

A Brain-Computer Interface Four-class Classification Algorithm Integrating a Custom Spiking Neural Network with Attention Mechanisms

Ying-Jie Jia, Tian-Wei Shi*, Ling Ren, Wen-Hua Cui

Abstract—This paper addresses the challenges of feature extraction and classification accuracy in brain-computer interface (BCI) systems based on motor imagery tasks. We propose the SAMPN-Network model, which integrates a custom spiking neural network with a soft attention mechanism (SoftAttention-Layer), alongside the Simplicityformer classifier that incorporates a multi-head attention mechanism (MHA). The resulting classification algorithm, named Custom Spiking Neural Network Layer with Soft Attention Mechanism and Multi-Head Attention for Classification (SNA-MHC), is specifically designed to optimize classification accuracy in BCI systems. In our approach, raw EEG signals corresponding to motor imagery (MI) tasks are first normalized and then transformed into discrete spike trains using threshold encoding to make them suitable for processing by Spiking Neural Networks (SNN). These spike signals are subsequently processed by the SAMPN-Network model, which performs feature extraction by integrating a soft attention mechanism with the SNN module. The SNN module utilizes pulse neurons to encode and enhance the temporal information in EEG signals. Concurrently, the soft attention mechanism calculates attention weights to automatically focus on critical segments of the EEG signals associated with MI tasks while suppressing background noise and irrelevant temporal information, thereby extracting more precise time-series features. Following time-sequence feature extraction, a Multi-Head Attention Mechanism performs parallel attention computation across time domain, frequency domain, and more abstract feature spaces. This approach captures interdependencies between features across different dimensions and enhances the discriminative power of the classifier. Finally, the integrated features are processed by a Softmax classifier to perform four-class classification of MI tasks. Experimental results demonstrate that the proposed SNA-MHC model outperforms existing state-of-the-art models in terms of classification accuracy on both the TechBrain and BCI Competition IV2a datasets. Specifically, SNA-MHC achieves an average classification accuracy improvement of 13.02%, 4.41%, 8.46%,

15.05%, and 15.88%, respectively, when compared to other algorithmic models. Furthermore, when compared to traditional CNN and SNN models, SNA-MHC exhibits superior energy efficiency while maintaining classification accuracy, further validating its robust performance.

Index Terms—Brain-Computer Interface (BCI), Motor Imagery (MI), Spiking Neural Network (SNN), Soft Attention Mechanism, Multi-Head Attention Mechanism, EEG Signal Classification

I. INTRODUCTION

BRAIN-Computer Interface (BCI) is a technology that enables communication between the brain and external devices by directly capturing and interpreting electrical signals (EEG), and has garnered significant attention in recent years [1]. BCI systems offer innovative solutions across various domains, including neurological rehabilitation, prosthetic control, and cognitive training. These systems have demonstrated significant utility in assisting patients with stroke or spinal cord injuries in rehabilitation and motor function restoration [2], enabling accurate control of prosthetics for complex tasks such as grasping and movement [3], and enhancing cognitive abilities through brainwave feedback systems in patients with attention deficits and cognitive decline.

Within the BCI domain, Motor Imagery (MI) has emerged as a prominent research area [4]. MI tasks require subjects to imagine specific movements without physical execution, thereby activating the brain's motor cortex and generating characteristic electroencephalogram (EEG) signals. This approach has been widely implemented in neural rehabilitation [5], prosthetic limb control [6], brain-controlled wheelchairs [7], drone navigation, and various other applications [8].

Despite significant advancements in MI-based BCI systems, feature extraction and classification remain fundamental challenges. Traditional feature extraction methods exhibit inherent limitations: wavelet transform, while capable of analyzing signals at different time scales to extract features across frequency ranges [9] is susceptible to noise interference, particularly when processing high-frequency components of low signal-to-noise ratio signals. Fourier transform effectively analyzes frequency components but erroneously assumes EEG signals are stationary, thereby failing to capture the non-stationary characteristics inherent in MI tasks [10]. Additionally, spectral entropy analysis and autoregressive (AR) models, though suitable for specific feature extraction [11], demonstrate insufficient feature expressivity when confronted with complex temporal signals.

Manuscript received January 31, 2025; revised May 19, 2025.

This work was supported by the Liaoning Province Science and Technology Major Project (2022JH1/10400005), Liaoning Province Department of Education university basic research funds and the Liaoning Province Intelligent Construction IoT Application technology Key laboratory open project (2024KFKT-08) and the National Natural Science Foundation of China (U1908218).

YingJie Jia is a graduate student at the School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan 114051, China (e-mail: 2676457511@qq.com).

TianWei Shi* is a Professor at the School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan 114051, China (Corresponding author to provide phone: +086-139-9805-3962; e-mail: tianweiabcc@163.com).

Ling Ren is a Lecturer at the School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan 114051, China (e-mail: renlingqianzhihe@163.com).

WenHua Cui is a Professor at the School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan 114051, China (e-mail: 176878392@qq.com).

To overcome these limitations, deep learning techniques have been integrated into BCI systems. Convolutional Neural Networks (CNNs) have shown promising results in EEG signal classification through their ability to learn high-level features [12]. However, CNNs exhibit limitations when processing EEG signals characterized by low signal-to-noise ratios and strong temporal dependencies. Long Short-Term Memory (LSTM) networks can capture extended temporal dependencies in time series data, making them more suitable for EEG signals with pronounced dynamic characteristics [13]. Nevertheless, their high computational complexity constrains their application in real-time BCI systems.

In the classification phase, traditional algorithms display various constraints. Linear Discriminant Analysis (LDA) performs well in low-dimensional, linearly separable feature spaces but struggles with highly nonlinear distributions. Support Vector Machines (SVM) face similar limitations with linear kernels, though their capacity for nonlinear pattern classification can be enhanced through nonlinear kernel implementations [14]. Decision tree algorithms encounter difficulties managing the strong temporal dependencies in EEG signals, adversely affecting classification performance [15].

Recent years have witnessed the emergence of Spiking Neural Networks (SNN) as a significant research focus in EEG signal processing due to their unique event-driven operational mechanism [16]. By emulating neuronal pulse discharge behavior, SNNs effectively capture the sparse and non-stationary characteristics of EEG signals, thereby enhancing temporal feature representation. However, SNNs continue to lag behind CNNs and LSTMs in high-precision classification tasks, primarily due to limitations in optimization methodologies and hardware support.

To address these challenges, researchers have proposed various enhancement strategies. Kheradpisheh et al. (2021) combined the spatial feature extraction capabilities of CNNs with the temporal processing strengths of SNNs, significantly improving EEG signal classification performance [17]. Zheng et al. (2022) explored the integration of SNN and LSTM architectures to further enhance classification accuracy by effectively capturing both long-term dependencies and sparse temporal features [18]. Additionally, Yao et al. (2023) proposed an architecture integrating SNN with a self-attention mechanism, substantially improving the model's ability to capture critical information in MI tasks [19].

While these advancements have yielded performance improvements in SNNs, achieving efficient EEG signal classification on low-power devices remains a significant challenge, particularly in environments with constrained computational resources and stringent power consumption requirements. Such contexts necessitate more efficient algorithms or hardware acceleration techniques that maintain classification accuracy while minimizing energy consumption.

To address these issues, we present a novel approach that combines a custom spiking neural network with a soft attention mechanism (SoftAttentionLayer). Additionally, we introduce SNA-MHC, a four-class classification model for BCI systems (Multi-CustomSpikeLayer with Soft Attention Mechanism and Multi-Head Attention). This model is specifically designed to

enhance the classification accuracy of EEG signals in motor imagery tasks.

The SAMPN-Network feature extraction model process is shown in Figure 5. SNA-MHC preprocesses the original EEG signals using the custom spiking neural network and applies threshold encoding to convert continuous EEG signals into discrete pulse signals, aligning with the processing requirements of SNN. This process effectively preserves the temporal characteristics of EEG signals while minimizing noise interference. In the feature extraction phase, SNA-MHC incorporates the SAMPN-network module, which integrates the SoftAttentionLayer and SNN. Leveraging the biological properties of spiking neurons, the SNN module encodes the timing information of EEG signals through pulse activation, enabling the effective capture of temporal features. The soft attention mechanism automatically prioritizes the critical signals related to the motor imagery task and suppresses irrelevant or noisy signals, thereby significantly enhancing the accuracy of time-sequence feature extraction. Furthermore, SNA-MHC employs the Multi-Head Attention Mechanism (MHA) to capture relationships between different dimensions by performing parallel computations across multiple feature spaces, including time and frequency domains. This approach enhances the model's global understanding of the data. MHA effectively models complex dependencies between feature dimensions by processing multiple attention heads in parallel. Simultaneously, it extracts crucial temporal information from diverse feature spaces, strengthens cross-dimensional feature representation capabilities, and provides comprehensive feature representation for classification tasks, substantially improving classification accuracy.

Finally, the integrated features processed by MHA are input to a Softmax classifier to perform four-class classification for the motor imagery task. Compared to traditional SNN or CNN models, our proposed SNA-MHC significantly reduces computational complexity while maintaining high classification accuracy.

II. METHOD

This paper proposes the SAMPN-network model, which integrates a custom spiking neural network and a soft attention mechanism (SoftAttentionLayer) for feature extraction, and a Simplicityformer classifier incorporating a multi-head attention mechanism (MHA). The SNA-MHC classification algorithm, based on the brain-computer interface (BCI) system, can efficiently process real-time EEG signals and perform accurate classification.

Initially, the pre-processed real-time EEG signals were transformed into discrete pulse sequences through threshold encoding. This encoding method effectively converts complex continuous signals into discrete pulse sequences, making them suitable for processing by spiking neural networks (SNN). Subsequently, the discrete pulse sequences were input to the SAMPN-network model, which includes an SNN module, where they were processed into standard pulse signals using leaky integrate-and-fire (LIF) neurons.

To fully exploit the key features of EEG signals, a soft attention mechanism is integrated into the SAMPN-network

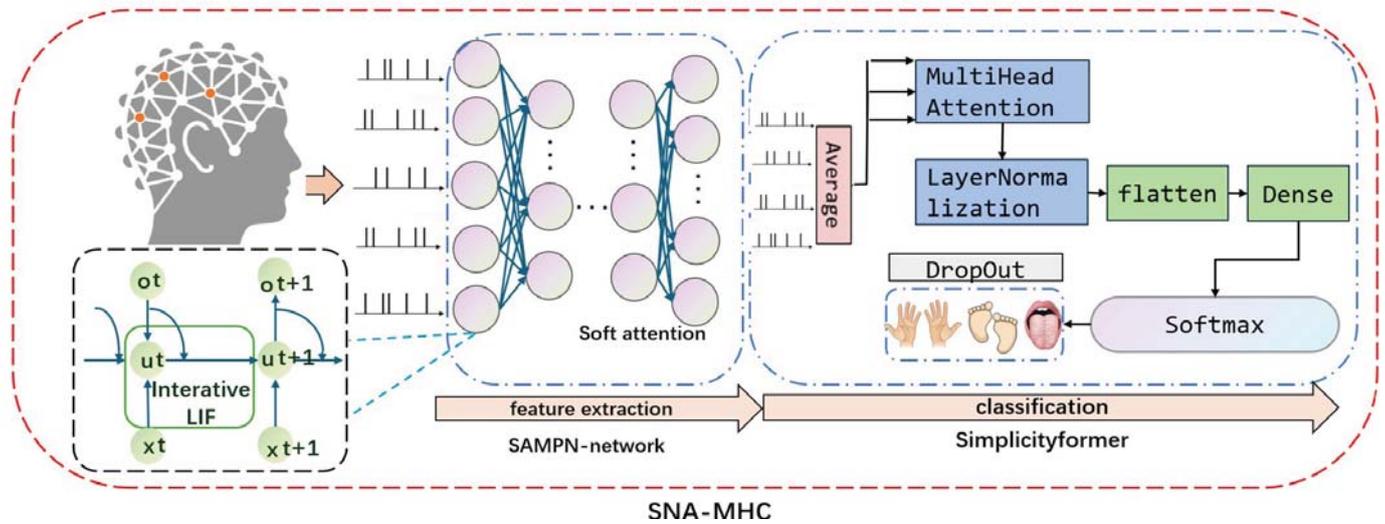


Fig. 1: Structure diagram of SNA-MHC model

module. This mechanism enables the extraction of multi-dimensional features by applying a weighted strategy that highlights important information, suppresses redundant features, and enhances the quality of feature representation. Finally, the extracted features were fed into the Simplicityformer classifier. The multi-head attention mechanism independently analyzes the input information by computing attention distributions for each attention head. The final classification is achieved through the Softmax layer after normalization and fully connected layer processing. The complete architecture of the SNA-MHC model is shown in Figure 1.

A. Introduction to Data sets

BCI Competition IV 2a: The BCI Competition IV 2a dataset consists of electroencephalographic (EEG) signals recorded from 22 healthy subjects, each performing four distinct motor imagery (MI) tasks: left hand, right hand, foot, and tongue. Data acquisition was carried out using 64 EEG channels with a sampling rate of 250 Hz. Each MI task lasted for 4 seconds, during which the EEG signals were recorded and labeled with their respective task categories. A total of 2 minutes of EEG data were collected per subject, incorporating multiple repetitions of each motor imagery task. The dataset provides a rich source for investigating brain-computer interface (BCI) applications related to motor imagery.

TechBrain Dataset: The TechBrain dataset is a laboratory-based electroencephalographic (EEG) dataset comprising data from 20 subjects, with a mean age of 23.4 ± 1.2 years (subjects 1-20). The dataset includes four motor imagery tasks: hand grip (task 1), hand open (task 2), wrist flexion (task 3), and wrist extension (task 4). All participants were recruited from Liaoning University of Science and Technology and had no history of neurological disorders. During the experiment, none of the subjects were under the influence of any medications. The study was conducted in strict accordance with the ethical guidelines of the Declaration of Helsinki, and all participants provided written informed consent prior to participation.

During the experiment, subjects wore Neuroscan (NuAmps) electrode caps and sat comfortably in an armchair, relaxing for 3 minutes. They were instructed to maintain a distance of 50 cm from a 21-inch LCD monitor. EEG data was recorded using 40-channel Ag/AgCl electrodes (FP1, FP2, F7, F3, FZ, F4, F8, FT7, FC3, FCZ, FC4, FT8, T3, C3, CZ, C4, T4, TP7, CP3, CPZ, CP4, TP8, T5, P3, PZ, P4, T6, O1, OZ, and O2) placed according to the international 10-20 system, with close contact to the scalp for real-time acquisition of EEG signals during motor imagery tasks. The signal sampling frequency was set to 50 Hz, and the data were recorded continuously in 32-bit precision. Bilateral mastoid electrodes served as reference electrodes. Throughout the experiment, subjects were instructed to remain still and avoid any noticeable movements or vocalizations. The experiment was conducted on a DELL XPS 8940 microserver equipped with an i7-11700 CPU, RTX 3060Ti graphics card, and 32 GB RAM. Each experimental session consisted of four motor imagery tasks, each repeated three times. Subjects completed 10 sets of experiments on the same day, with a 5-minute break between each set as depicted in Figure 2.



Fig. 2: Subjects doing the experiment

The timing scheme utilized by the brain-computer interface system is illustrated in Figure 3. The total sampling time for a single motor imagery task ranges from 0 to 5 seconds. To ensure the stability and accuracy of the signal, the period

from 1 to 4 seconds is selected for signal feature extraction and classification, as shown in Figure 3(a). Since some motor imagery tasks may have extended durations, the continuous acquisition of signals is segmented into multiple data segments, as depicted in Figure 3(b).

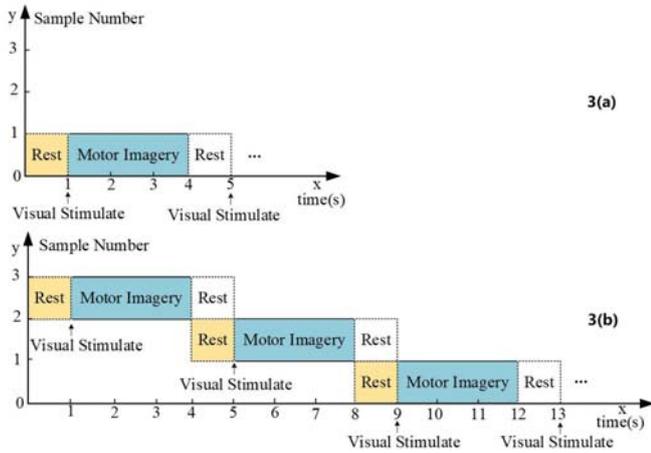


Fig. 3: Timing schemes of acquisition the effective single MI task and continuous MI tasks EEG signals

B. Data Preprocessing

To reduce common-mode noise between electrodes and enhance the overall quality of the data, the bilateral mastoid electrodes were used as the reference points for re-referencing. Independent component analysis (ICA) was then applied to remove artifacts, such as oculomotor and electrocardiogram (ECG) artifacts.

Furthermore, the original EEG data were subjected to random sampling using resampling data augmentation techniques, with the option to replace or retain the original sample. This process aimed to increase the diversity of the data and improve the robustness of model training. A function, `augment_data`, was defined to take the original data and labels as inputs and generate a specified number of augmented data samples through resampling. For each raw sample, the function randomly selects data points and determines whether the same data point can be selected multiple times based on the "replace" parameter. This method not only enhances the model's adaptability to various data variations but also effectively expands the size of the dataset.

C. SNA-MHC Classification Model

1) *SAMPN-Network Feature Extraction Module*: Initially, the EEG signals for a single motor imagery task were normalized using the Z-score method. The normalization formula is given as follows:

$$X_{\text{normalized}}^t = \frac{X^t - \min(X)}{\max(X) - \min(X)} \quad (1)$$

Additionally, the long-term EEG signal is divided into time windows of fixed length, with the data from each window being input to the pulse encoder as a sample. This process

transforms the entire time-series signal into a series of independently processed samples. Let the window length be denoted as L , and n represent the number of samples. Thus, each sample x_i corresponds to the signal within a specific time window, and the overall data structure is as follows:

$$X = \{x_1, x_2, x_3, \dots, x_n\} \quad (2)$$

To convert the continuous EEG signal into a pulse form, a thresholding method is employed for encoding. The trigger pulse threshold is set to $\theta = 1.0$ with a sampling frequency of 500 Hz. At each sampling time the current signal value $s(t)$ is recorded and compared with θ . If the signal exceeds the threshold, a pulse is triggered and set to 1; otherwise, no pulse is triggered, and it is set to 0. The formula is as follows:

$$S_i(t) = \begin{cases} 1, & \text{if } X_i(t) \geq \theta \\ 0, & \text{else} \end{cases} \quad (3)$$

After the motion imagery EEG signal is encoded, it is converted into a binary discrete pulse signal, which is used as the input to the SAMPN-network neural network containing leaky integrate-and-fire (LIF) neurons. LIF neurons process the input pulse signals by dynamically updating the membrane potential $V(t)$. Each input pulse causes an increase in the membrane potential $V(t)$, represented by the weighted accumulation of signal strength and frequency. Simultaneously, the membrane potential decays over time, simulating the charge leakage behavior of biological neurons. A pulse is triggered ($S_k(t) = 1$ for neuronal activation) only when $V(t)$ accumulates to the threshold $\theta = 1.0$, after which it resets to 0, returning to the integrative state. The processing of LIF neurons is as follows:

$$\frac{dV_i^j(t)}{dt} = -V_i^j(t) + \sum_{k=1}^N w_{ik} S_k(t) + I_j(t) \quad (4)$$

Where N is the number of neurons, w_{ik} is the input weight, and $I_j(t)$ is the external input.

Next, the timing pulse signals from the LIF neurons were input to the soft attention mechanism module. The soft attention mechanism computes the feature importance score for each time point by dynamically modeling the characteristics of the pulse signal in the time dimension. These scores were used to adjust the timing pulse weights in the feature map, emphasizing the contribution of critical moments to the overall task. Finally, the weighted features can capture task-related temporal patterns more effectively, enhancing the discriminative power of the model. The formula is as follows:

$$e_i = \tanh(W_e h_i + b_e) \quad (5)$$

The score result is converted into attention weight through Softmax function, the formula is as follows:

$$a_i = \frac{\exp(e_i)}{\sum_{j=1}^T \exp(e_j)} \quad (6)$$

Finally, the time series features output by LIF neurons were weighted and summed using attention weights to get the final features. The formula is as follows:

$$c = \sum_{i=1}^T a_i a h_i \quad (7)$$

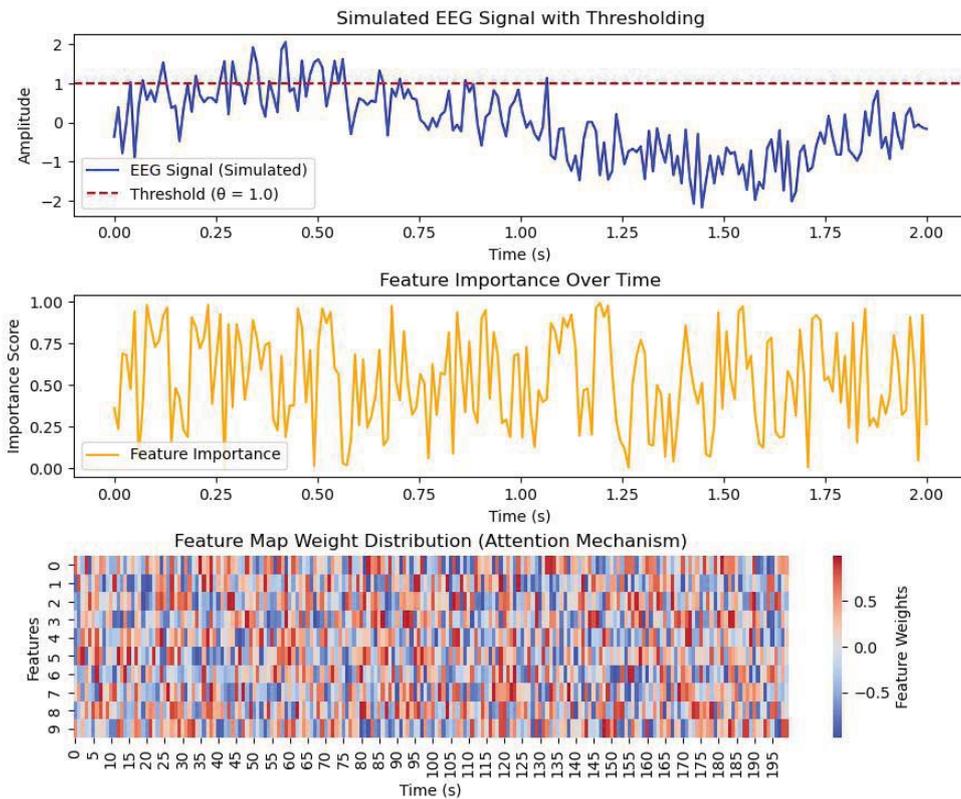


Fig. 4: Weight distribution after attention module scores

The SAMPN-Network feature extraction model process is shown in Figure 8.

2) *Simplicityformer Classification Module*: Multi-Head Attention (MHA) is a technique used to enhance a model's ability to capture dependencies over long distances. This mechanism decomposes the input signal into multiple attention "heads," each independently focusing on different characteristic dimensions of the input data. These heads perform attention calculations in parallel, allowing each to capture information from different feature subspaces of the input data. The outputs of the attention heads are then integrated to form a richer, more comprehensive representation, thereby improving the model's capacity to represent the input data. This section aims to utilize multi-head attention mechanisms to explore the complexity of electroencephalography (EEG) and enhance the comprehensiveness of feature extraction through the parallel processing of multi-dimensional subspaces. The formula can be expressed as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (8)$$

Among them, $Q, K,$ and V are Query, Key, and Value matrices respectively, and d_k is the dimension of the key vector.

During the training process, to ensure the stability of the network and the generalization ability of the optimized model, Layer Normalization is introduced after each multi-head atten-

tion module. This step adjusts and scales the activation values of the input layer. Its mathematical expression is as follows:

$$\text{Layer Norm}(x) = \gamma\left(\frac{x - \mu}{\sigma}\right) + \beta \quad (9)$$

Where μ and σ represent the mean and standard deviation of the data, respectively. and γ and β were learnable parameters. Through layer normalization, the internal covariate shift is reduced, leading to more stable network training.

The deep structure of the model consists of several fully connected layers, which introduce nonlinearity through the Rectified Linear Unit (ReLU) activation function. Specifically, after each fully connected layer performs the linear transformation, the ReLU activation function is applied to perform a nonlinear transformation of the output. The formula is as follows:

$$\text{ReLU}(x) = \max(0, x) \quad (10)$$

Where the linear output $y = W \cdot x + b = W$ of the fully connected layer (where W is the weight matrix, x is the input vector, and b is the bias vector) provides input for subsequent processing by the nonlinear activation function, mapping the input features to the target dimension. The ReLU activation function enhances the network's ability to capture complex patterns by preserving positive values and suppressing negative ones. It also mitigates the vanishing gradient problem, thus improving the training efficiency of deep models.

The output layer of the model uses a dense layer with a Softmax activation function to convert the final feature vector

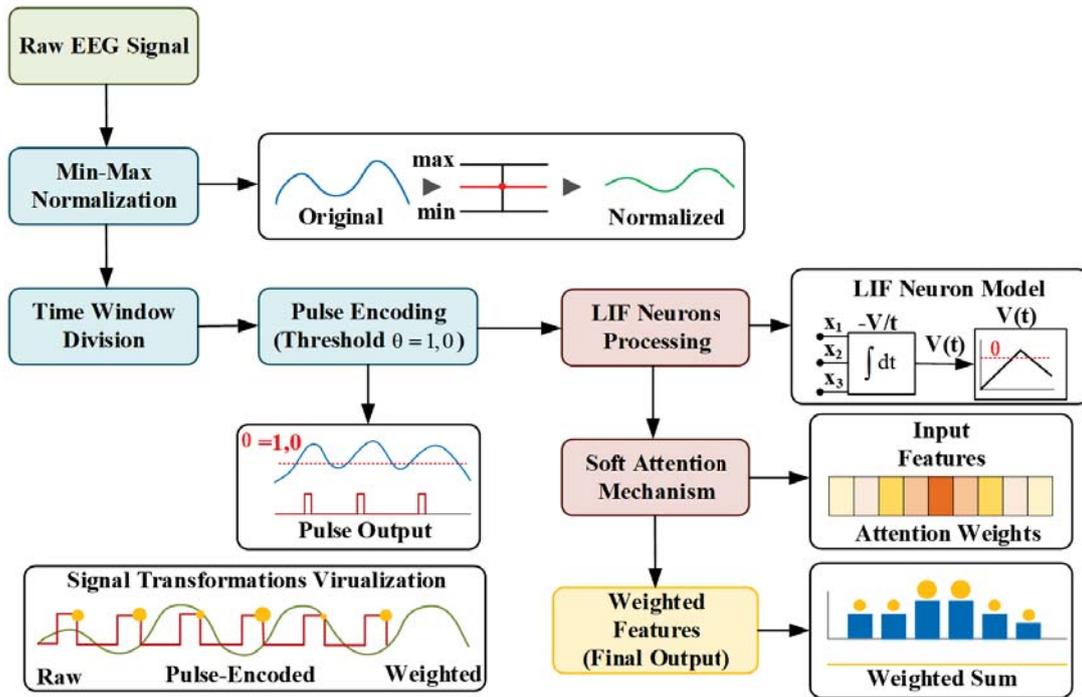


Fig. 5: SAMPN-Network feature extraction model process

into a probability distribution for the predicted class. The formula is as follows:

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (11)$$

Where x_i is the input of the i -th neuron in the output layer. The denominator is the sum of the exponents of all output categories, ensuring that the sum of the output probabilities equals 1.

The overall model is optimized using the Adam optimizer, which combines the advantages of momentum and adaptive learning rates. It employs update rules based on the estimation of the first-order and second-order moments of the gradients, effectively adjusting the update step size for each parameter. This ensures both high efficiency and stability during training.

III. RESULT AND ANALYSIS

A. Experimental Settings

In this experiment, the MNE library was used to preprocess the original motor imagery (MI) task EEG signals, and the TensorFlow framework was employed to construct a neural network for solving the EEG signal classification problem. The experiment was conducted on a DELL microserver equipped with an 11th-generation Intel® Core™ i7-11700 processor (2.50 GHz), an NVIDIA GeForce RTX 3060 Ti graphics card, and 80 GB of RAM. The experimental data for a single subject were run on the constructed network model for 200 epochs, with the batch size and learning rate set to 128 and 0.01, respectively. Each subject's data was divided into training and test sets with a ratio of 85:15. Stratified five-fold cross-validation was applied to split the training set to ensure that the proportion of each class in each fold remained consistent

with that of the original training set. The model proposed in this paper was experimentally compared with the 3D-CNN and LSTM classification model (3D-CLMI) [20], the self-attentional pulsed neural network time-channel joint attention (STCA-SNN) [21], the non-iterative pulsed neural network classification model with attention (NiSNN-A) [22], the end-to-end pulsed neural network model (HR-SNN) [23], and the feature coding-based CNN classification model (CNNs) [24].

B. Experimental Results

The highest classification accuracy of the SNA-MHC model proposed in this paper is 93.89% on the BCI Competition IV 2a dataset, with an average classification accuracy of 92.80%. Compared to the 3D-CLMI, STCA-SNN, NiSNN-A, HR-SNN, CNNs, DFI-HCNN, and CNN-LSTM models, the classification accuracy of SNA-MHC increased by 10.48%, 3.13%, 8.44%, 15.22%, 13.09%, 4.17% and 11.01% respectively. Additionally, the standard deviation of the SNA-MHC model is only 0.69, compared to the 3D-CLMI, STCA-SNN, NiSNN-A, HR-SNN, and CNNs models. This indicates that the SNA-MHC model not only achieves significant improvements in classification accuracy but also demonstrates more stable data adaptability and generalization.

In the TechBrain dataset collected from the laboratory, the highest classification accuracy of the SNA-MHC model was 93.83%, with an average classification accuracy of 93.37%. Compared to other methods, the classification accuracy of the SNA-MHC model improved by 13.02%, 4.41%, 8.46%, 15.05%, 15.88%, 5.62%, and 13.02% respectively. Additionally, the SNA-MHC model has a standard deviation of only 0.43, highlighting its significant advantages in adaptability and generalization.

TABLE I: Comparison of Classification Performance of Different Models on BCI Competition IV 2a Dataset

Subjects	3D-CLMI	STCA-SNN	NiSNN-A	HR-SNN	CNNs	DF-HCNN	CNN-LSTM	SNA-MHC (Ours)
1	83.36	90.32	83.27	79.04	77.69	86.23	83.24	93.66
2	83.48	92.00	82.58	77.58	79.07	90.22	82.37	91.88
3	83.74	89.86	85.85	76.16	78.45	88.64	80.55	93.75
4	81.91	91.29	82.75	77.91	78.04	92.85	80.73	92.50
5	81.51	87.64	83.30	76.30	81.12	91.09	81.14	93.83
6	82.83	86.04	85.49	77.74	81.74	88.34	83.76	92.90
7	81.50	90.30	85.51	79.50	81.44	88.97	81.92	92.79
8	81.42	89.00	83.48	76.39	79.20	88.47	80.17	92.30
9	82.12	91.12	85.14	78.24	79.96	86.38	82.75	93.25
10	80.70	87.05	86.81	75.59	79.42	88.16	81.35	93.89
11	82.47	89.58	86.13	78.76	77.72	88.26	80.91	91.77
12	82.97	91.12	82.57	77.55	78.05	87.66	81.72	91.50
13	82.78	91.54	84.40	76.78	82.24	88.61	82.48	92.36
14	81.81	87.34	84.92	79.28	80.61	88.36	83.12	91.77
15	82.20	90.85	83.20	76.87	80.89	87.47	79.83	93.83
AVG(%)	82.32	89.67	84.36	77.58	79.71	88.63	81.79	92.80
SD	0.71	3.20	1.88	1.39	2.25	2.53	1.73	0.69

TABLE II: Comparison of Classification Performance of Different Models on TechBrain dataset

Subjects	3D-CLMI	STCA-SNN	NiSNN-A	HR-SNN	CNNs	DF-HCNN	CNN-LSTM	SNA-MHC (Ours)
1	79.74	88.66	85.76	75.44	77.65	88.91	77.48	92.59
2	79.88	86.48	84.29	78.51	76.47	88.89	80.15	91.55
3	80.16	86.00	85.71	75.40	76.14	88.45	78.27	92.80
4	80.06	89.36	81.06	78.46	79.73	86.35	81.62	91.62
5	79.87	90.35	85.53	75.40	74.37	84.27	79.51	92.78
6	79.91	88.41	84.58	78.62	76.12	82.29	77.12	93.20
7	79.06	87.31	83.33	77.60	80.12	87.63	79.86	93.06
8	79.80	86.56	81.16	76.58	77.71	85.69	78.94	91.56
9	80.11	86.15	84.42	77.85	75.52	86.49	80.72	93.83
10	78.32	90.22	83.17	77.91	75.00	86.54	81.23	91.74
11	78.44	86.24	83.06	77.61	74.12	87.71	78.36	92.72
12	79.49	90.06	84.29	76.71	76.44	84.78	80.27	92.07
13	78.27	85.48	84.01	79.81	75.49	88.99	79.64	91.94
14	78.23	88.91	84.87	75.98	77.04	88.67	77.93	92.24
15	78.74	89.19	84.40	78.52	75.44	85.59	82.08	92.80
AVG(%)	79.35	87.96	83.91	77.32	76.49	86.75	79.35	92.37
SD	0.53	2.74	1.92	1.58	2.83	3.68	2.69	0.43

After 200 iterations of training, the classification accuracy of the model on the training set stabilized at approximately 92%, while the classification accuracy on the test set ultimately stabilized at about 92.80%. The average classification accuracy of the model is 92.37%. The experimental results demonstrate that, by combining the threshold-coded pulse signal conversion, the SNN integrated with the soft attention mechanism for feature extraction, and its suitability for multi-head attention mechanism classification tasks, the SNA-MHC model exhibits significant advantages in classification tasks on both the BCI Competition IV 2a dataset and the self-collected dataset. It demonstrates superior performance, adaptability, and generalization for the four-class motor imagery (MI) task.

C. Ablation experiment Settings

To explore the role of each component in the SNA-MHC model, an ablation experiment was conducted. This paper evaluates the contributions of resampling, the Spiking Neural Network (SNN), and multi-head attention mechanisms, establishing the following classification models:

1. TSNS (SNA-MHC without Soft Attention): The soft attention mechanism was removed from the original SNA-

MHC model, while the threshold coding, custom pulse neural network, and multi-head attention mechanism (MHA) modules were retained. This model was used to assess the ability of the soft attention mechanism to enhance the model's focus on sparse temporal features and the importance of feature weighting selection.

2. TSNM (SNA-MHC without MHA): The multi-head attention mechanism (MHA) was removed, with the threshold coding module, custom pulse neural network, and SoftAttentionLayer retained. While the model can still extract some features, its ability to focus on key information was reduced, especially when processing complex signals, leading to weakened classification performance.

3. TSND (SNA-MHC without Data Augmentation): The resampling techniques for data augmentation were removed, retaining the threshold coding module, soft attention mechanisms, custom pulse neural networks, and multi-head attention mechanisms. The model was trained directly with the original MI task EEG signals. The removal of the data enhancement strategy negatively impacted the model's generalization ability to different subjects, resulting in a decrease in classification accuracy.

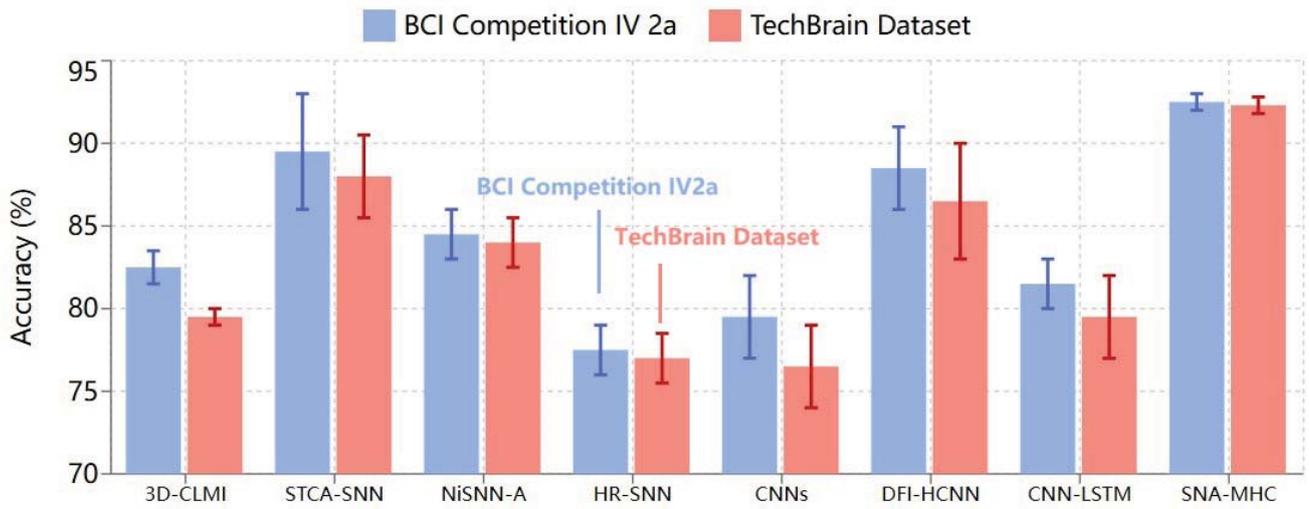


Fig. 6: Histogram of accuracy of two data sets

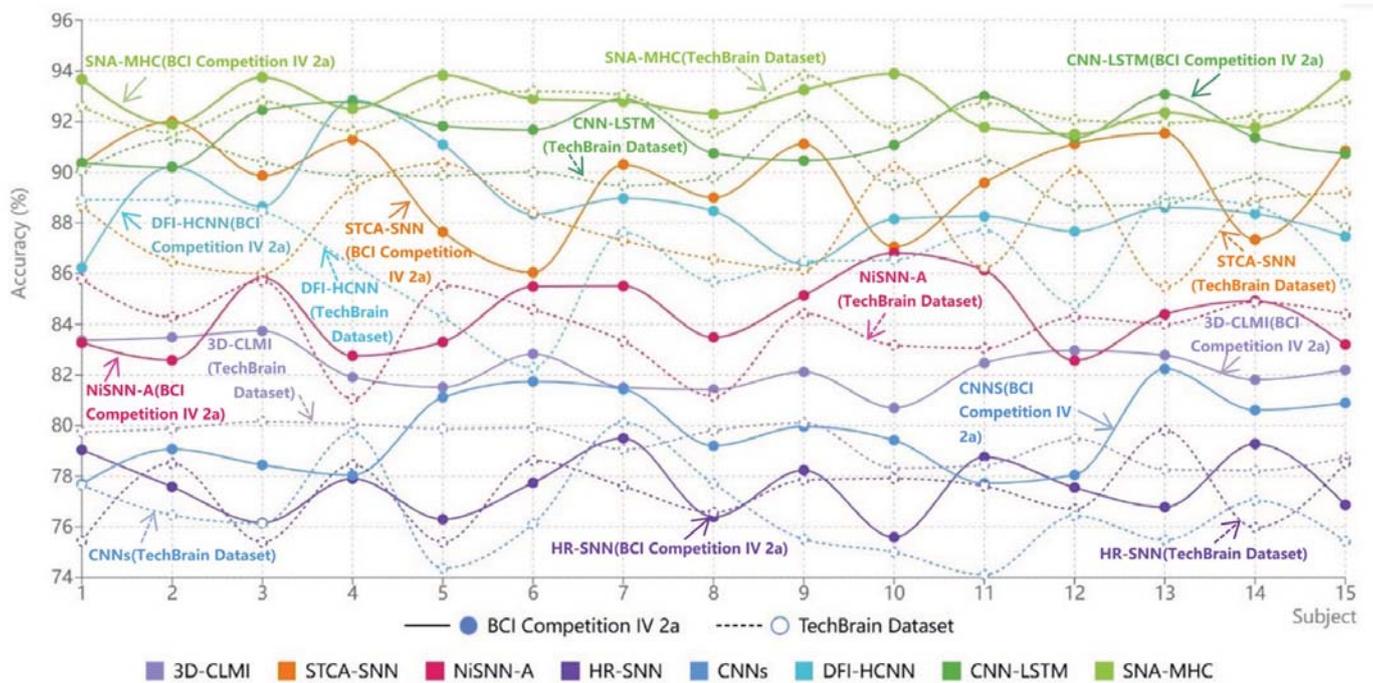


Fig. 7: Changes in the accuracy of the two data sets in different comparison experiments

4. TSNSE (SNA-MHC without Spike Encoding): The threshold coding module was removed, and the custom pulse neural network, SoftAttentionLayer, and multi-head attention mechanism (MHA) were retained. Without pulse encoding, the model's ability to capture the temporal features of EEG signals for MI tasks was significantly reduced, leading to a considerable drop in classification performance.

1) Ablation Results:

IV. CONCLUSION

This paper presents SNA-MHC, a four-class classification algorithm for brain-computer interfaces based on motor imagery, which integrates a custom spiking neural network with

the SoftAttentionLayer. The Simplicityformer classifier, based on the multi-head attention mechanism (MHA), is employed. Initially, the SAMPN-network architecture, which combines spiking neural networks and soft attention mechanisms, is used to extract features from pre-processed EEG signals. The spiking neural network module simulates the transmission of neural impulses, effectively capturing the temporal dependencies of EEG signals, while the soft attention mechanism enhances the focus on key features by dynamically weighting critical time segments. Subsequently, the Simplicityformer classifier, which leverages the multi-head attention mechanism, is employed to process multiple feature subspaces in parallel, capturing rich contextual information and improving the cor-

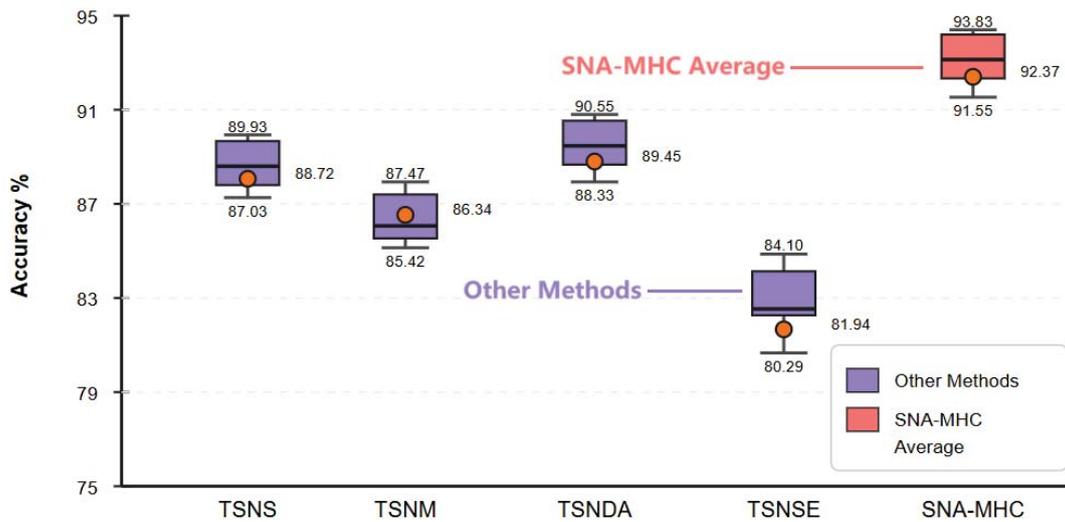


Fig. 8: Diagram of ablation chamber

TABLE III: Comparison of Classification Performance of Different Models in Ablation Experiments on Datasets

Subjects	TSNS	TSNM	TSNDA	TSNSE	SNA-MHC(Ours)
1	87.03	86.39	89.78	81.65	92.59
2	87.63	85.75	89.77	84.10	91.55
3	88.13	86.54	90.48	82.24	92.80
4	88.43	87.41	88.33	81.19	91.62
5	89.03	85.66	88.52	83.07	92.78
6	89.23	87.47	89.68	81.03	93.20
7	89.43	86.66	89.23	82.46	93.06
8	89.63	85.42	90.55	80.29	91.56
9	89.93	86.43	89.33	80.92	93.83
10	88.73	85.66	88.83	82.45	92.80
AVG(%)	88.72	86.34	89.45	81.94	92.37
SD	3.34	0.69	0.71	1.09	0.73

relation between features. The introduction of the multi-head attention mechanism not only enhances feature understanding but also improves the classification accuracy and robustness of the model. Finally, a Softmax classifier is used for the four-class classification of motor imagery tasks.

Experimental results on the BCI Competition IV 2a dataset and the TechBrain dataset demonstrate that the SNA-MHC algorithm significantly outperforms other comparison methods in classification tasks. On the BCI Competition IV 2a dataset, the average classification accuracy of SNA-MHC is 92.80%, with a standard deviation of 0.69, which is substantially higher than that of other comparison methods (3D-CLMI: 82.32%, STCA-SNN: 89.67%, NiSNN-A: 84.36%, HR-SNN: 77.58%, CNNs: 79.71%, DFI-HCNN: 88.63%, CNN-LSTM: 81.79%). Compared with other methods, the classification accuracy of SNA-MHC on the BCI 2a dataset improved by 10.48% (3D-CLMI), 3.13% (STCA-SNN), and 8.44% (NiSNN-A), respectively, and by 15.22% (HR-SNN) and 13.09% (CNNs). On the TechBrain dataset, the average classification accuracy of SNA-MHC is 92.37%, with a standard deviation of 0.43, again outperforming the comparison methods (3D-CLMI: 79.35%, STCA-SNN: 87.96%, NiSNN-A: 83.91%, HR-SNN: 77.32%,

CNNs: 76.49%). Compared to the other methods, the classification accuracy of SNA-MHC on the TechBrain dataset improved by 13.02% (3D-CLMI), 4.41% (STCA-SNN), and 8.46% (NiSNN-A), respectively, and by 15.05% (HR-SNN) and 15.88% (CNNs).

The results of the ablation studies show that the spiking neural network module in the SAMPN-network, the soft attention mechanism, and the multi-head attention mechanism in the Simplicityformer play crucial roles in enhancing the model’s performance. By analyzing each module individually, the results demonstrate that each component is essential in capturing temporal features, enhancing focus, and integrating information. The experimental findings confirm that the SNA-MHC algorithm exhibits excellent performance and generalization capabilities in the four-class motor imagery task, providing strong support for the practical application of brain-computer interface systems.

REFERENCES

- [1] Y. Sun et al., “A surgery-detection two-dimensional panorama of signal acquisition technologies in brain-computer interface,” Aug. 30, 2023, arXiv: arXiv:2308.16102.
- [2] M. F. Almufareh, S. Kausar, M. Humayun, and S. Tehsin, “Leveraging motor imagery rehabilitation for individuals with disabilities: A review,” Healthcare, vol. 11, no. 19, Art. no. 19, Jan. 2023.
- [3] J. Edwards, “Signal processing powers next-generation prosthetics: Researchers investigate techniques that enable artificial limbs to behave more like their natural counterparts [special reports],” IEEE Signal Process. Mag., vol. 35, no. 1, pp. 13–16, Jan. 2021.
- [4] L. Brusini, F. Stival, F. Setti, E. Menegatti, G. Menegaz, and S. F. Storti, “A systematic review on motor-imagery brain-connectivity-based computer interfaces,” IEEE Trans. Hum.-Mach. Syst., vol. 51, no. 6, pp. 725–733, Dec. 2021.
- [5] P. Marque, M. Prudhomme, C. Chotard, X. de Boissezon, and E. Castel-Lacacnal, “Treatment of central neuropathic pain by rTMS associated to self-rehabilitation of imagined movement or mirror therapy,” Ann. Phys. Rehabil. Med., vol. 60, pp. e76–e77, Sep. 2022.
- [6] J. Edwards, “Signal processing powers next-generation prosthetics: Researchers investigate techniques that enable artificial limbs to behave more like their natural counterparts [special reports],” IEEE Signal Process. Mag., vol. 35, no. 1, pp. 13–16, Jan. 2023.

- [7] O. R. Pinheiro, L. R. G. Alves, and J. R. D. Souza, "EEG signals classification: Motor imagery for driving an intelligent wheelchair," *IEEE Lat. Am. Trans.*, vol. 16, no. 1, pp. 254–259, Jan. 2018.
- [8] T.-W. Shi, G.-M. Chang, J.-F. Qiang, L. Ren, and W.-H. Cui, "Brain computer interface system based on monocular vision and motor imagery for UAV indoor space target searching," *Biomed. Signal Process. Control*, vol. 79, p. 104114, Jan. 2023.
- [9] J.-M. Kim et al., "A learnable continuous wavelet-based multi-branch attentive convolutional neural network for spatio-spectral-temporal EEG signal decoding," *Expert Syst. Appl.*, vol. 251, p. 123975, Oct. 2024.
- [10] N. Sharma et al., "An efficient approach for recognition of motor imagery EEG signals using the fourier decomposition method," *IEEE Access*, vol. 11, pp. 122782–122791, 2023.
- [11] G. Roy, A. K. Bhoi, and S. Bhaumik, "A comparative approach for MI-based EEG signals classification using energy, power and entropy," *IRBM*, vol. 43, no. 5, pp. 434–446, Oct. 2022.
- [12] Z. Hu, H. Wu, and L. He, "A neural network for EEG emotion recognition that combines CNN and transformer for multi-scale spatial-temporal feature extraction," vol. 51, no. 8, 2024.
- [13] X. Wang, Y. Wang, W. Qi, D. Kong, and W. Wang, "BrainGridNet: A two-branch depthwise CNN for decoding EEG-based multi-class motor imagery," *Neural Netw.*, vol. 170, pp. 312–324, Feb. 2024.
- [14] A. Echtioui, W. Zouch, M. Ghorbel, and C. Mhiri, "Convolutional neural network with support vector machine for motor imagery EEG signal classification," *Multimed. Tools Appl.*, vol. 82, no. 29, pp. 45891–45911, Dec. 2023.
- [15] S. Guan, K. Zhao, and S. Yang, "Motor imagery EEG classification based on decision tree framework and riemannian geometry," *Comput. Intell. Neurosci.*, vol. 2019, no. 1, p. 5627156, 2019.
- [16] C. D. Virgilio G., J. H. Sossa A., J. M. Antelis, and L. E. Falcón, "Spiking neural networks applied to the classification of motor tasks in EEG signals," *Neural Netw.*, vol. 122, pp. 130–143, Feb. 2020.
- [17] H. Ghaemi, E. Mirzaei, M. Nouri, and S. R. Kheradpisheh, "BioLCNet: Reward-modulated locally connected spiking neural networks," Jul. 07, 2022, arXiv: arXiv:2109.05539..
- [18] W. Zheng, P. Zhao, G. Chen, H. Zhou, and Y. Tian, "A hybrid spiking neurons embedded LSTM network for multivariate time series learning under concept-drift environment," *IEEE Trans. Knowl. Data Eng.*, pp. 1–1, May 2022.
- [19] M. Yao et al., "Attention spiking neural networks," Sep. 28, 2022, arXiv: arXiv:2209.13929.
- [20] S. Cheng and Y. Hao, "3D-CLMI: A motor imagery EEG classification model via fusion of 3D-CNN and LSTM with attention," Dec. 20, 2023, arXiv: arXiv:2312.12744.
- [21] X. Wu et al., "STCA-SNN: Self-attention-based temporal-channel joint attention for spiking neural networks," *Front. Neurosci.*, vol. 17, p. 1261543, Nov. 2023.
- [22] C. Zhang, W. Pan, and C. D. Santina, "NiSNN-a: Non-iterative spiking neural networks with attention with application to motor imagery EEG classification," Dec. 09, 2023, arXiv: arXiv:2312.05643.
- [23] Y. Li, L. Fan, H. Shen, and D. Hu, "HR-SNN: An end-to-end spiking neural network for four-class classification motor imagery brain-computer interface," *IEEE Trans. Cogn. Dev. Syst.*, vol. 16, no. 6, pp. 1955–1968, Dec. 2024.
- [24] Y. Hu, J. Yan, F. Fang, and Y. Wang, "Character encoding-based motor imagery EEG classification using CNN," *IEEE Sens. Lett.*, vol. 7, no. 10, pp. 1–4, Oct. 2023.