Efficient Spiking Neural Networks with Sparse Selective Activation for Continual Learning

Jiangrong Shen[1, 2], **Wenyao Ni**[2], **Qi Xu**[3], **Huajin Tang**[1, 2, 4*]
[1]State Key Lab of Brain-Machine Intelligence, Zhejiang University, China
[2]The College of Computer Science and Technology, Zhejiang University, China
[3]School of Computer Science and Technology, Dalian University of Technology, China
[4]MOE Frontier Science Center for Brain Science and Brain-Machine Integration, Zhejiang University, Hangzhou, China
jrshen@zju.edu.cn, 22221209@zju.edu.cn, xuqi@dlut.edu.cn, htang@zju.edu.cn

Abstract

The next generation of machine intelligence requires the capability of continual learning to acquire new knowledge without forgetting the old one while conserving limited computing resources. Spiking neural networks (SNNs), compared to artificial neural networks (ANNs), have more characteristics that align with biological neurons, which may be helpful as a potential gating function for knowledge maintenance in neural networks. Inspired by the selective sparse activation principle of context gating in biological systems, we present a novel SNN model with selective activation to achieve continual learning. The trace-based K-Winner-Take-All (K-WTA) and variable threshold components are designed to form the sparsity in selective activation in spatial and temporal dimensions of spiking neurons, which promotes the subpopulation of neuron activation to perform specific tasks. As a result, continual learning can be maintained by routing different tasks via different populations of neurons in the network. The experiments are conducted on MNIST and CIFAR10 datasets under the class incremental setting. The results show that the proposed SNN model achieves competitive performance similar to and even surpasses the other regularization-based methods deployed under traditional ANNs.

Introduction

Biological organisms are capable of learning continually from interactions with their environments throughout their lifespan (Kandel and Hawkins 1992). Machine intelligence with artificial neural networks (ANNs) has exhibited remarkable capabilities in computer vision and natural language processing applications (LeCun, Bengio, and Hinton 2015). However, the new generation of applications, such as self-driving cars and wearable devices, require new machine intelligence that can acquire new knowledge without forgetting the old one while conserving limited computing resources (Kudithipudi et al. 2022). Inspired by biological systems, spiking neural networks (SNNs) have exhibited more complex spatiotemporal dynamics in comparison to ANNs (Maass 1997; Subbulakshmi Radhakrishnan et al. 2021; Yin, Corradi, and Bohte 2023; Bu et al. 2023), and have the potential to implement the next generation of machine intelligence with low power consumption by combining with neuromorphic hardware (Imam and Cleland 2020; Pei et al. 2019; Deng et al. 2021). Hence, in this paper, we explore how to implement continual learning using SNNs by leveraging the spatiotemporal dynamics specificity in SNNs.

In biological systems, the context has been found to have a significant influence on modulating, filtering, and assimilating new information (Kay and Laurent 1999; Levinson et al. 2020). Context gating facilitates the selective activation of subpopulations of neurons, thereby encouraging the reduction of memory interference between similar experiences (Kudithipudi et al. 2022). Coincidentally, the SNNs model exhibits the relational property with the context gating mechanism. On the one hand, SNNs have significant sparsity owing to the expansion of the temporal dimension and the discrete nature of information transmission termed as "spike". This characteristic is advantageous in mitigating interference between neurons with distinct functionalities (Shen et al. 2023). On the other hand, the synaptic weight of each postsynaptic neuron undergoes update only when the corresponding connected presynaptic neuron's membrane potential reaches the firing threshold and emits a spike, while even small activation values in ANN can arise weight updates (Hammouamri, Masquelier, and Wilson 2022). Hence, the aforementioned sparsity in two different levels of neural activity and synapse plasticity could be helpful for SNNs model to reduce memory interference and mitigate the catastrophic forgetting (McCloskey and Cohen 1989) (French 1999) during the continual learning scenario.

In this paper, we explore how to alleviate catastrophic forgetting by enhancing the neural dynamics characteristics in SNN. The selective activation SNNs (SA-SNN) model for continual learning with trace-based K-Winner-Take-All (K-WTA) and variable threshold mechanisms is proposed to alleviate catastrophic forgetting by enhancing the neural dynamics characteristics in SNN, which does not need task labels or memory replay. In the SA-SNN model, we first adopt a biologically plausible, temporal trace-based K-WTA method to reduce interference between different tasks. The trace-based K-WTA method itself converges with the connectivity of many brain regions that utilize inhibitory interneurons and we further modify it to accommodate multi-step spiking neurons. Then we design a simple but effective variable threshold method to modify the threshold of spiking neurons, which enables to encourage the participation of

*Corresponding author.

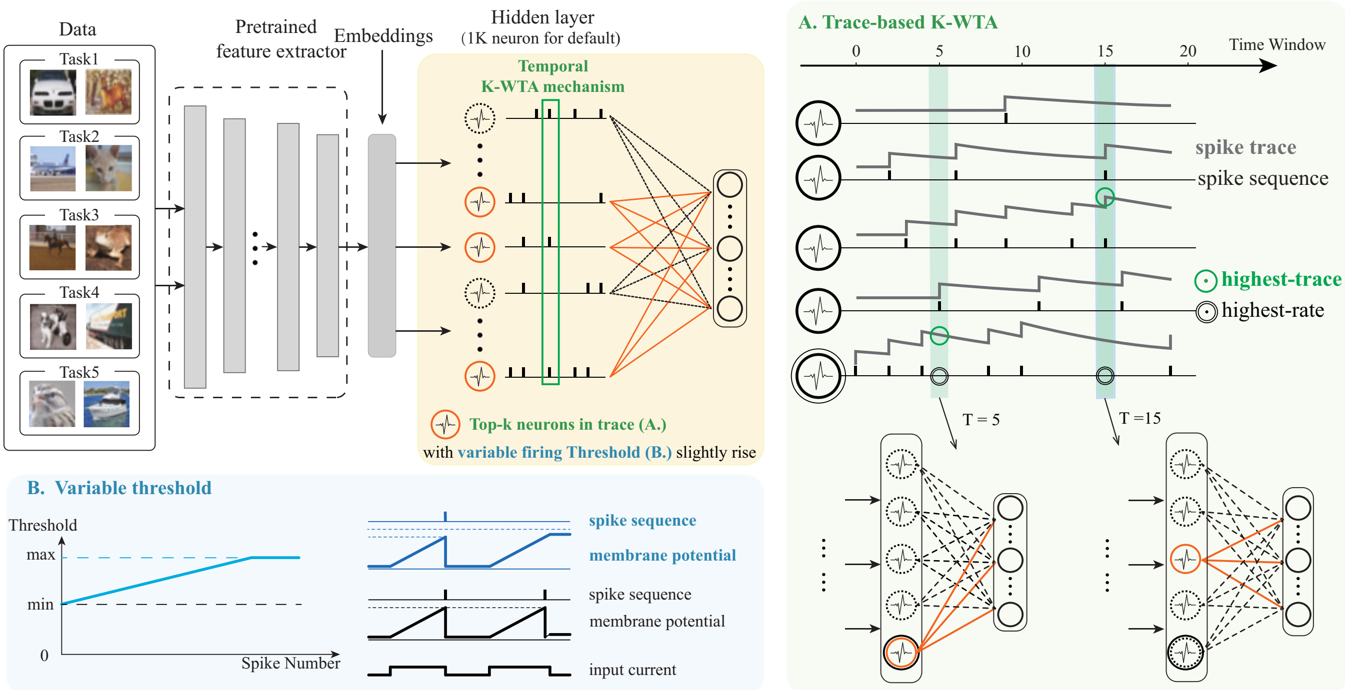

Figure 1: The architecture of the proposed SA-SNN model which incorporates trace-based K-WTA and variable threshold components. (A.) The process of obtaining the Top-K Mask in SNN via trace in a single time step and the rate over the whole time window. The dark blue curves and the black vertical line sequence below represent the traces and the spike sequence throughout the entire time window, respectively. (B.) The plotting of the relationship between the variable firing threshold of the neurons and their firing times.

all neurons and thus enhances the effect of gating at the population neuron level. The experiments are conducted in the class incremental (Class-IL) setting.

Our main contributions are summarized as follows:

- We propose the SNNs model with selective activation (SA-SNN) to reduce memory interference and mitigate catastrophic forgetting for continual learning scenarios without additional task information, inspired by the biological selective activation of context gating mechanism.
- The trace-based K-WTA and variable threshold components are developed with the aim of inducing sparsity in the selective activation of spiking neurons across both spatial and temporal dimensions, which in turn facilitates the activation of subpopulations of neurons that are specialized for different tasks.
- The Class-IL experiments are conducted on MNIST and CIFAR10 datasets. The proposed SA-SNN model achieves competitive performance similar to and even surpasses the other regularization-based methods deployed under ANNs.

Related Works

Continual learning aims to acquire knowledge in a sequential manner while ensuring that agents only have access to the data of the current task, without compromising their ability to recall previously learned tasks. The primary continual learning techniques can be broadly categorized into three distinct groups, namely regularization-based, rehearsal-based, and architecture-based techniques.

Regularization-based techniques preserver synaptic connections by adding a regularization term in loss function to consolidate the previously acquired knowledge, such as EWC(Kirkpatrick et al. 2017), MAS(Aljundi et al. 2018), SI(Zenke, Poole, and Ganguli 2017), which calculate the importance of each parameter through a certain rule, and generate a penalty term to limit the change of important parameters.

Rehearsal-based techniques aim to enhance knowledge retention by preserving several key samples (or intermediate representations) from each task and propagating them forward in the network mixed with current task's data. Most of these methods are oriented towards ANNs, such as (Van de Ven, Siegelmann, and Tolias 2020; Arani, Sarfraz, and Zonooz 2022; Rebuffi et al. 2017). As definition, such methods inevitably require additional storage space for extra information and extension of the model. The architecture-based methods (Kang et al. 2022; Yoon et al. 2017), which enhance the network's ability to perform different tasks by continuously adjusting its architecture, also encounter similar issues. As a result of the purposeful partitioning/expansion of the subnetworks, these methods often exhibit relatively stable performance between split tasks. However, in order to distinguish the usage range of different subnetworks, it tends to be necessary for these methods to have prior knowledge of task information.

As for other SNNs-based methods for continual learning, (Antonov, Sviatov, and Sukhov 2022) determines the importance of synaptic weights via stochastic Langevin dynamics with local STDP and achieved continual learning by unsupervised learning. (Skatchkovsky, Jang, and Simeone 2022) introduced an online rule base on the Bayesian SNN model. (Hammouamri, Masquelier, and Wilson 2022) achieves continual learning in SNNs by training an external network using evolutionary strategies to generate the firing threshold of classifer. (Tadros et al. 2022) implements the local plasticity to help the model to correct bias after the learning process of new tasks by using a conversion algorithm to switch between rate-coding and spike-coding. In addition, ANN-oriented methods, such as (Bricken et al. 2023) and (Shen, Dasgupta, and Navlakha 2021), have investigated the mechanisms similar to the selective activated Top-K function and provided efficient solutions for continual learning with ANNs model. These models are all based on rate-coding so it is not possible to explain the continual learning mechanism of neural networks from the level of temporal neural dynamics, nor to fully utilize the specific characteristics of biological neurons.

In brief, though the aforementioned techniques offer distinct performance benefits, they often necessitate intricate algorithms, additional storage usage, or task-specific knowledge, thereby deviating considerably from the innate sequential learning abilities of biological agents. Our approach attains the identical objective by only augmenting neural dynamics and neural-computational characteristics like context gating, which is more proximate to the possible biological learning process. By proposing meaningful neural computation features and temporal neural dynamics based on selective activation, we have developed a potential continual learning model of SNNs that is more akin to the learning processes of selective activation in biological systems.

Methods

The architecture of SNNs with selective activation is illustrated in Figure 1, which consists of a possible feature extractor for mining deep features in complex tasks and a multi-layer SNN to enable the network with the ability of continual learning. We apply the Temporary K-WTA mechanism in the dynamic of neurons in the hidden layer to reduce the mutual interference between different tasks. Meanwhile, the neurons with variable firing thresholds are employed to encourage silent neurons to participate in the learning process to some extent.

Task Setup

The task we perform in the is the continual learning task under the class-IL scenario. The following is the training pipeline for this type of task. Assume there are N learning phase (i.e. N sequential tasks in total). Task in each phase contains the data D_i with all training samples of c_i classes, where i devotes the index of the learning phase. During training stage, each learning phase (e.g. i-th phase, $i <= N$) can only access data related to the current task (i.e, all data belongs to class c_i in D_i) to train the model. When the training is done, the model will be evaluated on a test set with all classes $\sum_{j=1}^{i} c_j$ so far. The ultimate goal of the class-IL is to enable the model to distinguish all learned classes ever seen after training on all tasks in sequence. In the following, we introduce the components and trainning details of our solution.

Trace-Based K-WTA Mechanism

In neural circuit of animals, as mentioned in some researches (Lin et al. 2014) (Stevens 2015), there are inhibitory neurons which collect the excitation of some neurons and send feed-back inhibition and finally prevent most of the neurons from firing firing. This mechanism is often referred to as a winner-task-all(WTA) mechanism (Shen, Dasgupta, and Navlakha 2021), which is potentially helpful in maintaining network robustness (this mechanism is also refer to as K-WTA, where K devotes the number of winners). However, The homogeneous spike firing of neurons within a time window in SNNs makes it difficult to proceed the proper comparison among neurons directly.

It may not be rational enough if we directly integrate the spikes in the entire time window and then deploy the K-WTA mechanism. On the one hand, due to the discrete nature of the spikes, confusion is still easy to arise as there may be a lot of neurons that have the same number of spikes (especially with relatively small time steps). On the other hand, it seems to be unreasonable to use future information to generate gating signals for past events, which is somewhat lacking from the perspective of temporal rationality.

In the process of investigation, we noticed that Spike Timing-Dependent Plasticity (STDP) (Markram et al. 1997; Dan and Poo 2004), one kind of synaptic plasticity rule, updates the weight based on the spike time interval between the presynaptic and postsynaptic neurons. A internal variable called "trace" is introduced in (Morrison, Diesmann, and Gerstner 2008) for neurons to bridge the gap between time scale and action potential in the plasticity theory, which is updated with each spike and decays between spikes. Since this variable can give an online estimate of the mean firing rate in the spike train (Morrison, Diesmann, and Gerstner 2008), we can also use it as a indicator for temporal K-WTA. The "trace" is calculated as follows:

$$tr[t+1] = tr[t] - \frac{tr[t]}{\tau} + S[t+1], \tag{1}$$

where τ is the time constant, which decides the decay speed of the trace. $tr[t]$ is the trace of the neurons at time step t. $S[t+1]$ denotes the neurons' spike output at step t. This trace can be calculated at each time step and it is relatively easy to compare and obtain the Top-K value, so we apply this to deploy K-WTA computation step by step. As illustrated in Figure 1 (A), it is one example to show the TopK neuron activation difference between using trace or rate. The rate-based k-WTA chooses the 5th neuron when T=5 and 15 (high rate), while the trace-based K-WTA has better selective activation diversity with activating the 5th and 3rd neurons when T=5 and 15 (high trace). Another potential benefit

Methods	splitMNIST (h=1000)	splitCIFAR10 (h=1000)	splitCIFAR100 (h=2K, tasks=25)
None	19.96 (+/-0.01)	25.29 (+/-1.97)	4.28(+/-0.38)
Joint	97.72 (+/-0.04)	92.19 (+/-0.08)	65.28(+/-0.15)
EWC-SNN	19.89 (+/-0.00)	30.04 (+/-2.65)	3.93(+/-0.24)
MAS-SNN	19.91 (+/-0.01)	30.44 (+/-2.60)	3.09(+/-0.45)
SDMLP	46.67 (+/-3.92)	73.27 (+/-1.28)	-
SA-SNN(rate)	50.22 (+/-0.91)	76.88 (+/-2.12)	21.37(+/-0.77)
SA-SNN	**60.06** (+/- 2.16)	**77.73** (+/-1.95)	**22.86**(+/-0.64)
FlyModel	76.97 (+/-1.26)	70.09 (+/-0.51)	17.25 (+/-0.42)
SDMLP + EWC	79.61 (+/- 2.46)	78.64 (+/- 0.30)	21.31 (+/-0.72)
SA-SNN + EWC	**82.18** (+/- 1.14)	**80.39** (+/- 1.84)	**36.47** (+/- 2.13)

Table 1: The validation accuracy of the baseline and our methods in the experiment. All the methods listed here adopt the same classifier structure with one single hidden layer (h devote hidden size) and K is set to be 10. For split-CIFAR10, the model will have a pre-trained feature extractor. We highlight the final results of our methods. Note that SDMLP and Flymodel still use the calculation mode of ANN (and we skip the pre-training of the first layer of MLP in SDMLP), while EWC and MAS are applied to SNN networks.

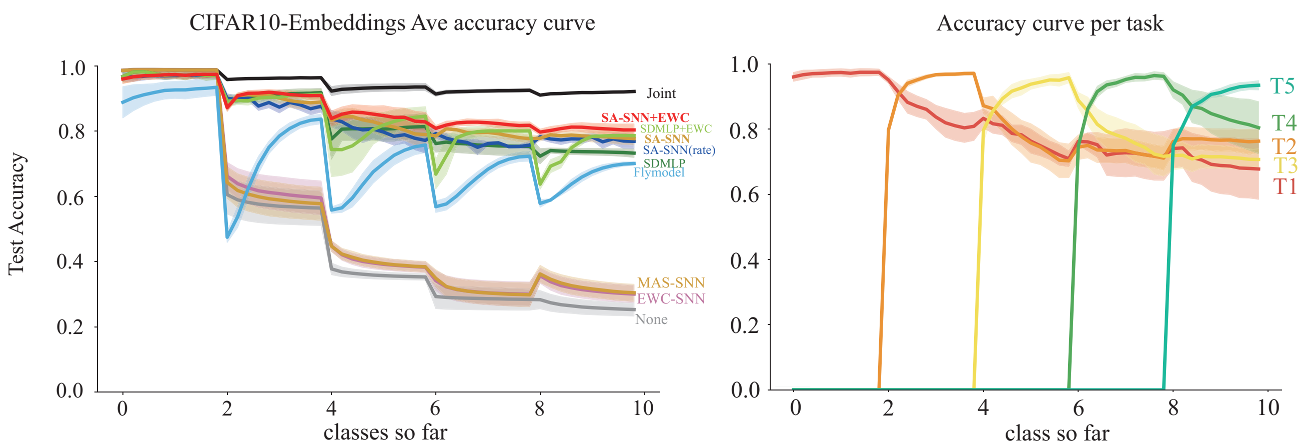

Figure 2: (Left)The performance comparison between our model and other baseline models on splitCIFAR10 dataset. The curve depicting the accuracy levels for the class ever learned over the learning process indicates that our model achieves competitive performance on splitCIFAR10. (Right) The average accuracy curves of our model across five tasks (T1 to T5).

of this trace-based K-WTA method is that it does not strictly adhere to the constraints of K from the perspective of the entire time window, which may improve the expression ability of the sub-networks under the Top-K function.

Variable Threshold

Top-K activation function usually leads to large groups of dead neurons (Ahmad and Scheinkman 2019; Fedus, Zoph, and Shazeer 2022). This is because the randomly initialized weight may allow a group of neurons easily activated and their synaptic weight continuously updated, leaving other neurons never activate and thus never receive feedback signals.

Inspired by the threshold variability mentioned in (Izhikevich 2004), we propose to use variable firing thresholds for the spiking neurons in the hidden layers. As shown in Figure 1 (B), we set the firing threshold to slowly increase as the activation times increase. It is worth noting that for convenience, the firing threshold won't decay in this paper, which enhances memory maintenance by the irreversible threshold changes. In that way, those neurons which were most frequently activated in the past have a decreasing probability of being reactivated, making those neurons with relatively low activation frequencies more likely to be activated and then gradually participate in the learning process of the network.

By adding the above two features to the basic neuron model, we can obtain the following neuronal dynamics:

$$
\begin{aligned}
H[t] &= f(V[t-1], X[t]), \\
S[t] &= \Theta(H[t] - V_{th}), \\
Mask[t] &= TopK(tr[t]), \\
S^*[t] &= S[t] \cdot Mask[t],
\end{aligned}
\tag{2}
$$

$$V[t] = H[t] - V_{th} \cdot S[t],$$

where $X[t]$ is the input to the neurons at the time step t. $S[t]$ represents the neurons' original spiking output. $V[t]$ and $H[t]$ are the membrane potential of the neurons before firing spikes (after charging) and after firing spikes, respectively. $\Theta(\cdot)$ is the spiking function. $TopK(\cdot)$ is the function used in trace-based K-WTA component to generate a mask that filters out the K largest trace which is mentioned in Equation 1. S^* denotes the actual activation output after the Top-K function. $f(\cdot)$ represents the state update of the neuron (such as integrate-and-fire (IF) neuron in Equation 3 or leaky integrate-and-fire (LIF) neuron model in Equation 4).

$$f(V[t-1], X[t]) = V[t-1] + X[t] \tag{3}$$

$$f(V[t-1], X[t]) = V[t-1] - \frac{(V[t-1] - V_{reset})}{\tau_t} + X[t] \tag{4}$$

we adopt the linear approach to implement the variable threshold as shown in Figure 1(B Left). Though it seems to be more biologically plausible for the firing threshold V_{th} to grow in a non-linear manner, we found the experimental results are similar between the non-linear and the linear approach. Hence, we choose the linear variable threshold for simplicity. The variable threshold is obtained by:

$$V_{th} = min(Th_{min} + \frac{C \cdot (Th_{max} - Th_{min})}{p}, Th_{max}) \tag{5}$$

where Th_{max} and Th_{min} are the upper and lower bounds of the firing threshold, C is the counter that record the accepted firing numbers of neurons and p is a hyper-parameter that controls the rate of threshold changing. This variable threshold mechanism can avoid certain neurons always to be selected for new coming classes but forget its response to the old classes. Through the irreversible variable threshold increasing, the activated neurons in the previous task would be more difficult to be activated again when training the following task.

Though the above two components can facilitate the formation of the sub-networks and guide neurons to participate in continual learning. Over-dense input may still lead to abnormal performance. That is because the greedy kernel of back-propagation as well as the limited precision and activation of SNN may lead to the excessively high firing rate of these neurons, which may harm the memory retention. So we adopt L^2 normalization method for the input layer and its correlated weight matrix to control the sparsity, inspired by (Bricken et al. 2023). Besides, we also applied some proven effective techniques in continual learning, such as Dale's rules (Dale 1935) and SGD optimizer to avoid the "stale momentum" problem (Zhu et al. 2023).

And In the training process of each learning stage, the training method used in SNN is consistent with the standard image classification algorithm of the SNN (Zhu et al. 2023; Xu et al. 2021; Shen et al. 2021). That is, we calculate the cross entropy between the outputs and the labels as loss, and use STBP (Wu et al. 2018; Xu et al. 2023; Guo et al. 2023) to train the learnable parameters in the network.

Results

To verify the effectiveness of our proposed framework in continual learning problems. The original CIFAR10 dataset is randomly divided into 5 tasks with 2 classes per task, referred to as "splitCIFAR10" datasets in the following. Each model is trained on these tasks sequentially as described in section 3.1. The pre-trained feature extractor from (Bricken et al. 2023) is applied to transform each sample in the CIFAR dataset into 256-dimensional latent embeddings. We also evaluate our model performance on the splitMNIST, splitN-MNIST (without the pre-trained feature extractor) and split-CIFAR100 (The experimental results on CIFAR100 dataset are illustrated in the supplementary materials). The experimental details other than splitCIFAR10 are recorded in the supplementary materials.

The models are trained 20000 batches on each sub-dataset (256 samples per batch, approximately 500 epochs for the whole CIFAR10 dataset) without any other pre-processing. We test the model performance on every class it has learned. The final accuracy is obtained by taking the average of three groups of experiments with multiple random seeds.

Basic Model Settings

All methods involved in the comparative studies are multi-layer neural networks with a single hidden layer. The hidden size is set to be 1000 and the hyper-parameter K is set to be 10. The time step is set to be 16 in splitCIFAR10.

The baseline models contain three different types: The first is the most basic situations, including the "without any measures" (None) and "training with all learned data" (Joint); The second is several regularization-based continual learning methods including EWC (Kirkpatrick et al. 2017) and MAS (Aljundi et al. 2018). Especially, the fixed version of EWC in (Bricken et al. 2023) is employed to improve the model performance. Since these methods are independent of the computational characteristics of ANNs, they can directly be transferred to SNNs. The third is some recently proposed biologically plausible methods based on ANNs, including SDMLP (Bricken et al. 2023) and Flymodel (Shen, Dasgupta, and Navlakha 2021), where SDMLP ignores its pre-training process of the first layer for convenience and fair, and Flymodel runs only one epoch for its particularity. In addition, the IF neuron model is employed in our model for its good compatibility with variable thresholds.

Performance Comparison

The comparative studies are conducted to evaluate the performance of the proposed model. As illustrated in Table 1 and Figure 2 (left), the proposed SA-SNN outperforms other methods on splitMNIST and splitCIFAR10 datasets, achieving an accuracy of 60.06 % and 77.73 %, respectively. Furthermore, it consistently maintains its accuracy advantage over other methods throughout the entirety of the learning process. Additionally, SA-SNN also exhibits superior performance compared to state-of-the-art ANNs models of SDMLP (73.27 %) and FlyModel (70.09 %) with similar principle for continual learning.

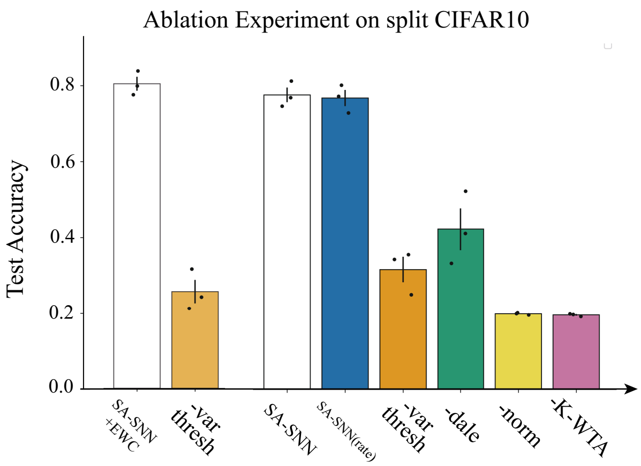

Figure 3: The validation accuracy of all ablated versions of SA-SNN. '-' means the removed corresponding component.

Besides, our model demonstrates a relatively high level of performance even by utilizing the rate mask directly (represented by SA-SNN(rate)), achieving an accuracy of 76.88% higher than the SDMLP algorithm on splitCIFAR10, despite both methods employing the K-WTA mechanism to mitigate the issue of forgetting. This observation suggests that the "integrate and fire" pattern of spiking neurons in the SNN may have advantages in the case of continual learning.

Moreover, we evaluate the accuracies of earliest classes after each classes training from T1 to T5 to directly compare the models' performance on CIFAR10 dataset. The SA-SNN achieves higher average accuracy with less variance than SDMLP and original SNN especially after last two tasks training. as shown in Figure 2 (Right), it is worth noting that despite the absence of any constraints on the direction of weight update or pre-training process in our model, the accuracy change curve of our model remains relatively balanced performance across various tasks. The balance among different tasks benefits from the effectiveness of selective activation of specific subpopulations of neurons and their associated connections for distinct tasks.

On this basis, since our method SA-SNN is not based on the inferred importance of weights to avoid forgetting, it's compatible with regularization-based methods such as EWC, MAS, SI(Zenke, Poole, and Ganguli 2017) that add penalty term in loss functions. As shown in Table 1, when combined with EWC, our model exhibits an enhancement of approximately 2.66 $\%$ and 22.12 $\%$ on splitCIFAR10 and splitMNIST dataset, compared with that of SDMLP plus EWC. This big gap between these two datasets appears because EWC applies regularization function to find a sub optimal local minima by minimizing the loss of new task around the local region of parameters space of old task, since the sequential classes in MNIST share more similar low-level features than CIFAR10, it is easier for SA-SNN with EWC to find a good joint probability distribution around the local region of parameters space of old task. Meanwhile, it is observed that the performance of EWC and MAS techniques on SNN in the conducted experiment is comparatively low, approximately 30 $\%$ on splitCIFAR10. And EWC methods exhibited a more noteworthy accuracy improvement by nearly 5 $\%$ when applied to ANNs on splitCIFAR10. This performance difference may be attributed to distinct neuron activation way between SNNs and ANNs. The above results are sufficient to demonstrate the effectiveness of our method.

Our SA-SNN model mainly introduce five hyper-parameter: p, τ, k, Th_{min} and Th_{max}. A small τ can lead to an overemphasis on temporal proximity by the Top-K function, thereby impeding the network's ability to segregate tasks into distinct subnetworks. Hence, τ is set to 10 and K is set to 10 through empirical observation.

Then, the influence of the other two hyper-parameters is illustrated in Figure 4 (a) and (b). Regarding the maximum threshold Th_{max}, its influence on the overall performance is relatively small when taking a value greater than 2.0. Since it is obvious that this parameter is correlated with p and the continuous change of threshold itself can lead to forgetting, we opt to establish a fixed value of 2.0 for it.

p controls the change speed of the firing threshold. When its value is too large, the continuously changing threshold will actually affect the ability to maintain knowledge. Meanwhile, it's not enough to fully mobilize the neurons when p is too small. Referring to the results in Figure 4 (b), we choose to set $p = 2000000$ under the experiment setting.

For hyperparameters in other methods for comparison, we use grid-search for parameter tuning and select relatively reasonable values for them.

Ablation Studies

In order to assess the efficacy of the incorporated components that facilitate the ability for continual learning, we conduct a series of ablation experiments and present the results in Figure 3. It is obvious that combining these components yield better performance than other versions without a certain component, indicating the necessity of these components. It is worth noting that apart from combining all components (78 $\%$), removing a single component alone can lead to a sharp decrease in the final results. That means it is synergistic effect between sparse input and isolated learning that results in the continual learning ability of SA-SNN.

As mentioned earlier, we can see that the Multi-step mask outperforms the rate-mask in terms of performance. This phenomenon is more notable in the splitMNIST experiment. It may be attributed to the fact that this method appropriately receives some ambiguous outputs, thereby indirectly expanding the network's ability to generalize tasks.

Regardless of the additional gains from the extended regularization-based method, the overall performance of the method we use under SNN is better than that of the similar method of SDMLP implemented under ANN, but where does the effort of this improvement come from? To explore the possible reasons, we refer to the definition of "task-selectivity" in (Flesch et al. 2023) and compare the relevant performance of the trained models.

We briefly define that when a neuron primarily responds to the input from only one category, it has selectivity for this

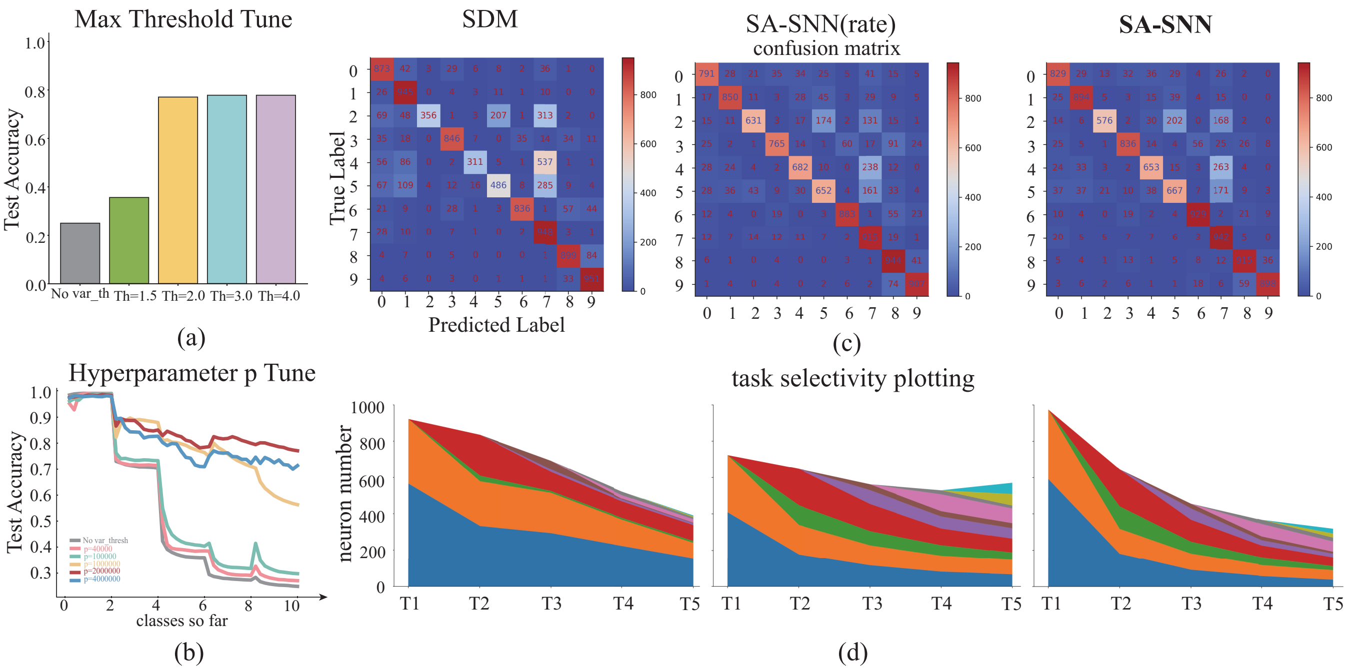

Figure 4: (a) Comparison of model performance with different hyper-parameter Th_{max}. (b) The accuracy curve of models with different hyper-parameter p. (c,d) The plotting of some statistical features of trained neural networks using SDMLP and our method in one experiment. (c) The confusion matrix of SDMLP, SA-SNN(rate), SA-SNN after training on splitCIFAR10. (d) The distribution of neurons with different task-selectivity during continual learning, starting from the end of Task1 learning.

category.To roughly classify neurons with different selectivity, we regress the hidden layer activity against 10 expected distinct selectivity (i.e. neurons only respond to a specific category). The 'activity' refers to the neurons' output (after $ReLu$), while in SNN it refers to the spikes count. And one of the results is shown in Figure 4 (c, d). In Figure 4 (d), we notice that the distribution of the neurons with different selectivity is relatively biased under the SDMLP method, while that under our method seems to be even during the learning process. This phenomenon is also directly reflected in the final confusion matrix of the model in Figure 4 (c): during the learning process of SDMLP models, the proportion of neurons with category 2, 4 and 5 selectivity is relatively small, so the ability to identify these categories is very easy to be disturbed in the subsequent learning process, which finally results in a relatively low accuracy in several categories. And this kind of contradiction is not so prevalent when using our method.

Besides, the number of neurons with specific selectivity is more when using rate-mask than using Multi-step mask. That may be attributed to the Multi-step mask's tolerance towards the neurons with similar functions, which encourages mixed selectivity. It still maintains a relatively uniform selectivity distribution even so.

Discussion

Compared to ANNs, SNNs share more biological features, and the spike firing mechanism within is thought to make the SNN networks more robust. In this paper, to explore the potential of intrinsic properties in SNNs, we introduce a set of neural features and combine the SNNs' computation mechanism with K-WTA mechanism in a biological plausibly way. The components we proposed in the paper ultimately form an SNN network that can automatically form relatively balanced subnetworks and achieve good performance in the class increment learning settings of continual learning with no need for task information and extra storage. The trace variable is innovatively introduced to express the degree of neuronal activities over a period of time, thus determine the superior K neurons when proceeding K-WTA process. It is different from membrane potential in SNNs, the membrane potential is only used to determine the spike firing in our neuron model.

Conclusions

Overall, inspired by the selective sparse activation principle of context gating in biological systems, we propose the SA-SNN with effective components of trace-base K-WTA, normalization and variable threshold, and reach a competitive performance. In addition to a new method for continual learning, we also investigate the potential advantages of the sparse selective activation mechanism in SNN during continual learning. The experimental results suggested the effectiveness of the proposed components of SNNs on continual learning tasks. What's more, our method also provides a possible way for augmenting continual learning capabilities in machine intelligence with limited computing resources through the integration of neuromorphic hardware.

Acknowledgements

This work was supported by National Key Research and Development Program of China (No. 2022YFB4500100), National Natural Science Foundation of China under Grant (No. 62306274, No.62236007, No. 62206037, No. 61925603), and Fellowship from the China Postdoctoral Science Foundation under Grant (No. 2023M733067 and No. 2023T160567).

References

Ahmad, S.; and Scheinkman, L. 2019. How can we be so dense? The benefits of using highly sparse representations. *arXiv preprint arXiv:1903.11257*.

Aljundi, R.; Babiloni, F.; Elhoseiny, M.; Rohrbach, M.; and Tuytelaars, T. 2018. Memory aware synapses: Learning what (not) to forget. In *Proceedings of the European conference on computer vision (ECCV)*, 139–154.

Antonov, D.; Sviatov, K.; and Sukhov, S. 2022. Continuous learning of spiking networks trained with local rules. *Neural Networks*, 155: 512–522.

Arani, E.; Sarfraz, F.; and Zonooz, B. 2022. Learning fast, learning slow: A general continual learning method based on complementary learning system. *arXiv preprint arXiv:2201.12604*.

Bricken, T.; Davies, X.; Singh, D.; Krotov, D.; and Kreiman, G. 2023. Sparse Distributed Memory is a Continual Learner. *arXiv preprint arXiv:2303.11934*.

Bu, T.; Ding, J.; Hao, Z.; and Yu, Z. 2023. Rate Gradient Approximation Attack Threats Deep Spiking Neural Networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7896–7906.

Dale, H. 1935. Pharmacology and nerve-endings.

Dan, Y.; and Poo, M.-m. 2004. Spike timing-dependent plasticity of neural circuits. *Neuron*, 44(1): 23–30.

Deng, L.; Wu, Y.; Hu, Y.; Liang, L.; Li, G.; Hu, X.; Ding, Y.; Li, P.; and Xie, Y. 2021. Comprehensive snn compression using admm optimization and activity regularization. *IEEE transactions on neural networks and learning systems*.

Fedus, W.; Zoph, B.; and Shazeer, N. 2022. Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity. *The Journal of Machine Learning Research*, 23(1): 5232–5270.

Flesch, T.; Nagy, D. G.; Saxe, A.; and Summerfield, C. 2023. Modelling continual learning in humans with Hebbian context gating and exponentially decaying task signals. *PLOS Computational Biology*, 19(1): e1010808.

French, R. M. 1999. Catastrophic forgetting in connectionist networks. *Trends in cognitive sciences*, 3(4): 128–135.

Guo, Y.; Peng, W.; Chen, Y.; Zhang, L.; Liu, X.; Huang, X.; and Ma, Z. 2023. Joint A-SNN: Joint training of artificial and spiking neural networks via self-Distillation and weight factorization. *Pattern Recognition*, 142: 109639.

Hammouamri, I.; Masquelier, T.; and Wilson, D. 2022. Mitigating Catastrophic Forgetting in Spiking Neural Networks through Threshold Modulation. *Transactions on Machine Learning Research*.

Imam, N.; and Cleland, T. A. 2020. Rapid online learning and robust recall in a neuromorphic olfactory circuit. *Nature Machine Intelligence*, 2(3): 181–191.

Izhikevich, E. M. 2004. Which model to use for cortical spiking neurons? *IEEE transactions on neural networks*, 15(5): 1063–1070.

Kandel, E. R.; and Hawkins, R. D. 1992. The biological basis of learning and individuality. *Scientific American*, 267(3): 78–87.

Kang, H.; Mina, R. J. L.; Madjid, S. R. H.; Yoon, J.; Hasegawa-Johnson, M.; Hwang, S. J.; and Yoo, C. D. 2022. Forget-free continual learning with winning subnetworks. In *International Conference on Machine Learning*, 10734–10750. PMLR.

Kay, L. M.; and Laurent, G. 1999. Odor-and context-dependent modulation of mitral cell activity in behaving rats. *Nature neuroscience*, 2(11): 1003–1009.

Kirkpatrick, J.; Pascanu, R.; Rabinowitz, N.; Veness, J.; Desjardins, G.; Rusu, A. A.; Milan, K.; Quan, J.; Ramalho, T.; Grabska-Barwinska, A.; et al. 2017. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*, 114(13): 3521–3526.

Kudithipudi, D.; Aguilar-Simon, M.; Babb, J.; Bazhenov, M.; Blackiston, D.; Bongard, J.; Brna, A. P.; Chakravarthi Raja, S.; Cheney, N.; Clune, J.; et al. 2022. Biological underpinnings for lifelong learning machines. *Nature Machine Intelligence*, 4(3): 196–210.

LeCun, Y.; Bengio, Y.; and Hinton, G. 2015. Deep learning. *nature*, 521(7553): 436–444.

Levinson, M.; Kolenda, J. P.; Alexandrou, G. J.; Escanilla, O.; Cleland, T. A.; Smith, D. M.; and Linster, C. 2020. Context-dependent odor learning requires the anterior olfactory nucleus. *Behavioral Neuroscience*, 134(4): 332.

Lin, A. C.; Bygrave, A. M.; De Calignon, A.; Lee, T.; and Miesenböck, G. 2014. Sparse, decorrelated odor coding in the mushroom body enhances learned odor discrimination. *Nature neuroscience*, 17(4): 559–568.

Maass, W. 1997. Networks of spiking neurons: the third generation of neural network models. *Neural networks*, 10(9): 1659–1671.

Markram, H.; Lübke, J.; Frotscher, M.; and Sakmann, B. 1997. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275(5297): 213–215.

McCloskey, M.; and Cohen, N. J. 1989. Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation*, volume 24, 109–165. Elsevier.

Morrison, A.; Diesmann, M.; and Gerstner, W. 2008. Phenomenological models of synaptic plasticity based on spike timing. *Biological cybernetics*, 98: 459–478.

Pei, J.; Deng, L.; Song, S.; Zhao, M.; Zhang, Y.; Wu, S.; Wang, G.; Zou, Z.; Wu, Z.; He, W.; et al. 2019. Towards artificial general intelligence with hybrid Tianjic chip architecture. *Nature*, 572(7767): 106–111.

Rebuffi, S.-A.; Kolesnikov, A.; Sperl, G.; and Lampert, C. H. 2017. icarl: Incremental classifier and representation learning. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 2001–2010.

Shen, J.; Xu, Q.; Liu, J. K.; Wang, Y.; Pan, G.; and Tang, H. 2023. ESL-SNNs: An Evolutionary Structure Learning Strategy for Spiking Neural Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(1): 86–93.

Shen, J.; Zhao, Y.; Liu, J. K.; and Wang, Y. 2021. Hybrid-SNN: Combining bio-machine strengths by boosting adaptive spiking neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 34(9): 5841 – 5855.

Shen, Y.; Dasgupta, S.; and Navlakha, S. 2021. Algorithmic insights on continual learning from fruit flies. *arXiv preprint arXiv:2107.07617*.

Skatchkovsky, N.; Jang, H.; and Simeone, O. 2022. Bayesian continual learning via spiking neural networks. *arXiv preprint arXiv:2208.13723*.

Stevens, C. F. 2015. What the fly's nose tells the fly's brain. *Proceedings of the National Academy of Sciences*, 112(30): 9460–9465.

Subbulakshmi Radhakrishnan, S.; Sebastian, A.; Oberoi, A.; Das, S.; and Das, S. 2021. A biomimetic neural encoder for spiking neural network. *Nature communications*, 12(1): 2143.

Tadros, T.; Krishnan, G. P.; Ramyaa, R.; and Bazhenov, M. 2022. Sleep-like unsupervised replay reduces catastrophic forgetting in artificial neural networks. *Nature Communications*, 13(1): 7742.

Van de Ven, G. M.; Siegelmann, H. T.; and Tolias, A. S. 2020. Brain-inspired replay for continual learning with artificial neural networks. *Nature communications*, 11(1): 4069.

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.

Xu, Q.; Li, Y.; Shen, J.; Liu, J. K.; Tang, H.; and Pan, G. 2023. Constructing deep spiking neural networks from artificial neural networks with knowledge distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7886–7895.

Xu, Q.; Shen, J.; Ran, X.; Tang, H.; Pan, G.; and Liu, J. K. 2021. Robust transcoding sensory information with neural spikes. *IEEE Transactions on Neural Networks and Learning Systems*, 33(5): 1935–1946.

Yin, B.; Corradi, F.; and Bohte, S. M. 2023. Accurate online training of dynamical spiking neural networks through forward propagation through time. *Nature Machine Intelligence*, 1–10.

Yoon, J.; Yang, E.; Lee, J.; and Hwang, S. J. 2017. Lifelong learning with dynamically expandable networks. *arXiv preprint arXiv:1708.01547*.

Zenke, F.; Poole, B.; and Ganguli, S. 2017. Continual learning through synaptic intelligence. In *International conference on machine learning*, 3987–3995. PMLR.

Zhu, Y.; Fang, W.; Xie, X.; Huang, T.; and Yu, Z. 2023. Exploring Loss Functions for Time-based Training Strategy in Spiking Neural Networks. In *Thirty-seventh Conference on Neural Information Processing Systems*.