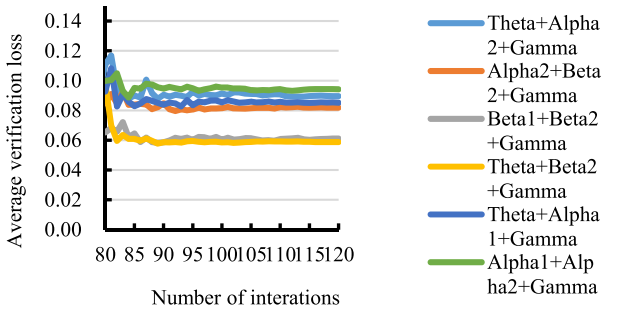

Fig. 9. Average verification accuracy curves under each frequency band combination.

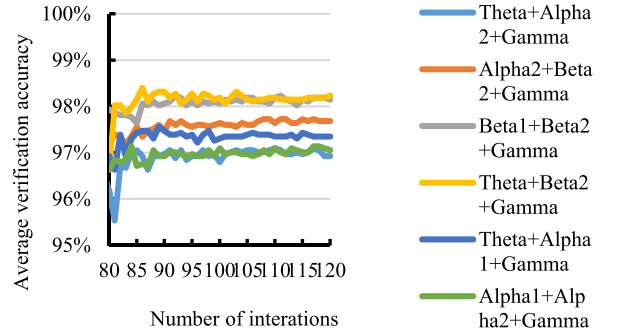

Fig. 10. Average verification accuracy curves under each frequency band combination.

TABLE VIII
AVERAGE EVALUATION INDEX UNDER THE CONDITIONS OF SIX FREQUENCY BAND COMBINATIONS

Band combination	Precision	Recall	F1-score	AUC
Theta-Alpha2-Gamma	96.9%	97.2%	97.2%	99.1%
Alpha2-Beta2-Gamma	97.6%	97.9%	97.9%	99.3%
Beta1-Beta2-Gamma	97.7%	98.0%	98.0%	99.4%
Theta-Beta2-Gamma	98.0%	98.1%	98.0%	99.4%
Theta-Alpha1-Gamma	97.4%	97.4%	97.5%	99.3%
Alpha1-Alpha2-Gamma	97.2%	97.0%	97.0%	99.2%

stabilize to 0.059 and 0.060, respectively. Alpha2-Beta2-Gamma, Theta-Alpha1-Gamma, and Theta-Alpha2-Gamma verification loss rates are the next. Alpha1-Alpha2-Gamma has the highest verification loss rate, which is stable between 0.09 and 0.1.

According to Fig. 10, the verification accuracy of Beta1-Beta2-Gamma and Theta-Beta2-Gamma is the highest, which stabilize to 0.981 and 0.982, respectively. Verification accuracy of Alpha2-Beta2-Gamma is the next, which is stable at 0.979. And Theta-Alpha1-Gamma is stable at 0.975 and 0.974, respectively. Verification accuracy of Theta-Alpha2-Gamma and Alpha1-Alpha2-Gamma is the lowest, which are stable at about 0.97.

Table VIII shows that under the combination of the Theta-Beta2-Gamma band and Beta1-Beta2-Gamma band, the four evaluation indexes obtained by CNN classification are all higher than the other four evaluation indexes obtained. Theta-Alpha2-Gamma is the lowest of the four evaluation indexes.

In conclusion, under the condition of the combination of Theta-Beta2-Gamma and Beta1-Beta2-Gamma bands, the

classification method combining MCMI and multi-scale high-density convolutional neural network for processing the data in this study achieves the optimal evaluation indexes. Therefore, the combination of Theta-Beta2-Gamma and Beta1-Beta2-Gamma bands can be preferred when using this classification method to deal with similar classification tasks.

IV. Discussion

A. Performance Analysis of MHCNN

In the field of image classification, CNN has become the most widely used deep learning classifier with better performance than other classifiers due to its ability to consider both temporal and spatial information [34]. To improve the performance of CNN, researchers have carried out optimization and improvement in many aspects. On the one hand, to change the structure of CNN, the multi-size convolution kernel could be used for each layer. In terms of convolutional layer connection, the residual connection could also be used to realize skipping transfer or make each layer integrate the feature output of other layers [35]–[37]. On the other hand, the algorithm adjusted the parameters. For example, the weight initialization method, activation function, optimization function/loss function, learning rate, and other improvements were made [38]–[40].

Studies showed that the training time of traditional convolutional neural networks was too long, the accuracy was not ideal, and could not get diversified image features. Gradient dispersion would be generated in the process of backpropagation [41]–[44]. Based on these studies, we designed a multi-scale high-density convolutional neural network based on the multi-scale convolutional neural network and Densenet structure. In terms of model optimization, we applied the adaptive stochastic gradient descent method to the multi-scale high-density convolutional neural network. This algorithm could effectively avoid the problems of slow training speed of deep learning model and ignoring the global optimal solution, and then better evaluate the effect of spatial cognitive training. Compared with the classical CNN model, our proposed model has been greatly improved because of the superiority of the CNN framework. It can be seen from the results in this paper that MHCNN has a faster convergence speed and higher accuracy under the combination of six frequency bands.

1) The Clinical Application Value of the MHCNN: In this experiment, the classification effect of image data constructed based on MCMI in three CNN models was compared under six different frequency band combinations. It was found that the multi-scale high-density convolutional neural network classification model based on MCMI had the best performance under whatever frequency band combinations. It had obvious advantages in the combination of six frequency bands.

In addition, the classification results of MCMI data in the multi-scale high-density convolutional neural network model under different frequency band combinations were separately observed. It was found that no matter in the average verification loss, average verification accuracy, or each evaluation value of the CNN model, the CNN model is based on the combination of Theta-Beta2-Gamma and Beta1-Beta2-Gamma bands had the best classification effect. Therefore, it can be concluded that the multi-scale high-density convolutional neural network based on the combination of Theta-Beta2-Gamma and Beta1-Beta2- Gamma bands is the most effective method to evaluate the effect of spatial cognitive training. This result is similar to some studies on spatial cognitive EEG [12], [45]–[50]. In these studies, the EEG signal analysis results of the combination of Theta-Beta2-Gamma and Beta1-Beta2- Gamma bands were closely related to spatial cognitive training.

Considering the shortcomings of the spatial cognitive ability recognition and analysis method based on EEG signals in this research, further improvements will be made in the future. The experimental data are from the BCI-VR spatial cognitive training and evaluation system independently developed by the laboratory, and the data set used is not as large as the data set usually used by deep learning. So more relevant EEG signals can be collected to verify the universality of the method proposed in this paper. In addition, it is also a good practice to effectively evaluate spatial cognition in virtual environments [50]–[53]. At the same time, our proposed model aims to verify the effectiveness of cognitive training and can be applied to clinical medicine.

For the improvement of network structure, the size and number of multi-scale convolution cores, and the combination with dense blocks can be optimized in the future. Actually, spatial information can be found by analyzing different EEG signals, but the current model did not deal with it yet. Therefore, we will take the spatial characteristics into consideration referring to the methods like TRCA [54] and CSP [55]. In addition to the EEG feature analysis, it will be of great significance to find the reasons for the transformation and key features, combined with the existing brain theory. Our proposed model, which is designed for clinical use, is utilized to verify the effectiveness of cognitive training. In the future, we can invite more subjects to participate in our experiment as the same as the reference [56], for the verification of reliability of the model. We hope that our proposed method can be combined with the online training, to monitor the changes of subjects' ability in real-time. Moreover, it will be of great significance to analyze the causes of EEG changes as in [57], combined with the existing knowledge of brain science.

V. Conclusion

The classification performance of the multi-scale high-density convolutional neural network method was better than that of the classical convolutional neural network method and the multi-scale convolutional neural network method. The new classification method could be used as an effective biological indicator of spatial cognitive training effect, which provided a new perspective for spatial cognitive ability evaluation and analysis.

Acknowledgment

The authors have no potential conflicts.

References

[1] E. A. Johnson, "A study of the effects of immersion on short-term spatial memory," M.S. thesis, College Technol., Purdue Univ., West Lafayette, IN, USA, 2010.

[2] J. L. Whitwell *et al.*, "3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease," *Brain, J. Neurol.*, vol. 130, no. 7, pp. 1777–1786, May 2007.

[3] R. Hornero, D. Abásolo, J. Escudero, and C. Gómez, "Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease," *Philos. Trans. Roy. Soc. A, Math. Phys. Eng. Sci.*, vol. 367, no. 1887, pp. 317–336, Jan. 2009.

[4] R. S. Astur, M. L. Ortiz, and R. J. Sutherland, "A characterization of performance by men and women in a virtual Morris water task: A large and reliable sex difference," *Behav. Brain Res.*, vol. 93, nos. 1–2, pp. 185–190, Jun. 1998.

[5] R. Masoomi and A. Khadem, "Enhancing LDA-based discrimination of left and right hand motor imagery: Outperforming the winner of BCI competition II," in *Proc. 2nd Int. Conf. Knowl.-Based Eng. Innov. (KBEI)*, Nov. 2015, pp. 392–398.

[6] M. Manjusha and R. Harikumar, "Performance analysis of KNN classifier and K-means clustering for robust classification of epilepsy from EEG signals," in *Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, Mar. 2016, pp. 2412–2416.

[7] R. S. S. Kumari and J. P. Jose, "Seizure detection in EEG using time frequency analysis and SVM," in *Proc. Int. Conf. Emerg. Trends Electr. Comput. Technol.*, Mar. 2011, pp. 626–630.

[8] K. D. Tzimourta *et al.*, "Machine learning algorithms and statistical approaches for Alzheimer's disease analysis based on resting-state EEG recordings: A systematic review," *Int. J. Neural Syst.*, vol. 31, no. 5, May 2021, Art. no. 2130002.

[9] K. Akrofi, M. C. Baker, M. W. O'Boyle, and R. B. Schiffer, "Clustering and modeling of EEG coherence features of Alzheimer's and mild cognitive impairment patients," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 1092–1095.

[10] K. Akrofi, R. Pal, M. C. Baker, B. S. Nutter, and R. W. Schiffer, "Classification of Alzheimer's disease and mild cognitive impairment by pattern recognition of EEG power and coherence," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2010, pp. 606–609.

[11] M. Baker, K. Akrofi, R. Schiffer, and M. W. O. Boyle, "EEG patterns in mild cognitive impairment (MCI) patients," *Open Neuroimag. J.*, vol. 2, pp. 52–55, Aug. 2008.

[12] X. Li, Y. Yan, and W. Wei, "Identifying patients with poststroke mild cognitive impairment by pattern recognition of working memory load-related ERP," *Comput. Math. Methods Med.*, vol. 2013, Sep. 2013, Art. no. 658501.

[13] J. C. McBride *et al.*, "Spectral and complexity analysis of scalp EEG characteristics for mild cognitive impairment and early Alzheimer's disease," *Comput. Methods Programs Biomed.*, vol. 114, no. 2, pp. 153–163, Apr. 2014.

[14] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," in *Proc. 4th Int. Conf. Learn. Represent. (ICLR)*, May 2016, pp. 1–15.

[15] J. R. Wolpaw, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, Jun. 2000.

[16] D. Wen *et al.*, "Estimating coupling strength between multivariate neural series with multivariate permutation conditional mutual information," *Neural Netw.*, vol. 110, pp. 159–169, Feb. 2019.

[17] A. Seth, J. M. Vance, and J. H. Oliver, "Virtual reality for assembly methods prototyping: A review," *Virtual Reality*, vol. 15, no. 1, pp. 5–20, 2011.

[18] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.

[19] R. Li, "Task state EEG recognition for spatial cognitive assessment," M.S. thesis, Dept. Comput. Sci. Technol., Yanshan Univ., Qinhuangdao, China, 2019.

[20] J. Pan *et al.*, "Prognosis for patients with cognitive motor dissociation identified by brain-computer interface," *Brain*, vol. 143, no. 4, pp. 1177–1189, Apr. 2020.

[21] H. Yang, J. Han, and K. Min, "A multi-column CNN model for emotion recognition from EEG signals," *Sensors*, vol. 19, no. 21, p. 4736, Oct. 2019.

[22] J. Yuan, "Research on spatial cognitive training and evaluation system integrating brain-computer interface and virtual reality," M.S. thesis, Dept. Comput. Sci. Technol., Yanshan Univ., Qinhuangdao, China, 2019.

[23] D. Wen *et al.*, "Resting-state EEG signal classification of amnestic mild cognitive impairment with type 2 diabetes mellitus based on multispectral image and convolutional neural network," *J. Neural Eng.*, vol. 17, no. 3, Jun. 2020, Art. no. 036005.

[24] D. Wen *et al.*, "Multi-dimensional conditional mutual information with application on the EEG signal analysis for spatial cognitive ability evaluation," *Neural Netw.*, vol. 148, pp. 23–36, Apr. 2022.

[25] C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, and F. Morabit, "A convolutional neural network approach for classification of dementia stages based on 2D-spectral representation of EEG recordings," *Neurocomputing*, vol. 323, pp. 96–107, Jan. 2019.

[26] O. Tsinalis, P. M. Matthews, Y. Guo, and S. Zafeiriou, "Automatic sleep stage scoring with single-channel EEG using convolutional neural networks," 2016, *arXiv:1610.01683*.

[27] F. C. Morabito *et al.*, "Deep convolutional neural networks for classification of mild cognitive impaired and Alzheimer's disease patients from scalp EEG recordings," in *Proc. IEEE 2nd Int. Forum Res. Technol. Soc. Ind. Leveraging Better tomorrow (RTSI)*, Sep. 2016, pp. 1–6.

[28] J. Jeong, J. C. Gore, and B. S. Peterson, "Mutual information analysis of the EEG in patients with Alzheimer's disease," *Clin. Neurophysiol.*, vol. 112, no. 5, pp. 827–835, May 2001.

[29] D. Li, J. Xu, J. Wang, X. Fang, and Y. Ji, "A multi-scale fusion convolutional neural network based on attention mechanism for the visualization analysis of EEG signals decoding," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 2615–2626, Dec. 2020.

[30] W. Ko, E. Jeon, S. Jeong, and H.-I. Suk, "Multi-scale neural network for EEG representation learning in BCI," *IEEE Comput. Intell. Mag.*, vol. 16, no. 2, pp. 31–45, May 2021.

[31] X. Yu, J. Fan, J. Chen, P. Zhang, and L. Han, "NestNet: A multi-scale convolutional neural network for remote sensing image change detection," *Int. J. Remote Sens.*, vol. 42, no. 13, pp. 4902–4925, 2021.

[32] X.-H. Han, Y. Zheng, and Y.-W. Chen, "Multi-level and multi-scale spatial and spectral fusion CNN for hyperspectral image super-resolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 1–10.

[33] Z. Dong, P. Zhao, J. Zheng, K. Ji, Y. Chen, and J. Fu, "Intelligent injection molding: Parameters self-learning optimization using iterative gradient-approximation adaptive method," *J. Appl. Polym. Sci.*, vol. 138, no. 29, p. 50687, Aug. 2021.

[34] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[35] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2016, pp. 1–7.

[36] M. Marsden, K. McGuinness, S. Little, and N. E. O'Connor, "ResnetCrowd: A residual deep learning architecture for crowd counting, violent behaviour detection and crowd density level classification," in *Proc. 14th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Aug. 2017, pp. 1–7.

[37] Z. Wu, C. Shen, and A. van den Hengel, "Wider or deeper: Revisiting the ResNet model for visual recognition," *Pattern Recognit.*, vol. 90, pp. 119–133, Jun. 2019.

[38] T. Jiang and J. Cheng, "Target recognition based on CNN with LeakyReLU and PReLU activation functions," in *Proc. Int. Conf. Sens., Diag., Prognostics, Control (SDPC)*, Aug. 2019, pp. 718–722.

[39] S. Koturwar and S. Merchant, "Weight initialization of deep neural networks(DNNs) using data statistics," 2017, *arXiv:1710.10570*.

[40] P. Doležel, P. Škrabánek, and L. Gago, "Weight initialization possibilities for feedforward neural network with linear saturated activation functions," *IFAC-PapersOnLine*, vol. 49, no. 25, pp. 49–54, 2016.

[41] H.-R. Xiang *et al.*, "Chinese character recognition based on residual separable convolutional neural network," in *Proc. 15th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. (ICCWAMTIP)*, Dec. 2018, pp. 49–53.

[42] S. Jiang *et al.*, "Dual attention dense convolutional network for intelligent fault diagnosis of spindle-rolling bearings," *J. Vib. Control*, vol. 27, nos. 21–22, pp. 2403–2419, Nov. 2021.

[43] X.-H. Han, Y. Zheng, and Y.-W. Chen, "Multi-level and multi-scale spatial and spectral fusion CNN for hyperspectral image super-resolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 1–10.

[44] J. Fu and J. Liang, "Virtual view generation based on 3D-dense-attentive GAN networks," *Sensors*, vol. 19, no. 2, p. 344, Jan. 2019.

[45] D. Wen, Q. Xue, C. Lu, X. Guan, Y. Wang, and X. Li, "A global coupling index of multivariate neural series with application to the evaluation of mild cognitive impairment," *Neural Netw., Off. J. Int. Neural Netw. Soc.*, vol. 56, pp. 1–9, Aug. 2014.

[46] D. J. White, M. Congedo, J. Ciorciari, and R. B. Silberstein, "Brain oscillatory activity during spatial navigation: Theta and gamma activity link medial temporal and parietal regions," *J. Cogn. Neurosci.*, 24, no. 3, pp. 686–697, 2011.

[47] Z. Dai *et al.*, "EEG cortical connectivity analysis of working memory reveals topological reorganization in theta and alpha bands," *Frontiers Hum. Neurosci.*, vol. 11, p. 237, May 2017.

[48] F. Bartolomei, F. Wendling, J.-J. Bellanger, J. Régis, and P. Chauvel, "Neural networks involving the medial temporal structures in temporal lobe epilepsy," *Clin. Neurophysiol.*, vol. 112, no. 9, pp. 1746–1760, Sep. 2001.

[49] D. Wen, Z. Bian, Q. Li, L. Wang, C. Lu, and X. Li, "Resting-state EEG coupling analysis of amnestic mild cognitive impairment with type 2 diabetes mellitus by using permutation conditional mutual information," *Clin. Neurophysiol.*, vol. 127, no. 1, pp. 335–348, Jan. 2016.

[50] D. Wen *et al.*, "The EEG signal analysis for spatial cognitive ability evaluation based on multivariate permutation conditional mutual information-multi-spectral image," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 10, pp. 2113–2122, Oct. 2020.

[51] D. Wen *et al.*, "Combining brain–computer interface and virtual reality for rehabilitation in neurological diseases: A narrative review," *Ann. Phys. Rehabil. Med.*, vol. 64, no. 1, Jan. 2021, Art. no. 101404.

[52] Y. Zhou *et al.*, "The current research of spatial cognitive evaluation and training with Brain–Computer interface and virtual reality," *Frontiers Neurosci.*, vol. 13, p. 1439, Feb. 2020.

[53] D. Wen, X. Lan, Y. Zhou, G. Li, S.-H. Hsu, and T.-P. Jung, "The study of evaluation and rehabilitation of patients with different cognitive impairment phases based on virtual reality and EEG," *Frontiers Aging Neurosci.*, vol. 10, p. 88, Apr. 2018.

[54] J. Jin, Z. Wang, R. Xu, C. Liu, X. Wang, and A. Cichocki, "Robust similarity measurement based on a novel time filter for SSVEPs detection," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 14, 2021, doi: 10.1109/TNNLS.2021.3118468.

[55] J. Jin, R. Xiao, I. Daly, Y. Miao, X. Wang, and A. Cichocki, "Internal feature selection method of CSP based on L1-norm and dempster-shafer theory," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4814–4825, Nov. 2020.

[56] J. Jin *et al.*, "The study of generic model set for reducing calibration time in P300-based brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 3–12, Jan. 2020.

[57] H. Jiang, D. Wu, R. Jiao, and Z. Wang, "Analytical comparison of two emotion classification models based on convolutional neural networks," *Complexity*, vol. 2021, Feb. 2021, Art. no. 6625141.