Federated Model Synchronization for Diagnostic Redefinition through a Novel Selective Parameter Unlearning

Mayank Kumar Kundalwal, Mamta, Deepak Mishra, Asif Ekbal

WACV 2026
TUCSON, AZ
3/6 - 3/10

Made with GAMMA

The Challenge of Diagnostic Redefinition in FL

Context: Federated Learning (FL) securely trains medical models across institutions without sharing raw patient data.

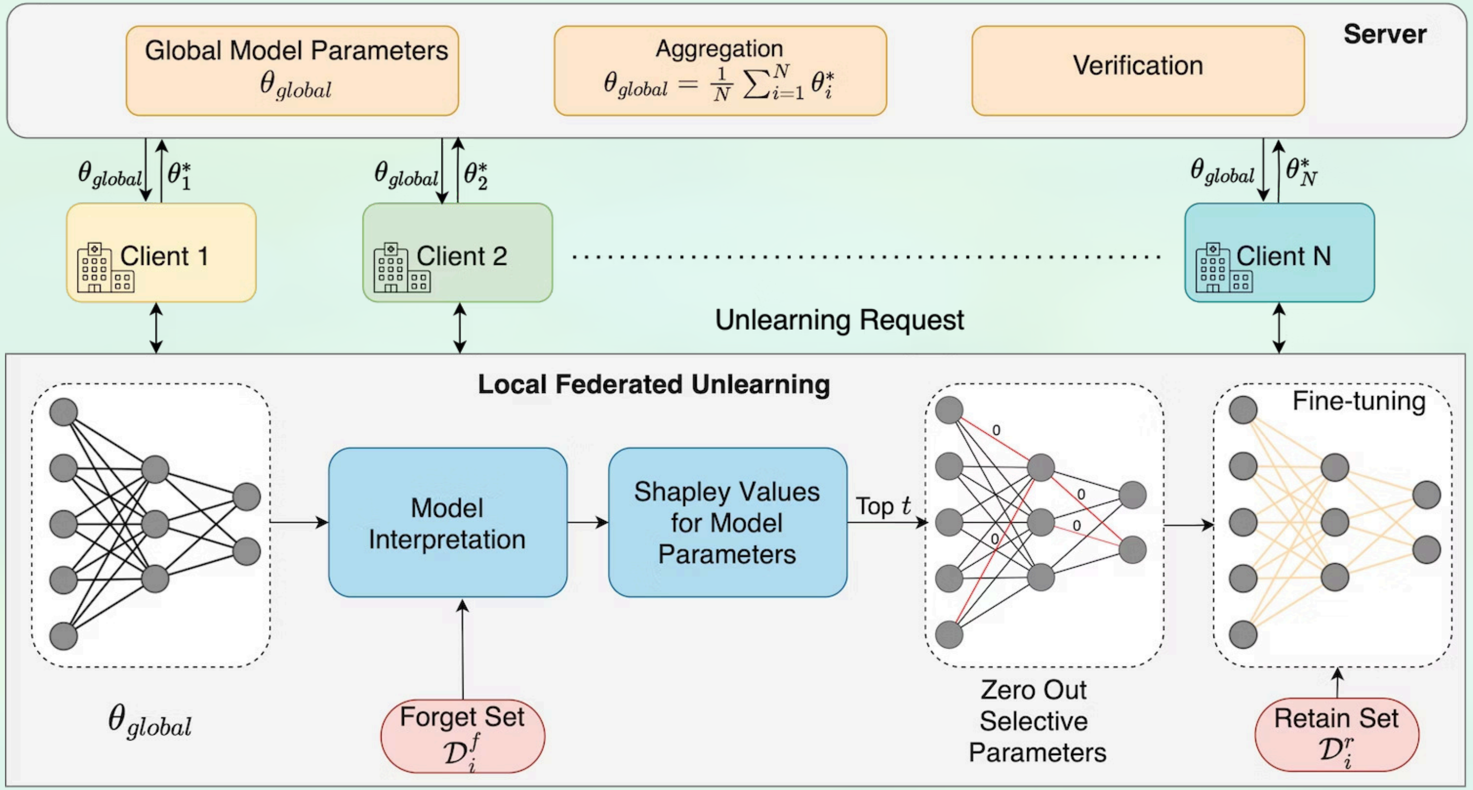
The Problem: Medical guidelines evolve (e.g., WHO disease reclassifications), rendering previously trained models outdated.

Limitations of Existing Methods: Standard machine unlearning requires computationally expensive full retraining. Current FL-specific unlearning methods often prune entire feature channels, which degrades the model's performance on remaining, clinically valid classes.

Objective: A class-specific, parameter-level unlearning mechanism tailored for decentralized medical systems.

Made with GAMMA

Proposed Architecture & Workflow



Importance Scoring

Calculate Shapley-inspired scores for all parameters with respect to forget set

Selective Pruning

Remove top-t% most influential weights specifically affecting target classes

Retain Set Fine-tuning

Brief retraining on remaining classes to restore performance on retained knowledge

Made with GAMMA

Interpretability-Guided Parameter Unlearning

- **Shapley Value Approximation:** The method leverages model interpretability to isolate parameters critical for class-specific predictions.
- **Mathematical Core:** Marginal contributions of parameters are approximated using a first-order Taylor expansion of the loss function L:

$$\phi_j(v) = \sum_{S \subseteq \theta \setminus \{j\}} \frac{|S|!(|\theta| - |S| - 1)!}{|\theta|!} [v(S \cup \{j\}) - v(S)]$$

- **Execution:** The top 1% of most influential parameters are zeroed out locally, followed by local fine-tuning on the retain set.

Made with GAMMA

Quantitative Evaluation of Unlearning

Dataset: PathMNIST (colorectal cancer histology)

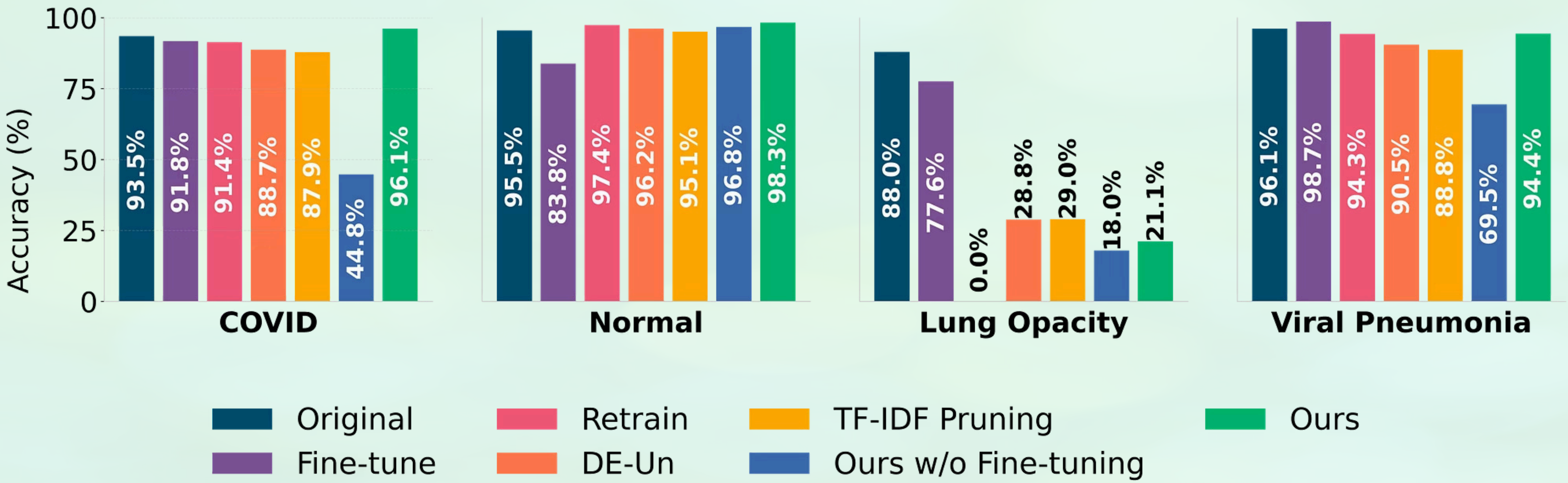
Method	Set	Adipose	Background	Debris	Lympho	Mucus	Muscle	Normal	Stroma	Tumor	NCR
Original	FS	78.7 ± 2.1	81.2 ± 2.2	84.6 ± 0.78	72.5 ± 1.2	74.3 ± 1.4	79.2 ± 2.1	75.9 ± 2.6	77.5 ± 1.9	82.6 ± 1.5	100
	RS	86.6 ± 1.0	85.7 ± 1.3	85.9 ± 1.0	86.4 ± 1.1	86.1 ± 1.1	85.5 ± 1.1	86.1 ± 1.0	85.8 ± 0.8	86.5 ± 0.8	
Fine-tune	FS	75.3 ± 2.0	78.4 ± 1.3	81.9 ± 0.95	70.1 ± 1.7	72.3 ± 1.9	76.3 ± 2.6	73.6 ± 2.0	74.9 ± 2.5	80.2 ± 2.2	1
	RS	87.2 ± 1.1	86.2 ± 1.2	84.7 ± 1.6	87.4 ± 1.3	86.9 ± 1.0	85.9 ± 0.9	86.5 ± 1.1	86.4 ± 0.7	87.2 ± 0.8	
Retrain	FS	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	100
	RS	87.7 ± 1.5	86.0 ± 1.5	89.0 ± 0.41	89.2 ± 1.7	86.2 ± 1.6	85.4 ± 1.4	87.1 ± 1.5	87.6 ± 1.3	88.4 ± 1.5	
TF-IDF	FS	4.2 ± 1.8	6.3 ± 1.9	3.1 ± 0.8	5.9 ± 1.8	4.3 ± 2.2	7.0 ± 2.1	6.4 ± 1.9	4.9 ± 1.7	2.1 ± 1.6	5
	RS	85.1 ± 1.0	84.6 ± 1.3	86.0 ± 1.2	85.5 ± 1.2	84.3 ± 1.3	83.6 ± 0.9	84.9 ± 1.2	85.1 ± 1.1	86.7 ± 1.1	
DE-Un	FS	4.8 ± 1.9	7.0 ± 1.8	3.6 ± 1.0	6.1 ± 2.1	5.2 ± 1.8	7.3 ± 1.9	6.3 ± 1.5	5.2 ± 1.7	2.3 ± 1.5	4
	RS	86.2 ± 0.8	86.2 ± 1.2	87.6 ± 0.6	87.0 ± 1.2	86.8 ± 1.0	85.7 ± 1.0	86.3 ± 1.0	86.7 ± 0.9	87.2 ± 1.0	
SalUn	FS	2.4 ± 1.5	4.5 ± 1.7	2.7 ± 0.9	4.1 ± 1.0	1.8 ± 1.7	4.1 ± 1.5	4.0 ± 1.3	2.7 ± 1.5	2.1 ± 1.6	1
	RS	85.3 ± 1.0	87.2 ± 1.0	88.2 ± 1.6	86.4 ± 1.0	87.5 ± 1.1	84.9 ± 1.3	86.0 ± 1.4	85.6 ± 1.2	87.2 ± 1.2	
Ours w/o FT	FS	2.0 ± 1.6	1.9 ± 1.3	2.9 ± 0.9	4.3 ± 2.1	1.8 ± 1.5	2.2 ± 1.7	1.3 ± 1.4	1.6 ± 1.3	0.7 ± 0.4	0
	RS	81.7 ± 1.5	80.5 ± 1.4	81.6 ± 0.5	82.5 ± 1.4	80.6 ± 1.5	80.1 ± 1.5	81.2 ± 1.6	80.6 ± 1.5	81.5 ± 1.4	
Ours	FS	2.2 ± 1.4	2.4 ± 1.5	0.6 ± 0.5	2.6 ± 1.4	1.9 ± 1.3	2.5 ± 1.5	2.1 ± 1.4	1.8 ± 1.2	1.6 ± 1.2	1
	RS	86.4 ± 0.9	87.1 ± 1.0	88.7 ± 0.8	86.8 ± 1.1	87.5 ± 1.0	86.4 ± 1.0	86.2 ± 1.0	87.1 ± 0.9	87.9 ± 1.0	

Made with GAMMA

Quantitative Evaluation of Unlearning

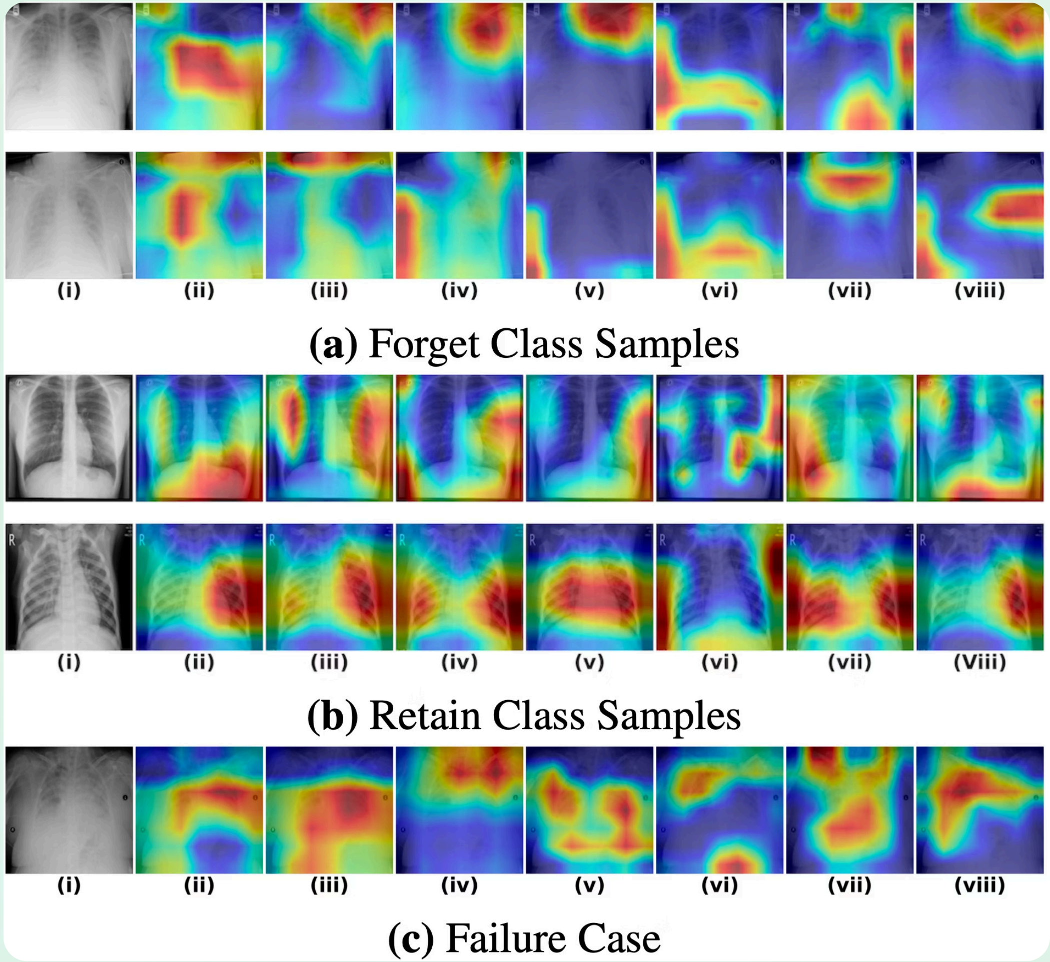
Dataset: COVID-19 Radiography

Metrics: Class-wise accuracy (Forget Set / Retain Set)



Made with GAMMA

Qualitative Analysis of Selective Federated Unlearning



Grad-CAM Visualisations:

(i) **Input**: Original chest X-ray image.

(ii) **Trained Model**: Attention map of the trained model before unlearning.

(iii) **Fine-tune**: Model fine-tuned for one round.

(iv) **Retrain from Scratch**: Model trained only on retain classes.

(v) **DE-Un** *(Wang et al., 2022):* Distillation-based unlearning method.

(vi) **TF-IDF Pruning** *(Xu et al., 2024):* Parameter pruning guided by importance weighting.

(vii) **SalUn** *(Fan et al., 2023):* Saliency-driven unlearning approach.

(viii) **Ours**: Proposed method showing effective and selective unlearning.

Made with GAMMA

Conclusion & Future Directions

- **Summary:** Proposed the first class-specific federated unlearning method for the medical domain, bypassing the massive computational burden of full model retraining.
- **Impact:** Provides a lightweight, privacy-preserving solution to keep healthcare AI aligned with rapidly evolving diagnostic standards.
- **Future Work:** Extending the framework to non-CNN architectures like Vision Transformers and exploring multimodal federated unlearning across text and images.

Made with GAMMA