

Received June 11, 2020, accepted July 13, 2020, date of publication July 16, 2020, date of current version July 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3009665

# S-EEGNet: Electroencephalogram Signal Classification Based on a Separable Convolution Neural Network With Bilinear Interpolation

WENKAI HUANG<sup>1</sup>, (Member, IEEE), YIHAO XUE<sup>2</sup>, LINGKAI HU<sup>2</sup>, AND HANTANG LIULI<sup>3</sup>

<sup>1</sup>Center for Research on Leading Technology of Special Equipment, School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China

<sup>2</sup>School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China

<sup>3</sup>Chien-Shiung Wu College, Southeast University, Nanjing 211189, China

Corresponding author: Wenkai Huang (16796796@qq.com)

This work was supported by the Guangzhou Science and Technology Planning Project under Grant 202002030279.

**ABSTRACT** As one of the most important research fields in the brain–computer interface (BCI) field, electroencephalogram (EEG) classification has a wide range of application values. However, for the EEG signal, it is difficult for the traditional neural networks to capture the characteristics of the EEG signal more comprehensively from the time and space dimensions, which has a certain effect on the accuracy of EEG classification. To solve this problem, we can improve the accuracy of classification via end-to-end learning of the time and space dimensions of EEG. In this paper, a new type of EEG classification network, the separable EEGNet (S-EEGNet), is proposed based on Hilbert–Huang transform (HHT) and a separable convolutional neural network (CNN) with bilinear interpolation. The EEG signal is transformed into time–frequency representation by HHT, which allows the EEG signal to be better described in the frequency domain. Then, the depthwise and pointwise elements of the network are combined to extract the feature map. The displacement variable is added by the bilinear interpolation method to the convolution layer of the separable CNN, allowing the free deformation of the sampling grid. The deformation depends on the local, dense, and adaptive input characteristics of the EEG data. The network can learn from the time and space dimensions of EEG signals end to end to extract features to improve the accuracy of EEG classification. To show the effectiveness of S-EEGNet, the team used this method to test two different types of EEG public datasets (motor imagery classification and emotion classification). The accuracy of motor imagery classification is 77.9%, and the accuracy of emotion classification is 89.91%, and 88.31%, respectively. The experimental results showed that the classification accuracy of S-EEGNet improved by 3.6%, 1.15%, and 1.33%, respectively.

**INDEX TERMS** Hilbert–Huang transform (HHT), electroencephalogram (EEG) classification, convolutional neural network (CNN), bilinear interpolation.

## I. INTRODUCTION

With the development of human–computer interaction (HCI) [1] and the brain–computer interface (BCI) [2], the application potential of hot technologies in the field of HCI technology has begun to emerge. It is highly important to integrate EEG classification technology into HCI to improve HCI ability. Comparing electrocardiogram (ECG) [3], electromyogram (EMG) [4], and electroencephalogram (EEG) signals, the EEG signal is the most intuitive and objective expression of emotion. EEG is a pattern obtained by amplifying and recording the spontaneous biopotential of the brain

from the scalp with precise electronic instruments. This represents the spontaneous and rhythmic electrical activity of the brain cell group recorded by the electrodes. There are routine, dynamic, and video EEG types. However, due to the characteristics of the dynamic time series data of the EEG signal, each observation value in the EEG sequence is the comprehensive result of various factors affecting the change simultaneously. The real change of EEG signal is the superposition or a combination of several changes, which leads to the correlation and mutual restriction between EEG sequences. The prediction of EEG signals should be studied as a whole to highlight trends and periodic changes. To solve the problem that the traditional neural network unilaterally analyzes a single EEG sequence, which leads to difficulty

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

in improving the classification accuracy and further improve the accuracy of EEG classification, our team fully considers its space and time dimensions when classifying the EEG signals, aiming to capture the characteristics of the EEG signals more comprehensively, and a method of EEG signal classification is proposed based on separate convolution with bilinear interpolation. Compared with the existing methods, this method can classify EEG data more accurately and has stronger robustness.

Before this paper, various preprocessing methods for EEG have been employed to improve the accuracy of EEG classification. For example, Samiee *et al.* [5] used short-time Fourier transform (STFT) to process EEG signals, while Alyasseri *et al.* [6] proposed an EEG denoising method based on wavelet transform (WT), and Yahya *et al.* [7] studied an EEG processing method based on continuous WT (CWT) to improve classification accuracy. In the above literature, STFT is suitable for processing linear nonstationary signals, while WT and CWT can process nonlinear nonstationary signals in theory, but they can only process linear nonstationary signals in the actual algorithm implementation. Compared with the above-mentioned EEG signal processing methods, the Hilbert–Huang transform (HHT) used in this paper is suitable for analyzing nonlinear and nonstationary EEG signals and has complete adaptability. It can transform one-dimensional (1D) EEG signals into two-dimensional (2D) signals on the complex plane, which is more conducive to capturing the dynamic correlation information of EEG signals. Cho and Kang [8] proposed an image denoising method based on CNN, which not only improves the denoising performance, but also uses the separable convolution and the gradient prior in this study to reduce the computational complexity. Compared with the existing CNN denoising methods, this method has better denoising quality and is suitable for a variety of image processing works including EEG.

Neural network technology has been widely used in EEG recognition and classification. Reference [9] proposed a deep learning method based on long short-term memory (LSTM) to recognize emotions from primitive EEG signals. Reference [10] analyzed multichannel EEG using an artificial neural network while [11] proposed a new network-based Alzheimer’s disease (AD) recognition machine learning method that is used in AD-EEG signal recognition, which improves the classification accuracy by changing the network type. Reference [12] proposed a method to train radial basis function neural networks by feature extraction based on the public space pattern of sub-bands selected by sequential back floating. Moreover, [13] proposed a nonlinear feature extraction method based on deep multiset canonical correlation analysis to train neural networks. The core of this method is improving the feature extraction method to enhance the classification accuracy of the neural network. However, [14] used the STFT algorithm to transform the EEG signal of a moving image into a two-dimensional image, then used CapsNet to learn all kinds of characteristics of EEG signals; this solution improved the EEG preprocessing method to enhance the

detection effect. Reference [15] sought to improve the neural network and integrate the weight splitting technology into the algorithm of the Back-Propagation (BP) neural network for EEG recognition and analysis, with the aim of enhancing the classification accuracy. In reference [16], the concept of general model set is introduced, and the model is trained by weighted linear discriminative analysis, which greatly shortens the training time and provides a valuable new strategy for improving the performance of BCI based on P300. The core of these methods is to extract the features of EEG signals effectively to improve the classification accuracy. However, it is still difficult to select the best features from a large number of time-domain and frequency-domain analysis.

In recent years, researchers have gradually found that convolutional neural networks (CNNs) have high efficiency for EEG signal feature extraction, with the characteristics of rapidity and time savings. Therefore, the CNN has become the main research direction in the field of EEG signal recognition and classification. Reference [17] proposed a deep CNN-based learning and classification method for EEG emotional features. Moreover, [18] proposed a long-term and short-term storage network combined with spatial CNN to simultaneously learn spatial information and temporal correlation from the original motor imagery EEG (MI-EEG) signals. Reference [19] proposes a model based on deep convolution network and autoencoders, called AE-CDNN. All the above studies have been devoted to improving the network complexity to enhance classification accuracy. Waytowich *et al.* [20] proposed a compressed CNN (compact-CNN) that only needs the original EEG signal for automatic feature extraction. The experiment showed that compact-CNN is superior to the currently used method of canonical correlation analysis (CCA) and combined-CCA. Wu *et al.* [21] proposed a multiscale filter bank CNN (MSF-BCNN) for MI-EEG classification. A network initialization and fine-tuning strategy was also proposed to train an individual model for topic classification on small datasets. The team compared the MSFBCNN with the latest methods on the open dataset, and the results showed that the accuracy of this method in the topic classification was higher than the baseline. Kwon *et al.* [22] studied the super-resolution (SR) technology of deep CNNs, simulated the EEG data of white Gaussian noise and real brain noise, and obtained experimental EEG data in auditory evoked potential tasks. The SR-EEG simulation data with Gauss white noise or brain noise showed lower mean square error and higher correlation of sensor information, and they could detect the signal source more clearly than was possible in low-resolution (LR) EEG. Nejedly *et al.* [23] developed a machine learning method using CNN to detect the artifacts of intracranial EEG (iEEG) signals under clinical control conditions, and the performance of this method is compared with that of expert notes. The results show that this method can be used as a general model of iEEG. Lawhern *et al.* [24] introduces a compact convolutional neural network (EEGNet) for EEG-based BCIs. Deep convolution and separable convolution are

used to construct EEG specific model, and through four kinds of BCI paradigms, the similarities and differences between EEGNet and the latest research methods in the classification of subjects and cross subjects are compared. The results show that EEGNet is robust enough. On the basis of EEGNet, this study adds the preprocessing steps of EEG data, and improves the convolution layer by using bilinear interpolation method, and gets better results in EEG classification. The final purpose of these studies is improving the accuracy of EEG recognition or classification by deepening or improving the network model and preprocessing the EEG signal. Deep learning models, especially CNNs, have been successfully used in EEG recognition and classification tasks. In recent years, the CNN has made great progress in the field of EEG, and many deep learning models have shown good performance in the detection and classification of various types of EEG signals [25]–[31]. Nowadays, there are many gaps in the application of deep learning technology in multiple aspects of EEG classification, and various deep learning models have great room for improvement. Researchers are still committed to improving the accuracy of EEG recognition and classification.

The above literature shows that, although many studies have developed EEG by various heuristic methods, scholars have also carried out extensive research on the classification of EEG. At the same time, there are still many challenges in the process of EEG classification. Compared with existing research, which has involved increasing the number of network layers, complicating the network, or designing various network structures to improve the accuracy of classification, we can consider improving the network structure and the convolution layer simultaneously to improve the classification accuracy of the model for EEG signals. Some researchers proposed using HHT + CNN method to automatically and accurately identify the type and severity of rolling bearing fault [32], using HHT to convert the time series of vibration signals into time-frequency images, and then CNN learned the fault sensitive features in time-frequency domain from these images and carried out fault classification. Compared with the method in [32], the team improved the EEG data preprocessing, network architecture and convolution layer, and proposed S-EEGNet for EEG classification. First, HHT was used to preprocess the EEG, and then, based on the use of depthwise separable convolution [33], bilinear interpolation was employed to add displacement variables to the convolution layer so the sampling of the input feature map of the convolution check was shifted and focused on the target area of interest. At the same time as improving the accuracy, the complexity of the S-EEGNet model needed to be avoided as much as possible. The method developed here has the following advantages:

- (1) The construction of the model is based on HHT and a separable CNN with bilinear interpolation. HHT is an adaptive signal processing method suitable for analyzing nonlinear and nonstationary signals, and it is highly suitable for processing EEG signals. The separable

convolution has strong feature extraction ability, combined with the displacement variable added based on the bilinear interpolation method. It can effectively improve the accuracy of EEG classification;

- (2) This study uses two different datasets (emotion classification and motion representation) to test the S-EEG, and this achieves satisfactory results. Compared with the existing models, S-EEG has a wider application range, and it can be applied to MI-EEG classification and EEG emotional classification; it also has a stronger application value; and
- (3) This study improves the separable CNN, and the method of adding a displacement variable based on bilinear interpolation is applied to EEG classification, while HHT is used to preprocess the EEG data. The network structure is simple and can achieve higher classification accuracy, which represents a certain level of innovation.

## II. TECHNICAL DETAILS

The team uses an unsupervised learning method to classify the EEG, allowing it to automatically obtain the characteristics that can better describe the identified and classified EEG signals. The S-EEGNet proposed in this paper involves not only the spatial dimension but also the depth dimension (number of channels). Many features can be extracted from the high-dimensional EEG data by dimension reduction algorithm, which is suitable for EEG classification. In this section, we introduce HHT and depthwise separable convolution, add a displacement variable via the bilinear interpolation method, describe the architecture of the S-EEGNet model proposed by our team, and provide a detailed formula of the S-EEGNet model. Figure 1 shows the flow chart of using S-EEGNet to classify EEG in this paper, which is divided into five parts. First, the original EEG data are preprocessed by HHT; then, the S-EEGNet model is trained by the preprocessed EEG signal in the second, third, and fourth parts, and the EEG signal-related features are extracted from the time domain, frequency domain, spatial dimension, and depth dimension. Finally, the test data are input into the trained S-EEGNet model for classification in the last part. After this stage, the team uses the dataset to test a variety of new models to verify the effectiveness of S-EEGNet in EEG classification in this study.

### A. HHT PREPROCESSING

HHT [34] is an adaptive signal processing method that is suitable for processing nonlinear and nonstationary signals. It mainly consists of two parts: The first part is empirical mode decomposition (EMD); the second is Hilbert transform (HT). In the first part, EMD adaptively decomposes any complex signal into a series of intrinsic mode functions (IMFs) according to the signal characteristics. This satisfies the two following conditions: (1) the average value of the mean value tends to 0; and (2) the difference between the number of extreme points of the original signal (including the number

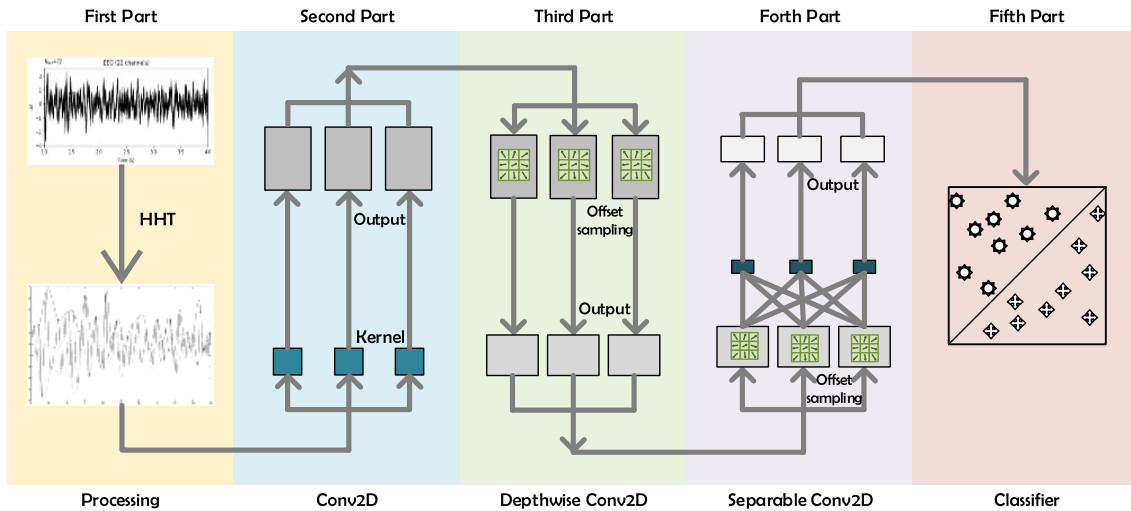


FIGURE 1. Flow chart of electroencephalogram (EEG) classification using S-EEGNet in this study.

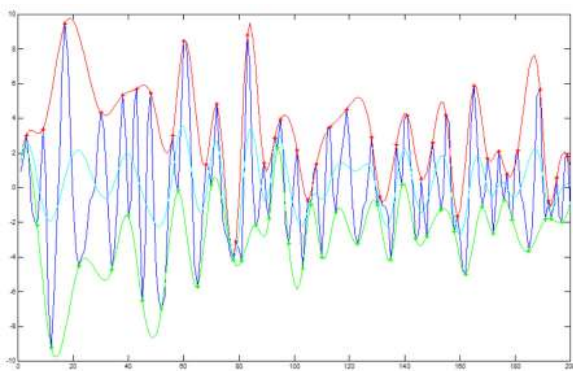


FIGURE 2. Diagram of 200 data points in DEAP dataset S01.

of maximum points + the number of minimum points) and the number of intersections of the original signal cannot be greater than 1 (less than or equal to 1). For the original signal  $x(t)$ , EMD can be used to decompose it into

$$x(t) = \sum_{i=1}^K IMF_{(i)}(t) + r_K(t). \quad (1)$$

where  $x(t)$  is the original signal, and  $IMF_{(i)}$  is  $K$  intrinsic mode functions  $r_K$  is the negligible residue of the signal, which is the remainder of the subtraction of the original signal and  $IMF_{(i)}$ .

HHT can transform the EEG signal into “linear steady state”. In order to better describe the EMD Algorithm in HHT, our team intercepted the first 200 data points of S01 in DEAP dataset [38], as shown in Figure 2.

As shown in Figure 2, the decomposition process of EMD is divided into four steps, the original signal is still  $x(t)$ , and the decomposition steps are as follows:

Step 1: the red dot on the red line is the maximum value of the EEG signal, and the green dot on the green line is the minimum value of the EEG signal. The team uses cubic spline to connect the upper and lower envelope lines, which are red envelope line and green envelope line. Then the mean line is made for the two lines, and the mean value of the envelope

is  $m(t)$ .  $h(t)$  can be obtained by subtracting the mean value from the input signal:

$$h(t) = x(t) - m(t) \quad (2)$$

Step 2: judge whether the  $h(t)$  obtained in step 1 meets the two conditions of the IMF. If not, take  $h(t)$  as the input signal and go back to step 1. If the conditions are met, get an IMF and go to step 3.

Step 3: set the  $k$ th IMF as  $h_k(t)$ , assign it to  $c_k(t)$ , obtained as follows:

$$c_k(t) = h_k(t) \quad (3)$$

$c_k(t)$  is separated from the original sequence and a new residual term is obtained:

$$r_k(t) = x_k(t) - c_k(t) \quad (4)$$

Step 4: judge whether the new remaining item meets the end condition of EMD, if not, bring the remaining item back to step 1; if it meets, end the EMD.

After decomposition, the original signal can be expressed in the form of ( $n$ IMF + 1 residual item):

$$x(t) = \sum_{i=1}^n c_i + r_n \quad (5)$$

The decomposed EEG signal is shown in Figure 3:

In the second part, HT is used to calculate the instantaneous frequency of each IMF. Each natural mode function  $c_i(t)$  can be expressed as:

$$c_i(t) = a_i(t) \cos\phi_i(t) \quad (6)$$

Then put the sampling point time, instantaneous frequency and instantaneous amplitude in three-dimensional space to obtain Hilbert spectrum  $H(\omega, t)$ , as shown in equation (7):

$$H(\omega, t) = Re \left\{ \sum_{i=1}^n a_i(t) e^{i \int \omega_i(t) dt} \right\} \\ = \sum_{i=1}^n a_i(t) \cos \left( \int \omega_i(t) dt \right) \quad (7)$$



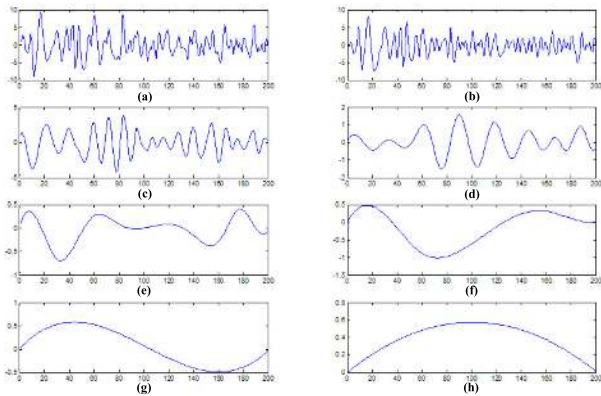


FIGURE 3. Schematic diagram of processed EEG signal, (a) is the original signal of EEG, and then (b) ~ (H) is IMF.

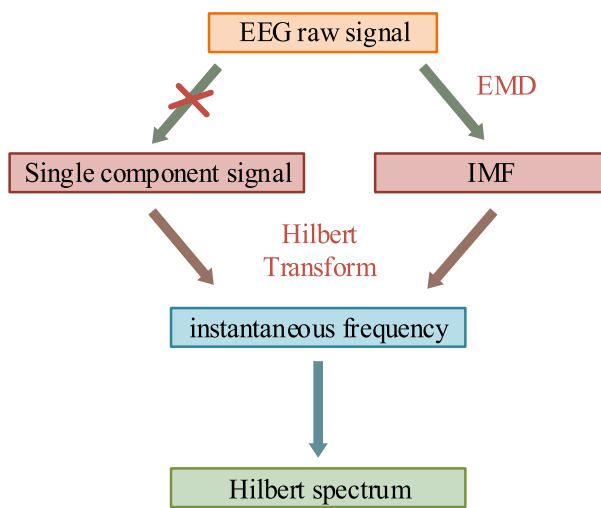


FIGURE 4. Flow chart of HHT processing of the EEG signal.

The team used HHT to transform a 1D EEG signal into a 2D time-frequency representation to characterize the time-frequency distribution of the instantaneous amplitude of the EEG signal. To better explain the HHT applied in this study, Figure 4 shows the flow chart of EEG signal processing with HHT.

### B. S-EEGNET MODEL BUILDING

In this paper, HHT is used to preprocess EEG signals, and a network model based on separable CNN with bilinear interpolation is established to classify EEG. The S-EEGNet model uses the idea of depthwise separable convolution. The core idea of depthwise separable convolution is to decompose a complete convolution operation into two steps, namely, depthwise and pointwise convolution. Depthwise convolution is completely carried out in a two-dimensional plane, and the number of filters is the same as the depth of the previous layer. The number of feature maps after depthwise convolution is the same as the depth of input layer, but this operation ends after each channel of input layer is convoluted independently, and there is no effective use of the information of different feature maps in the same space. Therefore, we need to add another step to combine these feature maps to generate a new

feature map, this requires the following pointwise convolution. The operation of pointwise convolution is very similar to that of conventional convolution. The difference is that the size of convolution kernel is  $1 \times 1 \times M$ , and  $M$  is the depth of the previous layer. Therefore, the convolution operation here will combine the previous map with weighting in the depth direction to generate a new feature map. The number of filters and feature maps is equal. The advantage of this is that it can greatly reduce the number of parameters and calculation while maintaining accuracy. Depthwise separable convolution is a lightweight, low latency network model, but its accuracy is not lower than that of a typical CNN [35]. The other innovation of this study is to improve the convolution layer of depthwise separable convolutional networks and add an offset by the bilinear interpolation method so that the sample of the convolution check input feature map is offset and focused on the target area of interest. After migration, each square corresponding to a convolution kernel can stretch and deform, changing the range of the receptive field and making it polygonal. The operation of including additional offset based on bilinear interpolation is divided into the two following steps:

- (1) Sampling from the input feature map  $x$  using the regular grid  $R$ ; and
  - (2) By the sum of  $w$ -weighted sampling values, defining the size and expansion of the receptive field via grid  $R$ .
- In the depthwise separable convolution, each positioning  $p$ , For each point  $p_0$  in the output feature map  $y$ , the calculation process is as shown in equation (8):

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n). \quad (8)$$

where  $p_n$  is the offset of each point on the convolution output relative to each point on the receptive field, which needs to be an integer. Grid  $R$  adds an enhanced offset, as shown in equation (9):

$$\{\Delta p_n | n = 1, \dots, N\}, \quad \text{where } N = |R|. \quad (9)$$

Then, in the action region of the convolution operation, a learnable parameter  $\Delta p_n$  is added. The output at this time is shown in equation (10):

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n). \quad (10)$$

This operation needs to derive the discontinuous position variables, so bilinear interpolation is used to convert the output of any position to the interpolation operation of the feature map. The specific operation is shown in equation (11):

$$x(p) = \sum_q G(q, p) \cdot x(q). \quad (11)$$

where  $x(q)$  represents the value of the point on all integer positions on the feature map. The specific calculation process of bilinear interpolation is shown in equation (12), where the value of  $q$  is an integer:

$$G(q, p) = g(q_x, p_x) \cdot g(q_y, p_y), \quad (12)$$

where  $g(a, b) = \max(0, 1 - |a - b|)$ .

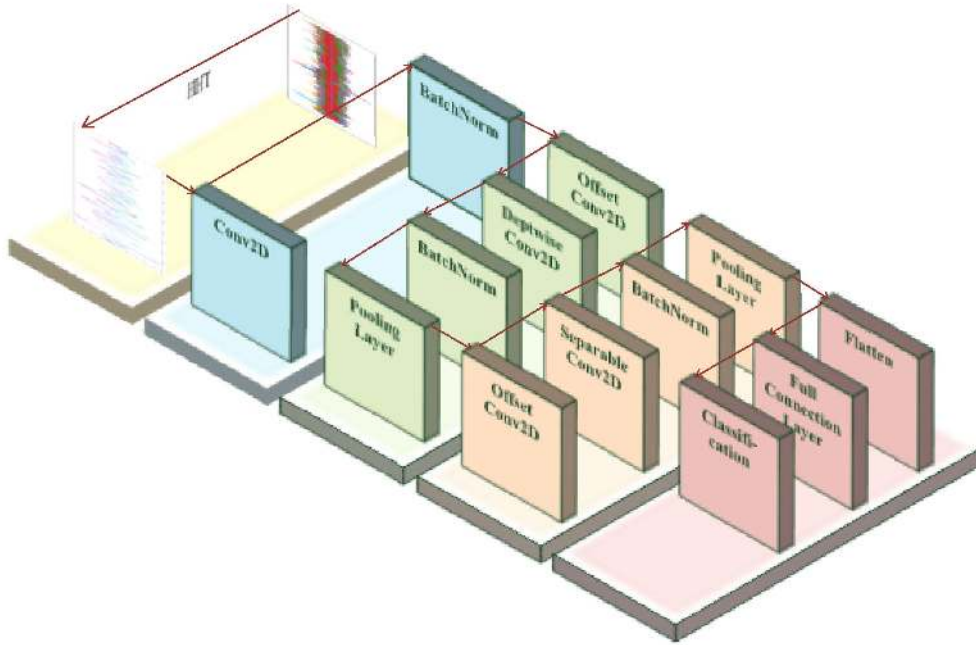


FIGURE 5. Model structure of S-EEGNet.

Figure 5 shows the flow of S-EEGNet in this study. We call the additional offset added in the convolution layer offset-conv2D.

The arrow in the graph represents the convolution kernel connectivity between the input and output, which is called a feature map. First, the original EEG data are preprocessed in the yellow part, and the filter is used to learn from the blue part. The sampling points are offset in the green part, and then the depthwise convolution is carried out to connect to each feature map. The number of feature maps after the depthwise convolution is completed is the same as the number of channels in the input layer, so the feature map cannot be extended. Moreover, the convolution operation is independent for each channel in the input layer, which does not effectively utilize the feature information of different channels in the same spatial position. Therefore, pointwise convolution is needed to combine these feature maps to generate a new feature map. Before this, the sample points are still offset. In the orange part, a separate convolution is employed to learn the time summary of each feature map and then use pointwise convolution to learn how to best combine the feature maps. Finally, the model classifies the EEG signals in the red part. Assuming that, in the feature map  $F$ , the size of the input EEG signal is  $(D_F, D_F, M)$ , the standard convolution  $K$  adopted is  $(D_K, D_K, M, N)$ , and the feature map  $G$ 's output size is  $(D_G, D_G, N)$ . The convolution calculation of the standard convolution is shown in equation (13):

$$G_{k,l,n} = \sum_{i,j,m} K_{i,j,m,n} F_{k+i-1,l+j-1,m}. \quad (13)$$

Assume that the number of input channels is  $M$  and the number of output channels is  $N$ . The corresponding amount of calculations is  $D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F$ . The standard

convolution  $(D_K, D_K, M, N)$  can be divided into depthwise convolution and pointwise convolution as follows:

- (1) The depthwise convolution is responsible for the filtering function; the size is  $(D_K, D_K, 1, M)$ , and the output characteristic is  $(D_G, D_G, M)$ ; and
- (2) The pointwise convolution is responsible for the conversion of channels; the size is  $(1, 1, M, N)$ , and the final output is  $(D_G, D_G, N)$ .

The convolution calculation of depthwise separable convolution is shown in equation (14):

$$\hat{G}_{k,l,n} = \sum_{i,j} \hat{K}_{i,j,m} F_{k+i-1,l+j-1,m}, \quad (14)$$

where  $\hat{K}$  is the depthwise separable convolution and the convolution kernel is  $(D_K, D_K, 1, M)$ . Here, the  $m_{th}$  convolution kernel is applied to the  $m_{th}$  channel in  $F$  to generate the output of the  $m_{th}$  channel on  $\hat{G}$ .

The number of matrix operations to be performed by depthwise separable solutions is  $D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + M \cdot N \cdot D_F \cdot D_F$ , and the convolution kernel parameter is  $D_K \cdot D_K \cdot M + N \cdot M$ . Compared with the ordinary convolution, the calculation amount is reduced:

$$\frac{D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + M \cdot N \cdot D_F \cdot D_F}{D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F} = \frac{1}{N} + \frac{1}{D_K^2}. \quad (15)$$

To describe the changes of network parameters more clearly in each layer of the S-EEGNet model, this paper gives a detailed description in Table 1.

The EEG classification network S-EEGNet proposed by our team is based on separable deformation convolution. The model has two main stages, which are as follows: (1) the depthwise convolution stage, which includes a deformable

**TABLE 1.** S-EEGNet architecture, where  $C$  = number of channels,  $T$  = number of timepoints,  $F_1$  = number of temporal filters,  $D$  = number of spatial filters,  $F_2$  = number of pointwise filters,  $N$  = number of classes.

Layer	Size	Output	Activation
Input		$(C, T)$	
Reshape		$(1, C, T)$	
Conv2D	$(5, 32)$	$(F_1, C, T)$	Linear
BatchNorm		$(F_1, C, T)$	
OffsetConv2D	$(4)$	$(F_1, C, T)$	
DepthwiseConv2D	$(C, 1)$	$(D * F_1, 1, T)$	Linear
BatchNorm		$(D * F_1, 1, T)$	
Activation		$(D * F_1, 1, T)$	Relu
AveragePool2D	$(2, 2)$	$(D * F_1, 1, T/2)$	
Dropout		$(D * F_1, 1, T/2)$	
OffsetConv2D	$(4)$	$(D * F_1, 1, T/2)$	
SeparableConv2D	$(5, 16)$	$(F_2, 1, T/2)$	Linear
BatchNorm		$(F_2, 1, T/2)$	
Activation		$(F_2, 1, T/2)$	Relu
AveragePool2D	$(2, 2)$	$(F_2, 1, T/8)$	
Dropout		$(F_2, 1, T/8)$	
Flatten		$(F_2 * T/8)$	
Dense		$N$	Softmax

layer, depthwise layer, BatchNorm layer, and average pooling layer; and (2) the separable convolution stage, which includes a deformable layer, separable layer, BatchNorm layer, and average pooling layer.

In S-EEGNet, with an increase in the depth of the network, the activation amplitude increases exponentially, and the sensitivity of the divergence information to the input of the non-normalized network decreases. This limits the possible expansion and further limits the learning rate of the control gradient decline speed in the non-normalized network. Therefore, the team used batch normalization [36] technology to normalize the input to accelerate the training speed. We normalize each dimension as equation (16):

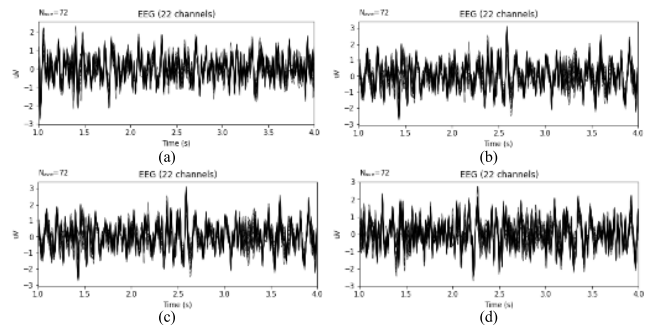
$$BN(X_i) = \frac{(X_i - E(X_i))}{\sqrt{Var(X_i)}}, \quad (16)$$

where  $E(X_i)$  refers to the average value of each batch of training data neurons  $X_i$  and the denominator refers to a standard deviation of the activation degree of each batch of data neurons  $X_i$ . To avoid affecting the feature distribution learned by the network in this layer, feature reconstruction is required:

$$E(X_i) = \frac{1}{m} \sum_{i=1}^L X_i, \quad (17)$$

$$Var(X_i) = \frac{1}{m} \sum_{i=1}^L [X_i - E(X_i)]^2. \quad (18)$$

In S-EEGNet, the team uses batch normalization technology to normalize the features learned in the convolution layer, so that the average value is 0 and the standard deviation is 1. The network is optimized using the Adam optimizer [37].



**FIGURE 6.** Examples of EEG signals in BCI competition IV dataset 2a.

### III. EXPERIMENTS AND RESULTS

#### A. EEG DATASET

##### 1) DATASET 1 (BCI COMPETITION IV DATASET 2A)

The first dataset used in this study is BCI competition IV dataset 2a, which consists of EEG data from nine subjects. This is an MI-EEG dataset with 22 scalp electrode positions. There are four different MI tasks, which are related to the movement imagination of the subjects, including the left and right hands, feet, and tongue. Two sessions were recorded on different days for each subject. Each session comprised six runs separated by short breaks. One run consisted of 48 trials (12 for each of the four possible classes), yielding a total of 288 trials per session. Figure 6 is an example of an EEG signal in BCI competition IV dataset 2a, where (a) is imagining left hand movement, (b) is imagining right hand movement, (c) is imagining leg movement, and (d) is imagining tongue movement.

##### 2) DATASET 2 (A DATABASE FOR EMOTION ANALYSIS USING PHYSIOLOGICAL SIGNALS)

The second dataset used in this study is a database for emotion analysis using physiological signals (DEAP) [38], which is a large open-source dataset used to analyze human emotional states. The dataset recorded EEG and peripheral physiological signals of 32 participants (half were male and half female). Each participant watched 40 music videos, each of which lasted about 1 minute. Then, the participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance, and familiarity. For 22 of the 32 participants, frontal face video was also recorded.

The dataset signals included 32 EEG signals (512 Hz), galvanic skin response (GSR) signals, electrooculogram (EOG) signals, Electromyography (EMG) signals, photoplethysmography (PPG) signals, skin temperature, and status signals. In this study, the team used the preprocessed dataset for the experiment. The EEG signal of the first 32 channels in the DEAP dataset was sampled down to 128 Hz, and the band-pass filter was 4.0–45.0 Hz. The duration of the denoised EEG signal in each track was 63 seconds, including 60 seconds of test signal plus 3 seconds of baseline. Figure 7 is an example of the EEG signal visualization of one subject in the DEAP dataset.



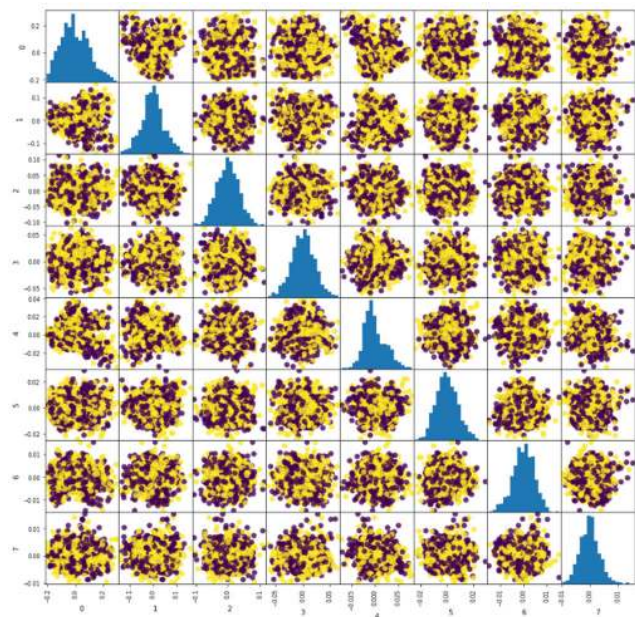


FIGURE 7. Visualization rendering of the DEAP dataset.

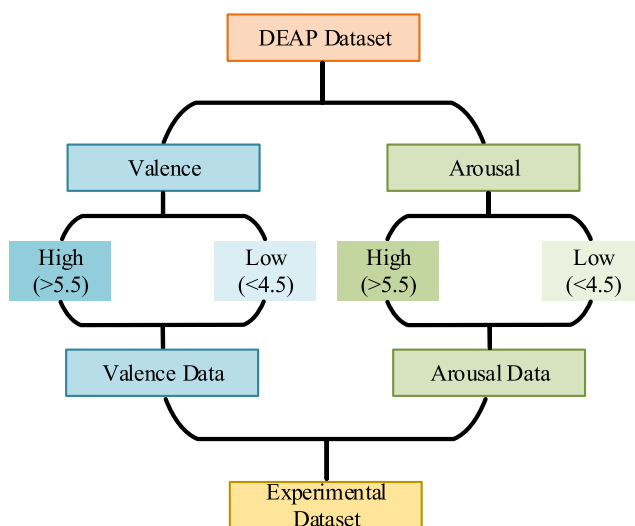


FIGURE 8. Diagram of the experimental dataset selection.

According to the emotional score of each video in arousal and effectiveness (score range: 1–9), the team used a common scoring standard [40] and set a threshold to segment the high and low states of both valence and arousal. A value greater than 5.5 was considered a high valence or arousal, while a value of less than 4.5 was considered a low valence or arousal. Then, S-EEGNet was used to perform an emotion classification task for valence or arousal and test the accuracy. Figure 8 shows the DEAP dataset selection diagram.

**B. EXPERIMENTAL RESULTS**

In the experiment, this paper used the EEG classification network model based on HHT and a separable convolution network with bilinear interpolation. To verify the performance of S-EEGNet in this study for EEG signal classi-

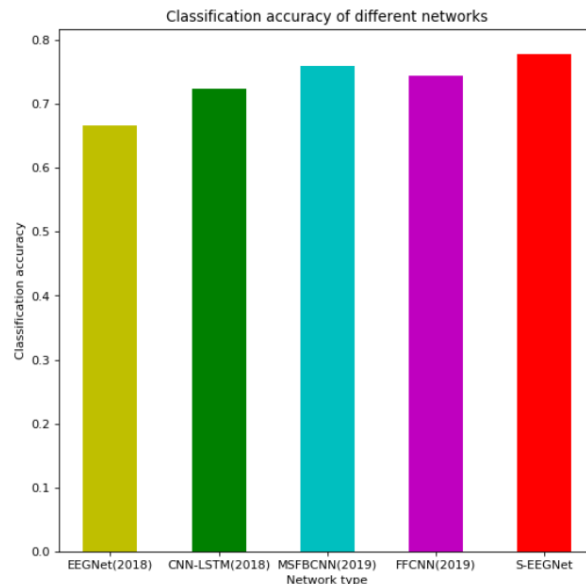


FIGURE 9. Comparison results between S-EEGNet and the four most recent MI-EEG classification networks.

fication, our team used S-EEGNet to test the system with BCI competition IV dataset 2a (motion image classification) and the DEAP dataset (emotion classification), compared with the latest technology. The team used accuracy as the evaluation standard of the EEG classification test to evaluate the performance of S-EEGNet for EEG signal classification. The definition of accuracy is given in equation (19):

$$accuracy = \frac{TP + TN}{TP + FN + FP + TN} \times 100\%, \quad (19)$$

where *TP* represents the EEG samples of the specified period correctly recognized by the model, *FN* represents the EEG samples of the specified period wrongly recognized by the model, *TN* represents the EEG samples of the nonspecified period correctly recognized by the model, and *FP* represents the EEG samples of the nonspecified period wrongly recognized by the model.

1) EXPERIMENT ON THE CLASSIFICATION OF S-EEGNET'S MOTION IMAGE

On the classification of EEG motion imagery, our team compared S-EEGNet with the other most recent MI-EEG classification methods, including CNN-LSTM [18], MSFBCNN [21], CNN based on feature fusion (FFCNN) [39], and EEGNet [24]. Since two different types of datasets were used in this study, the neural networks employed by our team for comparison in the two datasets were different. Figure 9 shows the comparison results between the S-EEGNet in this study and the four most recent MI-EEG classification networks.

In Figure 7, we compare S-EEGNet with the other latest EEG classification networks on BCI competition IV dataset 2a. Lawhern *et al.* [24] proposed a compact-CNN based model of EEG, which can be used in many types of EEG datasets. The accuracy of the model is 69.5% in BCI competi-



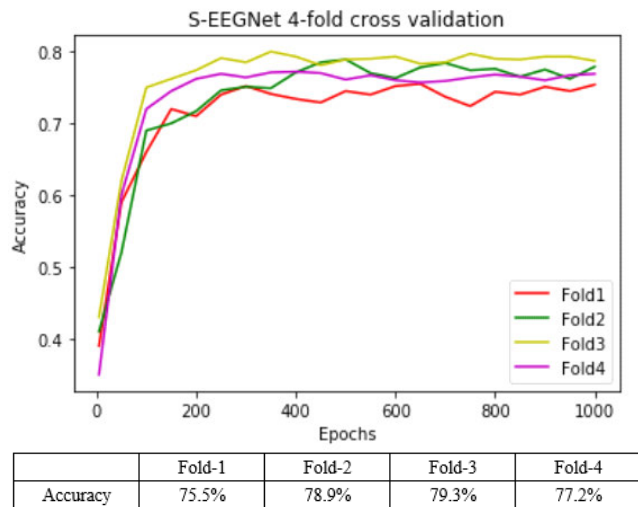


FIGURE 10. Accuracy results of S-EEGNet fourfold cross-validation.

tion IV dataset 2a used in this study. Yang et al. [18] proposed a CNN-LSTM neural network combining space CNN and LSTM to extract features from the original MI-EEG and tested the MI-EEG dataset in the study, reporting an accuracy of 72.4%. Wu et al. [21] proposed a parallel multiscale filter bank CNN for MI-EEG classification and achieved 75.9% classification accuracy. Amin et al. [39] proposed a method of extracting and fusing EEG features from different CNN layers to improve classification accuracy. Using the dataset in this study, 74.5% of the classification accuracy was achieved. The methods proposed by our team have improved the preprocessing and neural network. Compared with the existing methods to extract features from the original EEG data, this study has done HHT preprocessing for the EEG signal and input the processed data based on the separable CNN with bilinear interpolation after training in the neural network model of interpolation. The classification accuracy of the MI-EEG dataset is up to 77.9%. Compared with the latest methods, the proposed S-EEGNet improves the classification accuracy by 2%. To prove the robustness of S-EEGNet in MI-EEG classification, our team conducted fourfold cross-validation on the classification results of S-EEGNet, as shown in Figure 10.

As shown in Figure 10, the S-EEG proposed in this paper was tested in BCI competition IV dataset 2a, and the highest accuracy of MI-EEG classification was 79.5%, which is 3.6% higher than in the latest methods. After fourfold cross-validation, the average accuracy of MI-EEG classification obtained by S-EEG was 77.9%, which is 2% higher than the latest method. It can be seen that the proposed method based on the separable CNN and using the bilinear interpolation method to add displacement variables in the convolution layer can effectively improve the classification accuracy of the MI-EEG dataset.

In this study, the team used HHT to preprocess the EEG signal. To verify that the preprocessing method is superior to other commonly used preprocessing methods, the team

TABLE 2. Using BCI competition dataset 2a to compare various preprocessing methods.

	HHT+ S-CNN	FT+ S-CNN	STFT+ S-CNN	WT+ S-CNN	S-CNN
Acc	79.5%	74.1%	76.9%	75.9%	76.7%

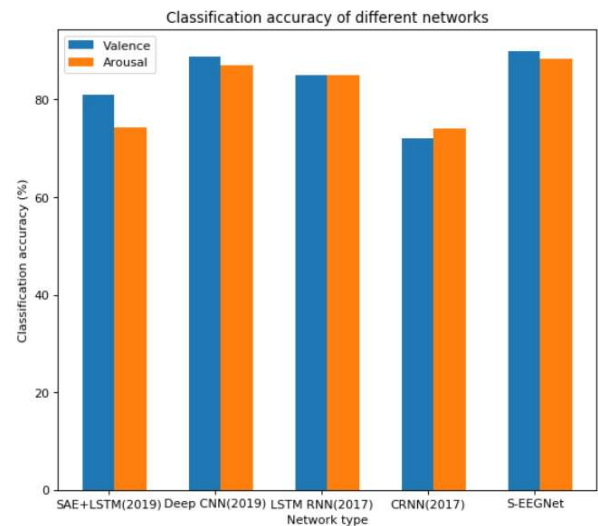


FIGURE 11. Comparison of accuracy of S-EEGNet and various advanced neural networks for EEG emotion classification.

conducted a comparative experiment. The team used HHT and Fourier transform (FT), STFT, and WT to carry out comparative experiments, using the above methods to preprocess the EEG signal, then input neural network for training and compare the accuracy results. Table 2 shows the experimental results of the comparative experiment.

In terms of EEG signal preprocessing, our team compared the HHT preprocessing method used in this study with several common methods. As shown in Table 2, compared with the three other preprocessing methods, HHT was used to preprocess the EEG signals, improving the classification accuracy by 2.6%. Compared with no preprocessing of the original signal, HHT could improve the classification accuracy by 2.8%. It can be seen that, based on not using HHT to preprocess the EEG signal, the original EEG signal was input into the neural network in this study, and the classification accuracy results obtained were 0.8% higher than they were using the latest methods. After preprocessing, the classification accuracy results were further improved, reaching 79.5% and 3.6% higher than the latest methods.

## 2) EMOTION CLASSIFICATION EXPERIMENT ON S-EEGNET

To show the effectiveness of S-EEG in EEG classification, the team also used the DEAP dataset (emotion classification) to test it and still employed fourfold cross-validation. The team compared S-EEGNet with other EEG emotion classification networks with the best effect, including Stack AutoEncoder (SAE)+LSTM [40], deep CNN [17], LSTM RNN [9], convolutional recurrent neural network (CRNN) [41], and compared with these methods, S-EEGNet still achieved satisfactory results in emotion classification tasks.

**TABLE 3.** Comparison with the most advanced models. All of these models employ the deap dataset.

Existing Models	Classifier	Accuracy (%)	
		Valence	Arousal
Xing et al. (2019)	SAE+LSTM	81.10	74.38
Chen et al. (2019)	Deep CNN	88.76	86.98
Alhagry et al. (2017)	LSTM RNN	85.00	85.00
Li et al. (2017)	CRNN	72.06	74.12
Ours	S-EEGNet	<b>89.91</b>	<b>88.31</b>

Figure 11 shows the accuracy comparison of S-EEGNet and other latest models for DEAP dataset classification tasks.

In Figure 11, we compare S-EEGNet with the latest EEG emotion classification network model on the DEAP dataset. For the high or low classification task of valence, S-EEGNet achieved good results, with a classification accuracy of 89.91%, which was 1.15% higher than the latest method. For the high or low arousal classification task, S-EEGNet still achieved good results; the classification accuracy was as high as 88.31%, which was 1.33% higher than the current latest methods. In general, S-EEGNet has some advantages in EEG emotion classification. Table 3 more intuitively introduces the results of S-EEGNet for DEAP dataset classification:

In terms of the selection of preprocessing methods, the team still used the control variable method, HHT, and various commonly employed EEG signal preprocessing methods to preprocess the dataset; it then input the neural network for training. The classification accuracy results are shown in Table 4.

The team still compared the HHT pretreatment method used in this study with several common methods. As shown in Table 4, compared with the other three preprocessing methods, using HHT to preprocess the dataset improved the classification accuracy by 1.89% and 1.05% respectively. This can be compared with no preprocessing, where the original signal was directly input into neural network; the classification accuracy improved by 2.77% and 2.19%, respectively with HHT preprocessing.

#### IV. DISCUSSION

In the EEG classification work, our team established an EEG classification network model S-EEGNet based on HHT

**TABLE 4.** Using deap dataset to compare various preprocessing methods.

	HHT+ S-CNN	FT+ S-CNN	STFT+ S-CNN	WT+ S-CNN	S-CNN
Valence	<b>89.91%</b>	84.2%	87.37%	88.02%	87.14%
Arousal	<b>88.31%</b>	84.69%	85.86%	87.26%	86.12%

and separable CNN with bilinear polarization. First, HHT was used to preprocess the original EEG signal, and then depthwise and pointwise methods were combined to extract the feature map of the EEG signal. On the convolution layer of a separable convolution network, the displacement variable was added by the bilinear interpolation method, allowing the free deformation of the sampling grid. The deformation depends on the local, dense, and adaptive input characteristics of the EEG data. S-EEGNet does not need any additional monitoring signals and can be directly obtained by learning the EEG signal. The displacement variable added by the bilinear interpolation method can easily replace several standard convolution elements in the network and carry out end-to-end training through standard back propagation. Compared with the traditional neural network, end-to-end learning from the spatial and temporal features of EEG signals is carried out to obtain its dynamic correlation, to improve the accuracy of EEG classification. The team tested this in BCI competition IV dataset 2a and the DEAP dataset. In BCI competition IV dataset 2a, the highest accuracy of MI-EEG classification was 79.5%, which was 3.6% higher than the latest methods. In the high or low classification task of value/arousal in the DEAP dataset, the S-EEGNet results were 1.15% and 1.33% higher than the latest research results. Most EEG classification networks in existing research are aimed at a specific field (motor imagery, emotion classification, error-related potential, visual-related potential, rest state, etc.), whereas our team used the motor imagery and emotion classification datasets to separate S-EEGNet. The test results are satisfactory. This study has a certain value for BCI and HCI.

#### V. CONCLUSION

In this paper, a neural network-based EEG signal classification model S-EEGNet was established. The S-EEGNet proposed by our team showed strong performance in various types of EEG classification tasks, effectively improving the accuracy and stability of EEG signal classification and providing a valuable method for more accurate HCI and further realization of computer-integrated intelligence.

Although S-EEGNet obtained good experimental results in this study, it still needs further research, including further research on music emotion classification based on EEG and prediction of epilepsy based on EEG. The team will further study how to build a more powerful and robust EEG classification model. All these points represent the direction of our next research work.

#### CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest regarding the publication of this paper.

## REFERENCES

- [1] H. Singh and J. Singh, "Object acquisition and selection using automatic scanning and eye blinks in an HCI system," *J. Multimodal User Interface*, vol. 13, no. 4, pp. 405–417, Dec. 2019.
- [2] Ç. I. Aci, M. Kaya, and Y. Mishchenko, "Distinguishing mental attention states of humans via an EEG-based passive BCI using machine learning methods," *Expert Syst. Appl.*, vol. 134, pp. 153–166, Nov. 2019.
- [3] C. Meisel and K. A. Bailey, "Identifying signal-dependent information about the preictal state: A comparison across ECoG, EEG and EKG using deep learning," *EBioMedicine*, vol. 45, pp. 422–431, Jul. 2019.
- [4] X. Xi, C. Ma, C. Yuan, S. M. Miran, X. Hua, Y.-B. Zhao, and Z. Luo, "Enhanced EEG–EMG coherence analysis based on hand movements," *Biomed. Signal Process. Control*, vol. 56, Feb. 2020, Art. no. 101727.
- [5] K. Samiee, P. Kovacs, and M. Gabbouj, "Epileptic seizure classification of EEG time-series using rational discrete short-time Fourier transform," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 541–552, Feb. 2015.
- [6] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, A. K. Abasi, and S. N. Makhadmeh, "EEG signals denoising using optimal wavelet transform hybridized with efficient metaheuristic methods," *IEEE Access*, vol. 8, pp. 10584–10605, 2020.
- [7] N. Yahya, H. Musa, Z. Y. Ong, and I. Elamvazuthi, "Classification of motor functions from electroencephalogram (EEG) signals based on an integrated method comprised of common spatial pattern and wavelet transform framework," *Sensors*, vol. 19, no. 22, Nov. 2019, Art. no. 4878.
- [8] S. I. Cho and S.-J. Kang, "Gradient prior-aided CNN denoiser with separable convolution-based optimization of feature dimension," *IEEE Trans. Multimedia*, vol. 21, no. 2, pp. 484–493, Feb. 2019.
- [9] S. Alhagry, A. A. Aly, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 10, pp. 355–358, Oct. 2017.
- [10] A. E. Hramov, V. A. Maksimenko, S. V. Pchelintseva, A. E. Runnova, V. V. Grubov, V. Y. Musatov, M. O. Zhuravlev, A. A. Koronovskii, and A. N. Pisarchik, "Classifying the perceptual interpretations of a bistable image using EEG and artificial neural networks," *Frontiers Neurosci.*, vol. 11, Dec. 2017, Art. no. 674.
- [11] H. Yu, X. Lei, Z. Song, C. Liu, and J. Wang, "Supervised network-based fuzzy learning of EEG signals for Alzheimer's disease identification," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 1, pp. 60–71, Jan. 2020.
- [12] M. H. Bhatti, J. Khan, M. U. G. Khan, R. Iqbal, M. Aloqaily, Y. Jararweh, and B. Gupta, "Soft computing-based EEG classification by optimal feature selection and neural networks," *IEEE Trans. Ind. Informat.*, vol. 15, no. 10, pp. 5747–5754, Oct. 2019.
- [13] Q. Liu, Y. Jiao, Y. Miao, C. Zuo, X. Wang, A. Cichocki, and J. Jin, "Efficient representations of EEG signals for SSVEP frequency recognition based on deep multiset CCA," *Neurocomputing*, vol. 378, pp. 36–44, Feb. 2020.
- [14] K.-W. Ha and J.-W. Jeong, "Motor imagery EEG classification using capsule networks," *Sensors*, vol. 19, no. 13, Jul. 2019, Art. no. 2854.
- [15] L. Liu, "Recognition and analysis of motor imagery EEG signal based on improved BP neural network," *IEEE Access*, vol. 7, pp. 47794–47803, 2019.
- [16] J. Jin, S. Li, I. Daly, Y. Miao, C. Liu, X. Wang, and A. Cichocki, "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.
- [17] J. X. Chen, P. W. Zhang, Z. J. Mao, Y. F. Huang, D. M. Jiang, and Y. N. Zhang, "Accurate EEG-based emotion recognition on combined features using deep convolutional neural networks," *IEEE Access*, vol. 7, pp. 44317–44328, 2019.
- [18] J. Yang, S. Yao, and J. Wang, "Deep fusion feature learning network for MI-EEG classification," *IEEE Access*, vol. 6, pp. 79050–79059, 2018.
- [19] T. Wen and Z. Zhang, "Deep convolution neural network and autoencoders-based unsupervised feature learning of EEG signals," *IEEE Access*, vol. 6, pp. 25399–25410, 2018.
- [20] N. Waytowich, V. J. Lawhern, J. O. Garcia, J. Cummings, J. Faller, P. Sajda, and J. M. Vettel, "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *J. Neural Eng.*, vol. 15, no. 6, Dec. 2018, Art. no. 066031.
- [21] H. Wu, Y. Niu, F. Li, Y. Li, B. Fu, G. Shi, and M. Dong, "A parallel multiscale filter bank convolutional neural networks for motor imagery EEG classification," *Frontiers Neurosci.*, vol. 13, Nov. 2019, Art. no. 1275.
- [22] M. Kwon, S. Han, K. Kim, and S. C. Jun, "Super-resolution for improving EEG spatial resolution using deep convolutional neural network–feasibility study," *Sensors*, vol. 19, no. 23, Dec. 2019, Art. no. 5317.
- [23] P. Nejedly, J. Cimbalnik, P. Klimes, F. Plesinger, J. Halamek, V. Kremen, I. Viscor, B. H. Brinkmann, M. Pail, M. Brazdil, G. Worrell, and P. Jurak, "Intracerebral EEG artifact identification using convolutional neural networks," *Neuroinformatics*, vol. 17, no. 2, pp. 225–234, Apr. 2019.
- [24] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013.
- [25] R. T. Schirmermeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggenberger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017.
- [26] A. R. Ozcan and S. Erturk, "Seizure prediction in scalp EEG using 3D convolutional neural networks with an image-based approach," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 11, pp. 2284–2293, Nov. 2019.
- [27] N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, S. Ippolito, and O. Kavehei, "Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram," *Neural Netw.*, vol. 105, pp. 104–111, Sep. 2018.
- [28] M. Wang, H. El-Fiqi, J. Hu, and H. A. Abbass, "Convolutional neural networks using dynamic functional connectivity for EEG-based person identification in diverse human states," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 12, pp. 3259–3272, Dec. 2019.
- [29] A. Antoniadou, L. Spyrou, D. Martin-Lopez, A. Valentin, G. Alarcon, S. Sanei, and C. C. Took, "Detection of interictal discharges with convolutional neural networks using discrete ordered multichannel intracranial EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 12, pp. 2285–2294, Dec. 2017.
- [30] M. Yu, Y. Sun, B. Zhu, L. Zhu, Y. Lin, X. Tang, Y. Guo, G. Sun, and M. Dong, "Diverse frequency band-based convolutional neural networks for tonic cold pain assessment using EEG," *Neurocomputing*, vol. 378, pp. 270–282, Feb. 2020.
- [31] S. Opalka, B. Stasiak, D. Szajerman, and A. Wojciechowski, "Multi-channel convolutional neural networks architecture feeding for effective EEG mental tasks classification," *Sensors*, vol. 18, no. 10, Oct. 2018, Art. no. 3451.
- [32] Z. Yuan, L. B. Zhang, L. X. Duan, and T. Li, "Intelligent fault diagnosis of rolling element bearings based on HHT and CNN," in *Proc. Prognostics Syst. Health Manage. Conf.*, P. Ding, C. Li, R. V. Sanchez, and S. Yang, Eds., 2018, pp. 292–296.
- [33] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," 2016, *arXiv:1610.02357*. [Online]. Available: <http://arxiv.org/abs/1610.02357>
- [34] M. Dätig and T. Schlurmann, "Performance and limitations of the Hilbert–Huang transformation (HHT) with an application to irregular water waves," *Ocean Eng.*, vol. 31, nos. 14–15, pp. 1783–1834, Oct. 2004.
- [35] M. X. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Nov. 2019, pp. 1–10. [Online]. Available: <http://arxiv.org/abs/1905.11946>
- [36] M. Liu, W. Wu, Z. Gu, Z. Yu, F. Qi, and Y. Li, "Deep learning based on batch normalization for P300 signal detection," *Neurocomputing*, vol. 275, pp. 288–297, Jan. 2018.
- [37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [38] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; Using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan./Mar. 2012.
- [39] S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Bencherif, and M. S. Hossain, "Multilevel weighted feature fusion using convolutional neural networks for EEG motor imagery classification," *IEEE Access*, vol. 7, pp. 18940–18950, 2019.
- [40] X. Xing, Z. Li, T. Xu, L. Shu, B. Hu, and X. Xu, "SAE+LSTM: A new framework for emotion recognition from multi-channel EEG," *Frontiers Neurobot.*, vol. 13, Jun. 2019, Art. no. 37.
- [41] X. Li, D. Song, P. Zhang, G. Yu, Y. Hou, and B. Hu, "Emotion recognition from multi-channel EEG data through convolutional recurrent neural network," in *Proc. IEEE Int. Conf. Bioinf. Biomed.*, Kansas City, MI, USA, Nov. 2017, pp. 352–359.

...