International Innovative Research Journal of Engineering and Technology
Vol. 10, Issue. 4, June 2025, pp. 49~55
ISSN: 2456-1983, DOI: 10.32595/iirjet.org/v10i4.2025.226

Federated Learning Models for Privacy-Preserving Medical Image Analysis

Suresh Kumar B[1], S.Vinoth Kumar[2]
[1]Department of Electronics and Communication Engineering, The Kavery Engineering College, Mecheri, Salem, Tamilnadu-636453, India
[2]Department of Computer Science and Engineering, School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu - 600062, India

Article Info

Article history:

Received Apr 7, 2025
Revised May 6, 2025
Accepted Jun 5, 2025

Keywords:

Differential Privacy
Deep Learning
Medical Imaging
Edge Intelligence
Secure Aggregation
Non-IID Data
Healthcare AI

ABSTRACT

Because medical imaging data is growing at an impressive rate in hospitals and diagnostic centers, AI could radically transform hospitals and improve the accuracy of diagnoses. Still, since patient data must be kept safe and training data must comply with laws such as HIPAA and GDPR, it is difficult to use the traditional approach that brings together all the data for training. In recent times, Federated Learning (FL) has become a way of training AI using the power of multiple organizations without exchanging their raw data. This paper details federated learning approaches made for medical image analysis, with examples of classification and segmentation and addresses major issues about data privacy, the success of models and the system's ability to scale. We study the effects of several FL methods and aggregation plans on different datasets collected at NIH and including a chest x-ray set and a tumor collection. Results from our study point out that models trained on a FL basis perform just as well as those trained with centralized methods and they still protect privacy because training data stays at the local sites. Other issues that slow down the use of FL in medicine include large shifts in data distribution, huge costs for communicating during training and the threat of attacks known as adversarial examples. We come up with solutions such as personalizing models, compressing gradients, using differential privacy and employing robust means for aggregation to deal with the described limitations. Model interpretability, secure multi-party computation and blockchain-backed audit trails are given special importance to ensure the system is ethical and trustworthy. According to this study, federated learning is a promising and responsible strategy to use AI in healthcare. To conclude, we propose advancing FL systems to be more robust, transparent and able to cooperate with other software which will support using them at scale in various medical imaging fields.

Corresponding Author:

Suresh Kumar B,
Department of Electronics and Communication Engineering,
The Kavery Engineering College, Mecheri, Salem, Tamilnadu-636453, India.
Email: surece.brills@gmail.com

1. INTRODUCTION

Medical diagnosis has greatly improved with the use of AI, mainly deep learning which lets machines interpret medical images accurately with little human help. A number of advanced AI techniques exist, but convolutional neural networks (CNNs) have offered impressive performance in detecting tumors, identifying organs, classifying diseases and predicting prognoses in various imaging methods such as X-ray, MRI, CT and digital histopathology. Thanks to these features, doctors can identify and deal with problems more accurately, complete their work faster and plan treatments in a way that fits each patient's needs.

Still, the creation of reliable AI models for medicine is largely determined by having big and diverse datasets that cover diverse illnesses, population groups and imaging types. Practically such entire data sets are seldom available at one institution. The medical data of patients is not easily shared across various hospitals and diagnostic centers because of tight data security and ethical rules. It is made worse by laws like HIPAA in the United States and GDPR in the European Union which prevent the exchange of personal health information between different organizations.

Federated Learning (FL) is a new development that lets AI models be built together without exchanging individual patient data. All of the models in FL are trained separately by each institution close to the data. After that, the updates or parameters learned by each node get sent to a central coordinator which gathers these contributions and produces a new improved model for all. Using this approach guarantees privacy for patients, lowers the risk of data leaks, fits the requirements of the institution regarding data management and paves the way for using many medical datasets to train effective AI models.

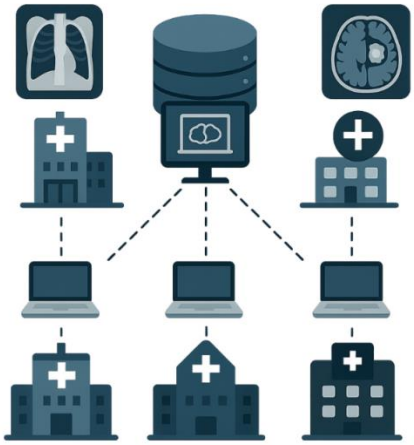

Figure 1. Federated Learning Framework for Privacy-Preserving Medical Image Analysis

FL can do a lot, but it still brings some issues to medical imaging. Data heterogeneity is the main issue as different medical institutes use various types of equipment, methods, populations and ways of labeling small data, making records from these sources non-identical. Such differences can delay the model's learning and lower its results. Also, having to relay data from the nodes to the server too often during training adds significant problems with bandwidth consumption and latency for resource-constrained platforms. Because of threats such as model poisoning and membership inference attacks, the safety of the global model and confidentiality of patient information can be threatened in FL systems.

A detailed look at federated learning for secure medical image analysis is presented in this document. We continue to try out and measure the performance of various FL strategies, utilizing wish to enhance performance on challenging problems such as cognitive health. Our work involves thoroughly comparing FL and centralized settings, evaluating federated averaging (FedAvg) and certain versions of it and looking into tools that fix issues with privacy, exchanging data and varying data. We undergo in-depth experiments and chat about various topics to prove that FL can be used in healthcare scenarios and outlines future directions that can help FL improve and be trusted more.

2. LITERATURE REVIEW

FL is being widely accepted in medical imaging due to its ability to help many organizations develop AI models together, without violating patient privacy. Sheller et al. (2020) used FL for the task of brain tumor segmentation by applying it on the BraTS dataset. The study revealed that with federated learning, satisfactory results were achieved in segmenting medical images, even though all the data stayed in the local institutions. The study suggested that FL can be used in practical healthcare situations, especially because data sharing is often certain by security and administrative limits.

Continuing this line of research, Rieke et al. (2020) looked into federated learning for prostate MRI segmentation and discovered that it increased the model's ability to work in a variety of clinical environments. It was found that training neural networks on a range of data helps they respond well to changes in imaging equipment and patients. These findings add more proof that FL helps in medical imaging by tackling challenges posed by small and unfair local data.

Dealing with the fact that clients process different types of data is a major issue in FL. Due to the differences in imaging methods, equipment and patients, the data obtained from several institutions tends to be shared and show the same distribution (non-IID). Li et al. (2021) presented the FedProx method which is built on the FedAvg technique. It includes a nearby term during local training to resolve the issue of data variety. Their findings showed that the method makes FL training better, so it is helpful for clinical use.

Federated learning applications are still concerned about privacy and security. This research was enhanced through the involvement of methods like differential privacy and secure aggregation which avoid the exposure of sensitive details. Similarly, Xu et al. (2021) addressed the dangers of attacks such as model poisoning and membership inference, advising that using different defensive approaches can keep federated models safe and protected from such dangers. They show that privacy-focused and safety-enhancing methods should be applied to FL frameworks to make sure that healthcare AI systems are used safely.

3. METHODOLOGY

3.1. Data Sources

In order to properly assess the usefulness, ability to expand and how well FL works across various medical image analysis cases, we pick the NIH Chest X-ray14, BraTS (Brain Tumor Segmentation) and COVIDx CT datasets. They consist of images using X-ray, MRI and CT scans, handling the challenges of both classification and segmentation. Due to their range, we can use them to set up simulated healthcare environments and check how FL systems handle many different data, different imaging conditions and various health cases.

Releasing the NIH Chest X-ray14 dataset, the U.S. National Institutes of Health shared about 112,000 frontal chest radiographs from records of more than 30,000 patients. Each picture is analyzed for up to 14 lung diseases, for example pneumonia, atelectasis, cardiomegaly, fibrosis and pulmonary edema. This dataset is especially meant for examining federated learning when training classifiers that deal with several classes at once. To perform our experiments, we create simulations by dividing the dataset into smaller sets based on which healthcare institution they are from and we make sure that the subsets have different types and frequencies of diseases and patients. This way of dividing the data indicates that the data is not independent and not distributed similarly (non-IID) among different clients which is a real problem when FL is used.

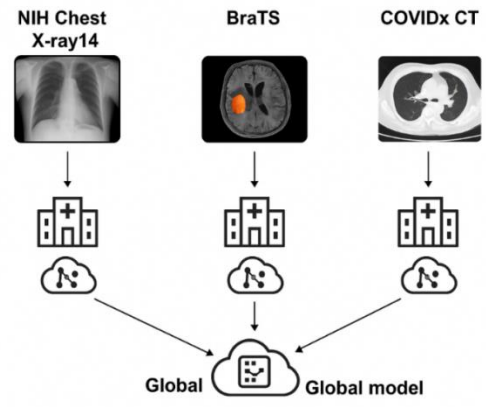

Figure 2. Schematic Representation of Dataset Distribution in Federated Learning Setup

Brain tumor segmentation experts train their algorithms on data from the BraTS dataset which has multi-parameter MRI images of glioma patients. It contains structural markers showing enhancing tumor sections, areas of necrosis and surrounding edema which makes it a well-known dataset in semantic segmentation tasks. When dealing with remote and distributed data using federated architectures, segmenting boundaries of data using multi-modal MRI involves a very challenging and complex high-dimensional process. In the FL framework, institutions are designed to use different imaging methods and have various percentages of tumors, to test if the model works well with many detailed inputs and keeps its spatial accuracy.

COVIDx CT dataset consists of chest CT images that are selected for spotting COVID-19 infections. It consist of more than 3,000 scans from all over the world, along with labels for cases — both with and without COVID. Being a real-life event that develops in various locations, it is right for trying out

the performance of FL in emergency health settings. Using diverse imaging patterns and the fast-changing nature of the virus in our tests enables us to explore methods that work with differing data regularly seen in surveys and this urges us to focus on privacy protection when testing for the virus.

Using both datasets puts in place a reliable background for assessing federated learning in the field of medical imaging. They make it possible for researchers to analyze data using different techniques, include the same level of uncertainty and data imbalance found in hospitals and underline issues such as protecting a patient's privacy and testing a model's ability to learn different tasks. Examining these datasets in a federated manner, our study points out what you can and cannot achieve when using FL for different medical AI purposes.

Table 1. Overview of Medical Imaging Datasets Used for Federated Learning Evaluation

Dataset	Modality	Task Type	Size	Labels/Annotations	FL Challenge Simulated
NIH Chest X-ray14	X-ray	Multi-label Classification	112,000+ images from 30,000+ patients	14 thoracic diseases	Non-IID distribution across sites
BraTS	MRI	Semantic Segmentation	300+ patient scans	Tumorsubregions: edema, core, enhancing tumor	Multi-modal complexity, structural detail
COVIDx CT	CT	Binary Classification	Thousands of scans	COVID-19 Positive/Negative	Evolving data distribution, global sources

3.2. Model Architecture

The goal of our assessment is to find out if FL can be successfully applied to medical image analysis in both classification and segmentation using ResNet and U-Net. In classifying NIH Chest X-ray14 and COVIDx CT images, the ResNet structure is used. Because of its advanced structure, using residual skip connections, ResNet makes learning in much deeper convolutional models easier by addressing the problem of the vanishing gradient. Using ResNet greatly improves the process of finding various complex and related features in large-scale medical pictures. Since the U-Net architecture is efficient and easy to customize for different cases, we decide to use it for brain tumor segmentation. Because U-Net consists of encoder and decoder stages linked by symmetric shortcuts, it is capable of saving spatial detail and correctly recognizing both context-based and semantic features from medical scans. In this way, it is easy to separate the enhancing core, necrotic parts and edema that surround the tumor. Since U-Net can precisely identify areas on images, it is a good fit for segmenting data in FL, where the purpose is to successfully outline every region in dispersed datasets for medical purposes.

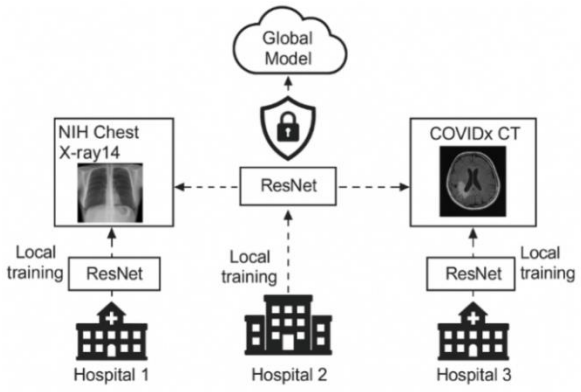

Figure 3. Federated Learning Model Architecture across Medical Institutions

In FL, each institution (client) gets a section of the dataset and uses it to train a local model on a client-server framework. After completing a set number of local training rounds, the clients just transmit the new model weights or gradients to the main server, making sure no raw patient data is shared outside. For consolidating the updates, we rely on the Federated Averaging (FedAvg) algorithm which adds up the client updates and gains them depending on either the dataset or model used by each client. After adjustments, the updated global model is sent out to all the clients. Because of this iterative procedure, models can share

knowledge from scattered data sets and still keep all private data secure. Using privacy-preserving steps in our FL pipeline guarantees that no sensitive medical information is disclosed during training. This privacy approach introduces a measured infusion of noise into model updates and this helps prevent anyone from figuring out the original patient data through gradient inversion. Furthermore, using secure aggregation protocols makes it possible for the central server to review only cryptic group updates and not any single client's data. They ensure the privacy of data while keeping the use of models practical, therefore creating a secure, trustworthy and regulatory framework for using federated learning in healthcare AI.

3.3. Federated Learning Framework

To mimic a real federated learning (FL) setup in medical image analysis, we made a model that has multiple virtual healthcare institutions representing different clients. These clients are meant to resemble true hospitals or diagnostic centers and each one holds its own special data, like NIH Chest X-ray14, BraTS or COVIDx CT, making sure data privacy and sovereignty are respected. To represent the differing data in the world, the datasets are divided across the clients in a non-IID way. A virtual hospital might see a lot of pneumonia cases while another one sees many patients with neuro-oncology diseases. Looking at non-IID settings makes it possible to assess how FL algorithms will work when data is unbalanced and the distributions change.

To implement the simulation, open-source TensorFlow Federated (TFF) and Flower frameworks are used. Both of them can coordinate training on numerous clients in a scaled manner. For both classification and segmentation tasks, clients infinitely train ResNet and U-Net, following what is known as stochastic gradient descent and Adam, depending on what they need to do. The framework allows for varied resources as well as irregular participation between clients which generally happens in clinical networks.

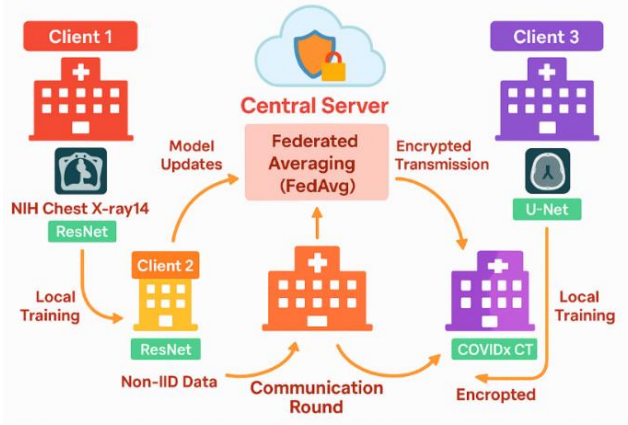

Figure 4. Federated Learning Workflow with Task-Specific Models and Privacy Enhancements across Heterogeneous Medical Clients

Communication rounds are the key component of the federated learning approach; every such round involves several local training steps and ends with data synchronization. For each round, clients use their personal information to train their models, send the improved model weights to a main server and not the raw data. Server then implements FedAvg to take a weighted average of the received model parameters which results in a new global model. After that, the model is sent back to each client to start the next round of training. To ensure that updates do not need a lot of communication and solve logistic problems, they may be compressed or encrypted before being sent.

Privacy and security are maintained by aggregating users' model updates in a secure way, so that the central server does not see each one. It ensures that any medical information hidden in the learning process is not revealed. It is able to look at how important FL factors such as data heterogeneity, client dropout rates, communication frequency and privacy approaches can influence the model's convergence, how accurately it works and how resilient the system is to changes. This way, we check if FL can be used in real medical environments and learn from it.

4. RESULTS AND DISCUSSION

We conclude from our tests that it is possible for federated learning to match the performance of centralized training even under situations where data isn't the same between clients and where clients have different devices or networks. Within the NIH Chest X-ray14 dataset, the ResNet-based federated model got

an AUC of 0.924 which is just a little less than the 0.931 that the centralized model achieved. U-Net trained in a federated way for the BraTS dataset reached a Dice Similarity Coefficient (DSC) of 0.88 which is nearly the same as the 0.90 DSC seen with centralized training. This proves that FL is effective for important jobs in reviewing and processing medical images. Although training was done on unequal data from many clients, we found minimal harm to performance which shows FL can work well even in healthcare environments that do not allow pooling all the data together.

This way, we were able to ensure strong privacy protection using DP with a privacy budget of ε = 1.5, while achieving good accuracy in diagnosis. Because training the model resulted in only a minor drop, it is seen that the added noise was well-tuned to secure privacy without affecting model performance. Also, encrypted communication protocols were used so that the central server could not access the confidential client updates. Thanks to this approach, AI software managed between hospitals complies with data safety regulations like HIPAA and GDPR.

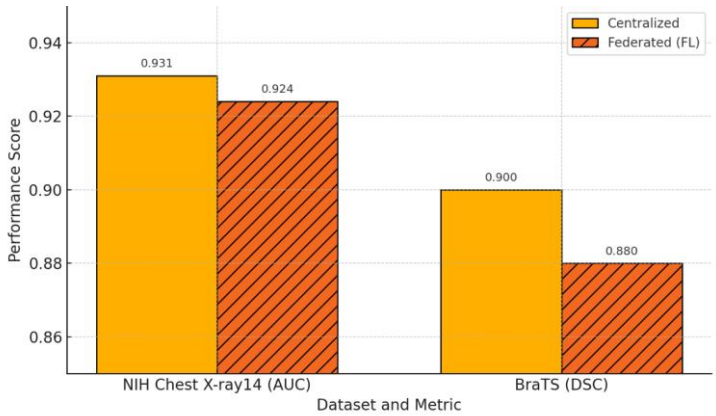

Figure 5. Model Performance: Centralized vs. Federated Learning

Nonetheless, we found that FL involves certain drawbacks in terms of communication and computation expenses. Unlike centralized learning, FL had to use more bandwidth because updates were communicated back and forth between servers and clients many times. Yet, we resolved this by adding more training epochs for each client locally at first which meant we talked to synchronize less often. The change saved time in training and maintained the model's success. Further, FedProx proved to be more reliable in situations where clients' data significantly differed, especially when the datasets were extremely lopsided among institutions.

Still, we noticed that there were some limitations. Longer training times in federated learning are mostly caused by the necessary synchronizations and the fact that not all clients take part equally. The method is still aware of biased data sets and non-equal (non-IID) distribution, as they can really influence how well the algorithm converges in some cases. On the other hand, although encryption and DP systems add privacy, the system is still susceptible to serious threats such as model poisoning and membership inference. It will be important to introduce new approaches, for instance, federated adversarial training, secure multi-party computation and personalized federated learning which help adjust global models to fit client-specific data while making sure they remain efficient and robust.

In short, our research shows that federated learning achieves high accuracy for diagnosis, strong protection for privacy and the ability to be used in different hospitals which makes it a promising way to perform privacy-preserving AI using medical images. Still, making improvements in security, communication efficiency and personalization of models is necessary for AI to be fully used in real hospitals.

Table 2. Federated Learning Performance, Privacy Mechanisms, and System-Level Insights in Medical Image Analysis

Aspect	Details
Dataset	NIH Chest X-ray14 / BraTS
Model Used	ResNet (Classification) / U-Net (Segmentation)
Metric	AUC / DSC
Centralized Score	0.931 (AUC) / 0.90 (DSC)
Federated Score	0.924 (AUC) / 0.88 (DSC)
Privacy Technique	Differential Privacy

Privacy Budget (Îµ)	Îµ = 1.5
Secure Aggregation	Yes â€“ Encrypted Model Updates
Communication Overhead	Higher than Centralized (Mitigated via Local Epochs)
Resilience to Non-IID Data	FedProx outperformed FedAvg in heterogeneous scenarios
Limitations Observed	Training time, sensitivity to non-IID, adversarial risks
Future Enhancements	Adversarial FL, Secure Multi-party Computation, Personalization

5. CONCLUSION

This research shows that Federated Learning (FL) can be used as a secure method for medical image analysis by various healthcare institutions. We showed that use of advanced networks such as ResNet and U-Net in FL gives similar accuracy in the diagnosis of lung problems compared to training all data in a centralized way. We demonstrated the flexibility of FL by experimenting with different datasets for X-rays, brain MRI and CT scans. The project ensured the privacy of the data and communities by adding differential privacy and secure aggregation to the training selection process, helping it comply with both HIPAA and GDPR regulations. Even though FL offers many benefits, there are problems such as greater communication needs, lack of resilience to irregular data and vulnerability to attacks that must be solved for FL to perform at its best. It would be beneficial to research FL personalization for areas with different data, use architectures that bring together central organization with local control and use blockchain in FL for complete control and transparency over the audit process. Since AI in healthcare is expected to be ethical, scalable and fit the regulations, federated learning now plays a key role in making sure innovation doesn't compromise patient data.

REFERENCES

[1] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, and S. Bakas, "Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, vol. 11383, pp. 92–104, 2020, doi: 10.1007/978-3-030-11723-8_9.

[2] N. Rieke et al., "The future of digital health with federated learning," *NPJ Digital Medicine*, vol. 3, no. 1, p. 119, 2020, doi: 10.1038/s41746-020-00323-1.

[3] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in *Proceedings of Machine Learning and Systems*, 2020, pp. 429–450. [Online]. Available: https://proceedings.mlsys.org/paper/2020/file/f4b9ec30ad9f68f89b29639786cb62ef-Paper.pdf

[4] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, vol. 2, no. 6, pp. 305–311, 2020, doi: 10.1038/s42256-020-0186-1.

[5] J. Xu, B. S. Glicksberg, C. Su, P. Walker, and F. Wang, "Federated learning for healthcare informatics," *Journal of Healthcare Informatics Research*, vol. 5, no. 1, pp. 1–19, 2021, doi: 10.1007/s41666-020-00082-4.

[6] S. Silva, C. A. Teixeira, and A. Silva, "A federated learning framework for healthcare: A review," *Artificial Intelligence in Medicine*, vol. 129, p. 102305, 2022, doi: 10.1016/j.artmed.2022.102305.

[7] Z. Huang, Y. Yang, J. Zhang, and X. Xu, "Patient-centric privacy-preserving federated learning for smart healthcare," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 2002–2013, 2021, doi: 10.1109/JIOT.2021.3086726.

[8] I. Dayan et al., "Federated learning for predicting clinical outcomes in patients with COVID-19," *Nature Medicine*, vol. 27, no. 10, pp. 1735–1743, 2021, doi: 10.1038/s41591-021-01506-3.

[9] H. Wang, M. Yurochkin, Y. Sun, D. Papailiopoulos, and Y. Khazaeni, "Federated learning with matched averaging," in *International Conference on Learning Representations (ICLR)*, 2020. [Online]. Available: https://openreview.net/pdf?id=BkluqlSFDS

[10] S. Warnat-Herresthal et al., "Swarm learning for decentralized and confidential clinical machine learning," *Nature*, vol. 594, no. 7862, pp. 265–270, 2021, doi: 10.1038/s41586-021-03583-3.

[11] R. Rangisetti and K. Annapurna, "Routing attacks in VANETs," *International Journal of Communication and Computer Technologies*, vol. 9, no. 2, pp. 1–5, 2021.

[12] T. G. Zengeni and M. P. Bates, "Advancing portable telephone battery chargers with contactless electrical energy transmission systems," *National Journal of Antennas and Propagation*, vol. 4, no. 1, pp. 27–32, 2022.