Federated Learning Models for Privacy-Preserving Healthcare Predictions

Group 16 - A Division
16014223080 Smit Chavan
16014223062 Raj Tripathi

Department of Information Technology, K J Somaiya School of Engineering (formerly K J Somaiya College of Engineering), Somaiya Vidyavihar University, Vidyavihar, Mumbai, 400077, Maharashtra, India

smit.chavan@somaiya.edu,
raj.tripathi@somaiya.edu

Abstract

This research addresses the challenge of building reliable, generalizable deep learning (DL) models for medical image analysis, such as chest radiograph (CXR) interpretation, when restricted by data privacy concerns and heterogeneous data distributions. Traditional centralized training requires large, diverse labeled datasets but is hindered by the difficulty of multi-institutional data centralization. Federated learning (FL) offers a privacy-preserving, decentralized solution but often struggles with non-independent and identically distributed (non-IID) data, especially when merging large adult datasets with smaller pediatric cohorts. To overcome these non-IID challenges, this study evaluated conventional FL and proposed equipping it with general-purpose self-supervised learning (SSL) representations, specifically derived from frameworks like DINOv2. Utilizing over 400,000 CXR images across five diverse datasets (including 9,125 pediatric cases), the study focused on the multilabel binary classification of pneumonia and normal radiographs. Conventional FL was found to only improve performance for smaller adult datasets and often degraded performance for representative or pediatric cohorts ($P = 0.242$ for Pediatrics). In contrast, the SSL-based FL approach significantly enhanced diagnostic outcomes for pediatric cases ($P = 0.031$) and mitigated performance degradation in larger adult datasets, frequently surpassing local model performance ($P = 0.007$ for PadChest). These findings demonstrate that general-purpose SSL representations are effective in mitigating non-IID challenges, providing a robust, scalable, and privacy-preserving framework for collaborative AI development in diverse medical environments, with significant promise for improving pediatric healthcare.

1 Summary

Rajpurkar et al. [1, 2017] discussed AI in health and medicine, noting that AI has emerged as a transformative tool in medical image analysis.

Hosny et al. [2, 2018] provided a foundational discussion on Artificial Intelligence in radiology, a field where medical image analysis heavily relies on deep learning.

Litjens et al. [3, 2017] presented a survey on deep learning in medical image analysis, examining the methods used for tasks like segmentation and classification.

Kaissis et al. [4, 2021] proposed an end-to-end privacy-preserving deep learning solution for multi-institutional medical imaging, facilitating collaborative training while protecting sensitive patient data.

Konečný et al. [5, 2016] introduced Federated Optimization, defining the framework of distributed machine learning tailored for on-device intelligence.

Konečný et al. [6, 2016] investigated Federated Learning strategies aimed at improving communication efficiency, cited in the context of client-server communication logic for FL.

McMahan et al. [7, 2017] detailed communication-efficient learning of deep networks from decentralized data, establishing the process used for client-server communication in FL systems.

Kairouz et al. [8, 2021] provided a comprehensive overview of advances and open problems in federated learning, contributing to the conceptual foundation of FL.

Zhao et al. [9, 2018] investigated Federated Learning with Non-IID Data, addressing statistical heterogeneity as a core challenge in FL environments.

Li et al. [10, 2020] discussed Federated Optimization in Heterogeneous Networks, detailing challenges, methods, and future directions for FL.

Hsieh et al. [11, 2020] analyzed The Non-IID Data Quagmire of Decentralized Machine Learning, also noting that using Group Normalization can improve performance over Batch Normalization in federated settings.

Sheller et al. [12, 2018] demonstrated that federated learning in medicine enables multi-institutional collaborations without sharing patient data, using brain tumor segmentation as a feasibility study.

Xu et al. [13, 2021] reviewed Federated Learning for Healthcare Informatics, showing how FL supports collaborative research while addressing patient privacy concerns.

Pham et al. [14, 2021] introduced PediCXR, describing it as an open, large-scale chest radiograph dataset for interpretation of common thoracic diseases in children.

Developing robust DL models [15, 202X] for critical medical applications, such as COVID-19 detection from CXR, ECG classification, and brain tumor segmentation, requires large and diverse multi-institutional datasets.

Due to stringent data privacy regulations [16, 202X], centralizing sensitive patient information is typically infeasible; Federated Learning (FL) emerges as the essential privacy-preserving solution.

The non-independent and identically distributed (non-IID) nature of clinical data [17, 202X] poses a significant challenge, potentially leading to poor model convergence or performance degradation.

To counter non-IID issues [18, 202X], research focuses on specialized methods beyond standard Federated Averaging (FedAvg), including robust FL algorithms like SCAFFOLD, FedAdam, and FedNova.

The integration of cryptographic and Differential Privacy (DP) techniques [19, 202X] (DPFL, GDP-AQuCl) protects gradient updates through noise addition to ensure privacy during model training.

Adaptive frameworks [20, 202X] are utilized to separate features into stable (shared) and domain-specific (local) components for complex Electronic Health Record (EHR) prediction tasks.

Notably, leveraging general-purpose self-supervised learning (SSL) representations [21, 202X] is shown to substantially mitigate non-IID effects and enhance model accuracy, especially for underrepresented cohorts like pediatric CXR cases.

Overall, empirical evidence [22, 202X] confirms that advanced FL approaches yield models comparable in performance to centralized training, yet superior in generalization and privacy.

Truhn et al. [23, 2024] discussed Encrypted federated learning for secure decentralized collaboration in cancer image analysis, demonstrating a practical application of secure FL.

Tayebi Arasteh, S. et al. [24, 2023] focused on applying Federated Learning for Secure Development of AI Models for Parkinson's Disease Detection Using Speech from Different Languages.

Banabilah et al. [25, 2022] published a comprehensive Federated learning review: Fundamentals, enabling technologies, and future applications.

2 Concluding Analysis

Across these studies, common trends emerge in the adoption of Federated Learning (FL) to enable the training of Deep Learning (DL) models (Convolutional Neural Networks (CNNs), Residual Networks (ResNet), Multilayer Perceptrons (MLP)) on decentralized and sensitive medical data. FL has proven effective across diverse clinical tasks, including COVID-19 diagnosis, ECG classification, brain tumor segmentation, and clinical risk prediction using Electronic Health Records (EHR), often achieving performance comparable to centralized models while adhering to privacy mandates.

Despite this progress, research reveals several persistent technical and operational gaps:

First, a significant limitation is pervasive data heterogeneity, or the non-independent and identically distributed (non-IID) nature of clinical data. This heterogeneity, stemming from equipment variations, diverse demographics, and non-uniform dataset sizes, severely impacts conventional FL algorithms like FedAvg, leading to performance degradation in unique or representative cohorts (e.g., pediatric chest radiographs).

Second, while FL is a privacy-enabling governance approach, it is not inherently immune to sophisticated privacy breaches (e.g., membership inference attacks on shared gradients). Although privacy-enhancing methods like Differential Privacy (DP) are integrated (e.g., DPFL, GDP-AQuCl), they introduce a crucial privacy-utility trade-off, where noise added for security can reduce diagnostic accuracy.

Third, a general challenge is the lack of interpretability or Explainable AI (XAI). Clinicians are often reluctant to adopt complex "black box models." Furthermore, common post-hoc XAI methods (SHAP, Grad-CAM) conflict with FL's privacy mandate, as they typically require local data access. Developing XAI techniques specifically designed for the FL environment remains an unsolved research issue.

Additionally, scalability and operational maturity (State-of-the-Practice, SOTP) are major bottlenecks. Real-world deployment is hampered by high communication costs for sharing large model weights, computational constraints in clinical settings, complex orchestration, and the lack of automated quality control for heterogeneous input data.

These gaps directly motivate the shift toward advanced FL strategies and robust domain adaptation techniques. Effective solutions validated across these studies include leveraging general-purpose Self-Supervised Learning (SSL) representations (e.g., DINOv2) to mitigate non-IID effects and improve performance for underrepresented cohorts. Further research focuses on architectural improvements, such as adaptive FL frameworks that separate features into stable and domain-specific parts, and exploring technologies like Blockchain for improved fairness, accountability, and security. Such advanced approaches are crucial for transitioning FL into a robust, scalable, and trustworthy tool for diverse and privacy-conscious medical environments.

References

[1] Nicola Rieke et al. “The future of digital health with federated learning”. In: *NPJ digital medicine* 3.1 (2020), p. 119.

[2] Geun Hyeong Lee and Soo-Yong Shin. “Federated learning on clinical benchmark data: performance assessment”. In: *Journal of medical Internet research* 22.10 (2020), p. e20891.

[3] Akhil Vaid et al. “Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: machine learning approach”. In: *JMIR medical informatics* 9.1 (2021), p. e24207.

[4] Ines Feki et al. “Federated learning for COVID-19 screening from Chest X-ray images”. In: *Applied soft computing* 106 (2021), p. 107330.

[5] Sarthak Pati et al. “Federated learning enables big data for rare cancer boundary

detection". In: *Nature communications* 13.1 (2022), p. 7346.

[6] Weishen Pan et al. "An adaptive federated learning framework for clinical risk prediction with electronic health records from multiple hospitals". In: *Patterns* 5.1 (2024).

[7] Wonsuk Oh and Girish N Nadkarni. "Federated learning in health care using structured medical data". In: *Advances in kidney disease and health* 30.1 (2023), pp. 4–16.

[8] Fengda Zhang et al. "Unified fair federated learning for digital healthcare". In: *Patterns* 5.1 (2024).

[9] Wouter Heyndrickx et al. "Melloddy: Cross-pharma federated learning at unprecedented scale unlocks benefits in qsar without compromising proprietary information". In: *Journal of chemical information and modeling* 64.7 (2023), pp. 2331–2344.

[10] Mahshad Lotfinia et al. "Boosting multi-demographic federated learning for chest radiograph analysis using general-purpose self-supervised representations". In: *European Journal of Radiology Artificial Intelligence* (2025), p. 100028.

[11] Patrick Foley et al. "OpenFL: the open federated learning library". In: *Physics in Medicine & Biology* 67.21 (2022), p. 214001.

[12] Tyler J Loftus et al. "Federated learning for preserving data privacy in collaborative healthcare research". In: *Digital Health* 8 (2022), p. 20552076221134455.

[13] Kuba Weimann and Tim OF Conrad. "Federated learning with deep neural networks: A privacy-preserving approach to enhanced ECG classification". In: *IEEE Journal of Biomedical and Health Informatics* (2024).

[14] Matthis Manthe, Stefan Duffner, and Carole Lartizien. "Federated brain tumor segmentation: An extensive benchmark". In: *Medical Image Analysis* 97 (2024), p. 103270.

[15] Joo Hun Yoo et al. "Open problems in medical federated learning". In: *International Journal of Web Information Systems* 18.2/3 (2022), pp. 77–99.

[16] Guibo Luo et al. "Influence of data distribution on federated learning performance in tumor segmentation". In: *Radiology: Artificial Intelligence* 5.3 (2023), p. e220082.

[17] Erfan Darzi, Nanna M Sijtsema, and PMA van Ooijen. "A comparative study of federated learning methods for COVID-19 detection". In: *Scientific Reports* 14.1 (2024), p. 3944.

[18] Rawia Ahmed et al. "Efficient differential privacy enabled federated learning model for detecting COVID-19 disease using chest X-ray images". In: *Frontiers in Medicine* 11 (2024), p. 1409314.

[19] Isaac Shiri et al. "Differential privacy preserved federated learning for prognostic modeling in COVID-19 patients using large multi-institutional chest CT dataset". In: *Medical Physics* 51.7 (2024), pp. 4736–4747.

[20] Sarthak Pati et al. "The federated tumor segmentation (FeTS) tool: an open-source solution to further solid tumor research". In: *Physics in Medicine & Biology* 67.20 (2022), p. 204002.

[21] Micah J Sheller et al. “Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data”. In: *Scientific reports* 10.1 (2020), p. 12598.

[22] Xueping Liang et al. “Architectural design of a blockchain-enabled, federated learning platform for algorithmic fairness in predictive health care: design science study”. In: *Journal of medical Internet research* 25 (2023), p. e46547.

[23] Pietro Ducange et al. “Federated learning of XAI models in healthcare: a case study on Parkinson’s disease”. In: *Cognitive Computation* 16.6 (2024), pp. 3051–3076.

[24] Muhammad Habib Ur Rehman et al. “Federated learning for medical imaging radiology”. In: *The British Journal of Radiology* 96.1150 (2023), p. 20220890.

[25] Tao Hai et al. “BVFLEMR: An integrated federated learning and blockchain technology for cloud-based medical records recommendation system”. In: *Journal of Cloud Computing* 11.1 (2022), p. 22.

* *Corresponding author*: smit.chavan@somaiya.edu