

Saliency-Guided Replay for Continual Learning in Deep Spiking Neural Networks

Francesco Panetta¹, Federico Di Valerio¹ (✉), and Michela Proietti²

¹ Sapienza University of Rome, Italy panetta.1950468@studenti.uniroma1.it,
divalerio@diag.uniroma1.it

² Goethe University Frankfurt, Germany m.proietti@em.uni-frankfurt.de

Abstract. Spiking Neural Networks (SNNs) are promising for energy-efficient neuromorphic computing, but they still suffer from catastrophic forgetting when learning tasks sequentially. This work investigates the integration of Experience Packing and Replay (EPR) as a memory-efficient continual learning strategy for deep SNNs under low-latency constraints. The proposed framework combines Batch Normalization Through Time-stabilized VGG9 SNNs with Spike Activation Maps (SAM), which provide forward-only explanations based on spike activity. These explanations are temporally aggregated into spatial attention maps and used to select the most informative image patches for replay, instead of storing full images. Experiments on Split CIFAR10 and Split CIFAR100 show that SAM-guided EPR improves the accuracy-memory trade-off under tight memory budgets, and remains competitive while requiring substantially less storage. The code is available at this link.

Keywords: Spiking Neural Networks · Continual Learning · Explainable AI · Experience Replay.

1 Introduction

Artificial intelligence systems deployed in real-world environments must adapt to evolving data while meeting strict computational and energy constraints. Yet standard neural networks still suffer from catastrophic forgetting, where learning new tasks disrupts previously acquired representations [13]. Continual learning (CL) addresses this by preserving prior knowledge while integrating new information over time. Among CL strategies, replay-based methods are particularly effective, storing representative samples from past tasks and mixing them with current data during training [3,21]. However, their performance depends on memory capacity: larger buffers improve retention but are impractical for resource-constrained systems. To reduce this cost, Experience Packing and Replay (EPR) stores saliency-guided image patches rather than full samples, substantially reducing memory requirements while maintaining competitive CL performance [23].

Spiking Neural Networks (SNNs) have emerged as a promising paradigm for energy-efficient machine intelligence, using sparse spike events rather than

continuous-valued activations [12]. Surrogate-gradient learning has enabled deep SNNs to achieve competitive performance on vision benchmarks [16,28], while Batch Normalization Through Time (BNTT) improves training stability and reduces inference latency [8]. However, continual learning in SNNs remains underexplored [14]. Replay-based approaches have only recently been adapted to SNNs, first through standard experience replay to mitigate catastrophic forgetting [18], and more recently through compact latent replay to reduce storage and computation costs in neuromorphic systems [4]. To the best of our knowledge, no prior work has used explainable artificial intelligence (XAI) to guide SNN learning and mitigate catastrophic forgetting, despite its effectiveness in artificial neural networks (ANNs) [19].

In this work, we investigate whether temporally grounded saliency signals in SNNs can guide memory-efficient replay. Specifically, we adapt EPR [23] to deep SNNs by replacing conventional ANN saliency mechanisms with Spike Activation Maps (SAM) [9], which produce forward-only explanations from inter-spike interval dynamics. Using BNTT-trained VGG9 SNNs [8] operating at low latency (25 timesteps) on Split CIFAR10 and Split CIFAR100 CL benchmarks, we show that saliency-guided replay can be effectively transferred to the spiking domain, improving performance under equal memory budgets and reducing forgetting in low-memory regimes.

2 Related Work

SNNs provide a sparse, event-driven alternative to conventional ANNs, making them particularly attractive for neuromorphic and energy-efficient learning systems. Modern SNNs can be trained either through ANN-to-SNN conversion or directly with surrogate-gradient methods, enabling competitive performance across a range of vision and learning tasks [6,16,25,28]. Continual learning in SNNs has recently attracted increasing attention, motivated by the need for adaptive neuromorphic systems that can learn over time under memory and energy constraints [15]. Existing approaches mostly focus on regularization-based mechanisms to mitigate catastrophic forgetting, including adaptive synaptic plasticity, Bayesian learning, and sparse selective activation [17,26,27]. In contrast, replay-based continual learning for SNNs remains under-explored [18]. Recent work on compressed latent replay for SNNs further confirms the relevance of rehearsal-based strategies in spiking models, while highlighting the need for memory-efficient replay mechanisms tailored to temporal spiking representations [4]. In parallel, recent continual learning methods for conventional ANNs have begun to exploit XAI to identify and preserve task-relevant information from previous experiences [1,5,7,22,23]. Among these, EPR [23] is particularly relevant to our work: it uses saliency maps to select discriminative image patches and stores only those patches in the replay buffer. This reduces memory requirements while maintaining competitive continual learning performance. However, most saliency-guided replay methods rely on explanation mechanisms originally designed for conventional ANNs. Applying these mechanisms directly to SNNs

overlooks the distinctive properties of spiking computation, including sparse event-driven activity and temporal dynamics. This motivates spike-compatible explanation strategies that can exploit temporal spiking activity [2,9]. Our work addresses this gap by adapting the principle of saliency-guided replay to SNNs. Instead of using Grad-CAM-based image patches as in EPR, we use SAM-based spatio-temporal saliency [9] to select compact replay information. In this way, explanation becomes a functional component of continual learning in SNNs, rather than only a diagnostic tool.

3 Methods

This section introduces a continual learning framework for SNNs that combines Spike Activation Maps (SAM) [9] with a replay strategy inspired by Experience Packing and Replay (EPR) [23]. The key idea is to use temporally grounded saliency to store only the most informative regions of past samples, enabling memory-efficient replay in neuromorphic continual learning.

3.1 Spike Activation Maps

Standard gradient-based explanation methods, such as Grad-CAM [24], are difficult to apply to SNNs because of their discrete and non-differentiable spike dynamics. We therefore adopt SAM, a forward-only explanation mechanism that extracts saliency directly from spike-train activity.

Given a spike tensor $S \in \{0, 1\}^{T \times C \times H \times W}$, where T is the number of timesteps, C the number of channels, and $H \times W$ the spatial resolution, SAM is computed using a causal temporal kernel:

$$K(\tau) = \begin{cases} \exp(-\gamma\tau), & \tau > 0, \\ 0, & \tau = 0, \end{cases} \quad (1)$$

where γ controls the decay rate. This kernel assigns larger weights to recent spikes while ensuring that only past activity contributes to the current timestep. Next, a neuronal contribution score is computed:

$$\text{NCS}_{c,h,w}(t) = \sum_{\tau=1}^{t-1} K(\tau) S_{c,h,w}(t - \tau). \quad (2)$$

Finally, the SAM heatmap at timestep t is obtained by combining the accumulated spike history with the current spike activity:

$$H_t(h, w) = \sum_{c=1}^C \text{NCS}_{c,h,w}(t) S_{c,h,w}(t). \quad (3)$$

The resulting tensor $H \in \mathbb{R}^{T \times H \times W}$ captures the spatio-temporal importance of neural activations during the forward pass. For replay selection, we reduce the

temporal dimension either by averaging all heatmaps, $H = \frac{1}{T} \sum_{t=1}^T H_t$, or by using only the final heatmap, $H = H_T$. Finally, SAMs can be extracted from different network layers. Based on findings on deep convolutional networks [29], we use the deepest convolutional layer, which provides lower spatial resolution but richer semantic representations and more discriminative saliency maps.

3.2 Experience Packing and Replay

We propose an SNN-adapted EPR strategy that uses SAMs to identify and store the most salient image patch instead of the full input. Each buffer exemplar contains the extracted patch, its original coordinates $(top, left)$, the class label, the aggregated SAM heatmap used for selection and optionally a SAM history for offline explanation-drift analysis. For each input, we compute an aggregated SAM heatmap $H \in \mathbb{R}^{H_{sam} \times W_{sam}}$ at the target layer and bilinearly upsample it to the input resolution, yielding $H_{full} \in \mathbb{R}^{32 \times 32}$ for CIFAR images. We then scan the image with a sliding window of size $P \times P$, with $P = 24$ by default, and compute the saliency score of each candidate patch:

$$s_{(top, left)} = \sum_{h=top}^{top+P} \sum_{w=left}^{left+P} H_{full}(h, w). \quad (4)$$

The most informative patch is selected by choosing the candidate location with the maximum saliency score:

$$(top^*, left^*) = \arg \max_{(top, left)} s_{(top, left)} \quad (5)$$

The selected patch is then extracted from the original image as:

$$patch = image[:, top^* : top^* + P, left^* : left^* + P]. \quad (6)$$

During replay, the patch is placed back into a zero-valued canvas of the original input size at its stored coordinates $(top^*, left^*)$. After each task, the memory buffer is updated using a temporary ring buffer M_t of recent samples from the current task. For Task 1, the model computes the SAM heatmap, extracts the most salient patch, reconstructs its padded version, and stores the exemplar only if the prediction is correct or if the ground-truth class appears among the top-3 predictions. For later tasks, the same procedure is applied to new samples, while stored exemplars may optionally be reprocessed with the updated model to track explanation drift. We refer to this EPR-inspired variant [23] as EPR-SP. As an ablation, we also consider BASE-EPR, where the samples in M_t are drawn uniformly at random from the current task training set.

At training stage i , the model is optimized using current-task mini-batches $B_i \subset D_i$ and replay minibatches B_{mem} sampled from the previously stored patches. Since replay samples are reconstructed as padded inputs $\tilde{x}$, the same supervised objective can be applied to both current and replay data:

$$L_{total} = L_{CE}^{(i)} + \lambda_{replay} L_{replay}, \quad (7)$$

where

$$L_{CE}^{(i)} = \frac{1}{|B_i|} \sum_{(x,y) \in B_i} \ell_{CE}(f_\theta(x), y), \quad L_{replay} = \frac{1}{|B_{mem}|} \sum_{(\tilde{x}, y) \in B_{mem}} \ell_{CE}(f_\theta(\tilde{x}), y). \quad (8)$$

Here λ_{replay} controls the stability-plasticity trade-off between retaining knowledge from previous tasks and adapting to the current one. This formulation avoids an explicit explanation-matching loss and instead uses explanations only to guide exemplar selection, making the replay mechanism compatible with the class-agnostic nature of SAM.

3.3 Evaluation Metrics

Let $a_{i,j}$ denote the accuracy on the test set of task j after the network has completed training on task i , and let N be the total number of tasks. We evaluate the models' final performance using the average accuracy and forgetting:

$$\text{AvgAcc} = \frac{1}{N} \sum_{j=1}^N a_{N,j}, \quad \text{AvgForg} = \frac{1}{N-1} \sum_{j=1}^{N-1} (a_{j,j} - a_{N,j}) \quad (9)$$

4 Experimental Setup

We perform experiments on Split-CIFAR10 (5 tasks) and Split-CIFAR100 (both the 10-task and the 20-task versions) [10]. For each task, 20% of the task training data is held out as a validation set used for early stopping. The used architecture is taken from [8] and consists of a VGG-9-inspired SNN. Convolutional layers are made of leaky integrate-and-fire neurons [11] with a fixed threshold of 1.0 and soft-reset dynamics, and are paired with per-timestep batch normalization [8]. Input images are encoded as Poisson spike trains over $T=25$ discrete timesteps. A two-layer classification head outputs task-specific logits via independent task-specific linear layers, each mapping the shared feature representation to the classes of its corresponding task. We consider both task- (TIL) and class-incremental (CIL) learning scenarios, where task identities are available during both training and testing and just during training, respectively. In the CIL scenario, all task heads are concatenated at inference time and the class with the highest logit across all seen tasks is selected. Since the heads from previous tasks are frozen while training on the current task, we add a separate replay phase after each training epoch, in which the heads are unfrozen and samples from previous tasks are replayed. We compare EPR variants, BASE-EPR and EPR-SP, with standard experience replay (identified as BASELINE), which stores full 32×32 images sampled randomly from the current task's training set.

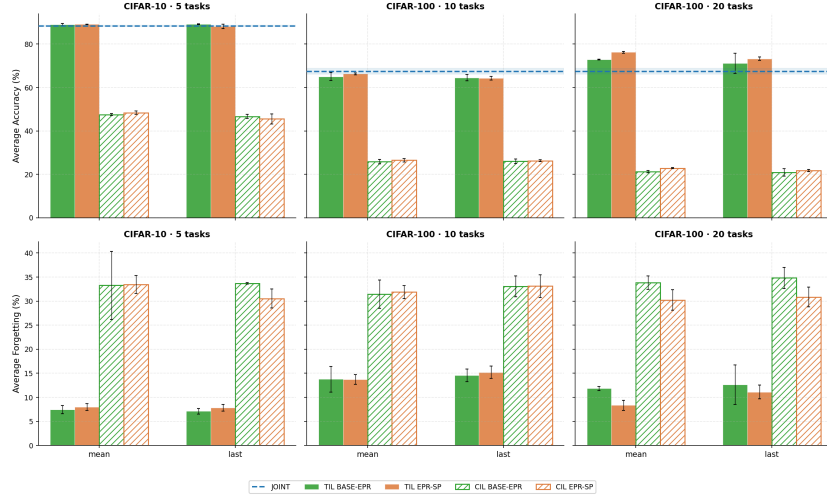


Fig. 1: Comparison of BASE-EPR and EPR-SP using mean and last temporal aggregation across benchmarks. Columns denote benchmark settings, with *AvgAcc* (top) and *AvgForg* (bottom). Solid and hatched bars indicate TIL and CIL, respectively; error bars show standard deviation over 3 seeds.

4.1 Hyperparameter Search

All models are trained using mini-batch SGD [20] with momentum 0.9, batch size 128, weight decay 5×10^{-4} , gradient clipping $\ell_2 \leq 1.0$, and initial learning rate $\eta \in \{0.1, 0.3\}$. These parameters were selected by performing a grid search on the BASELINE, varying membrane leak $\lambda_{\text{mem}} \in \{0.95, 0.99\}$, learning rate $\eta \in \{0.1, 0.3\}$, and dropout $\in \{\text{on}, \text{off}\}$ at a fixed memory budget ($M_t = 100$ samples for CIFAR10, $M_t = 30$ samples for CIFAR100). Subsequently, we run an ablation to choose sample selection strategy (BASE-EPR vs EPR-SP) and SAM aggregation method (mean vs last, as described in Section 3.1). Figure 1 shows that across all three benchmarks EPR-SP consistently outperforms or ties its BASE-EPR counterpart, confirming that model-validated selection improves the discriminative quality of the replay buffer. Regarding temporal aggregation, *mean* yields consistently higher or comparable accuracy across all benchmarks. Finally, replay-weight ablations ($\lambda_r = 0.3, 0.5, 0.7, 1.0$) showed that $\lambda_r = 1.0$ performs best on CIFAR10, whereas $\lambda_r = 0.3$ provides a better stability-plasticity trade-off on CIFAR100. These values are used in subsequent experiments. Therefore, in all subsequent experiments, we use EPR-SP with mean aggregation as the default configuration, with $\lambda_r = 1.0$ for CIFAR10 (in all scenarios) and CIFAR100 with tight memory budgets ($M_t = 30$), and $\lambda_r = 0.3$ for CIFAR100 with larger memory budgets ($M_t = 53$). The learning rate follows a cosine annealing schedule within each task, decaying from η to a minimum of 10^{-4} . At each task boundary, the schedule and the learning rate are reset to the initial value η . Early stopping monitors validation accuracy on the current task with

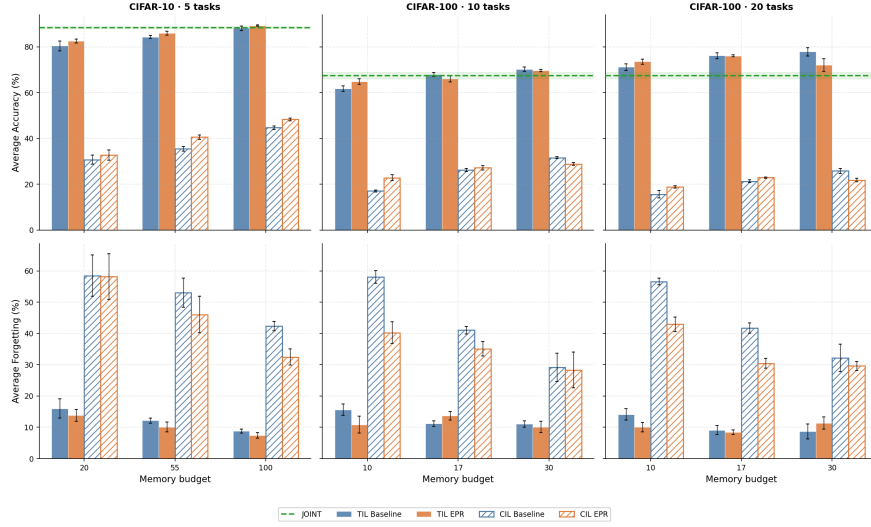


Fig. 2: TIL and CIL accuracy versus memory budget for BASELINE and EPR-SP across all benchmarks. At equal budgets, EPR-SP stores more patches than BASELINE stores full samples, yielding higher accuracy at tight and moderate budgets, with the advantage diminishing as memory increases.

patience 10 for CIFAR10 over 50 epochs and 20 for CIFAR100 over 100 epochs. Finally, concerning heatmap computation, we compute SAMs from layer 7 with temporal decay $\gamma=0.7$, which experimentally gave the best results.

5 Results

5.1 Comparison Under Equal Memory Budgets

A 24×24 patch requires $3 \times 24 \times 24 = 1,728$ bytes versus $3 \times 32 \times 32 = 3,072$ for a full image, yielding a storage cost factor of ≈ 0.56 ($\approx 44\%$ reduction). To enable fair comparisons between the BASELINE and EPR-SP we define:

$$\text{Memory budget } (M_B) = M_t \times \text{cost}, \quad \text{cost} = \begin{cases} 1.00 & \text{BASELINE} \\ 0.56 & \text{BASE-EPR, EPR-SP} \end{cases} \quad (10)$$

For instance, EPR-SP at $M_t = 177$ ($M_B=99$) is directly comparable to the BASELINE at $M_t = 100$ ($M_B=100$). Figure 2 shows the results under comparable memory budgets. EPR-SP outperforms the BASELINE at the tightest memory budget across all datasets. On CIFAR100/10 and CIFAR100/20 (10 and 20 tasks respectively) at $M_B \approx 10$, the *AvgAcc* increases by 3.1 and 2.4 points in TIL and 5.8 and 3.2 points in CIL, with CIL *AvgForg* reduced by 17.8 and 13.7 points. The gap narrows as the budget grows: at $M_B \approx 30$ on CIFAR100, the

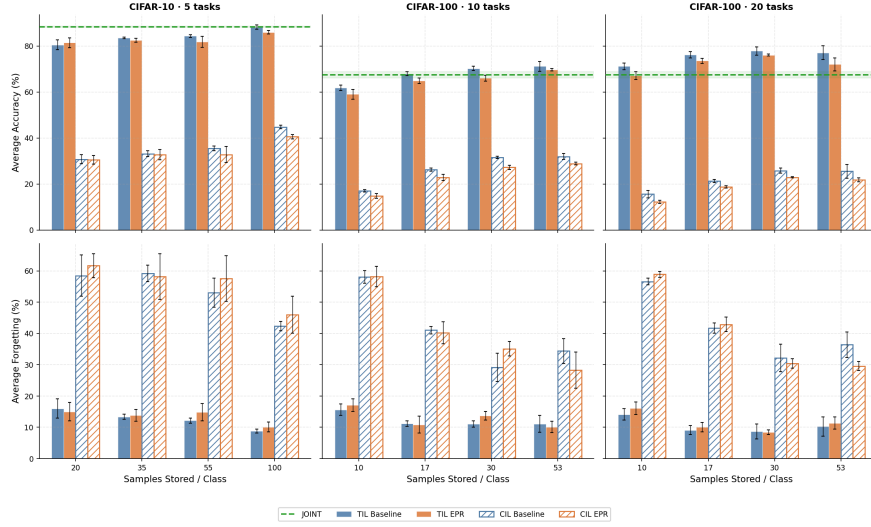


Fig 3: TIL and CIL accuracy versus exemplars per class K for BASELINE and EPR-SP on Split-CIFAR10 (5 tasks), CIFAR100/10 (10 tasks), and CIFAR100/20 (20 tasks). At the same K , EPR-SP uses 44% less memory than BASELINE. The gap narrows as K decreases, with EPR-SP outperforming BASELINE at the tightest CIFAR10 budget.

BASELINE recovers the lead on TIL accuracy while EPR-SP retains a forgetting advantage, suggesting the occurrence of a buffer-saturation effect in which additional low-resolution samples yield diminishing returns compared to full-resolution ones. These results confirm that in the presence of stringent memory constraints, as in the case of neuromorphic hardware, SAM-guided patch replay can obtain better results than standard replay.

5.2 Comparison Under Equal Sample Budgets

Figure 3 further highlights the practical benefit of patch-based replay when the memory budget is extremely limited. When performance is plotted against the number of stored samples per class, the BASELINE generally achieves higher accuracy than EPR-SP at the same nominal number of samples, especially as the buffer size increases. However, an EPR-SP sample requires only about 56% of the memory of a full image, therefore the resulting performance gap is relatively small compared to the memory savings. Overall, EPR-SP preserves competitive TIL and CIL accuracy while reducing the storage cost by 44%. This suggests that, in highly memory-constrained settings, EPR-SP offers a favorable trade-off: it sacrifices only a limited amount of accuracy while enabling a much more compact replay buffer. This behavior is particularly relevant for neuromorphic and embedded scenarios, where memory is often a stronger constraint than raw computational capacity.

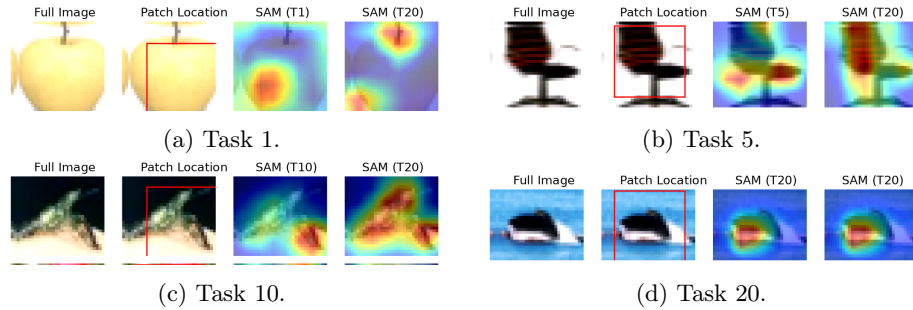


Fig. 4: Evolution of spatial attribution maps throughout continual learning. For Tasks 1, 5, 10, and 20, we show the input image, the selected patch location, the SAM generated immediately after learning the task from which the patch is selected, and the SAM obtained after training on all 20 tasks.

5.3 Explanation Drift Analysis

Figure 4 compares SAMs generated immediately after learning a task with those obtained after completion of the task sequence. Across all tasks, the salient regions remain largely consistent, indicating preservation of task-relevant visual representations. Interestingly, for earlier tasks, the final SAMs are often more concentrated on the target object and less affected by background regions. This suggests that continual learning not only preserves relevant evidence but can also refine shared feature representations as the network is exposed to additional visual concepts.

6 Conclusions

We presented SNN-adapted EPR, a SAM-guided replay strategy for CL in deep SNNs. Experiments on Split-CIFAR10 and Split-CIFAR100 showed favorable accuracy-memory trade-offs compared to standard replay, particularly under tight memory budgets. Qualitative analysis further indicated that task-relevant representations are preserved and may become more localized over time. Limitations include dependence on SAM quality, the use of a fixed patch size, and evaluation on small-scale image classification benchmarks. Future work could explore pre-trained backbones to improve early explanation quality, enabling better patch selection and memory initialization. Another direction is to make SAM more class-sensitive through output-conditioned scores or temporal attention, enabling the introduction of regularization terms to explicitly prevent explanation drift.

Acknowledgments. This work has been carried out while Michela Proietti and Federico Di Valerio were enrolled in the Italian National Doctorate on Artificial Intelligence and the PhD in Engineering in Computer Science, respectively, run by Sapienza University of Rome.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bai, G., Ling, C., Gao, Y., Zhao, L.: Saliency-augmented memory completion for continual learning (2023), <https://arxiv.org/abs/2212.13242>
2. Bitar, A., Rosales, R., Paulitsch, M.: Gradient-based feature-attribution explainability methods for spiking neural networks. *Frontiers in Neuroscience* **17** (2023). <https://doi.org/10.3389/fnins.2023.1153999>
3. Buzzega, P., Boschini, M., Porrello, A., Abati, D., Calderara, S.: Dark experience for general continual learning: a strong, simple baseline. *Advances in neural information processing systems* **33**, 15920–15930 (2020)
4. Dequino, A., Carpegna, A., Nadalini, D., Savino, A., Benini, L., Di Carlo, S., et al.: Compressed latent replays for lightweight continual learning on spiking neural networks. In: 2024 IEEE Computer Society Annual Symposium on VLSI (ISVLSI). pp. 240–245 (2024). <https://doi.org/10.1109/ISVLSI61997.2024.00052>
5. Di Valerio, F., Proietti, M., Ragno, A., Capobianco, R.: Cip-net: Continual interpretable prototype-based network. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 40, pp. 20772–20780 (2026)
6. Diehl, P.U., Neil, D., Binas, J., Cook, M., Liu, S.C., Pfeiffer, M.: Fast-classifying, high-accuracy spiking deep networks through weight and threshold balancing. In: 2015 International Joint Conference on Neural Networks (IJCNN). pp. 1–8 (2015). <https://doi.org/10.1109/IJCNN.2015.7280696>
7. Ebrahimi, S., Petryk, S., Gokul, A., Gan, W., Gonzalez, J.E., Rohrbach, M., et al.: Remembering for the right reasons: Explanations reduce catastrophic forgetting. In: *International Conference on Learning Representations* (2021), <https://openreview.net/forum?id=tHgJoMfy6nI>
8. Kim, Y., Panda, P.: Revisiting batch normalization for training low-latency deep spiking neural networks from scratch. *Frontiers in neuroscience* **15**, 773954 (2021)
9. Kim, Y., Panda, P.: Visual explanations from spiking neural networks using interspike intervals. *Scientific reports* **11**(1), 19037 (2021)
10. Krizhevsky, A.: Learning multiple layers of features from tiny images. Tech. rep., University of Toronto (2009)
11. Lapique, L.: Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation. *Journal de physiologie et de pathologie générale* **9**, 620–635 (1907)
12. Maass, W.: Networks of spiking neurons: The third generation of neural network models. *Neural Networks* **10**(9), 1659–1671 (1997). [https://doi.org/10.1016/S0893-6080\(97\)00011-7](https://doi.org/10.1016/S0893-6080(97)00011-7), <https://www.sciencedirect.com/science/article/pii/S0893608097000117>
13. McCloskey, M., Cohen, N.J.: Catastrophic interference in connectionist networks: The sequential learning problem (1989). [https://doi.org/10.1016/S0079-7421\(08\)60536-8](https://doi.org/10.1016/S0079-7421(08)60536-8), <https://www.sciencedirect.com/science/article/pii/S0079742108605368>
14. Minhas, M.F., Putra, R.V.W., Awwad, F., Hasan, O., Shafique, M.: Continual learning with neuromorphic computing: Foundations, methods, and emerging applications. *IEEE Access* (2025)

15. Mishra, R., Suri, M.: A survey and perspective on neuromorphic continual learning systems. *Frontiers in Neuroscience* **17**, 1149410 (2023). <https://doi.org/10.3389/fnins.2023.1149410>
16. Neftci, E.O., Mostafa, H., Zenke, F.: Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks. *IEEE Signal Processing Magazine* **36**(6), 51–63 (2019). <https://doi.org/10.1109/MSP.2019.2931595>, <https://arxiv.org/pdf/1901.09948.pdf>
17. Panda, P., Allred, J.M., Ramanathan, S., Roy, K.: Asp: Learning to forget with adaptive synaptic plasticity in spiking neural networks. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* **8**(1), 51–64 (Mar 2018). <https://doi.org/10.1109/jetcas.2017.2769684>
18. Proietti, M., Ragno, A., Capobianco, R.: Memory replay for continual learning with spiking neural networks. In: 2023 IEEE 33rd International Workshop on Machine Learning for Signal Processing (MLSP). pp. 1–6 (2023). <https://doi.org/10.1109/MLSP55844.2023.10285911>
19. Proietti, M., Ragno, A., Capobianco, R.: Xai-guided continual learning: Rationale, methods, and future directions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **15**(4), e70046 (2025)
20. Robbins, H., Monro, S.: A stochastic approximation method. *The Annals of Mathematical Statistics* **22**(3), 400–407 (1951). <https://doi.org/10.1214/aoms/1177729586>
21. Rolnick, D., Ahuja, A., Schwarz, J., Lillicrap, T., Wayne, G.: Experience replay for continual learning. *Advances in neural information processing systems* **32** (2019)
22. Rymarczyk, D., van de Weijer, J., Zieliński, B., Twardowski, B.: Icicle: Interpretable class incremental continual learning (2023), <https://arxiv.org/abs/2303.07811>
23. Saha, G., Roy, K.: Saliency guided experience packing for replay in continual learning. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. pp. 5273–5283 (2023)
24. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-cam: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision* **128**(2), 336–359 (Oct 2019). <https://doi.org/10.1007/s11263-019-01228-7>
25. Sengupta, A., Ye, Y., Wang, R., Liu, C., Roy, K.: Going deeper in spiking neural networks: VGG and Residual architectures. *Frontiers in Neuroscience* **13**, 95 (2019). <https://doi.org/10.3389/fnins.2019.00095>, <https://www.frontiersin.org/articles/10.3389/fnins.2019.00095>
26. Shen, J., Ni, W., Xu, Q., Tang, H.: Efficient spiking neural networks with sparse selective activation for continual learning. *Proceedings of the AAAI Conference on Artificial Intelligence* **38**(1), 611–619 (Mar 2024). <https://doi.org/10.1609/aaai.v38i1.27817>, <https://ojs.aaai.org/index.php/AAAI/article/view/27817>
27. Skatchkovsky, N., Jang, H., Simeone, O.: Bayesian continual learning via spiking neural networks. *Frontiers in Computational Neuroscience* **Volume 16 - 2022** (2022). <https://doi.org/10.3389/fncom.2022.1037976>, <https://www.frontiersin.org/journals/computational-neuroscience/articles/10.3389/fncom.2022.1037976>
28. Tavanaei, A., Ghodrati, M., Kheradpisheh, S.R., Masquelier, T., Maida, A.: Deep learning in spiking neural networks. *Neural Networks* **111**, 47–63 (2019). <https://doi.org/10.1016/j.neunet.2018.12.002>, <https://www.sciencedirect.com/science/article/pii/S0893608018303332>

29. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: European conference on computer vision. pp. 818–833. Springer (2014)