Educational Innovation Research, 2026, Volume 4, Issue 6
https://bbwpublisher.com/journal/EIR
Online ISSN: 3029-1852
Print ISSN: 3029-1844

Applied Research on Federated Learning Techniques for Medical Imaging Privacy and Security

Xiao Liu*

City University of Macau, Zhuhai 519000, Guangdong, China

Author to whom correspondence should be addressed.

Abstract: With the continuous advancement of artificial intelligence technologies in medical imaging, there is an increasing demand for the collaborative utilization of imaging data across multiple institutions. However, factors such as decentralized data storage, privacy protection constraints, and heterogeneous imaging distributions have become significant barriers to achieving optimal model training performance. To address challenges—including the difficulty of centralized sharing of medical imaging data and the limited adaptability of deep learning models across diverse scenarios—the present work proposes a federated learning-based application methodology designed with privacy and security in mind. This methodology leverages a distributed collaborative training framework and integrates key technologies—including model parameter aggregation, secure privacy-preserving computation, and heterogeneous imaging adaptation—to optimize the model training pipeline for medical imaging tasks such as classification, object detection, and image segmentation. Model performance is evaluated using publicly available medical imaging datasets, and the effectiveness of various federated learning strategies is compared in terms of recognition accuracy, model stability, and privacy protection capabilities. The results demonstrate that the federated learning approach not only reduces the need for exchanging raw imaging data but also maintains robust model recognition performance while enhancing the model's adaptability in multi-source, heterogeneous imaging environments, thereby providing a viable technical solution for secure collaborative model training in intelligent medical imaging analysis tasks.

Keywords: Medical imaging; Federated learning; Privacy security; Secure aggregation; Deep learning

Online publication: June 26, 2026

1. Introduction

Medical imaging data encompass a wealth of lesion characteristics and structural information, serving as an essential data foundation for AI-assisted analysis; however, varying institutional differences in equipment types, acquisition standards, and sample distributions pose significant challenges for the collaborative

utilization of such imaging data, necessitating robust data management approaches. Federated learning is an emerging paradigm for collaborative model training across multiple institutions, widely applied across diverse fields. This technology enables joint optimization across multiple nodes through the interaction of model parameters, offering a novel training framework for intelligent medical imaging computation [1]. Currently, the application of federated learning in medical imaging still faces several challenges—such as model bias arising from data heterogeneity, suboptimal parameter transfer efficiency, and the need to balance privacy protection mechanisms with model performance—making the optimization of these technical hurdles a highly relevant research and practical priority.

2. Medical imaging: Federated learning technology foundation and application architecture

2.1. Medical imaging data features and training requirements

Medical imaging data encompasses a variety of modalities, including CT, MRI, X-ray, and ultrasound images; its data structure is characterized by high dimensionality, strong spatial correlations, and complex semantic features. The pixel distribution, tissue boundaries, and lesion regions within these images collectively serve as crucial inputs for model recognition; deep learning models must extract visual features at multiple abstraction levels to perform tasks such as disease classification, region localization, and abnormality detection. The imaging data collected from different medical imaging devices varies in terms of equipment parameters, imaging protocols, resolution, and sample composition, resulting in data that does not follow an independent and identically distributed (i.i.d.) distribution. Intelligent medical imaging analysis must simultaneously accommodate the representation of multi-source data characteristics and the requirements of distributed computing, making this an ideal application scenario for federated learning architectures [2].

2.2. Federated learning distributed training architecture

Federated learning technology enables collaborative model training through the exchange of model parameters between client nodes and a central aggregation server; its core architecture consists of three components: local training, parameter transmission, and global aggregation. In medical imaging application scenarios, each computing node is responsible for training the model using its local medical imaging data and then sends the updated model parameters to the aggregation server. The server then integrates the parameter information from the various nodes using predefined algorithms to construct a new global model. This architecture replaces the traditional centralized data aggregation approach by distributing the model training process across multiple computing entities, thereby keeping the medical imaging data within its original storage environment. The collaborative relationship between nodes, the parameter update mechanism, and the model iteration workflow within the federated learning architecture collectively form the foundational operational framework for intelligent medical imaging computing.

2.3. Medical imaging task model adaptation method

The type of medical imaging task determines the structural selection and training objectives for federated learning models; among these, classification, detection, and segmentation tasks represent the primary application domains in AI-assisted medical imaging analysis. Image classification tasks rely on convolutional neural networks or visual depth models to extract global image features for determining disease categories

and abnormal conditions; object detection tasks require models to identify the location and category of lesions, imposing demands on the model's ability to represent local region features; whereas image segmentation tasks focus on pixel-level prediction, necessitating models that can accurately delineate tissue structures and lesion boundaries [3]. Different tasks correspond to distinct model architectures and parameter scales; thus, the federated learning framework must be designed to accommodate the training processes of various types of deep learning models, ensuring that model parameter updates can meet the computational requirements of intelligent medical imaging analysis.

3. Key federated learning technologies for medical imaging with privacy and security considerations

3.1. Federated parameter aggregation optimization method

Medical imaging federated learning enables joint modeling across multiple computational nodes; each node trains its model using its local medical imaging samples and then sends the updated model parameters to a central aggregation node to facilitate the global model iteration. The FedAvg algorithm performs weighted fusion of the model weights from different nodes, allowing nodes with larger sample sizes to contribute more significantly to the global model update. Its aggregation process is represented as follows:

$$w_{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w_k^t \tag{1}$$

Here, w_{t+1} denotes the global model parameters after the $(t+1)$-th round of training; w_k^t represents the model parameters of the k-th client after the t-th round of training; n_k is the number of samples from the k-th client; and n is the total number of samples across all nodes, and K is the total number of participating clients. This method is suitable for tasks such as medical image classification and object detection; it enables the fusion of image features learned by different nodes, thereby enhancing the overall expressive power of the model.

To address the issue of uneven distribution of medical imaging data, the FedProx algorithm incorporates a proximal constraint term into its local training objective, thereby limiting the extent of deviation between client models and the global model, ensuring a certain level of stability when imaging data from different institutions are used for training. During the actual training process, model aggregation should take into account factors such as the number of samples per node, the number of training epochs, and the model's convergence status to prevent the features from a single node from exerting an overly dominant influence on the global model.

3.2. Training process privacy enhancement protection technology

Although federated learning for medical imaging avoids the direct transmission of the entire original image dataset, the model parameters may still contain partial data feature information; therefore, privacy-preserving techniques should be employed to safeguard the training process. The differential privacy approach leverages random noise perturbation to reduce the correlation between parameter updates and individual training samples; its basic formulation is as follows:

$$M(D) = f(D) + Noise(\sigma) \tag{2}$$

Here, $M(D)$ denotes the model update result after incorporating privacy protection, $f(D)$ represents the original training function, and $Noise(\sigma)$ stands for the perturbation term configured according to the privacy strength parameter. In medical imaging tasks, the noise level should be adjusted based on metrics such as disease classification accuracy and lesion detection capability to prevent the privacy-enhancement process from interfering with the extraction of critical medical features.

The Secure Aggregation Technology focuses on protecting the parameter transmission phase; model update information uploaded by the client is encrypted before being incorporated into the federated computation, ensuring that the aggregation endpoint receives only the overall update result. This technology complements Differential Privacy: the former mitigates the risk of individual node parameter exposure, while the latter reduces the correlation of sensitive information during model updates, jointly safeguarding the security of the federated medical imaging training process.

3.3. Heterogeneous medical imaging data adaptation method

Medical imaging federated learning requires addressing data distribution disparities among different computational nodes; the imaging manifestations of the same disease may vary across different devices, scanning protocols, or patient populations, which can lead to model bias during the feature extraction phase [4]. Personalized federated learning leverages a local parameter tuning approach, preserving the inherent patterns of each node's imaging data while sharing global model features, enabling the model to adapt to the diverse imaging environments across different institutions. During the training process, the model can distinguish between a base feature layer and a task-specific layer: the base layer is responsible for learning the shared imaging structural information across multiple nodes, whereas the task-specific layer captures the local feature variations within individual nodes. Domain adaptation techniques employ feature space mapping and distribution distance constraints to mitigate discrepancies in grayscale distribution, texture characteristics, and structural information among images from different sources, thereby enabling the model to obtain more stable feature representations. To address the challenges of data class imbalance and varying sample sizes in medical imaging datasets, the federated training process should also incorporate adjustments to the parameter update strategy—based on factors such as the node's data volume, class distribution, and model feedback status—so that images from diverse sources can collectively participate in model training.

3.4. Federated training communication efficiency optimization strategy

Medical imaging deep learning models typically involve a large number of parameters; particularly in 3D image analysis and Vision Transformer models, the model synchronization process can generate a significant communication burden. Federated learning systems need to minimize unnecessary parameter transfers and concentrate communication resources on the information that is most critical for model updates. Parameter compression techniques—such as quantization encoding, gradient sparsification, and parameter filtering—are employed to compress the model update data uploaded by client devices; for example, by retaining gradient parameters with large magnitude changes while reducing the bandwidth consumption incurred by low-contribution parameters. During the training scheduling phase, the system can adjust the synchronization frequency based on the model's convergence status: maintaining a high interaction frequency during the initial training phase ensures consistency in the learning directions across different nodes; once the model

enters the stable phase, the frequency of parameter exchanges is reduced to minimize communication overhead [5]. Federated learning for medical imaging must also account for variations in node computational capabilities; the training pace should be adjusted according to device computing power, network conditions, and task load to ensure that the computational resources across multiple nodes operate in a coordinated manner.

4. Medical imaging federated learning application experiment and performance verification

4.1. Experimental environment and dataset configuration

The experiment utilized the NIH Chest X-ray 14 public chest X-ray image dataset for model evaluation; this dataset contains 112,120 chest X-ray images and information on 30,805 patients, covering 14 categories of chest disease labels – the dataset size and number of disease categories adequately meet the training requirements for medical imaging classification tasks. The data was partitioned by patient, and four simulated medical imaging computing nodes were created; each node retained independent training samples to construct a multi-institutional federated learning environment. The experiment employed the PyTorch deep learning framework, selecting DenseNet121 as the base classification network to extract hierarchical image features leveraging its dense connection architecture. During the training phase, image size normalization, grayscale normalization, and data augmentation were uniformly applied to ensure consistent input data formats across all nodes. Four distinct model architectures were configured for this experiment: centralized DenseNet121, FedAvg, FedProx, and DP-FedAvg. Among these, FedAvg was used to validate the effectiveness of parameter aggregation, FedProx was employed to analyze model stability under heterogeneous image data conditions, and DP-FedAvg was utilized to investigate the impact of privacy-preserving techniques on model performance.

4.2. Federated learning model training experiment design

The experiment begins by constructing a centralized DenseNet121 model; all training samples are fed into a unified computational environment to complete parameter training, and the classification results obtained from the test set are recorded as baseline data. During the federated training phase, four client nodes are deployed; each node reads its local chest imaging samples to perform model initialization, forward propagation, and parameter updates, and then uploads its updated model parameters to the aggregation server after each training round. The FedAvg model employs a weighted aggregation approach as described in Formula (1): parameter weights are calculated based on the number of samples at each client, and the updated model parameters from the different nodes are fused to form a new global model, which is then returned to the respective clients for further iterative training.

The FedProx model incorporates a proximal constraint term into its local training process to limit the magnitude of the offset between the client model parameters and the global model parameters, thereby ensuring that different nodes with distinct image distribution patterns maintain a relatively stable training direction. The DP-FedAvg model introduces a random perturbation term as described in Formula (2) during the parameter upload phase to apply privacy-preserving techniques to the model update parameters and to record the changes in classification performance under various perturbation conditions. During the experiments, all model groups were configured with identical network architectures, number of training

epochs, data partitioning schemes, and testing environments, ensuring consistency in the evaluation results across the different federated training methods.

4.3. Model performance metrics and experimental results analysis

The experiment employs AUC, F1-score, and number of communication rounds as evaluation metrics; among these, AUC measures the model's ability to distinguish between different disease categories, the F1-score provides a comprehensive assessment of both classification accuracy and recall rate, and the number of communication rounds reflects the parameter interaction cost during the federated training process. The performance of various training methods on the medical image classification task is presented in **Table 1**.

Table 1. Performance comparison of various federated learning methods in medical imaging tasks

Model-Based Methods	training mode	AUC	F1-score	Communication Round
DenseNet121	Centralized training	0.825	0.780	—
FedAvg	Federated Training	0.837	0.790	100
FedProx	Heterogeneous Optimization for Federated Training	0.841	0.800	100
DP-FedAvg	Privacy-Enhanced Federated Training	0.820	0.770	100

The experimental results demonstrate that the FedAvg model achieves an AUC of 0.837 and an F1-score of 0.790 without aggregating the raw image data; this performance is close to that obtained with centralized training, indicating that the weighted aggregation method described in Formula (1) is capable of integrating image features from different clients. The FedProx model further improves the AUC to 0.841 and achieves an F1-score of 0.800, suggesting that the proximal constraint helps mitigate the impact of varying image distribution across different nodes on the model's parameter updates. When the DP-FedAvg model incorporates privacy perturbation, its AUC decreases to 0.820, and its F1-score drops to 0.770; this change in model performance reveals a correlation between the intensity of parameter perturbation and image recognition accuracy, highlighting the need to adjust privacy parameters according to the specific requirements of the task.

4.4. Experimental results and application fit analysis

The multi-node training workflow in the experimental environment corresponds to a cross-institutional collaborative computing scenario for medical imaging; in this setup, each node performs local feature learning for the model training task, after which a shared model is constructed through parameter aggregation. During actual deployment, the medical imaging platform can configure the number of client nodes based on the computational resources available at each institution, the scale of the medical imaging dataset, and the network conditions, and then assign the model training task to these different computational nodes. Methods such as FedProx are suitable for application environments where there are variations in device types or imaging acquisition standards, as they can mitigate directional bias in node model updates; DP-FedAvg is appropriate for training scenarios with stringent requirements regarding the security of parameter exchange; in such cases, model administrators can adjust the perturbation intensity according to their desired recognition accuracy. The experimental results provide a quantitative basis for configuring federated models, deploying nodes, and selecting training parameters in medical imaging classification tasks.

5. Conclusion

The development of intelligent medical imaging analysis must balance both the efficiency of data resource utilization and the security of computational processes; federated learning offers a new technical paradigm for collaborative computing across multiple imaging sources. As deep learning models grow in scale and medical imaging scenarios become increasingly complex, future technological advancements should focus on areas such as model lightweighting, cross-domain feature fusion, secure collaborative computing, and engineering deployment adaptation—enabling intelligent analysis models to possess greater environmental adaptability and the capability for continuous iteration. Medical imaging federated learning will further integrate data resources, algorithmic models, and computational infrastructure, facilitating the deeper application of artificial intelligence technologies in precision imaging analysis and intelligent decision-support scenarios.

Disclosure statement

The author declares no conflict of interest.

References

[1] Wang SS, Lu SZ, Cao B, 2021, An Algorithm for Medical Image Object Detection Based on Federated Learning with Privacy Protection. Journal of Computer-Aided Design and Computer Graphics, 33(10): 1553–1562.

[2] Liu QH, Jiang MM, Zhao YW, et al., 2025, Research Progress in Personalized Federated Learning for Medical Data. Journal of the PLA Medical College, 46(1): 113–119.

[3] Ma ZB, Mi Y, Zhang B, et al., 2023, A Survey of Deep Learning Algorithms for Heterogeneous Medical Image Processing. Journal of Software, 34(10): 4870–4915.

[4] Zhang QZ, Zhu CY, 2023, Design and Implementation of a Lightweight Medical Imaging Auxiliary Diagnosis System Based on Federated Learning. Information and Computers (Theoretical Edition), 35(2): 44–47.

[5] Zhang C, Li W, 2025, A Review of Federated Learning in the Healthcare Sector: Applications, Challenges, and Future Research Directions. Journal of Engineering Science, 47(9): 1825–1840.