

FDPTU: Federated Attack-aware Differentially Privacy Tree Unlearning framework for Multimodal Biometric Privacy

Samyak Jain^{*†}, Simran Gupta^{†‡}, Pronaya Bhattacharya[‡], Sandip Roy[§],
Sachin Shetty[¶], Rutvij H. Jhaveri^{||}, Saed Alrabae^{**}

^{*†‡}Department of Computer Science and Engineering, Amity School of Engineering and Technology,
Amity University, Kolkata, India

[‡]Department of Computer Science and Engineering, School of Engineering and Technology,
Sharda University, Greater Noida, India.

^{§¶}Center for Secure & Intelligent Critical Systems, Old Dominion University, Virginia, USA

^{||}School of Technology, Pandit Deendayal Energy University, India

^{**}Department of Information Systems and Security, UAE University

Email: ^{*}samyak.jain18@s.amity.edu, [†]simran.gupta17@s.amity.edu, [‡]pbhattacharya@kol.amity.edu,

[§]sroy@odu.edu, [¶]sshetty@odu.edu, ^{||}rutvij.jhaveri@sot.pdpu.ac.in, ^{**}salrabae@uaeu.ac.ae

[†]Equally contributed as first authors.

Corresponding Authors: Rutvij H. Jhaveri (rutvij.jhaveri@sot.pdpu.ac.in), and Saed Alrabae (salrabae@uaeu.ac.ae)

Abstract—Multimodal biometric authentication increasingly relies on federated learning to reduce centralized exposure of sensitive traits such as face, iris, and fingerprint data. However, federated training remains vulnerable to gradient leakage, adversarial poisoning, deepfake or morph-based manipulation, and inefficient removal of client contributions. This paper presents *FDPTU*, a Federated Attack-aware Differentially Private Tree Unlearning framework for privacy-preserving multimodal biometric learning. The framework combines secure template generation, homomorphic encryption, Gaussian differential privacy, attack-aware update monitoring, and dependency-tree-guided federated unlearning. Suspicious or withdrawn contributions are localized through bounded traversal of update dependencies and removed at the shard level without full retraining. Cryptographic attestation further supports verifiable proof of removal. Experiments on multimodal biometric and deepfake datasets show that *FDPTU* preserves verification utility while reducing residual influence by up to 95% and lowering unlearning cost by approximately 35–45% compared with baseline forgetting strategies.

Index Terms—Biometric Authentication, Dependency Tree, Differential Privacy, Federated Unlearning, Multimodal Systems

I. INTRODUCTION

Multimodal biometric authentication has become central to identity management in healthcare, smart cities, financial services, and critical infrastructure. By combining facial, iris, and fingerprint traits, such systems improve recognition reliability. However, biometric data are inherently sensitive and non-revocable. Once leaked or misused, the same traits cannot be reissued like passwords or tokens, creating long-term identity and privacy risks [1], [2].

Federated Learning (FL) reduces centralized data exposure by allowing clients to train collaboratively while retaining raw biometric data locally [3]. Yet FL does not fully eliminate privacy and security risks. Model gradients and client updates

may still reveal sensitive information through membership inference, attribute inference, or reconstruction attacks [4], [5]. Differential Privacy (DP) can mitigate such leakage, but noise injection often affects convergence stability and recognition accuracy [6]. At the same time, Federated Unlearning (FU) has emerged as a way to remove client contributions without retraining from scratch [7]. Existing methods such as SISA [8], FedEraser [9], selective forgetting [10], [11], and sharding-based unlearning frameworks [12] provide partial solutions, but they remain limited in multimodal biometric settings because they are often attack-agnostic, storage-intensive, or unable to provide precise and verifiable removal.

This paper addresses the following research question: *How can multimodal biometric systems support collaborative learning under strict privacy constraints, resist adversarial contamination, and enable verifiable data removal without sacrificing recognition performance or scalability?*¹

To answer this question, we propose *FDPTU*, a Federated Attack-aware Differentially Private Tree Unlearning framework for multimodal biometric privacy. The framework protects biometric templates through local feature extraction, homomorphic encryption, and Gaussian DP during federated aggregation. It further introduces an attack-aware monitoring layer to detect suspicious deepfake and morph-based updates. Instead of globally retraining the model, *FDPTU* constructs a dependency tree over client updates and performs bounded traversal to identify affected shards. Selective unlearning then removes compromised or withdrawn contributions while preserving benign updates. A cryptographic audit trail provides

¹Initial implementation code is available at <https://github.com/sam-1409/FDPTU.git>

verifiable evidence of removal.

The major contributions of this work are summarized as follows:

- We formulate attack-aware, privacy-preserving multimodal biometric learning with certified federated unlearning under operational and regulatory constraints.
- We design an integrated framework combining secure template protection, Gaussian DP, encrypted aggregation, dependency-based attack tracing, and selective shard-level unlearning.
- We evaluate *FDPTU* on multimodal biometric and deepfake datasets, showing improved privacy resilience, lower residual influence, reduced unlearning cost, and stable recognition utility compared with existing FL and FU baselines.

The remainder of this paper is organized as follows. Section II presents the system model and problem formulation. Section III describes the proposed framework. Section IV reports the experimental evaluation. Section V concludes the paper and outlines future directions.

II. *FDPTU*: SYSTEM MODEL AND PROBLEM FORMULATION

This section presents the system model and optimization objectives of *FDPTU*, as shown in Fig. 1. The framework considers a federated multimodal biometric environment where distributed clients collaboratively train a global authentication model while preserving local data privacy, detecting adversarial updates, and supporting selective unlearning.

A. Federated Multimodal Biometric Setting

Let $\mathcal{C} = \{c_i\}_{i=1}^C$ denote a set of distributed clients, where each client stores biometric samples from multiple modalities such as face, iris, and fingerprint. To avoid direct exposure of sensitive traits, all modality-specific preprocessing is performed locally. For modality $m \in \{1, \dots, M\}$, client c_i generates a protected biometric template as

$$T_{c_i}^{(m)} = f_{\text{ext}}^{(m)}(c_i^{(m)}), \quad (1)$$

where $f_{\text{ext}}^{(m)}(\cdot)$ represents the corresponding feature extraction function. The resulting multimodal templates are encrypted before training, i.e., $\hat{T}_{c_i} = \mathcal{E}(T_{c_i})$, so that neither raw samples nor exposed biometric representations leave the client device.

During each communication round t , every selected client trains locally and sends only protected model updates to the server. To reduce information leakage from transmitted updates, Gaussian differential privacy is applied during aggregation:

$$w^{(t+1)} = w^{(t)} + \sum_{i=1}^C \Delta w_{c_i}^{(t)} + \mathcal{N}(0, \sigma^2 I), \quad (2)$$

where $\Delta w_{c_i}^{(t)}$ is the update submitted by client c_i and $\mathcal{N}(0, \sigma^2 I)$ denotes calibrated Gaussian noise. This formulation allows collaborative learning while limiting cross-client linkability and update-level privacy leakage.

B. Attack-Aware Dependency and Unlearning Model

Although DP reduces leakage, malicious clients may still inject poisoned, deepfake, or morph-based updates into the federated process. Therefore, *FDPTU* incorporates an update monitoring layer before aggregation. Each client update is assigned an anomaly score

$$s_{c_i} = f_{\text{det}}(\Delta w_{c_i}), \quad (3)$$

where $f_{\text{det}}(\cdot)$ denotes the attack detection function. If $s_{c_i} \geq \tau$, the update is marked suspicious and its influence is traced through a dependency tree constructed from historical aggregation records.

Let $\mathcal{U}$ denote the set of affected nodes identified through bounded Breadth-First Search over the dependency tree. Instead of retraining the entire global model, *FDPTU* performs selective rollback by removing only the affected contributions:

$$w^{(-\mathcal{U})} = w - \sum_{c_i \in \mathcal{U}} \Delta w_{c_i}. \quad (4)$$

The same mechanism also supports client-initiated unlearning requests. When a client c_i requests removal, its contribution is subtracted from the trained model as $w^{(-c_i)} = w - \Delta w_{c_i}$, followed by secure re-aggregation of the remaining updates. A cryptographic audit record is generated to provide verifiable proof of deletion.

C. Optimization Objectives

The goal of *FDPTU* is to jointly optimize selective attack removal, privacy preservation, and unlearning efficiency. First, the selectivity of rollback is defined as

$$O_1 = 1 - \frac{|\mathcal{U}|}{N_{\text{tot}}}, \quad (5)$$

where N_{tot} is the total number of participating updates. A larger O_1 indicates more precise unlearning with fewer benign contributions removed.

Second, the unlearning latency is modeled as

$$O_2 = \sum_{c_i \in \mathcal{U}} \left(\frac{D_{\text{proc}, c_i}}{C_{\text{edge}}} + \frac{D_{\text{trans}, c_i}}{R_{\text{edge}}} \right), \quad (6)$$

where D_{proc, c_i} and D_{trans, c_i} denote the processing and transmission loads associated with client c_i , while C_{edge} and R_{edge} denote available edge computation and communication capacity. Third, privacy preservation is expressed through the differential privacy budget as

$$O_3 = e^{-\epsilon}. \quad (7)$$

The joint optimization objective is therefore written as

$$\max_{\mathcal{U}, \epsilon} \lambda_1 O_1 + \lambda_2 O_3 - \lambda_3 O_2, \quad (8)$$

subject to

$$C1 : |\mathcal{U}| \ll N_{\text{tot}}, \quad (9)$$

$$C2 : \epsilon \rightarrow 0, \quad (10)$$

$$C3 : O_2 \leq T_{\text{max}}, \quad (11)$$

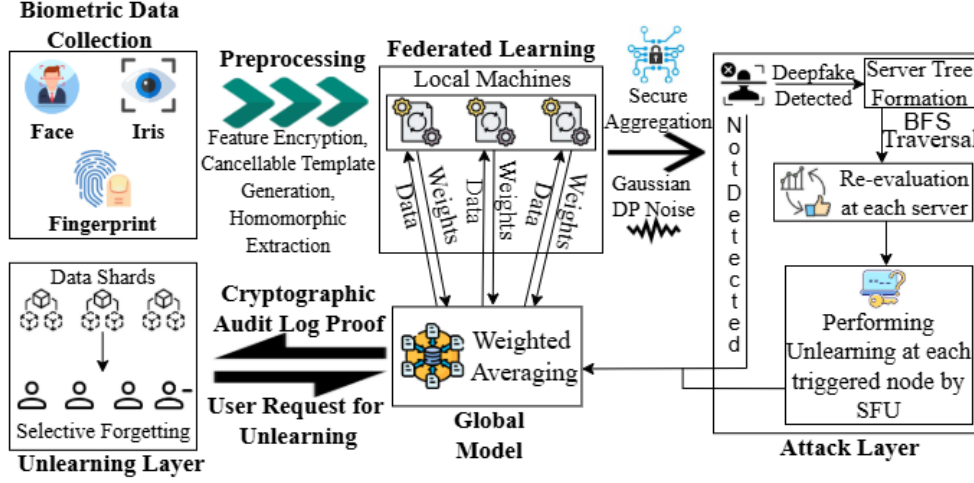


Fig. 1: System model of *FDPTU*. The framework supports privacy-preserving federated learning, attack-aware dependency tracking, and localized rollback of compromised or withdrawn biometric contributions.

where λ_1 , λ_2 , and λ_3 control the relative importance of selectivity, privacy, and latency. Constraint C1 enforces localized unlearning, C2 promotes stronger privacy protection, and C3 bounds the response time required for attack mitigation or client-requested data removal.

III. *FDPTU*: THE PROPOSED FRAMEWORK

This section presents the operational workflow of *FDPTU*. As shown in Fig. 2, the framework integrates secure local biometric processing, differentially private federated training, attack-aware monitoring, dependency-guided rollback, and cryptographic attestation. The proposed design differs from conventional FL and FU methods by linking adversarial detection with update-dependency tracing, so that only contaminated or withdrawn contributions are removed instead of retraining the full global model.

A. Secure Feature Extraction and DP Federated Training

Each client locally extracts modality-specific biometric features and converts them into protected templates before training. The templates are encrypted, and only protected model updates are transmitted to the server. Gaussian noise, calibrated to the privacy budget ϵ , is injected to limit update-level leakage, while secure aggregation prevents the exposure of individual client contributions. For N clients, R communication rounds, and E local epochs, the dominant training complexity is $O(RNE)$. Encryption, noise injection, and secure averaging add constant-factor overhead without affecting the asymptotic scalability. Algorithm 1 summarizes this training workflow.

Although Algorithm 1 protects biometric representations, privacy protection alone cannot prevent poisoned, synthetic, or morph-based updates from degrading the model. Therefore, *FDPTU* adds an attack-aware control stage before selective unlearning.

Algorithm 1: Secure Feature Extraction and Differentially Private Federated Training

Input: Client datasets $\{D_i\}_{i=1}^N$; privacy budget ϵ ; rounds R

Output: Trained global model G

foreach client $i \in \{1, \dots, N\}$ **do**

 Extract multimodal features F_i and generate protected templates T_i

 Encrypt templates $\hat{T}_i \leftarrow \mathcal{E}(T_i)$ using homomorphic encryption

Initialize global model $G^{(0)}$

for $r = 1$ **to** R **do**

foreach selected client i in parallel **do**

 Train locally on $\hat{T}_i$ and compute update $\Delta M_i^{(r)}$

 Perturb update with Gaussian DP noise calibrated to ϵ

 Send protected update to server

 Securely aggregate protected updates and obtain $G^{(r)}$

return $G = G^{(R)}$

B. Attack Detection and Threshold-Based Control

The attack monitoring layer evaluates submitted updates using complementary deepfake and morph detectors. For each evaluated sample, spatial and frequency-domain features are used to estimate deepfake confidence and morph likelihood. Their aggregated scores, S_d and S_m , are fused as

$$S = \alpha S_d + (1 - \alpha) S_m, \quad (12)$$

where $\alpha \in [0, 1]$ controls the contribution of the two detectors. Unlearning is triggered only when $S \geq \tau$, where τ is validation-calibrated to reduce false rollback under benign fluctuations. For an evaluation set of size $|E|$ and feature

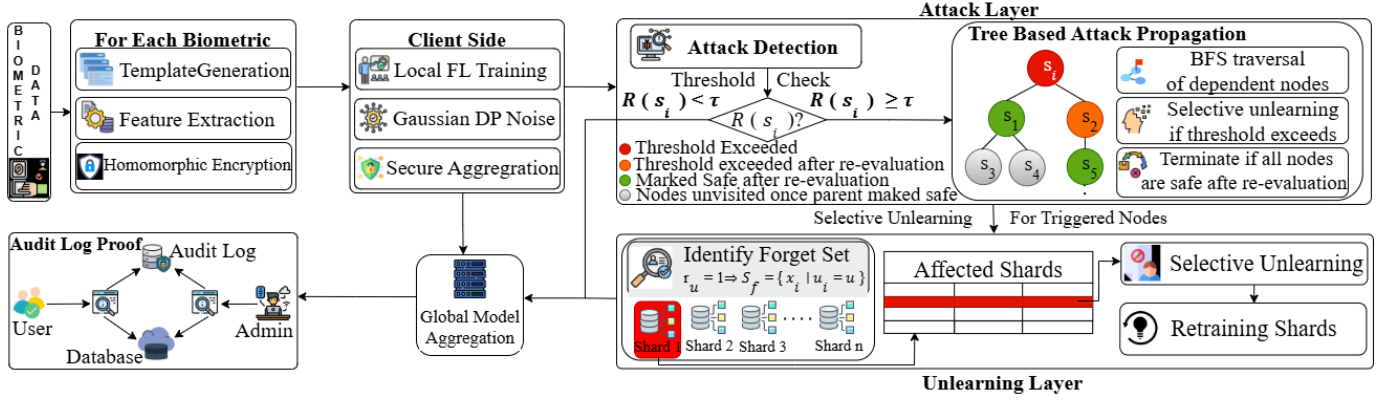


Fig. 2: Workflow of the proposed *FDPTU* framework for attack-aware, privacy-preserving, and selectively unlearnable federated biometric learning.

Algorithm 2: Deepfake and Morph Attack Detection with Threshold Evaluation

Input: Global model G ; evaluation dataset E ; threshold τ
Output: Attack flag; affected node set A
Initialize $S_d \leftarrow \emptyset$, $S_m \leftarrow \emptyset$
foreach sample $x \in E$ **do**
 Extract spatial and frequency-domain features F_x
 Compute detector scores $s_d(x) \leftarrow G_d(F_x)$ and $s_m(x) \leftarrow G_m(F_x)$
 Append $s_d(x)$ to S_d and $s_m(x)$ to S_m
Aggregate detector outputs to obtain S_d and S_m
Compute fused attack score $S \leftarrow \alpha S_d + (1 - \alpha) S_m$
if $S \geq \tau$ **then**
 $A \leftarrow$ affected client updates identified by contribution analysis
 return *TRUE*, A
else
 $A \leftarrow \emptyset$
 return *FALSE*, A

dimension d , the detection cost is $\mathcal{O}(|E|d)$. Algorithm 2 summarizes this control process.

Once an attack is detected, the affected updates must be removed without destabilizing benign learning. *FDPTU* therefore performs dependency-guided selective unlearning.

C. Dependency-Guided Selective Federated Unlearning

A directed dependency tree is constructed from historical aggregation records, where each node represents a client update and each edge captures its influence across communication rounds. Starting from the affected set A , bounded BFS identifies the final affected node set $\mathcal{U}$. Only the corresponding shard contributions are removed and recomputed, while benign updates are preserved. Let G and G^- denote the model states before and after unlearning. The rollback is constrained by $\|G - G^-\| \leq \eta$, where η bounds residual influence after

Algorithm 3: Threshold-Triggered Selective Federated Unlearning with Attested Proof

Input: Global model G ; affected node set A ; threshold τ
Output: Updated model G' ; attestation proof P
Construct dependency tree $\mathcal{T}$ from update history
Initialize BFS queue Q with affected nodes A
while $Q \neq \emptyset$ **do**
 Dequeue node n from Q
 if contribution score(n) $\geq \tau$ **then**
 Remove contribution of n from G and recompute dependent parameters
 Enqueue affected child nodes of n into Q
Re-aggregate benign updates to obtain G'
Generate attestation proof P from pre- and post-unlearning states
return G' and P

removal. A cryptographic proof is then generated by comparing pre- and post-unlearning states. Algorithm 3 presents the complete rollback procedure.

Let V and E_t denote the number of nodes and edges in the dependency tree. The traversal cost is $\mathcal{O}(V + E_t)$, while recomputation is restricted to affected shards rather than the entire federation. Thus, *FDPTU* satisfies the joint objective in Eq. (8): dependency-guided traversal supports selective rollback, Gaussian DP and secure aggregation preserve privacy, and shard-level recomputation bounds unlearning latency.

IV. FDPTU: PERFORMANCE EVALUATION

The proposed *FDPTU* framework is evaluated on two public datasets: (i) *Multi-modal Iris Fingerprint Biometric Data* [13], containing iris and fingerprint samples from 45 subjects, and (ii) a *Deepfake and Real Image Dataset* [14], containing authentic and manipulated facial images. To emulate a federated biometric environment, the facial dataset is partitioned into 45

TABLE I: Hyperparameter Configuration

Hyperparameter	Value
Number of Clients	45
Global Rounds	10
Client Fraction	0.2
Batch Size	32
Optimizer Learning Rate	10^{-3} (Adam)
Privacy Budget (ϵ)	5
DP Noise Multiplier	1.0
Gradient Clipping	1.0
Shards	5

virtual clients, each containing 150 real and 50 manipulated samples. The experiments analyze recognition utility, privacy resilience, residual influence after unlearning, and system overhead. The main experimental settings are summarized in Table I.

A. Unlearning Effectiveness and Utility Preservation

Fig. 3 shows the behavior of *FDPTU* during federated training and unlearning. In Fig. 3a, embedding similarity increases across training rounds, indicating stable biometric representation learning. After attack-triggered or client-requested unlearning, residual influence drops sharply, confirming that the dependency-guided rollback removes targeted contributions without disrupting benign updates. Fig. 3b further shows that *FDPTU* maintains a better balance between utility and attack suppression than FedAvg and DP-FedAvg. FedAvg preserves utility but lacks explicit removal capability, whereas DP-FedAvg improves privacy at the cost of recognition performance. In contrast, *FDPTU* combines stable utility with stronger residual influence reduction.

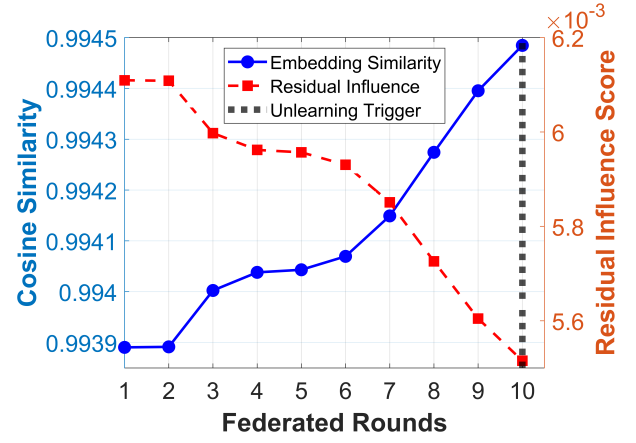
B. Attack Resilience and Privacy–Utility Trade-off

Membership inference attack (MIA) analysis is used to evaluate privacy leakage under authentic and manipulated samples [15]. As shown in Fig. 5a, *FDPTU* exhibits higher attack exposure before unlearning because client-specific contributions remain encoded in the trained model. After selective unlearning, the MIA AUC decreases substantially, showing that the removed updates no longer retain strong membership signals. The ROC curves in Fig. 5b move closer to the random baseline after unlearning, further confirming reduced inference success.

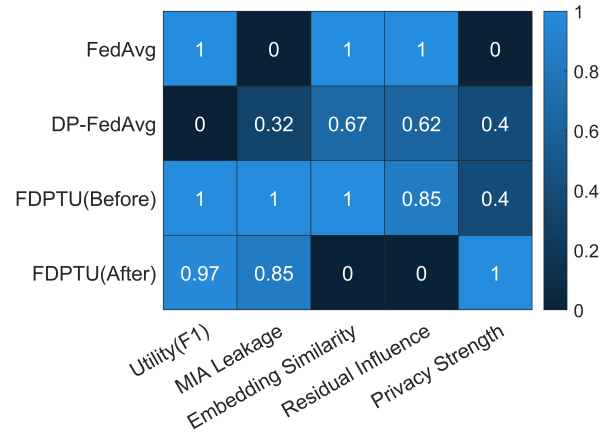
Fig. 4a reports the privacy–utility trade-off. As privacy protection increases, F1-score generally decreases due to DP-induced noise. However, *FDPTU* maintains higher F1-scores than competing methods at comparable privacy levels, indicating that selective rollback and multimodal fusion help preserve discriminative information. Fig. 4b and Fig. 4c show that *FDPTU* achieves balanced performance across utility, privacy, efficiency, and unlearning capability, while baseline methods are weaker in at least one dimension.

C. System Overhead

Table II compares normalized training, communication, and unlearning costs. *FDPTU* introduces moderate training and



(a) Unlearning effectiveness in FDPTU.



(b) Attack–utility dynamics.

Fig. 3: FDPTU performance and behavior analysis

TABLE II: Computational Cost Comparison

Method	Time/Round	Comm. Over-head	Unlearning Cost
FedAvg [16]	1.00	1.00	–
DP-FedAvg [17]	1.35	1.10	–
FedSF [9]	1.20	1.05	0.85
SISA [8]	1.15	1.00	0.70
FDPTU	1.42	1.18	0.45

communication overhead due to encryption, DP noise, and dependency tracking. However, it achieves the lowest unlearning cost because rollback is restricted to affected shards rather than the full model. This confirms that the additional overhead during training is offset by faster and more targeted removal during attack-triggered or client-requested unlearning.

V. CONCLUSION AND FUTURE WORKS

This paper presented *FDPTU*, an attack-aware, privacy-preserving FL and FU framework for multimodal biometric systems. By combining secure template protection, Gaussian DP, dependency-guided attack tracing, selective unlearning,

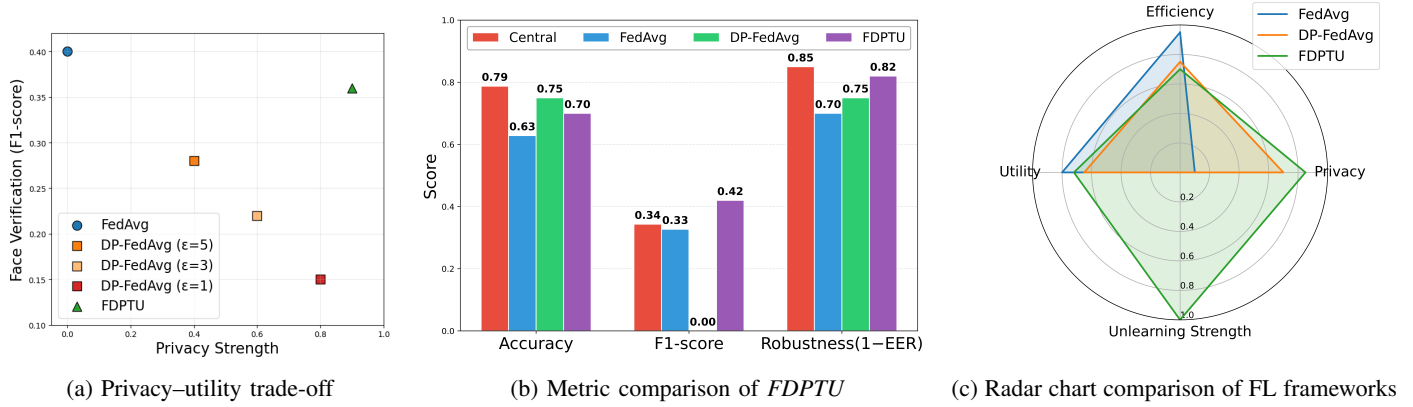


Fig. 4: Trade-off analysis and performance stability comparison across federated learning frameworks

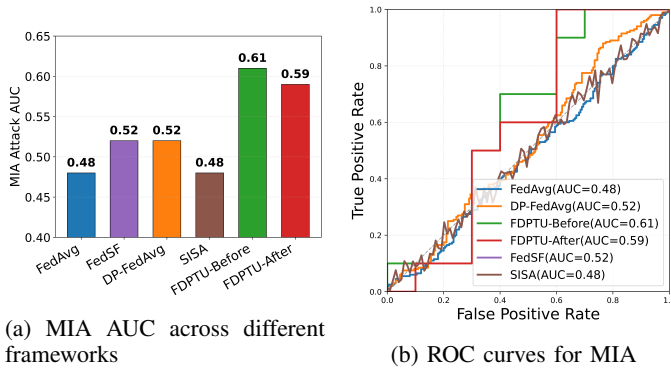


Fig. 5: Effectiveness of *FDPTU* in reducing attack success

and attested proof of removal, *FDPTU* enables compromised or withdrawn client contributions to be removed without global retraining. Experimental results show reduced residual influence and MIA success while preserving verification utility and multimodal representation stability. Future work will extend the framework to larger heterogeneous federations, adaptive and collusion-based attacks, and optimized edge-cloud deployment.

ACKNOWLEDGEMENTS

The manuscript is drafted and edited with the assistance of the ChatGPT-5.2 language model (OpenAI, <https://openai.com>) for grammatical refinement and language clarity. All technical content, analysis, and conclusions were verified and finalized by the authors.

REFERENCES

- [1] M. Balamurugan, "Biometric authentication: A double-edged sword for security," *International Journal of Science and Research (IJSR)*, vol. 13, no. 9, pp. 170–173, 2024.
- [2] M. Bhattacharya, S. Roy, S. Chattopadhyay, A. K. Das, and S. S. Jamal, "Aspa-mosn: An efficient user authentication scheme for phishing attack detection in mobile online social networks," *IEEE Systems Journal*, vol. 17, no. 1, pp. 234–245, 2023.
- [3] J. Guo, H. Mu, X. Liu, H. Ren, and C. Han, "Federated learning for biometric recognition: a survey," *Artificial Intelligence Review*, vol. 57, no. 8, p. 208, 2024.
- [4] A. K. Mishra, M. Wazid, D. P. Singh, A. K. Das, S. Roy, and S. Shetty, "Acks-ia: An access control and key agreement scheme for securing industry 4.0 applications," *IEEE Transactions on Network Science and Engineering*, vol. 11, no. 1, pp. 254–269, 2024.
- [5] Z. Zhang, Z. Tianqing, W. Ren, P. Xiong, and K.-K. R. Choo, "Preserving data privacy in federated learning through large gradient pruning," *Computers & Security*, vol. 125, p. 103039, 2023.
- [6] K. Wei, J. Li, C. Ma, M. Ding, W. Chen, J. Wu, M. Tao, and H. V. Poor, "Personalized federated learning with differential privacy and convergence guarantee," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 4488–4503, 2023.
- [7] N. Romandini, A. Mora, C. Mazzocca, R. Montanari, and P. Bellavista, "Federated unlearning: A survey on methods, design guidelines, and evaluation metrics," *IEEE Transactions on Neural Networks and Learning Systems*, 2024.
- [8] K. Koch and M. Soll, "No matter how you slice it: Machine unlearning with sisa comes at the expense of minority classes," in *2023 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, pp. 622–637, 2023.
- [9] G. Liu, X. Ma, Y. Yang, C. Wang, and J. Liu, "Federaser: Enabling efficient client-level data removal from federated learning models," in *2021 IEEE/ACM 29th International Symposium on Quality of Service (IWQOS)*, pp. 1–10, IEEE, 2021.
- [10] L. Wang, X. Zeng, J. Guo, K.-F. Wong, and G. Gottlob, "Selective forgetting: Advancing machine unlearning techniques and evaluation in language models," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, pp. 843–851, Apr. 2025.
- [11] J. Xu, Z. Wu, C. Wang, and X. Jia, "Machine unlearning: Solutions and challenges," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 8, no. 3, pp. 2150–2168, 2024.
- [12] S. Chatterjee, S. Jain, P. Bhattacharya, S. Roy, S. Banerjee, P. Rana, and S. Shetty, "Shard-unlearn: A sharded elastic sgd privacy preserving federated unlearning framework for 5g-assisted healthcare," in *Proceedings of the AAAI Fall Symposium Series*, vol. 7, pp. 481–487, 2025.
- [13] "Multimodal_iris_fingerprint_biometric_data," <https://www.kaggle.com/datasets/ninadmehendale/multimodal-iris-fingerprint-biometric-data>, 2024. [Accessed 2026-01-15].
- [14] "deepfake and real images," <https://www.kaggle.com/datasets/manjilkarki/deepfake-and-real-images>, 2022. [Accessed 2026-01-15].
- [15] L. Bai, H. Hu, Q. Ye, H. Li, L. Wang, and J. Xu, "Membership inference attacks and defenses in federated learning: A survey," *ACM Computing Surveys*, vol. 57, no. 4, pp. 1–35, 2024.
- [16] R. Li, H. Wang, Q. Lu, J. Yan, S. Ji, and Y. Ma, "Research on medical image classification based on improved fedavg algorithm," *Tsinghua Science and Technology*, vol. 30, no. 5, pp. 2243–2258, 2025.
- [17] S. S. Anwar and M. M. Hoque, "Optimizing privacy-accuracy trade-off in iot intrusion detection: An analysis of fedavg and fedprox with differential privacy," in *2024 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*, pp. 095–100, IEEE, 2024.