2025 4th International Conference on Computer, Electronic and Materials Engineering (ICCEME 2025)

Motor Imagery Recognition Based on Spiking Neural Networks

Dawei Yan

The Affiliated High School of Peking University, Beijing, China

yandawei2023@163.com

Keywords: Motor Imagery, Brain–Computer Interface, Electroencephalography, Spiking Neural Networks

Abstract: Motor Imagery (MI) is a key research direction in brain–computer interface (BCI) systems, with widespread applications in neural rehabilitation and brain-controlled prosthetics. To address the challenges of low signal-to-noise ratio and strong temporal dependencies in electroencephalography (EEG) signals, this paper proposes an MI classification method based on spiking neural networks (SNN). By leveraging temporal encoding and neural dynamic modeling, the proposed approach effectively captures time-dependent features in EEG data. Experiments conducted on the BCI Competition IV 2008 dataset demonstrate that the method achieves an accuracy of 79.6%, representing a 5.2% improvement over conventional convolutional neural networks (CNNs), thereby validating the effectiveness and advantages of SNNs for MI classification tasks.

1. Introduction

Brain–Computer Interface (BCI) is a technology that enables direct interaction between the human brain and external devices by interpreting electroencephalogram (EEG) signals (Figure 1). It has been widely applied in neural rehabilitation, intelligent prosthetic control, and assistive communication systems. Among various paradigms, Motor Imagery (MI) is one of the most commonly used BCI applications, where users can control external devices by imagining the movement of their left hand, right hand, or limbs, offering a novel communication channel for individuals with motor impairments (Figure 2). The core challenge of MI-based BCI lies in accurately decoding the user's intention from EEG signals to achieve efficient and real-time human–machine interaction[1].

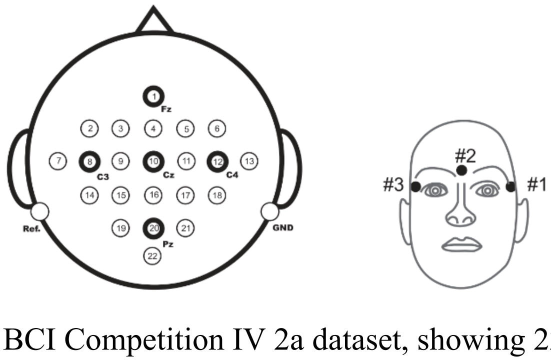

Figure 1 Electrode layout of BCI Competition IV 2a dataset, showing 22 EEG channels (left) and 3 EOG channels (right).

However, EEG signals are inherently noisy, exhibit strong temporal dependencies, and vary significantly across subjects, posing substantial challenges to MI classification tasks. Traditional approaches rely heavily on handcrafted features and conventional machine learning classifiers. For instance, Herman et al. proposed a method based on wavelet transform and spectral analysis, extracting frequency-domain features for classification using support vector machines (SVM) or linear discriminant analysis (LDA). While these methods can be effective to some extent, they often require extensive feature engineering, lack robustness across subjects, and struggle to perform well in complex tasks[2].

With the emergence of deep learning, models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been increasingly applied to MI classification, enhancing the

  DOI: 10.25236/icceme.2025.014

automation of feature extraction[3]. Nevertheless, these methods often suffer from high computational complexity and are still limited in modeling fine-grained temporal dependencies in EEG data, making it difficult to balance accuracy and efficiency.

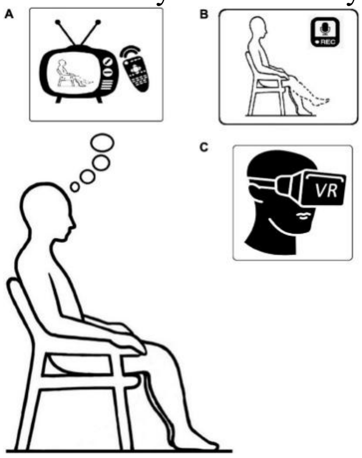

Figure 2 Illustration of different motor imagery paradigms: (A) passive visual guidance, (B) kinesthetic imagination with audio cues, and (C) immersive VR-based stimulation.

To address these challenges, this paper proposes a novel MI classification method based on spiking neural networks (SNNs). Leveraging biologically inspired neural dynamics and temporal encoding mechanisms, the proposed approach effectively captures temporal features in EEG signals while offering greater computational efficiency and interpretability. In this framework, continuous EEG signals are encoded into spike trains via a spike generation mechanism, and a gated neuronal dynamic regulation strategy is employed to enhance the model's sensitivity to subtle MI-related patterns. Experiments conducted on the BCI Competition IV 2008 Dataset 2a demonstrate that the proposed SNN-based method outperforms conventional CNN models in both classification accuracy and cross-subject generalization, validating its effectiveness and practical value in MI-based BCI applications.

2. Manuscript

2.1. Spiking Neural Networks

To address the challenges of temporal modeling and cross-subject generalization in motor imagery (MI) EEG signal classification, this paper proposes a classification method based on spiking neural networks (SNNs). As Figure 3 shows, the proposed approach integrates temporal encoding strategies with event-driven neural modeling mechanisms, enhancing the ability to capture temporal dependencies in EEG signals while maintaining low computational overhead. The overall framework comprises EEG signal preprocessing and temporal encoding, spiking neuron modeling, network architecture design, and end-to-end training strategy.

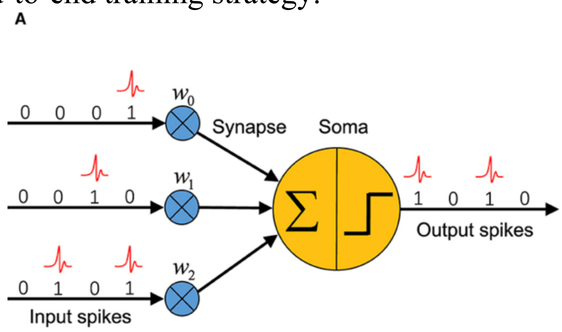

Figure 3 Schematic of a spiking neuron model: weighted input spikes are integrated at the soma, and output spikes are generated when the membrane potential exceeds the firing threshold.

We constructed experimental samples based on the BCI Competition IV 2008 Dataset 2a, which contains 22-channel EEG recordings associated with four types of motor imagery tasks: left hand, right hand, foot, and tongue movements. The EEG signals were recorded at a sampling rate of 250 Hz. All signals were first processed using a band-pass filter (4–38 Hz) to remove irrelevant frequency components, followed by z-score normalization applied to each channel. Subsequently, a sliding window of 2 seconds with a 0.5-second stride was applied to segment the signals, resulting in a balanced set of samples. To meet the input requirements of spiking neural networks, we adopted the Time-to-First-Spike (TTFS) encoding strategy, which converts continuous-valued EEG signals into sparse spike trains. For each channel, the signal value x_$x_i \in [0,1]$ was mapped to a spike firing time defined as:

$$t_i = T_{max} \cdot (1 - x_i) \tag{1}$$

Here, T_{max} denotes the duration of the temporal encoding window; a larger signal value results in an earlier spike firing time. As a result, each sample is represented as a spike-based temporal matrix of shape [C, T], where C denotes the number of EEG channels and T represents the number of discrete time steps.

In terms of network architecture design, we constructed a temporal spiking model based on Leaky Integrate-and-Fire (LIF) neurons. Each neuron maintains a membrane potential u_$u_i(t)$in the discrete time domain, which evolves over time according to the following update rule:

$$u_i(t) = \lambda \cdot u_i(t-1) + I_i(t), \text{if } u_i(t) < \theta \tag{2}$$

Here, $\lambda \in (0,1)$ denotes the membrane potential decay factor, $I_i(t)$ represents the synaptic input current at time step t, and θ is the fixed firing threshold. When the membrane potential $u_i(t)$ exceeds the threshold (i.e., $u_i(t) \geq \theta$), the neuron emits a spike and immediately resets its membrane potential to zero. Event propagation in the network is implemented through a hierarchical convolutional structure, where the spiking convolutional layers perform one-dimensional convolutions along the temporal axis to extract local spatiotemporal spike patterns. These are followed by spike pooling operations within defined temporal windows to reduce dimensionality and retain salient firing activity. Finally, the output layer applies a softmax function over the spike counts of each output neuron to produce probability distributions for the four-class motor imagery classification task.

Since the spiking function of spiking neurons is non-differentiable, traditional backpropagation algorithms cannot be directly applied. To address this issue, we adopt the surrogate gradient technique, which preserves the discrete spiking mechanism during the forward pass while employing a differentiable approximation to the gradient during the backward pass. Let $S(u)$ denote the spiking function of a neuron; during training, its derivative is approximated by a surrogate function as follows:

$$\frac{\partial S(u)}{\partial u} \approx \alpha \cdot max\left(0,1 - \left|\frac{u-\theta}{\beta}\right|\right) \tag{3}$$

Here, α is the gradient scaling factor, and β controls the range of the approximation. This strategy enables the spiking neural network to be trained in an end-to-end manner.

2.2. Experimental Results and Analysis

2.2.1. Dataset and Experimental Setup

To evaluate the effectiveness of the proposed motor imagery classification method based on spiking neural networks (SNNs), we conducted extensive experiments on the publicly available BCI Competition IV 2008 Dataset 2a. This dataset, released by Graz University of Technology in Austria, contains EEG recordings from nine healthy subjects. Each subject participated in two recording sessions, with each session consisting of 288 trials covering four types of motor imagery tasks: left hand, right hand, foot, and tongue movements. EEG signals were recorded using 22 electrode channels at a sampling rate of 250 Hz, and three additional EOG channels were simultaneously collected to facilitate ocular artifact correction. Each trial lasted for four seconds, with a cue presented at the 2-second mark to prompt the subject. The interval from 2 to 4 seconds was designated as the actual motor imagery execution phase. To ensure data quality, we only extracted the 2-second EEG segments following the cue for training and evaluation.

During preprocessing, all EEG signals were band-pass filtered within the range of 4–38 Hz to retain the μ (8–13 Hz) and β (13–30 Hz) rhythms that are closely associated with motor imagery. Subsequently, z-score normalization was applied to each channel. We employed a fixed 2-second sliding window to segment the signals and applied spike-based temporal encoding to each segment, resulting in sparse event-driven representations. To enhance the model's generalization ability across subjects, we introduced data augmentation techniques during training, including random channel dropout and temporal jittering.

For model training, mini-batch learning was used with a batch size of 64 and an initial learning rate of 0.001. The model was trained for 100 epochs, with parameters updated iteratively over the training set in each epoch. To improve convergence and stability, we adopted a cosine annealing schedule for dynamic learning rate adjustment, defined as follows:

$$\eta_t = \eta_0 \cdot \frac{1}{2}\left(1 + \cos\left(\frac{t}{T}\pi\right)\right) \tag{4}$$

Here, η_0 denotes the initial learning rate, T is the total number of training epochs, and t is the current epoch index. This strategy maintains a relatively high learning rate in the early stages to facilitate rapid convergence and gradually reduces it in later stages for fine-grained optimization of model parameters.

In terms of the loss function, considering that the motor imagery (MI) task is a multi-class classification problem, we adopt the standard cross-entropy loss, which is defined as follows:

$$\mathcal{L}_{CE} = \sum_{i=0}^{C} y_i \log \hat{y}_i \tag{5}$$

To evaluate the performance of the proposed method in motor imagery (MI) classification tasks, we employed the following evaluation metrics: Accuracy, Precision, Recall, and Macro-averaged F1 score (Macro-F1). Accuracy reflects the overall classification correctness, while precision and recall measure the model's ability to identify each class accurately and completely, respectively. The Macro-F1 score provides a balanced evaluation across all classes, making it especially suitable for assessing performance stability and generalization in multi-class classification scenarios. All metrics are computed on the test set, and evaluation is conducted independently for each subject. The final reported results represent the average performance across all subjects.

2.2.2. Experimental Results

To validate the effectiveness of our proposed spiking neural network (SNN) model for motor imagery classification, we compared it against four representative neural network baselines: CNN, RNN (LSTM), EEGNet, and Transformer[4]. These baselines span from traditional deep learning architectures to EEG-specific lightweight models, as well as advanced attention-based frameworks, enabling a comprehensive and fair comparison from multiple modeling perspectives.

CNNs are widely used in deep learning and are effective at extracting spatially invariant features from local regions. Our implemented CNN baseline consists of two 1D convolutional layers with 32 and 64 filters, respectively, each with a kernel size of 3. Batch normalization and max pooling follow each convolutional layer. The CNN extracts local temporal patterns from EEG signals through sliding convolution kernels, and classification is performed via two fully connected layers. While CNNs offer strong feature extraction capabilities, their fixed receptive fields limit their ability to capture long-range temporal dependencies, which is critical for modeling EEG signals in MI tasks.

Recurrent neural networks, particularly Long Short-Term Memory (LSTM) networks, are widely used for modeling sequential data[5]. We employ a two-layer LSTM baseline, with each layer comprising 64 hidden units, followed by a fully connected layer for four-class classification[6]. LSTM captures temporal dynamics by propagating hidden states through time, enabling it to model EEG signal evolution. However, RNNs often suffer from vanishing/exploding gradients, have lower training efficiency, and provide limited interpretability regarding which temporal segments influence the prediction.

EEGNet is a lightweight deep neural network architecture specifically designed for EEG-based classification[7]. It integrates depthwise separable convolutions with spatial filters, simulating Common Spatial Pattern (CSP) behavior in the initial layer, followed by pointwise convolutions to extract temporal features. EEGNet has a compact structure and low parameter count, making it well-suited for training on small-scale EEG datasets. It has achieved competitive results across various EEG-based BCI paradigms such as MI, ERP, and SSVEP. Due to its alignment with EEG signal characteristics, EEGNet is widely regarded as a strong baseline in BCI applications.

Transformer is a modern sequence modeling architecture built entirely on self-attention mechanisms, capable of capturing global dependencies between time steps. In this study, we implement a simplified Transformer model consisting of two encoder blocks, each with four attention heads and a hidden dimension of 128. Positional encodings are added to the EEG input sequences to preserve temporal order. Although Transformer has strong modeling capacity, its large number of parameters and reliance on extensive training data make it prone to overfitting in limited-data EEG scenarios. Moreover, its use of continuous activations lacks the event-driven sparsity inherent to SNNs, leading to reduced inference efficiency and interpretability.

Our proposed model is based on a spiking neural network architecture. It utilizes temporal encoding to convert continuous EEG signals into spike trains and employs Leaky Integrate-and-Fire (LIF) neurons to model membrane potential dynamics over time, forming a biologically inspired event-driven framework for temporal modeling. By incorporating surrogate gradient optimization, the model can be trained in a fully differentiable manner, enabling it to effectively capture fine-grained temporal dependencies in MI tasks. Compared to other baselines, SNN offers advantages in sparse computation, temporal sensitivity, and energy efficiency, making it particularly well-suited for deployment in resource-constrained embedded BCI systems.

To comprehensively assess the performance of the proposed SNN-based motor imagery classification method, we conducted comparative experiments with the aforementioned baseline models. The evaluation metrics include accuracy and macro-averaged F1 score. All experiments were conducted on the BCI Competition IV 2008 Dataset 2a. Table 1 summarizes the average classification performance of each method, based on test data collected across all subjects.

Table 1 Experimental results of each method.

Model	ACC (%)	F1 (%)
CNN	73.4	70.2
RNN (LSTM)	74.1	71.3
EEGNET	75.9	73.0
Transformer	76.8	74.5
Ours(SNN)	**79.6**	**77.1**

The results demonstrate that the proposed SNN method outperforms all baseline models across both evaluation metrics, indicating its superior capability in modeling the temporal dynamics inherent in EEG signals. Compared with traditional convolutional approaches such as CNN, the event-driven mechanism and temporal encoding of the SNN significantly enhance its responsiveness to local, task-relevant temporal features. Unlike RNNs, which rely on hidden state propagation for temporal modeling, the SNN explicitly encodes temporal structures through spike-based transmission, thereby mitigating issues such as vanishing gradients. Furthermore, compared with the strong EEG-specific baseline EEGNet, the proposed SNN achieves better performance while maintaining a lightweight architecture, validating the effectiveness and generalizability of biologically inspired spiking models in motor imagery recognition.

It is also noteworthy that the Transformer model performs reasonably well on this task; however, its performance gains are limited by its large parameter size and high data requirements. In contrast, the proposed SNN leverages sparse event-driven computation to achieve higher accuracy and greater stability during inference, making it especially suitable for energy-constrained and real-time embedded BCI systems.

To further investigate the contribution of each key component within the proposed model, we conducted a series of ablation studies. Specifically, we evaluated performance under the following modifications: (1) removing the temporal encoding module and directly inputting normalized analog signals; (2) replacing LIF neurons with standard ReLU activations, effectively converting the network into a conventional continuous-valued model; and (3) removing the surrogate gradient mechanism, training the model without gradient-based optimization and relying solely on spike statistics. The results of these ablation experiments are presented in Table 2.

Table 2 Experimental results of each method.

Model Variant	Accuracy (%)	Macro F1 (%)
Full Model (SNN + Temporal Encoding + SG)	79.6	77.1
w/o Temporal Encoding (use raw analog input)	75.3	72.5
Replace LIF with ReLU (non-spiking network)	74.8	71.9
w/o Surrogate Gradient (non-trainable spikes)	72.6	69.2
Model Variant	Accuracy (%)	Macro F1 (%)

3. Conclusion

The paper proposes an MI classification method based on SNNs for EEG signals. Experiments on BCI Competition IV 2008 Dataset 2a show the method achieves 79.6% accuracy, outperforming CNN, RNN, EEGNet and Transformer. Ablation studies confirm the importance of temporal encoding, LIF neurons and surrogate gradient.

References

[1] Altaheri, H., Muhammad, G., Alsulaiman, M., Amin, S.U., Altuwaijri, G.A., Abdul, W., Bencherif, M.A., and Faisal, M.: ‘Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: a review’, Neural Computing and Applications, 2023, 35, (20), pp. 14681-14722

[2] Lu, N., Li, T., Ren, X., and Miao, H.: ‘A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines’, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2017, 25, (6), pp. 566-576

[3] Ju, C., and Guan, C.: ‘Tensor-CSPNet: A Novel Geometric Deep Learning Framework for Motor Imagery Classification’, IEEE Transactions on Neural Networks and Learning Systems, 2023, 34, (12), pp. 10955-10969

[4] Zhu, H., Forenzo, D., and He, B.: ‘On the Deep Learning Models for EEG-Based Brain-Computer Interface Using Motor Imagery’, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022, 30, pp. 2283-2291

[5] Khademi, Z., Ebrahimi, F., and Kordy, H.M.: ‘A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals’, Computers in Biology and Medicine, 2022, 143, pp. 105288

[6] Zhang, R., Zong, Q., Dou, L., and Zhao, X.: ‘A novel hybrid deep learning scheme for four-class motor imagery classification’, Journal of Neural Engineering, 2019, 16, (6), pp. 066004

[7] Pérez-Velasco, S., Santamaría-Vázquez, E., Martínez-Cagigal, V., Marcos-Martínez, D., and Hornero, R.: ‘EEGSym: Overcoming Inter-Subject Variability in Motor Imagery Based BCIs With Deep Learning’, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022, 30, pp. 1766-1775