Research_Proposal_Samin-Yeasar_IARCO_Junior - Samin Yeasar.pdf

by Mr Adnan

Submission date: 14-Oct-2025 01:27PM (UTC+0300)
Submission ID: 2780774641
File name: Research_Proposal_Samin-Yeasar_IARCO_Junior_-_Samin_Yeasar.pdf (167.63K)
Word count: 1703
Character count: 10309

IARCO Research Proposal

Full Legal Name: Samin Yeasar

Institution: Birshreshtha Munshi Abdur Rouf Public College

Category: Junior

Class/Grade/Year: Grade/Class VII (2025)

Country: Bangladesh

Submission Date: September 22, 2025

Registered Email Address: sheditzofficial918@gmail.com

Research Topic: Privacy-Preserving Federated & Differentially Private Deep Learning for Multi Center Medical Imaging

Privacy-Preserving Federated & Differentially Private Deep Learning for Multi Center Medical Imaging

Background and Context

Recent advances in artificial intelligence (AI), deep learning (DL), and federated learning (FL) have significantly enhanced the potential of medical imaging. Diagnostic tasks such as X-ray or MRI classification now achieve expert-level accuracy when models are trained on large, diverse datasets. However, over 30% of healthcare organizations experienced data breaches last year, and regulations such as the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA) prohibit sharing raw medical data across hospital boundaries (Yahiaoui et al., 2024). Sharing data across hospitals is important for making AI models fair and reliable, but privacy laws and technical limits make this kind of collaboration very difficult today. The prevailing approach remains siloed: each hospital trains local models with limited generalizability or negotiates complex legal agreements for centralized research studies. A collaborative approach that enables hospitals to train AI systems without transferring patient images could address these challenges.

Federated learning addresses these issues by aggregating model weights centrally rather than sharing raw data. However, model updates can still leak patient information through inversion or reconstruction attacks (Zhou et al., 2025). Incorporating differential privacy (DP), typically by injecting noise into gradients, can limit the risk of individual re-identification and enhance legal and practical assurances (Liu et al., 2023). Nevertheless, excessive privacy noise can degrade model performance, particularly if implemented without careful calibration. Additionally, site-specific data distributions, such as variations in scanners or local demographics, further challenge model accuracy and necessitate algorithms that support both personalization and privacy protection.

This proposal aims to design and rigorously evaluate a comprehensive pipeline that integrates federated, differentially private deep learning with local personalization across diverse medical institutions. The central objective is to determine whether it is possible to improve diagnostic accuracy using global data while ensuring mathematical privacy guarantees for each patient and institution.

Review of Prior Research

The field has seen rapid developments across several domains, but integration remains an open challenge. Notably, Sheller et al. (2019) and Li et al. (2021) pioneered real-world FL deployments for brain tumor imaging, but relied mainly on vanilla federated averaging (FedAvg) without strong cryptographic or DP-based guarantees. The latest simulation studies demonstrate FL models approach or, at times, surpass centralized performance for segmentation. However, as Zhou et al. (2025) emphasize, even sharing weights can leak sensitive data, and single-server architectures create tampering and single-point-of-failure risks. Techniques like secure

aggregation (e.g., group verifiable secret sharing) have recently emerged to ensure servers cannot maliciously reconstruct patient data, while maintaining fault tolerance and practical overhead for real-world resource constraints (Zhou et al., 2025).

The use of DP in medical analytics comprehensively reviewed by Liu et al. (2023) shows strengths and trade-offs. DP-SGD and related DP techniques offer provable privacy, but real-world deployments often suffer from an accuracy-privacy trade-off. For example, Adnan et al. (2024) report that achieving ε-privacy budgets below 2 often reduces AUC by 1–3% even with extensive tuning. Adding homomorphic encryption has proven even more costly computationally, especially for high-dimensional image data. Recent frameworks, such as that evaluated by Elbachir Yahiaoui et al. (2024), combine 3D U-Net architectures with FL and DP for multi-institutional segmentation, reporting Dice scores above 86% while maintaining privacy through selective model sharing and noisy gradient updates.

Another direction in research focuses on personalization — letting each hospital adjust the shared model to better fit its own patients. Elhussein & Gürsoy (2023) explored this using local fine-tuning and adapter layers. Newer ideas group similar sites or patients together using secure methods, but these often require more communication and computing power.

In summary, although foundational elements have advanced, the literature identifies several key gaps. These include the need for systematic evaluation of the trade-off between privacy and clinical accuracy, scalable and practical aggregation protocols suitable for real-world multi-site deployment, and robust personalization strategies that adapt to data heterogeneity without compromising privacy.

Central Research Questions

This investigation is structured around three substantive, answerable questions:

1. Can a federated learning pipeline, combining differential privacy (DP) and secure, group-based aggregation, match or approach the diagnostic accuracy (e.g., AUC, Dice) of a centrally trained model for real-world multi-center medical imaging tasks, while holding privacy leakage below clinically/policy-relevant thresholds?
2. How do advanced personalization strategies (e.g., local fine-tuning, adapter layers, or cluster-based FL) interact with privacy noise, especially in scenarios of cross-site heterogeneity and data imbalance? What is the explicit trade-off among privacy guarantees, model utility, and per-site fairness?
3. What “utility/epsilon” frontier emerges for state-of-the-art differentially private FL across public multi-center imaging datasets? That is, how far can ε be reduced before clinical performance becomes unacceptable, and what practical DP settings (e.g., DP-SGD parameters) yield viable models for deployment?

These questions highlight a key tension: stronger privacy settings may protect patients better, but they could also reduce the accuracy of diagnoses, especially for smaller or underrepresented groups. This study aims to measure where that trade-off becomes unacceptable and suggest practical guidelines for real hospital deployments.

Methodology

Research Design

The study will use a combination of public, multi-center imaging datasets such as MIMIC-CXR (chest X-rays), BraTS/TCGA (multi-site brain MRI), and multi-institutional segmentation challenges. Baselines will include: (i) centrally trained models ("upper bound"), (ii) local-only models per site ("lower bound"), and (iii) varied FL pipelines with and without DP and personalization.

Federated Pipeline

The core framework includes:

- FL with Secure Aggregation: Utilizing group verifiable secret-sharing schemes (GVSA) as in Zhou et al. (2025) or SMPC protocols for cryptographic protection of model updates.
- Differential Privacy: Integrating DP-SGD at both client and/or server sides. Parameters (ε, δ) will be tuned, with experiments at ε=8, 2, 1 and matching baseline DP implementations (Liu et al., 2023).
- Personalization: Adapters/local-fine-tuning; for clustering-based personalization, SMPC will be employed to avoid leaking patient-level statistics (Elhussein & Gürsoy, 2023).

Experimental Protocol

- Model Architectures: Use of standard convolutional backbones (e.g., ResNet50, VGG, 3D U-Net) pre-trained on ImageNet or similar, finetuned for each imaging task.
- Evaluation: Metrics include accuracy/AUC for classification, Dice or IoU for segmentation, calibration error, communication cost (MB/transmission rounds), privacy parameters (ε, δ), and per-site fairness (disparity in performance).
- Statistical Analysis: Bootstrapped confidence intervals; ANOVA for multi-arm comparisons (e.g., FedAvg vs. FedProx vs. personalized FL).
- Compute & Resources: All experiments will be benchmarked on cloud GPUs, with 2D tasks requiring sub-16GB cards; segmentation evaluated at reduced resolution if needed for resource constraints.

Expected Challenges

- Utility-Privacy Trade-Off: Managing ε reduction without unacceptable AUC/Dice drop.
- Communication & Computation Overhead: Ensuring cryptographic protocols are feasible at scale (as GVSA aims for <10% overhead).
- Personalization–Privacy Tension: Fine-tuning risks overfitting to small/homogeneous local data, which may leak more about individuals if not properly controlled.
- Heterogeneity: Addressing data distribution shifts due to site-specific protocols, demographics, and scanners.

Timeline

- Month 1: Dataset preparation, initial baselines.
- Months 2–3: Pipeline implementation (FL, DP, secure aggregation).
- Month 4: Personalization module development.
- Month 5: Experimental runs and parameter sweeps.
- Month 6: Analysis and draft results.

Creativity, Innovation, Urgency

This project brings together three fast-moving areas: federated learning with stronger security, differential privacy methods adjusted for medical images, and techniques that let each hospital customize the shared model to its own data. Unlike much of the literature, which addresses these in isolation, this work will empirically evaluate the achievable privacy–utility boundary with practical, scalable open-source methods and public datasets, under clinically relevant settings. The approach is timely: new regulatory pressures (EU AI Act, U.S. executive orders) demand "privacy-by-design" and "explainable" AI. Moreover, global pandemics, like COVID-19, have accentuated the urgency and necessity for trustworthy AI tools that respect patient autonomy.

The goal is to show, with real experiments, that hospitals can build safer shared models without giving up performance or breaking privacy laws. If successful, this work could serve as a practical example for policymakers and engineers who want to use AI in healthcare responsibly.

References

[1] M. E. Yahiaoui, M. Derdour, R. Abdulghafor, S. Turaev, M. Gasmi, A. Bennour, A. Aborujilah, and M. A. Sarem, "Federated learning with privacy preserving for multi-institutional three-dimensional brain tumor segmentation," *Diagnostics*, vol. 14, no. 24, Art. no. 2891, Dec. 2024, doi: 10.3390/diagnostics14242891.

[2] A. Elhussein and G. Gürsoy, "Privacy-preserving patient clustering for personalized federated learning," in *Proc. Mach. Learn. Res.*, vol. 219, pp. 150–166, 2023.

[3] R. Haripriya, N. Khare, and M. Pandey, "Privacy-preserving federated learning for collaborative medical data mining in multi-institutional settings," *Sci. Rep.*, vol. 15, Art. no. 12482, Apr. 2025, doi: 10.1038/s41598-025-97565-4.

[4] W. Liu, Y. Zhang, H. Yang, and Q. Meng, "A survey on differential privacy for medical data analysis," *Ann. Data Sci.*, pp. 1–15, Jun. 2023, doi: 10.1007/s40745-023-00475-3.

[5] S. Zhou, L. Wang, L. Chen, Y. Wang, and K. Yuan, "Group verifiable secure aggregate federated learning based on secret sharing," *Sci. Rep.*, vol. 15, Art. no. 9712, Mar. 2025, doi: 10.1038/s41598-025-94478-0.

[6] V. Scheltjens, L. N. W. Momo, W. Verbeke, and B. De Moor, "Target informed client recruitment for efficient federated learning in healthcare," *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, Art. no. 380, Dec. 2024, doi: 10.1186/s12911-024-02798-4.

[7] A. Linardos, K. Kushibar, S. Walsh, J. M. M. Snoek, and J. Duchateau, "Federated learning for multi-center imaging diagnostics: a simulation study in cardiovascular disease," *Scientific Reports*, vol. 12, Art. no. 3551, Mar. 2022, doi: 10.1038/s41598-022-07186-4.

Research_Proposal_Samin-Yeasar_IARCO_Junior - Samin Yeasar.pdf

ORIGINALITY REPORT

11%	7%	8%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	arxiv.org Internet Source	1%
2	www.aal-persona.org Internet Source	1%
3	www.preprints.org Internet Source	1%
4	Corinne Allaart, Saba Amiri, Henri Bal, Adam Belloum, Leon Gommans, Aart van Halteren, Sander Klous. "Private and Secure Distributed Deep Learning: A Survey", ACM Computing Surveys, 2024 Publication	1%
5	Syed Hussain Ali Kazmi, Faizan Qamar, Rosilah Hassan, Kashif Nisar, Mohammed Azmi Al-Betar. "Security of Federated Learning in 6G Era: A Review on Conceptual Techniques and Software Platforms used for Research and Analysis", Computer Networks, 2024 Publication	1%
6	thesai.org Internet Source	1%
7	eprint.iacr.org Internet Source	1%
8	Submitted to Harokopio University Student Paper	1%

9	ctv.veeva.com Internet Source	1 %
10	www.ieee-jas.net Internet Source	1 %
11	Bora Bugra Sezer, Hasret Turkmen. "PPFLQB: A Privacy-Preserving Federated Learning enhanced Quantum-secure Blockchain layered framework", Internet of Things, 2025 Publication	1 %
12	pubmed.ncbi.nlm.nih.gov Internet Source	1 %
13	Noura A. Aldossary. "Artificial Intelligence in Optometry: Potential Benefits and Key Challenges: A Narrative Review", International Journal of Advanced Computer Science and Applications, 2025 Publication	<1 %
14	Yogeswar Reddy Thota, Tooraj Nikoubin. "EdgeAI with TinyML: Redefining Privacy and Security in Online Identity Management", IntechOpen, 2025 Publication	<1 %
15	Rufai Yusuf Zakari, Kassim Kalinaki, Zaharaddeen Karami Lawal, Najib Abdulrazak. "Federated learning for enhanced cybersecurity in modern digital healthcare systems", Institution of Engineering and Technology (IET), 2024 Publication	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	Off		