

# Incorporating Learnable Membrane Time Constant to Enhance Learning of Spiking Neural Networks

Wei Fang<sup>1,2</sup>, Zhaofei Yu<sup>1,2,3</sup>, Yanqi Chen<sup>1,2</sup>, Timothée Masquelier<sup>4</sup>, Tiejun Huang<sup>1,2,3</sup>, Yonghong Tian<sup>1,2\*</sup>

<sup>1</sup>Department of Computer Science and Technology, Peking University, China

<sup>2</sup>Peng Cheng Laboratory, China

<sup>3</sup>Institute for Artificial Intelligence, Peking University, China

<sup>4</sup>Centre de Recherche Cerveau et Cognition (CERCO), UMR5549 CNRS - Univ. Toulouse 3, France

f fwei, yuzf12, chyq g@pku.edu.cn, timothee.masquelier@cnrs.fr,

f tjuhuang, yhtian g@pku.edu.cn

## Abstract

Spiking Neural Networks (SNNs) have attracted enormous research interest due to temporal information processing capability, low power consumption, and high biological plausibility. However, the formulation of efficient and high-performance learning algorithms for SNNs is still challenging. Most existing learning methods learn weights only, and require manual tuning of the membrane-related parameters that determine the dynamics of a single spiking neuron. These parameters are typically chosen to be the same for all neurons, which limits the diversity of neurons and thus the expressiveness of the resulting SNNs. In this paper, we take inspiration from the observation that membrane-related parameters are different across brain regions, and propose a training algorithm that is capable of learning not only the synaptic weights but also the membrane time constants of SNNs. We show that incorporating learnable membrane time constants can make the network less sensitive to initial values and can speed up learning. In addition, we reevaluate the pooling methods in SNNs and find that max-pooling will not lead to significant information loss and have the advantage of low computation cost and binary compatibility. We evaluate the proposed method for image classification tasks on both traditional static MNIST, Fashion-MNIST, CIFAR-10 datasets, and neuromorphic N-MNIST, CIFAR10-DVS, DVS128 Gesture datasets. The experiment results show that the proposed method outperforms the state-of-the-art accuracy on nearly all datasets, using fewer time-steps. Our codes are available at <https://github.com/fangwei123456/Parametric-Leaky-Integrate-and-Fire-Spiking-Neuron>.

(a) Spiking neuron (b) The membrane potential of a LIF neuron

Figure 1. (a) A Leaky Integrate-and-Fire (LIF) neuron with membrane potential  $V$ , membrane time constant  $\tau$ , input  $I(t)$  and synaptic weight  $w$ . (b) The membrane potential of the LIF neuron when constant input is received. Increasing or decreasing  $\tau$  will stretch the  $v = f(t)$  curve in the  $t$  direction while increasing or decreasing  $w$  will stretch the  $v = f(t)$  curve in the  $V$  direction.

## 1. Introduction

Spiking Neural Networks (SNNs) are viewed as the third generation of neural network models, which are closer to biological neurons in the brain [38]. Together with neuronal and synaptic states, the importance of spike timing is also considered in SNNs. Due to their distinctive properties, such as temporal information processing capability, low power consumption [49], and high biological plausibility [16], SNNs increasingly arouse researchers' great interest in recent years. Nevertheless, it remains challenging to formulate efficient and high-performance learning algorithms for SNNs.

Generally, the learning algorithms for SNNs can be divided into unsupervised learning, supervised learning, reward-based learning, and Artificial Neural Network (ANN) to SNN conversion methodologies. Either way, we find that most existing learning methods only consider learning the synaptic-related parameters like synaptic weights and treat the membrane-related parameters as

\* Corresponding author

hyperparameters. These membrane-related parameters like [29], and STDP-based expectation-maximization algorithm [43, 17]. However, these methods are only suitable for shallow of a single spiking neuron, are typically chosen to be the low SNNs, and the performance is far below state-of-the-art same for all neurons. Note, however, there exist different ANN results.

membrane time constants for spiking neurons across brain regions [39, 9, 30], which are proved to be essential for the representation of working memory and formulation of learning [20, 53]. Thus simply ignoring different time constants in SNNs will limit the heterogeneity of neurons and thus the expressiveness of the resulting SNNs.

In this paper, we propose a training algorithm that is capable of learning not only the synaptic weights but also membrane time constants of SNNs. As illustrated in Fig. 1, we find that adjustments of the synaptic weight and the membrane time constants have different effects on neuronal dynamics. We show that incorporating learnable membrane time constants is able to enhance the learning of SNNs.

The main contributions of this paper can be summarized as follows:

- 1) We propose the backpropagation-based learning algorithm using spiking neurons with learnable membrane parameters, referred to as Parametric Leaky Integrate-and-Fire (PLIF) spiking neurons, which better represent the heterogeneity of neurons and thereby enhancing the expressiveness of the SNNs. We show that the SNNs made of PLIF neurons are more robust to initial values and can learn faster than SNNs made of neurons with a fixed time constant.
- 2) We reevaluate the pooling methods in SNNs and discredit the previous conclusion that max-pooling results in significant information loss. We find that compared to average-pooling, max-pooling is able to better preserve the asynchronous characteristic of neuron firing, as well as reduce the computation cost. Our experiments show that the performance of max-pooling is comparable to average-pooling.
- 3) We evaluate our methods on both traditional static MNIST [32], Fashion-MNIST [59], CIFAR-10 [31] datasets widely used in ANNs as benchmarks, and neuromorphic N-MNIST [44], CIFAR10-DVS [36], DVS128 Gesture [1] datasets that focus on verifying the network's temporal information processing capability. The proposed method exceeds state-of-the-art accuracy on nearly all tested datasets, using fewer time-steps.

## 2. Related Works

**Unsupervised learning of SNNs** The unsupervised learning methods of SNNs are based on biological plausible local learning rules, like Hebbian learning [22] and Spike-Timing-Dependent Plasticity (STDP) [3]. Existing approaches exploited the self-organization principle [56, 11,

**Reward-based learning of SNNs** Reward-based learning of SNNs mimics the way the human brain learns by taking advantage of the reward or punishment signals induced by dopaminergic, serotonergic, cholinergic, or adrenergic neurons [13, 6, 41]. Despite the methods that arise in reinforcement learning, like policy gradient [52, 28], temporal-difference learning [46, 14] and Q-learning [6], some heuristic phenomenological models based on STDP [15, 62] were proposed recently.

**ANN to SNN conversion** ANN to SNN conversion (ANN2SNN) converts a trained non-spiking ANN to an SNN by using the firing rate of each spiking neuron to approximate the corresponding ReLU activation of an analog neuron [24, 7, 50]. It can get near lossless inference results as an ANN [51, 10], but there is a trade-off between accuracy and latency. To improve accuracy, longer inference latency is needed [19]. ANN2SNN is restricted to rate-coding, which loses the processing capability in temporal tasks. As far as we know, ANN2SNN only works for static datasets, not neuromorphic datasets.

**Supervised learning of SNNs** SpikeProp [5] was the first supervised learning method for SNNs based on backpropagation, which used a linear approximation to overcome the non-differentiable threshold-triggered firing mechanism of SNNs. Subsequent works included Tempotron [18], ReSuMe [45], and SPAN [40], but they could only be applied to single-layer SNNs. Recently, the surrogate gradient method was proposed and provided another solution to training multi-layer SNNs [35, 26, 64, 57, 54, 34, 27]. It utilized surrogate derivatives to define the derivative of the threshold-triggered firing mechanism. Thus the SNNs could be optimized with gradient descent algorithms as ANNs. Zenke et al. [63, 42] systematically studied the remarkable robustness of surrogate gradient learning and showed that SNNs optimized by the surrogate gradient methods can achieve competitive performance with ANNs. Compared to ANN2SNN, the surrogate gradient method has no restrictions on simulating time-steps because it is not based on rate-coding [58, 63].

**Spiking neurons and layers of deep SNNs** Spiking neuron and layer models play an essential role in SNNs. Cheng et al. [8] added the lateral interactions between neighboring neurons and get better accuracy and stronger noise-robustness. Zimmer et al. [65] firstly adopt the learnable time constants in LIF neurons for the speech recognition task. Bellec et al. [2] proposed the adaptive threshold spiking neuron to enhance computing and learning capabilities of SNNs, which was improved by [61] with learnable time constants. Rathi et al. [47] suggested using a

learnable membrane leak and ring threshold to netune SNNs converted from ANNs. Despite this, no systematic research on the effects of learning membrane time constants to SNNs has been conducted so far, which is exactly the aim of this paper. Wu et al. [58] found that normalization layers are also critical for deep SNNs and proposed Neuron Normalization (NeuNorm) to balance each neuron's ring rate to avoid severe information loss. Ledinauskas, E et al. [33] rstly suggested that using Batch Normalization [25] in deep SNNs for faster convergence.

### 3. Methods

In this section, we rst brie y review the Leaky Integrate-and-Fire model in Sec. 3.1, and analyze the effect of synaptic weight and membrane time constant in Sec. 3.2. The Parametric Leaky Integrate-and-Fire model and the network structure of the SNNs are then introduced in Sec. 3.3 – Sec. 3.5. At last, we describe the spike max-pooling and the learning algorithm of SNNs in Sec. 3.6 and Sec. 3.7.

#### 3.1. Leaky Integrate-and-Fire model

The basic computing unit of an SNN is the spiking neuron. Neuroscientists have built several spiking neuron models for describing the accurate relationships between input and output signals of the biological neuron. The Leaky Integrate-and-Fire (LIF) model [16] is one of the simplest spiking neuron models used in SNNs. The subthreshold dynamics of the LIF neuron is defined as:

$$\frac{dV(t)}{dt} = -(V(t) - V_{rest}) + X(t); \quad (1)$$

where  $V(t)$  represents the membrane potential of the neuron at time  $t$ ,  $X(t)$  represents the input to neuron at time  $t$ ,  $\tau$  is the membrane time constant, and  $V_{rest}$  is the resting potential. When the membrane potential  $V(t)$  exceeds a certain threshold  $V_{th}$  at time  $t^f$ , the neuron will elicit a spike and then the membrane potential  $V(t)$  goes back to a reset value  $V_{reset} < V_{th}$ . The LIF neuron achieves a balance between computing cost and biological plausibility. We set  $V_{rest} = V_{reset}$  in this paper, and will not make a distinction between them in the rest of this paper.

#### 3.2. Function comparison of synaptic weight and membrane time constant

In most of the previous learning algorithms for SNNs made of LIF neurons, the membrane time constant is regarded as a hyper-parameter and chosen to be the same for all neurons before learning. The learning of SNNs is only to optimize the synaptic weights. However, it cannot be ignored that the behavior of a spiking neuron for given inputs depends not only on the weights of connected synapses but also on the neuron's inherent dynamics controlled by the membrane time constant

Figure 2. The membrane potential of a LIF neuron when instant spikes at  $t = 5; 80; 85; 90$  are received.

In order to compare the effects of synaptic weight and membrane time constant to the neuronal dynamics, we consider a simple case where the LIF neuron receives weighted input  $X(t) = wI(t)$  from a presynaptic neuron  $z_j$  (Fig. 1(a)). The rest potential  $V_{rest}$  is set to 0. When the input is constant, namely  $I(t) = 1$ , the membrane potential of the LIF neuron  $z_i$  changes over time is shown in Fig. 1(b) (blue curve), which is computed according to Eq. (1). Increasing or decreasing  $\tau$  as shown by the  $\tau +$  and  $\tau -$  curves, will stretch the  $v = f(t)$  curve in the  $V$  direction. On the contrary, increasing or decreasing  $w$  will stretch the  $v = f(t)$  curve in the  $t$  direction, and will not change the steady-state voltage of the neuron as  $V(+\infty) = wI$ . Fig. 2 illustrates the response of the neuron  $z_i$  to instant input spikes at time  $t = 5; 80; 85; 90$ ms, namely  $X(t) = w(\delta(t-5) + \delta(t-80) + \delta(t-85) + \delta(t-90))$ . The neuron's response to instant input spikes at  $t = 5$  indicates that a smaller (the  $\tau -$  curve) leads to faster charge to the steady-state voltage and faster decay to the resting value, making the LIF neuron more sensitive to an instant spike. This sensitivity helps the neuron to capture instant variety in the input. In contrast, a smaller (the  $w -$  curve) leads to a slower charge to the steady-state voltage without affecting decaying speed. When there are three successive input spikes, the membrane potential of the neuron with a smaller (the  $\tau -$  curve) will reach a higher value at a faster rate, which makes it easier to re.

To some extent, the effect of decreasing  $\tau$  is similar to that of increasing  $w$ . Nevertheless, adjusting both  $\tau$  and  $w$  can bring some superior additional benefits. As mentioned above, changing both  $\tau$  and  $w$  can stretch the  $v = f(t)$  curve, namely the neuron's response to a given input, in both  $t$  direction and  $V$  direction, which endows the neuron better fitting ability.

#### 3.3. Parametric Leaky Integrate-and-Fire model

We propose the Parametric Leaky Integrate-and-Fire (PLIF) spiking neuron model to learn both the synaptic weights and the membrane time constants of SNNs. The dynamics of the PLIF neuron can be described by Eq. (1). The SNNs with PLIF neurons follow the three rules:

<sup>1</sup>  $R_1(t)$  represents Dirac delta function.  $\forall t \notin 0$ , then  $R_1(t) = 0$ .  $\int_{-\infty}^{\infty} R_1(t) dt = 1$ .

Figure 3. The general discrete spiking neuron model.

(1). The membrane time constant is optimized automatically during training, rather than being set as a hyper-parameter manually before training.

(2). The membrane time constants shared within the neurons in the same layer in SNNs, which is biologically plausible as the neighboring neurons have similar properties.

(3). The membrane time constant of neurons in different layers are distinct, making diverse phase-frequency responsiveness of neurons.

In fact, the proposed rules are able to increase the heterogeneity of neurons and the expressiveness of the resulting SNNs while effectively controlling computation costs.

For numerical simulations of PLIF neurons in SNNs, we need to consider a version of the parameters dynamics that is discrete in time. Specifically, by including the threshold-triggered firing mechanism and the reset of the membrane potential after firing, we can describe the dynamics of all kinds of spiking neurons with the following equations:

$$H_t = f(V_{t-1}; X_t); \quad (2)$$

$$S_t = (H_t - V_{th}); \quad (3)$$

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

To avoid confusion, we use  $H_t$  and  $V_t$  to represent the membrane potential after neuronal dynamics and after the trigger of a spike at time-step  $t$  respectively.  $X_t$  denotes the external input, and  $V_{th}$  denotes the firing threshold.  $S_t$  denotes the output spike at time  $t$  which equals 1 if there is a spike and 0 otherwise. Eq. (3) describes the spike generative process, where  $(x)$  is the Heaviside step function and is defined by  $(x) = 1$  for  $x \geq 0$  and  $(x) = 0$  for  $x < 0$ . Eq. (4) illustrates that the membrane potential returns to  $V_{reset}$  after eliciting a spike, which is called hard reset and widely used in deep SNNs [33].

As shown in Fig. 3, Eqs. (2) - (4) build a general model to describe the discrete spiking neuron's action: charging, firing, and resetting. Specifically, Eq. (2) describes the neuronal dynamics, and different spiking neuron models have different functions  $f(\cdot)$ . For example, the function  $f(\cdot)$  for the LIF neuron and PLIF neuron is

$$H_t = V_{t-1} + \frac{1}{\tau}((V_{t-1} - V_{reset}) + X_t); \quad (5)$$

Figure 4. The general formulation of our networks and its unfolded formulation.  $N_{conv}$  indicates there are  $N_{conv}$  of Conv2d-Spiking Neurons connected sequentially.  $N_{down}$  and  $N_{fc}$  have the same meaning. Note that the network's parameters are shared at all time-steps.

For PLIF neurons, directly optimizing the membrane time constant in Eq. (5) may induce numerical instability as  $\tau$  is in the denominator. Besides, Eq. (5), as the discrete version of Eq. (1), is a valid approximation only when the time-step  $\Delta t$  is smaller than  $\tau$ , that is,  $\tau > 1$ , which is ignored by [47, 61]. To avoid the above problems, we reformulate Eq. (5) to the following equation with a trainable parameter  $a$ :

$$H_t = V_{t-1} + k(a)((V_{t-1} - V_{reset}) + X_t); \quad (6)$$

Here  $k(a)$  denotes the clamp function and  $k(a) \in (0; 1)$ , which ensures that  $\tau = \frac{1}{k(a)} \in (1; +\infty)$ . In our experiments,  $k(a)$  is the sigmoid activation function, that is,  $k(a) = \frac{1}{1 + \exp(-a)}$ .

#### 3.4. RNN-like Expression of LIF and PLIF

The LIF and PLIF neurons have a similar function as recurrent neural networks. Specifically, when  $V_{reset} = 0$ , the neuronal dynamics of the LIF neuron and PLIF neuron (Eq. (5)) can be written as:

$$H_t = (1 - \frac{1}{\tau})V_{t-1} + \frac{1}{\tau}X_t; \quad (7)$$

where the integration progress  $X_t$  makes the LIF and PLIF neurons able to remember current input information, while the leakage progress  $(1 - \frac{1}{\tau})V_{t-1}$  can be seen as forgetting some information from the past. Eq. (7) shows that the balance between remembrance and forgetting is controlled by the membrane time constant which plays an analogous role as the gates in Long Short-Term Memory (LSTM) networks [23].



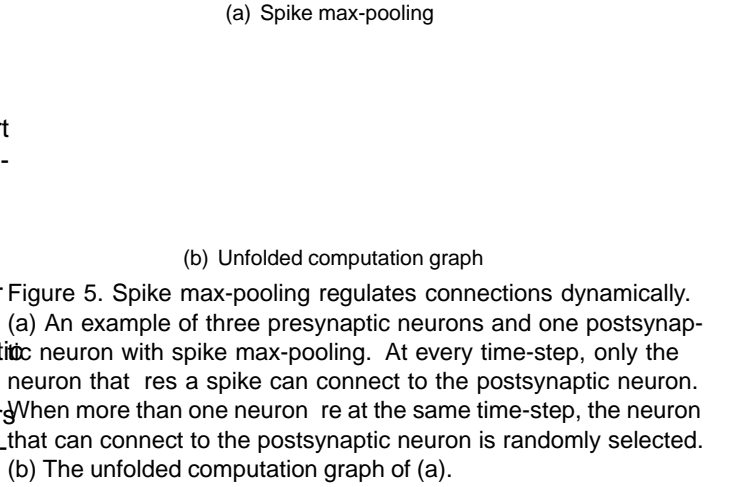
### 3.5. Network Formulation

We propose a general formulation to build SNNs in this paper, which is illustrated in Fig. 4. The SNN includes a spiking encoder network and a classifier network. The spiking encoder network consists of  $N_{\text{down}}$  down-sample modules, each of which contains  $N_{\text{conv}}$  repeated Conv2d-Spiking Neurons and a pooling layer. The spiking encoder can extract features from inputs and convert them into the ring spikes at different time-steps. The classifier network consists of  $N_{\text{fc}}$  repeated FC-Spiking Neurons. Here Conv2d denotes the 2D convolutional layer and FC denotes the fully connected layer. Many previous works [11, 34, 54, 64, 8, 19] used a Poisson encoder to convert images to spikes as input, while [50] suggested that this encoding introduces variability into the ring of the network and impairs its performance. Similar to [50, 58, 47], the input is directly fed to our network without being firstly converted to spikes. In this situation, the image-spike encoding is done by the first Conv2d-Spiking Neurons module, which can be seen as a learnable encoder. Note that synaptic connections, including convolutional layers and fully connected layers, are stateless, while the spiking neuron layers have self-connections in the temporal domain, as the unfolded network formulation shown in Fig. 4. All parameters are shared at all time-steps.

### 3.6. Spike Max-Pooling

The pooling layer is widely used to reduce the size of feature maps and to extract compact representation in convolutional ANNs, as well as SNNs. Most previous studies [51, 8, 48] preferred to use the average-pooling in SNNs as they found that max-pooling in SNNs leads to significant information loss. We argue that the max-pooling is consistent with the SNNs' temporal information processing ability and can increase SNNs' timing capability in temporal tasks and reduce the computation cost for the next layer.

Specifically, the max-pooling layers are behind spiking neuron layers in our model (Fig. 4), and the max-pooling operation is carried on spikes. Different from all neurons that transmit information to the next layer equally in the average-pooling window, only the neuron that fires a spike in the max-pooling window can transmit information to the next layer. Therefore, the max-pooling layer introduces the winner-take-all mechanism, allowing the fired neuron to communicate with the next layer and ignoring other neurons in the pooling window. Another attractive property is that the max-pooling layer will regulate connections dynamically (Fig. 5). The spiking neuron's membrane potential  $V_t$  will return to  $V_{\text{reset}}$  after firing a spike. It is hard for a spiking neuron to fire again as recharging needs time. However, if the neurons in the max-pooling window fire asynchronously, they will be connected to the postsynaptic neuron in turn, which makes the postsynaptic neuron re-



sembles to connect a continuously firing presynaptic neuron and easier to fire. The winner-take-all mechanism in the spatial domain and time-variant topology in the temporal domain achieved by max-pooling can increase SNNs' timing capability in temporal tasks, such as classifying the CIFAR10-DVS dataset. It is worth noting that the outputs of the max-pooling layer are still binary, while the outputs of the average-pooling layer are float. The matrix multiplication and element-wise multiplication operation on spikes can be accelerated by replacing multiplication with logical AND, which is also the advantage of SNNs compared with ANNs.

### 3.7. Training Framework

Here we combine the neuron model (Fig. 3) and network formulation (Fig. 4) to drive the backpropagation training algorithm for SNNs. Denote the simulating time-step as  $T$  and classes number as  $C$ , the output  $O = [o_{t,i}]$  is a  $C \times T$  tensor. For a given input with label  $l$ , we encourage the neuron that represents class  $l$  to have the highest excitatory level while other neurons should remain silent. So the target output is defined by  $Y = [y_{t,i}]$  with  $y_{t,i} = 1$  for  $i = l$  and  $y_{t,i} = 0$  for  $i \neq l$ . The loss function is defined by the mean squared error (MSE)  $\text{MSE}(O; Y) = \frac{1}{T} \sum_{t=0}^{T-1} L_t = \frac{1}{T} \sum_{t=0}^{T-1} \frac{1}{C} \sum_{i=0}^{C-1} (o_{t,i} - y_{t,i})^2$ . And the predicted label  $\hat{l}$  is regarded as the index of the neuron with the maximum firing rate  $\hat{l}_p = \arg \max_i \frac{1}{T} \sum_{t=0}^{T-1} o_{t,i}$ .

Here we suppose that  $\theta_t^i$  represents the learnable parameter of the PLIF neurons in the  $i$ -th layer in the network. At time-step  $t$ , the vector  $\mathbf{H}_t^i$  and  $\mathbf{V}_t^i$  represent the membrane potential after neuronal dynamics and after reset, the vector  $\mathbf{V}_{th}^i$  and  $\mathbf{V}_{reset}^i$  represents the threshold and reset potential, respectively. The weighted inputs from the previous layer are  $\mathbf{X}_t^i = \mathbf{W}^{i-1} \mathbf{I}_t^{i-1}$ .  $\mathbf{S}_t^i = [\mathbf{s}_{t,j}^i]$  denotes the output spike at time-step  $t$ , where  $\mathbf{s}_{t,j}^i = 1$  if the  $j$ -th neuron res a spike, else  $\mathbf{s}_{t,j}^i = 0$ . The gradients backward from the next layer are  $\frac{\partial \mathcal{L}}{\partial \mathbf{S}_t^i}$ . According to Fig. 3 and Fig. 4, we can calculate the gradients recursively:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{H}_t^i} = \frac{\partial \mathcal{L}}{\partial \mathbf{H}_{t+1}^i} \frac{\partial \mathbf{H}_{t+1}^i}{\partial \mathbf{H}_t^i} + \frac{\partial \mathcal{L}}{\partial \mathbf{H}_t^i} \quad (8)$$

$$\frac{\partial \mathbf{H}_{t+1}^i}{\partial \mathbf{H}_t^i} = \frac{\partial \mathbf{H}_{t+1}^i}{\partial \mathbf{V}_t^i} \frac{\partial \mathbf{V}_t^i}{\partial \mathbf{H}_t^i} \quad (9)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{H}_t^i} = \frac{\partial \mathcal{L}}{\partial \mathbf{S}_t^i} \frac{\partial \mathbf{S}_t^i}{\partial \mathbf{H}_t^i} \quad (10)$$

According to Eq. (6), Eq. (3), and Eq. (4) we can get

$$\frac{\partial \mathbf{H}_{t+1}^i}{\partial \mathbf{V}_t^i} = 1 - k(a^i) \quad (11)$$

$$\frac{\partial \mathbf{V}_t^i}{\partial \mathbf{H}_t^i} = 1 - \mathbf{S}_t^i + (\mathbf{V}_{reset}^i - \mathbf{H}_t^i) \frac{\partial \mathbf{S}_t^i}{\partial \mathbf{H}_t^i} \quad (12)$$

$$\frac{\partial \mathbf{S}_t^i}{\partial \mathbf{H}_t^i} = \phi(\mathbf{H}_t^i - \mathbf{V}_{th}^i) \quad (13)$$

$$\frac{\partial \mathbf{H}_t^i}{\partial \mathbf{X}_t^i} = k(a^i) \quad (14)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{a}} = (\mathbf{V}_{t-1}^i - \mathbf{V}_{reset}^i) + \mathbf{X}_t^i k'(a^i) + \frac{\partial \mathbf{H}_t^i}{\partial \mathbf{V}_{t-1}^i} \frac{\partial \mathbf{V}_{t-1}^i}{\partial \mathbf{H}_{t-1}^i} \frac{\partial \mathbf{H}_{t-1}^i}{\partial \mathbf{a}} \quad (15)$$

Finally, we can get the gradients of the learnable parameters:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{a}} = \sum_{t=0}^{T-1} \frac{\partial \mathcal{L}}{\partial \mathbf{H}_t^i} \frac{\partial \mathbf{H}_t^i}{\partial \mathbf{a}} \quad (16)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{W}^{i-1}} = \sum_{t=0}^{T-1} \frac{\partial \mathcal{L}}{\partial \mathbf{H}_t^i} \frac{\partial \mathbf{H}_t^i}{\partial \mathbf{X}_t^i} \mathbf{I}_t^{i-1} \quad (17)$$

Note that  $\frac{\partial \mathcal{L}}{\partial \mathbf{S}_t^i} = 0$  when  $t = T$ ,  $\mathbf{V}_{t-1}^i = \mathbf{V}_{reset}^i$ . We use derivative of the surrogate function  $\phi(x)$  to define the derivative of spiking function  $\phi(x)$  (see supplementary).  $k(x)$  is the clamp function.

## 4. Experiments

We evaluate the performance of SNNs with PLIF neurons and spike max-pooling for classification tasks on

Dataset	$N_{conv}$	$N_{down}$	$N_{fc}$
*MNIST	1	2	2
CIFAR-10	3	2	2
CIFAR10-DVS	1	4	2
DVS128 Gesture	1	5	2

Table 1. Network structures for different datasets.  $N_{conv}$ ,  $N_{down}$  and  $N_{fc}$  are defined in Fig. 4. \*MNIST denotes MNIST, Fashion-MNIST and N-MNIST datasets.

both traditional static MNIST, Fashion-MNIST, CIFAR-10 datasets, and neuromorphic N-MNIST, CIFAR10-DVS, and DVS128 Gesture datasets. More details of the training can be found in the supplementary.

### 4.1. Network Structure

The network structures of SNNs for different datasets are shown in Tab. 1. We set kernel size = 3, stride = 1 and padding = 1 for all Conv2d layers. The out channels of Conv2d layers is 256 for CIFAR-10 dataset and 128 for all other datasets. A batch normalization (BN) layer is added after each Conv2d layer. As the parameters of BN layer can be absorbed in its from Conv2d layer [50], we can remove BN in the SNNs for inference. All pooling layers set kernel size = 2 and stride = 2. For all networks, the out features of the first FC layer is a quarter of the in features, and the out features of the second FC layer is  $M/C$ , where  $C$  is the classes number and  $M$  is the neurons of a population to represent one class. A dropout layer [34] is placed before each FC layer. A voting layer after the output spiking neurons layer is used to boost classifying robustness. The voting layer is implemented by average-pooling with kernel size =  $M$  and stride =  $M$ . We set  $M = 10$  for all datasets. We use the average-pooling to implement democratic voting, such that the minority is subordinate to the majority. Using max-pooling to vote may result in a dictatorship, as the minority will not be involved in the computation graph (see Fig. 5) and using neurons to represent one class will degenerate into using one neuron.

### 4.2. Comparison with the State-of-the-Art

Tab. 2 shows the accuracies of the proposed methods (PLIF neurons with  $\phi_0 = 2$ , max-pooling) and other comparing methods on both traditional static MNIST, Fashion-MNIST, CIFAR-10 datasets, and neuromorphic N-MNIST, CIFAR10-DVS, DVS128 Gesture datasets. We set the same training hyperparameters for all datasets (see supplementary). As shown in Tab. 2, we achieve the highest accuracies on all datasets except for CIFAR-10. The accuracy on CIFAR-10 is slightly lower than [19], which is based on ANN2SNN conversion. However, they only applied to static images as ANN2SNN is ill-suited to neuromorphic datasets. Different from them, our method is also applicable to neuromorphic datasets and outperforms the spike-based

Model	Method	Accuracy MNIST	Accuracy Fashion-MNIST	Accuracy CIFAR-10	Accuracy N-MNIST	Accuracy CIFAR10-DVS	Accuracy DVS128 Gesture
[24]	ANN2SNN	98.37%	-	82.95%	-	-	-
[50]	ANN2SNN	99.44%	-	88.82%	-	-	-
[51]	ANN2SNN	-	-	91.55%	-	-	-
[19]	ANN2SNN	-	-	93.63%	-	-	-
[35]	Spike-based BP	99.31%	-	-	98.74%	-	-
[57]	Spike-based BP	99.42%	-	-	98.78%	50.7%	-
[54]	Spike-based BP	99.36%	-	-	99.2%	-	93.64%
[27]	Spike-based BP	-	-	-	96%	-	95.54%
[26]	Spike-based BP	99.49%	-	-	98.84%	-	-
[64]	Spike-based BP	99.62%	90.13%	-	-	-	-
[58]	Spike-based BP	-	-	90.53%	99.53%	60.5%	-
[34]	Spike-based BP	99.59%	-	90.95%	99.09%	-	-
[8]	Spike-based BP	99.5%	92.07%	-	99.45%	-	-
[37]	Spike-based BP	-	-	-	96.3%	32.2%	-
[60]	Spike-based BP	-	-	-	-	-	92.01%
[12]	Spike-based BP	99.46%	-	-	99.39%	-	96.09%
[21]	Spike-based BP	-	-	-	98.28%	-	(10 classes) 93.40%
[47]	ANN2SNN and Spike-based BP	-	-	92.64%	-	-	-
[55]	HATS	-	-	-	99.1%	52.4%	-
[4]	GCN	-	-	-	99.0%	54.0%	-
Ours	Spike-based BP	99.72%	94.38%	93.50%	99.61%	74.80%	97.57%

Table 2. Performance comparison between the proposed method and the state-of-the-art methods on different datasets. The highest accuracies of previous works are in bold.

Dataset	SOTA	SOTA's T	ours T
MNIST	[64]	400	8
Fashion-MNIST	[8]	20	8
CIFAR-10	[19]	2048	8
N-MNIST	[58]	59-64	10
CIFAR10-DVS	[58]	230-292	20
DVS128 Gesture	[27]	500(training) 1800(testing)	20

Table 3. The time-steps of previous SOTA works and ours on each dataset.

Neuron	Fashion-MNIST	CIFAR-10	CIFAR10-DVS	DVS128 Gesture
PLIF( $\tau_0 = 2$ )	94.38%	93.50%	74.80%	97.57%
LIF( $\tau = 2$ )	94.17%	93.03%	73.60%	96.88%
PLIF( $\tau_0 = 16$ )	94.65%	93.23%	70.50%	92.01%
LIF( $\tau = 16$ )	94.47%	47.50%	62.40%	76.74%

Table 4. Accuracy of using PLIF/LIF.

BP SOTA accuracy.

Tab. 3 compares the number of time-steps of our method and the previous works that achieve the best performance on each dataset. It can be found that the proposed method takes fewer time-steps than all the other methods. For example, our method uses up to 256 fewer inference time-steps compared to ANN2SNN conversion [19]. Thus our method can not only decrease the memory consumption and the training time but also increase inference speed greatly.

### 4.3. Ablation Study

We conduct extensive ablation studies to evaluate PLIF neurons and max-pooling on four challenging datasets. We

first study the effect of PLIF neurons. In this experiment, we train the same SNNs with PLIF neurons and LIF neurons respectively, and compare the test accuracy. As shown in Tab. 4, if the initial membrane time constant of PLIF neurons is set equal to the membrane time constant of LIF neurons, the test accuracy of the SNNs with PLIF neurons is always higher than that with LIF neurons. This is due to the membrane time constants of PLIF neurons in different layers can be different after learning, which better represents the heterogeneity of neurons. Fig. 6 illustrates the test accuracy of PLIF vs. LIF neurons during training. As can be seen, the accuracy and convergence speed of the SNNs with LIF neurons decrease seriously if the initial value of the membrane time constant is not reasonable (red curve). In contrast, the PLIF neurons can learn the appropriate membrane time constants and achieve better performance (green curve).

To analyze the influence of initial values in PLIF neu-

Pooling	Fashion-MNIST	CIFAR-10	CIFAR10-DVS	DVS128 Gesture
Average	94.74%	93.30%	72.70%	97.22%
Max	94.38%	93.50%	74.80%	97.57%

Table 5. Accuracy of using max-pooling/average-pooling.

(a) Fashion-MNIST

(b) CIFAR-10

(c) CIFAR10-DVS

(d) DVS128 Gesture

Figure 6. The test accuracy of PLIF vs. LIF neurons on different datasets during training. The shaded curves indicate the original data. The solid curves are 64-epoch moving averages.

(a) The change of  $\tau_i$  during training on CIFAR-10.

(b) The change of  $\tau_i$  during training on CIFAR10-DVS.

Figure 7. The change of membrane time constants in different layers during training with different initial values.  $\tau_i$  represents the membrane time constant of the  $i$ -th PLIF neurons layer.

rons, we show how the membrane time constants of the neurons in each layer change during learning with respect to different initial values. As shown in Fig. 7, the membrane time constants with different initial values in each layer tend to gather during training, which indicates that the

PLIF neurons are robust to initial values. Note that (4) in Fig. 7(a) and (4) in Fig. 7(b) tend to infinity. This could be explained as follows. The PLIF neurons with the membrane time constants (4) and (6) in two SNNs are behind the first FC layer with weight  $W_{fc}$ . We check the training logs and find that the distribution, mean and variance of  $\frac{W_{fc}}{\tau_i}$  ( $\tau_i = (4)$  or  $(6)$ ) converge after dozens of epochs (see supplementary). Refer to the dynamics of PLIF neurons (Eq. (5)) with  $X_t = W_{fc} I_t$  and  $\tau_i \neq 0$ , we can find  $H_t \neq V_{t-1} + \frac{W_{fc}}{\tau_i} I_t$ . It means that the PLIF neurons after the first FC layer are learning to become the Non-Leaky-Integrate-and-Fire neurons.

We further study the effect of max-pooling. Tab. 5 compares the accuracy of the proposed SNNs with max-pooling/average-pooling on four challenging datasets. The performance of max-pooling is similar to that of average-pooling, which indicates that the previous conclusion that max-pooling results in significant information loss in SNNs is not reasonable. Remarkably, the max-pooling gets slightly higher accuracies on CIFAR-10, CIFAR10-DVS, and DVS128 Gesture datasets, showing its better fitting capability in complex tasks.

## 5. Conclusion

In this work, we proposed the Parametric Leaky Integrate-and-Fire (PLIF) neuron to incorporate the learnable membrane time parameter into SNNs. We show that the SNNs with the PLIF neurons outperform state-of-the-art comparing methods on both static and neuromorphic datasets. Besides, we show that the SNNs made of PLIF neurons are more robust to initial values and can learn faster than SNNs consist of LIF neurons. We also reevaluate the performance of max-pooling and average-pooling in SNNs and find the previous works underestimate the performance of max-pooling. We recommend using max-pooling in SNNs for its lower computation cost, higher temporal fitting capability, and the characteristic to receive spikes and output spikes rather than outputting values as average-pooling.

## 6. Acknowledgment

This work is supported by grants from the National Natural Science Foundation of China under contracts No.62027804, No.61825101, and No.62088102.



## References

- [1] Arnon Amir, Brian Taba, David Berg, Timothy Melano, Jeffrey McKinstry, Carmelo Di Nolfo, Tapan Nayak, Alexander Andreopoulos, Guillaume Garreau, Marcela Mendoza, et al. A low power, fully event-based gesture recognition system. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7243–7252, 2017. [2](#)
- [2] Guillaume Bellec, Darjan Salaj, Anand Subramoney, Robert Legenstein, and Wolfgang Maass. Long short-term memory and learning-to-learn in networks of spiking neurons. In *Advances in Neural Information Processing Systems*, pages 787–797, 2018. [2](#)
- [3] Guo-qiang Bi and Mu-ming Poo. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience* 18(24):10464–10472, 1998. [2](#)
- [4] Yin Bi, Aaron Chadha, Alhabib Abbas, Eirina Bourtsoulatz, and Yiannis Andreopoulos. Graph-based object classification for neuromorphic vision sensing. *Proceedings of the IEEE International Conference on Computer Vision*, pages 491–501, 2019. [7](#)
- [5] Sander M Bohte, Joost N Kok, and Han La Poutre. Error-backpropagation in temporally encoded networks of spiking neurons. *Neurocomputing* 48(1-4):17–37, 2002. [2](#)
- [6] Matthew Botvinick, Jane X. Wang, Will Dabney, Kevin J. Miller, and Zeb Kurth-Nelson. Deep reinforcement learning and its neuroscientific implications. *Neuron* 107(4):603–616, 2020. [2](#)
- [7] Yongqiang Cao, Yang Chen, and Deepak Khosla. Spiking deep convolutional neural networks for energy-efficient object recognition. *International Journal of Computer Vision* 113(1):54–66, 2015. [2](#)
- [8] Xiang Cheng, Yunzhe Hao, Jiaming Xu, and Bo Xu. LISNN: Improving Spiking Neural Networks with Lateral Interactions for Robust Object Recognition. In *ICAI*, pages 1519–1525. International Joint Conferences on Artificial Intelligence Organization, 7 2020. [2](#), [5](#), [7](#)
- [9] Gustavo Deco, Josephine Cruzat, and Morten L Kringelbach. Brain songs framework used for discovering the relevant timescale of the human brain. *Nature Communications* 10(1):1–13, 2019. [2](#)
- [10] Shikuang Deng and Shi Gu. Optimal conversion of conventional artificial neural networks to spiking neural networks. In *International Conference on Learning Representations* 2021. [2](#)
- [11] Peter U Diehl and Matthew Cook. Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Frontiers in Computational Neuroscience* 9:99, 2015. [2](#), [5](#)
- [12] Haowen Fang, Amar Shrestha, Ziyi Zhao, and Qinru Qiu. Exploiting Neuron and Synapse Filter Dynamics in Spatial Temporal Learning of Deep Spiking Neural Networks. *arXiv preprint arXiv:2003.02944*, 2020. [7](#)
- [13] Nicolas F  maux and Wulfram Gerstner. Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Frontiers in Neural Circuits* 9:85, 2016. [2](#)
- [14] Nicolas F  maux, Henning Sprekeler, and Wulfram Gerstner. Reinforcement learning using a continuous time actor-critic framework with spiking neurons. *PLoS Computational Biology*, 9(4):e1003024, 2013. [2](#)
- [15] Johannes Friedrich, Robert Urbanczik, and Walter Senn. Spatio-temporal credit assignment in neuronal population learning. *PLoS Computational Biology* 7(6):e1002092, 2011. [2](#)
- [16] Wulfram Gerstner, Werner M Kistler, Richard Naud, and Liam Paninski. *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press, 2014. [1](#), [3](#)
- [17] Shangqi Guo, Zhaofei Yu, Fei Deng, Xiaolin Hu, and Feng Chen. Hierarchical bayesian inference and learning in spiking neural networks. *IEEE Transactions on Cybernetics* 49(1):133–145, 2017. [2](#)
- [18] Robert G  tig and Haim Sompolinsky. The tempotron: a neuron that learns spike timing-based decisions. *Nature Neuroscience* 9(3):420–428, 2006. [2](#)
- [19] Bing Han, Gopalakrishnan Srinivasan, and Kaushik Roy. RMP-SNN: Residual Membrane Potential Neuron for Enabling Deeper High-Accuracy and Low-Latency Spiking Neural Network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 13558–13567, 2020. [2](#), [5](#), [6](#), [7](#)
- [20] Michael E Hasselmo and Chantal E Stern. Mechanisms underlying working memory for novel information. *Trends in Cognitive Sciences* 10(11):487–493, 2006. [2](#)
- [21] Weihua He, YuJie Wu, Lei Deng, Guoqi Li, Haoyu Wang, Yang Tian, Wei Ding, Wenhui Wang, and Yuan Xie. Comparing SNNs and RNNs on Neuromorphic Vision Datasets: Similarities and Differences. *arXiv preprint arXiv:2005.02183*, 2020. [7](#)
- [22] Donald Olding Hebb. *The organization of behavior: A neuropsychological theory*. Psychology Press, 2005. [2](#)
- [23] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation* 9(8):1735–1780, 1997. [4](#)
- [24] Eric Hunsberger and Chris Eliasmith. Spiking deep networks with LIF neurons. *arXiv preprint arXiv:1510.08829*, 2015. [2](#), [7](#)
- [25] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015. [3](#)
- [26] Yingyezhe Jin, Wenrui Zhang, and Peng Li. Hybrid macro/micro level backpropagation for training deep spiking neural networks. In *Advances in Neural Information Processing Systems*, pages 7005–7015, 2018. [2](#), [7](#)
- [27] Jacques Kaiser, Hesham Mostafa, and Emre Neftci. Synaptic Plasticity Dynamics for Deep Continuous Local Learning (DECOLLE). *Frontiers in Neuroscience* 14:424, 2020. [2](#), [7](#)
- [28] David Kappel, Robert Legenstein, Stefan Habenschuss, Michael Hsieh, and Wolfgang Maass. A dynamic connectome supports the emergence of stable computational function of neural circuits through reward-based learning. *eNeuro* 5(2), 2018. [2](#)
- [29] Saeed Reza Kheradpisheh, Mohammad Ganjtabesh, Simon J Thorpe, and Timoth  e Masquelier. Stpd-based spiking deep convolutional neural networks for object recognition. *Neural Networks* 99:56–67, 2018. [2](#)

- [30] Christof Koch, Moshe Rapp, and Idan Segev. A brief history of time (constants) *Cerebral Cortex* 6(2):93–101, 1996. 2
- [31] Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. Technical report, 2009. 2
- [32] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11):2278–2324, 1998. 2
- [33] Eimantas Ledinauskas, Julius Ruseckas, Alfonsas Sedas, and Giedrius Buzas. Training Deep Spiking Neural Networks. *arXiv preprint arXiv:2006.04436* 2020. 3, 4
- [34] Chankyu Lee, Syed Shakib Sarwar, Priyadarshini Panda, Gopalakrishnan Srinivasan, and Kaushik Roy. Enabling spike-based backpropagation for training deep neural network architectures *Frontiers in Neuroscience* 14, 2020. 2, 5, 6, 7
- [35] Jun Haeng Lee, Tobi Delbruck, and Michael Pfeiffer. Training deep spiking neural networks using backpropagation. *Frontiers in Neuroscience* 10:508, 2016. 2, 7
- [36] Hongmin Li, Hanchao Liu, Xiangyang Ji, Guoqi Li, and Luping Shi. Cifar10-dvs: an event-stream dataset for object classification. *Frontiers in Neuroscience* 11:309, 2017. 2
- [37] Qianhui Liu, Haibo Ruan, Dong Xing, Huajin Tang, and Gang Pan. Effective AER Object Classification Using Segmented Probability-Maximization Learning in Spiking Neural Networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1308–1315, 2020. 7
- [38] Wolfgang Maass. Networks of spiking neurons: the third generation of neural network models *Neural Networks* 10(9):1659–1671, 1997. 1
- [39] Maurizio Mattia and Paolo Del Giudice. Population dynamics of interacting spiking neurons. *Physical Review E* 66(5):051917, 2002. 2
- [40] Ammar Mohemmed, Stefan Schliebs, Satoshi Matsuda, and Nikola Kasabov. Span: Spike pattern association neuron for learning spatio-temporal spike patterns. *International Journal of Neural Systems* 22(04):1250012, 2012. 2
- [41] Milad Mozafari, Mohammad Ganjtabesh, Abbas Nowzari-Dalini, Simon J Thorpe, and Timothy J Masquelier. Bio-inspired digit recognition using reward-modulated spike-timing-dependent plasticity in deep convolutional networks. *Pattern Recognition* 94:87–95, 2019. 2
- [42] Emre O Neftci, Hesham Mostafa, and Friedemann Zenke. Surrogate gradient learning in spiking neural networks. *IEEE Signal Processing Magazine* 36:61–63, 2019. 2
- [43] Bernhard Nessler, Michael Pfeiffer, Lars Buesing, and Wolfgang Maass. Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. *PLoS Computational Biology* 9(4):e1003037, 2013. 2
- [44] Garrick Orchard, Ajinkya Jayawant, Gregory K Cohen, and Nitish Thakor. Converting static image datasets to spiking neuromorphic datasets using saccades. *Frontiers in Neuroscience* 9:437, 2015. 2
- [45] Filip Ponulak and Andrzej Kasicki. Supervised learning in spiking neural networks with ReSuMe: sequence learning, classification, and spike shifting. *Neural Computation* 22(2):467–510, 2010. 2
- [46] Wiebke Potjans, Markus Diesmann, and Abigail Morrison. An imperfect dopaminergic error signal can drive temporal-difference learning. *PLoS Computational Biology* 7(5):e1001133, 2011. 2
- [47] Nitin Rathi and Kaushik Roy. DIET-SNN: Direct Input Encoding With Leakage and Threshold Optimization in Deep Spiking Neural Networks. *arXiv preprint arXiv:2008.03658* 2020. 2, 4, 5, 7
- [48] Nitin Rathi, Gopalakrishnan Srinivasan, Priyadarshini Panda, and Kaushik Roy. Enabling deep spiking neural networks with hybrid conversion and spike timing dependent backpropagation. *arXiv preprint arXiv:2005.01807* 2020. 5
- [49] Kaushik Roy, Akhilesh Jaiswal, and Priyadarshini Panda. Towards spike-based machine intelligence with neuromorphic computing. *Nature* 575(7784):607–617, 2019. 1
- [50] Bodo Rueckauer, Iulia-Alexandra Lungu, Yuhuang Hu, Michael Pfeiffer, and Shih-Chii Liu. Conversion of continuous-valued deep networks to efficient event-driven networks for image classification. *Frontiers in Neuroscience* 11:682, 2017. 2, 5, 6, 7
- [51] Abhronil Sengupta, Yuting Ye, Robert Wang, Chiao Liu, and Kaushik Roy. Going deeper in spiking neural networks: Vgg and residual architectures. *Frontiers in Neuroscience* 13:95, 2019. 2, 5, 7
- [52] H Sebastian Seung. Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron* 40(6):1063–1073, 2003. 2
- [53] Karthik H Shankar and Marc W Howard. A scale-invariant internal representation of time. *Neural Computation*, 24(1):134–193, 2012. 2
- [54] Sumit B Shrestha and Garrick Orchard. Slayer: Spike layer error reassignment in time. *Advances in Neural Information Processing Systems* 31:1412–1421, 2018. 2, 5, 7
- [55] Amos Sironi, Manuele Brambilla, Nicolas Bourdis, Xavier Lagorce, and Ryad Benosman. HATS: Histograms of averaged time surfaces for robust event-based object classification. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1731–1740, 2018. 7
- [56] Narayan Srinivasa and Youngkwan Cho. Self-organizing spiking neural model for learning fault-tolerant spatio-motor transformations. *IEEE Transactions on Neural Networks and Learning Systems* 23(10):1526–1538, 2012. 2
- [57] Yujie Wu, Lei Deng, Guoqi Li, Jun Zhu, and Luping Shi. Spatio-temporal backpropagation for training high-performance spiking neural networks. *Frontiers in Neuroscience* 12:331, 2018. 2, 7
- [58] Yujie Wu, Lei Deng, Guoqi Li, Jun Zhu, Yuan Xie, and Luping Shi. Direct training for spiking neural networks: Faster, larger, better. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1311–1318, 2019. 2, 3, 5, 7
- [59] Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747* 2017. 2

- [60] Yannan Xing, Gaetano Di Caterina, and John Soraghan. A new spiking convolutional recurrent neural network (scrnn) with applications to event-based hand gesture recognition. *Frontiers in Neuroscience* 14:1143, 2020. 7
- [61] Bojian Yin, Federico Corradi, and Sander M Böhnt Effective and Efficient Computation with Multiple-timescale Spiking Recurrent Neural Networks. arXiv preprint arXiv:2005.11633 2020. 2, 4
- [62] Mengwen Yuan, Xi Wu, Rui Yan, and Huajin Tang. Reinforcement Learning in Spiking Neural Networks with Stochastic and Deterministic Synapses. *Neural Computation*, 31(12):2368–2389, 2019. 2
- [63] Friedemann Zenke and Tim P Vogels. The remarkable robustness of surrogate gradient learning for instilling complex function in spiking neural networks. *BioRxiv*, 2020. 2
- [64] Wenrui Zhang and Peng Li. Spike-train level backpropagation for training deep recurrent spiking neural networks. In *Advances in Neural Information Processing Systems* pages 7802–7813, 2019. 2, 5, 7
- [65] Romain Zimmer, Thomas Pellegrini, Srisht Fateh Singh, and Timothée Masquelier. Technical report: supervised training of convolutional spiking neural networks with PyTorch. arXiv preprint arXiv:1911.10124 2019. 2











