International Journal of Research and Applied Innovations (IJRAI)

| ISSN: 2455-1864 | www.ijrai.com | editor@ijrai.com | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 7, Issue 5, September–October 2024||

DOI:10.15662/IJRAI.2024.0705002

Continual Learning Frameworks for Edge AI Applications

Rohan Vijay Verma

Government Dungar College, Bikaner, Rajasthan, India

ABSTRACT: Edge AI has emerged as a transformative paradigm enabling real-time intelligent processing close to data sources, thereby reducing latency, preserving privacy, and optimizing bandwidth usage. However, deploying AI models on edge devices introduces unique challenges due to resource constraints and dynamic environments. One critical challenge is enabling edge devices to adapt to new data over time without forgetting previously learned knowledge—a problem addressed by continual learning frameworks.

Continual learning (CL) enables models to learn from a stream of data incrementally, making them well-suited for edge AI applications where data distributions evolve and retraining on centralized servers is often infeasible. This paper presents a comprehensive overview of continual learning frameworks tailored for edge AI, analyzing various approaches including rehearsal, regularization, and architectural methods.

We examine the suitability of different CL techniques in edge contexts, focusing on memory efficiency, computational overhead, and robustness to concept drift. Additionally, the paper discusses strategies for overcoming catastrophic forgetting—a key issue where models lose prior knowledge when trained on new data. By leveraging lightweight replay buffers, knowledge distillation, and dynamic model expansion, recent frameworks have shown promise in maintaining accuracy over extended learning periods.

Our research methodology involves evaluating existing continual learning algorithms on edge-relevant benchmarks using resource-limited devices, analyzing performance trade-offs between accuracy, latency, and memory consumption. We further explore hybrid models integrating federated learning with continual learning to harness distributed edge data collaboratively while preserving privacy.

Results demonstrate that carefully designed continual learning frameworks significantly improve model adaptability and sustainability on edge devices. However, challenges remain around optimizing computation, communication overhead, and security.

This study provides valuable insights for researchers and practitioners seeking to deploy intelligent, adaptive AI on edge platforms. Future work directions include enhancing lightweight CL models, developing adaptive resource management strategies, and integrating privacy-preserving mechanisms to realize robust and scalable edge AI systems.

KEYWORDS: Continual Learning, Edge AI, catastrophic Forgetting, Rehearsal Methods, Regularization Methods, Architectural Methods, Federated Learning, Resource-Constrained Devices, Concept Drift, Model Adaptation,

I. INTRODUCTION

The proliferation of Internet of Things (IoT) devices and advances in AI have accelerated the deployment of intelligent applications at the network edge. Edge AI refers to running AI inference and learning tasks on local devices, such as smartphones, sensors, and embedded systems, to enable low-latency, privacy-sensitive, and bandwidth-efficient operations. This shift away from cloud-centric models requires novel AI frameworks capable of adapting to continuously changing environments with limited computational resources.

A major challenge in edge AI is enabling models to learn incrementally from new data without degrading previously acquired knowledge—a problem known as catastrophic forgetting. Unlike traditional batch learning, where models are trained on static datasets, edge devices must handle non-stationary data streams and evolving tasks. Continual learning

International Journal of Research and Applied Innovations (IJRAI)

| ISSN: 2455-1864 | www.ijrai.com | editor@ijrai.com | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 7, Issue 5, September–October 2024||

DOI:10.15662/IJRAI.2024.0705002

(CL) frameworks offer solutions by allowing models to update incrementally, preserving old knowledge while integrating new information.

Existing CL methods broadly fall into rehearsal (memory replay), regularization (penalty-based constraints), and architectural (dynamic model expansion) categories. These approaches help mitigate forgetting but vary widely in resource demands and effectiveness, which is critical in the constrained environments of edge devices.

Moreover, privacy and data communication concerns motivate combining CL with federated learning, enabling collaborative training across multiple edge nodes without centralized data aggregation.

This paper provides a comprehensive survey and evaluation of continual learning frameworks tailored for edge AI applications. We analyze their advantages, limitations, and suitability under typical edge constraints, including limited memory, compute power, and intermittent connectivity.

By identifying gaps and opportunities, this work aims to guide future research and deployment strategies for robust, adaptive AI systems capable of lifelong learning on edge platforms.

II. LITERATURE REVIEW

Continual learning has garnered significant attention as a key enabler of adaptive AI systems capable of learning from data streams. Early work by McCloskey and Cohen (1989) introduced the catastrophic forgetting problem, where neural networks forget old knowledge when trained on new tasks. Subsequent research developed several approaches to mitigate this effect.

Rehearsal methods, such as Experience Replay (Robins, 1995) and its variants, store a subset of past data to retrain alongside new data, effectively reducing forgetting. Later, frameworks like GEM (Gradient Episodic Memory) by Lopez-Paz and Ranzato (2017) improved memory efficiency by selectively replaying samples.
Regularization-based methods constrain model updates to preserve old knowledge. Elastic Weight Consolidation (EWC) by Kirkpatrick et al. (2017) penalizes changes to important parameters, enabling incremental learning with minimal memory overhead.

Architectural approaches, such as Progressive Neural Networks (Rusu et al., 2016), dynamically expand model capacity to accommodate new tasks without interference. However, these methods increase model complexity and may be unsuitable for resource-constrained edge devices.

Recent surveys (Parisi et al., 2019; De Lange et al., 2021) comprehensively categorize CL techniques and discuss their strengths and limitations, emphasizing the need for lightweight and efficient frameworks for edge AI.

Edge-specific continual learning introduces challenges including limited computational power, energy constraints, and privacy requirements. To address these, federated learning frameworks (McMahan et al., 2017) have been integrated with CL to enable decentralized, privacy-preserving learning across multiple edge devices (Chen et al., 2020).

Other research explores hardware-aware CL optimizations (Li et al., 2020) and adaptive memory management to suit edge scenarios. Despite progress, achieving a balance between learning efficacy, resource efficiency, and privacy remains an open challenge, motivating ongoing research into novel CL frameworks for edge AI.

III. RESEARCH METHODOLOGY

This study employs an experimental and analytical approach to evaluate continual learning frameworks in edge AI contexts. We select representative CL algorithms from rehearsal (e.g., Experience Replay, GEM), regularization (EWC), and architectural (Progressive Neural Networks) categories, implemented using popular deep learning libraries. We conduct experiments on standard continual learning benchmarks such as Split-MNIST, CIFAR-100 incremental tasks, and CORe50, focusing on scenarios with streaming data and class incremental learning to mimic real-world edge applications.

International Journal of Research and Applied Innovations (IJRAI)

| ISSN: 2455-1864 | www.ijrai.com | editor@ijrai.com | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 7, Issue 5, September–October 2024||

DOI:10.15662/IJRAI.2024.0705002

Edge device constraints are emulated on hardware platforms including Raspberry Pi 4 and NVIDIA Jetson Nano, measuring model accuracy, inference latency, memory footprint, and energy consumption during continual training and evaluation phases.

To assess privacy-preserving learning, we simulate federated learning settings where multiple edge nodes collaboratively train models with periodic aggregation, evaluating communication overhead and model convergence.
We further analyze the impact of different replay buffer sizes and regularization strengths on catastrophic forgetting and model adaptation.

Qualitative analysis includes studying model robustness under concept drift and varying data distributions, a typical challenge for deployed edge AI.

Security and privacy implications are reviewed, considering how continual learning frameworks can be integrated with encryption and secure hardware.

Our methodology balances quantitative metrics with practical deployment considerations, offering insights into trade-offs between learning performance and resource utilization in edge environments.

Advantages

- Enables adaptive AI that learns continuously without requiring centralized retraining.
- Reduces catastrophic forgetting, maintaining model performance on previous tasks.
- Facilitates deployment on resource-constrained edge devices.
- Supports privacy by limiting data transfer via decentralized learning.
- Allows real-time model updates to handle dynamic environments and concept drift.

Disadvantages

- Replay-based methods require memory buffers, which may be limited on edge devices.
- Regularization approaches may not fully prevent forgetting in complex tasks.
- Architectural expansion increases model size, unsuitable for devices with limited storage.
- Federated continual learning introduces communication overhead and synchronization challenges.
- Balancing accuracy, resource consumption, and privacy remains complex.

IV. RESULTS AND DISCUSSION

Experiments indicate that rehearsal methods such as Experience Replay provide strong baseline performance with manageable resource consumption, but require careful buffer management. GEM reduces forgetting further but at increased computational cost.

Regularization techniques like EWC achieve efficient learning with lower memory usage but show reduced accuracy on long task sequences. Architectural methods offer best retention but are impractical for most edge deployments due to model growth.

Federated learning integration shows promise in preserving privacy while enabling knowledge sharing, though network latency and communication bandwidth can degrade performance.
Energy profiling reveals that continual learning updates increase power consumption, emphasizing the need for energy-efficient algorithms.

Overall, hybrid approaches combining replay and regularization yield favorable trade-offs, but real-world edge scenarios demand further optimization.

V. CONCLUSION

Continual learning frameworks represent a crucial advancement for enabling adaptive AI on edge devices. By mitigating catastrophic forgetting and facilitating incremental updates, these techniques allow edge models to evolve with changing data distributions in resource-constrained environments.

Our study highlights the strengths and weaknesses of various CL methods and the importance of considering device limitations, privacy, and communication overhead in framework design.

Future research should focus on lightweight, energy-efficient continual learning algorithms tailored for heterogeneous edge hardware, as well as robust federated CL protocols that maintain privacy and scalability.

VI. FUTURE WORK

- Developing memory-efficient replay strategies optimized for edge hardware constraints.
- Exploring dynamic regularization methods adaptive to changing data distributions.
- Designing modular architectures enabling controlled model expansion on edge devices.
- Enhancing federated continual learning frameworks for asynchronous and heterogeneous networks.
- Integrating continual learning with secure hardware enclaves for privacy and integrity.
- Investigating energy-aware continual learning algorithms to extend edge device battery life.

REFERENCES

1. Kirkpatrick, J., Pascanu, R., Rabinowitz, N., et al. (2017). *Overcoming catastrophic forgetting in neural networks*. Proceedings of the National Academy of Sciences, 114(13), 3521-3526.
2. Lopez-Paz, D., & Ranzato, M. (2017). *Gradient Episodic Memory for Continual Learning*. Advances in Neural Information Processing Systems (NeurIPS).
3. Rusu, A. A., Rabinowitz, N. C., Desjardins, G., et al. (2016). *Progressive Neural Networks*. arXiv preprint arXiv:1606.04671.
4. Parisi, G. I., Kemker, R., Part, J. L., et al. (2019). *Continual lifelong learning with neural networks: A review*. Neural Networks, 113, 54-71.
5. McMahan, H. B., Moore, E., Ramage, D., et al. (2017). *Communication-efficient learning of deep networks from decentralized data*. Proceedings of AISTATS.
6. Chen, M., Yang, T., Zhang, Z., et al. (2020). *FedWeIT: A Continual Learning Framework for Federated Learning*. Proceedings of AAAI.
7. De Lange, M., Aljundi, R., Masana, M., et al. (2021). *A Continual Learning Survey: Defying Forgetting in Classification Tasks*. IEEE Transactions on Pattern Analysis and Machine Intelligence.
8. Li, Z., Hoi, S. C. H., & Wang, J. (2020). *Learning to remember: A comprehensive survey on memory-augmented neural networks*. arXiv preprint arXiv:1907.06650.
9. Robins, A. (1995). *Catastrophic forgetting, rehearsal and pseudorehearsal*. Connection Science, 7(2), 123-146.
10. McCloskey, M., & Cohen, N. J. (1989). *Catastrophic interference in connectionist networks: The sequential learning problem*. Psychology of Learning and Motivation, 24, 109-165.