SCIENCE CHINA
Information Sciences

• LETTER •

November 2026, Vol. 69, Iss. 11, 219103:1–219103:2
https://doi.org/10.1007/s11432-025-5127-7

Spatial-temporal knowledge fusion for continual learning in spiking neural networks

Jiashuo WANG[1,2], Qing'ao CHAI[1,2], Rui YAN[3] & Huajin TANG[1,2*]

[1]*College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China*
[2]*The State Key Lab of Brain-Machine Intelligence, Zhejiang University, Hangzhou 311113, China*
[3]*College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China*

Received 19 October 2025/Revised 6 May 2026/Accepted 18 August 2026/Published online 19 October 2026

Citation Wang J S, Chai Q A, Yan R, et al. Spatial-temporal knowledge fusion for continual learning in spiking neural networks. Sci China Inf Sci, 2026, 69(11): 219103, https://doi.org/10.1007/s11432-025-5127-7

Spiking neural networks (SNNs), characterized by their event-driven and energy-efficient nature, have emerged as a promising computational paradigm for edge computing, real-time processing, and neuromorphic hardware deployment. Different from traditional artificial neural networks (ANNs) that operate on continuous activations, SNNs transmit and process information through discrete spikes, enabling implementation on neuromorphic chips. The sparse activation, local learning rules, and spatiotemporal dynamics of SNNs make them structurally and functionally closer to biological neural systems, which has fueled interest in deploying SNNs for real-world applications. In practical scenarios, real-world data exhibits non-stationarity and sequential arrival, requiring models to continuously incorporate new knowledge. However, traditional deep learning methods typically suffer from catastrophic forgetting, where learning new data overwrites previously learned information. Continual learning methods aim to address this by learning new tasks while minimizing forgetting, with class incremental learning (CIL) being among the most challenging settings.

Recent studies have explored continual learning in SNNs from perspectives such as structural plasticity [1], sparse activation [2], and logit alignment [3]. Despite these advances, existing SNN-based CIL methods still lag behind ANN counterparts, particularly in longer-step scenarios. Moreover, a critical yet underexplored issue is task confusion, where new classes are highly similar to previously learned ones and the model struggles to distinguish between them, leading to systematic misclassification.

To address this challenge, we draw inspiration from complementary learning systems (CLS) theory, which posits that the brain balances rapid learning and long-term consolidation through hippocampal fast episodic encoding and cortical generalization across related episodes. Crucially, such generalized knowledge is then fed back to the hippocampus via the medial temporal lobe (MTL) [4], thereby achieving a balance between rapid response to novel information and long-term knowledge generalization.

Based on these insights, we focus on designing a continual learning method for SNNs that comprises a fast learner and a slow learner. The slow learner, analogous to the cortex, focuses on acquiring general representations through self-supervised learning, capturing the structural relationships within the data that are invariant across tasks. The fast learner, analogous to the hippocampus, rapidly adapts to new information through a dynamically expandable architecture. To emulate the cortical feedback mechanism, general representations from the slow learner are fused with features extracted by the fast learner across spatial and temporal scales. The fused representations are then fed back to the fast learner, providing contextual information that enables the model to discriminate between similar classes from different tasks, thereby alleviating task confusion. Furthermore, to mitigate the computational overhead caused by the continual addition of new modules in the expandable architecture, we draw inspiration from the brain's sparse coding mechanism and introduce a k-winners-take-all sparse activation module to improve computational efficiency. Experiments on CIFAR100 and ImageNet100 show that our method outperforms existing SNN-based CIL methods and achieves performance comparable to ANN counterparts with only ∼10% energy consumption based on theoretical estimates.

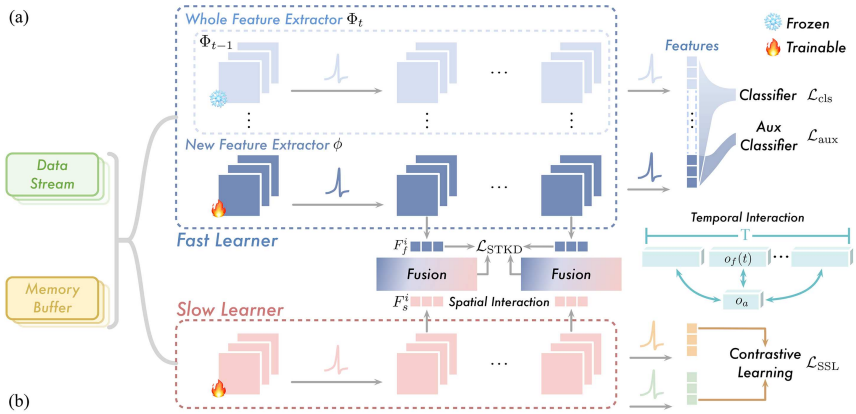
Methods. Our model comprises two SNN modules (Figure 1(a)). The slow learner acquires general representations through self-supervised contrastive learning with the InfoNCE loss. The fast learner adapts to new tasks through a dynamically expandable architecture following DER [5]. Upon task t, the previously trained feature extractor Φ_{t-1} is frozen and a new module ϕ is appended to form $\Phi_t(\boldsymbol{x}) = [\Phi_{t-1}(\boldsymbol{x}), \phi(\boldsymbol{x})]$. A classification loss $\mathcal{L}_{\text{cls}}$ and an auxiliary loss $\mathcal{L}_{\text{aux}}$ jointly enhance the new module's recognition of new classes. In addition, a memory buffer stores representative samples from previous tasks for rehearsal.

To emulate the cortical feedback mechanism in complementary learning systems, general representations learned by the slow learner are fused with features extracted by the fast learner across both spatial and temporal dimensions, and then fed back to the fast learner. Along the spatial dimension, in complex scenarios, features from different classes may exhibit high similarity at certain spatial scales, leading to inter-class interference. To address this, features from the i-th block of the fast learner $\boldsymbol{F}_f^i$ and slow learner $\boldsymbol{F}_s^i$ are integrated through a cross attention module, where the query is derived from $\boldsymbol{F}_f^i$ and the key and value are computed from the concatenation of $\boldsymbol{F}_f^i$ and $\boldsymbol{F}_s^i$. By integrating multi-scale

* Corresponding author (email: htang@zju.edu.cn)

© Science China Press 2026

(a)



(b)

Model	Methods	CIFAR100-B0-5steps		CIFAR100-B0-10steps		CIFAR100-B0-20steps	
		Avg(%)	Last(%)	Avg(%)	Last(%)	Avg(%)	Last(%)
ANN	DER	76.80	68.30	75.36	65.22	74.09	62.48
	TCIL	77.70	69.60	77.30	66.40	75.10	63.50
	DNE-1head	-	-	74.03	69.10	73.27	67.61
SNN	ALADE-SNN	**78.67**	69.67	75.42	63.13	72.73	59.43
	Ours	77.02	**70.69**	**76.85**	**68.34**	**75.54**	**66.20**

Figure 1 (Color online) (a) Overview of our proposed method; (b) comparison on CIFAR100-B0 benchmarks (best SNN results in bold).

features from different stages of the feature extractor, this spatial fusion mechanism exploits both low-level spatial details and high-level semantic abstractions, enhancing the model's sensitivity to local variations and alleviating the interference caused by inter-class feature similarity. The logits obtained from the fused features then serve as the teacher to guide the training of the fast learner through KL-divergence-based knowledge distillation.

Along the temporal dimension, considering that the output of SNNs is dynamic across timesteps, existing distillation methods that focus on spatial feature alignment are insufficient. We therefore extend the distillation over the temporal dimension by aligning the fused logits $\boldsymbol{o}_a^i$ of the i-th spatial block with the fast learner's logits $\boldsymbol{o}_f^i(t)$ at each timestep t, yielding the spatial-temporal knowledge distillation loss $\mathcal{L}_{\text{STKD}} = \frac{1}{T}\sum_i \sum_t \mathcal{L}_{\text{KD}}(\boldsymbol{o}_a^i, \boldsymbol{o}_f^i(t))$. This temporal alignment leverages the temporal attributes of SNNs and facilitates consistent model performance across timesteps. The fused representations are fed back to the fast learner, providing complementary contextual information that enables the model to better discriminate between similar classes from different tasks, thereby alleviating task confusion.

Although the dynamic structural expansion helps reduce task interference, it introduces more parameters and, an increasing computational burden as modules are added. To mitigate this overhead, we draw inspiration from the brain's sparse coding mechanism and introduce a k-winners-take-all mechanism. Before the input current is passed to each spiking neuron layer, a binary mask retains only the top-K neurons with the highest membrane currents while suppressing the rest. This sparse activation strategy reduces overall synaptic operations while maintaining sufficient representational capacity, enabling more energy-efficient continual learning with only a moderate accuracy trade-off. For a more detailed description of our method, please refer to Appendix A.

Experiments. Figure 1(b) presents results on the CIFAR100-B0 benchmark. In most cases, our method significantly outperforms competing approaches and achieves performance comparable to ANN-based methods. In the 10-step setting, our method achieves 1.43% higher average accuracy and 5.21% higher last step accuracy over the state-of-the-art spiking CIL method ALADE-SNN [3]. Under the more challenging 20-step setting, the improvements further increase to 2.81% and 6.77%, respectively. As the number of tasks grows, catastrophic forgetting becomes more pronounced, leading to substantial performance drops in most existing methods, whereas our approach maintains strong performance even in longer-step scenarios, suggesting more effective accumulation of general knowledge. Detailed results are provided in Appendix B.

Conclusion. Motivated by the CLS theory, we develop a brain-inspired continual learning method for SNNs to address catastrophic forgetting. By fusing generalized and task-specific representations across spatial and temporal scales, our method alleviates task confusion. Experiments on CIFAR100 and ImageNet100 demonstrate superior performance over existing SNN-based baselines, especially in longer-step scenarios, while achieving accuracy comparable to ANN counterparts with only ∼10% of their energy consumption based on theoretical estimates (Appendix B.5). These results demonstrate that our framework effectively narrows the performance gap between SNN-based and ANN-based CIL methods while preserving the energy efficiency advantage of SNNs, contributing to advances in class incremental learning for SNNs.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 32441113, 62236007, 62276235).

Supporting information Appendixes A–C. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Han B, Zhao F, Zeng Y, et al. Enhancing efficient continual learning with dynamic structure development of spiking neural networks. In: Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, 2023. 2993–3001
2 Shen J, Ni W, Xu Q, et al. Efficient spiking neural networks with sparse selective activation for continual learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2024. 611–619
3 Ni W, Shen J, Xu Q, et al. Alade-snn: adaptive logit alignment in dynamically expandable spiking neural networks for class incremental learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2025. 19712–19720
4 Shi Q, Liu F, Li H, et al. Hybrid neural networks for continual learning inspired by corticohippocampal circuits. Nat Commun, 2025, 16: 1272
5 Yan S, Xie J, He X. Der: dynamically expandable representation for class incremental learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021. 3014–3023