Adaptive Continual Learning Tsetlin Machines

Chong Tang
University of Southampton
chong.tang@soton.ac.uk

Neelam Singh
University of Southampton
n.singh@soton.ac.uk

Jagmohan Chuahan
University of Southampton
j.chauhan@soton.ac.uk

I. Introduction

Continual learning (CL) is crucial for achieving human-like adaptability in machine intelligence, allowing models to learn from new experiences without forgetting previous knowledge. This capability is critical in fields such as robotics and natural language processing, for example, enhancing efficiency and user experience in real-time applications like home assistive robots. However, challenges like catastrophic forgetting, where a model loses knowledge in previous tasks when learning new ones, complicate CL [1]. Techniques like Elastic Weight Consolidation (EWC), Synaptic Intelligence (SI), Experience Replay (ER), Deep Generative Replay (DGR), and iCaRL have been developed to mitigate forgetting by regularizing parameters or using past examples [2]. However, these methods face limitations in computational efficiency and the ability to incorporate new class categories.

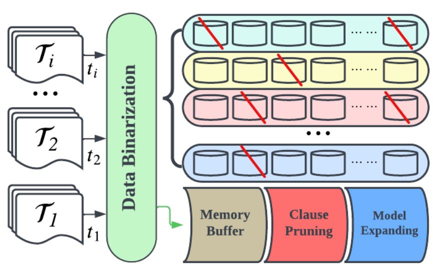

Fig. 1: Overview of the AdaTM architecture

II. Adaptive Tsetlin Machine

We propose the Adaptive Tsetlin Machine (AdaTM) as a solution for CL tasks, offering a novel approach that contrasts with traditional neural networks by leveraging the logic-based structure of the Tsetlin Machine (TM) [3] for edge computing efficiency. AdaTM dynamically adjusts its architecture for new tasks, incorporates a class-balance memory buffer, and uses clause confidence-based pruning to maintain knowledge without the need for intensive recalibrations. This adaptability, coupled with computational efficiency, makes AdaTM especially suitable for resource-constrained environments like edge computing.

A. Experimental Setup

AdaTM's evaluation performs across MNIST, FashionMNIST, AudioMNIST, TESS, and PAMAP2 datasets in class-incremental CL tasks. It compares traditional networks (MLP, LeNet5, ResNet18, and their binary versions) and TM variations against benchmarks like EWC, LwF, SI, DGR, ER, iCaRL, and GDumb. Metrics include accuracy, memory efficiency, and energy use, with tests on laptops and Raspberry Pis.

B. Performance and Impact

AdaTM demonstrated superior performance and low forgetting measures across multiple datasets, including MNIST (97.29%), FashionMNIST (82.67%), TESS (96.26%), and PAMAP2 (74.80%), surpassing benchmark algorithms in accuracy. This contrasts with the significantly lower performance of strategies like EWC, LwF, and SI, which often fell below 50% accuracy in similar settings, underscoring AdaTM's effective learning and knowledge retention in class-incremental continual learning scenarios.

For efficiency tests, AdaTM and AdaTM+Prune showed marked advantages in memory use and processing speed on datasets like MNIST, FashionMNIST, and PAMAP2, achieving up to 35x improvements over GDumb's ResNet and better memory efficiency than iCaRL's MLP. AdaTM, especially its pruned version, significantly reduced latency (up to 52.37% on FashionMNIST) while maintaining accuracy. Energy consumption analyses on a Raspberry Pi highlighted AdaTM's ability to notably decrease energy use, particularly with pruning, balancing speed, and energy efficiency without sacrificing performance, making it ideal for edge computing applications.

III. Conclusion

The AdaTM framework outperforms traditional neural networks in accuracy, memory, and speed in CL tasks. It also introduces an effective pruning mechanism, promising for real-world use. Future efforts will aim to boost efficiency and develop advanced pruning methods, enhancing AdaTM for complex datasets. This work advances sustainable, scalable CL, especially in edge AI.

References

[1] C. V. Nguyen, A. Achille, M. Lam, T. Hassner, V. Mahadevan, and S. Soatto, "Toward understanding catastrophic forgetting in continual learning," *ArXiv*, vol. abs/1908.01091, 2019.

[2] M. De Lange, R. Aljundi, M. Masana, S. Parisot, X. Jia, A. Leonardis, G. Slabaugh, and T. Tuytelaars, "A continual learning survey: Defying forgetting in classification tasks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 7, pp. 3366–3385, 2021.

[3] O.-C. Granmo, "The tsetlin machine–a game theoretic bandit driven approach to optimal pattern recognition with propositional logic," *arXiv preprint arXiv:1804.01508*, 2018.