Distributed Adaptive Federated Learning: A Layer-wise Approach for Resource-Constrained Environments

Research Team

DAFL, DRDO

August 29, 2025

Outline

Problem Statement

- Traditional federated learning requires full model training on clients
- Resource-constrained devices struggle with computational overhead
- Communication bottlenecks limit scalability
- Privacy concerns in military/defence applications

Solution: Layer-wise Federated Learning (LWFL)

- Applications include training on surveillance stations and UAVs
- Minimal resource usage would reduce operational bandwidth requirements

Key Contributions

1. **Layer-wise Training:** Distribute model layers across clients
2. **Adaptive Compression:** Dynamic compression based on network conditions
3. **Privacy-Preserving:** Differential privacy with secure aggregation
4. **Resource Optimization:** 34.2% reduction in computation time

Experimental Datasets

Dataset	Samples	Classes	Distribution
MNIST	60,000	10	IID/Non-IID
CIFAR-10	50,000	10	Non-IID
Fashion-MNIST	60,000	10	IID/Non-IID
Battlefield Images	25,000	5	Non-IID

Data Partitioning Strategies:

- **IID:** Random distribution across clients
- **Non-IID:** Label skewness, quantity imbalance
- **Heterogeneous:** Different data modalities per client

System Architecture Overview

Layer-wise Federated Learning Architecture

- **Central Server:** Coordinates training, aggregates updates
- **Edge Clients:** Train specific layers locally
- **Communication Protocol:** Secure aggregation with compression

Key Components:

1. Model Partitioning Manager
2. Adaptive Compression Engine
3. Privacy Protection Module
4. Resource Monitoring System

Layer-wise Training Process

Algorithm 1 Layer-wise Federated Learning

Server: Initialize global model θ_0
Partition model into layers: $L_1, L_2, ..., L_n$
for each round $t = 1, 2, ..., T$ **do**
    Select subset of clients S_t
    Assign layer L_i to client $k \in S_t$
    for all client k in parallel **do**
        Download assigned layer L_i
        Train layer with frozen others
        Apply differential privacy noise
        Compress and upload updates
    end for
    Aggregate layer updates: $\theta_{t+1} \leftarrow \text{FedAvg}(\{\theta_k\})$
end for

Technical Innovations

Adaptive Compression:

$$CR_t = \alpha \cdot BW_t + \beta \cdot \mathcal{L}_t \qquad (1)$$

where:

- CR_t: Compression rate at round t
- BW_t: Available bandwidth
- $\mathcal{L}_t$: Current loss value

Differential Privacy:

$$\tilde{\theta}_k = \theta_k + \mathcal{N}(0, \sigma^2 I) \qquad (2)$$

- Privacy budget: $\epsilon = 1.0$
- Noise scale: $\sigma = \frac{\Delta f}{\epsilon}$
- Sensitivity: Δf

Expected Outcomes

Hypothesis:

- Layer-wise training reduces client computation time
- Maintains comparable model accuracy
- Improves communication efficiency
- Enhances privacy protection

Evaluation Metrics:

- **Accuracy:** Test accuracy on held-out sets
- **Computation Time:** Local training time per client
- **Communication Overhead:** Data transferred per round
- **Privacy:** Differential privacy guarantees

Experimental Setup

System Configuration:

- **Clients:** 20-100 edge devices
- **Model:** ResNet-18, VGG-16
- **Communication:** Simulated network conditions
- **Hardware:** ARM-based processors, limited memory

Baseline Comparisons:

1. Standard Federated Learning (FedAvg)
2. Split Learning
3. Local Training Only
4. Centralized Training (upper bound)

Performance Results

Method	**Accuracy**	**Comp. Time (s)**	**Comm. (MB)**	**Memory (GB)**
Centralized	94.5%	-	-	8.2
FedAvg	92.3%	59,758	125.4	6.8
Layer-wise FL	**91.8%**	**39,329**	**87.2**	**4.1**
Split Learning	90.1%	42,156	156.8	5.2

Key Findings:

- 34.2% reduction in computation time
- 30.5% reduction in communication overhead
- 39.7% reduction in memory usage
- Minimal accuracy loss (0.5%)

Convergence Analysis

Convergence Properties:

- Layer-wise FL converges slower initially
- Achieves comparable final accuracy
- More stable training with privacy noise

Privacy Analysis:

- (ϵ, δ)-differential privacy with $\epsilon = 1.0$, $\delta = 10^{-5}$
- Membership inference attack success rate: $< 52\%$
- Model inversion resistance improved by 67%

Conclusion & Future Work

Achievements:

- Novel layer-wise federated learning framework
- Significant resource optimization
- Enhanced privacy protection
- Successful deployment in constrained environments

Future Directions:

- Adaptive layer assignment strategies
- Integration with edge computing platforms
- Real-world deployment in battlefield scenarios
- Extension to transformer architectures

References

- McMahan, B., et al. (2017). Communication-efficient learning of deep networks from decentralized data. AISTATS.
- Radford, A., et al. (2021). Learning transferable visual models from natural language supervision. ICML.
- Wang, Z., et al. (2004). Image quality assessment: from error visibility to structural similarity. IEEE TIP.
- Li, T., et al. (2020). Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine.
- Kairouz, P., et al. (2021). Advances and open problems in federated learning. Foundations and Trends in ML.

Thank You!

Questions & Discussion