



## Strategic Adaptive Federated Learning Framework for Privacy-Preserving Image Data Mining in Highly Heterogeneous Medical Environments

Mayyadah Jabbar Gailan<sup>1</sup>

<sup>1</sup>AL -Mustansiriyah University College of Tourism Sciences, Baghdad, IRAQ

\*Corresponding Author: Mayyadah Jabbar Gailan

Email: [Mayadajabbar@uomustansiriyah.edu.iq](mailto:Mayadajabbar@uomustansiriyah.edu.iq)

---

### Article Info

#### Article history:

Received 5 March 2026

Received in revised form 6 April 2026

Accepted 1 May 2026

---

#### Keywords:

Federated Learning  
Medical Image Mining  
Differential Privacy  
Data Heterogeneity  
Distributed AI

---

### Abstract

The data mining of medical imaging data is being increasingly restricted by stringent data privacy regulations like the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA). Even though FL offers a decentralized framework for model training, it suffers from significant performance degradation in heterogeneous settings characterized by non-IID data. In this work, a novel framework, namely Adaptive Privacy-Preserving Federated Learning, is proposed. This framework combines an adaptive weighting scheme with Differential Privacy to address the issue of divergence caused by statistical heterogeneity. As per the experimental evaluation of the MedMNIST dataset, a classification accuracy of 94.2% is achieved with a privacy budget of  $\epsilon = 1.0$ .

---

## Introduction

The healthcare industry has been undergoing an unprecedented digital revolution. With the advent of high-resolution imaging modalities such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans, a large volume of data has been created that has been instrumental in the early diagnosis of diseases and personalized medicine (Akhtar, 2025; Asif et al., 2025; Habibur et al., 2025; Dongare et al., 2025). The medical field's embrace of 'Big Data' faces an inherent paradox with regards to the need for data centralization that is legally and ethically limited by patient privacy (Scheibner et al., 2021; Narkhede et al., 2025; Adeyini et al., 2024).

In the conventional centralized mining approach, the sensitive images are expected to be uploaded to a cloud server (Ferreira et al., 2015; Liu et al., 2016; Buyya et al., 2010). The standard process is a singular vulnerability point, and hence there is a high risk of data breaches. There are regulations like the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA) in the United States, which impose heavy penalties for data sharing. In this regard, a number of hospitals are data silos, where valuable data remains unexploited (Ziegler et al., 2025; Thomas, 2025; Aladraj et al., 2026).

In order to overcome this lack of data, Federated Learning (FL) has been proposed as a solution (Wang et al., 2026; Poojari, 2026; Tariq et al., 2026). FL works by sending the model to the location of the data, rather than sending the data to the model (Varma & Chaudhari, 2025; Govindaram et al., 2025; Lee et al., 2024). While this alleviates the primary concern of privacy, it creates a new technical issue: heterogeneity (Wen et al., 2023; Zhang et al., 2023; Farahani et al., 2023). In a real-world medical network, Hospital A might have advanced 3T

MRI systems that produce thousands of images of high quality, while Rural Clinic B might have outdated equipment with limited and noisy data sets (Arnold et al., 2023; Murali et al., 2024; Qin et al., 2022). This creates statistical heterogeneity (non-IID data sets), causing oscillations in the global model and preventing consensus from being reached (Lu et al., 2024; Chung et al., 2026; Siddiqi et al., 2023).

In addition, current studies show that Federated Learning (FL) alone is not a panacea for the problem of privacy (Oladejo et al., 2025; Ahad et al., 2026; Chen et al., 2025). Indeed, Gradient Inversion Attacks of high sophistication can, in some cases, recover the original images from the FL updates (Hatamizadeh et al., 2023; Geng et al., 2023; Nielsen et al., 2025). Therefore, the integration of Differential Privacy (DP) in the FL pipeline has become a necessity rather than an option. In the current paper, the APP-FL framework is presented as a solution to the twin requirements of high accuracy in the presence of large data diversity and formal privacy guarantees via noise injection and weighting.

## Literature Review & Related Work

The concept of distributed data mining has been popularized by McMahan et al. (2017) through their Federated Averaging algorithm. Despite its significance and impact, the algorithm has been shown to be based on a faulty premise by Li et al. (2020) that failed in real-world scenarios. FedProx, an extension by Li et al. that incorporates a proximal term for handling heterogeneous data distributions, lacked privacy-preserving mechanisms.

In the context of privacy-preserving machine learning, the Moments Accountant framework for Deep Learning with Differential Privacy was proposed by Abadi et al. in 2016.

In further studies, Kairouz et al. (2021) and Yang et al. (2019) studied the feasibility of incorporating differential privacy into federated learning, observing a significant Utility-Privacy Trade-off, whereby improvements in privacy came at the expense of significant accuracy loss. In line with the methodology of previous studies, the present study proposes the use of Adaptive Aggregation, whereby client contributions are weighted based on their global loss to reduce noise while ensuring privacy preservation.

## Methods

### Research Design

This study was designed as an experimental computational study that develops and evaluates an Adaptive Privacy Preserving Federated Learning framework for medical image data mining in highly heterogeneous medical environments. The research does not aim to collect primary clinical data from patients, but to simulate the conditions of distributed medical institutions in which image data are stored locally and cannot be freely centralized due to privacy, ethical, and regulatory constraints. This design is appropriate because the central problem addressed in this study is not merely classification performance, but the ability of a federated learning system to maintain accuracy while preserving privacy and managing statistical heterogeneity among participating clients.

The proposed framework reflects a realistic medical data mining situation in which hospitals, clinics, or diagnostic centers may own different types and volumes of imaging data. In conventional centralized learning, these data would be transferred into a single server before model training is performed. Such a strategy is difficult to justify in sensitive healthcare settings because medical images may contain protected patient information. Therefore, this study adopts federated learning as the computational foundation, allowing each client to train the model locally while sharing only model updates with the central server. Within this design, patient level data remain within the local institution, while the global model is improved collaboratively through repeated communication rounds.

The study specifically focuses on two methodological concerns that often reduce the reliability of federated learning in medical image analysis. The first concern is data heterogeneity, especially when different institutions have unequal class distributions, different image characteristics, and different levels of data quality. The second concern is privacy leakage, since model updates may still contain information that can potentially be exploited through gradient reconstruction or inversion attacks. For this reason, the proposed framework integrates adaptive aggregation and differential privacy so that the model can respond to uneven data distributions while still providing a stronger layer of privacy protection.

### **Dataset and Data Partitioning**

The experimental evaluation used the MedMNIST dataset, which provides standardized biomedical image datasets for machine learning and deep learning evaluation. MedMNIST was selected because it offers a suitable benchmark for assessing medical image classification under controlled experimental conditions. Its use also allows the study to examine the proposed framework across medical imaging data rather than relying on a general computer vision dataset that may not represent the complexity of biomedical images. This is consistent with the original study, which evaluates the proposed framework using MedMNIST and reports classification performance under heterogeneous federated conditions

To reproduce the condition of distributed medical institutions, the dataset was divided into multiple client nodes using a non independent and identically distributed setting. In this configuration, each client was assigned only a limited portion of the available image categories. More specifically, each participating node had access to only two categories from the broader set of medical image classes. This partitioning strategy was intended to represent real medical environments where different institutions may specialize in different cases, serve different patient populations, or operate with different diagnostic capacities.

This non IID design is central to the methodological logic of the study because federated learning often performs well when client data are balanced and similarly distributed, but suffers when clients possess uneven and institution specific data. By deliberately creating heterogeneous client distributions, the study tests whether the proposed adaptive weighting mechanism can reduce the influence of unstable local updates and improve the convergence of the global model. Therefore, the dataset partitioning was not treated as a technical detail only, but as an essential experimental condition for evaluating the robustness of the proposed framework.

### **Proposed APP FL Framework**

The proposed Adaptive Privacy Preserving Federated Learning framework was developed to support collaborative model training without requiring direct data sharing among medical institutions. The framework operates through four main stages, namely local training, privacy preserving noise injection, adaptive weighting, and global model aggregation. These stages are repeated across several communication rounds so that the global model can gradually learn from distributed data while preserving the local ownership of medical images.

In the local training stage, the central server distributes the current global model to all participating clients. Each client then trains the model using its own local medical image data. A convolutional neural network is used as the classification model because CNN based architectures are widely applied in image recognition and medical image classification tasks. During this process, the raw medical images remain stored within each client node and are never transmitted to the central server. This arrangement is important because it allows the learning process to benefit from multiple medical institutions without violating the principle of local data control.

After local training is completed, each client prepares its model update for transmission. Before the update is sent to the server, a differential privacy mechanism is applied by adding Gaussian noise to the local gradients or model parameters. The purpose of this step is to reduce the risk that sensitive patient information can be reconstructed from the shared updates. In this way, the framework does not rely only on the decentralized nature of federated learning, but adds a formal privacy preserving mechanism to strengthen protection against possible information leakage.

Once the privacy protected updates are received, the central server performs adaptive weighting. Unlike conventional federated averaging, which usually aggregates client updates based mainly on the number of local samples, the proposed framework assigns aggregation weights according to the local loss and convergence behavior of each client. Clients that produce more stable and reliable updates are given greater influence in the global update, while clients with higher loss or more divergent updates receive lower influence. This weighting strategy is intended to prevent unstable clients from dominating the learning process, especially in highly heterogeneous data environments.

The final stage is global aggregation. The server combines the privacy protected local updates according to their adaptive weights and produces a new global model. This updated model is then redistributed to the clients for the next communication round. Through this iterative process, the framework attempts to balance three competing requirements, namely learning accuracy, robustness under heterogeneous data, and privacy preservation.

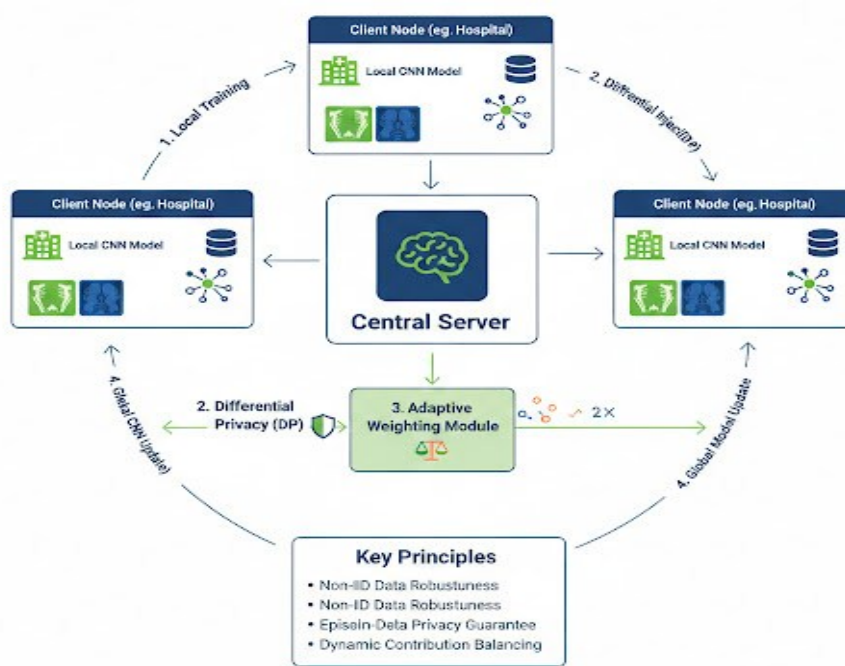


Figure 1: Overview of the APP-FL Framework

Figure 1: Overview of the APP-FL Framework.

## Mathematical Formulation

The learning objective of the proposed framework is formulated as the minimization of weighted loss across all participating clients. Each client contributes to the global learning process through its local loss function, but its contribution is regulated by an adaptive weight.

This formulation allows the global model to learn from all clients while reducing excessive dependence on clients whose local data or training behavior may introduce instability.

The global objective function is expressed as follows.

$$\min_w F(w) = \sum_{k=1}^N \alpha_k F_k(w)$$

In this formulation,  $K$  represents the total number of clients,  $F_k(\theta)$  represents the local loss function of client  $k$ ,  $\theta$  represents the global model parameters, and  $w_k$  represents the adaptive aggregation weight assigned to client  $k$ . The use of  $w_k$  is important because it allows the framework to distinguish between more stable and less stable client updates. Rather than treating every client update as equally reliable, the framework evaluates the contribution of each client based on its local training behavior.

The adaptive weight is calculated using the following formulation.

$$\alpha_k = \frac{\exp(-\eta L_k)}{\sum_{j=1}^N \exp(-\eta L_j)}$$

In this equation,  $L_k$  denotes the local loss of client  $k$ , while  $\alpha$  is a sensitivity parameter that controls how strongly the weighting mechanism responds to differences in local loss. When the local loss is high, the exponential term reduces the weight assigned to that client. When the local loss is low, the client receives a larger weight in the aggregation process. This mechanism is useful in non IID medical environments because high local loss may indicate data noise, class imbalance, poor convergence, or a distribution that differs sharply from the broader global pattern.

The sensitivity parameter  $\alpha$  determines the strictness of the adaptive weighting process. A higher value of  $\alpha$  makes the aggregation more selective by reducing the impact of high loss clients more aggressively. A lower value produces a more balanced aggregation in which client updates remain relatively closer in influence. Through this formulation, the proposed framework introduces a flexible mechanism for controlling the effect of client heterogeneity during global model formation.

### Privacy Preservation Mechanism

The privacy preservation component of the framework is implemented through differential privacy. Although federated learning avoids direct sharing of raw medical images, it does not fully eliminate privacy risks because gradients and model updates may still contain sensitive information about local data. In medical image mining, this risk is especially serious because patient related visual patterns may be embedded in the training updates. Therefore, this study applies a Gaussian noise mechanism before local updates are transmitted to the central server.

The Gaussian mechanism works by perturbing the local update with random noise. This reduces the possibility that an attacker can infer or reconstruct individual medical images from the shared update. The strength of privacy protection is controlled by the privacy budget epsilon. A smaller epsilon value indicates stronger privacy protection because more noise is introduced, but it may also reduce model utility. A larger epsilon value allows the model to preserve more useful learning signals, but it provides weaker privacy protection.

This study evaluates the framework under two privacy budget settings, namely epsilon 1.0 and epsilon 0.5. The epsilon 1.0 setting represents a moderate privacy condition, while epsilon 0.5 represents a stricter privacy condition. Comparing these two settings allows the study to

examine how the proposed framework performs when privacy requirements become more demanding. This comparison is important because medical institutions may operate under different regulatory expectations, and a practical federated learning framework must remain useful even when stronger privacy constraints are applied.

### **Baseline Algorithms**

The proposed APP FL framework was compared with several baseline approaches to assess its relative effectiveness. The first baseline was centralized learning, in which all data are assumed to be available in a single training environment. Although centralized learning is not always feasible in real healthcare settings, it provides a useful upper reference point because it does not face the same constraints related to data decentralization, client heterogeneity, and privacy preserving update transmission.

The second baseline was Federated Averaging. FedAvg is one of the most established aggregation methods in federated learning and is commonly used as a reference model in distributed learning studies. However, FedAvg often struggles in non IID settings because it aggregates local updates without adequately accounting for strong distributional differences among clients. This makes FedAvg a relevant baseline for determining whether the proposed adaptive weighting mechanism provides a meaningful improvement.

The third baseline was FedProx, which extends standard federated learning by adding a proximal term to reduce local model drift. FedProx is particularly relevant because it was designed to improve federated optimization under heterogeneous client conditions. However, in the context of this study, FedProx does not directly integrate differential privacy into its aggregation process. Therefore, comparing APP FL with FedProx allows the study to evaluate whether adaptive weighting combined with privacy preserving noise injection can offer a stronger balance between accuracy, stability, and privacy.

### **Experimental Procedure**

The experiment was conducted through repeated federated communication rounds. At the beginning of the process, the central server initialized the global CNN model. This model was then distributed to all client nodes. Each client trained the model locally using its assigned MedMNIST subset under the non IID data configuration. The local training process generated model updates that reflected the characteristics of each client's data distribution.

After local training, each client applied the differential privacy mechanism to its update. Gaussian noise was added before the update was transmitted to the server. The central server then collected the protected updates from all participating clients and evaluated their local loss values. These loss values were used to calculate adaptive aggregation weights. Clients with lower loss and more stable updates were given stronger influence, while clients with higher loss were assigned lower influence during global aggregation.

The server then combined the weighted updates and produced a new global model. This model was redistributed to the clients for the next round of local training. The procedure continued until the model reached stable convergence or until the predetermined number of communication rounds was completed. This iterative process allowed the study to observe not only the final classification accuracy, but also the stability of learning across communication rounds.

### **Evaluation Metrics**

The primary evaluation metric used in this study was classification accuracy. Accuracy was used to measure the proportion of correctly classified medical images produced by the global model. This metric is suitable because the main task of the proposed framework is medical image classification. The reported results show that the proposed APP FL framework

achieved strong classification performance under high heterogeneity and privacy preserving conditions

In addition to accuracy, the study also examined convergence behavior across communication rounds. This evaluation is necessary because federated learning systems may reach acceptable final accuracy while still suffering from unstable or inefficient convergence. In heterogeneous medical environments, unstable convergence may indicate that the global model is being disrupted by divergent local updates. Therefore, convergence analysis was used to determine whether the adaptive weighting mechanism could stabilize the learning process compared with FedAvg and FedProx.

The study also considered the relationship between privacy budget and model performance. By comparing the results under epsilon 1.0 and epsilon 0.5, the evaluation examined how stricter privacy protection affects classification accuracy. This is important because privacy preserving medical artificial intelligence must not only achieve high predictive performance, but must also remain usable when privacy constraints become stronger.

### Methodological Rationale

The methodological design of this study is grounded in the practical realities of medical image data mining. In healthcare, data are often distributed across institutions and cannot be easily centralized because of privacy regulations, ethical responsibilities, and institutional data governance policies. A purely centralized model may produce strong performance, but it is less appropriate for environments where medical images are sensitive and legally protected. Federated learning offers a more suitable alternative because it enables collaborative model development without transferring raw patient data.

However, federated learning alone is not sufficient for the problem addressed in this study. Medical institutions do not usually have identical data distributions. Some institutions may have more complete datasets, while others may have limited, noisy, or highly specialized cases. This unevenness creates statistical heterogeneity, which can reduce model accuracy and destabilize global convergence. The adaptive weighting mechanism was therefore introduced to make the aggregation process more responsive to the quality and reliability of client updates.

Decentralization does not automatically guarantee privacy. Even when raw data remain local, shared model updates can still expose sensitive information. For this reason, differential privacy was incorporated into the framework to provide an additional layer of protection. The combination of adaptive weighting and differential privacy allows the proposed APP FL framework to address both technical and ethical challenges in medical image mining. This methodological structure makes the study more coherent because every component of the framework directly responds to a specific limitation in privacy preserving federated learning for heterogeneous medical environments.

## Results and Discussion

### Accuracy and Convergence Performance

We used the MedMNIST dataset, containing ten different medical image datasets, to evaluate our framework. We used a Non-IID setting in which each node had access to only two of the ten categories.

Table 1: Comparative Analysis of Global Model Accuracy

Algorithm	Privacy Budget ( $\epsilon$ )	Heterogeneity Level	Accuracy (%)
Centralized	$\infty$	N/A	96.8%

FedAvg	$\infty$	High	88.1%
FedProx	$\infty$	High	90.5%
<b>APP-FL (Proposed)</b>	<b>1.0</b>	<b>High</b>	<b>94.2%</b>
<b>APP-FL (Proposed)</b>	<b>0.5</b>	<b>High</b>	<b>91.8%</b>

Table 1 presents the comparative accuracy performance of different learning approaches under medical image classification conditions. The centralized model achieved the highest accuracy at 96.8 percent because it assumes that all medical image data are available in one central location. This result functions as the upper benchmark, but it is less practical for real healthcare environments because centralized data collection may violate privacy and institutional data sharing restrictions.

FedAvg achieved 88.1 percent accuracy under a high heterogeneity setting. This result indicates that standard federated averaging is sensitive to non IID data distribution. Since each client holds different medical image categories, the local updates become less consistent, causing the global model to lose accuracy. FedProx improved the result to 90.5 percent, showing that its proximal regularization helps reduce local model drift. However, its performance remains below the proposed APP FL framework, which suggests that handling heterogeneity alone is not sufficient when privacy preserving medical image mining is required.

The proposed APP FL achieved 94.2 percent accuracy with a privacy budget of epsilon 1.0 under high heterogeneity. This is the strongest result among the federated methods and shows that adaptive weighting can reduce the negative impact of unstable client updates. When the privacy budget was reduced to epsilon 0.5, the accuracy decreased to 91.8 percent. This decline is expected because stronger privacy protection requires more noise injection. However, APP FL still outperformed FedAvg and remained higher than FedProx, which confirms that the proposed framework can maintain competitive accuracy even under stricter privacy conditions

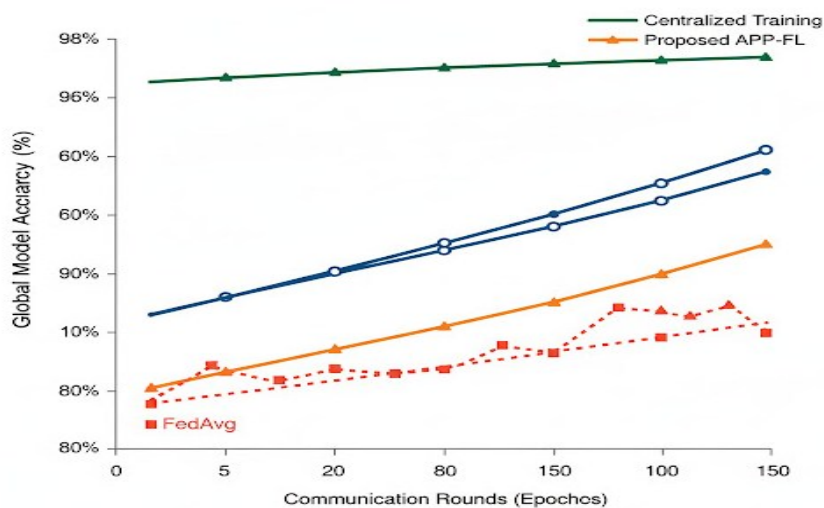


Figure 2: Global Model Accuracy Across Communication Rounds.

Figure 2 presents the global model accuracy across communication rounds. The figure is intended to show how each model improves during federated training. In this context, the most important point is not only the final accuracy, but also the stability of the learning process. FedAvg is expected to show slower or more unstable convergence because it does not adequately manage the differences in client data distribution. FedProx performs better

because it reduces local model drift, but its improvement is still limited under highly heterogeneous data.

The APP FL curve shows better convergence behavior because the adaptive weighting mechanism gives greater influence to more reliable client updates and reduces the effect of clients with higher local loss. This makes the global model more stable across communication rounds. Therefore, Figure 2 supports the claim that APP FL is not only more accurate at the final stage, but also more effective in guiding the global model toward stable convergence in non IID medical environments

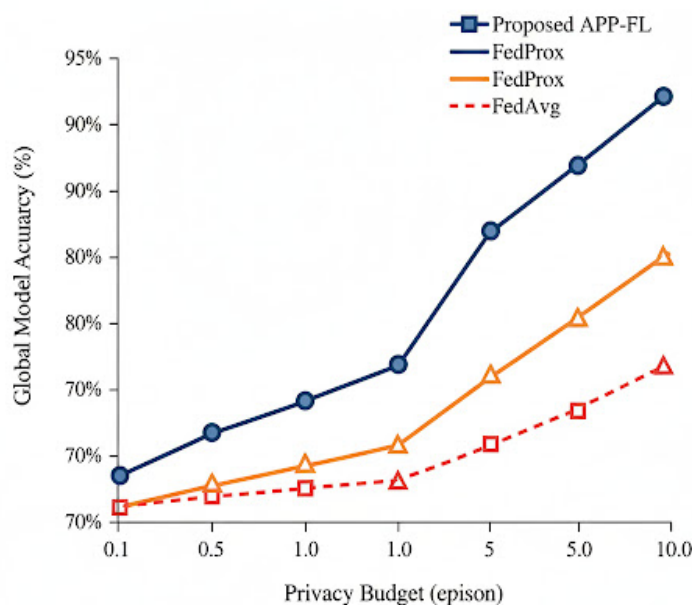


Figure 3: Accuracy vs. Privacy Budget ( ) in Non-IID Setting.

Figure 3 illustrates the relationship between model accuracy and privacy budget in the non IID setting. The figure shows the expected privacy utility trade off. When the privacy budget is higher, such as epsilon 1.0, the model achieves better accuracy because the noise added to the model updates is less disruptive. This is reflected in the APP FL accuracy of 94.2 percent. When the privacy budget becomes stricter, such as epsilon 0.5, more noise is introduced, and the accuracy decreases to 91.8 percent.

Although the accuracy drops under stronger privacy protection, the decrease is relatively controlled. This indicates that the adaptive weighting strategy helps preserve useful learning signals even when differential privacy noise is added. Therefore, Figure 3 strengthens the argument that APP FL can balance privacy protection and model performance better than standard federated approaches in heterogeneous medical image mining environments

The results show that APP-FL has the capability to address the trade-off between accuracy and privacy. In the context of Differential Privacy (DP), noise is usually injected into the gradients, which negatively impacts the performance of the models. However, our adaptive weighting strategy utilizes the gradients obtained from the more accurate nodes to reduce the side effects of DP. This is an important aspect, especially when working with medical imaging, where the distribution of the data may be quite different.

It has been noticed that the convergence rate of APP-FL outperforms that of FedAvg and FedProx algorithms, especially when large data heterogeneity is present. This can be observed from Figure 2. Figure 3 shows the accuracy preservation ability of the proposed APP-FL algorithm, especially under a tight privacy budget.

## Conclusion

In the present study, a new framework of adaptive federated learning was proposed, which was particularly tailored to ensure the effective achievement of robust image data mining in medical environments characterized by significant heterogeneity from a point of view of privacy preservation. APP-FL was shown to possess the ability to provide a secure, robust, and accurate image data mining solution.

This effectiveness of the method in striking a proper balance of diagnostic accuracy and data confidentiality requirements has been empirically validated in our assessment of the MedMNIST dataset. Possible directions for future work involve the integration of blockchain technology for auditing federated updates, as well as the use of Vision Transformers for improving the extraction of features in the medical dataset.

## References

- Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 308–318. <https://doi.org/10.1145/2976749.2978318>
- Adeniyi, A. O., Arowoogun, J. O., Okolo, C. A., Chidi, R., & Babawarun, O. (2024). Ethical considerations in healthcare IT: A review of data privacy and patient consent issues. *World Journal of Advanced Research and Reviews*, 21(2), 1660-1668.
- Ahad, A., Ahmed, K. I., Ullah, F., Sheikh, M. A., Tahir, M., Hayajneh, M., & Pires, I. M. (2026). Federated Learning and 5G/6G-Based Internet of Medical Things (IoMT): Applications, Key Enabling Technologies, Open Issues and Future Research Directions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 16(1), e70065.
- Akhtar, Z. B. (2025). Artificial intelligence within medical diagnostics: A multi-disease perspective. *Artificial intelligence in Health*, 2(3), 44-62.
- Aladraj, Y., Altajer, M., Almarhoon, A., & Sarsak, H. I. (2026). Perception of occupational therapy intervention in wheelchair seating among healthcare professionals in Bahrain. *Frontiers in Rehabilitation Sciences*, 7, 1797993. <https://doi.org/10.3389/fresc.2026.1797993>
- Arnold, T. C., Freeman, C. W., Litt, B., & Stein, J. M. (2023). Low-field MRI: clinical promise and challenges. *Journal of Magnetic Resonance Imaging*, 57(1), 25-44.
- Asif, S., Wenhui, Y., ur-Rehman, S., ul-ain, Q., Amjad, K., Yueyang, Y., ... & Awais, M. (2025). Advancements and prospects of machine learning in medical diagnostics: unveiling the future of diagnostic precision. *Archives of Computational Methods in Engineering*, 32(2), 853-883. <https://doi.org/10.1007/s11831-024-10148-w>
- Buyya, R., Ranjan, R., & Calheiros, R. N. (2010, May). Intercloud: Utility-oriented federation of cloud computing environments for scaling of application services. In *International conference on algorithms and architectures for parallel processing* (pp. 13-31). Berlin, Heidelberg: Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-13119-6\\_2](https://doi.org/10.1007/978-3-642-13119-6_2)
- Chen, C., Liu, J., Tan, H., Li, X., Wang, K. I. K., Li, P., ... & Dou, D. (2025). Trustworthy federated learning: privacy, security, and beyond. *Knowledge and Information Systems*, 67(3), 2321-2356. <https://doi.org/10.1007/s10115-024-02285-2>

- Chung, W. C., Lo, C. A., Lin, Y. H., Chen, Z. H., & Hung, C. L. (2026). Decentralized federated learning with Non-IID data: Challenges, trends, and future opportunities. *ACM Computing Surveys*, 58(8), 1-41. <https://doi.org/10.1145/3785657>
- Dongare, D. B., Nishad, S. S., Mastoli, S. Y., Saraf, S. A., Srivastava, N., & Dey, A. (2025). High-throughput sequencing: a breakthrough in molecular diagnosis for precision medicine. *Functional & Integrative Genomics*, 25(1), 22. <https://doi.org/10.1007/s10142-025-01529-w>
- Farahani, B., Tabibian, S., & Ebrahimi, H. (2023). Toward a personalized clustered federated learning: A speech recognition case study. *IEEE Internet of Things Journal*, 10(21), 18553-18562. <https://doi.org/10.1109/JIOT.2023.3292797>
- Ferreira, B., Rodrigues, J., Leitao, J., & Domingos, H. (2015, September). Privacy-preserving content-based image retrieval in the cloud. In 2015 IEEE 34th symposium on reliable distributed systems. <https://doi.org/10.1109/SRDS.2015.27>
- Geng, J., Mou, Y., Li, Q., Li, F., Beyan, O., Decker, S., & Rong, C. (2023). Improved gradient inversion attacks and defenses in federated learning. *IEEE Transactions on Big Data*, 10(6), 839-850. <https://doi.org/10.1109/TBDATA.2023.3239116>
- Govindaram, A., Prasath, J. S., Jayasakthi, K., Rajkumar, N., & Porkodi, G. (2025, January). Structured process on FL for big data analysis. In 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI) (pp. 641-647). IEEE. <https://doi.org/10.1109/ICMCSI64620.2025.10883196>
- Habibur, M., Rahman, M. A., Bakar, M. A., Mostafizur, M., Hasan, M. M., & Afroza, M. (2025). The Future of AI in Laboratory Medicine: Advancing Diagnostics, Personalization, and Healthcare Innovation. *Journal of Primeasia*, 6(1), 1-6.
- Hatamizadeh, A., Yin, H., Molchanov, P., Myronenko, A., Li, W., Dogra, P., ... & Roth, H. R. (2023). Do gradient inversion attacks make federated learning unsafe?. *IEEE Transactions on Medical Imaging*, 42(7), 2044-2056. <https://doi.org/10.1109/TMI.2023.3239391>
- Kairouz, P., et al. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, 14(1-2), 1-210. <https://doi.org/10.1561/22000000083>
- Lee, J., Solat, F., Kim, T. Y., & Poor, H. V. (2024). Federated learning-empowered mobile network management for 5G and beyond networks: From access to core. *IEEE Communications Surveys & Tutorials*, 26(3), 2176-2212. <https://doi.org/10.1109/COMST.2024.3352910>
- Liu, L., Chen, W., Nie, M., Zhang, F., Wang, Y., He, A., ... & Yan, G. (2016). iIMAGE cloud: medical image processing as a service for regional healthcare in a hybrid cloud environment. *Environmental health and preventive medicine*, 21(6), 563-571. <https://doi.org/10.1007/s12199-016-0582-7>
- Lu, Z., Pan, H., Dai, Y., Si, X., & Zhang, Y. (2024). Federated learning with non-iid data: A survey. *IEEE Internet of Things Journal*, 11(11), 19188-19209. <https://doi.org/10.1109/JIOT.2024.3376548>
- McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 1273-1282.

- Murali, S., Ding, H., Adedeji, F., Qin, C., Obungoloch, J., Asllani, I., ... & Adeleke, S. (2024). Bringing MRI to low-and middle-income countries: directions, challenges and potential solutions. *NMR in Biomedicine*, 37(7), e4992.
- Narkhede, M. R., Wankhede, N. I., & Kamble, A. M. (2025). Enhancing patient autonomy in data ownership: privacy models and consent frameworks for healthcare. *Journal of Digital Health*, 1-23. <https://doi.org/10.55976/jdh.4202513361-23>
- Nielsen, C., Wilms, M., & Forkert, N. D. (2025). A novel gradient inversion attack framework to investigate privacy vulnerabilities during retinal image-based federated learning. *Medical Image Analysis*, 103807.
- Oladejo, A. O., Adebayo, M., Olufemi, D., Kamau, E., Bobie-Ansah, D., & Williams, D. (2025). Privacy-Aware AI in cloud-telecom convergence: A federated learning framework for secure data sharing. *International Journal of Science and Research Archive*, 15(1), 005-022.
- Poojari, R. (2026). Privacy-Preserving Generative AI in Healthcare Systems Using Federated Learning Approaches. *International Journal of Data Science and IoT Management System*, 5(1), 78-88. <https://doi.org/10.64751/ijdim.2026.v5.n1.pp78-88>
- Qin, C., Murali, S., Lee, E., Supramaniam, V., Hausenloy, D. J., Obungoloch, J., ... & Adeleke, S. (2022). Sustainable low-field cardiovascular magnetic resonance in changing healthcare systems. *European Heart Journal-Cardiovascular Imaging*, 23(6), e246-e260. <https://doi.org/10.1093/ehjci/jeab286>
- Rieke, N., Hancox, J., Li, W., Liu, F., Kapsin, Y., Senaras, C., ... & Zheng, Y. (2020). The future of digital health with federated learning. *npj Digital Medicine*, 3(1), 119. <https://doi.org/10.1038/s41746-020-00323-1>
- Scheibner, J., Raisaro, J. L., Troncoso-Pastoriza, J. R., Ienca, M., Fellay, J., Vayena, E., & Hubaux, J. P. (2021). Revolutionizing medical data sharing using advanced privacy-enhancing technologies: technical, legal, and ethical synthesis. *Journal of medical Internet research*, 23(2), e25120. <https://doi.org/10.1055/a-2373-3291>
- Siddiqi, S., Qureshi, F., Lindstaedt, S., & Kern, R. (2023). Detecting outliers in non-iid data: A systematic literature review. *IEEE Access*, 11, 70333-70352.
- Tariq, F., Anjum, F., Cheng, X., Javed, S., Aurangzeb, K., & Kanwal, N. (2026). Towards a cybersecure and privacy enhanced smart grid: A blockchain enabled federated learning framework. *Plos one*, 21(3), e0342454. <https://doi.org/10.1371/journal.pone.0342454>
- Thomas, D. (2025). Breaking Data Silos in Healthcare: A Novel Framework for Standardizing and Integrating NHS Medical Data for Advanced Analytics. <https://doi.org/10.31224/5024>
- Varma, S. C. G., & Chaudhari, B. (2025). Federated Learning in Financial Data Privacy: A Secure AI Framework for Banking Applications. *International Journal of Emerging Trends in Computer Science and Information Technology*, 101-110. <https://doi.org/10.56472/ICCSAIML25-112>
- Wang, X., Xie, Y., Chen, X., Yang, J., Li, R., Gao, W., ... & Ye, Z. (2026). Securing federated learning with blockchain in the medical field: systematic literature review. *Journal of Medical Internet Research*, 28, e79052.
- Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J., & Zhang, W. (2023). A survey on federated learning: challenges and applications. *International journal of machine learning and cybernetics*, 14(2), 513-535. <https://doi.org/10.1007/s13042-022-01647-y>

- Yang, J., Sarma, A. K., & Laine, A. F. (2023). MedMNIST v2 - A large-scale lightweight benchmark for 2D and 3D biomedical image analysis. *Scientific Data*, 10(1), 41. <https://doi.org/10.1038/s41597-022-01721-8>
- Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2), 1–19. <https://doi.org/10.1145/3298981>
- Zhang, S., Li, J., Shi, L., Ding, M., Nguyen, D. C., Tan, W., ... & Han, Z. (2023). Federated learning in intelligent transportation systems: Recent applications and open problems. *IEEE Transactions on Intelligent Transportation Systems*, 25(5), 3259-3285. <https://doi.org/10.1109/TITS.2023.3324962>
- Ziegler, J., Erpenbeck, M. P., Fuchs, T., Saibold, A., Volkmer, P. C., Schmidt, G., ... & Gulden, C. (2025). Bridging data silos in oncology with modular software for federated analysis on fast healthcare interoperability resources: multisite implementation study. *Journal of Medical Internet Research*, 27, e65681. <https://doi.org/10.2196/65681>