Oligodendrocyte-Driven Spiking Neural Model

Mengqiao Han[1], **Liyuan Pan**[2†], **Xiabi Liu**[2†], **Hongming Zhang**[1]

[1]College of Information Engineering, Northwest Agricultural and Forestry (A&F) University
[2]School of Computer Science and Technology, Beijing Institute of Technology
hanmq@nwafu.edu.cn, {liyuan.pan, liuxiabi}@bit.edu.cn, zhm@nwsuaf.edu.cn

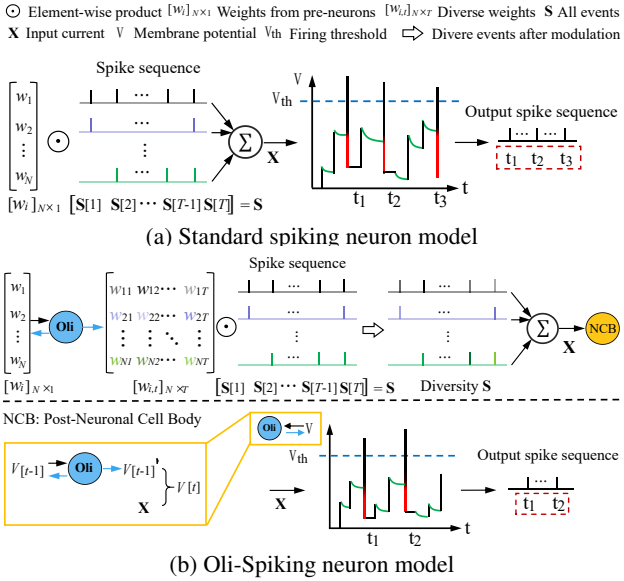
Abstract

The spiking neuron model (SNM) mimics the processing paradigm of synaptic and membrane potentials in the cerebral cortex. However, existing SNMs are limited by two issues. First, they lack spike diversity. Although a spiking neuron perceives temporally varying input currents, SNMs only use identical synaptic weights for regulation. Second, they are insensitive to weak spikes. The potential accumulation in SNMs is solely driven by external inputs, ignoring the internal dynamics of potential. Oligodendrocytes, a recent revelation in neuroscience, enhance neural signaling by forming bidirectional communication. This offers the potential to alleviate the aforementioned issues. In this paper, we first propose the mechanism of the oligodendrocyte-spiking neuron (Oli-N) model. Subsequently, using the Oli-N model, we develop our Oli-inspired spiking neural network (Oli-SNN), which broadens the diversity of spike representations and enhances neurons' firing precision through improved sparse coding to enhance weak spikes. Experiments show that our Oli-SNN achieves state-of-the-art performance in the classification task on both static and neuromorphic datasets.

Introduction

Spiking neural networks (SNNs) are a promising way to achieve energy-efficient intelligence, they encode information into spike patterns by modeling spike transmission and nonlinear activation inherent in the neurodynamics of biological neurons (Mainen and Sejnowski 1995; Eshraghian et al. 2023; Ghosh-Dastidar and Adeli 2009).

As the basic unit of existing SNN, the spiking neuron model (SNM) temporally transmits information from N pre-neurons to a post-neuron via spikes (Fig. 1a). Given T timestamps and N pre-neurons, each pre-neuron generates either a spike or not per timestamp. This outputs $\mathbf{S}$ with NT elements, where $\mathbf{S} = [s_{i,t}]_{N\times T}$ and $s_{i,t} \in \{0, 1\}$. Typically, spikes generated by all pre-neurons at a timestamp $t \in [1, T]$ constitute an event $\mathbf{S}[t]$, where $\mathbf{S}[t] \in \mathbb{R}^{N\times 1}$ is the t^{th} column of $\mathbf{S}$. These T events are transmitted along the axon of pre-neuron and undergo modulation by the same set of neurotransmitters (referred to as synaptic weights in a network, *i.e.*, $\mathbf{W} = [w_i]_{N\times 1}$), resulting in currents $\mathbf{X}$, where



(a) Standard spiking neuron model

(b) Oli-Spiking neuron model

Figure 1: Illustration of the spiking neuron model in existing works and ours. (a) Spiking neuron model, where $t \in [1, T]$ denotes the t^{th} timestamp, V is the membrane potential. Here, $\mathbf{S}$ is a matrix that records whether a neuron has spiked at t. Each event $\mathbf{S}[t]$ is regulated by an identical set of synaptic weights, $\mathbf{W} = [w_i]_{N\times 1}$. (b) Our Oli-spiking neuron model, the oligodendrocyte (Oli) is a key support element in neural function via bidirectional communication with spiking neurons. The Oli assigns specific synaptic weights, $\mathbf{W} = [w_{i,t}]_{N\times T}$, to spikes at each timestamp and correlates the V with its previous activity, which leads the neuron to output sparser spike sequences.

$\mathbf{X} = [\mathbf{S}^\mathsf{T}\mathbf{W}]_{T\times 1}$ and $^\mathsf{T}$ is the transposed operator. Essentially, the weight for a spike sequence from the same pre-neuron remains constant at different timestamps. Then, these currents simulate the post-neuron's membrane potential at various timestamps, generating output spikes with temporal characteristics that stimulate the post-neuron.

As a simplified representation of neurodynamics between neurons, the current SNM raised two issues. **First**, neuronal synapses exhibit plasticity, leading to sustained functional changes upon repeated stimulation by spikes at differ-

† Corresponding author.

ent timestamps, which can regulate neurotransmitter types, thereby improving the precision of temporal information processing (Du and Dreyfus 2002). However, existing SNMs uniformly process spikes across entire timestamps using the same weights, limiting the representation of diverse characteristics in spikes. **Second**, changes in neuronal membrane potential influence the subsequent evolution of its potentials (Fields 2015), thereby enhancing the temporal correlation between spike sequences. In existing SNM, membrane potential accumulation depends solely on input currents, without accounting for the potential itself, which reduces neuronal sensitivity to weak spikes.

Fortunately, oligodendrocytes (Oli) in neurodynamics theory (Xin and Chan 2020), with their bidirectional communication, promisingly enhance the biological plausibility of current SNN functions. Oli enables a temporal response to both neurotransmitters and membrane potentials. Consequently, they regulate synaptic plasticity to encourage sparse diversity and provide feedback on membrane potentials, enhancing sensitivity to weak spikes. This prompts neurons to generate sparser spike sequences over time (Fig. 1b). Therefore, we introduce Oli to SNN and formulate the Oli-spiking neuron (Oli-N) model in the following aspects.

1) Synaptic plasticity. Oli, when stimulated by neurotransmitters, produces neurotrophins that regulate the subsequent release of neurotransmitters in both quantity and type (Du and Dreyfus 2002). Oli assigns different weights, *e.g.*, $\mathbf{W} = [w_{i,t}]_{N\times T}$, to spikes at timestamp $t-1$ based on the weight at t, through bidirectional communication with pre-neurons' weights. This diversity in spike-weight combinations enhances the representation of SNN-encoded.

2) Membrane potential feedback. The neuronal potential activity at timestamp $t-1$ influences Oli's behavior, prompting it to form myelin sheaths around the axon, which accelerates the conduction of action potentials (spikes) along the axon, enhancing the synchronization of post-neuronal firing at timestamp t (Fields 2015). In existing SNMs, neurons and axons are typically modeled as a unified entity, thus Oli plays a crucial role in regulating the precision of neuronal firing. Based on this, Oli controls the potential activity at timestamp t, *e.g.*, $V[t]$, in response to changes in potential at $t-1$. This mechanism enables potential changes at timestamp t beyond reliance solely on input currents, enabling even weak spikes to directly influence the neuron's firing.

3) Collaboration. Oli is typically not isolated, but groups with multiple Olis to collectively optimize neural communication. This allows each member to focus on regulating neural functions within a specific region, improving optimization efficiency (Simons and Trajkovic 2006). Therefore, we extend the concept of Oli from a single entity to a group.

Additionally, we also propose a **4)** connection integration mechanism. This allows Oli to process the weights from multiple neurons. With the formulated Oli-N model, we construct our Oli-inspired spiking neural network (Oli-SNN) to be directly trainable. Our main contributions include:

- We first introduce oligodendrocytes into the spiking neuron model and then formulate the Oli-N model.
- We construct Oli-SNN based on the Oli-N model, achieving accuracy improvements on the classification task alongside sparser spike sequences.
- Compared to SOTA SNN methods, Oli-SNN achieves an accuracy improvement of $0.6\% \sim 2.3\%$ on neuromorphic datasets, and $0.48\% \sim 1.18\%$ on static datasets.

Related Work

Spiking Neuron Models. Hodgkin and Huxley pioneered the modeling of action potentials in biological neural networks (Hodgkin and Huxley 1952), offering extensive biological details and high computational costs (Kistler, Gerstner, and Hemmen 1997). Subsequently, several neuron models have emerged, designed to receive continuous inputs and convert them into spike sequences, such as spike response model (SRM) (Jolivet and Gerstner 2003), Izhikevich neuron model (Izhikevich 2003), Leaky Integrate-and-Fire neuron (Dayan and Abbott 2005; Burkitt 2006; Wu et al. 2018), and PLIF (Fang et al. 2021).

Spiking Neural Networks. Unlike traditional deep learning models that convey information using continuous decimal values, SNNs use discrete spike sequences to calculate and transmit information (Cao, Chen, and Khosla 2015). There are three ways to get SNN models: 1) unsupervised learning; 2) indirect supervised learning; and 3) supervised learning. The first one originated from the weight modification of biological synapses, *e.g.*, spike timing-dependent plasticity (Querlioz et al. 2013). Due to its primary dependency on local neuronal activities without a global supervisor, achieving high performance is quite difficult (Kheradpisheh, Ganjtabesh, and Masquelier 2016). The second one firstly trains an artificial neural network (ANN), and then the pre-trained ANN is converted to SNN by replacing ReLU activation layers with spiking neuron layers (Meng et al. 2022), which requires extremely high latency to achieve satisfying accuracy (Han, Srinivasan, and Roy 2020). The last one indicates that gradient descent is a very popular optimization method for this type. However, the event-triggered mechanism in spiking neurons is non-differentiable, and the surrogate gradient is used for backpropagation (Yao et al. 2023a). Unlike the existing SNN, our Oli-SNN, inspired by the Oli concept, spikes diversity and sensitivity to weak spikes.

Neural Communication. Oligodendrocytes temporally respond to synapse-released neurotransmitters and neuronal potential activity, thereby providing continuous feedback to optimize neural function (Du and Dreyfus 2002; Fields 2015). This suggests that changes in neural function combine external diversity injections and intrinsic potential status (Duncan, Simkins, and Emery 2021). Building on this, we establish a more reasonable spiking neuron model with diverse inputs and precise firing.

Methodology

Preliminaries

The spiking neuron accumulates membrane potential from input currents, which is compared against a threshold to determine spike generation. In this work, we transform the Leaky Integrate-and-Fire (LIF) spiking neuron (Wu et al.

2018), whose dynamic model is described as follows:

$$\mathbf{X} = (\mathbf{S})^{\mathsf{T}}\,\mathbf{W} = \Big[\sum_{i=1}^{N} s_{i,t}\; w_i\Big]_{T\times 1}\,, \tag{1}$$

$$H[t] = V[t-1] + \frac{1}{\tau}\big(X[t] - (V[t-1] - V_{reset})\big)\,, \tag{2}$$

$$S_o[t] = \Phi(H[t] - V_{th})\,, \tag{3}$$

$$V[t] = H[t](1 - S_o[t]) + V_{reset} S_o[t]\,, \tag{4}$$

where N is the number of pre-neurons and $t \in [1, T]$ is the timestamp. Here, $\mathbf{S} = [s_{i,t}]_{N\times T}$ denotes an $N * T$ binary matrix of 0 and 1. Specifically, $s_{i,t}$ is a spike at timestamp t produced by the i^{th} pre-neurons. Note, $\mathbf{W} = [w_i]_{N\times 1}$ is the weights (modulation amplitudes) that regulate all events from the same pre-neurons, which lacks specific consideration for each event concerning the timestamps. We denote the input currents as $\mathbf{X} \in \mathbb{R}^{T\times 1}$, and $X[t] \in \mathbf{X}$ represents the input current at timestamp t. When the membrane potential $H[t]$ exceeds the firing threshold V_{th}, the neuron outputs a spike $S_o[t]$. Here, τ is the membrane time constant, and $\Phi(V)$ is the Heaviside step function, which is 1 when $V \geq 0$ and 0 otherwise. Following an event, the membrane potential $V[t]$ resets to reset potential V_{reset} if a spike is produced, while it remains unchanged as $H[t]$ if no spike occurs.

Oligodendrocyte-Spiking Neuron Model

With the inclusion of Olis, the weights originating from pre-neurons and the membrane potential from post-neurons are initially transmitted to the Oli. Subsequently, the Oli, through bidirectional communication, temporally assigns specific weights to individual events and facilitates the accumulation of correlated membrane potential. We define the bidirectional communication introduced by Oli from four aspects to formulate our Oli-N model.

First, the bidirectional communication endows neuronal synapses with plasticity, enabling Oli to assign diverse weights to regulate events at different timestamps (Du and Dreyfus 2002). We take a unit 'N_i – Oli' from Fig. 2a as an example. The pre-neuron $\mathbf{N}_i$ initially releases the weight $w_{i,1} = w_i$ to trigger the Oli, which then produces neurotrophins as feedback to regulate $\mathbf{N}_i$, inducing the re-release of new neurotransmitters $w_{i,2} = w_i$, which can also re-trigger the Oli. In general, upon receiving the weight $w_{i,t-1}$, the Oli then produces temporal output $w_{i,t} = f^{\mathrm{A}}_{i,t}(w_{i,t-1})$, providing distinct regulation amplitudes for each timestamp, thereby amplifying the diversity of spatio-temporal information. Based on Eq. (1), the diversity of spikes is generated due to temporal weights

$$\mathbf{X} = (\mathbf{S} \odot \mathbf{W}^{\mathrm{n1}})^{\mathsf{T}}\mathbf{1}_N = \Big[\sum_{i=1}^{N} s_{i,t}\; w_{i,t}\Big]_{T\times 1}\,, \tag{5}$$

$$\mathbf{W}^{\mathrm{n1}}[t] = \begin{cases} F^{\mathrm{A}}_t(\mathbf{W}^{\mathrm{n1}}[t-1]) = [f^{\mathrm{A}}_{i,t}(w_{i,t-1})]_{N\times 1}, & 1 < t \leq T \\ \mathbf{W}, & t = 1\,, \end{cases}$$

where $\odot$ denotes the element-wise product, and $\mathbf{W}^{\mathrm{n1}} \in \mathbb{R}^{N\times T}$ is the time-dependent weight generated by the bidirectional communication between pre-neurons and the Oli. Here, $w_{i,t} \in \mathbf{W}^{\mathrm{n1}}$ denotes the weight for the i^{th} neuron at timestamp t, imparting each spike with a specific modulation amplitude. We defined $\mathbf{W}^{\mathrm{n1}}[t]$ as the t^{th} column of the weight matrix $\mathbf{W}^{\mathrm{n1}}$. Note, $f^{\mathrm{A}}_{i,t}(\cdot)$ is a regulation function, and $F^{\mathrm{A}}_t(\cdot) = f^{\mathrm{A}}_{i,t}\mathbf{1}_N$.

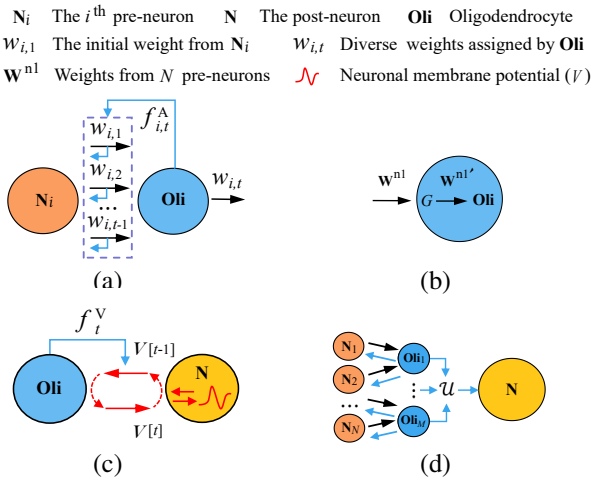


Figure 2: Illustrations of the bidirectional communication between the neuron and the oligodendrocyte. (a) Synaptic plasticity. The Oli responds to the weights from the pre-neuron $\mathbf{N}_i$, assigning specific weights to regulate individual events. (b) Connection integration mechanism. The Oli integrates the weights from N pre-neurons. (c) Membrane potential feedback. The Oli correlates the membrane potential accumulation at timestamp t with its intrinsic potential activity at $t-1$. (d) Collaboration. Multiple Olis collaboratively engage in optimizing the neural function.

Second, following the concept of Olis (Bergles and Richardson 2016), the Oli integrates the weights from N pre-neurons at each timestamp, such as $\mathbf{W}^{\mathrm{n1}}[t]$, to regulate the weights based on global considerations (Fig. 2b). Consequently, the Oli in our Oli-N model demonstrates a connection integration mechanism. Based on Eq. (5), we have

$$\begin{aligned} \mathbf{X} &= (\mathbf{S} \odot \mathbf{W}^{\mathrm{n2}})^{\mathsf{T}}\mathbf{1}_N = \Big[\sum_{i=1}^{N} s_{i,t}\; w_{i,t}\Big]_{T\times 1} \\ \mathbf{W}^{\mathrm{n2}}[t] &= \begin{cases} F^{\mathrm{A}}_t(G^t(\mathbf{W}^{\mathrm{n1}}[t-1])), & 1 < t \leq T \\ G^1(\mathbf{W}), & t = 1\,, \end{cases} \end{aligned} \tag{6}$$

where $\mathbf{W}^{\mathrm{n1}'}[t] = G^t(\mathbf{W}^{\mathrm{n1}}[t-1])$, $G^t(\cdot)$ is a integration function for timestamp t, $\mathbf{W}^{\mathrm{n1}'}[t] \in \mathbb{R}^{N\times 1}$ resulting in a weight $w_{i,t} \in \mathbf{W}^{\mathrm{n2}}$ is related to weights from all N pre-neurons at the same timestamp.

Third, Oli establishes correlations between membrane potential activities at different timestamps (Madorran et al. 2022), such that the membrane potential changes at timestamp t are influenced by both the input currents and its previous potential activity. This process enhances the neuron's sensitivity to weak spikes. As depicted in Fig. 2c, the Oli is activated by the membrane potential recorded from timestamp $t-1$, *i.e.*, $V[t-1]$, and then influences the potential update at timestamp t. Then, the correlated membrane potential, in conjunction with the input current, promotes the

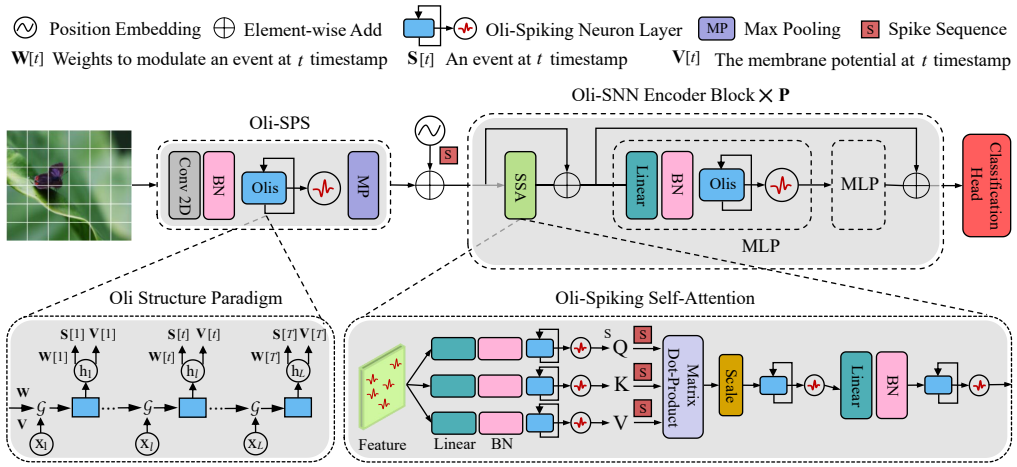


Figure 3: The schematic of Oli-SNN, in which our Oli-N model is used to construct the spiking neuron layer. Oli-SNN comprises an Oli-Spiking patch splitting module (Oli-SPS), an Oli-SNN encoder module, and a linear classification head module. In the zoom-in picture, the structure of each oligodendrocyte (Oli) is configured as an LSTM with L layers.

potential accumulation. As pre Eq. (2), the resultant correlated membrane potential is

$$H[t] = f_t^{\mathrm{V}}(V[t-1]) + \frac{1}{\tau}\Big(X[t] - \big(f_t^{\mathrm{V}}(V[t-1]) - V_{reset}\big)\Big)\,, \quad (7)$$

where $f_t^{\mathrm{V}}(\cdot)$ is a control function of the Oli, allowing neurons to fire by incorporating their historical potential activity. The output of the Oli comprises two shared-parameter output forms, denoted as $F_t(\cdot) = \{F_t^{\mathrm{A}}(\cdot), f_t^{\mathrm{V}}(\cdot)\}$. This configuration allows the Oli to refine the precision of neuronal firing by updating the membrane potential while considering external input currents.

Fourth, Oli typically does not exist in isolation; rather, multiple Olis form a collaborative group that optimizes neural communication through dynamic interactions among them. Therefore, we expand the model to involve multiple Olis, enabling each member's function to focus on a specific region, thereby improving overall optimization capability. Moreover, the state of each Oli is derived from the optimization results involving the other $M-1$ Olis, enhancing its dynamic regulation capacity.

Considering M number of Olis, denoted as $\mathbf{O}_{\{M\}}$, where each corresponds to a subset of N_k pre-neurons, *e.g.*, the k^{th} Oli, $\mathbf{O}_k$, establishes bidirectional communication with N_k pre-neurons along with the correlated post-neuron. Formally, we reformulate Eq. (6) as

$$\mathbf{X} = (\mathbf{S} \odot \mathbf{W}^{\mathrm{n3}})^{\intercal}\mathbf{1}_N = \Big[\sum_{i=1}^{N} s_{i,t}\, w_{i,t}\Big]_{T\times 1} \quad (8)$$

$$\mathbf{W}^{\mathrm{n3}}[t] = \begin{cases} \mathcal{U}\Big(F_{t,k}^{\mathrm{A}}\Big(G^{t}\big(\mathbf{e}_k \odot \mathbf{W}^{\mathrm{n1}}[t-1]\big)\Big)\Big)_{k=1}^{M}, & 1 < t \le T \\ G^{1}(\mathbf{e}_k \odot \mathbf{W}), & t = 1\,, \end{cases},$$

and (7) as

$$H[t] = \mathcal{U}\Big(F_{t,k}^{\mathrm{V}}\big(V[t-1]\big)\Big)_{k=1}^{M} + \frac{1}{\tau}\Big(\mathbf{X}[t] - \Big(\mathcal{U}\big(F_{t,k}^{\mathrm{V}}(V[t-1])\big)_{k=1}^{M} - V_{reset}\Big)\Big)\,, \quad (9)$$

where $F_{t,k}(\cdot) = \{F_{t,k}^{\mathrm{A}}(\cdot), F_{t,k}^{\mathrm{V}}(\cdot)\}$ is the function of the k^{th} Oli to simultaneously assign weights and update membrane potential. Here, $\mathbf{e}_k \in \mathbb{R}^{N\times 1}$ is a binary vector with value one in the established connection between pre-neurons and Olis, and $\sum_{k=1}^{M}\mathbf{e}_k = \mathbb{1}$, $|\mathbf{e}_k| = N_k$. Each $\mathbf{e}_k$ can be treated as a mask to select pre-neurons' weights that participate in bidirectional communication with the k^{th} Oli. Note, $\mathcal{U}(\cdot)$ is a joint function among multiple Olis (see Fig. 2d). Finally, we formulate our Oli-N model incorporating Eq. (3), Eq. (4), Eq. (8), and Eq. (9).

Framework of Oli-SNN

We construct a spiking neuron layer, denoted as Oli-L, based on our proposed Oli-N model. By replacing the standard network layer in a Transformer with Oli-L, we develop a directly trainable oligodendrocyte-inspired spiking neural network (Oli-SNN). We further evaluate the advantages of the Oli-N model over the traditional LIF model. Fig.3 provides an overview of our Oli-SNN.

Given a 2D image sequence $\mathbf{I} \in \mathbb{R}^{T\times C\times H\times W}$, with height H, width W, and channel C. The Oli-Spiking Patch Splitting (Oli-SPS) module first splits $\mathbf{I}$ into spike-form patches $\mathbf{x}$. Then, the conditional position embedding generator (CPEG) (Chu et al. 2021) generates spike-form relative position embedding and adds it to $\mathbf{x}$ for getting $\mathbf{X}_0$. After passing the $\mathbf{X}_0$ to the P-block Oli-SNN encoder, which contains an Oli-Spiking Self Attention (Oli-SSA) and an MLP block, we get $\mathbf{X}_p$ with $p \in \{1, \cdots, P\}$. Finally, the classification head (CH) is utilized on $\mathbf{X}_p$ and outputs prediction $\hat{\mathbf{Y}}$, whose described as follows,

$$\begin{aligned} \mathbf{X}_0 &= \text{Oli-SPS}\big(\mathbf{I}\big) + \text{CPRG}\Big(\text{Oli-SPS}\big(\mathbf{I}\big)\Big)\,, \\ \mathbf{X}_p' &= \text{Oli-SSA}\big(\mathbf{X}_{p-1}\big) + \mathbf{X}_{p-1}\,, \\ \mathbf{X}_p &= \text{MLP}\big(\mathbf{X}_p'\big) + \mathbf{X}_p'\,, \\ \hat{\mathbf{Y}} &= \text{CH}(\mathbf{X}_p)\,. \end{aligned} \quad (10)$$

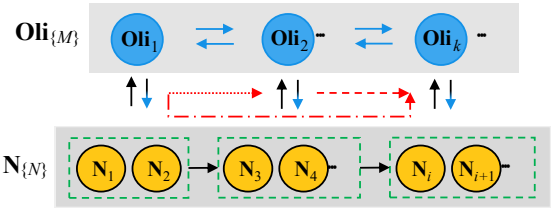


Figure 4: The chain optimization for Olis, $\mathbf{O}_M$. Straight arrows are data flow direction, while curved arrows are the impact of preceding Olis on subsequent ones.

With the Oli-L, Oli-SNN function, $\hat{\mathcal{N}}$ is formulate as

$$\hat{\mathbf{Y}} = \hat{\mathcal{N}}\Big(\mathbf{I}, \mathcal{O}(\mathbf{W}, \mathbf{V}, \mathbf{\Phi})\Big) \,, \tag{11}$$

where optimization of weights $\mathbf{W}$ and membrane potentials $\mathbf{V}$ is driven by the Oli function $\mathcal{O}(\cdot)$, endowing neurons with the capacity to process diverse inputs and exhibit precise firing. Here, $\mathbf{\Phi}$ is the weights of Olis, and $\mathbf{\Phi} = \{\Phi_F, \Phi_G, \Phi_{\mathcal{U}}\}$ are implied in $\{F_{t,k}(\cdot)\}_{t,k=1}^{T,M}$, $\{G^t(\cdot)\}_{t=1}^{T}$, and $\mathcal{U}(\cdot)$.

Bidirectional communication enables each Oli, $F_{t,k}(\cdot)$, to temporally optimize neural communication, endowing it with the characteristics of processing time series data. Consequently, as depicted in Fig. 3, this function is constructed using a recursive LSTM, denoted as $F_{t,k}(\cdot) = [\text{LSTM}_l]_{l=1}^{L}$, where l denotes the number of layers within the LSTM. This structure enables the capture of temporal correlations between event features and potential changes across the entire timestamp. In each layer of the LSTM, the input integrates the output from all preceding LSTM layers, *i.e.*, $F_{t,k}(w,l) = \sum_{l=1}^{l-1} F_{t,k}(w,l)$. This amplifies the responsiveness of Olis to the historical optimization of events and potentials at each timestamp.

Given that the weights $\mathbf{W}$ can involve a substantial number of parameters, *e.g.*, a 101-layer ResNet has approximately 44.5 million parameters (He et al. 2016). Thus, the function $G^t(\cdot)$ needs to compress $\mathbf{W}$, while retaining more key information. We use attention and ViT modules to built this function. The attention module first compresses the input-weight of Oli into channel descriptors via global average pooling. Descriptors are then fused through a lightweight fully connected layer, resulting in the global feature for each Oli input. The input features of each Oli are treated as a patch fed into the ViT module, incorporating positional information to represent their associations. Due to the compression operation, ViT can be designed as an extremely lightweight model.

The collaboration between Olis is established through the joint function $\mathcal{U}(\cdot)$, with dynamic interaction realized via a chain optimization. As depicted in Fig. 4, bidirectional communication is established between Olis, $\mathbf{O}_{\{M\}}$, and N number of neurons, $\mathbf{N}_{\{N\}}$. During each training epoch, the optimization progresses sequentially from $\mathbf{O}_1$ to $\mathbf{O}_M$ until $\mathbf{O}_{\{M\}}$ is fully optimized. Hence, employing $\mathcal{U}(\cdot)$ involves optimizing $\mathbf{O}_{\{M\}}$ across all timestamps. The chain optimization introduces temporal interactions among neurons by propagating spikes from optimized neurons to influence others. The joint function $\mathcal{U}(\cdot)$ and the integrative connection function $G^t(\cdot)$ within Olis ensure a global-temporal optimization of neurons by each Oli across the entire network.

Training Scheme

Building upon insights from (Simons and Nave 2016), the dynamic interaction between neurons and Olis provides an alternative optimization strategy for Oli-SNN. Constructing our Oli-SNN with Oli-N, we implement a two-step optimization to simulate this interaction procession (Bradl and Lassmann 2010). Initially, we optimize the spiking neuron module (SNM), followed by optimization of the Oli module (OM). The optimal parameters $\mathbf{W}$ of SNM are determined by minimizing the following function

$$\mathbf{W}^* = \arg\min_{\mathbf{W}} \mathcal{L}_{\text{task}}\Big(\hat{\mathcal{N}}\big(\mathbf{I}, \mathcal{O}(\mathbf{W}, \mathbf{V}, \mathbf{\Phi})\big), \mathbf{Y}^{\text{gt}}\Big) \,, \tag{12}$$

where $\mathcal{L}_{\text{task}}(\cdot)$ is the task specific loss function and $\mathbf{Y}^{\text{gt}}$ is the ground-truth label.

Subsequently, utilizing the current optimal $\mathbf{W}^*$, we optimize the OM by incorporating both a task term $\mathcal{L}_{\text{task}}(\cdot)$ and a regularization term $\mathcal{L}_{\text{reg}}(\cdot)$. The task term $\mathcal{L}_{\text{task}}(\cdot)$ drives the OM to optimize the activation status of neurons in the SNM, aiming for superior performance. Simultaneously, the regularization term $\mathcal{L}_{\text{reg}}(\cdot)$ employs $\mathcal{KL}$ divergence to enhance spike diversity, distinguishing the output distribution of Olis at each timestamp. This is defined as

$$\mathcal{L}_{\text{reg}}(\Phi) = -\sum_{k=1}^{M}\sum_{t=2}^{T} F_{t-1,k}(\mathbf{W}_k) \log \frac{F_{t-1,k}(\mathbf{W}_k)}{F_{t,k}(\mathbf{W}_k)} \,, \tag{13}$$

where $\mathbf{W}_k = \mathbf{e}_k \odot \mathbf{W}$ represents the input to the k^{th} Oli. Taking Eq. (12) and (13), the final learning objective of OM is formulated as

$$\begin{aligned} \mathbf{\Phi}^* &= \arg\min_{\mathbf{\Phi}} \big(\mathcal{L}_{\text{task}}(\mathbf{\Phi}) + \lambda \mathcal{L}_{\text{reg}}(\mathbf{\Phi})\big) \\ &= \arg\min_{\mathbf{\Phi}} \big(\mathcal{L}_{\text{OM}}(\mathbf{\Phi})\big) \,, \end{aligned} \tag{14}$$

where λ is a balancing factor. After training, only a small transformed vector output by Oli needs to be saved, which is used to convert the original weights to temporal weights. Thus, the final number of parameters in our Oli-SNN is only slightly increased compared to the baseline.

Experiments

Implementation Details. The experimental models are developed using PyTorch and SpikingJelly. We utilize the AdamW optimizer during training. The configuration is specified as follows: the number of Olis is set to $M = 4$, the diversity coefficient $\lambda = 5e\text{-}4$, and LSTM layers within each Oli is set to $l = 4$. All reported figures represent the averages of five repeated experiments. Please refer to the supplementary material for more details.

Dataset. We conduct experiments on both neuromorphic datasets (DVS128 and CIFAR10-DVS) and static datasets (CIFAR and ImageNet) to evaluate the performance of our Oli-SNN. DVS128 with a 128×128, encompassing 11 hand gesture categories from 29 individuals under 3 illumination conditions. CIFAR10-DVS, with a 128×128, provides

Methods	CIFAR10-DVS		DVS128	
	T	Acc (%)	T	Acc (%)
LIAF-Net	10	70.4	60	97.6
TA-SNN	10	72.0	60	98.6
DSR	10	77.3	20	98.1
Att MS PLIF	20	75.1	20	98.3
Spikformer	16	80.9	16	98.3
Oli-SNN	16	**82.4** ↑ 1.5	16	**98.9** ↑ 0.6
QKFormer	16	84.0	16	98.6
Oli-SNN	16	**85.8** ↑ 1.8	16	**99.3** ↑ 0.7
SpikingResformer	10	81.2	10	91.8
Oli-SNN	10	**83.5** ↑ 2.3	10	**93.2** ↑ 1.4

Table 1: Performance comparison to the SOTA SNN methods on two neuromorphic datasets. 'Acc (%)' denotes Top-1 accuracy, 'T' denotes Time Steps. The best number in **bold**.

$9,000$ training samples and $1,000$ test samples. CIFAR provides $50,000$ training and $10,000$ test images with a 32×32 resolution. ImageNet with a 224×224 resolution, contains around 1.25 million training images across $1,000$ classes and $50,000$ images for validation.

Baselines. We train our Oli-SNN from scratch and conduct comparative against the state-of-the-art (SOTA) direct-training SNN baselines, including Att MS (Yao et al. 2023b), Spikformer (Zhou et al. 2023), DSR (Meng et al. 2022), QKFormer (Zhou et al. 2024) and SpikingResformer (Shi, Hao, and Yu 2024).

Note that in the following experiments, we reconstruct the spiking layer of baselines by replacing their LIF model with our proposed Oli-N model, in order to evaluate the advantages of Oli-N over the traditional spiking neuron model.

Results of Neuromorphic Datasets

The comparison between our Oli-SNN and the SOTA SNN methods is shown in Tab. 1. SNNs typically require more timestamps to capture temporal dependencies in neuromorphic data. For fairness, we follow the timestamp settings used in the SOTA baselines.

For CIFAR10-DVS, our accuracy outperforms the corresponding baseline by 1.5%, 1.8%, and 2.3%, respectively. Additionally, our method shows a substantial 8.5% increase in accuracy compared to DSR with binary spikes (77.3%). For DVS128, our method consistently outperforms the compared SOTA SNN methods. Although TA-SNN employs floating-point spikes during forward propagation, our best result (99.3%) exceeds its performance (98.6%) by 0.7%.

Results of Static Datasets

CIFAR. The experimental results of our method on CIFAR are presented in Tab. 2, alongside the results obtained from the ANN method. For CIFAR10, our Oli-SNN achieves SOTA accuracy. By replacing the LIF model in these baselines—Spikformer, QKFormer, SpikingResFormer—with our Oli-N model, their accuracy is further improved by 0.53%, 0.58% and 0.92%, under the same time step setting. For CIFAR100, the performance improvement

Methods	Architecture	Params (M)	T	C10 Acc (%)	C100 Acc (%)
ANN	Trans-4-384	9.32	-	96.73	81.02
DSR	ResNet18	11.17	20	95.40	78.50
Att MS ResNet	ResNet-34	22.12	1	94.83	76.62
Spikformer	S-Trans-4-384	9.32	4	95.19	77.86
Oli-SNN	Oli-S-Trans-4-384	10.40	4	**95.72** ↑ 0.53	**78.78** ↑ 0.96
QKFormer	HST	6.74	4	96.18	81.15
Oli-SNN	Oli-HST	10.40	4	**96.76** ↑ 0.58	**82.21** ↑ 1.06
SpikingResformer	SResformer-Ti	11.14	4	95.53	79.92
Oli-SNN	Oli-SResformer-Ti	12.36	4	**96.45** ↑ 0.92	**80.86** ↑ 0.94

Table 2: Performance comparison with existing SOTA models on CIFAR. 'Oli-SNN-L-D' denotes L encoder blocks and D feature embedding dimensions. 'S-Trans' denotes Spiking Transformer. 'HST' denotes Hierarchical Spiking Transformer. 'SResformer' denotes SpikingResformer. C10 and C100 denote results on CIFAR10 and CIFAR100.

Methods	Architecture	Params (M)	T	Power (mJ)	Acc (%)
ANN	Trans-8-512	29.68	1	8.33	80.80
DSR	ResNetT18	11.17	20	8.48	67.86
Att MS ResNet	ResNetT34	22.12	1	0.55	69.15
Spikformer	S-Trans-8-512	29.68	4	**11.58**	73.38
Oli-SNN	Oli-S-Trans-8-512	31.06	4	11.82 ↑ 0.24	**73.86** ↑ 0.48
QKFormer	HST-10-768	64.96	4	38.91	84.22
Oli-SNN	Oli-HST-10-768	67.98	4	**36.75** ↓ 2.46	**85.40** ↑ 1.18
SpikingResformer	SResformer-M	35.52	4	5.46	77.24
Oli-SNN	Oli-SResformer-M	37.24	4	**5.43** ↓ 0.03	**78.11** ↑ 0.87

Table 3: Performance comparison of Oli-SNN with existing SOTA models on ImageNet. Power is calculated as the theoretical energy consumption (TEC) when predicting an image from the ImageNet test set.

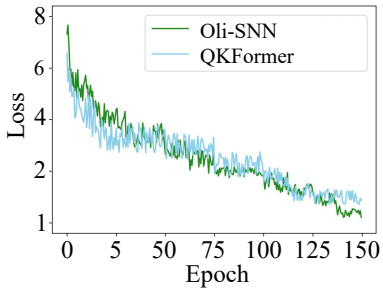


Figure 5: Training loss curve of Oli-SNN over epochs, showing steady decrease in loss and stabilization during training.

achieved by our Oli-SNN is more pronounced, surpassing the competing methods by 0.96%, 1.06% and 0.94%.

ImageNet. We present the comparative results on ImageNet in Tab. 3. Our Oli-SNN, trained from scratch on ImageNet, achieves accuracy improvements of 0.48%, 1.18% and 0.87% by reconstructing the spiking layers of baselines using the proposed Oli-N model.

Note that although our framework introduces a few additional parameters to enhance weight diversity, the power during the testing phase remains reduced for larger models (in terms of parameter count). This is attributed to the lower firing rates of the query, key, value components in Oli-SSA (as illustrated in Fig. 5), which compensate for the overhead introduced by the Oli mechanism.

Model	Training time (1 batch)	Inference time (1 batch)
QKFormer(29.08M, T=4), BS=1	2.72s	1.33s
Oli-SNN (30.45M, T=1), BS=1	3.35s	1.20s
QKFormer(29.08M, T=4), BS=6	3.62s	1.82s
Oli-SNN (30.45M, T=6), BS=6	4.58s	1.59s

Table 4: The training and inference time between Oli-SNN and QKFormer on ImageNet. 'BS' means Batch Size.

Discussion and Ablation Studies

This section explains the rationale behind our method by comparing it to the previous second-best method, QKFormer, and evaluates its performance on ImageNet.

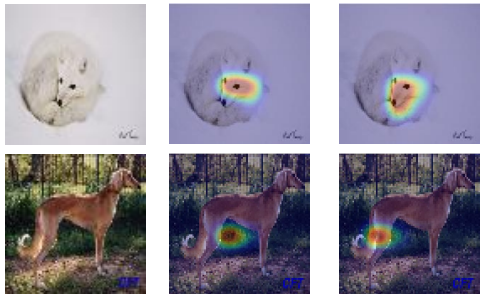
Training curves. The training loss trajectory of Oli-SNN on ImageNet, illustrated in Fig. 5, highlights the convergence of our framework and training scheme. Initially, our Oli-SNN exhibits a relatively high loss value attributed to the dynamic optimization of neurons by Olis in Oli-L has not yet reached stability. As Olis gradually assigns more appropriate modulation amplitudes to events and simultaneously enables adaptive firing in neurons. Our method establishes a lower bound on the QKFormer loss value after 100th epochs. Beyond the 150th epoch, our method reduces oscillatory.

The Training and Inference Time. In terms of training time, the per-batch training time of Oli-SNN is 1.23 times that of QKFormer, primarily due to the overhead introduced by the Oli mechanism. The training epochs of QKFormer on ImageNet are 200th, while our framework's training time is 150th due to its faster reduction of oscillations (as mentioned in Fig. 5). Thus, the total training time cost of Oli-SNN on ImageNet is close to QKFormer.

In terms of inference time, Oli-SNN outperforms QKFormer because Oli is not involved in the calculations during inference (only the transformation vector produced by Oli is stored after training), and Oli-SNN is inherently more sparse (as shown in Fig. 5).

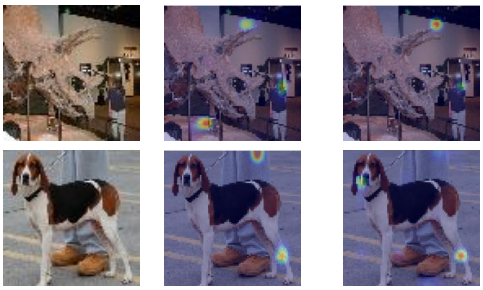
Ablation of Spike Diversity. Spike diversity is reflected in assigning specific weights to neurons at various timestamps. We only introduce the synaptic plasticity mechanism into the spiking neuron model, labeled as 'Ours-Weight', to evaluate its effectiveness. First, we analyze the overlay of attention maps for all events within 4 timestamps on the input image to highlight the benefits of using multiple amplitudes to regulate spikes, rather than a single amplitude for all. This visualization follows Grad-CAM (Selvaraju et al. 2017). Spike diversity encourages the SNN to focus more attention on target regions rather than the surroundings. For instance, in the 1^{st} row of Fig. 6, labeled 'fox', our focus grows within the target region, whereas Spikformer either overlooks or pays less attention.

Ablation of Weak Spike Sensitivity. We utilize only the membrane potential feedback mechanism in our model, labeled as 'Ours-Potential', to evaluate its ability to enhance neuron sensitivity to weak spikes. First, we explore the reflection of attention maps with weights below $5e$-4 in the last encoder block of the model on the input image. This



(a) Input (b) QKFormer (c) Ours-Weight

Figure 6: The attention map results on images for QKFormer and 'Ours-Weight' with only spike diversity.



(a) Input (b) QKFormer (c) Ours-Potential

Figure 7: The attention map results on images for QKFormer and 'Ours-Potential' with only sensitivity to weak spikes.

is achieved by associating membrane potentials at different timestamps through Oli, increasing focus on specific locations in the input image. For instance, in the first row of Fig. 7 labeled 'skeleton', we can highlight the osseous within the target.

Conclusions

In this paper, we propose an oligodendrocyte-spiking neuron model, focusing on the bidirectional communication between Olis and neurons, thereby constructing Oli-SNN. An analysis of the communication mechanism reveals that Oli-SNN exhibits characteristics encompassing more diverse spikes and heightened sensitivity to weak spikes. Experimental results highlight that our Oli-SNN achieves SOTA accuracy through more precise neuron firing.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (62506303, 82171965, 62302045, 42377341), National Key Research and Development Program (2023YFD1300101), Shenzhen Science and Technology Innovation Commission Sustainable Development Technology Special Projects (NO. KCXST20221021111609022), and the Northwest A&F University Scientific Research Startup Foundation.

References

Bergles, D. E.; and Richardson, W. D. 2016. Oligodendrocyte development and plasticity. *Cold Spring Harbor perspectives in biology*, 8(2): a020453.

Bradl, M.; and Lassmann, H. 2010. Oligodendrocytes: biology and pathology. *Acta neuropathologica*, 119: 37–53.

Burkitt, A. N. 2006. A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. *Biological cybernetics*, 95: 1–19.

Cao, Y.; Chen, Y.; and Khosla, D. 2015. Spiking deep convolutional neural networks for energy-efficient object recognition. *International Journal of Computer Vision*, 113: 54–66.

Chu, X.; Tian, Z.; Wang, Y.; Zhang, B.; Ren, H.; Wei, X.; Xia, H.; and Shen, C. 2021. Twins: Revisiting the design of spatial attention in vision transformers. *Advances in Neural Information Processing Systems*, 34: 9355–9366.

Dayan, P.; and Abbott, L. F. 2005. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. MIT press.

Du, Y.; and Dreyfus, C. F. 2002. Oligodendrocytes as providers of growth factors. *Journal of neuroscience research*, 68(6): 647–654.

Duncan, G. J.; Simkins, T. J.; and Emery, B. 2021. Neuron-oligodendrocyte interactions in the structure and integrity of axons. *Frontiers in Cell and Developmental Biology*, 9: 653101.

Eshraghian, J. K.; Ward, M.; Neftci, E. O.; Wang, X.; Lenz, G.; Dwivedi, G.; Bennamoun, M.; Jeong, D. S.; and Lu, W. D. 2023. Training spiking neural networks using lessons from deep learning. *Proceedings of the IEEE*.

Fang, W.; Yu, Z.; Chen, Y.; Masquelier, T.; Huang, T.; and Tian, Y. 2021. Incorporating learnable membrane time constant to enhance learning of spiking neural networks. In *Proceedings of the IEEE/CVF international conference on computer vision*, 2661–2671.

Fields, R. D. 2015. A new mechanism of nervous system plasticity: activity-dependent myelination. *Nature Reviews Neuroscience*, 16(12): 756–767.

Ghosh-Dastidar, S.; and Adeli, H. 2009. Spiking neural networks. *International journal of neural systems*, 19(04): 295–308.

Han, B.; Srinivasan, G.; and Roy, K. 2020. Rmp-snn: Residual membrane potential neuron for enabling deeper high-accuracy and low-latency spiking neural network. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 13558–13567.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Hodgkin, A. L.; and Huxley, A. F. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4): 500.

Izhikevich, E. M. 2003. Simple model of spiking neurons. *IEEE Transactions on neural networks*, 14(6): 1569–1572.

Jolivet, R.; and Gerstner, W. 2003. The spike response model: a framework to predict neuronal spike trains. In *International Conference on Artificial Neural Networks*, 846–853. Springer.

Kheradpisheh, S. R.; Ganjtabesh, M.; and Masquelier, T. 2016. Bio-inspired unsupervised learning of visual features leads to robust invariant object recognition. *Neurocomputing*, 205: 382–392.

Kistler, W. M.; Gerstner, W.; and Hemmen, J. L. v. 1997. Reduction of the Hodgkin-Huxley equations to a single-variable threshold model. *Neural computation*, 9(5): 1015–1045.

Madorran, E.; Stožer, A.; Arsov, Z.; Maver, U.; and Rožanc, J. 2022. A promising method for the determination of cell viability: the membrane potential cell viability assay. *Cells*, 11(15): 2314.

Mainen, Z. F.; and Sejnowski, T. J. 1995. Reliability of spike timing in neocortical neurons. *Science*, 268(5216): 1503–1506.

Meng, Q.; Xiao, M.; Yan, S.; Wang, Y.; Lin, Z.; and Luo, Z.-Q. 2022. Training high-performance low-latency spiking neural networks by differentiation on spike representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12444–12453.

Querlioz, D.; Bichler, O.; Dollfus, P.; and Gamrat, C. 2013. Immunity to device variations in a spiking neural network with memristive nanodevices. *IEEE transactions on nanotechnology*, 12(3): 288–295.

Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; and Batra, D. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, 618–626.

Shi, X.; Hao, Z.; and Yu, Z. 2024. SpikingResformer: bridging ResNet and vision transformer in spiking neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5610–5619.

Simons, M.; and Nave, K.-A. 2016. Oligodendrocytes: myelination and axonal support. *Cold Spring Harbor perspectives in biology*, 8(1): a020479.

Simons, M.; and Trajkovic, K. 2006. Neuron-glia communication in the control of oligodendrocyte function and myelin biogenesis. *Journal of cell science*, 119(21): 4381–4389.

Wu, Y.; Deng, L.; Li, G.; Zhu, J.; and Shi, L. 2018. Spatio-temporal backpropagation for training high-performance spiking neural networks. *Frontiers in neuroscience*, 12: 331.

Xin, W.; and Chan, J. R. 2020. Myelin plasticity: sculpting circuits in learning and memory. *Nature Reviews Neuroscience*, 21(12): 682–694.

Yao, M.; Hu, J.; Zhou, Z.; Yuan, L.; Tian, Y.; Bo, X.; and Li, G. 2023a. Spike-driven Transformer. In *Thirty-seventh Conference on Neural Information Processing Systems*.

Yao, M.; Zhao, G.; Zhang, H.; Hu, Y.; Deng, L.; Tian, Y.; Xu, B.; and Li, G. 2023b. Attention spiking neural networks. *IEEE transactions on pattern analysis and machine intelligence*.

Zhou, C.; Zhang, H.; Zhou, Z.; Yu, L.; Huang, L.; Fan, X.; Yuan, L.; Ma, Z.; Zhou, H.; and Tian, Y. 2024. Qkformer: Hierarchical spiking transformer using qk attention. *Advances in Neural Information Processing Systems*, 37: 13074–13098.

Zhou, Z.; Zhu, Y.; He, C.; Wang, Y.; Shuicheng, Y.; Tian, Y.; and Yuan, L. 2023. Spikformer: When Spiking Neural Network Meets Transformer. In *The Eleventh International Conference on Learning Representations*.