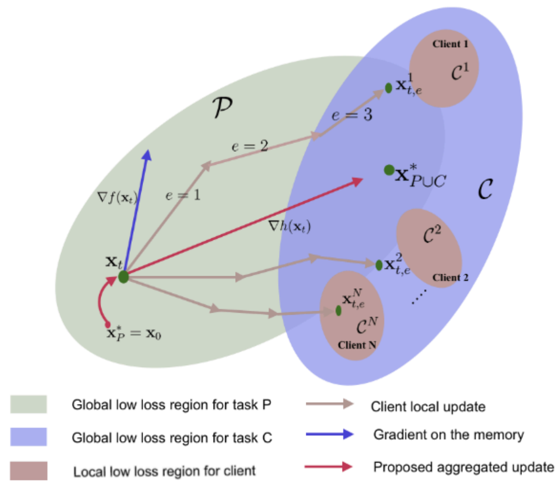
# REPLAY-BASED CONTINUAL FEDERATED LEARNING USING INCREMENTALLY AGGREGATED GRADIENTS

Keywords: Continual Federated Learning, Catastrophic Forgetting, Lifelong Learning, Privacy-Preserving AI, Replay Memory




The holy grail of machine learning is to build systems that learn continuously, adapt seamlessly, and preserve past knowledge without retraining from scratch. While Continual Learning (CL) enables lifelong adaptation [1], and Federated Learning (FL) ensures privacy-preserving collaboration, combining the two in Continual Federated Learning (CFL) introduces a fundamental roadblock: global catastrophic forgetting. Models trained on streaming client data often fail to retain past expertise when learning new tasks, limiting the adaptability and efficiency of CFL systems as an open-world learner. We propose C-FLAG (Continual Federated Learning with Aggregated Gradients), a novel replay-based strategy that addresses global and client-level catastrophic forgetting while ensuring fast convergence. C-FLAG leverages two key innovations: *Effective Gradient Updates* – Each client performs a *single gradient step* on replay memory and *multiple steps* on current data, balancing stability and plasticity. Incremental Aggregated Gradients (IAG) [2,3] – a computation cost-effective scheme that also reduces variance. These two features together enable convergence at a rate of $O(1/\sqrt{T})$ in non-convex settings. We provide a theoretical analysis showing that C-FLAG minimizes forgetting through an adaptive optimization sub-problem, translating CFL into an iterative algorithm with adaptive learning rates. Empirically, across task-incremental benchmarks, C-FLAG consistently outperforms state-of-the-art baselines in both accuracy (Fig 1) and forgetting reduction. A motivating use case is edge streaming analytics (e.g., real-time surveillance, industrial IoT, and autonomous systems), where memory and communication constraints make retraining infeasible. C-FLAG enables these systems to learn continually from streaming private data while preserving prior knowledge. Key contributions of this work are as follows: 1) First replay-based CFL framework with formal convergence guarantees in the non-convex regime. 2) A new gradient composition strategy that jointly mitigates bias, drift, and forgetting. 3) Empirical evidence of superior performance under heterogeneous data, client variability, and replay buffer constraints. C-FLAG demonstrates that continual, privacy-preserving, and scalable federated learning is possible without sacrificing past knowledge, bringing us one step closer to lifelong AI at the edge.

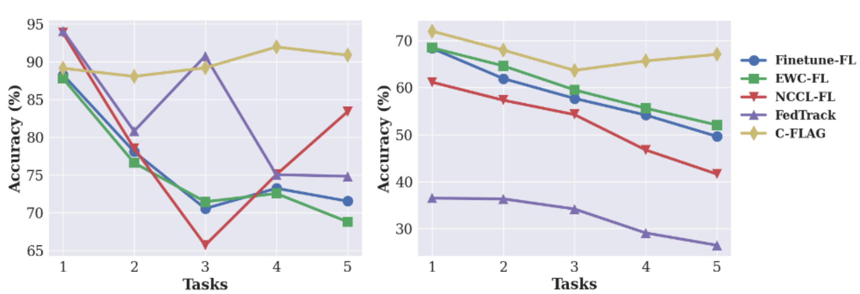



**Fig 1: Average accuracy across tasks for Split-CIFAR10 (left) and Split-CIFAR100 (right)**

[1] *Liyuan Wang, Xingxing Zhang, Hang Su, and Jun Zhu. A comprehensive survey of continual learning: Theory, method and application. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(8):5362–5383, 2024*

[2] *Aritra Mitra, Rayana Jaafar, George J Pappas, and Hamed Hassani. Federated learning with incrementally aggregated gradients. In 2021 60th IEEE Conference on Decision and Control (CDC), pages 775–782. IEEE, 2021.*

[3] *Seungyub Han, Yeongmo Kim, Taehyun Cho, and Jungwoo Lee. On the convergence of continual learning with adaptive methods. In Uncertainty in Artificial Intelligence, pages 809–818. PMLR, 2023.*