Federated Learning for Privacy-Preserving Healthcare Data

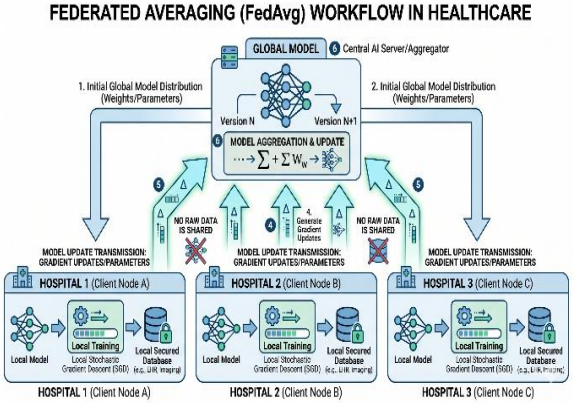
Shrinidhi Anchan[1], Prajwal K S[2], Shishir R Kulal[3], Swasthik Rai[4], Saket L Kumbla[5]
[1,3,4,5]*Student, Department of Computer Science and Design, Srinivas Institute of Technology, Mangalore, Karnataka, India*
[2]*Assistant Professor, Department of Computer Science and Design, Srinivas Institute of Technology, Mangalore, Karnataka, India*
doi.org/10.64643/IJIRTV12I12-206689-459

***Abstract*—The healthcare industry deals with large amounts of confidential patient information which can enhance the performance of machine learning algorithms to diagnose diseases and provide better treatments. Due to privacy and ethical issues, the sharing of such data is not possible. In recent times, Federated Learning (FL) has emerged as an excellent framework for training models in a cooperative manner without using the actual data. This work provides a review of FL in healthcare along with its problems and future research opportunities.**

***Index Terms*—Communication Overhead, Data Heterogeneity, Deep Leakage from Gradients, Federated Learning, Privacy-Preserving Healthcare.**

I. INTRODUCTION

Healthcare digitization has caused to the generation of huge amounts of medical data, even the electronic health records (EHRs) and imaging and genetic datasets. Models created using machine learning on such data have proved their diagnostic potential. Nevertheless, there exists a critical contradiction, which is legal and ethical restriction.



Legislation such as HIPAA in the USA and GDPR in Europe stipulates how patient data should be stored, transferred, and used. Consequently, hospitals and organizations conduct research in isolation due to the sensitive nature of information they possess. However, Federated Learning (FL) provides a way out of this dilemma. Proposed by McMahan et al. in 2017, FL allows several parties to jointly train a single model without moving the data to a centralized server.

II. LITARETURE REVIEW

A. The Move Towards Distributed Healthcare Systems

The move towards federated learning (FL) in clinical settings is mainly motivated by the inherent conflict between the requirement for large datasets and patient privacy protection legislation. Early surveys, as well as initial frameworks developed on the subject, indicated FL to be an extremely promising architectural approach for healthcare informatics, as it enables medical organizations to train models together without transferring any sensitive information contained within the dataset. Modern systematic reviews further support the viability of the federated learning architecture as the cornerstone for future healthcare systems [1][2][3][4].

B. Application Areas: Medical Images and Cancer Diagnostics

The most studied application areas of federated learning are image analysis, especially for medical images. This approach becomes invaluable for smaller hospitals that work with locally labeled, limited sets of data since they can use an internationally-trained

model to improve upon their data. For example, multi-organization big data has been used in FL networks for very complicated oncological diagnostics such as cancer margin detection [5][6][7][8][9].

C. EHR Data Processing and Pandemic Response

In addition to its use in imaging applications, FL has shown promise in data processing in electronic health records (EHRs). Because of heterogeneity of EHRs in each respective hospital network, state-of-the-art approaches, such as augmented temporal graph attention networks, have emerged to facilitate data processing in EHRs. Indeed, the practicality of these predictive algorithms has been tested in times of crisis; multiple organizations worldwide relied on FL technology to accurately predict patient outcomes in hospitals and calculate mortality rates based on varying EHRs [10][11].

D. Vulnerabilities and Alternative Architectures

FL has seen widespread application due to its unique potential in the healthcare sector; however, this technology itself is by no means an entirely flawless solution. In particular, there are several algorithmic and infrastructural issues associated with the technology, such as server–client communication barriers and the non-IID nature of global health data. Far more importantly, the idea that FL provides full protection from leaks has been refuted in recent research, which found malicious parties able to invert gradient updates to reveal private data [12][13][14].

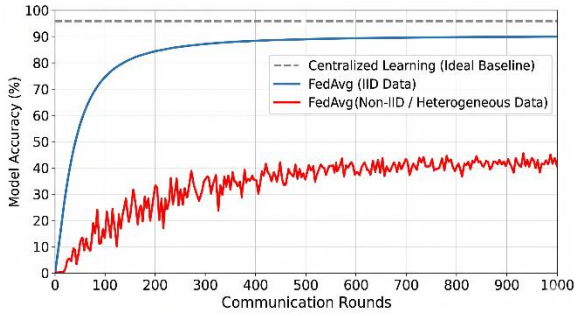
E. The Future Direction

The current direction of federated learning studies involves many concentrations on shifting beyond proof of concept and artificial data. With advancements in algorithm development, the central goal of both researchers and developers is the practical challenges involved in making the system viable in real world clinical settings [15].

III. BACKGROUND AND RELATED WORK

A. Traditional Centralized Learning vs. Federated Learning

The system requires data to be gathered at a single central location for traditional machine learning processes. The technology shows significant benefits across all domains except for areas which deal with confidential information like health care. The technology of Federated Learning solves this issue by enabling model distribution to multiple data sources.



In federated learning the system first establishes a global model which the server distributes to all connected devices. The devices build the model with their own data and transmit the gradient information back to the server to improve the model until it reaches the target accuracy level.

B. Core Algorithms & Effectiveness

The widely used aggregation technique is Federated Averaging (FedAvg), which was introduced by McMahan et al. This technique makes it possible for the global model to be updated based on a federated average of the local model based on the weight of the dataset of each client. In recent medical imaging research, it has been observed that federated learning works very well. It has been demonstrated that federated models trained in 10 different institutions could achieve 99% of centralized performance levels.

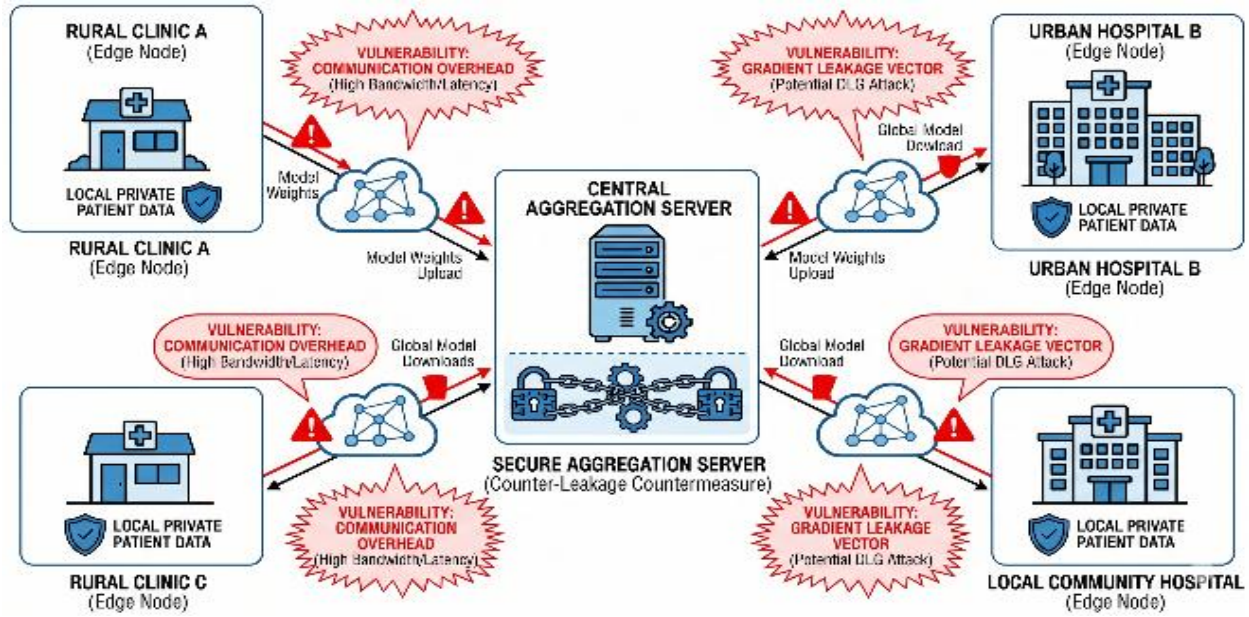
IV. APPLICATIONS IN HEALTHCARE

A. The Medical Imaging

FL system has achieved high success rates in both radiology and pathology departments. The research demonstrates that the FL-based model can perform complex brain tumor segmentation tasks at the same level as pooled data models which require complete patient image access.

B. The confidentiality of Electronic Health Record

Data requires healthcare organizations to create their own unique data standards. Through federated learning investigators can build predictive models which evaluate hospital readmission risks and patient mortality risks together with disease progression pathways while keeping patient information secure.



C. Pandemic Response and Global Health
FL established international collaboration during the COVID-19 pandemic. The system enabled multiple organizations to develop a worldwide diagnostic system through data sharing while protecting their national data sovereignty rights.

V. CHALLENGES AND VULNERABILITIES

A. Healthcare data exhibits substantial diversity in its real-world applications because patients at rural clinics present different demographic and epidemiological characteristics than patients at urban hospitals. Because of the non-iid property of the data, there is a disparity in the local models, which negatively impacts the global model's performance. FedAvg algorithm performs poorly under conditions of high heterogeneity.

B. Communication Burden The training process in FLL involves communication of model updates across possibly several hundred institutions. Using deep learning algorithms in connection with images from medicine can mean that there are GBs of updates for the model, hence resulting in very costly communication in terms of latency and bandwidth in hospitals.

C. The privacy protection system of FLL does not provide secure privacy protection. The "Deep Leakage from Gradients" (DLG) attack enables attackers to extract confidential information from secured data through the use of gradient updates. The malicious server uses gradient inversion techniques that depend on gradient updates to produce complete visual images and their matched text tokens.

The methods for stopping gradient leakage provide users with three options which include gradient pruning and differential privacy and secure aggregation yet users must select between maintaining accuracy or protecting their privacy.

D. Different educational institutions possess distinct IT systems which create training challenges because their systems experience delays when operating with slow nodes that act as stragglers. The central machine learning methods have numerous benchmark datasets but federated learning in healthcare applications lacks both benchmark datasets and evaluation systems.

VI. DIRECTIONS FOR FUTURE RESEARCH

To overcome these challenges, future research should work on these areas:

1. HLS-PFL:
Designing local models that will be able to take advantage of collaborative learning but do not require the enforcement of a common global model.

2. PF-HFL:
Mathematically-proving privacy bounds that could effectively defend against gradient inversion attacks.

3. HL-benchmark:
Designing universal standards of evaluation in order to promote reproducibility, specifically focusing on the problem of the lack of HL benchmarks.

4. Regulation for Healthcare ML:
Multidisciplinary studies involving machine learning and health care regulation with the focus on gradient sharing within the constraints of HIPPA and GDPR.

VII. CONCLUSION

Federated Learning signifies an enormous paradigm shift for the implementation of collaborative and privacy-conscious machine learning in healthcare. It offers a highly elegant solution to the inherent paradox of having the capability to utilize large volumes of data while abiding by the laws of patient privacy. Still, crucial issues relating to data heterogeneity, exorbitant communication overheads, and severe threats such as gradient leakage persist as unresolved problems that must be tackled through interdisciplinary collaboration.

Future Scope:
Although the underlying surveillance engine is highly capable, future extensions can be done in the following manner:

1) Database Migration:
The flat JSON file-based storage should be migrated to a more scalable relational database, such as PostgreSQL or MySQL, for managing large campuses.

2) AI Extensions:
Implementation of the delayed fall detection algorithms, even integration of multimodal language models (e.g., Google Gemini) for generating natural summaries for security.

3) Secure Deployment:
Secure deployment on the cloud using HTTPS protocol (API authentication).

REFERENCES

[1] N. Rieke, J. Hancox, W. Li, F. Milletari, H. R. Roth, S. Albarqouni, *et al*., "The Future of Digital Health with Federated Learning," *NPJ Digit. Med.*, vol. 3, no. 1, pp. 1–7, 2020.

[2] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, "Federated Learning for Healthcare Informatics," J. Healthcare Inform. Res., vol. 5, no. 1, pp. 1–19, 2021.

[3] D. C. Nguyen, Q. V. Pham, P. N. Pathirana, M. Ding, A. Seneviratne, Z. Lin, et al., "Federated Learning for Smart Healthcare: A Survey," ACM Comput. Surv., vol. 55, no. 3, 2022.

[4] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, et al., "Advances and Open Problems in Federated Learning," Found. Trends Mach. Learn., vol. 14, nos. 1–2, pp. 1–210, 2021.

[5] M. J. Sheller, B. Edwards, G. A. Reina, J. Martin, S. Pati, A. Kotrotsou, et al., "Federated Learning in Medicine: Facilitating Multi-Institutional Collaborations Without Sharing Patient Data," Sci. Rep., vol. 10, no. 1, Art. no. 12598, 2020.

[6] S. Pati, U. Baid, B. Edwards, M. Sheller, S. H. Wang, G. A. Reina, et al., "Federated Learning Enables Big Data for Rare Cancer Boundary Detection," Nat. Commun., vol. 13, no. 1, Art. no. 7346, 2022.

[7] D. Ng, X. Lan, J. Yao, W. P. Chan, and M. Feng, "Federated Learning: A Collaborative Effort to Achieve Better Medical Imaging Models for Individual Sites That Have Small Labelled Datasets," Quant. Imaging Med. Surg., vol. 11, no. 2, p. 852, 2021.

[8] I. Dayan, H. R. Roth, A. Zhong, A. Harouni, A. Gentili, A. Z. Abidin, et al., "Federated Learning for Predicting Clinical Outcomes in Patients with COVID-19," Nat. Med., vol. 27, no. 10, pp. 1735–1743, 2021.

[9] A. Vaid, S. K. Jaladanki, J. Xu, T. Kurc, et al., "Federated Learning of Electronic Health Records to Improve Mortality Prediction in Hospitalized Patients with COVID-19," JMIR Med. Inform., vol. 9, no. 1, Art. no. e24207, 2021.

[10] R. S. Antunes, C. André da Costa, A. Küderle, I. A. Yari, and B. Eskofier, "Federated Learning for Healthcare: Systematic Review and Architecture

Proposal," ACM Trans. Comput. Healthcare, vol. 3, no. 4, pp. 1–36, 2022.

[11] S. Warnat-Herresthal, H. Schultze, K. L. Shastry, S. Manamohan, S. Mukherjee, V. Garg, et al., "Swarm Learning for Decentralized and Confidential Clinical Machine Learning," Nature, vol. 594, no. 7862, pp. 265–270, 2021.

[12] J. Geiping, H. Bauermeister, H. Dröge, and M. Moeller, "Inverting Gradients—How Easy Is It to Break Privacy in Federated Learning?" in Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 33, pp. 16937–16947, 2020.

[13] M. I. Sharif, M. Mehmood, M. P. Uddin, K. Siddique, Z. Akhtar, and S. Waheed, "Federated Learning for Analysis of Medical Images: A Survey," J. Comput. Sci., vol. 20, no. 12, pp. 1610–1621, 2024.

[14] S. Molaei, A. Thakur, G. Niknam, A. Soltan, H. Zare, and D. A. Clifton, "Federated Learning for Heterogeneous Electronic Health Records Utilising Augmented Temporal Graph Attention Networks," in Proc. Mach. Learn. Res., vol. 238, 2024.

[15] F. Gomez et al., "Federated Learning in Healthcare: From Research to Real-World Deployment," Annu. Rev. Biomed. Eng., vol. 28, 2026.