

RESEARCH ARTICLE

A Visual Characteristic Land-Scape Design for EEG Signal Based on LSTM-GAN

JUAN TAN¹, BAOCHEN WU², AND YUNLONG MA³¹College of Fine Arts and Design, Hunan University of Humanities, Science and Technology, Loudi 417000, China²China Forestry Group Zhonglin Times Holding Company Ltd., Shanghai 200000, China³Hunan University of Humanities, Science and Technology, Loudi 417000, China

Corresponding author: Juan Tan (15273858188@163.com)

ABSTRACT The ultimate goal of artificial intelligence is to endow machines with human intelligence. Studying and simulating electroencephalogram signals is a way to achieve this goal. The human brain has similar representation abilities for similar visual stimuli. By utilizing this feature, a visual stimulus electroencephalogram signal decoding model based on Long Short-Term Memory Network Bagging was proposed to decode and classify human brain signals. And based on this extracted classification model, a generative adversarial network based on a bi-directional short-term memory network was proposed. It could generate similar visual stimulus images similar to the human brain and represent the visual signals of the human brain. These experiments confirmed that the classification accuracy of the research method in the decoding of electroencephalogram signals reached 91.17%. In terms of extracting visual characteristics and land-scape features from the electroencephalogram, this research model had the highest classification accuracy and recall rates, with 98.38% and 97.94%, respectively. This stimulation image generation model studied had the best actual image generation performance, with an Inception score of 7.27. The study not only improves the accuracy of electroencephalogram signal classification, but also completes the re-construction of brain signals into images. It improves the collaborative representation ability of human-machine collaborative visual cognitive systems and has important significance in brain computer interaction.

INDEX TERMS Electroencephalogram visual decoding, Bi-LSTM, LSTM-B-GAN, generate adversarial re-construction, brain computer interaction.

I. INTRODUCTION

As computers rapidly develop, research on Artificial Intelligence (AI) is also receiving increasing attention. The ultimate goal of its research is to make machines closer to human thinking and capable of possessing human intelligence. To achieve this goal, it is necessary to study and simulate the human brain. The correspondence between Electroencephalogram (EEG) activity and visual characteristics is currently the main frontier field of brain science [1], [2]. Therefore, it is of great significance in Brain Computer Interaction (BCI) to study an effective EEG signal processing technology and visual representation techniques

for decoding EEG signals. However, there are still various problems with existing brain computer visual representation technologies. For example, the accuracy of EEG signal classification is low, the understanding of the machine's brain like land-scape is unreasonable and lacks biological basis, the collaborative representation ability is poor during BCI, and the resolution of generated images is poor [3]. In response to the low accuracy of EEG signal classification, a visual stimulation EEG signal decoding model based on spatiotemporal features is proposed for high-precision EEG signal decoding. Unreasonable understanding of the brain like land-scape of machines can lead to lower classification accuracy compared to EEG classification algorithms. In this regard, a method based on Long Short-Term Memory-Bagging (LSTM-B) regression classification is proposed

The associate editor coordinating the review of this manuscript and approving it for publication was Gustavo Olague .

in this experiment based on spatiotemporal features. This method maps human brain visual information onto machines, enabling machines to better annotate land-scapes based on human brain vision. In response to the poor representation ability of BCI, a Long Short-Term Memory-B-Generate Adversarial Network (LSTM-B-GAN) based on spatiotemporal features is proposed in this experiment to constrain the image generation process and improve the resolution of the generated image.

The LSTMS-B analysis method is proposed to guide joint decision-making among multiple networks to improve performance. In response to the weak representation ability in human-machine collaborative visual cognitive systems, a transformation of “image EEG signal image” is implemented. The attention mechanism is added to the Bi-LSTM network through attention gates and attention weighting, and the Bi-LSTM-AttGW model is proposed to decode EEG signals.

The combination of GAN and visual EEG signal image re-construction model can convert EEG signals into corresponding land-scape design images, providing technical reference for communication and BCI between hearing impaired and mute people.

First, the research status of EEG visual decoding related research is introduced. Secondly, the specific design of EEG visual feature land-scape mapping method based on LSTM-GAN is introduced. The third part is to verify the algorithm performance and practicality of the research method through simulation experiments. Finally, a summary and analysis of the entire content is conducted.

II. RELATED WORKS

With the continuous development of AI, the related fields of BCI are also receiving increasing attention. Many scholars have conducted relevant research on EEG signal decoding. Robustness and computational complexity are key challenges in developing motion imagination based on EEG signal decoding in practical brain computer interface systems. In response, M T Sadiq et al. proposed an automatic multi-variate empirical wavelet transform algorithm to decode different motion imagination tasks. These experiments confirmed that the classification performance of this method was better than existing methods, and the classification accuracy had been improved by 23.50% [4]. Noise and other signal sources can interfere with large EEG capacity, making it difficult for EEG classifiers to improve and have limited generalization ability. J F Hwaidi and T M Chen proposed a new classification method for EEG signals and motion imagination signals. This method eliminated noise in the signal by using a variational autoencoder. Then, in the experiment, a deep automatic encoder and convolutional framework were combined to classify EEG motion imagination signals, and the feasibility of the research model was confirmed through simulation experiments [5]. D Li et al. proposed a multi-scale fusion Convolutional Neural Network (CNN) based on Attention Mechanism (AM), considering

the insufficient variability of simple network frameworks to meet complex EEG decoding tasks. This network extracted spatiotemporal and multi-scale features from signals represented by multiple brain regions, which was supplemented by a dense fusion strategy to preserve the maximum information flow. The effectiveness of the research method was verified through experiments, and it was also demonstrated that AM had a positive role in analyzing EEG signal decoding [6]. D H Lee et al. attempted to classify pilots' mental states using only EEG signals during the continuous decoding process. They proposed a multi-feature block-based CNN and spatiotemporal EEG filter to identify pilots' current mental states, demonstrating the feasibility of classifying various types of mental states in real environments through experiments [7].

In terms of research related to EEG vision, X Geng D et al. considered that EEG signals based on brain computer interface devices had weak, non-linear, non-stationary, and time-varying characteristics. Therefore, they proposed a combined EEG signal processing method based on Independent Component Analysis (ICA), Wavelet Transform (WT), and Common Spatial Pattern (CSP). The high accuracy of this method was verified through cross comparison experiments [8]. Ahirwal MK et al. proposed a new channel selection technique for the recognition and characterization of visual stimulation EEG signals. This method extracted three types of features from EEG signals through EEG channels and used Support Vector Machine (SVM), artificial Neural Network (NN), and Naive Bayesian algorithm to classify visual scenes. These experiments confirmed that the average accuracy of this method was 90.53% [9]. P Nagabushanam et al. considered the combined effect of feature extraction and classification availability in deep learning algorithms. To improve the performance of EEG classification, they proposed a two-layer LSTM and a four-layer improved NN deep learning algorithm. Compared to other related architectures, the improved model provided better performance. However, the accuracy of visual image classification for this model was still below 90%, and further research was needed [10]. Long-term attention to repetitive visual stimuli can cause physiological and psychological fatigue, making subjects unable to concentrate enough, resulting in increased difficulty in decoding visual evoked potential EEG. In response, Z Gao et al. proposed a parallel multi-scale CNN based on AM. This network extracted advanced feature representations through spatial and temporal fusion of two consecutive convolutional blocks. These experiments confirmed the classification performance of this method [11].

In summary, there are still issues in current BCI related research that do not consider EEG spatial information, such as a lack of biological basis and weak BCI collaborative representation ability. Based on this, the decoding and classification of EEG visual signals under visual stimulation, the construction of EEG visual characteristic land-scape understanding models, and EEG visual characteristic land-scape re-construction are studied based on the spatiotemporal

EEG visual signal features. This can improve the accuracy of EEG signal decoding and classification, as well as the quality of land-scape re-constructed images.

III. EEG VISUAL CHARACTERISTIC LAND-SCAPE DESIGN BASED ON LSTM-GAN

The decoding and classification of visual stimulation EEG signals, the construction of EEG visual characteristic land-scape understanding models, and EEG visual characteristic land-scape re-construction are studied. This can improve the accuracy of EEG signal decoding and classification, as well as the quality of land-scape re-construction images, enabling machines to better represent the recognized land-scape.

A. A METHOD FOR FEATURE EXTRACTION AND CLASSIFICATION OF VISUAL STIMULATION EEG SIGNALS BASED ON LSTM-B

Due to the fact that most machines collect EEG signals as one-dimensional signals, neglecting the temporal and spatial information of their own brain signals, traditional visual stimulation EEG signal extraction methods have a problem of low feature recognition accuracy [12]. A classification model for decoding visual stimulation EEG signals is established based on the spatiotemporal features of EEG signals to improve the classification accuracy of visual stimulation brain signals. Figure 1 shows a spatiotemporal feature extraction method for EEG visual signals.

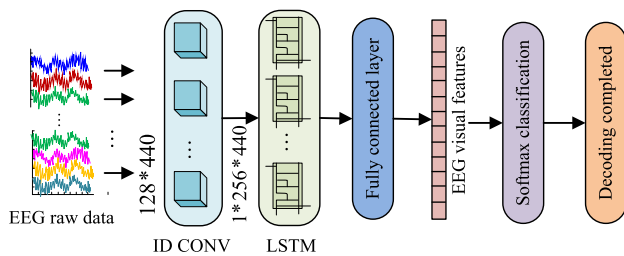


FIGURE 1. A spatiotemporal feature based classification model for EEG sensory signals.

In Figure 1, the model mainly consists of two modules: spatial convolution and temporal LSTM. The general process of this model is as follows. Firstly, the raw EEG signal with a size of 128*440 (128 EEG signal leads and 440 time sampling points) is input into the spatial convolution module of the model. The spatial convolution module consists of four one-dimensional convolutional layers connected in series. Through one-dimensional convolution, the correlations of different leads at corresponding spatial scales are obtained within corresponding time intervals, ensuring the temporal integrity of EEG signal. Secondly, the EEG signals processed by the spatial convolution module is input into the temporal LSTM module. On the basis of completing spatial correlation processing, this module processes each tensor data in the temporal dimension to ensure the integrity of visual stimulation EEG signal features. Finally, Softmax classification is

performed on the processed EEG visual features to complete the decoding and extraction of visual stimulation EEG feature signals.

After decoding and extracting feature information from EEG signals, there is a major challenge in completing BCI that how to represent the extracted visual information on the machine so that the machine can understand visual land-scape information remains. The existing EEG signal classification methods related to visual land-scape understanding also have the problem of low classification accuracy. Research will combine deep learning based on LSTM and ensemble learning based on Bagging on the basis of extracted EEG visual signal feature information. Furthermore, an LSTMS-B human brain activity decoding model is established in this experiment to implement a visual object classification model based on images, improving classification accuracy and generalization ability. This can enable machines to annotate scenes according to human visual abilities, understanding EEG visual characteristics and land-scapes. LSTMS-B mainly consists of improved LSTMS and improved Bagging. LSTMS is a model that combines LSTM and Swish activation functions, which introduces a new reinforcement learning-based Swish activation function to improve LSTM. Equation (1) is the Swish activation function [13].

$$f(x) = x * \text{sigmoid}(\beta x) \quad (1)$$

In Equation (1), $f(x)$ represents the Swish activation function. β represents any trainable hyperparameter or constant. The main function of Bagging is to improve the model's complexity training ability. This study will improve the generalization performance and prevent overfitting problems of the model by improving Boosting in Figure 2.

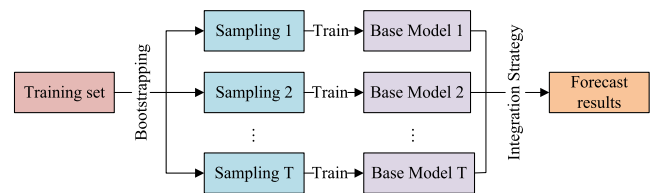


FIGURE 2. Schematic diagram of the improved Boosting algorithm.

In Figure 2, Boosting model has the following steps. First, T random samples are taken through Bootstrapping in the training set, and T sampling sets are extracted. Then T basis classifiers are trained using T sampling sets. Finally, the trained base classifier is integrated into a strategy to obtain strong classifier prediction results. Bootstrapping is a random sampling method, and the data that are not selected by Bootstrapping in the original data are Out of Bag (OOB). This study will improve Boosting's voting strategy on the foundation of OOB data, which uses OOB data to calculate the reliability coefficient of base classifiers and select base classifiers for different categories of data. This can avoid the algorithm's excessive reliance on high-performance classifiers and enhance the algorithm's generalization

performance. Equation (2) is the reliability coefficient.

$$C_{in} = \frac{OOB_F_{in}}{\sum_{n=1}^N OOB_F_{in}} \quad (2)$$

In Equation (2), C_{in} represents the n -th weak classifier's reliability coefficient on class i . i represents different categories. N represents the base classifiers number. OOB_F_{in} represents the F1 score of the n -th weak classifier's OOB data in class i . The high reliability coefficient of Boosting indicates good performance of the classifier on this dataset. This experiment combines LSTMS with improved Boosting to obtain the human visual land-scape activity decoding model LSTMS-B. Figure 3 shows the decoding schematic diagram of LSTMS-B's brain visual land-scape activity.

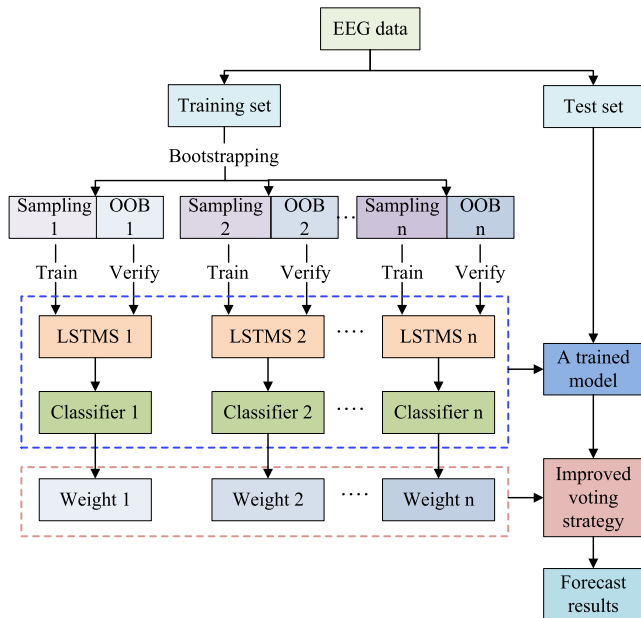


FIGURE 3. Schematic diagram of LSTMS-B brain visual decoding.

In Figure 3, LSTMS-B introduces LSTMS as a classifier to decode human visual activities based on the traditional Bagging. And the training results of classifier are weighted and verified using OOB data, and an improved voting strategy is used for decoding and classification. The general process of LSTMS-B is as follows. First, EEG stimulus image category data are divided into a training set and a testing set, and T sampling is performed on the training set through Bootstrapping to obtain T sampling sets. Secondly, LSTMS classifier is used to train the base classifier on sample set in T, and the OOB data corresponding to the training samples are collected to calculate the reliability coefficient. The corresponding weight of the classifier is obtained through the reliability coefficient. Finally, the base classifier is connected with the corresponding weights of the classifier, and an improved voting strategy is used to predict test set's visual land-scape category. Due to the improved voting algorithm and majority voting strategy in LSTMS-B increasing the complexity of the model, the study will introduce asymptotic time complexity

for algorithm complexity analysis. Equation (3) represents the computational complexity of the specific model.

$$T_{LSTMS-B} = \sum_{i=1}^{10} [4(h^*h + d^*h + h) + h^*I] \quad (3)$$

In Equation (3), $T_{LSTMS-B}$ represents the computational complexity of LSTMS-B. h represents the hidden layer's size. d represents EEG channels number, which is the input size. I represents the number of the categories. Equation (4) represents the number of the parameters.

$$T_{LSTMS} = 4(h^*h + d^*h + h) + h^*1 \quad (4)$$

In Equation (4), T_{LSTMS} represents the number of the parameters.

B. DESIGN OF EEG CHARACTERISTIC STIMULATION IMAGE GENERATION MODEL BASED ON LSTM-B IMPROVED GAN

This study completes the extraction of EEG visual specialties by constructing an LSTM-B model, enabling machines to understand and recognize EEG visual characteristic landscape. However, it is necessary to express the human brain visual landscape extracted by the machine to achieve BCI, that is, reconstructing the EEG visual landscape. Through human-machine interaction, machines can better represent the cognitive landscape. Therefore, a conditional spectral normalization generation adversarial network for visual stimulation re-construction will be established based on the decoding and classification of visual stimulation EEG signals using LSTM-B. AM and bi-directional LSTM are combined to analyze visual stimulation brain signals. While ensuring the accuracy of EEG signal classification, the number of the parameters of the model will be appropriately reduced and the training efficiency will be improved. This study introduces neural AM into a deep learning framework and proposes an impact separation method based on AM and LSTM. Figure 4 shows the internal schematic diagram of LSTM neurons introducing attention gates.

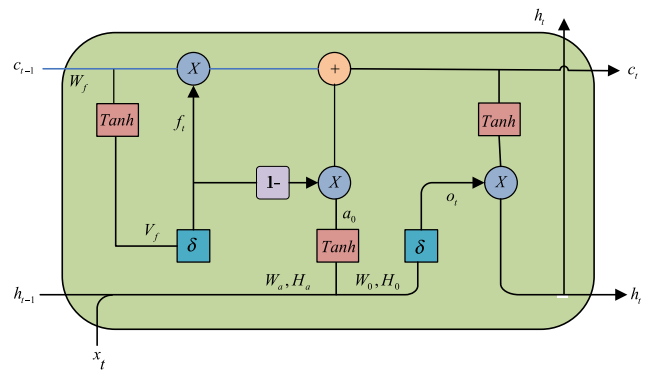


FIGURE 4. Internal schematic diagram of attention gate LSTM neurons.

In Figure 4, AM is used to replace the forget gate. The main function of the forget gate is to determine the location where historical information is deleted and add information

at that location, which plays a role in updating the cell state. Equation (5) represents the state update of the forget gate [14].

$$c_t = f_t^* c_{t-1} + (1 - f_t)^* a_t \quad (5)$$

In Equation (5), c_t represents the cell output state at the current t time. c_{t-1} represents the output state of the cell at the previous moment. a_t represents the cell candidate value. f_t represents the weight parameter of cell candidate values. In Figure 4, the research method uses attention gates instead of forgetting gates, while coupling the input gate with attention gates to reduce the internal parameters of NN, enabling neurons to capture more correlated historical location information. Equation (6) represents the improved neuron cell state update operation.

$$f_t = \delta(V_f^* \tanh(W_f^* c_{t-1})) \quad (6)$$

In Equation (6), f_t represents the update status of cell at the current t time. δ represents the weight of the candidate value. V_f and W_f are both parameters in attention gates. c_{t-1} represents the output state of neuron cell at the previous moment. This method has the effect of reducing the dimensionality of training parameters compared to ordinary long short-term attention networks. To increase the cognitive saturation of the model and focus too much attention on important EEG features, research will set attention weight coefficients in the output stage of model. Figure 5 shows a bi-directional LSTM with attention weighted layers.

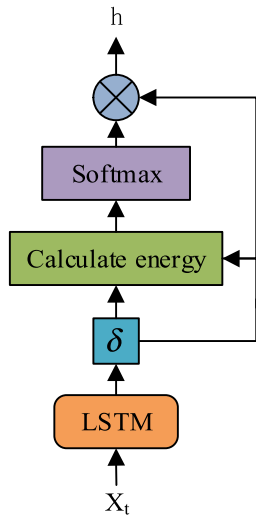


FIGURE 5. Bi-LSTM network with attention weighted layer.

Figure 5 defines the matrix $H = [h_1, h_2, \dots, h_T]$, which includes the output of all hidden layers. The energy of each feature is calculated by setting the attention weight matrix. Finally, the weight coefficients for feature selection are determined through Softmax operator calculation. Equation (7) represents the energy and weight coefficients.

$$\begin{cases} \text{Energy} = \text{ReLU}(H^* W_1)^* W_2 \\ \text{Weight} = \text{soft max}(\text{Energy}) \end{cases} \quad (7)$$

In Equation (7), *Energy* represents the attention feature energy. W_1 and W_2 represent the sample weights with the dimensions of $N \times N$ and $N \times 1$, respectively. *Weight* represents the attention weight, which is the contribution value of each time point to visual recognition. After attention weighting, Equation (8) is the final weight matrix used for the representation of the visual output.

$$h = H^* \text{Weight} \quad (8)$$

In Equation (8), h represents the weight matrix used for the representation of the visual output. H' represents the transposed matrix of the matrix H .

On the basis of improving the bi-directional LSTM classification of AM, this study will introduce the SNGAN method to analyze visual stimulation brain signals, that is, establishing LSTM-GAN to stimulate machine image generation. This model introduces spectral normalization technology to make discriminator stabilized and controls Lipschitz constant by constraining each layer's spectral norm of discriminator. Equation (9) represents the output and output relationship of the model.

$$x_n = a_n(W_n x_{n-1} + b_n) \quad (9)$$

In Equation (9), x_n represents the output value of the n -th layer network. x_{n-1} represents the input value of the n -th layer network. $a_n(\cdot)$ represents the n -th layer network's nonlinear activation function. W_n represents a parameter matrix. b_n represents the offset. The output gradient of this model after Lipschitz constraints is represented by Equation (10).

$$\begin{aligned} \|\nabla_x(f(x))\|_2 \\ = \|C_N W_N \dots C_1 W_1\|_2 \leq \|C_N\|_2 \|W_N\|_2 \dots \|C_1\|_2 \|W_1\|_2 a \end{aligned} \quad (10)$$

In Equation (10), $f(x)$ represents the output of the model. $\|\cdot\|$ represents a differential operator. $\|C_N\|_2$ represents the spectral norm of the matrix W_n . C_N represents the diagonal matrix of the corresponding layer. Due to the maximum spectral norm of the diagonal matrix being 1, Equation (10) can be simplified to ensure that the model output satisfies Lipschitz constraint. Equation (11) represents the output gradient of the simplified model.

$$\begin{aligned} \|\nabla_x(f(x))\|_2 \\ = \left\| C_N \frac{W_N}{\sigma(W_N)} \dots C_1 \frac{W_1}{\sigma(W_1)} \right\|_2 \leq \prod_{i=1}^N \frac{W_i}{\sigma(W_i)} = 1 \end{aligned} \quad (11)$$

In Equation (11), $\sigma(W_N)$ represents the maximum singular value of the matrix W . Lipschitz constraint = 1 can be satisfied by dividing the network parameters with the layer parameter matrix's spectral norm. In Figure 6, the spectral normalization technique is introduced to stabilize the discriminator for stimulating the image generation of LSTM-GAN.

In Figure 6, LSTM-GAN is mainly composed of two modules, including the LSTM-based EEG signal visual feature encoding and classification module and the GAN-based

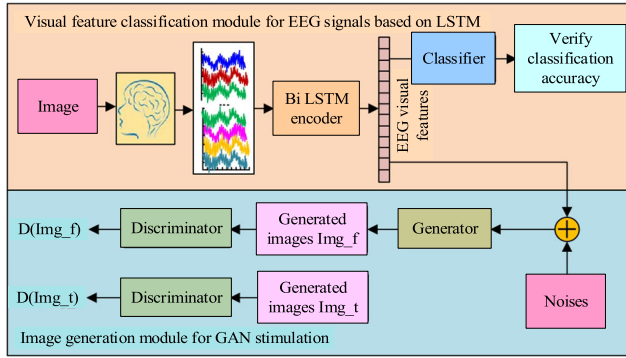


FIGURE 6. Schematic diagram of LSTM-GAN model.

stimulus image generation module. In the LSTM-based EEG signal visual feature classification module, the bi-directional LSTM fused with AM is studied as an encoder for this model. Low-dimensional visual EEG features are extracted from the original visual EEG signals of the human brain, processed, and finally encoded into EEG visual feature signals containing directional content. At this time, the eye EEG signals are associated with specific image categories. And this research will also verify and analyze the classification effectiveness of EEG features. The stimulation image generation module based on GAN will encode EEG visual features and influence noise in the experiment. And a GAN generation network will be added to generate images that are consistent with brain stimulation response. GAN generation network is a normalized generation adversarial network composed of discriminators and generators, corresponding to EEG feature categories, to generate images corresponding to EEG visual features. Due to the limited number of image samples used for EEG signal collection, this model will train GAN in two stages to fully utilize the selected images and their categories. In the first stage, GAN will be trained using ImageNet images that are not used for EEG signal collection. After the iteration is completed, the pre-trained model will be finely tuned using the corresponding EEG features in the second stage to generate higher image quality. To prevent excessive optimization of GAN by penalty gradients, this study will introduce Hinge function as objective loss function to evaluate the degree of optimization. Equations (12)-(13) represent the loss functions.

$$V_D(G, D) = E_{Z \sim P_{date}(x)}[\min(0, -1 + D(x|y))] + E_{Z \sim P_Z(x)}[\min(0, -1 - D(G(z|y)|y))] \quad (12)$$

In Equation (12), $V_D(G, D)$ represents the loss of the generated images. $P_{date}(x)$ represents the true sample distribution. $D(x|y)$ represents the discrimination result of a real image under the constraint of the conditional vector y . $P_Z(x)$ represents the distribution of the generated false samples. $G(z|y)$ represents the generation of the noise z and the condition vector y . E represents the expected operator. $D(G(z|y)|y)$ represents the discrimination result of the generated image

under y .

$$V_G(G, D) = E_{Z \sim P_Z(x)}[\min(0, -1 - D(G(z|y)|y))] \quad (13)$$

In Equation (13), $V_G(G, D)$ represents the generated image. $P_Z(x)$ represents the distribution of the generated false samples. $G(z|y)$ represents the generation of the noise z and the condition vector y . E represents the expected operator. $D(G(z|y)|y)$ represents the discrimination result of the generated image under y .

IV. EXPERIMENTAL AND ANALYSIS OF EEG VISUAL FEATURE DECODING CLASSIFICATION AND LAND-SCAPE GENERATION BASED ON LSTM

Experiments were conducted on different methods to verify the feasibility and superiority of this research model. They include LSTM EEG signal decoding and classification model based on spatiotemporal EEG visual signal features, LSTM-B image land-scape cognitive classification model based on EEG visual features, and LSTM-B-GAN generation adversarial image re-construction model considering AM.

A. VISUAL EEG SIGNAL FEATURE EXTRACTION AND CLASSIFICATION RESULTS BASED ON SPATIOTEMPORAL FEATURES

The visual stimulus image dataset for this study was derived from ImageNet dataset, and a subset of 20 image classes was selected as the dataset for this experiment. This data set contains 20 distinct and easily recognizable object images, each includes 50 images, or a total of 1000 images [15]. The specific information of the 20 image categories in the ImageNet data set is shown in Table 1. The images in this data set can be divided into several categories: food, animals, daily necessities, musical instruments, and transportation tools. This study used four experimenters to collect brain visual signals, with a total of 4000 visual feature EEG signals used for algorithm validation experiments. The hardware device configuration used in the study is as follows: processor Intel Core i7-8700k, main frequency 3.70GHz, and graphics processor NVIDIA GeForce GTX 1080. The software used in the study is Tensorflow1.8, Python 3.0 and MySQL5.7. The ratio of training data set to validation data set is 7:3. The layer number of LSTMS_B neural network is 1, the input dimension is 128, the hidden layer output dimension is 128, the output layer dimension is 40, and the batch size is 16, the optimizer is Adam, the learning rate is 0.001, and the attenuation factor is 0.5. Table 1 shows the selected ImageNet subset image categories.

Classification accuracy and F1 score were used as the evaluation criteria for the visual feature classification performance of EEG signals. And relevant models and research models were introduced for comparative experiments to more intuitively analyze the classification performance of this research model. Specifically, Linear Discriminant Analysis (LDA), LSTM+ReLU, EEGNet, CNN-Visual Geometry Group16 (CNN-VGG16), and CNN-ResNet101 were introduced [16].

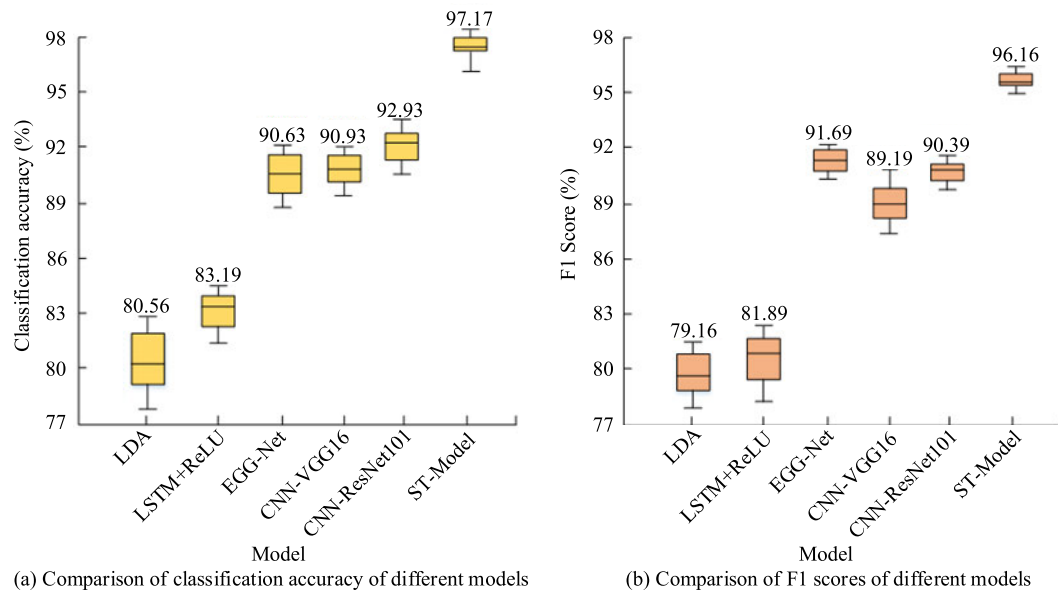


FIGURE 7. Comparison of classification accuracy and F1 scores of different models.

TABLE 1. ImageNet data set image category display.

Category number	Name	Category number	Name
0	Banana	10	Capuchin
1	Pizza	11	Hairstreak
2	Shoes	12	Coffee Machine
3	Mushroom	13	Electric guitar
4	Daisy	14	Aircraft
5	Glove	15	Piano
6	Bicycle	16	Pajamas
7	Convertible	17	Radio
8	Mug	18	Jack-o'-lantern
9	Panda	19	Tent

Figure 7 shows the classification results of different models.

In Figure 7 (a), the proposed EGG signal classification model based on spatiotemporal features had a relatively high classification accuracy of 97.17%. Among other classification methods for EEG signals, LDA had a relatively low classification accuracy of 80.56%, while CNN ResNet101 had a higher classification accuracy of 92.93%. Therefore, compared with common classification methods for EEG signals, the research algorithm had a significant improvement in classification accuracy. In Figure 7 (b), the F1 score of the research method was also relatively highest at 96.16%, and the relevant algorithms had F1 scores below 92%, indicating that the classification quality of the research model was relatively highest. Compared with common EEG

signal classification methods, there was also a significant improvement in the classification quality of the model. After the decoding and classification of EEG signals are completed, feature extraction was performed on EEG signals. The original LSTM, LSTM-ReLU (LSTM), LSTM-Swish (LSTMS), and LSTM-Minority Subject to Majority Principle Voting (LSTMS-MV) were introduced for comparative experiments to verify the generalization performance of the research model. The performance of EEG feature extraction model was evaluated in Figure 8 through two evaluation indicators: classification accuracy and recall rate.

In the model comparison in Figure 8, LSTMS had the highest classification accuracy and recall rate, with 91.39% and 90.89%, respectively, among the models that did not use voting strategies. Among the models considering voting strategies, LSTMS_B had the highest classification accuracy and recall rate, with classification accuracy and recall rates of 98.38% and 97.94%, respectively. The classification accuracy of the research model was 11.51% higher than LSTM, 3.09% higher than LSTMS-MV, and 6.99% higher than LSTMS, indicating that the classification performance of this research model was higher than that of the relevant models. Meanwhile, this research model had the highest recall rate, followed by LSTMS-MV, with LSTMS_B having a 2.56% higher recall rate than LSTMS-MV, indicating that the voting strategy of the research model was superior to traditional voting strategies. To further validate the feasibility and superiority of the research model, the model was trained under the stimulation of 20 types of images in the data set. Figure 9 shows the training results after 50 iterations.

In Figure 9, as the iteration increased, the classification accuracy of each model gradually increased, and finally reached a stable state in a relatively short time

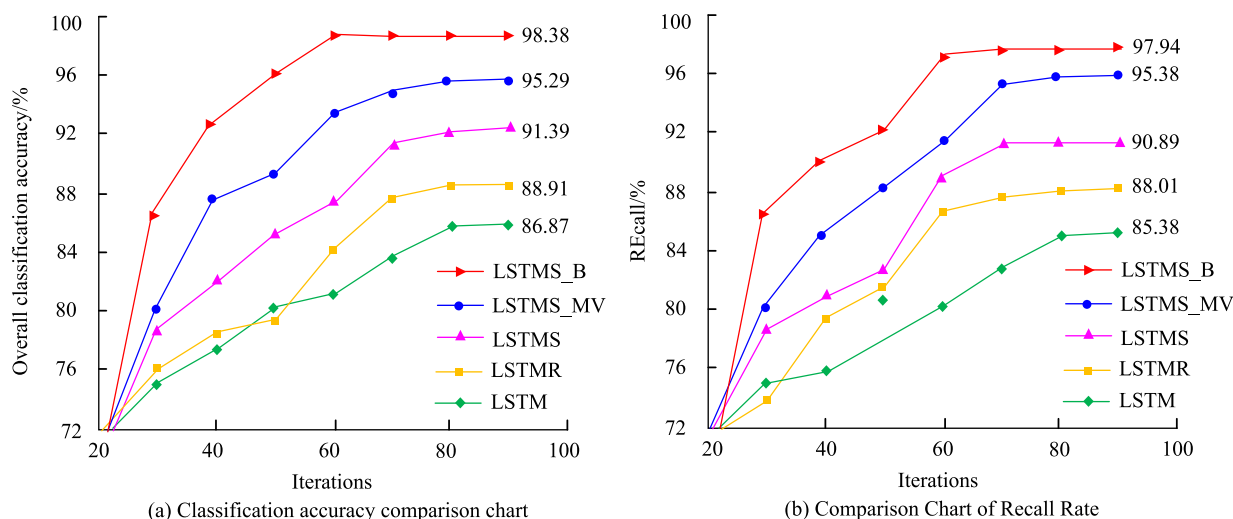


FIGURE 8. Comparison of classification results of different models.

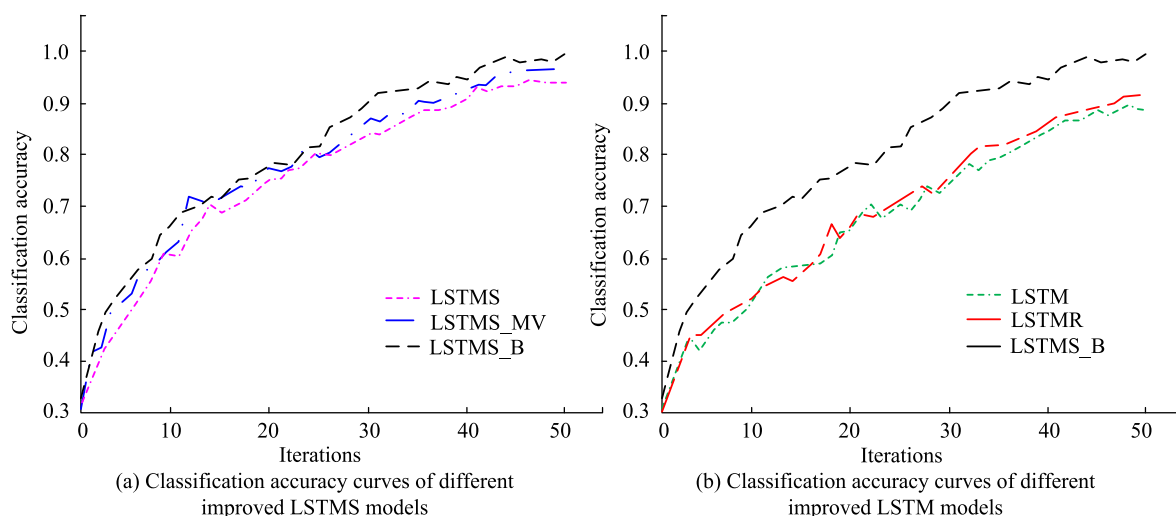


FIGURE 9. Changes in classification accuracy of the training process model.

after 50 iterations. After 50 iterations, the original LSTM had the lowest classification accuracy, with a stable accuracy of around 85%. The classification accuracy of LSTMS using the Swish activation function was about 92%, which was significantly higher than the original LSTM and the LSTMR model using ReLU activation function. This indicated that Swish activation function played an important role in LSTM training and solved the gradient dispersion problem to some certain. Meanwhile, the accuracy of LSTMS was lower than that of LSTMS-B and LSTMS-MV, indicating that the combination of basic learning and multi-classifier decision-making could improve classification performance. The accuracy of LSTMS-B was higher than that of LSTMS-MV model, further indicating that studying improved Bagging could effectively improve the model performance. To further demonstrate the advantages of the

research algorithm, the EEG feature extraction algorithm with higher neutral energy in the current research was introduced for comparative experiments in Figure 10.

In Figure 10, the classification accuracy of RNN-based model and SVM-based model was relatively low, with accuracy rates of 81.96% and 84.5%, respectively. Moreover, SVM classifier could only classify the EEG signals in both biological and non-biological categories. Pyramid Match Kernel (PMK) network visual discrimination method containing EEG information had a classification accuracy of 90.8% for three types of EEG signals. BiLSTM+ICA+SVM method used a combination of ICA and SVM to characterize EEG signals, with a classification accuracy of 96.68%. This research model had the highest classification accuracy of 98.38% when it could also classify 20 types of images, indicating that the LSTMS-B algorithm proposed in the study

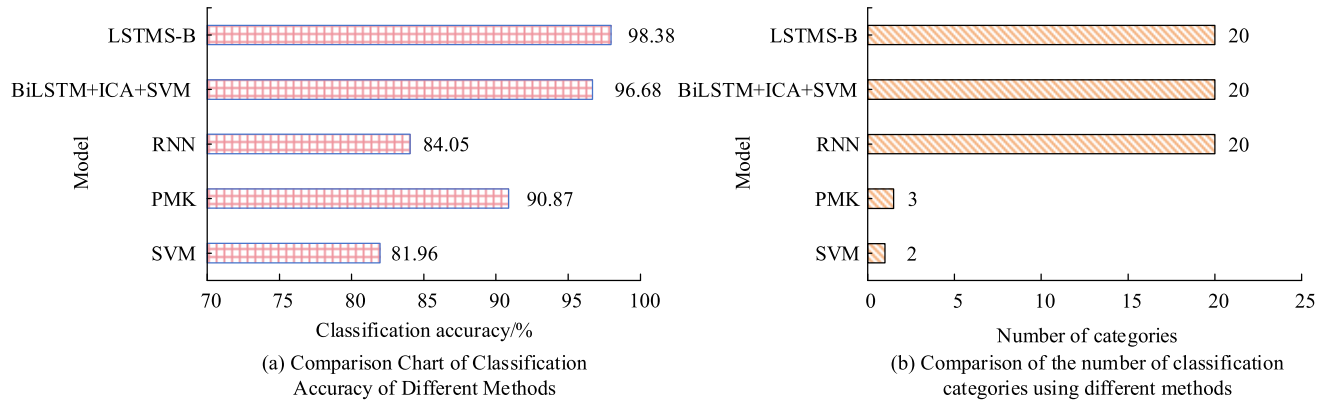


FIGURE 10. Comparison with other literature research results.

TABLE 2. K-fold cross validation evaluation.

K	Variational autoencoder combined recurrent neural network		Algorithm in this article	
	Accuracy (%)	Mean squared error	Accuracy (%)	Mean squared error
5	95.66	0.043	96.88	0.020
6	93.98	0.037	94.95	0.013
7	95.55	0.058	96.77	0.036
8	93.84	0.054	94.81	0.027
9	95.82	0.020	97.04	0.011
10	94.26	0.009	95.23	0.009

had good classification ability for the EEG signals. To further evaluate the robustness of the algorithm model in this article, K-Fold cross validation (K values between 5 and 10) was used. Table 2 shows the results of K-Fold cross-validation evaluation.

From Table 2, the EEG feature classification algorithm proposed by the research had higher accuracy and smaller error when the k-value changed from 5 to 10, with an average accuracy of 95.94% and an average error of 0.0193. However, the average accuracy and error of the variational autoencoder combined with the recurrent neural network were 94.85% and 0.0368, respectively. Therefore, this indicated that the proposed algorithm had more advantages in classification performance and efficiency. Da Costa AZ et al. used a deep residual neural network to construct a classifier for extracting external features from images. After experiments on a self-built food image data set, the optimal classification accuracy of the model was 91.7%. Due to the use of GTX 1080 Ti GPU to improve training speed, its calculation time was reduced to less than 200 seconds [17]. Wang J et al. proposed an improved deep learning neural network based on Hough transform to achieve high detection accuracy and stability in the design of product image detection and analysis tools. The experimental results on bottle image data sets with and without defects showed that the classification accuracy of this model was between 85% and 90%, but the computational

efficiency was not high [18]. Mahajan H B et al. designed an image and geographic information capture model based on a LSTM classifier combined with the Internet of Things. The simulation results of publicly available research data sets showed that the proposed model was efficient and robust. The overall accuracy of the model had been improved by about 5%, and the computational complexity had been reduced by about 84% [19]. The classification accuracy of the EEG signal classification model proposed in this study was relatively high, reaching 97.17%, with high computational efficiency. It tended to converge after 60 iterations (less than 60 seconds). Therefore, this indicated that the proposed model had advantages in classification performance and computational efficiency.

B. GENERATION RESULTS OF EEG CHARACTERISTIC STIMULATION IMAGES BASED ON LSTM-GAN

After extracting EEG visual feature signals, it is necessary for the robot to generate corresponding images of the extracted visual EEG land-scape features, that is, to represent the visual features of the human brain and achieve BCI. Relevant models were introduced for comparative experiments to evaluate the classification performance of the designed EEG visual feature classification model considering AM through classification accuracy in Figure 11.

In Table 3, Bi-LSTM-AttG represents Bi-LSTM with attention gates. Bi-LSTM-AttW represents Bi-LSTM with attention weight. BiLSTM+ICA+SVM utilizes Bi-LSTM as EEG signal feature encoder. The model using ICA and SVM as classifiers is currently the best in EEG signal classification research. In the comparative experiment, Bi-LSTM-AttGW with 2 LSTM layers and 128 hidden layers had the highest classification accuracy, with an accuracy of 98.67%, which was 2% higher than BiLSTM+ICA+SVM. This indicated that the research method had high classification performance and can effectively decode human brain activities related to vision. And before the number of hidden layers reached 128, the classification accuracy also increased as the number of layers increased. When the hidden layer parameter was

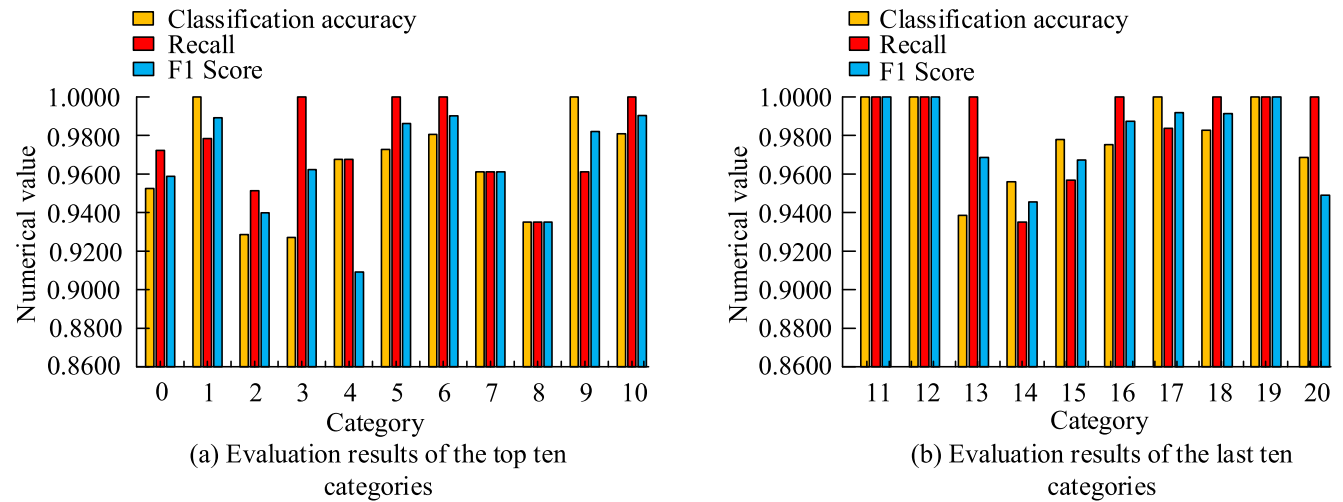


FIGURE 11. Classification results of 20 types of image stimulation EEG signals tested.

TABLE 3. Comparison of classification accuracy of related models.

Model	Classification accuracy/%
LDA	79.69
LSTM	83.93
EEG-Net	89.16
Bi-LSTM	91.68
BiLSTM+ICA+SVM	96.67
Bi-LSTM-AttG(1,18)	94.99
Bi-LSTM-AttW(1,18)	93.68
Bi-LSTM-AttGW(1,64)	96.34
Bi-LSTM-AttGW(1,128)	98.67
Bi-LSTM-AttGW(1,256)	96.16
Bi-LSTM-AttGW(2,128)	87.95

greater than 128 layers, the classification accuracy gradually decreased. Due to the increasing number of layers, the multi-layer NN structure was too complex, resulting in the model not being able to learn EEG features well. The highest performing Bi-LSTM-AttGW (1,128) model was applied to 20 types of image data in the data set for model testing in Figure 11.

In Figure 11, firstly, in the classification test results of EEG signals stimulated by 20 types of images, the classification accuracy of this research model was greater than 92.5%, and the classification accuracy for images of categories 1, 9, 11, 12, 17, and 19 reached 100%. Secondly, the classification recall rate of the model was higher than 93.5%, and the classification recall rate for images in categories 3, 5, 6, 10, 11, 12, 13, 16, 18, 19, and 20 reached 100%. Meanwhile, in the model testing results, the F1 scores of this model were all higher than 91%, and the F1 scores for images of categories 11, 12, and 19 reached 100%. These results confirmed that the visual image magnetic buckle EEG signal classification effect of the research model was good, with high accuracy, and the

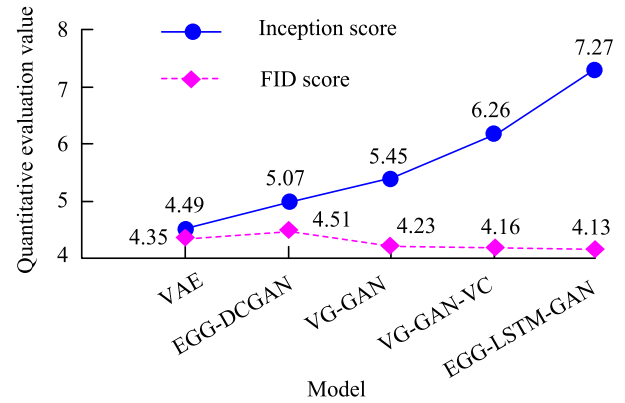


FIGURE 12. Inception and FID scores of images generated by different methods.

model had good classification performance. The Inception score and the Frechet Inception Distance (FID) score were introduced to evaluate the quality of machine stimulus image generation in Figure 12.

In Figure 12, the Inception and FID scores of several current EEG signal decoding and visual stimulus re-construction methods were compared with the research methods. They include stimulus image generation method based on Variational Auto Encoder (VAE), conditional stimulus image generation method based on Deep Convolutional Generative Adversarial Networks (EEG-DCGAN), visual stimulus re-construction method based on visual guided EEG feature representation and Conditional Adversarial Nets (VG-GAN), and visual stimulus re-construction method with visual consistency preserving term (VG-GAN-VC) based on VG-GAN. Among them, the highest Inception score of EGG-based LSTM-GAN was 7.27, and its lowest FID score was 4.13, indicating that the research method performed well in generating objective image quality, and these experimental results had been improved to a certain extent. Figure 13 shows the

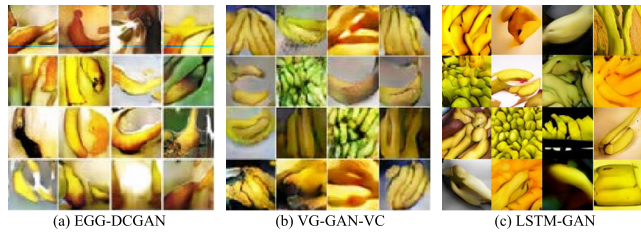


FIGURE 13. Shows the generation results of banana category thorn images using different models.

visual stimulus generation results of the research method, VG-GAN-VC, and EEG-DCGAN on category 0 images.

The image generation effect of the EEG-DCGAN model was the worst, and some of the banana images generated were not realistic. The image generated by VG-GAN-VC had a good effect and was more realistic. The image is a banana, but it is still relatively blurry. The stimulation image generation effect of the research model was the best, generating relatively real banana image information with high image clarity. This indicates that the research model can monitor EEG signals to enable computers to draw images corresponding to things or features that the human brain pays attention to in specific scenes, achieving BCI in the visual land-scape.

V. CONCLUSION

The continuous development of AI has brought machines closer to human thinking. Achieving BCI has become an important research goal for AI. However, the existing BCI still has problems such as low accuracy in classifying EEG signals, poor collaborative representation ability during BCI, and poor resolution in generating images. This study will address the low accuracy in EEG signal classification and propose a visual stimulus EEG signal decoding model based on spatiotemporal features. A method based on LSTM-B regression classification was proposed in this experiment to address the un-reasonable understanding of the machine's brain like land-scape, enabling the machine to better annotate the land-scape according to human brain vision. A generative adversarial network based on LSTM-B was proposed in this experiment to address the issue of poor resolution in generating images from the representation model to improve the quality of generated stimulus images. And relevant algorithms were introduced for comparative experiments to verify the feasibility and advantages of the research model. These experiments confirmed that the proposed EEG signal classification model based on spatiotemporal features had the highest classification accuracy and F1 score compared to other related models, with 91.17% and 96.16%, respectively. The comparison results of EEG visual feature image land-scape cognitive classification models showed that among the models without voting strategy, LSTMS had the highest classification accuracy and recall rate, with 91.39% and 90.89%, respectively. Among the models considering voting strategies, LSTMS-B had the highest classification accuracy and recall rate, with classification

accuracy and recall rates of 98.38% and 97.94%, respectively. In the EEG visual feature classification model for land-scaper-construction of machine EEG visual characteristics, the highest classification accuracy of the research model was 98.67%. The LSTM-GAN stimulation image generation model had the best actual image generation effect, with an Inception score of 7.27 and a FID score of 4.13. These results confirmed that while improving the classification performance of EEG signals, research could also generate stimulus images that matched the corresponding category labels and had higher quality, achieving BCI in the visual land-scape. Research aims to reconstruct the original image by generating images with the same visual category of stimulus. Future research can utilize more effective deep learning methods to directly extract neural representation visual features from EEG data.

REFERENCES

- [1] R. Yu, L. Qiao, M. Chen, S.-W. Lee, X. Fei, and D. Shen, "Weighted graph regularized sparse brain network construction for MCI identification," *Pattern Recognit.*, vol. 90, pp. 220–231, Jun. 2019, doi: [10.1016/j.patcog.2019.01.015](https://doi.org/10.1016/j.patcog.2019.01.015).
- [2] P. Wang, M. Wang, Y. Zhou, Z. Xu, and D. Zhang, "Multiband decomposition and spectral discriminative analysis for motor imagery BCI via deep neural network," *Frontiers Comput. Sci.*, vol. 16, no. 5, pp. 165–178, Jan. 2022, doi: [10.1007/s11704-021-0587-2](https://doi.org/10.1007/s11704-021-0587-2).
- [3] N. Cudlenco, N. Popescu, and M. Leordeanu, "Reading into the mind's eye: Boosting automatic visual recognition with EEG signals," *Neurocomputing*, vol. 386, pp. 281–292, Apr. 2020, doi: [10.1016/j.neucom.2019.12.076](https://doi.org/10.1016/j.neucom.2019.12.076).
- [4] M. T. Sadiq, X. Yu, Z. Yuan, F. Zeming, A. U. Rehman, I. Ullah, G. Li, and G. Xiao, "Motor imagery EEG signals decoding by multi-variate empirical wavelet transform-based framework for robust brain-computer interfaces," *IEEE Access*, vol. 7, pp. 171431–171451, 2019, doi: [10.1109/ACCESS.2019.2956018](https://doi.org/10.1109/ACCESS.2019.2956018).
- [5] J. F. Hwaidi and T. M. Chen, "Classification of motor imagery EEG signals based on deep autoencoder and convolutional neural network approach," *IEEE Access*, vol. 10, pp. 48071–48081, 2022, doi: [10.1109/ACCESS.2022.3171906](https://doi.org/10.1109/ACCESS.2022.3171906).
- [6] 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, doi: [10.1109/TNSRE.2020.3037326](https://doi.org/10.1109/TNSRE.2020.3037326).
- [7] D.-H. Lee, J.-H. Jeong, K. Kim, B.-W. Yu, and S.-W. Lee, "Continuous EEG decoding of pilots' mental states using multiple feature block-based convolutional neural network," *IEEE Access*, vol. 8, pp. 121929–121941, 2020, doi: [10.1109/ACCESS.2020.3006907](https://doi.org/10.1109/ACCESS.2020.3006907).
- [8] X. Geng, D. Li, H. Chen, P. Yu, H. Yan, and M. Yue, "An improved feature extraction algorithms of EEG signals based on motor imagery brain-computer interface," *Alexandria Eng. J.*, vol. 61, no. 6, pp. 4807–4820, Jun. 2022, doi: [10.1016/j.aej.2021.10.034](https://doi.org/10.1016/j.aej.2021.10.034).
- [9] M. K. Ahirwal and M. R. Kose, "Audio-visual stimulation based emotion classification by correlated EEG channels," *Health Technol.*, vol. 10, no. 1, pp. 7–23, Jan. 2020, doi: [10.1007/s12553-019-00394-5](https://doi.org/10.1007/s12553-019-00394-5).
- [10] Z. Gao, X. Sun, M. Liu, W. Dang, C. Ma, and G. Chen, "Attention-based parallel multiscale convolutional neural network for visual evoked potentials EEG classification," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 8, pp. 2887–2894, Aug. 2021, doi: [10.1109/JBHI.2021.3059686](https://doi.org/10.1109/JBHI.2021.3059686).
- [11] P. Nagabushanam, S. T. George, and S. Radha, "EEG signal classification using LSTM and improved neural network algorithms," *Soft Comput.*, vol. 24, no. 13, pp. 9981–10003, Jul. 2020, doi: [10.1007/s00500-019-04515-0](https://doi.org/10.1007/s00500-019-04515-0).
- [12] Y. Zhang, S. Q. Xie, H. Wang, and Z. Zhang, "Data analytics in steady-state visual evoked potential-based brain-computer interface: A review," *IEEE Sensors J.*, vol. 21, no. 2, pp. 1124–1138, Jan. 2021, doi: [10.1109/JSEN.2020.3017491](https://doi.org/10.1109/JSEN.2020.3017491).

- [13] D. Chakraborty, V. Narayanan, and A. Ghosh, "Integration of deep feature extraction and ensemble learning for outlier detection," *Pattern Recognit.*, vol. 89, pp. 161–171, May 2019, doi: [10.1016/j.patcog.2019.01.002](https://doi.org/10.1016/j.patcog.2019.01.002).
- [14] S. N. Amin, P. Shivakumara, T. X. Jun, K. Y. Chong, D. L. L. Zan, and R. Rahavendra, "An augmented reality-based approach for designing interactive food menu of restaurant using Android," *Artif. Intell. Appl.*, vol. 1, no. 1, pp. 26–34, Oct. 2022, doi: [10.47852/bonviewaia2202354](https://doi.org/10.47852/bonviewaia2202354).
- [15] 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, doi: [10.1016/j.patcog.2019.01.006](https://doi.org/10.1016/j.patcog.2019.01.006).
- [16] T. A. Waziri and B. M. Yakasai, "Assessment of some proposed replacement models involving moderate fix-up," *J. Comput. Cognit. Eng.*, vol. 2, no. 1, pp. 28–37, Apr. 2022, doi: [10.47852/bonviewjce2202150](https://doi.org/10.47852/bonviewjce2202150).
- [17] A. Z. da Costa, H. E. H. Figueroa, and J. A. Fracaroli, "Computer vision based detection of external defects on tomatoes using deep learning," *Biosyst. Eng.*, vol. 190, pp. 131–144, Feb. 2020, doi: [10.1016/j.biosystemseng.2019.12.003](https://doi.org/10.1016/j.biosystemseng.2019.12.003).
- [18] J. Wang, P. Fu, and R. X. Gao, "Machine vision intelligence for product defect inspection based on deep learning and Hough transform," *J. Manuf. Syst.*, vol. 51, pp. 52–60, Apr. 2019, doi: [10.1016/j.jmsy.2019.03.002](https://doi.org/10.1016/j.jmsy.2019.03.002).
- [19] H. B. Mahajan, N. Uke, P. Pise, M. Shahade, V. G. Dixit, S. Bhavsar, and S. D. Deshpande, "Automatic robot manoeuvres detection using computer vision and deep learning techniques: A perspective of Internet of Robotics Things (IoRT)," *Multimedia Tools Appl.*, vol. 82, no. 15, pp. 23251–23276, Jun. 2023, doi: [10.1007/s11042-022-14253-5](https://doi.org/10.1007/s11042-022-14253-5).



JUAN TAN was born in Loudi, Hunan, in November 1985. She received the bachelor's degree in industrial design from Tianjin Academy of Fine Arts, in 2008, the master's degree in software engineering from Shandong University, in 2011, and the Ph.D. degree in space design from Huxi University, in August 2022. Since July 2008, she has been teaching environmental design with the College of Fine Arts and Design, Hunan University of Humanities, Science and Technology, an Associate Professor, the Director of Design Discipline, the Director of Environmental Design Discipline, and a Secretary of the Teaching Party Branch of Environmental Design Discipline. Her academic situation projects are "Research on the Cultural Elements and Imagery Expression of Homesickness with Hunan Characteristics in Rural Construction," Hunan Provincial Social Science Fund Project, approved in 2019; Research on Contextual Experiential Teaching Reform of the Course "Indoor Soft Decoration Design," a research project on teaching reform in ordinary higher education institutions of Hunan Provincial Department of Education, approved in 2018; Research on the Development and Application of Home Products in the Art Form of Meishan Folk Nuo Mask," a scientific research project initiated by the Education Department of Hunan Province, in 2018; the Overall Plan for Implementing the Rural Revitalization Strategy in Shuiyuan, Lanrong, Chengbu; the Detailed Planning, Design, and Implementation of the Comprehensive Improvement Project for the living environment in the demonstration area of rural revitalization in Shuiyuan; the Chengbu County People's Government approved the Project, in 2022; the Overall Plan for Implementing the Rural Revitalization Strategy in Luojiashui, Shaoyang, from 2021 to 2030; the Detailed Landscape Plan for the Millennium Ginkgo Leisure Resort in Luojiashui, Hunan Provincial Department of Civil Affairs, approved in 2022; the Design of the Interior Decoration Design and

Surrounding Landscape Improvement and Renovation Project of Zhuzhou Zhongxing Industrial Company Ltd., approved by Zhuzhou Zhongxing Industrial Company Ltd., in 2021; the Overall Plan for the Red Pastoral Complex of Shuangjiang He Guozhong Middle School in Loudi, approved by the People's Government of Shuangjiang Township, in 2019; and the Overall Plan for Rural Revitalization Strategy in Xinshi, Louxing, approved by the Louxing District Rural Revitalization Bureau, in 2022. She has published conferences papers and journal articles, such as "A Convergence Study on the Satisfaction of Visitors by Exhibition Type in Chinese Urban Planning Museum," a well-known Korean academic journal (KCI—published academic journal), Korean Society for the Integration of Science and Art, in June 2022, Core Journal; Discussion on Modular Combination Furniture Design Based on Consumer Behavior, Forestry Industry, Vol. 57, 2020. Volume 6, Core Journal; Doing a Good Job in the Integration of Agriculture and Tourism to Assist in Poverty Alleviation in Ethnic Areas, Theoretical Edition of Hunan Daily, in August 2020, Theoretical Edition Article; and the Study on Landscape Planning and Design of Jielong, Xinhua, against the Background Rural Revitalization, the Third International Conference on Economic Development and Education Management, in 2019. She holds EI patents, such as A Kind of Ecological Garden Landscape Design Green Belt Structure (CN2021114220 04.5), in May 2023, China National Intellectual Property Administration, invention patent; and An Agronomic Plant Protection Device ZL202110219363, in July 2022, China National Intellectual Property Administration Invention Patent. Her research interest includes landscape design.



BAOCHEN WU was born in Suqian, Jiangsu, in May 1982. He received the bachelor's degree in law from the School of Law, Nanjing University, in June 2006, and the master's degree in retail management from Nord Business School, France, in December 2022. From June 2006 to December 2019, he worked at the local government of Jiangsu Province. Since January 2021, he has been with China Forestry Group Zhonglin Times Holdings Company Ltd. His research interest includes ecological landscape.



YUNLONG MA was born in Zhucheng, Shandong, in November 1995. He is currently pursuing the Graduate degree in agriculture with Hunan University of Humanities and Technology. From September 2015 to September 2017, he was enlisted in the military. From 2018 to 2020, he was a Regional Manager with Lunan Pharmaceutical Company Ltd. From September 2020 to May 2022, he was the Office Director of China Railway 12th Bureau Construction and Installation Engineering Company Ltd. From 2022 to 2023, he was a Pharmaceutical Representative with Shenzhen Kangzhe Pharmaceutical Company Ltd. His research interest includes landscape design.

...