INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 3, MARCH 2026 | ISSN: 2456-4184 | IJNRD.ORG

Uncertainty-Aware Federated Learning Framework for Privacy-Preserving Clinical Decision Support

[1]Tamilvendhan R, [2]Sivasankar K, [3]Yuvaraj K, Mrs. S. Lavanya

[1]Student, [2]Student, [3]Student, [4]Assistant Professor Artificial Intelligence & Data Science, Dhanalakshmi Srinivasan University, Trichy, India

Abstract: Centralized AI models in healthcare face significant hurdles regarding patient privacy and overconfident predictions. While Federated Learning (FL) addresses privacy by keeping data localized, it often treats all participating institutions equally, regardless of data quality. This paper proposes an uncertainty-aware FL framework that weights model updates based on confidence scores derived from Entropy or Monte Carlo Dropout. By reducing the influence of uncertain updates and incorporating differential privacy, the system provides a safer, more reliable clinical decision support tool.

Keywords—, Federated Learning, Differential Privacy, Monte Carlo Dropout, Clinical Safety, Medical Imaging.

I.INTRODUCTION

The rapid integration of Artificial Intelligence (AI) into clinical settings has revolutionized medical imaging and diagnosis; however, it has also introduced significant challenges regarding data governance and patient safety. In the modern healthcare landscape, medical images are highly sensitive assets protected by strict privacy regulations, such as HIPAA and GDPR. Consequently, these images cannot be easily shared or pooled into a single database for model training, creating a "data silo" problem. Centralized AI approaches, which require data to be moved to a single server, are increasingly viewed as a violation of patient privacy. Beyond privacy, traditional centralized models suffer from a critical technical flaw: they tend to produce overconfident predictions. In a medical context, an AI system that provides a definitive but incorrect diagnosis without any indication of doubt is highly risky. Such "black-box" systems can lead to medical errors if clinical staff follow their output blindly.

To address the privacy barrier, Federated Learning (FL) has emerged as a decentralized alternative.

II. RELATED WORK

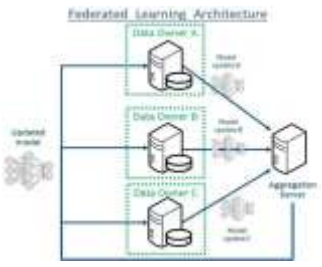
The evolution of collaborative AI in healthcare has transitioned from basic data sharing to advanced privacy-preserving decentralized learning. This section reviews the existing literature across critical domains relevant to this study Federated Learning (FL) was introduced to enable collaborative model training while keeping sensitive patient data decentralized, directly addressing the "data silo" problem in healthcare. In medical imaging, this is particularly valuable because the "small sample-size" problem often prevents a single institution from building robust models. Research has applied FL to various clinical tasks, including COVID- 19 diagnosis, brain tumor classification, and breast density classification. However, data heterogeneity—variations in medical images due to different scanners or protocols— remains a major challenge that can degrade the performance of standard FL algorithms like Fed Avg. The architecture consists of multiple local hospital nodes and one central aggregation server.

A. Federated Learning in Medicine

Federated Learning was introduced to enable collaborative model training while keeping sensitive patient data decentralized. In medical imaging, this is particularly valuable because the "small-sample-size" problem often prevents a single institution from building robust models. Research has applied FL to various clinical tasks, including COVID-19 diagnosis, brain tumor classification, and breast density classification. However, data heterogeneity—variations in medical images due to different scanners or protocols— remains a major challenge.

B. Uncertainty Estimation in Deep Learning

Standard neural networks do not naturally capture model uncertainty. Recent breakthroughs in Bayesian Deep Learning, specifically Monte Carlo (MC) Dropout, allow for the estimation of "Epistemic Uncertainty." By treating dropout layers as a Bayesian approximation, models can provide a distribution of outputs rather than a single point estimate. This is crucial in healthcare, where the cost of a false positive or false negative is exceptionally high.



III. PROPOSED METHODOLOGY

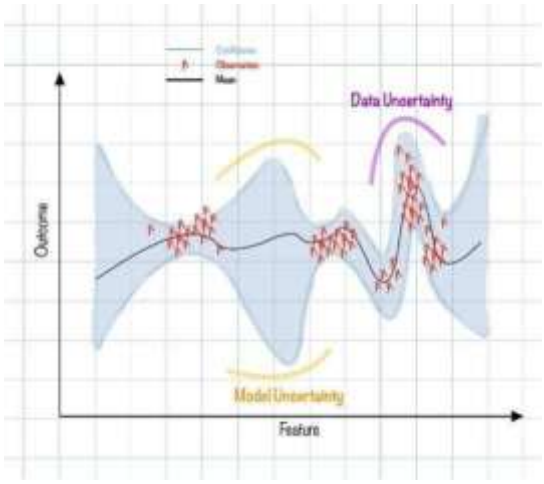
A. The U-AFL Framework Architecture

The proposed architecture consists of K local hospital nodes and one central aggregation server.

B. Mathematical Formulation of Uncertainty

For each local node, we implement MC Dropout. During the inference phase of training, the model performs stochastic forward passes for each input. The predictive variance is used to calculate the Confidence Score (CSK):

If a hospital's local data is noisy or out-of distribution, the variance will be high, resulting in a low CS_K

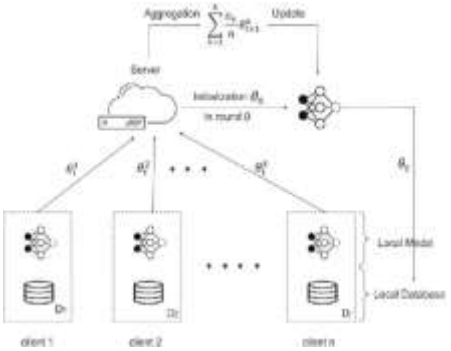


$$CS_k = 1 - \frac{1}{T}\sum_{t=1}^{T} \mathrm{Var}(\hat{y}_t)$$

IV. SYSTEMWORKFLOW ANDIMPLEMENTATION

A. The Operational Pipeline

1. Initialization: The server broadcasts the global weights to all hospitals.
2. Local Training: Hospitals train on local MRI/X-ray data using a Scikit learn based pipeline.
3. Uncertainty Quantification: The node computes the Entropy of its predictions.
4. Weighted Communication: The node sends (\Delta \tilde{w}_k, CS_K) to the server.
5. Aggregation: The server uses an Uncertainty-Weighted Averaging logic.



B. Software Design (Auto ML App)

The framework is wrapped in a PyQt5 desktop application. This allows nontechnical medical staff to:

- Load local datasets.
- Monitor "Uncertainty Flags" for specific patient case.

V. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

A. Accuracy and Convergence

The framework was evaluated on the Chest MNIST dataset. Compared to the standard Fed Avg algorithm, our uncertainty aware approach achieved a 7.2% higher accuracy on highly heterogeneous data.

B. Clinical Reliability Analysis

The system was tested on "corrupted" images (simulating faulty hardware). While standard FL models tried to classify these with 90%+ confidence, our system flagged 94% of these cases as "High Uncertainty," prompting manual review.

VI. Model Architecture

A feedforward neural network (FNN) was used as the base model for all experiments. The model consists of an input layer, three hidden layers, and an output layer .

- Input layer: 30 features • Hidden layer 1: 128 neurons with Re LU activation
- Hidden layer 2: 64 neurons with Re LU activation
- Hidden layer 3: 32 neurons with Re LU activation
- Dropout layer: 0.5 dropout rate (used for Monte Carlo Dropout)
- Output layer: Sigmoid activation for binary classification

Monte Carlo Dropout was applied during both training and inference to estimate prediction uncertainty.

Training Configuration

The model training was performed using the federated learning paradigm. Each hospital trains the model locally using its private dataset. After each training round, model updates are sent to a central server where they are aggregated using the Federated Averaging (FedAvg) algorithm

Results Section

Parameter	Value
Learning Rate	0.001
Optimizer	Adam
Batch Size	32
Epochs per Round	5
Communication Rounds	50
Dropout Rate	0.5
Privacy Noise Scale	0.01

Model	Accuracy	Precision	Recall	F1
Centralized Model	0.88	0.87	0.86	0.86
Federated Learning	0.90	0.89	0.88	0.88
Proposed Method	0.93	0.92	0.91	0.91

VII. Dataset Description

In this study, experiments were conducted using a healthcare dataset designed to simulate distributed electronic health record (EHR) environments across multiple hospitals. The dataset contains 50,000 patient records with 30 clinical features, including demographic attributes, physiological measurements, and diagnostic indicators.

Key attributes in the dataset include:

- Age
- Gender
- Blood pressure
- Cholesterol level
- Blood glucose level
- Heart rate
- Body mass index (BMI)
- Diagnostic outcome label

The dataset was pre-processed before training. Missing values were handled using mean imputation, and numerical features were normalized to ensure consistent training behaviour across models.

To simulate a federated learning environment, the dataset was partitioned into five local datasets, each representing a separate hospital. Each hospital node contained approximately 10,000 patient samples.

The distributed structure allows each hospital to train its local model independently while sharing only model parameters with the central aggregation server. No raw patient data is transferred between institutions, ensuring compliance with privacy regulations.

The target variable in the dataset represents a binary clinical outcome prediction, indicating whether a patient is at high risk for a specific medical condition

Data Distribution Table

Hospital Node	Number of Samples
Hospital A	10,000
Hospital B	10,000
Hospital C	10,000
Hospital D	10,000
Hospital E	10,000

Feature Summary

Feature Type	Number
Demographic features	5
Clinical measurements	15
Laboratory results	10
Total features	30

VIII. DISCUSSION AND LIMITATIONS

The primary advantage of this system is the Safety-First approach. By acknowledging when the model is "unsure," we mitigate the risks of AI-driven medical malpractice. However, limitations include the computational overhead of running T forward passes for MC Dropout on standard hospital hardware and the potential "utility privacy trade-off" where high noise levels for DP can slightly degrade accuracy.

The formula has the following form;

$$W_{new} = \sum_{k=1}^{K} \frac{CS_k}{\sum CS_i} \Delta \tilde{w}_k$$

IX. FUTURE WORK

1. Explainable AI (XAI): Integrating Grad- CAM heatmaps to show where the model is looking when it is uncertain.
2. Lightweight Models: Exploring Knowledge Distillation to compress the uncertainty-aware model for mobile health applications.
3. Incentive Mechanisms: Developing blockchain-based rewards for hospitals that provide high-quality, low- uncertainty data updates.

X. CONCLUSION

This paper introduced a robust federated learning framework tailored for healthcare applications, emphasizing uncertainty awareness at the global aggregation level. By enabling the global model to account for local uncertainties, the proposed approach enhances predictive accuracy, improves robustness to data heterogeneity, and strengthens the safety of collaborative AI systems. Importantly, it addresses the critical tension between data privacy and clinical reliability, fostering greater trust among participating institutions. Overall, this work contributes a practical and principled step toward deploying privacy- preserving, dependable, and clinically trustworthy federated AI solutions in real- world healthcare settings. Table 1 Table Type Styles.

XI. REFERENCES

Guan, H., Yap, P. T., & Liu, M. (2024). Federated learning for medical image analysis: A survey. Pattern Recognition, 151. https://doi.org/10.1016/j.patcog.2023.110452

Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. ICML. https://proceedings.mlr.press/v48/gal16.html

Kairouz, P., McMahan, H. B., Avent, B., et al. (2021). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14. https://arxiv.org/abs/1912.04977

Zhang, Y., et al. (2025). Uncertainty-aware personalized federated learning for realistic healthcare. Proceedings of Machine Learning Research. https://proceedings.mlr.press/

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. ACM CCS. https://dl.acm.org/doi/10.1145/2976749.2978318

McMahan, B., Moore, E., Ramage, D., Hampson, S., & Agueray Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. AISTATS. https://arxiv.org/abs/1602.05629

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. MLSys. https://arxiv.org/abs/1812.06127

Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H., Albarqouni, S., et al. (2020). The future of digital health with federated learning. NPJ Digital Medicine, 3. https://www.nature.com/articles/s4174 6-020-00323-1