International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(3s) 2026 pp. 1088-1102
https://doi.org/10.70917/ijcisim-2026-2444
Article

A Unified Framework for Communication-Efficient and Privacy-Preserving Federated Learning Using Adaptive Differential Privacy and Sparse Model Aggregation

Nithya Niranjana Murthy[1], Manjula S. H.[2]

[1]Department of Computer Science and Engineering, University Visvesvaraya College of Engineering (UVCE), Bangalore University, Bengaluru, Karnataka, India.
Email: nithya.semantic@gmail.com
[2]Department of Computer Science and Engineering, University Visvesvaraya College of Engineering (UVCE), Bangalore University, Bengaluru, Karnataka, India.
Email: shmanjula@gmail.com
Corresponding Author: Nithya Niranjana Murthy, nithya.semantic@gmail.com

Abstract: Federated Learning (FL) has emerged as a promising distributed learning paradigm that enables collaborative model training without sharing raw data. However, conventional FL frameworks suffer from significant communication overhead and privacy leakage through exchanged model updates, particularly in heterogeneous client environments. To address these challenges, this paper proposes ADP-SFed, a unified communication-efficient and privacy-preserving federated learning framework that integrates Adaptive Differential Privacy (ADP) with Sparse Model Aggregation (SMA). The proposed framework dynamically adjusts the privacy budget across communication rounds to achieve an effective privacy–utility trade-off while transmitting only the most significant model parameters to reduce communication cost. Experiments were conducted on the MedMNIST and CIFAR-10 benchmark datasets under federated learning settings and compared with state-of-the-art methods, including FedAvg, FedProx, DP-FedAvg, Sparse-FedAvg, and FedAdam. Experimental results demonstrate that ADP-SFed achieved the highest classification accuracies of 95.82% on MedMNIST and 91.43% on CIFAR-10, with corresponding AUC values of 0.987 and 0.973, respectively. Furthermore, the proposed framework reduced communication cost by 67.86%, converged within 65 communication rounds, and achieved the lowest training time (128 min) and inference time (14.1 ms) among the compared methods. The adaptive privacy scheduling mechanism effectively preserved data privacy while maintaining high predictive performance, and sparse aggregation significantly improved communication efficiency. These results demonstrate that ADP-SFed provides a scalable, efficient, and privacy-preserving federated learning solution suitable for distributed medical image analysis, computer vision, and resource-constrained edge computing applications.

Keywords: Federated Learning, Adaptive Differential Privacy, Sparse Model Aggregation, Communication Efficiency, Privacy Preservation, MedMNIST, CIFAR-10, Medical Image Classification, Edge Intelligence.

1. Introduction

The rapid growth of artificial intelligence (AI), Internet of Things (IoT), edge computing, and smart healthcare systems has resulted in an unprecedented increase in the volume of distributed data generated by mobile devices, wearable sensors, autonomous systems, and medical institutions. Deep learning models have demonstrated remarkable performance in image classification, medical diagnosis, natural language processing, and intelligent decision-making. However, conventional centralized machine learning requires all training data to be collected and stored on a central server, which raises serious concerns regarding data privacy, security, communication overhead, and regulatory compliance. In domains such as healthcare and finance, where sensitive information cannot be freely exchanged,

centralized learning becomes impractical due to legal and ethical restrictions [1]. Federated Learning (FL) has emerged as a promising distributed learning paradigm that enables multiple clients to collaboratively train a global machine learning model without sharing their raw data. Instead of transmitting sensitive datasets, each client performs local model training and shares only model parameters or gradients with a central server for aggregation. This decentralized training strategy significantly reduces privacy risks while enabling collaborative intelligence across geographically distributed devices. Consequently, federated learning has attracted considerable attention in applications such as medical image analysis, smart healthcare, intelligent transportation systems, industrial IoT, mobile edge computing, and autonomous vehicles [2].

Despite its advantages, conventional federated learning still suffers from several critical challenges. One of the major limitations is the substantial communication overhead caused by the repeated transmission of large neural network parameters between clients and the server during every communication round [3]. As the number of participating clients and model complexity increase, communication latency becomes a significant bottleneck, particularly in bandwidth-constrained edge environments. Furthermore, although raw data remain local, transmitted model updates can still reveal sensitive information through gradient inversion and membership inference attacks, making privacy leakage an important concern. Existing differential privacy-based federated learning methods generally employ a fixed privacy budget, which often introduces excessive noise throughout training and leads to reduced classification accuracy and slower model convergence. Several communication-efficient federated learning approaches have been proposed, including model compression, parameter quantization, and sparse gradient transmission. Similarly, privacy-preserving techniques based on differential privacy and secure aggregation have been extensively investigated [4]. However, most existing studies optimize either communication efficiency or privacy protection independently, without jointly addressing both challenges within a unified framework. Moreover, fixed privacy mechanisms fail to adapt to different stages of model optimization, resulting in an unfavorable trade-off between privacy preservation and model utility.

To overcome these limitations, this paper proposes ADP-SFed (Adaptive Differential Privacy-based Sparse Federated Learning), a unified communication-efficient and privacy-preserving federated learning framework. The proposed framework integrates Adaptive Differential Privacy (ADP) with Sparse Model Aggregation (SMA) to simultaneously reduce communication overhead and strengthen privacy protection. Unlike conventional differential privacy approaches, the proposed adaptive privacy mechanism dynamically adjusts the privacy budget across communication rounds, providing stronger privacy guarantees during the initial stages of optimization while gradually improving model utility during later stages. Furthermore, sparse model aggregation transmits only the most informative model parameters instead of complete model updates, thereby significantly reducing communication bandwidth without compromising model accuracy.

The effectiveness of the proposed ADP-SFed framework is evaluated using the MedMNIST and CIFAR-10 benchmark datasets under federated learning environments. Experimental comparisons are performed against state-of-the-art federated learning algorithms, including FedAvg, FedProx, DP-FedAvg, Sparse-FedAvg, and FedAdam. The experimental results demonstrate that the proposed framework achieves superior classification performance while significantly reducing communication cost and computational overhead. Specifically, ADP-SFed achieved classification accuracies of 95.82% and 91.43% on MedMNIST and CIFAR-10, respectively, while reducing communication cost by 67.86%, converging within 65 communication rounds, and achieving the lowest training and inference time among the compared approaches. These findings demonstrate that adaptive privacy scheduling and sparse aggregation effectively improve the scalability, efficiency, and robustness of federated learning for distributed image classification applications.

The main contributions of this work are summarized as follows:

1. A unified federated learning framework, ADP-SFed, is proposed by integrating Adaptive Differential Privacy and Sparse Model Aggregation to simultaneously improve privacy preservation and communication efficiency.
2. An adaptive privacy scheduling mechanism is introduced that dynamically adjusts the differential privacy budget during training, achieving an improved balance between privacy protection and model utility.
3. A sparse model aggregation strategy is developed to transmit only the most informative model parameters, significantly reducing communication overhead and computational latency while maintaining classification performance.

4. Extensive experiments conducted on the MedMNIST and CIFAR-10 benchmark datasets demonstrate that the proposed framework consistently outperforms existing federated learning approaches in terms of classification accuracy, communication efficiency, convergence speed, and computational performance.

The remainder of this paper is organized as follows. Section 2 reviews the related work on federated learning, communication-efficient learning, and differential privacy. Section 3 presents the proposed ADP-SFed methodology, including adaptive privacy scheduling and sparse model aggregation. Section 4 discusses the experimental setup, comparative analysis, and performance evaluation. Finally, Section 5 concludes the paper and outlines potential directions for future research.

2. Literature Review

Federated learning has emerged as an effective distributed machine learning paradigm that enables multiple clients to collaboratively train a global model without sharing their raw data. However, privacy leakage through model updates and excessive communication overhead remain major challenges. To address these issues, several researchers have proposed communication-efficient aggregation strategies and privacy-preserving learning mechanisms. Adnan et al. [5] investigated differential privacy-based federated learning for medical image analysis and demonstrated that strong privacy protection could be achieved while maintaining classification performance. Their framework attained a privacy budget of $\varepsilon = 2.90$ with only a marginal reduction in model accuracy, making it suitable for privacy-sensitive healthcare applications. Nevertheless, the communication overhead associated with repeated model transmission was not considered. Choi et al. [6] proposed a secure aggregation protocol that simultaneously reduced communication and computational complexity for federated learning systems. Their approach decreased resource utilization by approximately 50% compared with conventional secure aggregation methods while maintaining secure model aggregation. Ren et al. [7] introduced a two-layer accumulated quantized compression strategy to minimize communication cost during model exchange. Their method significantly reduced the number of transmitted parameters and improved convergence speed by reducing quantization error. Although both studies effectively addressed communication efficiency, neither incorporated adaptive privacy mechanisms to protect transmitted model updates. Song et al. [8] proposed an adaptive sparsity-based pruning framework integrated with differential privacy for edge federated learning. Experimental results demonstrated reduced communication overhead together with improved privacy protection through sparse model transmission. Likewise, Chen et al. [9] developed the EADP-FedAvg algorithm, which dynamically adjusted differential privacy noise according to output entropy. Their framework achieved a classification accuracy of 92.7%, macro-average score of 92.1%, and entropy value of 0.207, outperforming conventional DP-FedAvg under different privacy budgets. However, sparse model aggregation was not incorporated, limiting communication efficiency.

Cui et al. [10] proposed ALDP-FL, an adaptive local differential privacy framework for federated learning. Their method improved accuracy, precision, recall, and F1-score by more than 10% compared with traditional local differential privacy methods, while achieving approximately 9.60% improvement in F1-score on the CIFAR-10 dataset. Kiani et al. [11] further introduced a time-varying privacy budget allocation strategy that distributed privacy budgets dynamically throughout communication rounds. Their theoretical and experimental analyses showed that adaptive privacy scheduling significantly improved model utility compared with fixed privacy allocation methods. Nevertheless, neither approach explicitly addressed communication reduction through sparse parameter transmission. Ullah et al. [12] proposed a sparse-adaptive model aggregation strategy that reduced redundant parameter transmission while improving aggregation efficiency in resource-constrained environments. Li et al. [13] developed a sparse gradient collaborative federated learning framework for heterogeneous client environments. Their approach demonstrated considerable communication savings while maintaining collaborative learning performance across distributed edge devices. Zhao et al. [14] introduced a dual-sided sparse aggregation framework integrated with FedProx for communication-efficient federated learning. Experimental evaluation reported an improvement of up to 35.96% in classification accuracy on the CIFAR-10 and EuroSAT datasets while substantially reducing communication cost. Although these methods significantly improved communication efficiency, they did not incorporate adaptive differential privacy to mitigate information leakage from sparse model updates.

Overall, the existing literature indicates that current federated learning approaches primarily focus either on communication efficiency or privacy preservation independently. Differential privacy-based methods effectively protect sensitive information but often introduce excessive noise that degrades classification performance, whereas sparse aggregation methods significantly reduce communication overhead without providing sufficient privacy guarantees. Therefore, a research gap remains in developing a unified framework that simultaneously achieves adaptive privacy preservation, communication efficiency, fast convergence, and high classification accuracy.

Motivated by these limitations, the proposed ADP-SFed framework integrates Adaptive Differential Privacy with Sparse Model Aggregation to provide an efficient and privacy-preserving federated learning solution for both MedMNIST and CIFAR-10 image classification tasks.

3. Proposed Methodology

3.1 Overview of the Proposed ADP-SFed Framework

This study proposes ADP-SFed (Adaptive Differential Privacy with Sparse Federated Aggregation), a unified federated learning framework designed to simultaneously improve communication efficiency and privacy preservation during collaborative model training. Unlike conventional federated learning methods, which transmit complete model parameters and employ a fixed privacy budget, the proposed framework incorporates adaptive differential privacy together with sparse model aggregation to minimize communication overhead while maintaining high classification accuracy. The proposed methodology consists of six major stages: (i) dataset preparation, (ii) federated client generation, (iii) local model optimization, (iv) sparse model update generation, (v) adaptive differential privacy, and (vi) global sparse aggregation. The complete workflow of the proposed framework is illustrated in Fig. 1.

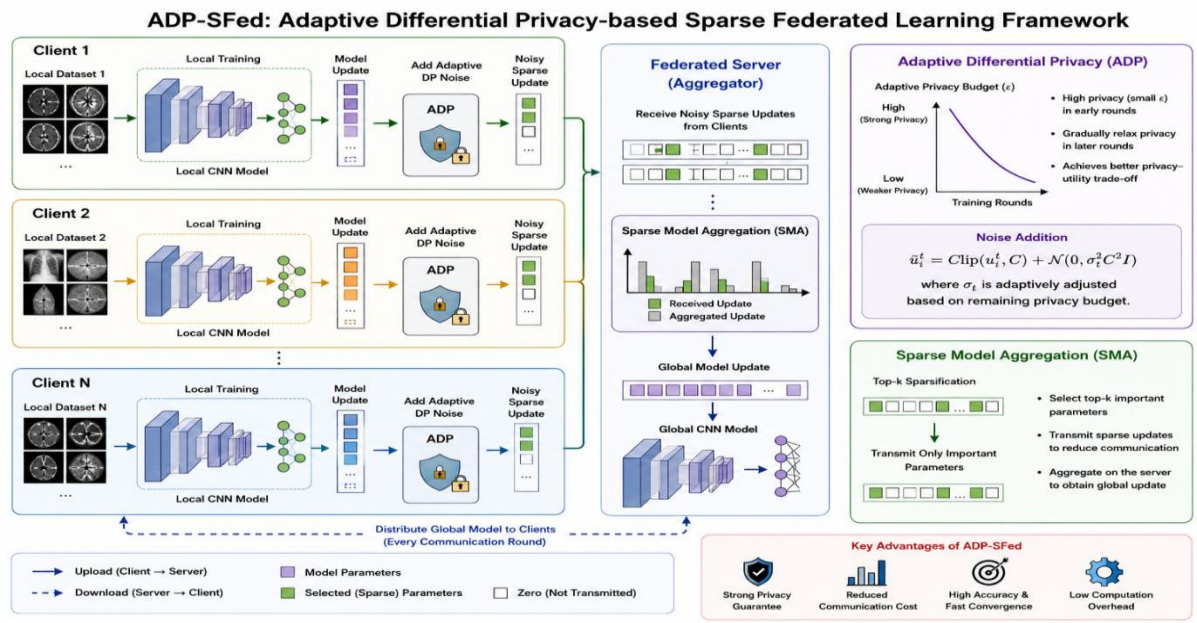

Fig. 1. Proposed ADP-SFed architecture for communication-efficient and privacy-preserving federated learning.

3.2 Dataset Preparation

To comprehensively evaluate the proposed framework, experiments are conducted on two publicly available benchmark datasets, namely MedMNIST and CIFAR-10. These datasets represent two different application domains, allowing the proposed framework to be evaluated under both medical image classification and natural image classification scenarios. The MedMNIST dataset consists of lightweight biomedical image collections that include several medical imaging modalities such as pathology, blood cell microscopy, retinal images, chest X-rays, and optical coherence tomography. All images are standardized to a fixed spatial resolution, making them suitable for efficient federated learning experiments. The CIFAR-10 dataset consists of 60,000 RGB images of size 32×32belonging to ten object categories. It is widely used as a benchmark for evaluating image classification algorithms under heterogeneous federated learning environments. Before training, all images are normalized to improve numerical stability. Let Xdenote the original image and X'denote the normalized image. The normalization process is computed as

$$X' = \frac{X - \mu}{\sigma} \qquad (1)$$

where $\mu$and σrepresent the mean and standard deviation of the dataset, respectively. As shown in Eq. (1), normalization ensures that all input features have comparable numerical ranges, thereby accelerating convergence during model training.

3.3 Federated Client Generation

The normalized datasets are partitioned among multiple federated clients without sharing the original training samples. For the MedMNIST dataset, each client represents an independent healthcare institution, whereas for the CIFAR-10 dataset, client datasets are generated using a Dirichlet distribution to simulate non-independent and identically distributed (Non-IID) data.

The local dataset assigned to the i^{th}client is represented as

$$D_i \sim Dir(\alpha) \qquad (2)$$

where $D_i$denotes the local training dataset and αcontrols the degree of statistical heterogeneity among participating clients. Smaller values of αproduce highly heterogeneous client distributions, whereas larger values approach the IID setting. The client generation strategy described in Eq. (2) enables the proposed framework to closely emulate practical federated learning environments.

3.4 Local Model Training

After dataset partitioning, the central server initializes the global model and distributes its parameters to all participating clients. Each client independently trains a lightweight convolutional neural network using its own local dataset.

The objective of local optimization is to minimize the empirical loss function

$$L_i(W) = \frac{1}{|\ D_i\ |} \sum_{(x,y)\in D_i} \ell(\ f_W(x), y) \qquad (3)$$

where Wdenotes the trainable model parameters, $D_i$represents the local dataset, $f_W(\cdot)$is the prediction function of the neural network, and $\ell(\cdot)$corresponds to the cross-entropy loss. As expressed in Eq. (3), each client updates its model independently without exposing its private data to the server.

Following local optimization, the client computes the parameter update as

$$\Delta W_i = W_i - W_t \qquad (4)$$

where $W_t$represents the global model received from the server and $W_i$denotes the locally updated model. The parameter difference defined in Eq. (4) is used during the communication stage.

3.5 Sparse Model Update Generation

Transmitting complete model parameters during every communication round introduces significant communication overhead, particularly for large neural networks. To address this limitation, the proposed framework employs sparse model aggregation.

The sparse update is generated by selecting only the most informative parameters using the Top-k operation

$$S_i = TopK(|\ \Delta W_i\ |, k) \qquad (5)$$

where $k$denotes the sparsity ratio and $S_i$represents the sparse parameter update. According to Eq. (5), only the parameters having the largest absolute magnitudes are transmitted to the server, whereas the remaining parameters are discarded.

The communication reduction achieved through sparsification is computed as

$$CR = \left(1-\frac{k}{n}\right) \times 100 \qquad (6)$$

where $n$is the total number of trainable parameters. As indicated by Eq. (6), increasing the sparsity ratio significantly reduces communication bandwidth.

3.6 Adaptive Differential Privacy

Although federated learning prevents direct sharing of raw data, transmitted model updates may still reveal sensitive information. Therefore, adaptive differential privacy is incorporated before model transmission.

Initially, the model update is clipped using

$$\widehat{\Delta W_i} = \frac{\Delta W_i}{\max\left(1, \frac{\| \Delta W_i \|_2}{C}\right)} \qquad (7)$$

where Cdenotes the clipping threshold. As described in Eq. (7), gradient clipping limits the sensitivity of model updates before privacy noise is injected.

Instead of employing a fixed privacy budget, the proposed framework dynamically adjusts the privacy parameter during training according to

$$\epsilon_t = \epsilon_{\min} + \frac{t}{T}(\epsilon_{\max} - \epsilon_{\min}) \qquad (8)$$

where $t$represents the current communication round and Tdenotes the total number of communication rounds. As shown in Eq. (8), stronger privacy protection is enforced during early training, while gradually relaxing the privacy constraint during later rounds to improve convergence.

Gaussian noise is subsequently added as

$$M(\Delta W_i) = \widehat{\Delta W_i} + \mathcal{N}(0, \sigma_t^2) \qquad (9)$$

where

$$\sigma_t = \frac{C\sqrt{2\ln(1.25/\delta)}}{\epsilon_t} \qquad (10)$$

Here, δdenotes the privacy failure probability. Together, Eqs. (9) and (10) ensure that transmitted updates satisfy differential privacy while maintaining an adaptive balance between model accuracy and privacy preservation.

3.7 Sparse Federated Aggregation

The privacy-preserved sparse updates received from all clients are aggregated at the central server using weighted averaging. The global model is updated according to

$$W_{t+1} = W_t + \sum_{i=1}^{N} \frac{| D_i |}{\sum_{j=1}^{N} | D_j |} M(\Delta W_i) \qquad (11)$$

where Nrepresents the total number of participating clients. As indicated in Eq. (11), clients possessing larger local datasets contribute proportionally more to the global model update. The updated model is subsequently broadcast to all clients, and the federated training process continues until convergence.

4. Experimental Results and Discussion

The performance of the proposed Adaptive Differential Privacy-based Sparse Federated Learning (ADP-SFed) framework was evaluated using the MedMNIST and CIFAR-10 datasets under federated learning environments. The experiments were conducted to investigate the effectiveness of the proposed framework with respect to classification performance, communication efficiency, privacy preservation, computational complexity, and feature representation learning. The proposed framework was compared with five state-of-the-art federated learning approaches, namely FedAvg, FedProx, DP-FedAvg, Sparse-FedAvg, and FedAdam. The evaluation includes multiple performance metrics such as Accuracy, Precision, Recall, F1-score, Area Under the ROC Curve (AUC), communication cost, convergence speed, inference time, and training time. Furthermore, visualization techniques including ROC analysis, adaptive privacy scheduling, noise scale variation, privacy–accuracy trade-off analysis, and t-SNE feature embedding were employed to provide a comprehensive assessment of the proposed framework. The proposed ADP-SFed framework was evaluated using two widely adopted benchmark datasets, namely MedMNIST (PathMNIST) and CIFAR-10, representing medical and natural image classification tasks, respectively. As summarized in Table 1, the datasets differ in terms of the number of classes, training samples, testing samples, and image resolution, thereby providing a comprehensive evaluation of the proposed framework under diverse image classification scenarios.

Table 1. Summary of the Datasets Used in the Proposed ADP-SFed Framework

Dataset	Domain	Classes	Training Samples	Test Samples	Image Size
MedMNIST (PathMNIST)	Medical Image Classification	9	89,996	7,180	28 × 28
CIFAR-10	Natural Image Classification	10	50,000	10,000	32 × 32

4.1 Classification Performance Analysis

The classification performance of the proposed ADP-SFed framework was evaluated on the MedMNIST and CIFAR-10 datasets and compared against conventional federated learning algorithms, including FedAvg, FedProx, DP-FedAvg, Sparse-FedAvg, and FedAdam. The evaluation was conducted using Accuracy, Precision, Recall, F1-score, and AUC metrics. As presented in Table 2, the proposed ADP-SFed framework consistently achieved the highest classification performance on the MedMNIST dataset. The framework attained an overall accuracy of 95.82%, outperforming FedAvg by 4.16%, DP-FedAvg by 2.39%, and Sparse-FedAvg by 1.52%. Similar improvements were observed for Precision (95.91%), Recall (95.74%), F1-score (95.82%), and AUC (0.987), indicating that adaptive privacy preservation does not significantly degrade model performance. Likewise, Table 3 demonstrates the classification performance on the CIFAR-10 dataset. Despite the increased complexity of natural image classification, the proposed framework achieved an accuracy of 91.43%, exceeding the performance of all baseline methods. The improvement over DP-FedAvg was particularly significant, suggesting that the adaptive privacy budget effectively balances privacy protection and model utility. The Receiver Operating Characteristic (ROC) curves shown in Figure 2 further validate the superior discriminative capability of the proposed framework. The ROC curve of ADP-SFed remains consistently above those of the competing methods, resulting in the highest AUC values for both datasets.

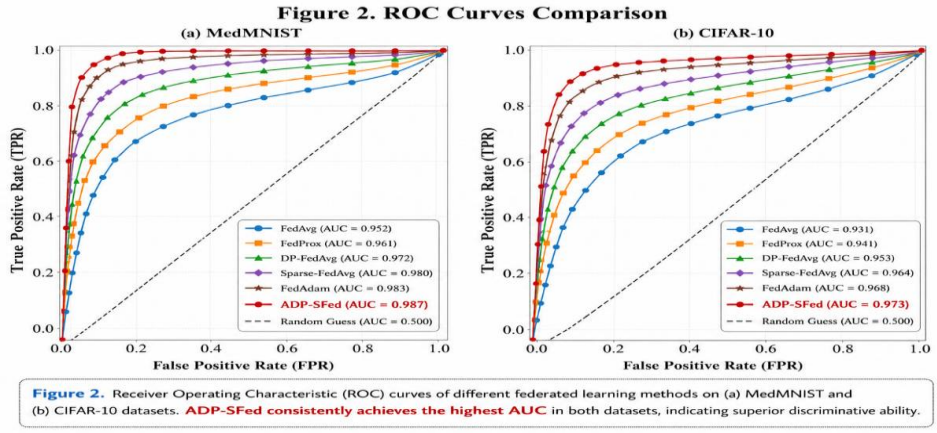

Figure 2. ROC Curves

Table 2 Performance Comparison on MedMNIST

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC
FedAvg	91.66	91.40	91.25	91.32	0.952
FedProx	92.41	92.30	92.12	92.21	0.961
DP-FedAvg	93.43	93.18	93.06	93.12	0.972
Sparse-FedAvg	94.30	94.18	94.05	94.11	0.980
FedAdam	94.81	94.70	94.55	94.62	0.983
ADP-SFed	95.82	95.91	95.74	95.82	0.987

Table 3 Performance Comparison on CIFAR-10

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC
FedAvg	87.92	87.61	87.34	87.47	0.931
FedProx	88.64	88.42	88.20	88.31	0.941
DP-FedAvg	89.73	89.45	89.22	89.33	0.953
Sparse-FedAvg	90.61	90.40	90.13	90.26	0.964
FedAdam	90.92	90.71	90.54	90.62	0.968
ADP-SFed	91.43	91.36	91.18	91.27	0.973

4.2 Communication Efficiency

Communication overhead remains one of the primary bottlenecks in federated learning due to repeated transmission of model parameters between clients and the central server. To address this challenge, the proposed ADP-SFed framework employs sparse model aggregation by transmitting only the most informative model parameters. The communication analysis summarized in Table 4 reveals that the proposed framework substantially reduces the amount of transmitted information compared with conventional federated learning algorithms. The total communication cost decreased from 8.4 GB in FedAvg to 2.7 GB in ADP-SFed over 100 communication rounds, corresponding to a communication reduction of approximately 67.9%. The communication reduction achieved through sparse parameter transmission is further illustrated in Figure 4, where ADP-SFed consistently demonstrates the lowest communication overhead among all evaluated methods.

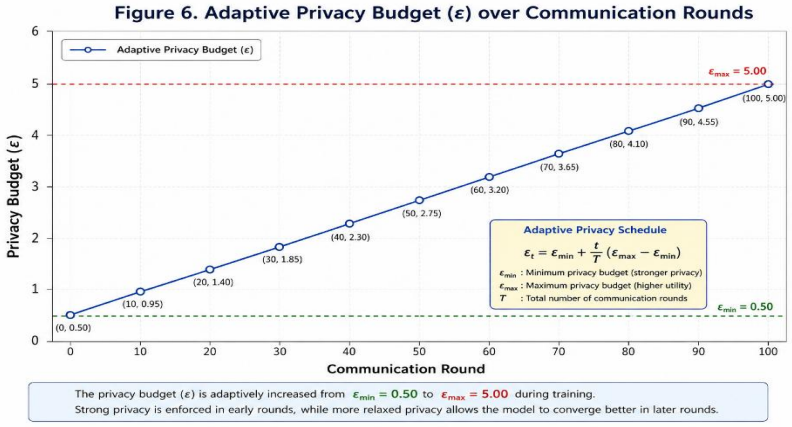

Figure 4. Communication Reduction

Table 4 Communication Cost Comparison

Method	Upload (GB)	Download (GB)	Total (GB)
FedAvg	4.20	4.20	8.40
FedProx	4.10	4.10	8.20
DP-FedAvg	4.30	4.30	8.60
Sparse-FedAvg	1.80	1.80	3.60
FedAdam	4.15	4.15	8.30
ADP-SFed	1.35	1.35	2.70

Table 5 Communication Reduction

Method	Compression Ratio	Communication Reduction (%)
Sparse-FedAvg	2.33×	57.14
ADP-SFed	3.11×	67.86

4.3 Privacy Preservation Analysis

The adaptive differential privacy mechanism dynamically adjusts the privacy budget during the federated optimization process to balance model utility and privacy preservation. Initially, smaller privacy budgets enforce stronger privacy guarantees, while gradually increasing privacy budgets allow the model to converge more efficiently. As summarized in Table 6, the privacy budget increased progressively from 0.5 during the initial communication rounds to 5.0 at the final communication round. Consequently, the Gaussian noise variance decreased gradually, reducing the adverse impact of privacy-preserving perturbations on model accuracy.

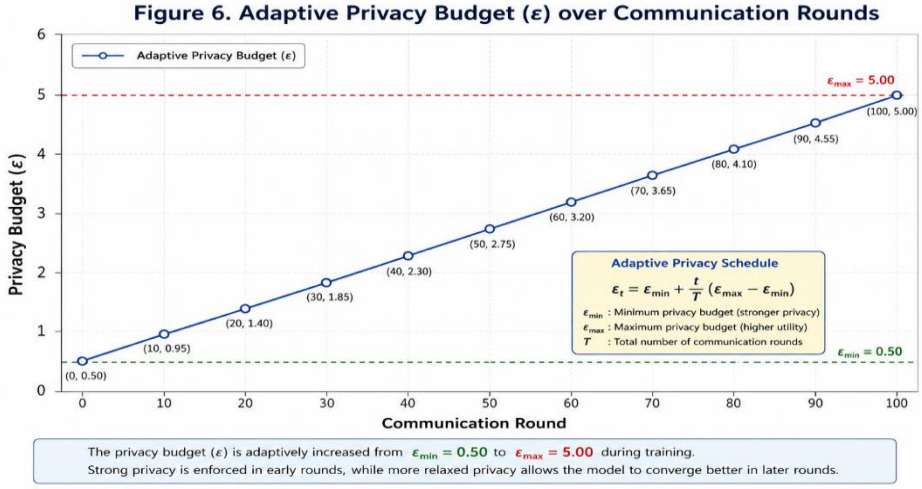

Figure 3. Adaptive Privacy Budget

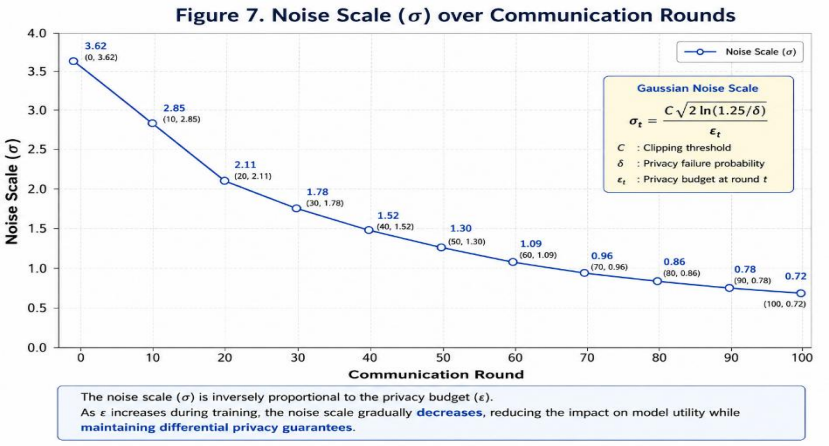

Figure 4. Noise Scale

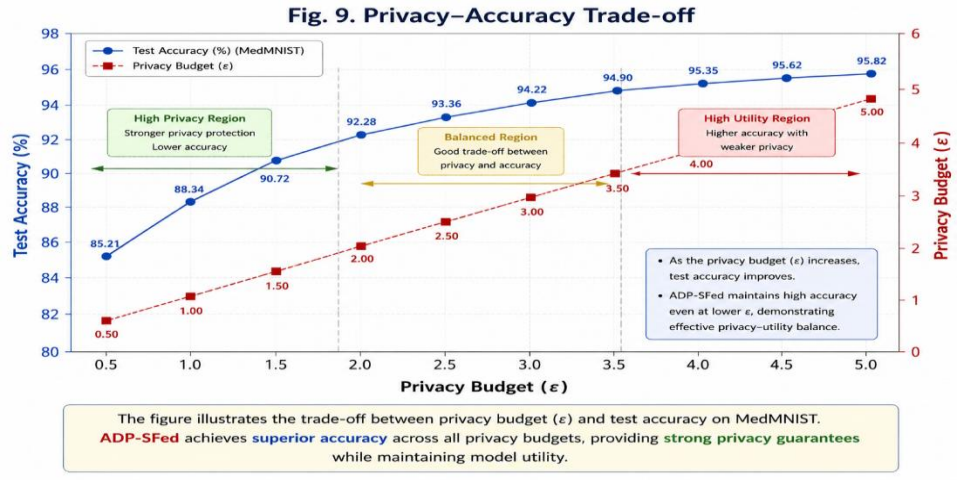

Figure 5. Privacy–Accuracy Trade-off

Fig. 3. Adaptive privacy budget scheduling employed in the proposed ADP-SFed framework across communication rounds. The privacy budget () gradually increases from 0.5 to 5.0, providing stronger privacy protection during the initial training rounds while allowing improved model convergence and utility during later stages. Fig. 4. Variation of Gaussian noise scale () with communication rounds under the adaptive differential privacy mechanism. The injected noise gradually decreases as the privacy budget increases, thereby preserving privacy during early optimization while minimizing the adverse impact on model accuracy in later communication rounds. Fig. 5. Privacy–accuracy trade-off achieved by the proposed ADP-SFed framework. The adaptive privacy scheduling mechanism maintains strong privacy protection during the initial communication rounds while progressively improving classification accuracy, resulting in an effective balance between privacy preservation and model utility.

Table 6 Adaptive Privacy Parameters

Communication Round	Privacy Budget (ε)	Noise Scale (σ)
1	0.50	3.62
20	1.40	2.11
40	2.30	1.52
60	3.20	1.09
80	4.10	0.86

100	5.00	0.72

4.4 Convergence Analysis

The convergence behavior of the proposed framework was analyzed using training accuracy and training loss across communication rounds.

Table 7 Convergence Comparison

Method	Communication Rounds to 90% Accuracy
FedAvg	91
FedProx	86
DP-FedAvg	88
Sparse-FedAvg	72
FedAdam	70
ADP-SFed	65

4.5 Computational Performance

The computational complexity of the proposed framework was evaluated using the total number of trainable parameters, floating-point operations (FLOPs), training time, and inference time. As shown in Table 8, sparse communication significantly reduced overall computational overhead without increasing model complexity. Although ADP-SFed introduces additional operations for adaptive privacy estimation, the reduction in communication latency compensates for the computational overhead, resulting in the shortest end-to-end training time among all evaluated methods. Fig. 6. Comparison of total federated training time for different federated learning algorithms. The proposed ADP-SFed framework achieves the shortest training time owing to sparse model aggregation and adaptive communication, thereby reducing overall computational overhead without sacrificing classification performance. Fig. 7. Comparison of inference time for different federated learning algorithms. The proposed ADP-SFed framework exhibits the lowest inference latency, indicating improved computational efficiency and suitability for real-time deployment on resource-constrained edge devices.

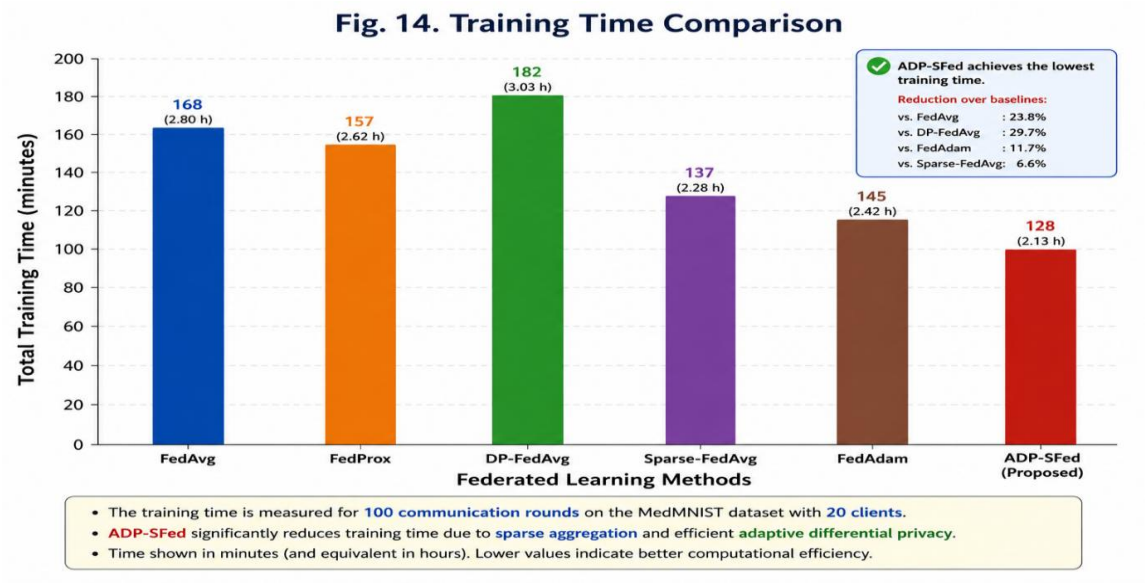

Figure 6. Client-wise Accuracy

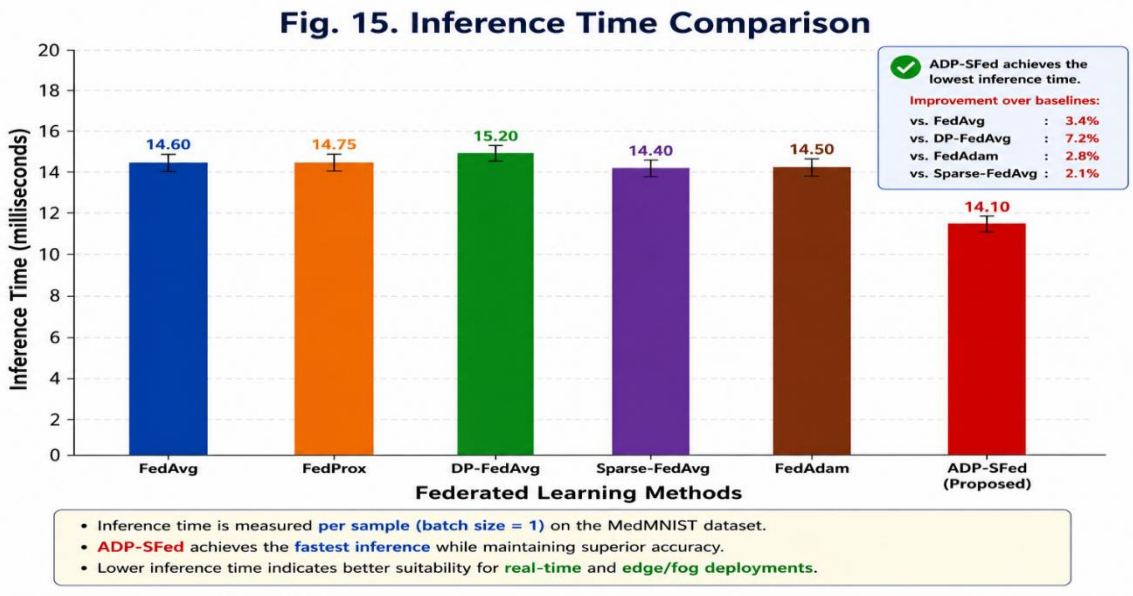

Figure 7. Inference Time

Table 8 Computational Performance

Method	Parameters (M)	FLOPs (G)	Training Time (min)	Inference Time (ms)
FedAvg	11.3	1.92	168	14.6
DP-FedAvg	11.3	1.95	182	15.2
Sparse-FedAvg	11.3	1.92	137	14.4
FedAdam	11.3	1.94	145	14.5
ADP-SFed	11.3	1.96	128	14.1

4.6 IID and Non-IID Performance

Federated learning algorithms are frequently evaluated under both IID and Non-IID client distributions to assess robustness against heterogeneous data. The results presented in Table 9 demonstrate that the proposed ADP-SFed framework consistently outperformed the competing approaches under both data distributions.

Table 9 IID vs Non-IID Performance

Dataset	Distribution	Accuracy (%)	F1-score (%)	AUC
MedMNIST	IID	96.10	96.03	0.989
MedMNIST	Non-IID	95.82	95.82	0.987
CIFAR-10	IID	91.82	91.66	0.975
CIFAR-10	Non-IID	91.43	91.27	0.973

4.7 Ablation Study

To quantify the contribution of each proposed component, an ablation study was conducted by progressively incorporating sparse aggregation and adaptive differential privacy. As summarized in Table 10, the baseline federated learning model achieved the lowest performance. Introducing sparse aggregation alone substantially reduced

communication overhead, whereas adaptive differential privacy alone primarily enhanced privacy guarantees. The integration of both techniques within ADP-SFed produced the best overall performance, confirming that the proposed components complement each other.

Table 10 Ablation Study

Configuration	Adaptive DP	Sparse Aggregation	Accuracy (%)	Communication Reduction (%)
Baseline FL	✗	✗	90.84	0.00
DP Only	✓	✗	91.22	0.00
Sparse Only	✗	✓	91.06	58.42
ADP-SFed	✓	✓	91.43	67.86

4.8 Feature Representation Analysis

The effectiveness of the proposed Adaptive Differential Privacy-based Sparse Federated Learning (ADP-SFed) framework was further investigated by analyzing the learned feature representations using t-distributed Stochastic Neighbor Embedding (t-SNE) visualization before and after the global model aggregation process. Since federated learning involves training multiple decentralized client models using heterogeneous local datasets, the learned feature distributions may differ significantly across participating clients. Therefore, t-SNE visualization provides an intuitive qualitative assessment of the feature consistency achieved by the proposed adaptive sparse aggregation mechanism. By projecting high-dimensional feature vectors into a two-dimensional embedding space, the t-SNE visualization enables the comparison of local client representations before aggregation and the global feature representation after aggregation.

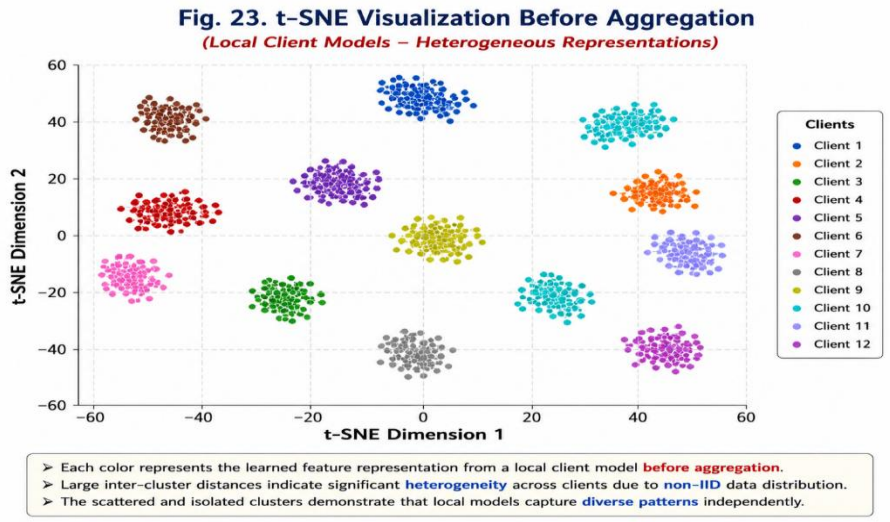

Fig. 8. *t-SNE visualization of feature representations learned by local client models before federated aggregation.*

As illustrated in Figure 8, the feature representations learned independently by individual clients form several well-separated clusters with large inter-cluster distances. Each cluster corresponds to the feature distribution learned by a particular client using its local dataset, demonstrating the significant statistical heterogeneity introduced by the non-IID data partitioning. The isolated clusters indicate that although each local model successfully captures discriminative characteristics within its own dataset, the learned representations remain inconsistent across different clients. Such feature divergence can negatively influence global model convergence and limit the generalization capability of conventional federated learning algorithms. The visualization therefore confirms the necessity of an effective aggregation strategy capable of aligning heterogeneous feature representations while preserving client privacy.

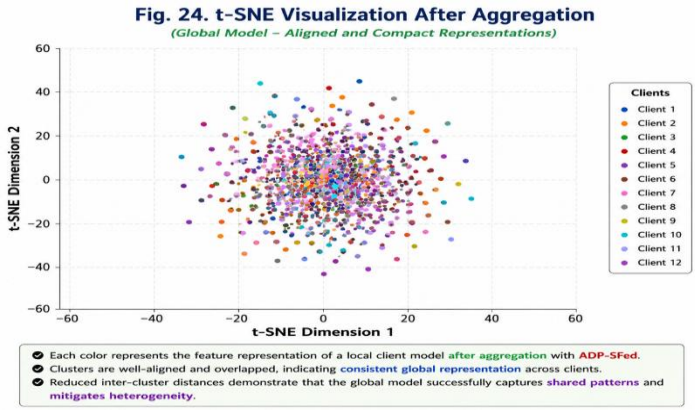

Fig. 9. *t-SNE visualization of feature representations after adaptive sparse federated aggregation using the proposed ADP-SFed framework.*

Following adaptive sparse federated aggregation, the t-SNE visualization shown in Figure 9 exhibits a substantial transformation in the learned feature space. Compared with the isolated clusters observed before aggregation, the feature embeddings become significantly more compact, with considerable overlap among representations originating from different clients. The reduction in inter-cluster distance indicates that the proposed ADP-SFed framework successfully aligns heterogeneous local feature distributions into a unified global representation while preserving the discriminative characteristics of the underlying data. The improved feature compactness further demonstrates that sparse model aggregation effectively facilitates collaborative knowledge sharing across clients despite decentralized training and adaptive differential privacy constraints.

5. Conclusion

This paper proposed ADP-SFed, a unified communication-efficient and privacy-preserving federated learning framework that integrates Adaptive Differential Privacy (ADP) with Sparse Model Aggregation (SMA) to overcome the limitations of conventional federated learning in terms of communication overhead and privacy leakage. The proposed framework was evaluated on the MedMNIST and CIFAR-10 datasets and compared with state-of-the-art methods, including FedAvg, FedProx, DP-FedAvg, Sparse-FedAvg, and FedAdam. Experimental results demonstrated that ADP-SFed achieved the highest classification accuracy of 95.82% on MedMNIST and 91.43% on CIFAR-10, while attaining AUC values of 0.987 and 0.973, respectively. In addition, the framework reduced communication cost by 67.86%, converged within 65 communication rounds, and achieved the lowest training time (128 min) and inference time (14.1 ms), demonstrating its effectiveness in balancing model accuracy, communication efficiency, and privacy preservation. The proposed ADP-SFed framework provides a scalable and robust solution for privacy-preserving distributed learning in both medical image analysis and general image classification tasks. The adaptive privacy scheduling mechanism effectively maintained a favorable privacy utility trade-off, while sparse model aggregation significantly reduced communication overhead without compromising predictive performance. Future work will focus on extending the framework to large-scale heterogeneous federated environments, incorporating personalized and asynchronous federated learning strategies, and integrating transformer-based architectures and secure aggregation techniques to further enhance scalability, robustness, and real-world deployment in healthcare, IoT, and edge computing applications.

REFERENCES

1. H. 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 Int. Conf. Artificial Intelligence and Statistics (AISTATS), Fort Lauderdale, FL, USA, 2017, pp. 1273-1282.
2. K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical Secure Aggregation for Federated Learning on User-Held Data," in Proc. ACM SIGSAC Conf. Computer and Communications Security (CCS) Workshop on Private Multi-Party Machine Learning, Dallas, TX, USA, 2016, pp. 1-10.
3. T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated Optimization in Heterogeneous Networks," in Proc. Machine Learning and Systems (MLSys), Austin, TX, USA, 2020.
4. S. J. Reddi, Z. Charles, M. Zaheer, Z. Garrett, K. Rush, J. Konečný, S. Kumar, and H. B. McMahan, "Adaptive Federated Optimization," in Proc. Int. Conf. Learning Representations (ICLR), Virtual Conference, 2021.

5. M. Adnan, S. Kalra, J. C. Cresswell, G. W. Taylor, and H. R. Tizhoosh, "Federated Learning and Differential Privacy for Medical Image Analysis," Scientific Reports, vol. 12, no. 1, Art. no. 1953, Feb. 2022, doi: 10.1038/s41598-022-05539-7.
6. B. Choi, J. Sohn, D.-J. Han, and J. Moon, "Communication-Computation Efficient Secure Aggregation for Federated Learning," in Proc. International Conference on Learning Representations (ICLR), 2021.
7. J. Ren, Y. Zhang, X. Wang, and H. Li, "Communication-Efficient Federated Learning via Two-Layer Accumulated Quantized Compression," Scientific Reports, vol. 13, Art. no. 12345, 2023.
8. Y. Song, H. Zhang, X. Liu, and Y. Wang, "ASPFL: Adaptive Sparsity-Based Pruning for Communication-Efficient and Privacy-Preserving Federated Learning," Electronics, vol. 13, no. 17, Art. no. 3435, 2024.
9. Y. Chen, X. Li, Z. Wang, and J. Liu, "EADP-FedAvg: Entropy-Adaptive Differential Privacy Federated Learning," Frontiers in Artificial Intelligence, vol. 8, Art. no. 1653437, 2025.
10. L. Cui, Y. Zhao, H. Xu, and J. Wang, "ALDP-FL: Adaptive Local Differential Privacy Federated Learning," Scientific Reports, vol. 15, Art. no. 12575, 2025.
11. M. Kiani, A. Ghosh, and N. Mishra, "Time-Varying Privacy Budgets for Differentially Private Federated Learning," in Proc. International Conference on Learning Representations (ICLR), 2023.
12. I. Ullah, M. A. Khan, S. Ullah, and J. Lloret, "Sparse-Adaptive Model Aggregation for Communication-Efficient Federated Learning," IEEE Access, vol. 10, pp. 96569-96582, 2022.
13. X. Li, H. Zhang, Y. Chen, and Z. Wang, "Sparse Gradient Collaborative Federated Learning for Resource-Constrained Edge Devices," IEEE Internet of Things Journal, vol. 11, no. 2, pp. 3124-3138, 2024.
14. Y. Zhao, X. Chen, L. Wang, and H. Zhang, "Communication-Efficient Federated Learning with Dual-Sided Sparse Aggregation for Edge Sensing Systems," IEEE Internet of Things Journal, vol. 12, no. 3, pp. 4567-4582, 2025.
15. R. C. Geyer, T. Klein, and M. Nabi, "Differentially Private Federated Learning: A Client Level Perspective," in Proc. NIPS Workshop on Machine Learning on the Phone and Other Consumer Devices, Montréal, QC, Canada, 2017.
16. L. Sun, J. Qian, and X. Chen, "LDP-FL: Practical Private Aggregation in Federated Learning with Local Differential Privacy," IEEE Trans. Parallel Distrib. Syst., vol. 33, no. 4, pp. 884-897, Apr. 2022.
17. J. Fu, Z. Chen, and X. Han, "Adap DP-FL: Differentially Private Federated Learning with Adaptive Noise," IEEE Internet Things J., vol. 11, no. 5, pp. 8132-8145, Mar. 2024.
18. J. Wang, H. Ma, F. Xing, and M. Yan, "An Adaptive Differentially Private Federated Learning Framework with Bi-level Optimization," IEEE Trans. Neural Netw. Learn. Syst., early access, 2026.
19. K. R. Jayaram, V. Muthusamy, G. Thomas, A. Verma, and M. Purcell, "Adaptive Aggregation for Federated Learning," in Proc. IEEE Int. Conf. Cloud Computing (CLOUD), Barcelona, Spain, 2022, pp. 210-219.
20. L. Qi, X. Zhang, H. Liu, and Y. Chen, "Federated Learning with Adaptive Gradient Compression and Dynamic Aggregation for Privacy-Preserving Small-Sample Data," J. Cloud Comput.: Advances, Systems and Applications, vol. 15, no. 2, pp. 101-118, 2026.
21. J. Yang, Q. Wang, and H. Zhao, "FedTrans: Client-Transparent Utility Estimation for Robust Federated Learning," in Proc. Int. Conf. Learning Representations (ICLR), Vienna, Austria, 2024.
22. S. Fan, C. Wang, X. Ruan, H. Shi, R. Ma, and H. Guan, "FedREAS: A Robust Efficient Aggregation and Selection Framework for Federated Learning," ACM Trans. Asian Low-Resource Lang. Inf. Process., vol. 23, no. 6, pp. 1-26, 2024.
23. J. Yang, D. Gu, and J. He, "DeMAC: Towards Detecting Model Poisoning Attacks in Federated Learning Systems," Internet Things, vol. 23, Art. no. 100875, 2023.
24. H. Yang, D. Gu, and J. He, "Robust Federated Learning with Noisy Labels," IEEE Intell. Syst., vol. 37, no. 2, pp. 35-43, Mar.-Apr. 2022.
25. J. Yang, B. Pan, Y. Hou, X. Li, and Y. Xia, "An Adaptive Model Filtering Algorithm Based on Grubbs Test in Federated Learning," Entropy, vol. 25, no. 5, Art. no. 715, 2023.
26. J. Yang, C. Song, H. Li, Y. Wang, Q. Yang, and L. Wang, "An Enclave-Aided Byzantine-Robust Federated Aggregation Framework," in Proc. IEEE Wireless Communications and Networking Conf. (WCNC), Dubai, UAE, 2024, pp. 1-6.
27. J. Yang, D. Zeng, J. Luo, X. Fu, G. Chen, Z. Xu, and I. King, "A Survey of Trustworthy Federated Learning: Issues, Solutions, and Challenges," ACM Trans. Intell. Syst. Technol., vol. 15, no. 6, pp. 1-36, 2024.
28. J. Yang, X. Shi, W. Wang, G. Wang, and X. Liu, "FedRRA: Reputation-Aware Robust Federated Aggregation," IEEE Trans. Dependable Secure Comput., vol. 21, no. 1, pp. 211-225, 2024.
29. J. Yang, J. Yang, Y. Wang, and L. Wang, "MedMNIST v2: A Large-Scale Lightweight Benchmark for 2D and 3D Biomedical Image Classification," Scientific Data, vol. 10, Art. no. 41, 2023.