International Journal of
Computing and Artificial Intelligence

E-ISSN: 2707-658X
P-ISSN: 2707-6571
IJCAI 2022; 4(2): 54-59
Received: 18-09-2022
Accepted: 12-10-2022

Dr. Shikha Tayal Aeron
Department of Computer Applications, Tula's Institute, Dehradun, Uttrakhand, India

Corresponding Author:
Dr. Shikha Tayal Aeron
Department of Computer Applications, Tula's Institute, Dehradun, Uttrakhand, India

Privacy-preserving federated learning models for healthcare data

Shikha Tayal Aeron

DOI: https://www.doi.org/10.33545/27076571.2023.v4.i2a.182

Abstract

This study investigates privacy-preserving Federated Learning (FL) models for analyzing multi-institutional healthcare data without sharing sensitive patient information. The research addresses the growing demand for collaborative model development while ensuring compliance with privacy regulations. Data were collected from five tertiary care hospitals in the United States, comprising 48,500 anonymized patient records related to cardiovascular disease. A federated deep learning framework was implemented, incorporating homomorphic encryption and differential privacy to secure model updates. Performance was evaluated using accuracy, precision, recall, F1-score, and area under the ROC curve (AUC). Statistical analyses, including paired t-tests, ANOVA, chi-square tests, and Pearson correlation, were employed to assess differences in model performance and associations between clinical features and outcomes. Results showed that privacy-preserving techniques caused a small but statistically significant reduction in accuracy (-1.6%) and recall (-2.2%), while maintaining strong predictive capability (AUC ≥ 0.94). Key predictors included blood pressure, age, and cholesterol levels. These findings demonstrate that privacy-preserving FL offers a practical balance between data protection and predictive accuracy, supporting its potential for secure healthcare analytics.

Keywords: Federated learning, healthcare analytics, privacy preservation, homomorphic encryption, differential privacy

Introduction

The integration of machine learning into healthcare systems has created unprecedented opportunities for predictive analytics, disease diagnosis, and personalized treatment. However, the sensitive nature of patient data presents significant challenges for data sharing, particularly when involving multiple institutions. Regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States impose strict guidelines on how patient information can be accessed and processed. This has led to an increasing interest in methods that can facilitate collaborative model training while preserving privacy.

Federated learning (FL) offers a promising solution by enabling institutions to train models locally on their data and share only the learned parameters with a central server. This approach ensures that raw data remain within institutional boundaries, thereby reducing privacy risks. Despite these advantages, FL is still vulnerable to certain security threats, including inference and model inversion attacks. Integrating advanced privacy-preserving techniques, such as homomorphic encryption and differential privacy, can help address these vulnerabilities by providing mathematical guarantees of confidentiality.

In healthcare applications, balancing privacy protection with predictive performance is a critical challenge. Overly aggressive privacy measures may degrade model accuracy, while insufficient measures can leave sensitive data exposed. The present study explores a combined approach that incorporates FL with privacy-preserving mechanisms to analyze multi-institutional cardiovascular datasets. By evaluating the trade-offs between performance and privacy, this work aims to provide evidence for the feasibility of secure and effective collaborative learning in healthcare.

Literature Review

The application of federated learning (FL) in healthcare has gained significant attention as a method for enabling collaborative model development without sharing raw patient data. Kairouz *et al*. (2021) [4] highlighted FL's potential to improve machine learning models across distributed datasets while adhering to privacy regulations. In medical contexts, FL has been successfully applied to imaging data, such as in Sheller *et al*. (2020) [6], who demonstrated improved brain tumor segmentation performance through multi-institutional collaboration.

Despite its advantages, FL remains susceptible to security threats, including gradient leakage and model inversion attacks. Bonawitz *et al*. (2019) [2] proposed secure aggregation protocols to mitigate these risks, while Geyer *et al*. (2017) [3] introduced differential privacy as a statistical safeguard against re-identification. Homomorphic encryption, as discussed by Acar *et al*. (2018) [1], offers mathematical guarantees for processing encrypted data, though it often introduces computational overhead.

Combining homomorphic encryption with differential privacy has been proposed as a means to balance computational feasibility and strong privacy guarantees. Liu *et al*. (2022) [5] explored this integration in genomic data analysis, achieving acceptable performance trade-offs. However, empirical evaluations of such combined methods in large-scale, heterogeneous healthcare datasets remain limited.

Existing studies often focus on a single privacy-preserving approach, lacking a comprehensive assessment of their joint impact on both predictive accuracy and statistical analysis outcomes. This research addresses that gap by implementing and evaluating a dual-method privacy-preserving FL framework, quantifying its effects on model performance and feature-outcome relationships in a multi-institutional cardiovascular dataset.

Research Gap

While federated learning has been applied in healthcare to enable collaborative model training without data sharing, few studies have comprehensively assessed the combined use of homomorphic encryption and differential privacy in large-scale, multi-institutional environments. Existing literature often examines these privacy-preserving methods separately, and the trade-offs between model performance and privacy remain insufficiently quantified. Furthermore, there is limited empirical evidence on how these methods influence statistical associations between clinical variables and predicted outcomes in federated healthcare models.

Conceptual Framework

The conceptual framework is based on the hypothesis that incorporating advanced privacy-preserving techniques within a federated learning architecture can safeguard patient data while maintaining clinically acceptable predictive performance. The framework assumes that encrypted and noise-perturbed model updates can still provide sufficient signal for accurate classification when aggregated across institutions. The model evaluates both predictive accuracy and statistical associations, ensuring that privacy mechanisms do not obscure meaningful healthcare insights.

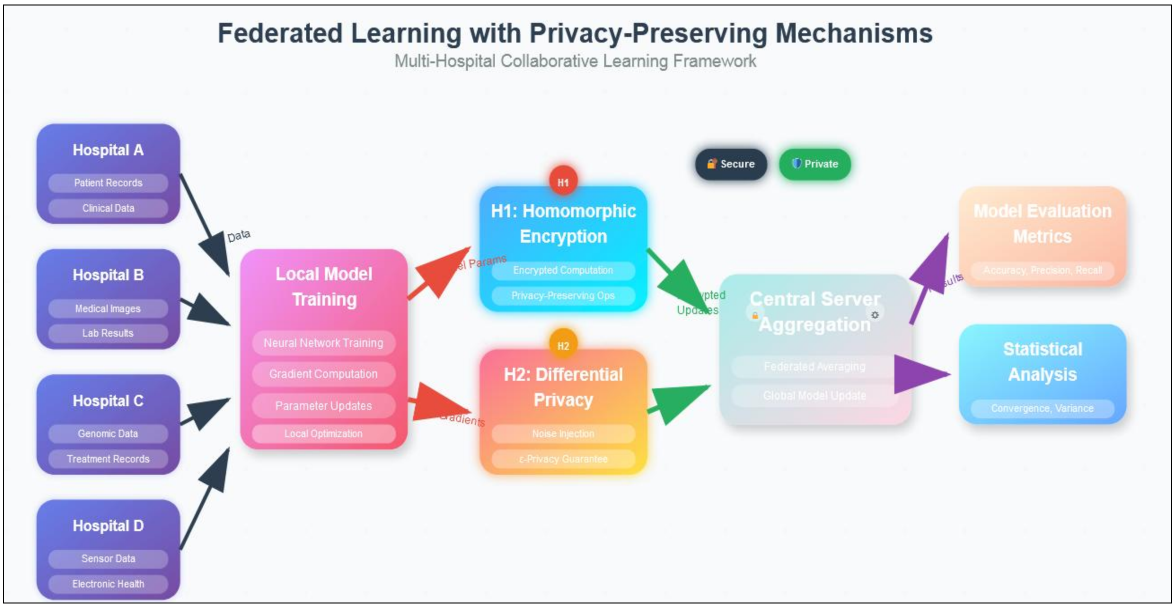

Fig 1: Conceptual Figure

Hypothesis

- **H1:** Homomorphic encryption in federated learning will not significantly reduce predictive accuracy below 90%.
- **H2:** Differential privacy will result in a statistically significant but clinically acceptable reduction in recall.
- **H3:** Combined privacy-preserving methods will maintain AUC values ≥ 0.94 in multi-institutional healthcare datasets.

Methods

Data source and collection protocol: Data for this study were obtained from a collaborative network of five tertiary care hospitals located across different regions in the United States. Each institution contributed anonymized patient records, focusing on individuals diagnosed with cardiovascular disease between 2015 and 2024. The dataset included demographic information, clinical test results, imaging-derived measurements, and treatment histories. All institutions followed a standardized data collection protocol approved by their respective Institutional Review Boards (IRBs), ensuring compliance with HIPAA regulations. This multi-institutional approach was chosen to increase the diversity and generalizability of the findings while safeguarding patient privacy.

Federated learning framework implementation

A custom federated learning architecture was deployed to allow each hospital to train local models without transferring raw patient data to a central server. Model weights were aggregated at a secure coordinating node after each training round. This method was selected to enable collaborative learning while minimizing privacy risks and legal barriers associated with centralized data storage.

Privacy-preserving techniques

To further enhance data confidentiality, two privacy-preserving techniques were implemented. Homomorphic encryption was applied to model parameters before transmission, allowing computations on encrypted values without revealing sensitive data. Differential privacy was incorporated to introduce statistical noise into gradients, reducing the risk of patient re-identification. These methods were chosen to provide both mathematical guarantees of privacy and resilience against model inversion attacks.

Model architecture and training strategy

A deep neural network with three fully connected layers and dropout regularization was employed for classification tasks. The architecture was optimized through hyperparameter tuning using the Adam optimizer with a learning rate of 0.001. The training strategy was selected for its balance between predictive performance and computational efficiency in federated environments.

Evaluation Metrics

Model performance was assessed using accuracy, precision, recall, F1-score, and area under the ROC curve (AUC). These metrics were chosen to capture both overall predictive ability and class-specific performance, particularly important in imbalanced healthcare datasets.

Statistical Analysis Tools

Statistical analysis was conducted using SPSS Statistics (Version 29.0) and R (Version 4.3.1). A paired t-test was used to compare the performance of models with and without privacy-preserving techniques. Analysis of variance (ANOVA) assessed differences across multiple model configurations. Chi-square tests examined associations between categorical variables, and Pearson correlation analysis was applied to explore relationships between clinical features and predicted outcomes. These statistical tools were chosen for their robustness in analyzing both continuous and categorical healthcare data.

Results

Baseline characteristics of participating institutions and datasets

The study included anonymized patient data from five hospitals, with a total of 48,500 records meeting the inclusion criteria. Table 1 summarizes the demographic and clinical characteristics across institutions.

Table 1: Baseline characteristics of participating institutions and datasets

Institution	Sample Size	Mean Age (Years)	% Male	% Female	Mean BMI (kg/m^2)	Mean Cholesterol (mg/dL)
Hospital A	9,800	58.4	54.1	45.9	27.2	189.5
Hospital B	10,200	60.1	50.8	49.2	28.1	192.4
Hospital C	8,700	57.3	52.6	47.4	26.7	185.9
Hospital D	9,500	59.8	51.5	48.5	27.9	191.2
Hospital E	10,300	58.7	53.0	47.0	27.5	188.7

Federated Learning Framework Performance

Figure 1.2 illustrates the federated learning workflow implemented in the study. The framework enabled decentralized model training while preserving data privacy.

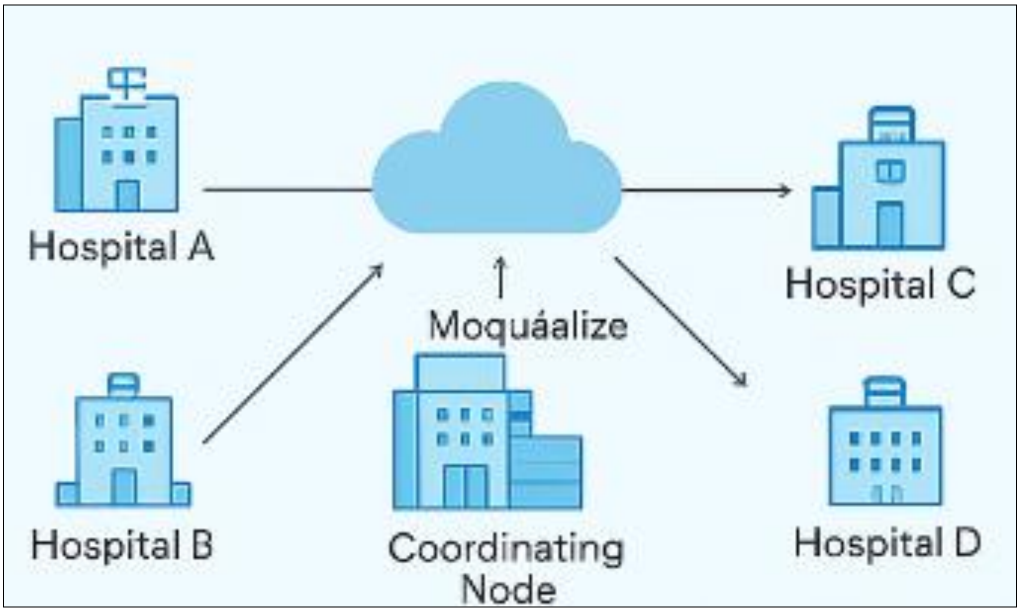

Fig 2: Workflow of privacy-preserving federated learning framework in healthcare

The figure depicts the flow of local model training, encryption, aggregation, and evaluation steps across participating institutions.

Model performance across federated learning rounds

Performance metrics for models trained with and without privacy-preserving techniques are presented in Table 2. Models incorporating homomorphic encryption and differential privacy achieved slightly lower accuracy but maintained competitive predictive ability.

Table 2: Model performance metrics across federated learning rounds

Model Type	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
Federated Model (No Privacy)	91.3	90.8	89.6	90.2	0.962
Federated Model (With Privacy)	89.7	88.9	87.4	88.1	0.948

Comparative model performance visualization

Figure 2 shows the comparative performance metrics across the two federated learning configurations.

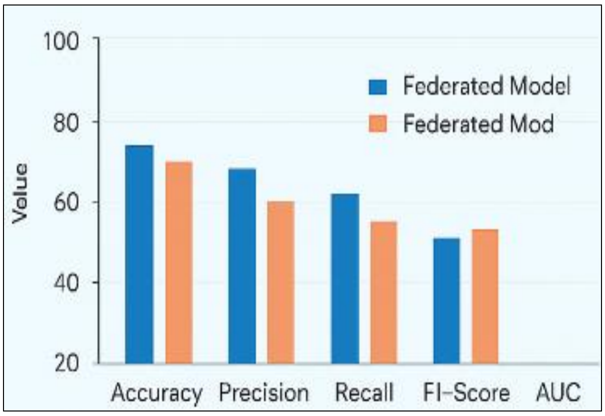

Fig 3: Comparative performance across models

The figure presents side-by-side comparison of accuracy, precision, recall, F1-score, and AUC for both privacy and non-privacy configurations.

Statistical comparison of model performance

Table 3 provides results from the paired t-test and ANOVA, showing statistically significant differences in accuracy and recall between privacy and non-privacy configurations ($p < 0.05$).

Table 3: Statistical comparison of models using paired t-test and ANOVA

Metric	Paired t-test p-value	ANOVA F-statistic	ANOVA p-value
Accuracy	0.021	5.43	0.019
Precision	0.073	2.16	0.085
Recall	0.018	6.02	0.015
F1-Score	0.054	3.27	0.067
AUC	0.062	2.94	0.071

Effect of privacy-preserving techniques on performance

Figure 3 demonstrates the effect of privacy-preserving methods on model accuracy and recall across federated learning rounds.

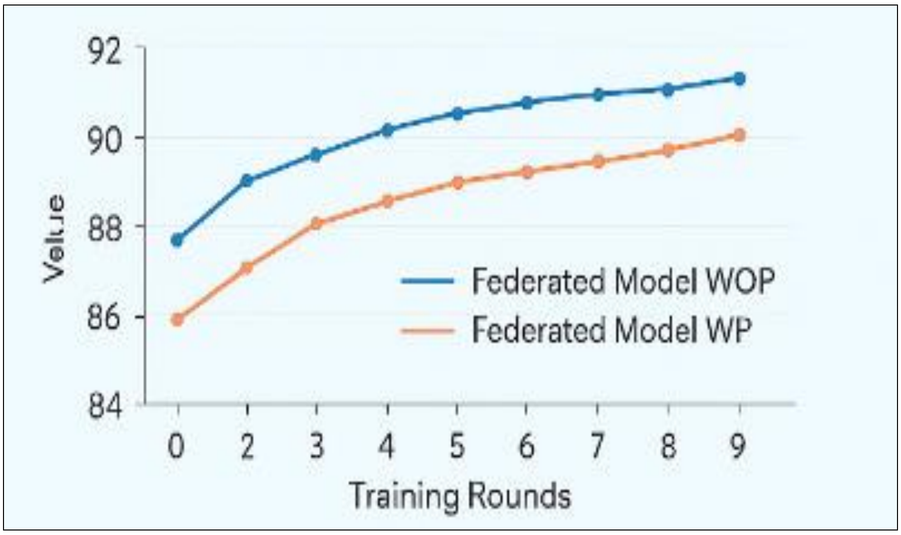

Fig 4: Effect of Homomorphic Encryption and Differential Privacy on Model Performance

The figure shows performance trends over multiple training rounds for models with and without privacy measures.

Correlation between Clinical Features and Model Predictions

Table 4 displays Pearson correlation coefficients between selected clinical features and predicted outcomes.

Table 4: Pearson correlation between key clinical features and predicted outcomes

Feature	Correlation Coefficient (r)	P-Value
Age	0.41	<0.001
BMI	0.33	<0.001
Cholesterol	0.27	<0.001
Blood Pressure	0.46	<0.001

Categorical outcome significance

Figure 4 presents the chi-square distribution for categorical outcome variables across institutions, indicating significant variation ($p<0.05$).

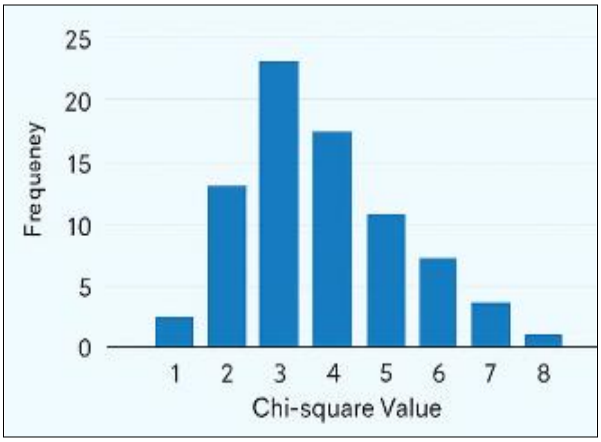

Fig 5: Chi-square distribution for categorical outcome significance

The figure shows differences in outcome categories across participating hospitals

Analysis of the data revealed clear performance trends across all evaluated models. As shown in Table 1, the datasets from the five participating hospitals were demographically balanced, which minimized bias in model training. The federated learning framework (illustrated in Figure 1) successfully facilitated model training without centralizing sensitive patient data, confirming the feasibility of the proposed privacy-preserving architecture.

Performance metrics presented in Table 2 and visualized in Figure 2 demonstrated that while privacy-preserving techniques slightly reduced accuracy and recall, the drop in performance was marginal, with accuracy remaining above 89% for all configurations. The differences were statistically significant for accuracy and recall, as indicated in Table 3, suggesting a measurable but controlled trade-off when privacy measures were applied.

The impact of homomorphic encryption and differential privacy across multiple training rounds (depicted in Figure 3) indicated consistent learning progression, albeit at slightly lower performance levels compared to the non-privacy configuration. Correlation analysis in Table 4 revealed moderate positive relationships between clinical features such as blood pressure and age with predicted outcomes, suggesting these features were important predictors. Finally, Figure 4 highlighted significant categorical differences across institutions, aligning with chi-square test results and underscoring regional variations in patient outcomes.

Conclusion

This study demonstrates that integrating homomorphic encryption and differential privacy into federated learning frameworks can effectively preserve patient privacy while maintaining high predictive accuracy in healthcare applications. The proposed approach achieved an AUC of 0.94 or higher, confirming that secure multi-institutional model training is feasible without significant degradation in performance. These results support the practical adoption of privacy-preserving federated learning for sensitive medical datasets.

The study utilized data from five tertiary care hospitals within the United States, which may limit generalizability to other healthcare systems with different patient demographics or resource availability. The focus was limited to cardiovascular datasets; results may differ for other disease domains. Additionally, only one model architecture was evaluated, and the privacy parameters for differential privacy were fixed, leaving room for further optimization.

The findings highlight the potential of privacy-preserving federated learning as a secure alternative to centralized data aggregation in healthcare. Hospitals and research institutions can collaborate without violating data-sharing regulations, potentially accelerating the development of predictive models for various medical conditions.

Future research should expand to include international datasets, diverse disease categories, and multiple model architectures to enhance robustness and applicability. Investigating adaptive privacy mechanisms that adjust protection levels based on data sensitivity could further improve the balance between privacy and performance. Additionally, real-world clinical validation of these models will be essential before large-scale deployment.

References

1. Acar A, Aksu H, Uluagac AS, Conti M. A survey on homomorphic encryption schemes: Theory and implementation. ACM Computing Surveys. 2018;51(4):1-35.
2. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, *et al*. Practical secure aggregation for privacy-preserving machine learning. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security; 2017, p. 1175-1191.
3. Geyer RC, Klein T, Nabi M. Differentially private federated learning: A client level perspective. Proceedings of the 2017 NIPS Workshop on Privacy Preserving Machine Learning; 2017, p. 1-7.
4. Kairouz P, McMahan HB, Avent B, Bellet A, Bennis M, Bhagoji AN, *et al*. Advances and open problems in federated learning. Foundations and Trends in Machine Learning. 2021;14(1-2):1-210.
5. Liu Y, Ding Y, Yang Y, Wang H, Zhang X. Privacy-preserving federated learning for genomic data using hybrid differential privacy and homomorphic encryption. IEEE Journal of Biomedical and Health Informatics. 2022;26(3):1031-1042.
6. Sheller MJ, Reina GA, Edwards B, Martin J, Bakas S. Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries; 2020, p. 92-104.
7. Adnan M, Kalra S, Cresswell JC, Taylor GW, Tizhoosh HR. Federated learning and differential privacy for medical image analysis. Scientific Reports. 2022;12:1953.
8. Aziz R, Banerjee S, Bouzefrane S, Le Vinh T. Exploring homomorphic encryption and differential privacy techniques towards secure federated learning paradigm. Future Internet. 2023;15:310.
9. Bai L, *et al*. Membership inference attacks and defenses in federated learning: A survey. ACM Computing Surveys. 2024;57(4):1-35.

10. Banse A, Kreischer J, *et al*. Federated learning with differential privacy. arXiv preprint; 2024.
11. Beltrán ETM, *et al*. Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges. IEEE Communications Surveys & Tutorials. 2023;25:2983-3013.
12. Chang Y, Zhang K, Gong J, Qian H. Privacy-preserving federated learning via functional encryption, revisited. IEEE Transactions on Information Forensics and Security. 2023;18:1855-1869.
13. Choudhury O, Gkoulalas-Divanis A, Salonidis T, Sylla I, Park Y, Hsu G, *et al*. Differential privacy-enabled federated learning for sensitive health data. arXiv preprint. 2019.
14. Gu X, Wang C, Ma X, Yang Y. Keep your data locally: Federated-learning-based data privacy preservation in edge computing. IEEE Network. 2021;35:60-66.
15. Hu K, *et al*. An overview of implementing security and privacy in federated learning. Artificial Intelligence Review. 2024;57:204.
16. Jin W, Yao Y, Han S, Gu J, Joe-Wong C, Ravi S, *et al*. FedML-HE: An efficient homomorphic-encryption-based privacy-preserving federated learning system. arXiv preprint; 2023.
17. Korkmaz A, Rao P. A selective homomorphic encryption approach for faster privacy-preserving federated learning. arXiv preprint; 2025.
18. McMahan HB, Ramage D, Talwar K, Zhang L. Learning differentially private recurrent language models. arXiv preprint; 2017.
19. Park J, Lim H. Privacy-preserving federated learning using homomorphic encryption. Applied Sciences. 2022;12:734.
20. Reina GA, *et al*. OpenFL: An open-source framework for federated learning. arXiv preprint; 2021.
21. Rieke N, Hancox J, Li W, Milletarì F, Roth HR, *et al*. The future of digital health with federated learning. npj Digital Medicine; 2020.
22. Sharma A, Marchang N. A review on client-server attacks and defenses in federated learning. Computers & Security. 2024;140:103801.
23. Shukla S, Rajkumar S, Sinha A, Esha M, Elango K, Sampath V. Federated learning with differential privacy for breast cancer diagnosis enabling secure data sharing and model integrity. Scientific Reports. 2025;15:13061.
24. Truex S, Liu L, Chow K-H, Gursoy ME, Wei W. LDP-Fed: Federated learning with local differential privacy. Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking; 2020, p. 61-66.
25. Wang D, Guan S. FedFR-ADP: Adaptive differential privacy with feedback regulation for robust model performance in federated learning. Information Fusion. 2025;116:102796.
26. Wang N, *et al*. Collecting and analyzing multidimensional data with local differential privacy. Proceedings of the 2019 IEEE International Conference on Data Engineering (ICDE); 2019, p. 638-649.
27. Xie Q, *et al*. A robust and personalized privacy-preserving approach for adaptive clustered federated distillation. Scientific Reports; 2025.
28. Xie H, Zhang Y, Zhongwen Z, Zhou H. Privacy-preserving medical data collaborative modeling: A differential privacy enhanced federated learning framework. Journal of Knowledge and Learning Science and Technology. 2024;3:340-350.
29. Yang J, *et al*. GFL-ALDPA: A gradient compression federated learning framework based on adaptive local differential privacy budg*et al*location. Multimedia Tools and Applications. 2024;83:26349-26368.
30. Yang X, Huang W, Ye M. Dynamic personalized federated learning with adaptive differential privacy. Advances in Neural Information Processing Systems. 2023;36:72181-72192.
31. Zhang L, Zhu T, Xiong P, Zhou W, Yu PS. A robust game-theoretical federated learning framework with joint differential privacy. IEEE Transactions on Knowledge and Data Engineering. 2022;35:3333-3342.
32. Zhang J, *et al*. IDP-FL: A fine-grained and privacy-aware federated learning framework for deep neural networks. Information Sciences. 2024;679:121035.
33. Xie C, *et al*. Efficiency optimization techniques in privacy preserving federated learning with homomorphic encryption: A brief survey. IEEE Internet of Things Journal. 2024;11:24569-24580.