Computer Methods and Programs in Biomedicine 243 (2024) 107927

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

EESCN: A novel spiking neural network method for EEG-based emotion recognition

FeiFan Xu [a], Deng Pan [a], Haohao Zheng [a], Yu Ouyang [a], Zhe Jia [a], Hong Zeng [a,b,*]

[a] Hangzhou Dianzi University, School of Computer Science and Technology, HangZhou, ZheJiang, China
[b] Key Laboratory of Brain Machine Collaborative of Zhejiang Province, HangZhou, ZheJiang, China

ARTICLE INFO

Keywords:
Convolutional neural network (CNN)
EEG emotion recognition
Neuromorphic
Recurrent neural network (RNN)
Spiking neural network (SNN)

ABSTRACT

Background and Objective: Although existing artificial neural networks have achieved good results in electroencephalograph (EEG) emotion recognition, further improvements are needed in terms of bio-interpretability and robustness. In this research, we aim to develop a highly efficient and high-performance method for emotion recognition based on EEG.

Methods: We propose an Emo-EEGSpikeConvNet (EESCN), a novel emotion recognition method based on spiking neural network (SNN). It consists of a neuromorphic data generation module and a NeuroSpiking framework. The neuromorphic data generation module converts EEG data into 2D frame format as input to the NeuroSpiking framework, while the NeuroSpiking framework is used to extract spatio-temporal features of EEG for classification.

Results: EESCN achieves high emotion recognition accuracies on DEAP and SEED-IV datasets, ranging from 94.56% to 94.81% on DEAP and a mean accuracy of 79.65% on SEED-IV. Compared to existing SNN methods, EESCN significantly improves EEG emotion recognition performance. In addition, it also has the advantages of faster running speed and less memory footprint.

Conclusions: EESCN has shown excellent performance and efficiency in EEG-based emotion recognition with potential for practical applications requiring portability and resource constraints.

1. Introduction

In recent years, emotion recognition has played an important role in affective Brain-computer Interface (aBCI). Since emotion is a complex mental state that reflects the high-level activities of the brain, it is still a challenge to accurately identify different emotion states. Although previous research on emotion recognition based on facial expression, speech, and text has made important progress, these methods often struggle with the issue of "deliberate hiding of emotional states" which makes it difficult to identify the genuine emotional states. Electroencephalogram (EEG) contains rich emotional information and is an objective physiological signal that is low-cost and easy to be collected. Therefore, EEG-based emotion recognition is receiving more and more attention [1,2].

Artificial neural networks (ANNs), especially deep learning methods have demonstrated excellent performance in EEG-based emotion identification [3] for the ability to learn high-level features from EEG automatically. Yang et al. [3] proposed a deep neural network topology called Spatio-spectral representation learning (SSRL) for decoding cortical processes associated with different walking conditions using EEG signals. Xiao et al. [4] introduced a 4D attention-based neural network, which combined attention-based convolutional neural network (CNN) and recurrent neural network (RNN), for emotion recognition. Song et al. [5] proposed a multi-channel EEG-based emotion recognition method called the dynamical graph convolutional neural networks (DGCNN), using a graph to learn the intrinsic relationship between different EEG channels dynamically. Inspired by the neuroscience study that reveals the discrepancy of emotion expression, Li et al. [6] proposed a bi-hemispheric discrepancy model (BiHDM) for EEG emotion recognition, employing four directed RNNs to traverse electrode signals on two brain regions and captures the discrepancy information between two hemispheres with a pairwise subnetwork. Yang et al. [7] presented a hierarchical network structure with subnetwork nodes to discriminate three emotions (positive, neutral and negative) based on EEG data.

* Corresponding author at: Hangzhou Dianzi University, School of Computer Science and Technology, HangZhou, ZheJiang, China.
E-mail addresses: 212050125@hdu.edu.cn (F. Xu), dpan@hdu.edu.cn (D. Pan), zhenghaohao@hdu.edu.cn (H. Zheng), 222050269@hdu.edu.cn (Y. Ouyang), 222050129@hdu.edu.cn (Z. Jia), jivon@hdu.edu.cn (H. Zeng).

https://doi.org/10.1016/j.cmpb.2023.107927
Received 18 May 2023; Received in revised form 16 October 2023; Accepted 13 November 2023
Available online 20 November 2023

Similar works were also conducted in this field [8–10]. Although the existing deep learning methods have achieved a certain success, most of them do not consider the intrinsic neurophysiological characteristics of EEG, resulting in poor robustness and generalization ability, as well as huge computational resource consumption.

As the new generation of artificial neural networks, spiking neural networks (SNNs) have demonstrated their capability in signal processing and target recognition by mimicking the behavior of biological neural systems [11], especially in accurately modeling temporal dynamics data. EEG signals are time-varying, and are more suitable for SNNs in emotion recognition to conduct EEG analysis with high temporal precision. To optimize SNNs in EEG-based analysis, existing SNNs have been explored from the perspective of spiking neurons and the spiking layers. Among them, the optimization of spiking neurons mainly used NeuCube [12] to simulate the biological properties and spatial structure of neurons to construct SNNs, so as to assist them in processing complex EEG data. The optimization of the spiking layers focuses on earnable synaptic weights or membrane time constants of spiking layer nodes to enhance the performance of SNNs with the nature of multi-layer interaction [13]. However, the potential of SNNs has not been fully explored in the current EEG-based analysis. One of the reasons is the lack of unified and robust training methods in EEG data processing and analysis, and the performance of SNNs still cannot be compared with high-performance deep neural networks (DNNs). The second reason is that SNNs lack efficient feature extraction methods to extract EEG non-associated relationships in spatial-temporal dimensions, which will also affect the performance of SNN classifier representations [14–16,12]. The third reason is that EEG signals have significant individual differences and non-stationary characteristics [17,18], most conventional SNNs have insufficient ability to comprehensively analyze both global and local features of EEG, which limits the further improvement of their robustness and generalization ability. The reason for this issue is that conventional SNNs tend to only analyze either the global features or the local features of EEG like SNNs based on Spike-Timing-Dependent Plasticity (STDP) [19] or the rank order rule [20,21]

To address the above issues, inspired by the recent work of deep SNNs, which has been shown to be effective on neuromorphic datasets [13,22], we propose Emo-EEGSpikeConvNet (EESCN) method for EEG emotion identification that combines CNN with spiking neuron mechanisms. To tackle the first issue, we integrated convolutional and fully-connected modules into the EESCN. Through parameter optimization [23], we enhanced network performance by adjusting parameters like kernel size, padding size, and hidden layer neuron scale, etc. For the second issue, we initially transform EEG signals into a neuromorphic 2D frame series using the Neuromorphic data generation module. The EESCN's continuous spiking neurons then transmit spiking signals, extracting the temporal features of 2D frame series. In response to the third issue, the Neuromorphic data generation module employs the convolution module and spiking neurons to facilitate global spatial and temporal feature extraction, respectively. In addition, the spiking neurons will obtain the local features through the mutual interaction in the training stage of EESCN, enhancing the local feature extraction capabilities.

The brief description of the process is as follows: First of all, we transform EEG signals into a neuromorphic 2D frame series format consisting of continuous 2D map temporal frames. Each of these frames mines the spatial correlation of different EEG channels and the continuous frames contain temporal information. It can capture the 2D spatial data in each time step frame and use the temporal data from the whole continuous frames. And the binary nature of the spiking mechanism has the added benefit of lowering computational costs.

Our contributions are summarized as follows: First, we propose a novel emotion recognition method based on Spiking Neural Networks (SNNs), called EESCN. This method includes a processing module for generating neuromorphic data and an SNN called NeuroSpiking framework for spatial-temporal feature extraction and EEG-based emotion classification.

Second, our EESCN method utilizes the unique characteristics of the binary nature of the spiking mechanism to reduce floating-point operations and memory usage during computation. Consequently, it is approximately three times faster on average and requires a substantially smaller memory footprint when compared to a conventional CNN+RNN with an analogous structure.

Third, we evaluate the EESCN method on inter-subject and cross-video with the DEAP and SEED-IV datasets. The experimental results demonstrate average emotion recognition accuracies of 94.56%, 94.81% and 94.73% on the valence, arousal and dominance of DEAP and a mean accuracy of 79.65% on SEED-IV. Compared to existing SNN methods, our method shows a significant improvement in classification performance in EEG emotion recognition.

2. Related works

In this section, we review current SNNs in the efforts of improving their abilities of robustness and generalization, such as unsupervised learning of SNNs, ANN to SNN conversion methods, supervised learning of SNN methods and the relevant research in SNN-based EEG analysis.

Unsupervised learning of SNNs: The unsupervised learning methods of SNNs are based on biologically plausible local learning rules, like Hebbian learning [24] and STDP. In [12,25,26], the authors presented the classification method of EEG data based on STDP. However, these methods are only applicable to shallow SNNs, and their performance is lower than the state-of-the-art ANN. In addition, although the learning approaches have excellent biological plausibility, they still need to be run in a specific simulation environment or a better hardware configuration, which limits the further performance improvement of current SNNs.

ANN to SNN conversion: The conversion from ANN to SNN ($ANN2SNN$) has also received much attention. This is due to the extremely strong correlation between the nonlinear activation of ReLU neurons in ANN and the issuance rate of Integrate-and-Fire (IF) neurons in SNN, which can be used to convert the trained ANN to the corresponding SNN while ensuring that there is not too much performance loss [27–29], but there is a trade-off between accuracy and latency. In recent years, $ANN2SNN$ has made some progress [30,31,27,32,33], their single-layer network structure is difficult to fully exploit the temporal dynamic features of EEG, which hinders the performance improvement of SNNs as well.

Supervised learning of SNNs: The SpikeProp [34] was the first supervised learning method for SNNs using gradient descent strategy based on backpropagation. To overcome SNNs' non-differentiable threshold-triggered firing mechanism, it adopted a linear approximation. Subsequent works included Tempotron [35], Re-SuMe [36], and SPAN [37], but they could only be applied to single-layer SNNs. For training multi-layer SNNs, a surrogate gradient method has recently been proposed. For example, Wu et al. proposed Spatio-Temporal Back-Propagation (STBP) [38] utilizes both the spatial domain and the temporal domain during the training phase, building a more user-friendly iterative Leaky Integrate and Fire (LIF) model for gradient descent training.

This kind of SNNs can use surrogate derivatives to define the derivative of the threshold trigger firing mechanism [11,39], thus this kind of SNN could be optimized with gradient descent algorithms as ANNs. Compared with $ANN2SNN$, the surrogate gradient method has no restrictions on simulating time steps and has high potential in dynamic data.

EEG analysis based on SNNs: At present, many studies have applied the optimized SNN models to EEG-based tasks. Luo et al. [12] proposed a Neucube SNN model with a variance method to extract EEG features and classify emotions. Capecci et al. [16] also applied a Neucube SNN model to analyze EEG data of patients with Alzheimer's disease. Antelis

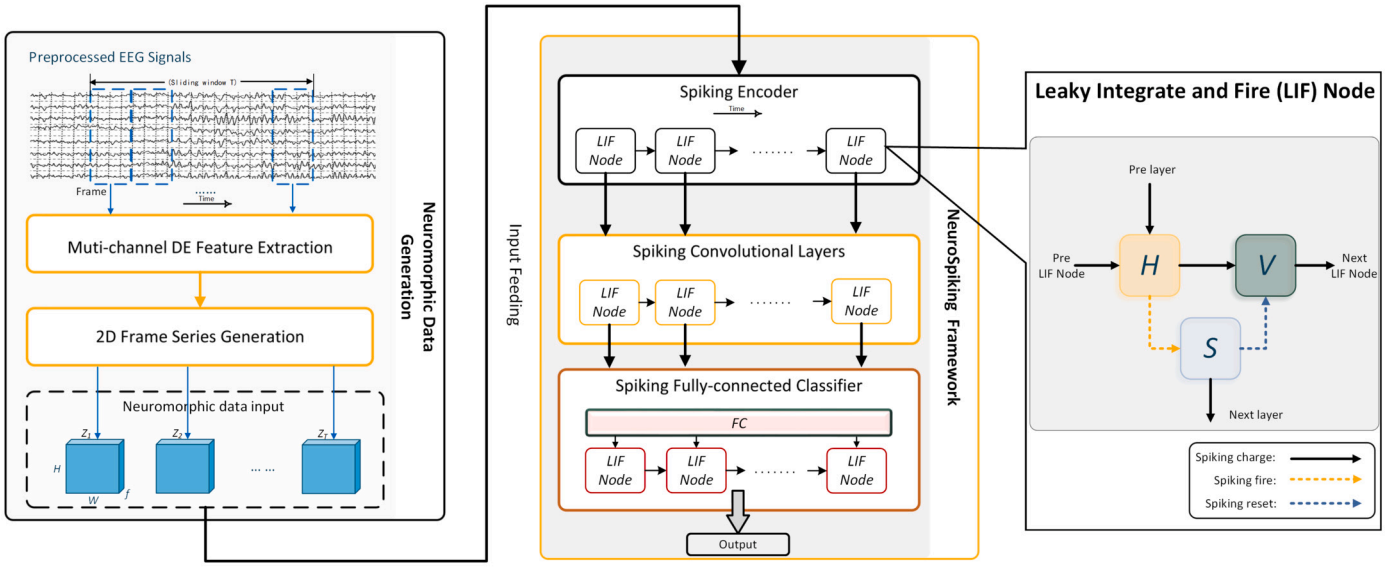

Fig. 1. The schematic diagram of EESCN.

et al. [40] reported the progress of SNN methods in EEG motor imagery recognition tasks and pointed out that SNNs have better recognition ability for the tasks.

In summary, unsupervised learning methods of SNN require a unique simulation environment and that cannot be optimized using existing deep learning methods, leading to low classification accuracy. However, EESCN improves classification by including deep learning techniques like convolutional and fully-connected modules. Unlike the ANN2SNN approach, EESCN uses LIF nodes and spikes for data transmission, mirroring brain neuron structures for, which is better biologically plausible and helps to study brain activities and neural patterns.

3. Methods

The schematic diagram of our proposed EESCN is shown in Fig. 1, consisting of neuromorphic data generation and NeuroSpiking framework. In neuromorphic data generation, it is responsible for converting EEG into a 2D neuromorphic data format and generating the frame series to input into the corresponding LIF nodes in the Spiking Encoder layer of the NeuroSpiking framework, so as to achieve the temporal interaction between nodes and learning of temporal features of neuromorphic data. The NeuroSpiking framework conducts EEG emotion classification through three spiking layers containing different LIF nodes, which are the Spiking Encoder, the Spiking Convolutional Layers and the Spiking Fully-connected Classifier.

3.1. The neuromorphic data generation module

First, EEG is preprocessed by down-sampling and bandpass filtering, then the preprocessed EEG is divided into segments by sliding window of length T, and each sequence of length T is further segmented into D frames with length t (Fig. 2). Next, according to previous studies [4,23], 4 frequency bands: $\theta(4-8Hz)$, $\alpha(8-13Hz)$, $\beta(13-30Hz)$ and $\gamma(30-45Hz)$ of the EEG signal have a specific effect on EEG emotion recognition. Therefore, Difference Entropy (DE) features of the 4 frequency bands on the D frames of all C channels are extracted. The DE features from each frame in time period t can be represented by 1D vector $X_{d,c} = [x^{\theta}_{d,c}, x^{\alpha}_{d,c}, x^{\beta}_{d,c}, x^{\gamma}_{d,c}]$, $d \in (1,2,...,D)$ represents the d-th frame in a segment and $c \in (1,2,...,C)$ represents the c-th channel.

Since the EEG signal collected on an electrode will be affected by multiple electrodes around it, in a specific brain region, but the 1D EEG feature vector only reflects the correlation between two adjacent electrodes. Therefore, we convert the 1D vectors in a frame into a 2D

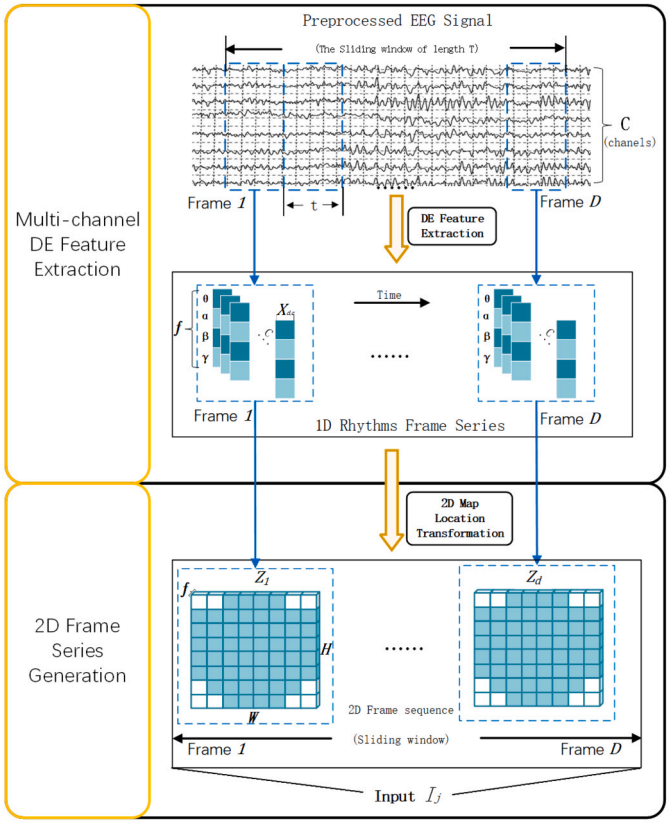

Fig. 2. Neuromorphic data generation, including Multi-Channel DE feature extraction and 2D frame series generation.

frame Z_d, $d \in (1,2,...,D)$. Each vector based on its index c (c-th channel) is then mapped onto an EEG 2D topographical electrode location. The channels on the 2D map correspond to the channels of the 1D vector, which all derive from the electrode channels of the EEG recording device. The shape of 2D frame is (f,H,W), where the f includes DE features of 4 frequency bands, H and W represent the height and width of the EEG 2D topographical electrode location map. In this way, the obtained 2D frame Z_d contains the temporal, spatial and frequency domain information of the EEG signal.

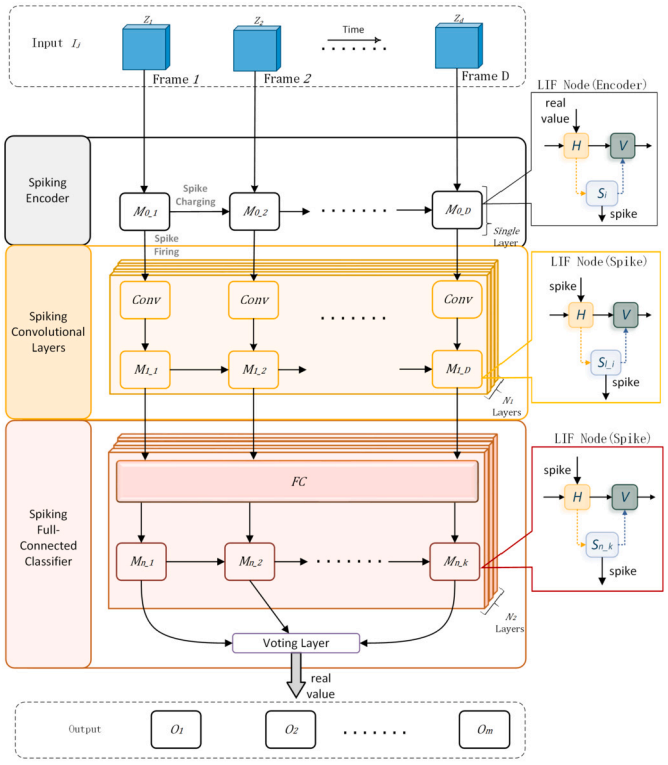

Fig. 3. The NeuroSpiking framework. $O_1 \sim O_m$ represent the output of the network and m is the number of categories in the classification task.

Finally, the $[Z_1, Z_2, ..., Z_D]$ in each segment is used as the subsequent framework's input $I_j, j \in (1, 2, ..., q)$, q is the segment number of one EEG trial. In this way, we can better extract spatial-temporal features in EEG signals.

3.2. The NeuroSpiking framework

Fig. 3 shows the detail of the proposed NeuroSpiking framework, including the Spiking Encoder, the Spiking Convolutional Layers, and the Spiking Fully-connected Classifier, respectively. The Spiking Encoder converts real-valued data inputs into spiking-formed binarized data for spike interaction in SNN. The Spiking Convolutional Layers consist of N_1 identical layers with convolution (Conv) modules and LIF nodes, extracting spatial and temporal features from spiking signals. The Spiking Fully-connected Classifier consists of N_2 identical hidden layers with a fully-connected layer and LIF nodes.

3.2.1. Spiking encoder

EEG 2D frame sequence I_j needs to be encoded into binarized spikes before feeding into SNN. The Spiking Encoder is designed to convert real-valued inputs into spiking-formed binarized spikes by the LIF Node (Encoder) at the beginning of NeuroSpiking framework. We take I_j as the input, then directly feed it to the Spiking Encoder. The encoding is done by mutual charging interaction between LIF nodes, which can be seen as a learnable encoder.

3.2.2. Spiking convolutional layers

Motivated by the generalizable feature extraction solutions like CNNs in EEG tasks [4,41], we use the Spiking Convolutional Layers in our NeuroSpiking framework for spatial-temporal feature extraction. As shown in Fig. 3, the Spiking Convolutional Layers consist of N_1 identical layers. Here, taking one layer as an example, the spike input first undergoes convolutional operations through a convolution (Conv) module to extract spatial features before being fed into the LIF nodes. Then, the spike input is fed to mutually connected LIF nodes, as the identical M_{l_D}, $d \in (1, 2, ..., D)$, $l \in (1, 2, ..., N_1)$. In this way, the connective LIF nodes can extract temporal features by spike charging and firing between LIF nodes and Layers. The detail of the mechanism related to LIF nodes will be explained in section 3.3.

3.2.3. Spiking fully-connected classifier

The Spiking Fully-connected Classifier is used for learning and decoding. It consists of N_2 identical hidden layers. Here, we also take one of the hidden layers as an example. It consists of a fully-connected layer and a sequence of LIF nodes. In detail, the fully-connected layer uses a fully-connected neural network for learning, then input into the connective LIF nodes. At the end of the Spiking Fully-connected Classifier, there is a voting layer using average-pooling to implement democratic voting to decode the spike output by LIF nodes into the real-valued. It integrates the $m * n$ output of hidden layers into m pattern classes, where m is the number of classes and n is the number of nodes representing one class.

3.3. The LIF neuron node in the NeuroSpiking framework

3.3.1. Subthreshold dynamics differential equation

The basic computing unit of the NeuroSpiking framework is the neuron node. The LIF model is one of the typical spiking neuron models used in SNNs. The subthreshold dynamics of the LIF neuron can be defined as Eq. (1):

$$c\frac{dV_t}{dt} = -(V_t - V_{reset}) + X(t) \tag{1}$$

For the convenience of understanding, the $[t]$ time step in this section represents the d-th time step in section 3.1 and 3.2. Where V_t represents the membrane potential of the neuron at the time $[t]$, $X(t)$ represents the input to the neuron at the time t, c is the membrane time constant, and V_{reset} is the resting potential. When the membrane potential V_t exceeds a certain threshold V_{th} at time t, the neuron will elicit a spike and then the membrane potential V_t goes back to a reset value V_{reset}. It is a simple model, which can achieve a balance between computing cost and biological plausibility.

3.3.2. Dynamic model of LIF node

The dynamics of the LIF neuron node can be described with the following equations:

$$H_t = (1 - \frac{1}{c})V_{t-1} + \frac{1}{c}X_t \tag{2}$$

$$S_t = \Theta(H_t - V_{th}) \tag{3}$$

$$V_t = H_t(1 - S_t) + V_{reset}S_t \tag{4}$$

They include the threshold-triggered firing mechanism and the reset of the membrane potential after firing, which can be used for numerical simulations. H_t and V_t represent the membrane potential after neuronal dynamics and after the trigger of a spike at time step t, respectively. X_t denotes the external input, and H_{th} denotes firing threshold. S_t denotes the output spike at $[t]$, which equals 1 if there is a spike and 0 otherwise.

- Eq. (2) describes the neuronal dynamics of the LIF neuron node when $V_{reset} = 0$. The integration progress $\frac{1}{c}X_t$ makes the LIF nodes able to remember current input information, while the leakage progress $(1 - \frac{1}{c})V_{t-1}$ can be seen as forgetting some information from the past. This equation shows that the balance between remembrance and forgetting is controlled by the membrane time constant c.
- Eq. (3) describes the spiking generative process, where $\Theta(x)$ is the Heaviside step function, denoting as $\Theta(x) = 1$ when $x \geq 0$ and $\Theta(x) = 0$ when $x < 0$.

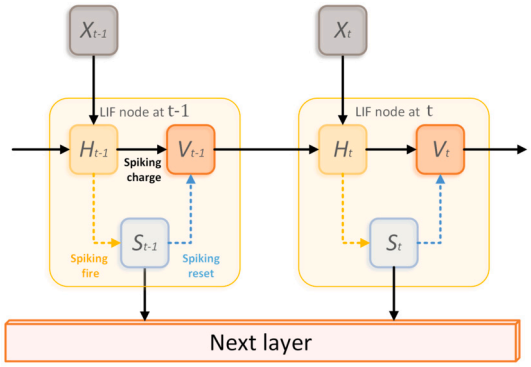

Fig. 4. The Leaky Integrate-and-Fire node model.

- Eq. (4) describes that the membrane potential returns to V_{reset} after eliciting a spike, which is called hard reset and widely used in deep SNNs.

3.3.3. LIF learning mode

As shown in Fig. 4, the LIF model is built for describing the discrete spiking neuron's learning activities, which include the activities of charging, firing, and resetting. The node on the left side indicates the state of the LIF node at $[t-1]$, and the right one indicates the state at $[t]$. The membrane potential after neuronal dynamics H_t is charged by the V_{t-1}. When H_{t-1} gets the input $X(t-1)$ at $[t-1]$, H_{t-1} is changed according to Eq. (1). If the potential is greater than the threshold it generates a spike S_{t-1} according to Eq. (1), and the H_{t-1} is reset to the after-trigger membrane potential V_{t-1} according to Eq. (3). If there is no spiking action, V_{t-1} can be obtained directly. Then, the generated spike will be used as the input of the next layer, i.e., the after-trigger membrane potential V_{t-1} will become the pre-excitation potential H_{t-1} at the next time step $[t]$. This process takes place sequentially between LIF nodes in each layer to transmit temporal information.

It should be noted that the membrane time constant c is optimized automatically during training, rather than being set as a hyperparameter manually before training. Furthermore, c is shared within the nodes in the same layer in NeuroSpiking, which is biologically plausible as the neighboring nodes have similar properties. Since c is different in different layers, so the diversity of the phase frequency response can be obtained.

4. Experiments

4.1. Datasets

In this paper, we validate the proposed EESCN based on two publicly available EEG emotion datasets DEAP [42] and SEED-IV [43], as they are widely used for testing the EEG emotion recognition performance of SNNs and ANNs [44], [45], [46], it is convenient for us to conduct the baseline comparisons.

DEAP: The DEAP is a public EEG emotion dataset. The dataset was recorded from 32 subjects who participated in 40 60-second trials. The emotional EEG signals evoked when each subject was watching the relevant videos were acquired by a 32-channel electrode device of BioSemi ActiveTwo. Then, the subjects were asked to fill in the questionnaire, and their emotional status was scored on dimensional state labels (Arousal, Dominance, and Valence), which possess global psychological functions, forming the basis for important physiological and psychological responses in humans [47,48].

SEED-IV: The SEED-IV is also a publicly available EEG dataset in which four different emotions of happy, sad, fearful and neutral are evoked by watching a video consisting of 168 movie clips, and the corresponding EEG signals are simultaneously captured as well. This dataset contains data collected from 15 subjects, with each subject participating in 3 sessions. There are 24 trials in each session and a total of approximately 1080 trials. The EEG signals were acquired by a 62-channel electrode device of the ESI NeuroScan System. Then, the raw EEG was also downsampled to 200 Hz and filtered to a bandpass of 0-70 Hz. Moreover, as a dataset using discrete state labels, it's crucial to consider factors like individual experiences, state and gender when using the dataset [49].

Table 1
Layer details.

Layer	Details	
Spiking Conv Layer 3x3	Conv 2d Layer	kernel size = 3 padding = 1
	BatchNorm 2d	—
	Dropout	p = 0.25
	MultiStep LIFNode	—
Spiking Conv Layer 1x1	Conv 2d Layer	kernel size = 1 padding = 1
	BatchNorm 2d	—
	Dropout	p = 0.25
	MultiStep LIFNode	—

p represents the ratio of dropout.

4.2. Dataset preprocessing

DEAP dataset preprocessing: To preprocess the DEAP dataset, we use the neuromorphic data generation method in 3.1 to extract the DE features of all trails in the DEAP dataset. Then, we divide the DE features of each trial into 8 non-overlapping T = 7.5 s segments, each segment consisting of 15 non-overlapping frames of t = 0.5 s, and finally, each segment is used as an input sample. The labels for each sample were based on the subjects' evaluation points of the videos, which were classified into two categories based on a threshold of 4: values greater than 4 were labeled as positive 1 and values less than or equal to 4 were labeled as negative 0, then, the labels are mapped to the one-hot format. Finally, the number of each category of samples is equaled. In addition, we follow the settings in the previous works [23,50] for our experiment on the DEAP.

SEED-IV dataset preprocessing: To preprocess the SEED-IV dataset, we also use the neuromorphic data generation module to extract the DE features of all trails in the DEAP. Then, the DE features of each trial are segmented with a non-overlapping window length of T = 4 s, each segment consisting of 4 non-overlapping frames of t = 1 s, and finally, each segment is used as an input sample. Since the four emotions of EEG data have already labeled four classes in SEED-IV, we just map the labels to the one-hot format. In the experimental settings on the SEED-IV, we follow a cross-video classification protocol proposed by Zheng and Zhong [6,51]. In detail, for each subject in a session, the samples of 16 videos are used as the training set and the remaining samples of 8 videos (2 videos per emotion type) as the test set.

4.3. NeuroSpiking framework settings

The NeuroSpiking framework was trained using optimized parameters: batch size of 16, epoch of 300, and Ir of 0.001. We used PyTorch on an AMD 5800X CPU and NVIDIA RTX3090 GPU. Table 1 and 2 list the specific parameters used to configure the framework. For the convenience of understanding, the "MultiStep LIFNode" represents the sequence of the connective LIF nodes mentioned in section 3.2.

4.4. Results

As shown in Table 3 and Table 4, EESCN achieves a stable and outstanding performance in the cross-video classification task on SEED-IV and the inter-subject classification task on DEAP.

Results on SEED-IV: The optimal average accuracy of all subjects in all sessions on SEED-IV is 79.65%. As shown in Table 3, we can find

Table 2
NeuroSpiking framework setting.

Layer		Input Chanels	Output Chanels
Spiking Encoder	MultiStep LIFNode	4	4
Spiking Convolutional Layers	Spiking Conv layer 3x3	4	64
	Spiking Conv Layer 3x3	64	128
	Spiking Conv Layer 3x3	128	256
	Spiking Conv Layer 1x1	256	64
Spiking Fully-Connected Classifier	Fully-connected layer	64*9*9	64*6*6
	MultiStep LIFNode	64*6*6	64*6*6
	Dropout	—	—
	Fully-connected layer	64*6*6	m*10
	MultiStep LIFNode	m*10	m*10
	VotingLayer	m*10	m

m represents the number of categories in classification task.

that the highest average accuracy is 92.41% and the lowest average accuracy is 63.31%, respectively. The reason for the low accuracy might be attributed to the low quality of EEG signals and the fact that this kind of discrete emotions can induce similar brain activities for some subjects [4,48]. In EESCN, when inputting these easily confused EEG signals, the neuromorphic data generation module will generate confusing frames, which in turn affects the classification results. This phenomenon is also noted in the literature [6].

Results on DEAP: The optimal average accuracies of all subjects in valence, arousal and dominance on DEAP are 94.56%, 94.81% and 94.73%, respectively. As shown in Table 4, we can find that EESCN performs well in most cases, e.g. the accuracies in subject #15 reach 99.37%, 97.33%, 99.37% and subject #27 reaches 99.37%, 96.87%, 100.0%. However, there are still low accuracies on a few subjects, such as the accuracies of subject #11 reaches 85.93%, 82.18%, 88.75% and subject #22 reaches 83.96%, 86.84%, 84.50%. The reason might be that some subjects may exhibit differing brain activities in the temporal dimension [47] and the stimuli may not fully evoke certain emotions [4], which result in fewer valid temporal emotional features being included in the Neuromorphic data generation module. This will impact the learning of temporal dimensional features by the LIF Nodes in the NeuroSpiking Framework.

4.5. Baseline comparison

The comparative analysis of EESCN is carried out with those methods of linear support vector machine (Linear-SVM) [52], NeuroSense [44], SNN with transfer learning method (SNN+TL) [53], NeuCube-based SNN [45], functional connectivity network (FCN) [1], SNN with infinite impulse response filters (SNN+IIR) [56], bi-hemisphere domain adversarial neural network (BiDANN) [46], NeuCube SNN [12], bi-hemispheric discrepancy model (BiHDM) [6] regularized graph neural networks (RGNN) [51], spatial temporal feature fusion neural network, attention-based CNN RNN (ACRNN) [54] and Fractal SNN (Fra-SNN) [55]. These methods use the spiking mechanisms or features that are similar to EESCN, so we choose them as representatives for comparison, the comparison results are shown in Table 5.

On the DEAP dataset, EESCN achieves valence, arousal, and dominance classification accuracies of 94.56%, 94.91%, and 94.73%, respectively. It outperforms SNN+IIR, NeuroSense, NeuCube SNN, SNN+TL, ACRNN, Fra-SNN and NeuCube-based SNN methods with a difference in accuracy ranging from 0.84% to 40.95%. Specifically, when compared to the methods that use frequency domain features PSD, such as SVM, SNN+TL and FNC, EESCN achieves at least 11.81%, 10.59%, and 35.13% improvements in valence, arousal, and dominance, respectively. Accordingly, when compared to the SNN methods of NeuroSense, SNN+TL, NeuCube-based SNN, NeuCube SNN and Fra-SNN, EESCN can also achieve at least 9.94%, 10.59%, 14.73% and 21.53% improvements in valence, arousal, and dominance, respectively.

On the SEED-IV dataset, we compared our method with ANN and GNN methods that also use DE features, such as BiDANN, BiHDM and RGNN. The results shown in Table 5 demonstrate that the classification accuracy of EESCN also outperforms those methods ranging from 0.28% to 9.39%.

4.6. Ablation study

To demonstrate the effectiveness of the neuromorphic data generation module and the spiking neuron (LIF Nodes) mechanism, we conducted an ablation study. We conducted experiments by removing the LIF Nodes and Neuromorphic data generation module from the EESCN respectively to demonstrate the effectiveness of these two parts we proposed for our EESCN. We kept the input for each sub-model similar to the proposed EESCN. The results of this study are shown in Table 6.

As shown in Table 6, both the neuromorphic data generation module and the LIF Nodes are crucial to the performance of EESCN. The addition of the neuromorphic data generation module improves the performance of EESCN by at least 17.1%. This is because the neuromorphic data generated by the neuromorphic data generation module can be used to obtain more effective features for the followed LIF Nodes in the NeuroSpiking framework, which can enhance the performance of the EESCN. While the LIF nodes in NeuroSpiking framework could effectively capture the temporal dimension features in the data, which also brings at least 9.33% performance improvement.

5. Discussion

The above-mentioned comparison analysis shows that EESCN could achieve better performance in EEG emotion recognition, it is because unlike the compared ANN and GNN methods that only consider spatial features, our method can also extract temporal features by LIF node sequences and combine them with spatial features for more effective classification. Meanwhile, the existing SNN methods only utilize spiking neuron nodes for spatio-temporal feature extraction and learning, making them ineffective in extracting spatial features. To address this limitation, EESCN not only utilizes LIF Node sequences to extract temporal features, but also uses a high-performance convolution module to extract spatial features, which significantly improves the performance of EESCN.

5.1. Impacts of sliding window length T of EESCN

According to previous research [57], the length of the sliding window T in SNN will affect the range of information extracted by LIF nodes and the ability of LIF nodes to extract temporal features, finally affecting the performance of the method.

We conduct experiments on SEED-IV and DEAP using different T values within a reasonable emotion duration time [58]. As shown in Fig. 5 and Fig. 6, the change of T does affect the performance of EESCN.

Firstly, the continuous increase of T does not lead to continuous performance improvement. For instance, the highest accuracies on SEED-IV and DEAP are achieved at T = 12 s and T = 7.5 s, respectively. When increasing T beyond these Ts, the increase in T will result in a decrease in accuracies. The possible reason is that the LIF node memory length cannot be increased indefinitely, similar to the fact that RNNs store a limit number of temporal features in memory and beyond a certain range, some significant temporal features will be lost, leading to performance degradation.

Secondly, although the change of T does affect the performance, the robustness of our method can overcome this limitation to some extent. For example, on the SEED-IV dataset, the largest gap is 1.62% when comparing T = 12 s and T = 16 s. On the DEAP dataset, when comparing the best result at T = 7.5 s with the worst result at T = 10 s, the largest gaps are 0.39%, 3.99% and 1.91% in valence, arousal and dominance, respectively. The difference between the accuracies of Ts is relatively

Table 3
EESCN performance on SEED-IV.

Subject Accuracy (%)															
Subject	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*
session 1	70.93	82.56	65.12	79.07	58.14	68.60	65.12	83.72	84.88	82.56	66.28	68.60	83.72	76.74	86.05
session 2	86.73	90.82	75.51	94.90	81.63	86.73	88.78	100.0	70.41	85.71	82.65	55.10	72.45	77.55	100.0
session 3	85.71	94.81	67.53	79.22	88.31	100.0	76.62	93.51	75.32	96.10	74.03	66.23	83.12	61.04	83.12
ACC AVG	81.13	89.39	69.39	84.40	76.03	85.11	76.84	**92.41**	76.87	88.13	74.32	**63.31**	79.76	71.78	89.72
ACC STD	8.55	**5.87**	5.95	8.13	**14.06**	13.52	10.0	7.51	7.13	6.03	8.11	10.75	9.51	9.37	7.69
Macro-F1 Score								0.7948							

The EESCN achieved optimal performance on SEED-IV cross-video classification on 3 sessions when using a T = 12 s sliding window.

Table 4
EESCN performance on DEAP.

Subject Accuracy (%)																
Subject	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	16
V	98.70	95.96	89.69	94.06	93.44	95.31	98.44	**99.37**	95.83	94.69	**80.31**	88.70	98.71	95.62	**99.37**	99.06
A	97.19	96.87	93.70	90.62	89.69	97.19	98.12	98.07	93.90	98.70	85.94	95.62	95.62	95.31	97.33	98.70
D	91.87	90.31	90.38	95.62	91.20	95.20	98.09	98.40	98.04	95.31	**82.19**	93.44	90.94	90.00	99.37	94.37
Subject Accuracy (%)																
Subject	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*
V	91.87	97.02	97.00	98.12	94.37	84.00	96.79	94.90	89.42	92.19	99.37	92.19	95.31	90.62	98.31	95.94
A	94.57	95.62	94.37	95.62	**99.37**	**83.96**	99.32	96.20	95.82	91.56	96.87	94.83	96.73	88.47	94.23	93.57
D	98.10	98.44	96.78	93.44	96.81	86.84	98.40	91.56	**100.0**	91.87	**100.0**	93.60	96.56	92.81	96.87	99.69
V Average Accuracy (%)		V Average STD (%)			A Average Accuracy (%)			A Average STD (%)			D Average Accuracy (%)			D Average STD (%)		
94.56		4.18			95.81			3.62			94.73			4.12		
V AUC		V F1 Score			A AUC			A F1 Score			D AUC			D F1 Score		
0.9521		0.9484			0.9567			0.9514			0.9404			0.9514		

Valence (V), Arousal (A), Dominance (D).
The optimal performance of EESCN on DEAP dataset, which T = 7.5 s under a five-fold cross-validation for each subject.

Table 5
The accuracies of the compared methods.

Methods	DEAP-Valence Accuracy (%)	DEAP-Arousal Accuracy (%)	DEAP-Dominance Accuracy (%)	SEED-IV Accuracy (%)	Features
Linear-SVM [52]	66.47	60.33	—	—	PSD
NeuroSense [44]	67.76	78.97	—	—	Simple spike-based features
SNN+TL [53]	82.75	84.22	—	—	PSD
NeuCube-based SNN [45]	84.62	61.54	—	—	Preprocessed EEG + EOG
ACRNN [54]	93.72	93.38	—	—	Processed EEG
Fra-SNN [55]	69.84	69.61	73.20	—	DE
FCN [1]	45.55	62.73	59.60	—	PSD+ENP
SNN+IIR [56]	61.15	53.86	67.50	—	Processed EEG
NeuCube SNN [12]	78.00	74.00	80.00	—	Variance+FFT+DWT
BiDANN [46]	—	—	—	70.29	DE
BiHDM [6]	—	—	—	74.35	DE
RGNN [51]	—	—	—	79.37	DE
EESCN (Ours)	**94.56 ± 4.18**	**94.81 ± 3.62**	**94.73 ± 4.12**	**79.65 ± 8.22**	DE

Differential Entropy – DE, Power Spectral Density – PSD, Fast Fourier Transform – FFT, Discrete Wavelet Transformation – DWT.

Table 6
Accuracy results for EESCN with different variations.

	DEAP-V Accuracy (%)	DEAP-A Accuracy (%)	DEAP-D Accuracy (%)	SEED-IV Accuracy (%)
EESCN without NDGM	73.16 ± 7.16	75.25 ± 4.12	71.23 ± 8.91	62.55 ± 9.21
EESCN without LN	80.66 ± 9.33	84.85 ± 5.74	79.62 ± 6.51	70.32 ± 9.82
EESCN	94.56 ± 4.18	94.81 ± 3.62	94.73 ± 4.12	79.65 ± 8.22

NDGM: Neuromorphic data generation module.
LN: LIF Nodes.

small on EESCN. This is because the multiple states of LIF nodes mentioned in 3.3 incorporate more effective temporal features, ensuring the stability of the method.

5.2. The efficiency of EESCN

Due to the binary nature of the spiking mechanism of SNN, the performance of EESCN is improved compared with conventional CNNs and RNNs in terms of running speed and memory footprint. To prove this, we compared the performance of EESCN and CNN+RNN with the same settings on SEED-IV. The settings of CNN+RNN are similar to Table 1, except for replacing the MultiStep LIFNode layer with RNN(T) to simulate temporal propagation of LIF nodes within a sliding window of length T.

As shown in Fig. 7, we use the *Estimated Total Size* to compare the memory footprint, which is the amount of data that the entire model costs in memory. It is observed that an increase in T results in an increase in the model size due to more LIF nodes interacting with each

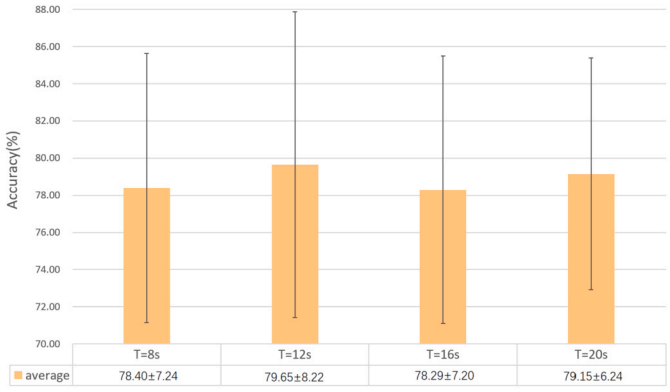

Fig. 5. The performance of different sliding window length T on SEED-IV.

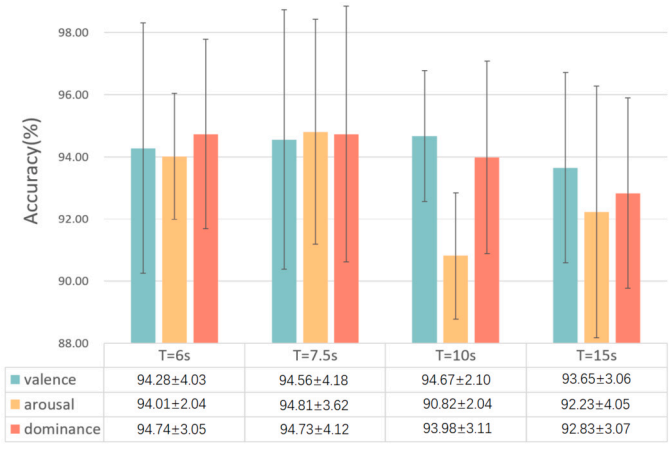

Fig. 6. The performance of different sliding window length T on DEAP.

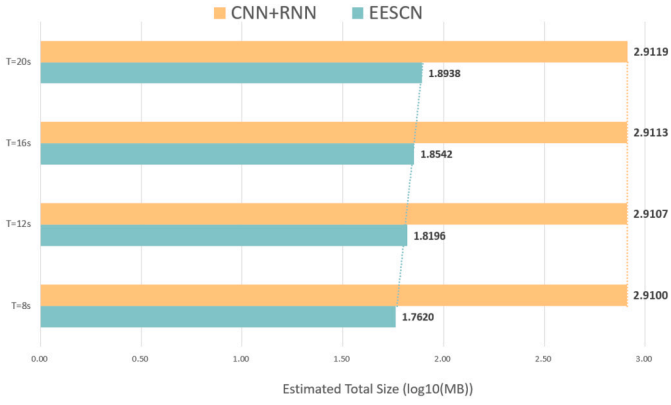

Fig. 7. Comparison of memory footprint for EESCN and CNN+RNN with different sliding window length T.

other within a longer sliding window. Notably, the EESCN saves about 50 times more memory footprint compared to CNN+RNN. This highlights the superior memory efficiency of EESCN, making it a more efficient choice for memory-constrained scenarios.

As mentioned in [53], SNN has the prospect of being applied to portable, low-power devices, which generally does not use high-power GPUs, so we chose to compare the running time of EESCN and CNN+RNN on the CPU. As shown in Fig. 8, it can be found that the running speed of EESCN has been greatly improved on the CPU compared with CNN+RNN, where the average improvement is more than 3 times and up to 4.3 times at T = 20 s. More importantly, as the T increases, the model size will certainly increase, which usually leads to a significant

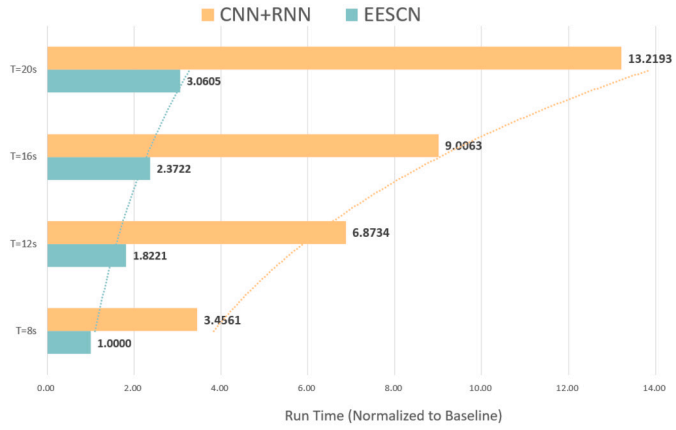

Fig. 8. Comparison of CPU running time for EESCN and CNN+RNN with different sliding window length T.

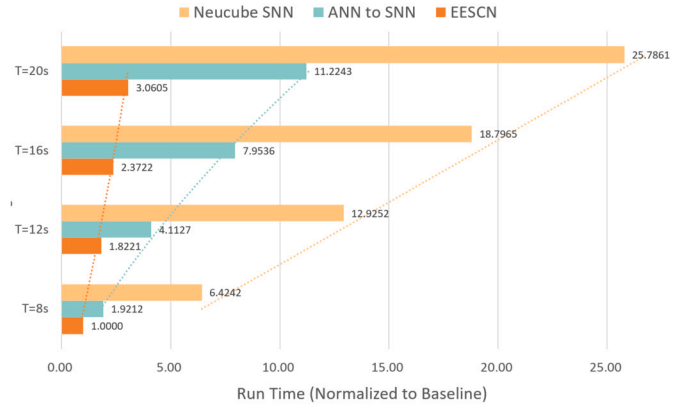

Fig. 9. Comparison of running time for different SNNs with different sliding window length T.

increase in the running time. However, the running time of EESCN does not increase too much and is less than that of CNN+RNN, which demonstrates the advantage of EESCN in terms of running efficiency, allowing it to be applied in a range of fields, such as the application scenarios with great restrictions on response time and memory consumption [59].

In terms of time comparison, we have added comparisons with two other SNNs under SEED-IV dataset to further demonstrate the efficiency of our proposed EESCN. First, we select the recently proposed ANN to SNN method [53] and added the same RNN structure as in the CNN+RNN method to ensure that the structure is consistent with EESCN, and the other selected SNN is the NeuCube SNN based on STDP [12]. All comparison uses the same dataset and running environment. The Fig. 9 shows the comparable running time performance at different sliding window lengths T.

As can be seen from Fig. 9, our proposed EESCN runs faster in classification tasks. Compared to the ANN to SNN method, the EESCN has improved 1.9 to 3.3 times. This is similar to the comparison experiment with CNN+RNN in our text, and the ANN to SNN method is slightly superior to CNN+RNN. Due to the pulse mechanism, floating-point operations are replaced by integer operations, which speeds up the calculation. This is consistent with past research that SNNs have been demonstrated to be more computationally efficient than ANNs both theoretically and in some practical applications [60]. Compared to the STDP based NeuCube SNN, its complex simulation process greatly reduces its running efficiency in classification tasks, making it run much slower than our proposed EESCN.

6. Conclusion

In this paper, we propose EESCN for EEG emotion recognition, which consists of a neuromorphic data generation module and a NeuroSpiking framework. The neuromorphic data generation module is a convenient input generation approach for SNNs. In addition, different from existing SNN methods, the proposed NeuroSpiking framework incorporates convolution modules with LIF neuron node sequences, allowing LIF neuron nodes to extract temporal features while efficiently extracting spatial features, thus improving the performance of EESCN. The experimental results on two public datasets demonstrate the effectiveness of our method. The further analysis of the impact of different sliding windows demonstrates the robustness of EESCN by setting different length T of sliding window, and the efficiency comparison between EESCN and CNN+RNN also shows the improvement of our method in terms of running efficiency and memory footprint. However, the EESCN is based on the conventional convolutional modules and fully-connected modules, which limits the biological plausibility compared to the STDP based SNNs. In addition, the classification capability of EESCN still need to be further improved. In our future work, we will attempt to transform the convolutional and fully-connected modules into structures that are more biologically plausible and enhance the classification capability of EESCN.

Declaration of competing interest

The authors declared no potential conflicts of interest with respect to the research, author-ship, and publication of this article.

Acknowledgement

This work was partially supported by the National Science Foundation of China under Grant 62076083, the National Key Research and Development Program of China under Grant 2022YFE0199300, the Hangzhou Artificial Intelligence Major Technological Innovation Project under Grant 2022AIZD0159, and the Key Laboratory of Brain Machine Collaborative Intelligence of Zhejiang Province under Grant No. 2020E10010.

References

[1] Peiyang Li, Huan Liu, Yajing Si, Cunbo Li, Fali Li, Xuyang Zhu, Xiaoye Huang, Ying Zeng, Dezhong Yao, Yangsong Zhang, et al., EEG based emotion recognition by combining functional connectivity network and local activations, IEEE Trans. Biomed. Eng. 66 (10) (2019) 2869–2881.

[2] Yu-Liang Hsu, Jeen-Shing Wang, Wei-Chun Chiang, Chien-Han Hung, Automatic ECG-based emotion recognition in music listening, IEEE Trans. Affect. Comput. 11 (1) (2017) 85–99.

[3] Sim Kuan Goh, Hussein A. Abbass, Kay Chen Tan, Abdullah Al-Mamun, Nitish Thakor, Anastasios Bezerianos, Junhua Li, Spatio–spectral representation learning for electroencephalographic gait-pattern classification, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (9) (2018) 1858–1867.

[4] Guowen Xiao, Meng Shi, Mengwen Ye, Bowen Xu, Zhendi Chen, Quansheng Ren, 4D attention-based neural network for EEG emotion recognition, Cogn. Neurodyn. (2022) 1–14.

[5] Tengfei Song, Wenming Zheng, Peng Song, Zhen Cui, EEG emotion recognition using dynamical graph convolutional neural networks, IEEE Trans. Affect. Comput. 11 (3) (2020) 532–541.

[6] Yang Li, Lei Wang, Wenming Zheng, Yuan Zong, Lei Qi, Zhen Cui, Tong Zhang, Tengfei Song, A novel bi-hemispheric discrepancy model for EEG emotion recognition, IEEE Trans. Cogn. Develop. Syst. 13 (2) (2020) 354–367.

[7] Yimin Yang, Q.M. Jonathan Wu, Wei-Long Zheng, Bao-Liang Lu, EEG-based emotion recognition using hierarchical network with subnetwork nodes, IEEE Trans. Cogn. Develop. Syst. 10 (2) (2017) 408–419.

[8] Salma Alhagry, Aly Aly Fahmy, Reda A. El-Khoribi, Emotion recognition based on EEG using LSTM recurrent neural network, Int. J. Adv. Comput. Sci. Appl. 8 (10) (2017).

[9] Zixing Zhang, Bingwen Wu, Björn Schuller, Attention-augmented end-to-end multi-task learning for emotion prediction from speech, in: ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2019, pp. 6705–6709.

[10] Germano R. Figueiredo, Wagner L. Ripka, Eduardo Félix Ribeiro Romaneli, Leandra Ulbricht, Attentional bias for emotional faces in depressed and non-depressed individuals: an eye-tracking study, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 5419–5422.

[11] Emre O. Neftci, Hesham Mostafa, Friedemann Zenke, Surrogate gradient learning in spiking neural networks: bringing the power of gradient-based optimization to spiking neural networks, IEEE Signal Process. Mag. 36 (6) (2019) 51–63.

[12] Yuling Luo, Qiang Fu, Juntao Xie, Yunbai Qin, Guopei Wu, Junxiu Liu, Frank Jiang, Yi Cao, Xuemei Ding, EEG-based emotion classification using spiking neural networks, IEEE Access 8 (2020) 46007–46016.

[13] Wei Fang, Zhaofei Yu, Yanqi Chen, Timothée Masquelier, Tiejun Huang, Yonghong Tian, Incorporating learnable membrane time constant to enhance learning of spiking neural networks, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 2661–2671.

[14] Samaneh Alsadat Saeedinia, Mohammad Reza Jahed-Motlagh, Abbas Tafakhori, Nikola Kasabov, Design of MRI structured spiking neural networks and learning algorithms for personalized modelling, analysis, and prediction of EEG signals, Sci. Rep. 11 (1) (2021) 12064.

[15] Georgios Ioannides, Ioannis Kourouklides, Alessandro Astolfi, Spatiotemporal dynamics in spiking recurrent neural networks using modified-full-force on EEG signals, Sci. Rep. 12 (1) (2022) 2896.

[16] Elisa Capecci, Zohreh Gholami Doborjeh, Nadia Mammone, Fabio La Foresta, Francesco C. Morabito, Nikola Kasabov, Longitudinal study of Alzheimer's disease degeneration through EEG data analysis with a NeuCube spiking neural network model, in: 2016 International Joint Conference on Neural Networks (IJCNN), IEEE, 2016, pp. 1360–1366.

[17] Pengbo Zhang, Xue Wang, Weihang Zhang, Junfeng Chen, Learning spatial–spectral–temporal EEG features with recurrent 3D convolutional neural networks for cross-task mental workload assessment, IEEE Trans. Neural Syst. Rehabil. Eng. 27 (1) (2018) 31–42.

[18] Alexander Craik, Yongtian He, Jose L. Contreras-Vidal, Deep learning for electroencephalogram (EEG) classification tasks: a review, J. Neural Eng. 16 (3) (2019) 031001.

[19] Mahima Milinda Alwis Weerasinghe, Josafath I. Espinosa-Ramos, Grace Y. Wang, Dave Parry, Incorporating structural plasticity approaches in spiking neural networks for EEG modelling, IEEE Access 9 (2021) 117338–117348.

[20] Simei Gomes Wysoski, Lubica Benuskova, Nikola Kasabov, Fast and adaptive network of spiking neurons for multi-view visual pattern recognition, Neurocomputing 71 (13–15) (2008) 2563–2575.

[21] Shirin Dora, K. Subramanian, S. Suresh, N. Sundararajan, Development of a self-regulating evolving spiking neural network for classification problem, Neurocomputing 171 (2016) 1216–1229.

[22] Yujie Wu, Rong Zhao, Jun Zhu, Feng Chen, Mingkun Xu, Guoqi Li, Sen Song, Lei Deng, Guanrui Wang, Hao Zheng, et al., Brain-inspired global-local learning incorporated with neuromorphic computing, Nat. Commun. 13 (1) (2022) 65.

[23] Fangyao Shen, Guojun Dai, Guang Lin, Jianhai Zhang, Wanzeng Kong, Hong Zeng, EEG-based emotion recognition using 4D convolutional recurrent neural network, Cogn. Neurodyn. 14 (2020) 815–828.

[24] Donald Olding Hebb, The Organization of Behavior: A Neuropsychological Theory, Psychology Press, 2005.

[25] Wael Alzhrani, Maryam Doborjeh, Zohreh Doborjeh, Nikola Kasabov, Emotion recognition and understanding using EEG data in a brain-inspired spiking neural network architecture, in: 2021 International Joint Conference on Neural Networks (IJCNN), IEEE, 2021, pp. 1–9.

[26] Nikola K. Kasabov, NeuCube: a spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data, Neural Netw. 52 (2014) 62–76.

[27] Abhronil Sengupta, Yuting Ye, Robert Wang, Chiao Liu, Kaushik Roy, Going deeper in spiking neural networks: VGG and residual architectures, Front. Neurosci. 13 (2019) 95.

[28] T. Yorozu, M. Hirano, K. Oka, Y. Tagawa, Electron spectroscopy studies on magneto-optical media and plastic substrate interface, IEEE Transl. J. Magn. Jpn. 2 (8) (1987) 740–741.

[29] Shikuang Deng, Shi Gu, Optimal conversion of conventional artificial neural networks to spiking neural networks, arXiv preprint, arXiv:2103.00476, 2021.

[30] William Severa, Craig M. Vineyard, Ryan Dellana, Stephen J. Verzi, James B. Aimone, Whetstone: a method for training deep artificial neural networks for binary communication, arXiv preprint, arXiv:1810.11521, 2018.

[31] Yongqiang Cao, Yang Chen, Deepak Khosla, Spiking deep convolutional neural networks for energy-efficient object recognition, Int. J. Comput. Vis. 113 (2015) 54–66.

[32] Peter U. Diehl, Daniel Neil, Jonathan Binas, Matthew Cook, Shih-Chii Liu, Michael Pfeiffer, Fast-classifying, high-accuracy spiking deep networks through weight and threshold balancing, in: 2015 International Joint Conference on Neural Networks (IJCNN), IEEE, 2015, pp. 1–8.

[33] Jianhao Ding, Zhaofei Yu, Yonghong Tian, Tiejun Huang, Optimal ANN-SNN conversion for fast and accurate inference in deep spiking neural networks, arXiv preprint, arXiv:2105.11654, 2021.

[34] Sander M. Bohte, Joost N. Kok, Han La Poutré, Error-backpropagation in temporally encoded networks of spiking neurons, Neurocomputing 48 (1) (2002) 17.

[35] Robert Gütig, Haim Sompolinsky, The tempotron: a neuron that learns spike timing–based decisions, Nat. Neurosci. 9 (3) (2006) 420–428.
[36] Filip Ponulak, Andrzej Kasiński, Supervised learning in spiking neural networks with resume: sequence learning, classification, and spike shifting, Neural Comput. 22 (2) (2010) 467–510.
[37] Ammar Mohemmed, Stefan Schliebs, Satoshi Matsuda, Nikola Kasabov, Span: spike pattern association neuron for learning spatio-temporal spike patterns, Int. J. Neural Syst. 22 (04) (2012) 1250012.
[38] Yujie Wu, Lei Deng, Guoqi Li, Jun Zhu, Luping Shi, Spatio-temporal backpropagation for training high-performance spiking neural networks, Front. Neurosci. 12 (2018) 331.
[39] Friedemann Zenke, Tim P. Vogels, The remarkable robustness of surrogate gradient learning for instilling complex function in spiking neural networks, Neural Comput. 33 (4) (2021) 899–925.
[40] Javier M. Antelis, Luis E. Falcón, et al., Spiking neural networks applied to the classification of motor tasks in EEG signals, Neural Netw. 122 (2020) 130–143.
[41] Aniket Singh Rajpoot, Mahesh Raveendranatha Panicker, et al., Subject independent emotion recognition using EEG signals employing attention driven neural networks, Biomed. Signal Process. Control 75 (2022) 103547.
[42] Sander Koelstra, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, Ioannis Patras, Deap: a database for emotion analysis; using physiological signals, IEEE Trans. Affect. Comput. 3 (1) (2011) 18–31.
[43] Wei-Long Zheng, Wei Liu, Yifei Lu, Bao-Liang Lu, Andrzej Cichocki, Emotionmeter: a multimodal framework for recognizing human emotions, IEEE Trans. Cybern. 49 (3) (2018) 1110–1122.
[44] Clarence Tan, Marko Šarlija, Nikola Kasabov, Neurosense: short-term emotion recognition and understanding based on spiking neural network modelling of spatio-temporal EEG patterns, Neurocomputing 434 (2021) 137–148.
[45] Abeer Al-Nafjan, Khulud Alharthi, Heba Kurdi, Lightweight building of an electroencephalogram-based emotion detection system, Brain Sci. 10 (11) (2020) 781.
[46] Yang Li, Wenming Zheng, Yuan Zong, Zhen Cui, Tong Zhang, Xiaoyan Zhou, A bi-hemisphere domain adversarial neural network model for EEG emotion recognition, IEEE Trans. Affect. Comput. 12 (2) (2018) 494–504.
[47] Eddie Harmon-Jones, Cindy Harmon-Jones, Elizabeth Summerell, On the importance of both dimensional and discrete models of emotion, Behav. Sci. 7 (4) (2017) 66.
[48] Jaak Panksepp, The Foundations of Human and Animal Emotions, 1998.
[49] Minchao Wu, Wei Teng, Cunhang Fan, Shengbing Pei, Ping Li, Zhao Lv, An investigation of olfactory-enhanced video on EEG-based emotion recognition, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 1602–1613.
[50] Wei Tao, Chang Li, Rencheng Song, Juan Cheng, Yu Liu, Feng Wan, Xun Chen, EEG-based emotion recognition via channel-wise attention and self attention, IEEE Trans. Affect. Comput. (2020).
[51] Peixiang Zhong, Di Wang, Chunyan Miao, EEG-based emotion recognition using regularized graph neural networks, IEEE Trans. Affect. Comput. 13 (3) (2020) 1290–1301.
[52] Fatemeh Zareayan Jahromy, Atena Bajoulvand, Mohammad Reza Daliri, Statistical algorithms for emotion classification via functional connectivity, J. Integr. Neurosci. 18 (3) (2019) 293–297.
[53] Zhanglu Yan, Jun Zhou, Weng-Fai Wong, EEG classification with spiking neural network: smaller, better, more energy efficient, Smart Health 24 (2022) 100261.
[54] Dongdong Li, Li Xie, Zhe Wang, Hai Yang, Brain emotion perception inspired EEG emotion recognition with deep reinforcement learning, IEEE Trans. Neural Netw. Learn. Syst. (2023) 1–14.
[55] Li Wei, Cheng Fang, Zhihao Zhu, Chuyi Chen, Aiguo Song, Fractal spiking neural network scheme for EEG-based emotion recognition, IEEE J. Transl. Eng. Health Med. (2023) 1.
[56] Haowen Fang, Amar Shrestha, Ziyi Zhao, Qinru Qiu, Exploiting neuron and synapse filter dynamics in spatial temporal learning of deep spiking neural network, arXiv preprint, arXiv:2003.02944, 2020.
[57] Rachmad Vidya Wicaksana Putra, Muhammad Shafique, FSpiNN: an optimization framework for memory-efficient and energy-efficient spiking neural networks, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. 39 (11) (2020) 3601–3613.
[58] Wei-Long Zheng, Bao-Liang Lu, Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks, IEEE Trans. Auton. Ment. Dev. 7 (3) (2015) 162–175.
[59] Matthijs Pals, Rafael J. Pérez Belizón, Nicolas Berberich, Stefan K. Ehrlich, John Nassour, Gordon Cheng, Demonstrating the viability of mapping deep learning based EEG decoders to spiking networks on low-powered neuromorphic chips, in: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2021, pp. 6102–6105.
[60] Yujie Wu, Lei Deng, Guoqi Li, Jun Zhu, Yuan Xie, Luping Shi, Direct training for spiking neural networks: faster, larger, better, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 1311–1318.