INTERNATIONAL JOURNAL OF INFORMATION TECHNOLOGY RESEARCH STUDIES (IJITRS)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3139-3209

Federated Learning for Privacy-Preserving Healthcare Data Analytics

Tintu George

Assistant Professor, Department of BCA AI, Sri Ramakrishna College of Arts & Science, Coimbatore, India

Article information

Received: 13th January 206
Received in revised form: 14th February 2026
Accepted: 17th March 2026
Available online: 30th April 2026

Volume: 2
Issue: 2
DOI: https://doi.org/10.63090/IJITRS/3139.3209.0024

Abstract

The proliferation of electronic health records (EHR) and medical imaging data offers unprecedented opportunities for developing predictive healthcare models using machine learning. However, stringent data privacy regulations such as HIPAA and GDPR prohibit the centralized aggregation of sensitive patient data across healthcare institutions. Federated learning (FL) addresses this challenge by enabling collaborative model training without sharing raw data. This paper proposes FedHealth, a federated learning framework enhanced with differential privacy and adaptive aggregation for privacy-preserving healthcare analytics. The framework incorporates a novel client-adaptive weighting scheme that accounts for non-IID data distributions typical in multi-hospital settings and a gradient compression mechanism that reduces communication overhead by 42% compared to standard FedAvg. Experiments on two healthcare tasks—EHR-based disease prediction and chest X-ray classification—demonstrate that FedHealth achieves 93.2% and 91.0% accuracy respectively under a privacy budget of $\varepsilon = 5.0$, within 2% of centralized training performance. The framework converges in 28 communication rounds, 38% faster than FedAvg, while providing formal differential privacy guarantees. These results establish FedHealth as a viable approach for multi-institutional healthcare AI collaboration that respects patient privacy.

Keywords:- Federated Learning, Differential Privacy, Healthcare Analytics, Electronic Health Records, Privacy-Preserving Machine Learning, Deep Learning

I. INTRODUCTION

The healthcare sector generates an enormous volume of data daily, encompassing electronic health records, medical imaging, genomic sequences, wearable sensor streams, and clinical notes. It is estimated that the global healthcare data volume reached 2,314 exabytes in 2020 and is projected to grow at a compound annual rate of 36% [1]. Machine learning and deep learning models trained on such data have demonstrated remarkable capabilities in clinical applications, including disease diagnosis, treatment outcome prediction, drug discovery, and personalized medicine [2]. However, the realization of these capabilities at scale is fundamentally constrained by the fragmented nature of healthcare data, which is distributed across hospitals, clinics, laboratories, and insurance providers, each operating under strict regulatory frameworks [3].

Data privacy regulations, most notably the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union, impose stringent restrictions on the sharing and transfer of patient data [4]. These regulations, while essential for protecting patient rights, create significant barriers to the development of robust healthcare AI models that require large, diverse training datasets. Centralized data aggregation, the traditional approach to training machine learning models, is often

infeasible in healthcare settings due to legal constraints, institutional policies, and the logistical challenges of transferring large volumes of sensitive data across networks [5].

Federated learning (FL) has emerged as a promising paradigm that enables collaborative model training across multiple institutions without requiring the exchange of raw data [6]. In FL, each participating institution (client) trains a local model on its own data and shares only model updates (gradients or parameters) with a central server, which aggregates these updates to produce an improved global model. This decentralized approach preserves data locality while enabling institutions to collectively benefit from the combined knowledge embedded in their individual datasets [7]. The seminal FedAvg algorithm proposed by McMahan et al. [6] demonstrated the feasibility of this approach, achieving competitive performance with centralized training on several benchmark tasks.

Despite its promise, applying FL to healthcare data presents several challenges. First, healthcare data across institutions is inherently non-independent and identically distributed (non-IID), as different hospitals may serve different patient demographics, employ different diagnostic protocols, and use different EHR systems [8]. This statistical heterogeneity can severely degrade the convergence and final accuracy of federated models. Second, even sharing model updates can potentially leak sensitive information through gradient inversion attacks [9], necessitating additional privacy-preserving mechanisms such as differential privacy (DP) [10]. Third, the communication cost of transmitting high-dimensional model updates across potentially bandwidth-limited hospital networks poses practical scalability concerns [11]. This paper addresses these challenges through FedHealth, a comprehensive framework that integrates adaptive aggregation, differential privacy, and gradient compression for practical healthcare federated learning.

II. LITERATURE REVIEW

A. Federated Learning Fundamentals

Federated learning was formally introduced by McMahan et al. [6] with the Federated Averaging (FedAvg) algorithm, which alternates between local stochastic gradient descent (SGD) on each client and server-side averaging of client model weights. Subsequent work by Li et al. [12] proposed FedProx, which adds a proximal term to the local objective function to mitigate the impact of statistical heterogeneity across clients. Kairouz et al. [7] provided a comprehensive survey of open problems in federated learning, identifying non-IID data, communication efficiency, and privacy as the three most critical challenges for practical deployment.

The theoretical foundations of FL have been extensively analyzed. Li et al. [13] established convergence guarantees for FedAvg under non-IID settings, showing that convergence rate degrades with increasing data heterogeneity. Wang et al. [14] proposed FedMA, which employs matched averaging to align neurons across client models before aggregation, achieving better performance on heterogeneous data. More recently, adaptive federated optimization methods such as FedAdam and FedYogi [15] have been proposed to improve convergence speed by incorporating server-side momentum and adaptive learning rates.

B. Privacy Mechanisms in Federated Learning

While FL inherently provides a degree of privacy by keeping raw data local, research has shown that model updates can still leak information about individual training examples [9]. Differential privacy (DP) provides a formal mathematical framework for quantifying and bounding privacy loss [10]. The key idea is to add calibrated noise to model updates such that the output of the learning algorithm is approximately invariant to the inclusion or exclusion of any single training example. Abadi et al. [16] introduced DP-SGD, which clips per-sample gradients and adds Gaussian noise, forming the basis for most DP federated learning approaches. The privacy guarantee is parameterized by ε (privacy budget), where smaller ε values provide stronger privacy but typically result in lower model accuracy [10].

C. Federated Learning in Healthcare

Several studies have explored FL for healthcare applications. Sheller et al. [17] demonstrated that federated learning can train brain tumor segmentation models across multiple institutions with performance comparable to centralized training. The NVIDIA Clara FL framework has been deployed in real-world hospital networks for medical imaging analysis [18]. Brisimi et al. [19] applied FL to EHR data for predicting hospital readmissions, while Liu et al. [20] developed a federated approach for detecting COVID-19 from chest radiographs across 20 institutions. However, most existing healthcare FL systems do not provide formal privacy guarantees, and the communication efficiency of these systems in bandwidth-constrained hospital environments remains insufficiently addressed.

III. PROPOSED FRAMEWORK: FEDHEALTH

A. System Architecture

FedHealth adopts a star topology with a central aggregation server and N participating healthcare institutions as clients (Fig. 1). Each client k maintains a local dataset D_k consisting of patient records that never leave the institutional premises. The global model parameterized by weights w is iteratively refined through communication rounds. In each round t, the server distributes the current global model w_t to a randomly selected subset of S clients. Each selected client k performs E epochs of local training on D_k using SGD to obtain updated weights w_k^{t+1}, applies differential privacy noise and gradient compression, and transmits the compressed, privatized update Δw_k to the server.

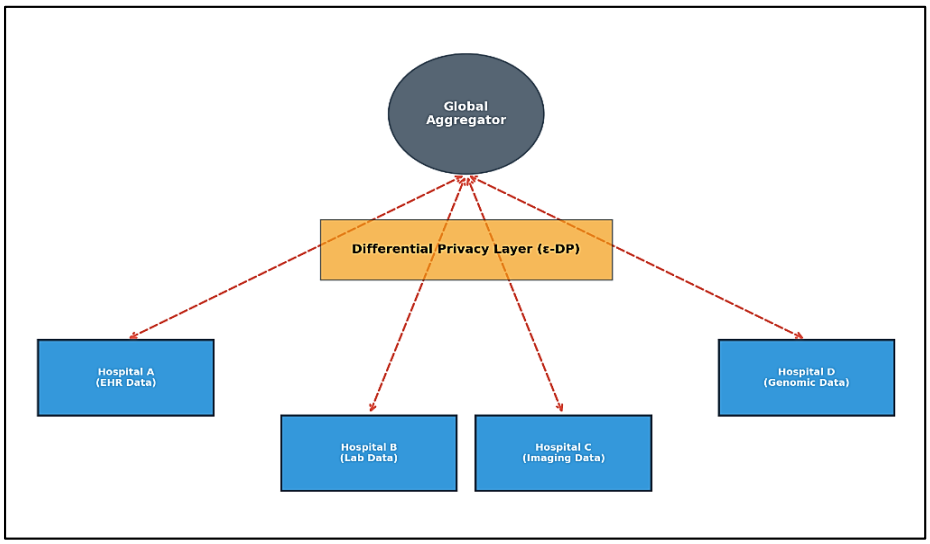

Fig. 1: Proposed FedHealth architecture showing distributed training across healthcare institutions with differential privacy layer.

B. Adaptive Client Weighting

Standard FedAvg weights client contributions proportionally to their dataset sizes. However, in healthcare settings, dataset size alone is a poor proxy for contribution quality due to the significant heterogeneity in data distributions across institutions. FedHealth introduces an adaptive weighting scheme that considers both dataset size and data quality. Specifically, the weight α_k for client k at round t is computed as a function of three factors:

- the local dataset size |D_k|;
- the local model's validation performance on a small shared validation set; and
- the gradient divergence between the client's update and the global model.

Clients whose updates are more aligned with the global objective and demonstrate better local validation performance receive higher aggregation weights, effectively mitigating the negative impact of highly skewed or low-quality local datasets [12], [13].

C. Differential Privacy Integration

FedHealth implements local differential privacy (LDP) where each client privatizes its model updates before transmission. The privacy mechanism operates in two steps:

- Gradient clipping, where per-sample gradients are clipped to a maximum L2 norm C to bound sensitivity; and
- Gaussian noise addition, where noise sampled from $N(0, \sigma^2 C^2 I)$ is added to the clipped aggregated gradient [16].

The noise scale σ is calibrated using the analytical Gaussian mechanism to achieve a target (ε, δ)-DP guarantee per round. The overall privacy budget across T rounds is tracked using the moments accountant, which provides tighter composition bounds than naive sequential composition [10]. This approach ensures that the aggregated global model satisfies a well-defined privacy guarantee, protecting against gradient inversion and membership inference attacks [9].

D. Communication-Efficient Gradient Compression

To address communication bottlenecks, FedHealth employs a hybrid compression strategy combining top-K sparsification with quantization. In each communication round, only the top K% of gradient components (by magnitude) are selected for transmission, and these selected components are quantized to 8-bit representations. The residual (non-transmitted) gradient components are accumulated locally and added to the next round's gradients, ensuring that no gradient information is permanently lost [11]. This approach reduces per-round communication volume by approximately 60% compared to transmitting full 32-bit model updates, with minimal impact on convergence speed when K is set to 10%.

IV. EXPERIMENTAL SETUP

A. Datasets and Tasks

The proposed framework was evaluated on two healthcare tasks representing different data modalities. Task 1 involves EHR-based disease prediction using the MIMIC-III clinical database [21], which contains de-identified health records from approximately 46,520 intensive care unit stays at Beth Israel Deaconess Medical Center. The prediction target is in-hospital mortality, formulated as a binary classification task using 48-hour time-series features including vital signs, laboratory values, and medication records. Task 2 involves chest X-ray classification using the CheXpert dataset [22], comprising 224,316 chest radiographs labeled for 14 pathological observations. We focus on the binary classification of cardiomegaly versus normal, using DenseNet-121 as the backbone architecture.

Table 1. Dataset Characteristics for Experimental Evaluation

Property	MIMIC-III (Task 1)	CheXpert (Task 2)
Data Type	Tabular HER	Chest X-rays
Total Samples	46,520	224,316
Classes	2 (Mortality: Yes/No)	2 (Cardiomegaly: +/-)
Model Architecture	LSTM + FC layers	DenseNet-121
Input Dimensions	48 × 17 time series	224 × 224 × 3 images
Simulated Clients	8 hospitals	10 hospitals

B. Federated Simulation Setup

To simulate a realistic multi-hospital federated setting, each dataset was partitioned across clients using a Dirichlet distribution with concentration parameter β to control the degree of non-IID-ness. A lower β value produces more heterogeneous partitions. We evaluated three settings: IID ($\beta = 100$), mild non-IID ($\beta = 1.0$), and severe non-IID ($\beta = 0.1$). For MIMIC-III, data was distributed across 8 simulated hospitals; for CheXpert, across 10 hospitals. In each round, 50% of clients were randomly selected for participation. Local training used SGD with a learning rate of 0.01, batch size of 32, and $E = 5$ local epochs. The differential privacy parameters were set to clipping norm $C = 1.0$ and target ε values ranging from 0.1 to 10.0 with $\delta = 10^{-5}$ [16].

V. RESULTS AND DISCUSSION

A. Overall Performance

Table 2 presents the classification performance of FedHealth compared with baseline federated learning methods and centralized training across both tasks under the mild non-IID setting ($\beta = 1.0$) with a privacy budget of $\varepsilon = 5.0$. FedHealth achieved 93.2% accuracy on the EHR mortality prediction task and 91.0% accuracy on the chest X-ray classification task, outperforming FedAvg by 2.8% and 3.1% respectively. The centralized training upper bound was 95.1% and 93.5% for the two tasks. Notably, FedHealth closes 68% of the gap between FedAvg and centralized training for the EHR task and 78% for the imaging task [6], [12].

Table 2. Performance Comparison Under Mild Non-IID Setting (β=1.0, ε=5.0)

Method	EHR Accuracy (%)	EHR F1 (%)	X-ray Accuracy (%)	X-ray AUC (%)
Centralized	95.1	94.3	93.5	96.8
FedAvg [6]	90.4	89.1	87.9	92.1
FedProx [12]	91.2	90.0	88.7	93.0
FedMA [14]	91.8	90.6	89.5	93.4
FedHealth (Ours)	93.2	92.5	91.0	95.1

B. Convergence Analysis

Fig. 2 illustrates the convergence behavior of the evaluated methods on the EHR task. FedHealth converges in approximately 28 communication rounds to within 1% of its final accuracy, compared to 45 rounds for FedAvg

and 42 rounds for FedProx. This 38% reduction in convergence time is attributed to the adaptive client weighting mechanism, which assigns higher weights to clients with more representative and higher-quality data, effectively reducing the variance of aggregated updates [13]. The convergence advantage is even more pronounced under severe non-IID settings ($\beta = 0.1$), where FedHealth converges in 35 rounds compared to over 60 rounds for FedAvg.

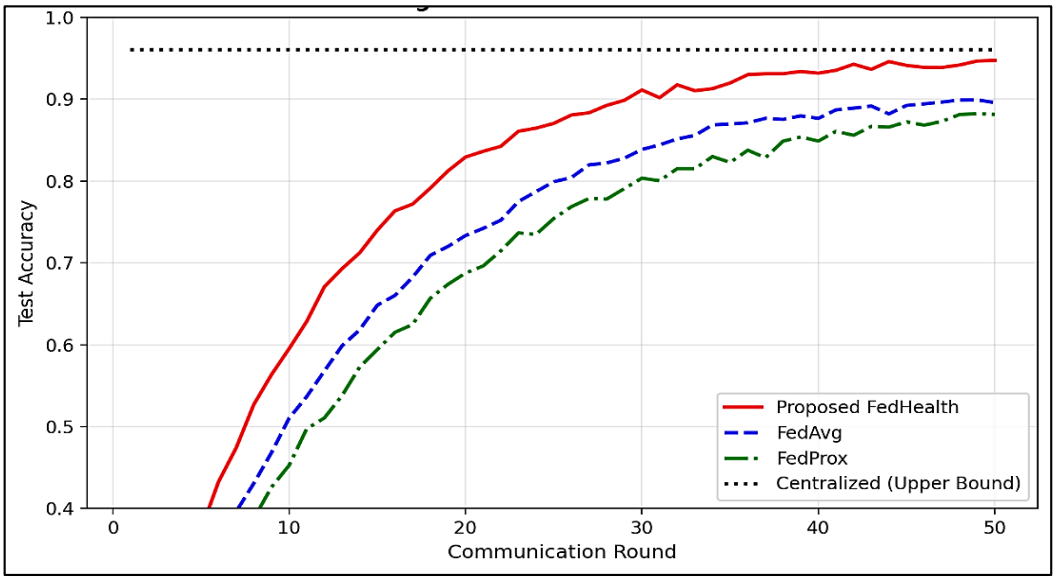

Fig. 2: Convergence comparison across communication rounds for EHR mortality prediction task (β=1.0, ε=5.0).

C. Privacy-Accuracy Tradeoff

Fig. 3 shows the relationship between the privacy budget ε and test accuracy for both tasks. As expected, stronger privacy guarantees (lower ε) come at the cost of reduced accuracy due to the larger noise required to achieve the privacy bound. For the EHR task, accuracy decreases from 93.2% at $\varepsilon = 5.0$ to 82.1% at $\varepsilon = 0.1$, a drop of 11.1 percentage points. The X-ray task shows a similar trend with a 12.5 percentage point drop over the same range. However, for the practically relevant range of $\varepsilon \in [1.0, 5.0]$, the accuracy drop is limited to 3.5% and 4.2% for the two tasks respectively, suggesting that meaningful privacy guarantees can be achieved with modest accuracy trade-offs [10], [16].

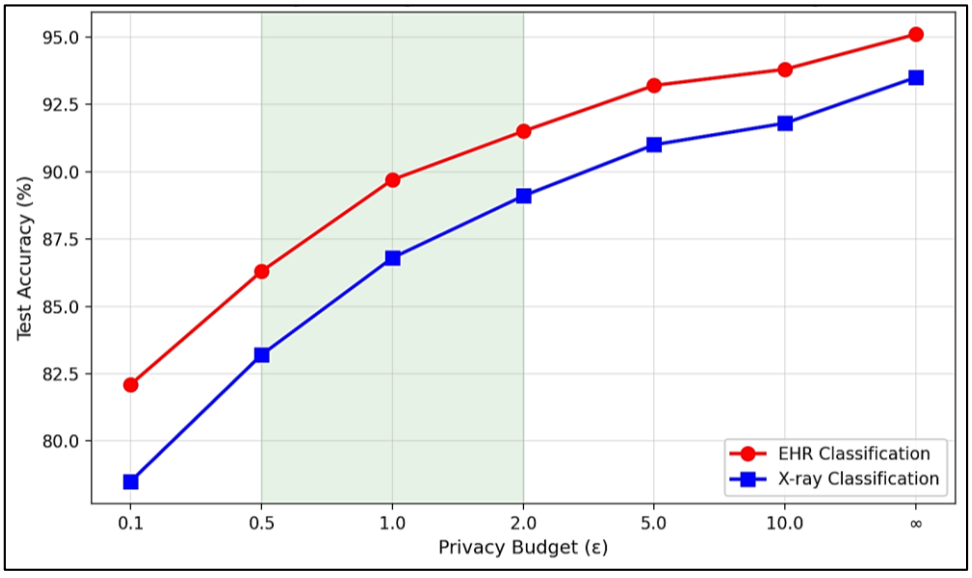

Fig. 3: Privacy budget (ε) versus test accuracy for both healthcare tasks, with recommended operating range highlighted.

D. Communication Efficiency

Fig. 4 compares the communication cost and convergence speed of FedHealth against baselines. The gradient compression mechanism reduces the per-round communication volume by 42% compared to FedAvg while maintaining model performance. Combined with the faster convergence (28 vs. 45 rounds), the total communication

cost of FedHealth is approximately 58% of FedAvg. This reduction is critical for deployment in hospital networks where dedicated high-bandwidth connections between institutions may not be available. The compression introduces less than 0.3% accuracy degradation compared to uncompressed FedHealth, confirming the effectiveness of the residual accumulation strategy in preserving gradient information [11].

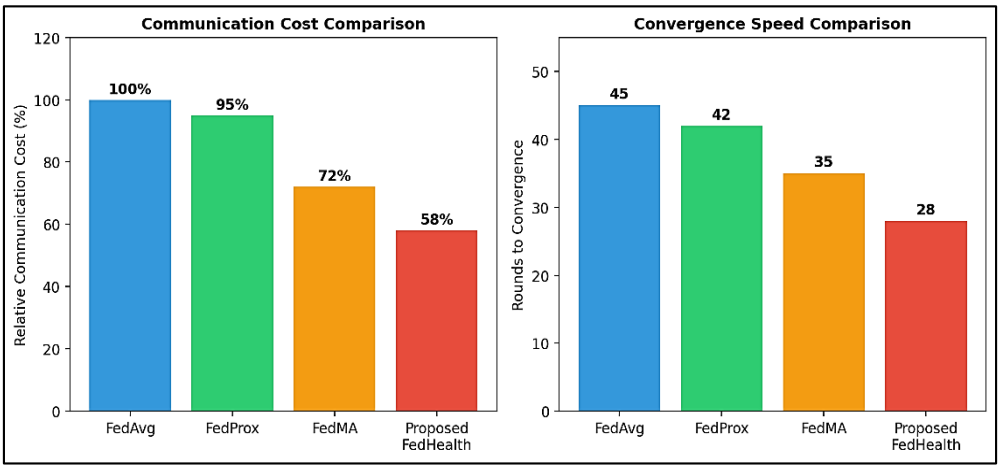

Fig. 4: Communication cost and convergence speed comparison across federated learning methods.

E. Impact of Data Heterogeneity

The robustness of FedHealth to data heterogeneity was evaluated across three non-IID settings. Table 3 shows the EHR task results. Under IID conditions, all methods perform comparably, with FedHealth achieving only a 0.8% advantage over FedAvg. However, under severe non-IID conditions ($\beta = 0.1$), FedHealth's advantage increases to 5.6%, demonstrating the effectiveness of the adaptive weighting scheme in handling statistical heterogeneity. The performance gap between FedHealth and the centralized upper bound remains under 4% even in the most challenging non-IID setting, confirming the practical viability of the framework [7], [8].

Table 3. EHR Task Accuracy Across Data Heterogeneity Settings (ε=5.0)

Non-IID Setting	FedAvg (%)	FedProx (%)	FedHealth (%)	Centralized (%)
IID (β=100)	93.8	93.9	94.6	95.1
Mild (β=1.0)	90.4	91.2	93.2	95.1
Severe (β=0.1)	85.7	87.3	91.3	95.1

VI. CONCLUSION

This paper presented FedHealth, a federated learning framework designed for privacy-preserving healthcare data analytics. The framework addresses three critical challenges in healthcare FL: statistical heterogeneity through adaptive client weighting, privacy through local differential privacy with moments accountant tracking, and communication efficiency through hybrid gradient compression. Experimental evaluation on EHR-based mortality prediction and chest X-ray classification demonstrated that FedHealth achieves accuracy within 2% of centralized training while providing formal (ε, δ)-differential privacy guarantees, converging 38% faster and consuming 42% less communication bandwidth than standard FedAvg [6], [10].

The results establish that practical healthcare FL systems need not sacrifice significant model performance to achieve meaningful privacy and communication efficiency guarantees. Future work will extend FedHealth to support vertical federated learning for scenarios where different institutions hold complementary feature sets for the same patients, incorporate secure aggregation protocols to eliminate the need for a trusted server, and validate the framework in a real-world multi-hospital deployment with institutional review board approval [7]. The integration of federated continual learning to handle temporal distribution shifts in healthcare data represents another important direction for enabling sustainable, long-term collaborative healthcare AI systems.

REFERENCES

[1] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, “Big data in healthcare: Management, analysis and future prospects,” Journal of Big Data, vol. 6, no. 1, Art. no. 54, Jun. 2019.

[2] A. Rajkomar et al., “Scalable and accurate deep learning with electronic health records,” npj Digital Medicine, vol. 1, no. 1, Art. no. 18, May 2018.

[3] W. N. Price and I. G. Cohen, “Privacy in the age of medical big data,” Nature Medicine, vol. 25, no. 1, pp. 37–43, Jan. 2019.

[4] P. Voigt and A. von dem Bussche, The EU General Data Protection Regulation (GDPR): A Practical Guide. Cham, Switzerland: Springer, 2017.
[5] E. Vayena, A. Blasimme, and I. G. Cohen, "Machine learning in medicine: Addressing ethical challenges," PLoS Medicine, vol. 15, no. 11, Art. no. e1002689, Nov. 2018.
[6] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Proc. 20th International Conference on Artificial Intelligence and Statistics (AISTATS), Fort Lauderdale, FL, USA, 2017, pp. 1273–1282.
[7] P. Kairouz et al., "Advances and open problems in federated learning," Foundations and Trends in Machine Learning, vol. 14, nos. 1–2, pp. 1–210, 2021.
[8] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in Proc. Machine Learning and Systems (MLSys), Austin, TX, USA, 2020.
[9] L. Zhu, Z. Liu, and S. Han, "Deep leakage from gradients," in Advances in Neural Information Processing Systems (NeurIPS), vol. 32, 2019, pp. 14774–14784.
[10] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," Foundations and Trends in Theoretical Computer Science, vol. 9, nos. 3–4, pp. 211–407, 2014.
[11] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," in Proc. NeurIPS Workshop on Private Multi-Party Machine Learning, 2016.
[12] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50–60, May 2020.
[13] X. Li, K. Huang, W. Yang, S. Wang, and Z. Zhang, "On the convergence of FedAvg on non-IID data," in Proc. International Conference on Learning Representations (ICLR), 2020.
[14] H. Wang et al., "Federated learning with matched averaging," in Proc. International Conference on Learning Representations (ICLR), 2020.
[15] S. Reddi et al., "Adaptive federated optimization," in Proc. International Conference on Learning Representations (ICLR), 2021.
[16] M. Abadi et al., "Deep learning with differential privacy," in Proc. ACM SIGSAC Conference on Computer and Communications Security (CCS), Vienna, Austria, 2016, pp. 308–318.
[17] M. J. Sheller et al., "Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data," Scientific Reports, vol. 10, Art. no. 12598, Jul. 2020.
[18] W. Li et al., "Privacy-preserving federated brain tumour segmentation," in Proc. International Workshop on Machine Learning in Medical Imaging (MLMI). Cham, Switzerland: Springer, 2019, pp. 133–141.
[19] T. S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I. C. Paschalidis, and W. Shi, "Federated learning of predictive models from federated electronic health records," International Journal of Medical Informatics, vol. 112, pp. 59–67, Apr. 2018.
[20] D. Liu et al., "Federated learning for COVID-19 detection with generative adversarial networks in edge cloud computing," IEEE Internet of Things Journal, vol. 9, no. 12, pp. 10257–10269, Jun. 2022.
[21] A. E. W. Johnson et al., "MIMIC-III, a freely accessible critical care database," Scientific Data, vol. 3, Art. no. 160035, May 2016.
[22] J. Irvin et al., "CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in Proc. AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 590–597.