37th Conference on Neural Information Processing Systems (NeurIPS 2023)

NEURAI

TriRE: A Multi-Mechanism Learning Paradigm for Continual Knowledge Retention and Promotion

Preetha Vijayan*, Prashant Bhat*, Bahram Zonooz†, Elahe Arani†

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

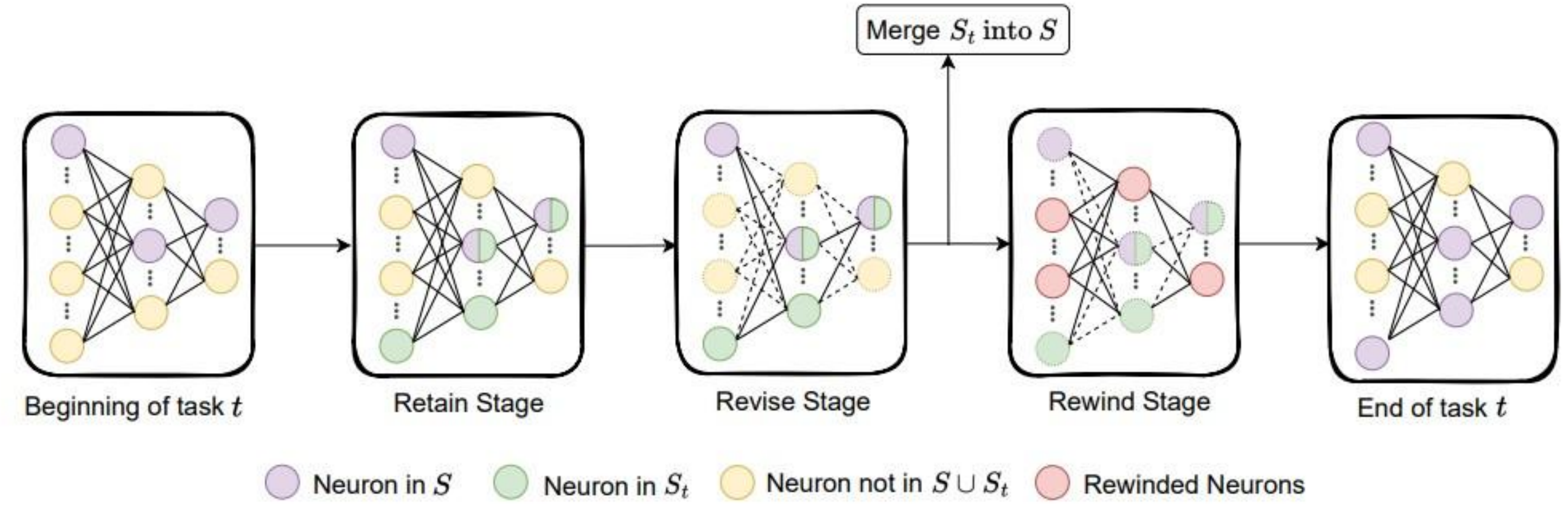
NEURAL INFORMATION PROCESSING SYSTEMS

Introduction

- Continual learning systems often struggles with **Stability-Plasticity Dilemma.** [1]
- Existing Approaches: Parameter Isolation, Weight Regularization, and Experience Rehearsal. [2]
  - Limitations: Capacity Saturation and Scalability Issues, Class Discrimination Challenges, and Overfitting on Buffered Data.
- The human brain orchestrates CL through the dynamic interplay of neurophysiological processes,[3] encompassing
  - Metaplasticity
  - Neurogenesis
  - Experience replay
  - Active forgetting, etc.

Hypothesis: By holistically combining these neurophysiological aspects instead of treating them as competing methods, a more comprehensive solution can be developed to address the stability-plasticity dilemma in continual learning.

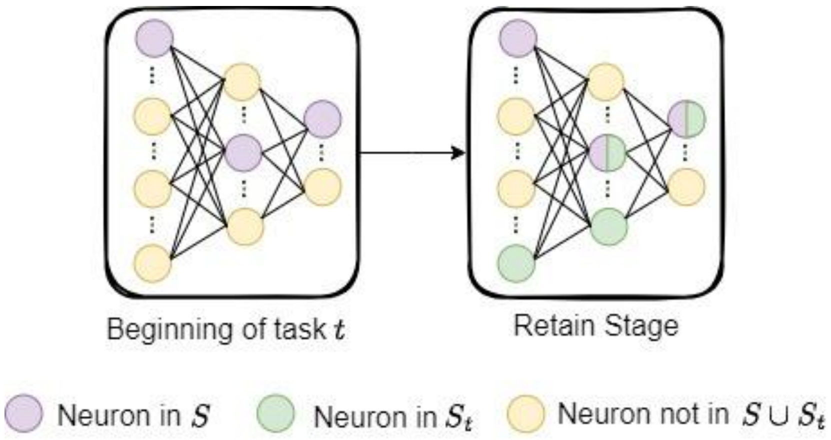
Proposed Methodology - TriRE



Inspired by the biological underpinnings of the CL mechanisms in the brain, we propose ‘REtain, REvise & REwind’ (TriRE), a novel CL paradigm to mitigate catastrophic forgetting.

TriRE - Retain, Revise, Rewind

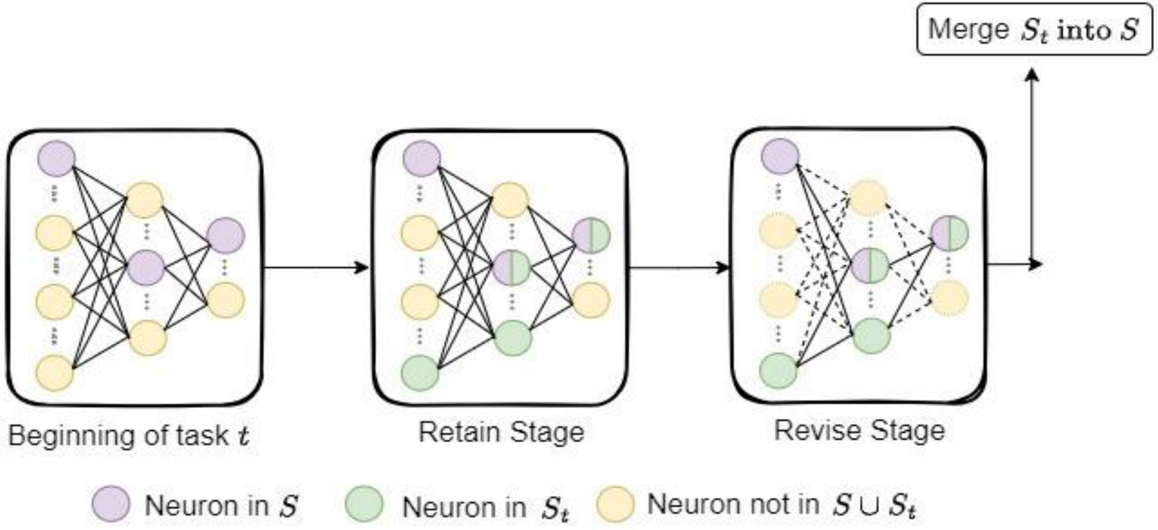
- Retain :
  - Inspired by the brain's use of context-dependent gating for selective filtering of neural information.
  - Induces modularity by training a hyper-network and extracting a subnetwork, S_t representing the current task's knowledge.
  - Achieved through activation pruning followed by weight pruning.



TriRE - Retain, Revise, Rewind

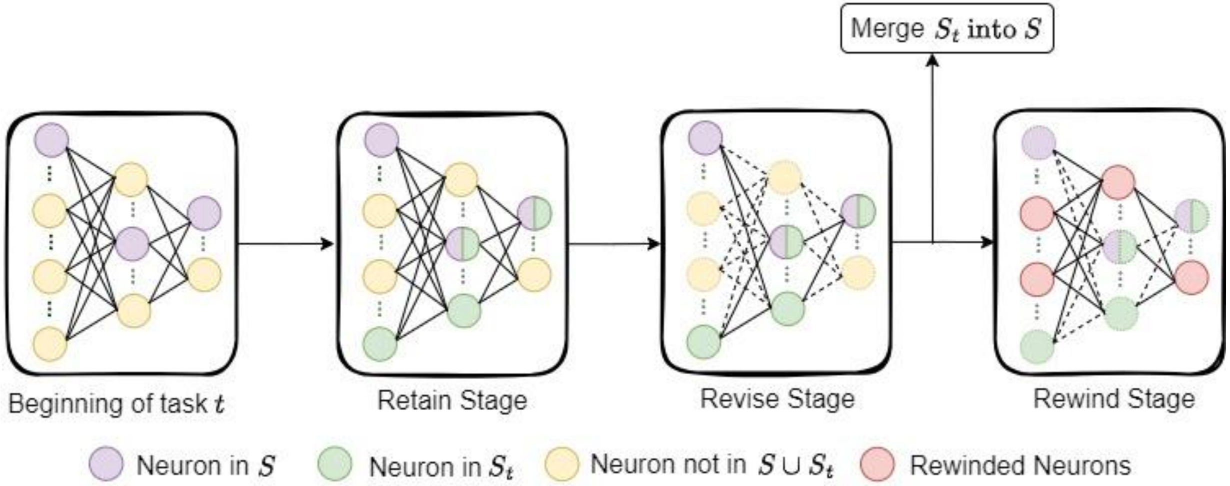
- Revise :

    - Draws inspiration from biological processes such as neurogenesis and metaplasticity.
    - Jointly fine-tunes the task-specific subnetwork (S_t) and the cumulative subnetwork from past tasks (S)
    - Extracted subnetwork is integrated with the cumulative mask from past tasks.



TriRE - Retain, Revise, Rewind

- Rewind :
    - Draws inspiration from the brain's active forgetting mechanism.
    - The weights not in the cumulative subnetwork is rewound to a state where it has learned essential features.
    - These weights are then fine-tuned for a few epochs using current task data.
    - This reactivates less active neurons and readies them for subsequent learning tasks.



Experimental Results

Table 1: Comparison of prior methods across various CL scenarios. We provide the average top-1 (%) accuracy of all tasks after training. † Results of the single EMA model.

Buffer size	Methods	Seq-CIFAR10		Seq-CIFAR100		Seq-TinyImageNet	
		Class-IL	Task-IL	Class-IL	Task-IL	Class-IL	Task-IL
-	SGD	19.62±0.05	61.02±3.33	17.49±0.28	40.46±0.99	07.92±0.26	18.31±0.68
	Joint	92.20±0.15	98.31±0.12	70.56±0.28	86.19±0.43	59.99±0.19	82.04±0.10
-	LwF	19.61±0.05	63.29±2.35	18.47±0.14	26.45±0.22	8.46±0.22	15.85±0.58
	oEWC	19.49±0.12	68.29±3.92	-	-	7.58±0.10	19.20±0.31
	SI	19.48±0.17	68.05±5.91	-	-	6.58±0.31	36.32±0.13
200	ER	44.79±1.86	91.19±0.94	21.40±0.22	61.36±0.35	8.57±0.04	38.17±2.00
	DER++	64.88±1.17	91.92±0.60	29.60±1.14	62.49±1.02	10.96±1.17	40.87±1.16
	CLS-ER†	61.88±2.43	**93.59**±0.87	43.38±1.06	**72.01**±0.97	17.68±1.65	52.60±1.56
	ER-ACE	62.08±1.44	92.20±0.57	35.17±1.17	63.09±1.23	11.25± 0.54	44.17±1.02
	Co^{2}L	65.57±1.37	93.43±0.78	31.90±0.38	55.02±0.36	13.88±0.40	42.37±0.74
	GCR	64.84±1.63	90.8±1.05	33.69±1.40	64.24±0.83	13.05±0.91	42.11±1.01
	DRI	65.16±1.13	92.87±0.71	-	-	17.58±1.24	44.28±1.37
	TriRE	**68.17**±0.33	92.45±0.18	**43.91**±0.18	71.66±0.44	**20.14**±0.19	**55.95**±0.78

Ablation Analysis

Table 2: Comparison of the contribution of each phase in TriRE. Note that the combination of *Revise* alone or *Revise & Rewind* has not been considered, as it is not feasible without the *Retain* phase.

Retain	Revise	Rewind	Seq-CIFAR100		Seq-TinyImageNet	
			Class-IL	Task-IL	Class-IL	Task-IL
✓	✗	✗	38.01	66.23	11.54	40.22
✓	✓	✗	33.08	60.03	8.44	31.90
✓	✗	✓	43.03	**72.09**	16.25	48.89
✓	✓	✓	**43.91**	71.66	**20.14**	**55.95**

- Retain focuses on reducing task interference but lacks in forward transfer and weight reuse.
- Combining Retain and Revise solidifies knowledge but encounters capacity issues.
- Retain and Rewind together encourage efficient knowledge delimitation but sacrifice forward transfer.
- Synergistic integration of all three stages consistently delivers the most robust results in both datasets.

Conclusion and Future Work

- TriRE is an innovative CL paradigm inspired by various neurophysiological mechanisms in the brain.
- Each task in TriRE is divided into stages, including the retention of active neurons, knowledge revision, and promotion of less active neurons for future tasks.
- TriRE significantly reduces task interference and outperforms individual CL methods.
- In the Seq-TinyImageNet dataset, TriRE achieves a 14% improvement over rehearsal-based baselines, surpasses the best parameter isolation method by 7%, and nearly doubles the performance of the best weight regularization approach.
- Future research directions include reducing computational and memory overhead, adapting TriRE for task-free CL with recurring classes, and leveraging intrinsic data structures within tasks.

Thank You

Contact: Preetha Vijayan
Email: preethai35@gmail.com

References

1. Mermillod, M., Bugaiska, A., & Bonin, P. (2013). The stability-plasticity dilemma: Investigating the continuum from catastrophic forgetting to age-limited learning effects. *Frontiers in psychology*, *4*, 504.

2. Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural networks*, *113*, 54-71.

3. Dhireesha Kudithipudi et al. “Biological underpinnings for lifelong learning machines”. In: Nature Machine Intelligence 4.3 (2022), pp. 196–210.