Multi Class Emotion Classification using PSD and CNN

Ramprasad Kumawat, Virendra Jain, Tapan Das Bairagi

Department of Electrical and Electronics Engineering (Mandsaur University) Mandsaur, India

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150800087

Received: 30 August 2026; Accepted: 04 September 2026; Published: 16 September 2026

ABSTRACT

Emotion identification based on electroencephalography (EEG) data has become a popular area of study in human-machine interactions. Conventional machine learning techniques employ carefully created classifiers using hand-made features that may be restricted to domain expertise. We suggested a convolutional neural network (CNN) model to automatically extract the spatiotemporal information from power spectral density (PSD) features extracted from EEG signals, motivated by the exceptional performance of deep learning techniques in recognition tests. The proposed model achieves a high accuracy rate of 90% using the preprocessing method with baseline signals, for the four-class classification task in the SEED_IV dataset. We have achieved a standard deviation of 4.3% in multiple trials

Keywords: CNN, EEG, emotion recognition, PSD, SEED_IV

INTRODUCTION

Physiological signals and non-physiological signals are the two groups into which current emotion detection techniques fall. Physiological signals accurately reflect human emotional states because they are not influenced by subjective influences, in contrast to non-physiological signals. This leads to a major improvement in the utility and reliability of emotion identification based on physiological data. Physiological signalling is receiving more and more attention from scholars. Emotions are identified using electrocardiogram (ECG), facial expression, eye movement (EM), EEG, and other physiological data [1]. These physiological indicators allow us to accurately assess the people's actual feelings. The EEG is an unmasked, objective, spontaneous physiological signal [2]. EEG-based emotion identification has therefore become more significant. Niemic [3] shown a strong relationship between the EEG activity and various emotional states.

The two most popular methods for evoking emotion are subject and external event evocation. The method by which people bring up certain emotional memories or experiences in order to arouse comparable emotions is known as subject evocation. The technique of employing external stimuli, such as music, films, and pictures, to evoke the appropriate feelings in participants is known as external event evocation. A 10-20 system is used to measure the EEG signal [5]. These broadcasts are divided into five frequency bands. Based on different conscious behaviours, different band signals are important [4].

Feature extraction is a crucial step in developing BCI systems. A popular technique for extracting spatial information is known as power spectral density (PSD) [5], [6], which efficiently creates an optimal solution to distinguish between different emotion classes. Typically, a bandpass filter is applied before PSD to focus on a specific frequency band, which significantly influences PSD's performance.

Recent advances in deep learning algorithms have accelerated the development of brain-computer interface (BCI) systems based on EEG. Convolutional neural networks (CNNs) have gained popularity in EEG-based emotion classification because of their capacity to extract both temporal and spatial characteristics from EEG data. Schirrmeister et al. [7] proposed two CNN topologies, Shallow and Deep, for emotion-based EEG recognition. Notably, the suggested models excel at utilising spectral power features across several frequency bands, as indicated by CNN visualisation findings.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

In section II we discuss previous work done by various researchers. Section III describes material and method

used in the paper. Section IV elaborates experimental results and compare results with other methods. In section V we conclude in brief.

RELATED WORK

It is necessary to identify significant elements from the EEG data in order to obtain appropriate classification results. Features are often retrieved in the frequency, time, and time-frequency domains. Zheng et al. [8] used brain asymmetry, differential entropy (DE), and power spectral density (PSD) to investigate important frequency bands and channels. They have used deep belief networks for three emotions, achieving an accuracy of 61% with PSD and 86% with DE. Zhong et al. [9] described the inter-channel interaction in the EEG data using an adjacency matrix in a graph neural network. A robust model was constructed using emotion-aware distribution learning and node-wise domain adversarial training. Song et al. introduced a novel emotion categorisation system based on dynamical graph convolutional neural networks (DGCNNs) [10]. They had characterised multichannel EEG characteristics using a graph. In order to get more distinct characteristics, the DGCNN algorithm was utilised to train neural networks and dynamically learn the inherent relationships between EEG channels. In order to train a hierarchical convolution neural network (HCNN) and classify three emotions from the SEED dataset, Li et al. [11] computed differential entropy at certain time intervals and organised those features in two-dimensional feature maps. The revelation that emotional reactions produced by two hemispheres are asymmetric led to the development of the bi-hemisphere domain adversarial neural network (BiDANN) [12]. Two local domain discriminators and a global domain discriminator were included in the model. In order to extract distinct emotional characteristics from each classifier, these two operated antagonistically with the classifier. As a result, it has tried to create a more universal model. In order to extract distinct emotional characteristics from each classifier, these two operated antagonistically with the classifier. As a result, it has tried to create a more universal model. Since differential entropy components of the EEG signal are best suited for emotion identification, Bhandari and Jain [13] took them into consideration. To determine the positional link between electrodes, they arranged data from 62 channels in a two-dimensional map. In order to determine electrode correlation in emotion detection, they have provided this data as an input to the CNN model and achieved an accuracy of 72.83%.

MATERIALS AND METHOD

In this section we discuss the dataset used, features extracted and the CNN model used in the experiment.

SEED_IV dataset

15 participants (7 men and 8 women) participated in the SEED-IV dataset research. Throughout the trial, each participant's 62-channel EEG data was recorded as they saw movie clips with different emotion classifications. Three separate sessions' worth of video were presented to each participant. There are 1,080 samples in the SEED-IV dataset [14] (72 samples x 15 individuals). Each of the four primary emotional categories—happy, sad, afraid, and neutral—has a total of 18 potential classes. As a result, the number of samples per topic or class is balanced.

The signals were recorded concurrently at a sampling rate of 1,000 Hz. Bandpass frequency filters with cut-off frequencies of 1-75 Hz were applied to the SEED-IV dataset in order to eliminate the unnecessary artefacts. The raw EEG signals are down-sampled to data with a 200 Hz sampling rate in order to prepare the data for analysis. We manually exclude data that is significantly tainted by EMG and EOG after visually inspecting the EEG data. The EEG data are processed using a bandpass filter between 0.5 Hz and 70 Hz to remove noise and artifacts. For feature extraction, a sample size of one second per channel is selected. Five frequency bands—delta (1-3 Hz), theta (4-7 Hz), alpha (8-13 Hz), beta (14-30 Hz), and gamma (31-50 Hz)—are used to calculate the characteristics.

Feature Extraction

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

For a given time sequence, the power spectrum of each band is obtained using the Welch approach [15].

Assume that the series is y d(n), where the interval length is B and the signal intervals are d=1,2,3,...,n. Consequently, power spectral density is described using the Welch technique as,

$$\hat{P}d(f) = \frac{1}{BV}\left|\sum_{n=0}^{B-1} yd(n)w(n)e^{-j2\pi f}\right|^2 \quad (1)$$

where V stands for normalization factor for power in window function and f is frequency band.

$$v = \frac{1}{B}\sum_{n=0}^{B-1}|w(n)|^2 \quad (2)$$

where w(n) is windowed data. This PSD is based on periodograms obtained by converting the signal from the time to the frequency domain.

$$\hat{p}\text{Welch (f)} = \frac{1}{l}\sum_{n=0}^{l-1}\hat{p}\ d(f) \quad (3)$$

We also use the linear discriminant system for feature smoothing [11]. Calculating the average of a predetermined window of successive data points from the signal is known as moving average smoothing. A smoother version of the original signal is produced by recalculating the average as the window advances along the data.

Let x[t] be the raw EEG signal, and y[t] be the smoothed signal. A simple moving average is given by:

$$\text{y[t]} = \frac{\text{i}}{N}\sum_{i=0}^{N-1} X\ [\text{t} - \text{i}] \quad (4)$$

where t is the current time point, and N is the window size (number of points included in the average).

Modeling the relationship between observations and hidden states as well as state transitions is the goal of the linear dynamic system (LDS) algorithm. The definition of a linear dynamic system is as follows:

$$X_t = z_t + w_t \quad (6)$$

$$z_t = A\ z_{t-1} + v_t, \quad (7)$$

where A is a transition matrix, w_t is Gaussian noise with mean $\bar{\omega}$ and variance Q, v_t is Gaussian noise with mean $\bar{v}$ and variance R, and xt and zt stand for the observed and concealed emotion variables, respectively. The corresponding form of these equations can also be written in terms of Gaussian conditional distributions.

$$\boldsymbol{p(x_t|z_t) = N(x_t|z_t)} +\ \bar{\omega}\ , \text{Q}\ , \quad (8)$$

$$\boldsymbol{p(Z_t|z_{t-1}) = N(z_t|Az_{t-1})} +\ \bar{v}\ , \text{R}. \quad (9)$$

It is believed that the initial state is

$$p(z_1) = \text{N}\ ((\boldsymbol{Z_1|\pi_0}), \mathbf{S_0}) \quad (10)$$

Based on the observation sequence$\boldsymbol{x_t}$,, the Expectation Maximization technique can be used to calculate θ using maximum likelihood. It is necessary to compute the marginal distribution, $\boldsymbol{p(Z_t|X)}$ in order to deduce the latent states $\boldsymbol{Z_t}$ from the observation sequence $\boldsymbol{x_t}$. One way to express the latent state is as

$$\boldsymbol{Z_t = E(z_t|X)} \quad (11)$$

where the expectation is indicated by E. The message propagation approach can be used to attain this marginal distribution [5]. To ascertain the LDS algorithm's previous parameters, we employ the cross-validation technique.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

Proposed CNN model

Two 1D Convolution layers [16] are used in the development of the proposed CNN Model for the SEED IV dataset. For jobs involving sequential or time-series data, like EEG signals, a 1D Convolutional Layer is very useful. From sequential data, it is able to automatically extract temporal characteristics or local patterns. It is effective for huge time-series data since it is quicker and requires fewer parameters [17]. Local linkages, such short-term dependencies in time-series data, can be learnt using a sliding window with filters. Weights are shared by 1D Convolution layers throughout the sequence, which lowers the possibility of overfitting and enhances the ability to generalise on unknown facts.

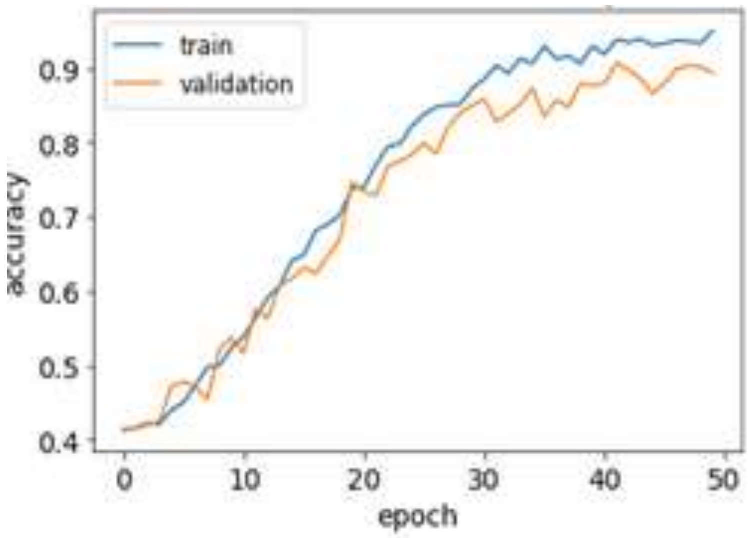
Max pooling is a pooling procedure that picks the maximum element from the filter-covered region of the feature map. Thus, the result of the max-pooling layer would be a feature map with the most prominent features from the prior feature map. Sergey Ioffe and Christian Szegedy [18] introduced it to solve the issue of "internal covariate shift," which occurs when the distribution of each layer's inputs changes during training when the parameters of the preceding layers change. Batch normalisation in CNN [19] solves a number of training-related difficulties. Internal covariate shift happens when the distribution of network activations changes due to parameter updates during training. Batch normalisation tackles this by normalising each layer's activations, ensuring that the mean and variance of inputs remain stable during training.

This model contains two ID convolution layers, a batch normalization layer, two dropout layers two maxpooling layers and two dense layers. The first convolution layer of the model applies 32 filters, each of size 1, across the input data. The filter size of 1 means the model looks at each individual feature separately and tries to extract patterns. ReLu activation introduces non-linearity. It ensures that only positive values pass through, helping the model learn complex features. Second convolution layer applies 64 filters capturing more complex patterns from the down-sampled output of the previous layer. A larger number of filters helps the model to extract more diverse and higher-level features. ReLu keeps only positive values, enhancing sparsity and reducing the risk of vanishing gradients.

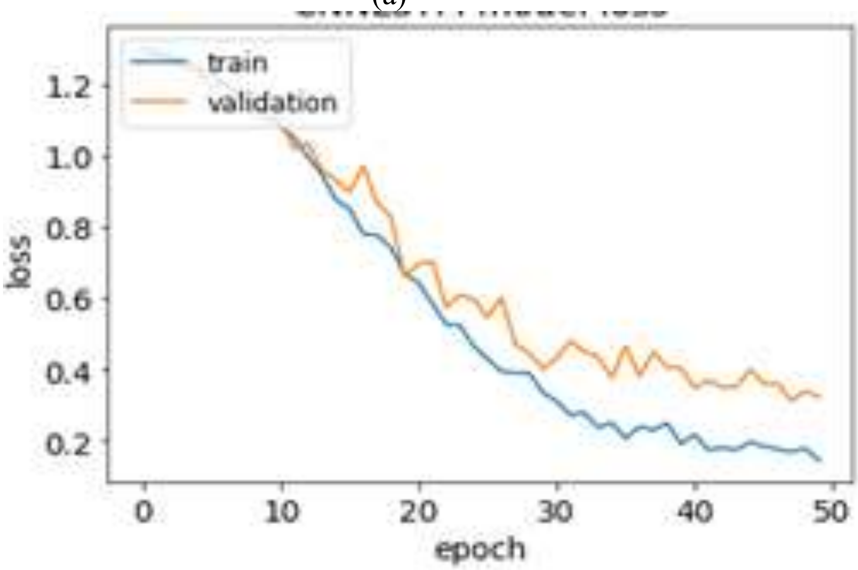
RESULT AND DISCUSSION

In this part, we discuss the success of detecting PSD features based on accuracy, precision, recall, and F1 scores. The recommended approach is implemented in Python. We have Windows 2010, 8GB of RAM, an Intel i5 CPU, and an NVIDIA Tesla K40 GPU. Out of the total data, we utilised 70% for training and 30% for testing emotion identification approaches to choose the best feature. Classification is performed using the suggested CNN model. To prevent overfitting, we set the batch size to 32, the number of epochs to 5, and the dropout rate at 0.2. For all sorts of features, the Adam optimiser is configured with a learning rate of 0.001, beta1=0.9, and beta2=0.999.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026



(a)



(b)

Fig.1: Training and validation curves for (a) accuracy and (b) loss

Table I. EEG emotion recognition classification performance for SEED IV dataset with different methods.

Method	Description	Feature extraction method	Accuracy/STD. (%)
SVM [21]	Support vector machine	DE (Differential Entropy)	56.6/20.25
RF [22]	Random forest	DE	50.91/16.18
CCA [23]	Canonical correlation analysis	PSD	54.5/18.38
GSCCA [24]	Group sparse canonical correlation analysis	PSD	69.18/16.25
DBN [13]	Deep belief network [13]	PSD	66.97/7.28
GRSLR [25]	Graph regularized sparse linear regression	DE	69.12/15.32
GCNN [26]	Graph convolution neural network	EEG graph features	68.25/15.0
DGCNN [30	Dynamic graph convolution neural network	Dynamic Graph of DE	70.18/16.09
DANN [27]	Domain adversarial neural network	DE	63.27/12.46
BiDANN [28]	Bi-hemisphere domain adversarial	Differential asymmetry of DE	70.19/12.53

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

	neural network		
Emotion-Meter [20]	Emotion-Meter	Multimodal feature	70.49/17.21
BiHDM [29]	Bi-hemispheric discrepancy model	DE	74.15/14.29
CNN [13]	Separable depthwise CNN	DE	72.83/5.72
Proposed CNN	Convolution Neural Network	PSD	90/4.3

The graph in the figure 1 depicts the accuracy of a model during training and validation across 50 epochs. This demonstrates how well the model fits the training data. It continuously grows and reaches around 90%. Validation Accuracy demonstrates the model's ability to generalise to new data. At first, validation accuracy rises in tandem with training accuracy, but after around 20 epochs, it plateaus at 88%. In conclusion, the model gradually increases its accuracy on both training and validation data.

1D CNNs carry out automated feature selection directly on the PSD, in contrast to techniques like GNNs or CCA, where the feature selection stage is distinct. Hierarchical features that capture minute temporal variations in EEG data are learnt by convolutional layers. EEG signals are time-series data, and identifying emotions depends on the temporal relationships between successive samples. 1D CNNs successfully learn temporal patterns by applying sliding convolutional filters over the signal.

On the other hand, conventional machine learning techniques such as Random Forest and SVM depend on static characteristics and are unable to identify temporal correlations. In 1D CNNs, the convolution process reduces the chance of overfitting while increasing computing efficiency by sharing parameters across the signal. Because of its efficiency, 1D CNNs can analyse huge EEG datasets with additional layers, which enhances performance. 1D CNNs are deep learning models that are able to identify nonlinear correlations in EEG signals that linear techniques such as CCA or sparse linear regression frequently overlook. 1D CNNs operate directly, whereas graph-based models (like GNN) require graph building as input, which increases complexity. GNNs are good at capturing spatial connections, but without other methods (like temporal GNNs), they might not be able to capture temporal dynamics well. CCA ignores deep learning's ability to recognise intricate patterns in favour of concentrating on correlations. Hemisphere Models just take into account hemisphere asymmetry and can overlook other crucial connections in EEG data, such as temporal variations. Random Forest and SVM are conventional machine learning techniques which are ineffective with high-dimensional. Thus For EEG-based emotion detection tasks, 1D CNNs are now the go-to option because of their efficiency, accuracy, and capacity to record rich temporal patterns.

CONCLUSIONS

This article proposes a unique multichannel EEG-based emotion identification technique based on CNN and PSD feature representation. The unique information of each EEG channel and the correlation information across brain channels are extracted from the PSD feature matrix for emotion recognition using a deep CNN, which has 1D convolution layers. The findings demonstrate that multichannel EEG signals' emotion-related properties may be efficiently represented by a PSD feature matrix, and the suggested CNN can fully utilise these features for emotion identification. Additionally, when compared to other current approaches, the suggested method's emotion identification performance demonstrates benefits, demonstrating the method's viability and efficacy.

REFERENCES

1. R. W. Picard. Affective Computing. Cambridge, MA: MIT Press, 1997.
2. Xiang. Li, Dawei. Song, Peng. Zhang, Yazhou. Zhang, Yuexian. Hou, and Bin. Hu, “Exploring EEG features in cross-subject emotion recognition,” Frontiers in Neuroscience, 12:162, 2018.
3. C. Niemic, “Studies of emotion: A theoretical and empirical review of psychophysiological studies of emotion.” Journal of Undergraduate Research, 2004.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

4. H. Mohamed, B. Zaidan and A. A. Zaidan, "A systematic review for human EEG brain signals based emotion classification, feature extraction, brain condition, group comparison," Journal of Medical systems, vol.42, pp. 162, 2018, doi: 10.1007/s10916-018-1020-8.
5. Duan R.-N, Wang X.-W, Lu B.-L. EEG-based emotion recognition in listening music by using support vector machine and linear dynamic system[C]// International Conference on Neural Information Processing. Springer. [S.l.]: [s.n.], 2012: 468–475.
6. Casella G, Fienberg S, Olkin I. Time Series Analysis and Its Applications - With R Examples [J]. Journal of the American Statistical Association, 2006, 97(458): 656–657.
7. Schirrmeister R.T, Springenberg J.T, Fiederer L.D.J, Glasstetter M, Eggensperger K, Tangermann M, Hutter F, Burgard W, Ball T. "Convolutional neural networks for deep learning in EEG decoding and visualization," 2017; 38:5391–420; Hum Brain Mapp.
8. W. L. Zheng and B. L. lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," IEEE Transaction on Autonomous mental development, vol. 7, no. 3, pp. 162-175, 2015.
9. P. Zhong, Di Wang and C. Miao, "EEG-based emotion recognition using regularized graph neural network," IEEE transaction on Affective Computing, vol.11, no.2, pp. 99-111, 2020.
10. T. Song, W. Zheng, P. Song and Z. Cui' "EEG emotion recognition using dynamical graph convolutional neural networks," IEEE transactions on Affective Computing, vol. 11, no. 3, pp. 532-541, 2020.
11. Jinpeng Li, Z. Zhang and H. He, "Implementation of EEG emotion recognition system based on hierarchical convolutional neural network," International conference on Brain inspired Cognitive Systems, Nov. 2016.
12. Y. Li, W. Zheng, Y. Zong, T. Zhang and X. Zhou, "A bi-hemisphere domain adversarial neural network model for EEG emotion recognition," IEEE transaction on Affective Computing, vol. 12, no. 2, pp. 494-504, 2021.
13. Nandini K. Bhandari, and Manish Jain, "EEG Emotion Recognition using Convolution Neural Network," 4th International Conference on Machine Intelligence and Signal Processing, NIT Raipur, 12-14 March, 2022.
14. https://bcmi.sjtu.edu.cn/home/seed/seed-iv.html
15. S. Noshadi, V. Abootalebai, M. T. Sadeghi, and M. S. Shahvazian, "Selection of an efficient feature space for EEG-based mental task discrimination," Bio cybernetics and biomedical Eng., vol. 34, no. 3, pp. 159-168, 2014.
16. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015). https://doi.org/10.1038/nature14539
17. P. Amini, S. A. Zahiri Motlagh and M. Nezhadpour, "A Large-Scale Infrastructure for Serious Games Services," 2nd National and 1st International Digital Games Research Conference: Trends, Technologies, and Applications (DGRC), Tehran, Iran, pp. 27-33, 2018.
18. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift", 2015.
19. T. Sledevic, "Adaptation of Convolution and Batch Normalization Layer for CNN Implementation on FPGA," 2019 Open Conference of Electrical, Electronic and Information Sciences (eStream), Vilnius, Lithuania, 2019, pp. 1-4, doi: 10.1109/eStream.2019.8732160.
20. W.L. Zheng, W. Liu, Y. Lu, B. L. Lu and A. Cichocki, "EmotionMerer: A multimodal framework for recognizing human emotions," IEEE transactions on Cybernetics, vol. 49, no. 3, pp. 1110-1122, 2019.
21. J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," Neural Processing Letters, vol. 9, no. 3, pp. 293-300, 1999.
22. L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.
23. B. Thompson, "Canonical correlation analysis," Encyclopedia of Statistics in Behavioral Science, 2005.
24. W. Zheng, "Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis," IEEE transaction on Cognitive and development systems, vol. 9, no. 3, pp. 281-290, 2017
25. Y. Li, W. Zheng, Z. Cui, Y. Zong and S. Ge, "EEG emotion recognition based on graph regularized sparse linear regression," Neural Processing Letters, pp. 1-17, 2018.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VIII, August 2026

26. M. Defferrard, X. Bresson and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," Conference on Neural information processing systems (NIPS), pp. 3844-3852, 2016.
27. Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F.Laviolette, M. Marchand and V. Lampitsky, "Domain-adversarial training of neural networks," journal of Machine Learning Research, vol.17, no. 59, pp. 1-35, 2016.
28. Y. Li, W. Zheng, Z. Cui, T. Zhang and Y. Zong, "A novel neural network model based on cerebral hemispheric asymmetry for EEG emotion recognition," International joint Conference on Artificial intelligence, pp. 1561-1567, 2018.
29. Y. Li, L.Wang, T. Song, W. Zheng, Y. Zong, "A novel Bi-hemispheric discrepancy model for EEG emotion recognition," IEEE transaction on Cognitive and Development systems, vol. 13, no. 2, pp. 354-367, 2021.
30. T. Song, W. Zheng, P. Song and Z. Cui' "EEG emotion recognition using dynamical graph convolutional neural networks," IEEE transactions on Affective Computing, vol. 11, no. 3, pp. 532-541, 2020