International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

PRIVACY-ENHANCED FEDERATED LEARNING FOR BRAIN TUMOR CLASSIFICATION: AN EVALUATION OF ACCURACY AND PRIVACY

Nada Farhan[1*], Jinan Qamar[2], Ledan Alyahya[3], and Sahar Jambi[4]

[1*,2,3,4] Department of Information Systems, King Abdulaziz University, 21589 Jeddah, Saudi Arabia.
Emails: nabdulazizfarhan@stu.kau.edu.sa (N.F.)[1], 2110001@stu.kau.edu.sa (J.Q.)[2], laalyahya@stu.kau.edu.sa (L.A.)[3]

***Corresponding Author:** Sahar Jambi
*shjambi@kau.edu.sa (S.J.)

Abstract

Federated Learning (FL) is increasingly recognized as a valuable approach in medical imaging. By allowing multiple healthcare institutions to collaboratively train machine learning models on their local data, FL addresses the challenges of data sharing while preserving patient privacy. This decentralized method not only mitigates privacy concerns but also enhances the diversity and robustness of models by leveraging heterogeneous data from different sources, leading to improved performance in critical applications such as disease diagnosis and treatment planning. This work implements FL to classify brain tumor images, enabling model training across multiple sites without direct data exchange. To further enhance privacy, FL is integrated with Differential Privacy (DP), which introduces controlled noise to protect individual data points. The study evaluates three configurations: centralized deep learning (as a benchmark), federated learning, and federated learning integrated with the DP-SGD technique. The dataset used to train the model is brain tumor MRI images, consisting of four distinct classes. The results indicate that the centralized DL model achieved a 97% testing accuracy, comparable to the FL model, which also reached 97% accuracy at the 18th round—confirming the efficiency of collaborative training without data sharing. However, integrating differential privacy (FL+DP) reduced the testing accuracy to 64%, reflecting the expected trade-off between privacy preservation and model performance.

Index Terms- Accuracy, Brain Tumor Classification, CNN, Differential Privacy, Federated Learning

I. Introduction

Brain tumors are among the most serious and life-threatening medical conditions, and early diagnosis of these tumors can improve patient outcomes and increase their chances of survival [2]. In this context, machine learning (ML) uses medical imaging techniques such as magnetic resonance imaging (MRI) to build diagnostic ML models that assist with tumor classification and detection. Building such models requires pooling multi-domain data from different sites [3], which is challenging due to the lack of public datasets, the sensitivity and confidentiality of medical data [4], and the strict privacy protection policies and guidelines followed in Saudi Arabia, such as the Saudi Health Information Exchange (SeHE) Policies, the Personal Data Protection Law (PDPL), and the Saudi Central Board of Accreditation of Healthcare Institutions (CBAHI) [5].

Deep learning (DL) techniques have shown promising results in medical image analysis, particularly in the classification of brain tumors using MRI scans. However, these models depend on collecting and sharing medical data in a centralized repository, which raises significant privacy, security, and regulatory concerns [6].

To address these challenges, Federated Learning (FL) has emerged as a promising distributed ML approach that enables multiple healthcare institutions to collaboratively train a shared model while preserving the privacy of their patients' data [7]. While FL holds promises for maintaining data locality, it is also susceptible to attacks and the risk of data leakage. Implementing privacy-preserving

Received: August 14, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

techniques such as Differential Privacy (DP) ensures robust protection of individual patient data [8]. Therefore, this research aims to develop a privacy-preserving model for classifying brain tumor using an FL framework integrated with DP. By combining these two technologies, we aim to protect patient data while still achieving high classification accuracy on MRI images.

Federated Learning (FL) facilitates collaborative model training among several institutions without necessitating the exchange of raw patient data. Every client locally trains a copy of the centralized model and reports its updates back to the server for aggregation across clients, without disclosing private local data [8]. Although FL maintains data locality, it does not inherently safeguard against privacy issues. In fact, even in cases where attackers do not have direct access to the dataset, users' privacy is at risk — such as in model inversion attacks, where an attacker uses shared model updates to reconstruct training samples [9]. This highlights a significant drawback of FL as a stand-alone privacy-preserving strategy.

Differential Privacy (DP) can be incorporated into the FL process to formally ensure privacy by adding noise to model updates. The implementation of DP demands a compromise between model accuracy and data privacy, as stronger privacy guarantees through noise addition may result in performance degradation [9]. Thus, developing a strategy that effectively integrates FL and DP is essential, but the challenge lies in balancing privacy preservation with high diagnostic accuracy in medical imaging models.

II. Literature Review

1.2 Background

Google first used the term "federated learning" in a paper published in 2016 addressing the issue of data sharing across multiple domains [10]. The paper's suggested solution was enabling multiple devices or clients to collaboratively train a global model without sharing their raw data. The server sends a copy of the global model, represented by the model weights, to each client, and they train it locally using their own private data, then share the model updates for aggregation. Figure 1 illustrates the working process of FL.

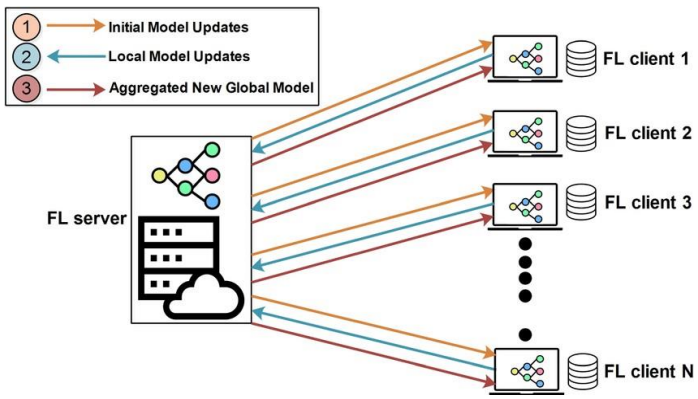

Figure 1: General working process of federated learning [11]

In this context, model updates refer to the learned weights of the neural network, which are adjusted during the training process based on the client's local data.

Incorporating differential privacy (DP) is a novel approach, since FL models alone face the risk of privacy leaks via shared model updates [6]. DP provides a strong mathematical guarantee by adding noise to the data or the computed results before transmission from client to global server. This limits the ability to infer information about individual data points and offers an additional layer of protection. Without incorporating DP, model inversion attacks can be performed, and adversaries reconstruct original training data, such as MRI images, from gradients. Similarly, membership inference attacks

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

reveal whether a specific individual's data contributed to training, while property inference attacks expose hidden attributes or patterns within the dataset without direct access [12]. These risks demonstrate that even without data exchange, sensitive medical information can still be inferred from model parameters.

2.2 Federated Learning in Medical Imaging Classification

A comparative study on the MNIST dataset [13] demonstrated that centralized Convolutional Neural Network (CNN) training achieved 96% accuracy, whereas Federated Learning (FL) utilizing the FedAvg algorithm attained 76% accuracy with a total loss of 1.18, This indicates that FL performance is contingent upon the number of clients and the distribution of data among them.

Another analysis of eight studies [14] using public data demonstrated that FL could achieve performance and interpretability comparable to centralized models, highlighting its efficiency even on basic hardware. In [15], a pre-trained EfficientNet-based FL approach improved medical image classification accuracy to 97.4% for MRI and 98.8% for CT scans, outperforming baseline models in terms of Diagnostic Odds Ratio. Although it highlighted the effectiveness of pre-trained models, it lacks the evaluation of privacy–efficiency trade-offs, leaving room for further research. Similarly, [16] employed a modified VGG16 architecture for brain tumor MRI classification, achieving 98% accuracy while enhancing data privacy. Despite promising results, challenges such as non-IID data, dataset diversity, and scalability remain key areas for future improvement.

2.3 Differential Privacy in Federated Learning

The study [6] applied DP to FL using the BraTS 2018 dataset, examining the trade-off between model performance and privacy. FL achieved segmentation accuracy comparable to centralized models, while DP techniques such as noise addition and gradient clipping enhanced privacy at the cost of slightly reduced accuracy and longer training time. Similarly, [17] proposed a privacy-preserving FL framework for brain tumor segmentation using multi-modal MRI data, achieving a Dice coefficient of 0.89 and specificity of 0.96 as the number of clients increased. Despite strong results, computational overhead, communication cost, and dependency on client data quality and diversity remained challenges.

A survey in [9] reviewed DP applications in DL and FL, covering approaches like noise addition, secure multiparty computation, and homomorphic encryption. It highlighted the ongoing difficulty of balancing privacy strength with model accuracy, as higher noise levels often reduce performance. Lastly, [18] introduced a decentralized FL framework for healthcare networks that uses MQTT protocols for real-time collaboration and has the potential to be used in fully decentralized medical applications that can grow.

2.4 Gap and Contribution

While significant progress has been made in applying Federated Learning (FL) within healthcare, a research gap remains in medical imaging, highlighting the need for further exploration. Limited studies have focused on FL for brain tumor classification, particularly regarding the integration of privacy-preserving techniques to analyze the trade-off between accuracy and privacy. Moreover, although several works demonstrate the potential of FL in healthcare, few examine how fundamental design choices—such as network architecture and update aggregation—affect model performance and privacy in practice. In addition, many studies rely on complex pre-trained models that demand high computational and communication resources, making them less suitable for realistic FL environments. In this research, we developed an FL setup optimized for limited resources by employing a simple CNN architecture designed for brain tumor classification using MRI images. The study employed three methodologies to assess the balance between privacy and model accuracy by adjusting parameters and simulating real-world federated learning environments. Lastly, we examined how the

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

proposed FL model performed compared to a centralized deep learning model, emphasizing the strengths and limitations of each.

III. Methodology

This section outlines the methodology adopted to develop and evaluate a privacy-preserving brain tumor classification framework using MRI images. The methodology involves three learning configurations:

i. Centralized Deep Learning (DL),
ii. Federated Learning (FL), and
iii. Federated Learning integrated with Differential Privacy (FL+DP).

Figure 2 shows the sequential phases of the proposed solution.

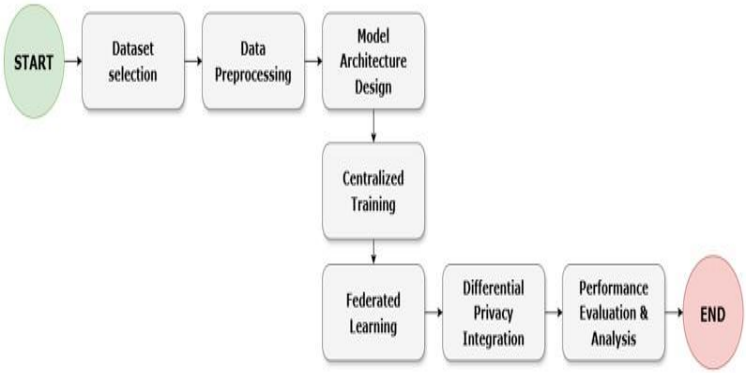

Figure 2: Research Methodology

3.1 Dataset Selection

In the initial phase of our methodology, we focused on gathering data that aligns with our goal of privacy-preserving, multi-institutional FL for brain tumor classification. The Brain Tumor MRI Dataset is publicly available for use on the Kaggle platform [1]. It consists of 7,023 MRI scans, categorized into four distinct classes for the tumor classification task (see Figure 3):

1. Glioma: Malignant tumors originating in glial cells.
2. Meningioma: Typically, benign tumors arising from the meninges.
3. Pituitary: Tumors of the pituitary gland, which can vary from benign to malignant.
4. No Tumor: Normal brain scans without detectable abnormalities

To establish a robust classification task, we chose public data to ensure compliance with ethical standards and avoid using sensitive or proprietary data from specific hospitals.
The dataset is divided into two sets following the 80% training and 20% testing rule, and each set consists of:

- Training set: This set has 5,712 images distributed into the four distinct classes: Pituitary, Meningioma, Glioma, and No Tumor.
- Testing set: This set has 1,311 images that are also distributed into the four distinct classes for tumor classification tasks.

Figure 4 shows the distribution of classes across the two sets.

Received: August 14, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

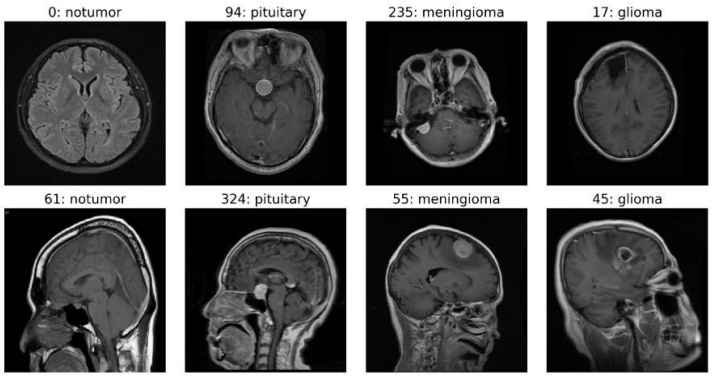

Figure 3: Sample MRI images from the dataset

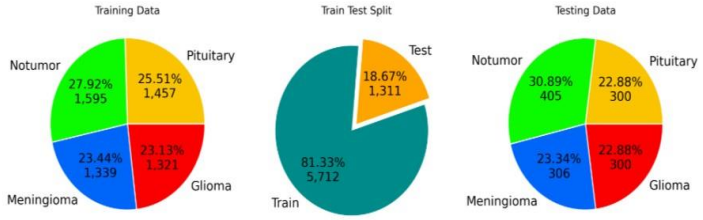

Figure 4: Dataset and class distribution

3.2 Data Preprocessing

To address the variations in imaging formats and quality, we applied standard preprocessing techniques to ensure consistency:

- Resizing: Each image is resized to a fixed dimension (H, W) to standardize input size for the CNN (H = 224 and W = 224 pixels).
- Normalization: Standardized pixel values using ImageNet's RGB mean and standard deviation with mean = [0.485, 0.456, 0.406] and std = [0.229, 0.224, 0.225].
- Data Augmentation: For each image, we applied a horizontal flip with a probability of 0.5 to increase dataset diversity and rotation within a $\pm 10^{o}$ range to improve rotation invariance.
- Conversion to Tensor: In the preprocessing pipeline, the images are converted into PyTorch tensors and scaled pixel values to a [0, 1] range.

3.3 Model Architecture Design

A CNN model was designed for multi-class classification of brain tumor MRI images into four classes: Glioma, Meningioma, Pituitary, and No Tumor. The choice of a simple, lightweight CNN was motivated by previous work in [19], which studied the performance of pre-trained models, like ResNet18 and VGG16, and custom CNNs. Although pre-trained models showed slightly higher accuracy and metrics, custom CNNs offer advantages in terms of computational simplicity, efficiency, and resource usage. These advantages make custom CNNs suitable for tasks with limited computational resources or requirements for faster training and inference.

3.4 Centralized Training

The custom CNN model was first trained in a centralized DL manner to serve as a benchmark for comparison against the FL setup. With centralized training, the model is directly trained on the entire dataset because all the data examples are located in a single location, simulating a traditional, fully accessible data scenario. Figure 5 illustrates the workflow for setting up a centralized DL.

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

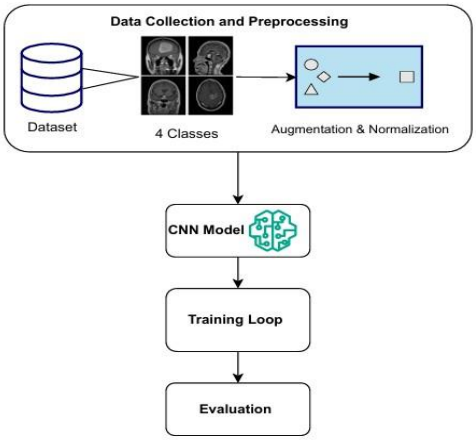

Figure 5: Workflow of the proposed model in a centralized setup

3.5 Federated Learning Setup

The main objective is to study the performance of the distributed learning approach through training our custom CNN model under the FL setup. In the FL simulation, the training data was split across a number of clients using Dirichlet distribution-based sampling, since it is a suitable method used when simulating non-IID (non-independent and identically distributed) data across clients. Specifically mimicking class imbalance—a type of non-IID scenario in which the distribution of the four tumor classes is deliberately made uneven among clients. This feature simulates real-world distributed learning settings in which participating nodes typically do not have equal access to data and class distributions.

After partitioning data among the clients, each client trains their dataset locally, sending the model updates (weights) to a central server for aggregation afterwards.

The Federated Averaging (FedAvg) method is used to aggregate these updates. It works at the end of each round, where it calculates a weighted average of the updated weights of all clients, where the weight is proportional to the number of data samples held by each client. This guarantees that clients with larger datasets have a higher impact on the updated global model.

This approach protects data privacy by preserving the data locality while operating like real-world scenarios in which gathering data at a central point is impractical or restricted. Figure 6 illustrates the workflow of the proposed model in the FL setup.

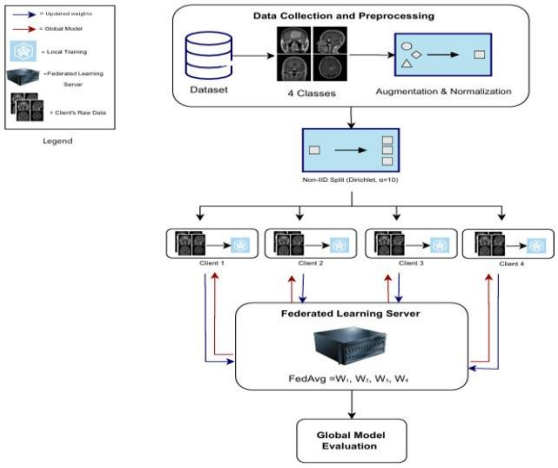

Figure 6: Workflow of the proposed model in a distributed FL setup

Received: August 14, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.6 Differential Privacy Integration

Integrating DP reduces the risk of data leakage and attacks, providing formal privacy guarantees. We integrated Differentially Private Stochastic Gradient Descent (DP-SGD) into our FL setup.

DP-SGD is a privacy-preserving optimization method that helps prevent data leakage through gradients (model updates), and it includes two main steps:

1. Gradient Clipping (regularization technique)
   Each gradient is minimized to a maximum norm before the gradients' averaging. This ensures the limited influence of any single data point on the overall model update.
2. Adding Gaussian Noise (to a group of clients)
   After all gradients are clipped and aggregated, a specific amount of Gaussian noise is added before sending them to the model. The added noise makes it statistically difficult for adversaries to reverse-engineer and infer details about specific data points from the shared gradients.

Figure 7 illustrates the workflow of the proposed model in the FL+DP setup, with the addition of an extra layer of protection based on differential privacy.

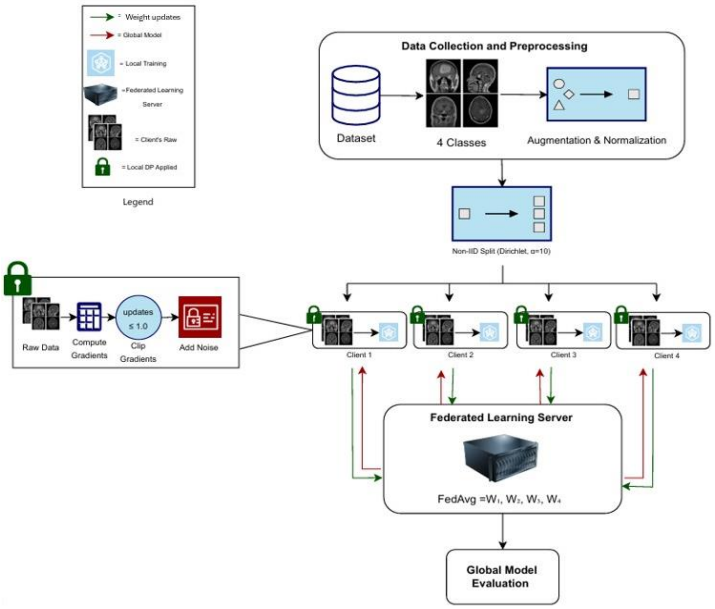

Figure 7: Workflow of the proposed model in a distributed FL+DP setup

3.7 Performance Evaluation

To evaluate the performance and privacy guarantees of our custom CNN architecture across the three setups—DL, FL, and FL+DP—we employed several general evaluation metrics and privacy-specific metrics, applying them to all setups to ensure a fair and clear comparison.

3.7.1 Performance Metrics

- Accuracy: Measures the percentage of images that are correctly classified out of the total number of images.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

- Precision: Assesses the proportion of true positive predictions out of all positive predictions. It indicates how accurately the model identifies a specific tumor type.

Received: August 14, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

- Recall (Sensitivity): Reflects the model's ability to correctly identify all actual positives. It emphasizes detecting instances of a particular tumor type without missing any.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- F1-score: Combines precision and recall, providing a balanced measure, especially useful in cases with class imbalance.

$$\text{F1-score} = \frac{2\times(\text{Precision}\times\text{Recall})}{\text{Precision}+\text{Recall}}$$

- Loss: This is a numerical value that represents the error occurring during training. It quantifies how well or poorly the model is performing; lower loss values indicate better predictions, and vice versa.

3.7.2 Privacy Metrics

For the FL+DP setup, we evaluated the privacy guarantee by tracking the privacy budget, epsilon (ε), which measures the privacy loss occurring throughout the training process:

- Clients (ε) a value recorded after each local training epoch to track the privacy loss of each individual client.
- Global (ε) a value recorded after each round of training to track the overall privacy loss of the global model.

Lower epsilon represents a higher privacy guarantee, and vice versa. This metric is used to analyze privacy and performance trade-off between FL and FL+DP setups.

IV. Experiment and Results

This section presents the experimental setup and results of applying Deep Learning (DL), Federated Learning (FL), and Federated Learning combined with Differential Privacy (FL+DP) for brain tumor classification.

Firstly, outlining the experimental design, including the implementation environment, data preparation, model architecture, hyperparameter settings, and training processes for each setup.

Secondly, presenting the experimental results, visualizations, and performance metrics, followed by a comparative analysis between the three configurations.

Finally, discussing the trade-off between privacy and performance, highlighting the effectiveness of FL and the impact of integrating differential privacy.

4.1 Experiment Design

4.1.1 Environment Setup

The research was implemented in Python using a mix of local devices, Kaggle, and Google Colab due to limited hardware resources. Some procedures were run on CPU and others on GPU, depending on availability and computational requirements. The PyTorch framework was used to build the custom CNN model, with the Opacus library integrated to support Differential Privacy (DP). Additional Python modules such as scikit-learn, matplotlib, seaborn, pandas, and tqdm were used for data processing, performance evaluation, and result visualization.

Received: August 14, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4.1.2 Data Transformation

As previously mentioned in the methodology section, the data preprocessing pipeline was applied to all MRI images using PyTorch's transforms module to standardize input dimensions and enhance model generalization.

4.1.3 CNN Brain Tumor Classification Model

The proposed custom CNN model consists of the following components:

- Input Layer: Accepts color MRI images resized to 224 × 224 pixels with three channels (RGB).
- Convolutional Layers: Two convolutional layers for hierarchical feature extraction, each followed by ReLU activation and max-pooling to reduce spatial dimensions.
  - Conv1: 16 filters, kernel size 3 × 3, padding 1.
  - Conv2: 32 filters, kernel size 3 × 3, padding 1.
- Max-Pooling Layers: Each convolutional layer is followed by MaxPool2d (kernel size 2 × 2, stride 2), halving feature map dimensions.
- Fully Connected Layers: A two-layer fully connected classifier processes the extracted features to perform the classification.
  - FC1: 128 units with ReLU activation.
  - FC2: 4 units (output layer for four tumor classes).

This lightweight yet sufficiently deep architecture was designed for the FL simulation, capturing key spatial features in MRI images while enabling fast training under limited computational resources.

4.1.4 Centralized DL Setup

In this stage, the proposed CNN model—BrainTumorCNN— was trained in a centralized deep learning setup to serve as a benchmark for comparison with the FL configuration. The model was trained on the entire dataset, where all data was stored in a single location.

- Data Loaders and Batching:
  PyTorch's DataLoader utility was used to shuffle and divide the dataset into manageable batches before training.
  - Batch Size: 32, to balance training stability and computational efficiency.
  - Shuffling: Applied at the start of each epoch to improve generalization and prevent the model from learning from data order.
- Training Configuration:
  - Loss Function: We used Cross-Entropy Loss as the loss function that calculates the classification error between the model's predicted labels and the true labels.
  - Optimizer: Adam, commonly used for image classification.
  - Learning Rate: 0.001 (lower values such as 0.0001 slowed training without improvement).
  - Epochs: The number of epochs represents the number of times that the model passes through the entire dataset. The model was trained for 10 epohcs to balance computational cost and model learning.

Table 1 summarizes the training hyperparameters for centralized deep learning.

Table 1: Centralized Deep Learning Training Hyperparameters

Parameter	Value
Optimizer	Adam
Learning Rate	0.001
Batch Size	32
Epochs	10
Loss Function	Cross-Entropy Loss

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Input Image Size	224 224

4.1.5 Federated Learning Setup

In this setup, we used our custom BrainTumorCNN model and trained it in a simulated FL setting, with several clients sharing the training data and a central server aggregating model updates. This approach was designed to reflect real-world distributed scenarios that protect data privacy by protecting data locality.

- Number of Clients: Four clients were selected due to limited computational resources, but experiments were also conducted with 8, 16, and 32 clients to evaluate the scalability and stability of the FL setup.
- Data Splitting: Training data were partitioned among clients using a Dirichlet distribution-based sampling method to simulate non-IID conditions by introducing class imbalance. The alpha parameter controlled the degree of imbalance—lower values produced more skewed distributions, while higher values yielded more balanced ones. A moderate value of $\alpha = 10$ was used as the main setting, with $\alpha = 3$ and $\alpha = 5$ tested for comparison under varying heterogeneity levels.
- Data Loaders and Batching:
  Two data loaders were created.
  - Federated Clients: Each client's loader handled shuffling and batching its local dataset with a reduced batch size of 16 to fit resource limitations.
  - Testing Dataset: Used for evaluating the global model after each federated round.
- Training Configuration:
  We made sure that all other training parameters from the model, optimizer, learning rate, loss function, and local epochs are equivalent to the ones used in the centralized DL setup to ensure fair comparison conditions.
  Table 2 summarizes the hyperparameters used during the FL process.
- Federated Training Process:
  The FL training was performed for 20 rounds and consisted of two main steps:
  - Local Training: At the beginning of each round, each client received a copy of the global model weights and trained the model locally on its dataset for 10 epochs. This allowed each client to update model weights based on its unique data characteristics.
  - Model Aggregation (FedAvg): After local training, clients sent their updated weights to the central server. The server then applied the Federated Averaging (FedAvg) algorithm to compute a weighted average of all client updates, producing an updated global model that was redistributed to the clients for the next round.
- Evaluation:
  Model evaluation was conducted at the end of each round in two stages:
  - Before Aggregation: Each client was evaluated on its local dataset.
  - After Aggregation: The global model was evaluated on the test set using key performance metrics, including accuracy, precision, recall, and F1-score.

Table 2: FL Training Hyperparameters

Parameter	Value
Optimizer	Adam
Learning Rate	0.001
Batch Size	16
Local Epochs per Round	10
Federated Rounds	20
Loss Function	Cross-Entropy Loss
Number of Clients	4, 8, 16, 32

Received: August 14, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Data Split Method	Dirichlet
Alpha Values	3, 5, 10
Aggregation Method	FedAvg
Input Image Size	224 224

4.1.6 Federated Learning + Differential Privacy Setup

To add an additional privacy-preserving layer to the Federated Learning (FL) setup, Differentially Private Stochastic Gradient Descent (DP-SGD) was integrated at the client side using Opacus, a library developed for training PyTorch models with differential privacy.

- Training Configuration:
  The optimizer algorithm was changed from Adam to DP-SGD with a momentum value of 0.9. All other hyperparameters were kept consistent with the standard FL setup. Table 3 summarizes all hyperparameters used in this configuration.
- DP Mechanism Application:
  Each client trained its model locally using DP-SGD, where the privacy engine operated as follows:
  - Gradient Clipping: Each gradient was clipped to a maximum L2 norm of 1.0. If any data sample's gradient exceeded this threshold, it was scaled down to ensure that no single data point could cause a significant model update.
  - Gaussian Noise: After clipping, Gaussian noise was added to the gradients before transmission to the server. The noise scale was set by a noise multiplier of 0.8. A higher noise multiplier increases privacy protection but may reduce model performance.
  - Target Delta (δ): Set to 1e−5, representing the acceptable probability of violating the differential privacy guarantee.
- Evaluation and Privacy Budget Tracking:
  Model evaluation was conducted at the end of each round in two stages:
  - Before Aggregation: Each client was evaluated on its own local dataset.
  - After Aggregation: The global model was evaluated on the test set using accuracy, precision, recall, and F1-score metrics.

  In addition, the privacy budget (ε) was tracked as follows:
  - Local Epsilon (ε): Recorded after each client's local training to measure per-client privacy loss.
  - Global Epsilon (ε): Calculated after each round to represent the overall privacy loss of the aggregated global model.

  Epsilon (ε) quantifies the privacy loss, where smaller values indicate stronger privacy guarantees. These values were later used to analyze the trade-off between privacy and performance and to compare the FL+DP setup with the standard FL configuration.

Table 3 summarizes the hyperparameters used during the FL+DP process. The model with the best performance at a specific round was saved for comparison with the centralized DL setup.

Table 3: FL+DP Training Hyperparameters

Parameter	**Value**
Optimizer	DP-SGD
Momentum	0.9
Learning Rate	0.001
Batch Size	16
Local Epochs per Round	10
Federated Rounds	20
Loss Function	Cross-Entropy Loss

Received: August 14, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Max Gradient Norm (Clipping)	1.0
Noise Multiplier (σ)	1.2
Target Delta (δ)	1×10^{-5}
Data Split Method	Dirichlet
Aggregation Method	FedAvg

4.1 Experiment Results

4.1.1 Centralized DL Results

The centralized model achieved 98% training and 97% testing accuracy, showing stable learning and satisfactory generalization (Figure 8). The precision, recall, and F1-score were consistently high across all tumor classes, confirming strong overall performance. As shown in the confusion matrix (see Figure 9), the model accurately classified No Tumor and Pituitary cases, while most misclassifications occurred between Glioma and Meningioma, likely due to visual similarities. These results establish a solid baseline for comparison with the FL setups.

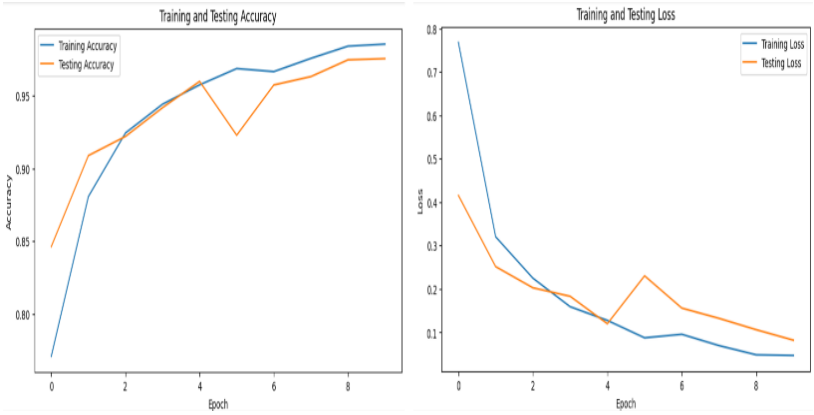

Figure 8: DL model's accuracy and loss over 10 epochs

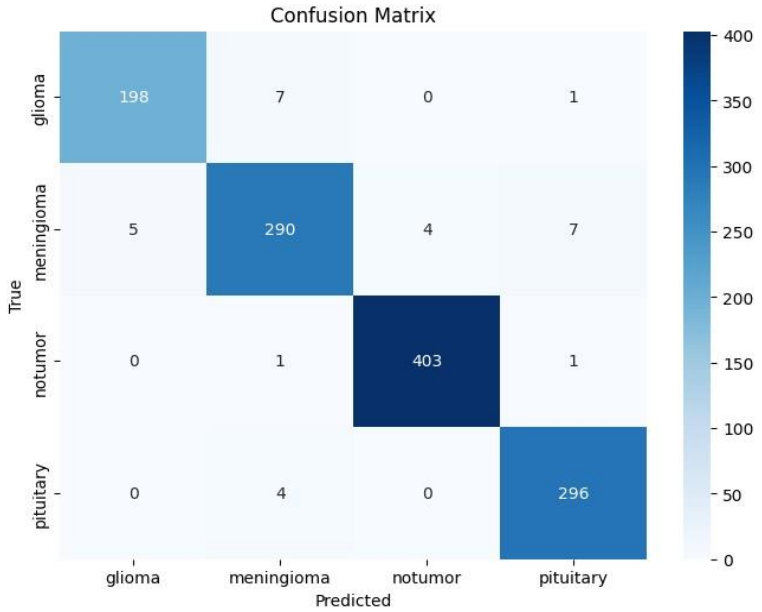

Figure 9: Confusion matrix of DL model over 10 epochs

Received: August 14, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4.1.2 FL Results

The global model in the federated setup achieved its best results after several communication rounds, with 96% training accuracy and 97% testing accuracy. As shown in Figure 10, the model's accuracy and loss improved steadily over 20 rounds, indicating effective aggregation and consistent learning across clients. The confusion matrix of the FL model over 20 rounds (see Figure 11) demonstrates strong classification performance with high precision, recall, and F1-scores, closely aligning with the centralized model.

The impact of data imbalance was examined using the Dirichlet alpha parameter. As illustrated in Figure 12 and detailed in Table 4, class distributions became more balanced as the alpha value increased, resulting in improved model stability and higher overall performance. Performance across different numbers of clients is summarized in Table 5, which shows that increasing the number of clients led to a gradual decline in accuracy due to smaller local datasets and stronger non-IID effects.

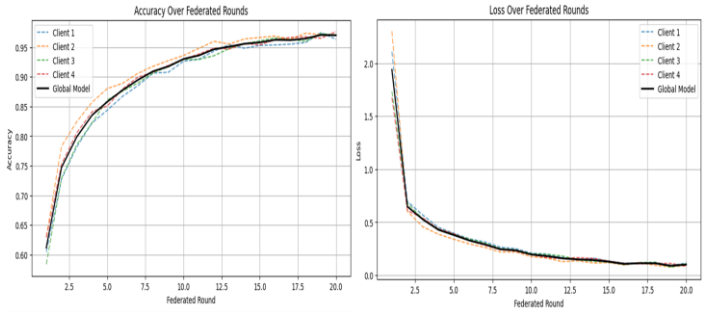

Figure 10: FL model's accuracy and loss over 20 rounds

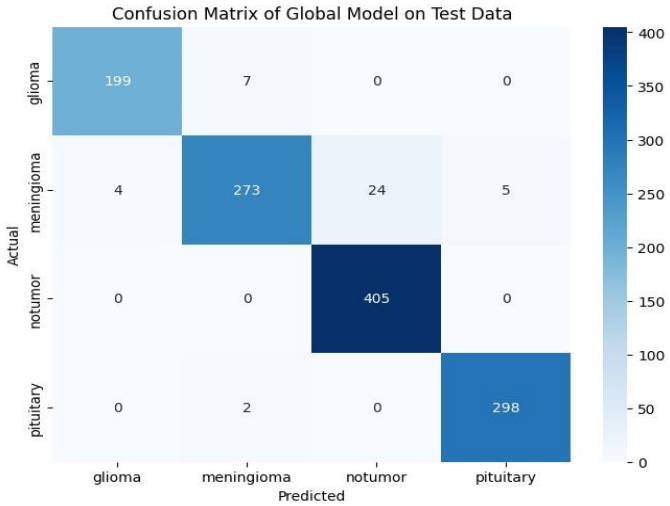

Figure 11: Confusion matrix of FL model over 20 rounds

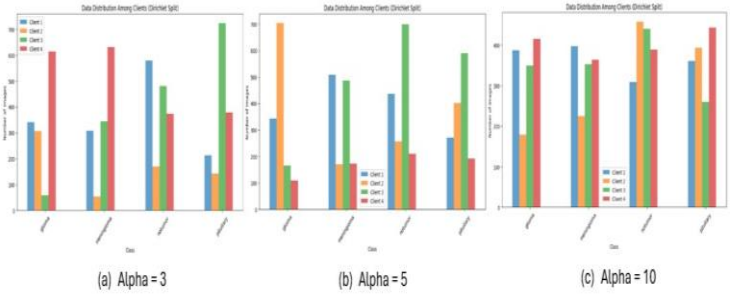

Figure 12: Class distribution at different alpha values

Received: August 14, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 4: Performance across different alpha values

Alpha	Accuracy	Precision	Recall	F1-Score
3	0.9542	0.9540	0.9533	0.9566
5	0.9565	0.9570	0.9565	0.9564
10	0.9729	0.9732	0.9728	0.9726

Table 5: Performance across different numbers of clients

Clients	Accuracy	Precision	Recall	F1-Score
4	0.9729	0.9732	0.9728	0.9726
8	0.9359	0.9357	0.9359	0.9354
16	0.8726	0.8694	0.8726	0.8692
32	0.7429	0.7456	0.7429	0.7449

4.1.3 FL+DP Results

To add an extra privacy-preserving layer, Differentially Private Stochastic Gradient Descent (DP-SGD) was integrated into the FL model with a maximum gradient norm (clipping) of 1.0 and a Gaussian noise multiplier of 1.2. At its best training round, the model achieved 69% training accuracy and 64% testing accuracy, showing the expected decline in performance caused by noise addition.

As illustrated in Figure 13, the model's accuracy and loss over 20 rounds improved gradually but remained lower than the standard FL setup, with slight fluctuations after round 10.

The confusion matrix of the FL+DP model over 20 rounds (Figure 14) shows that while the model performed well on No Tumor and Pituitary classes, it struggled with Meningioma, reflecting the trade-off between privacy and accuracy.

To assess the impact of different noise levels, the noise multiplier was reduced from 1.2 to 0.8. As shown in Figure 15, the accuracy remained almost unchanged, while the privacy budget (ε) increased—indicating weaker privacy protection. The model with noise = 1.2 achieved stronger privacy (lower ε) but reduced accuracy, whereas noise = 0.8 offered weaker privacy without performance gains.

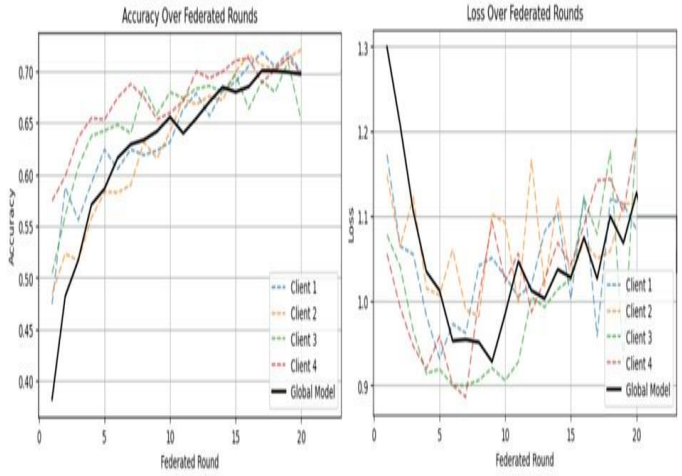

Figure 13: FL+DP model's accuracy and loss over 20 rounds

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

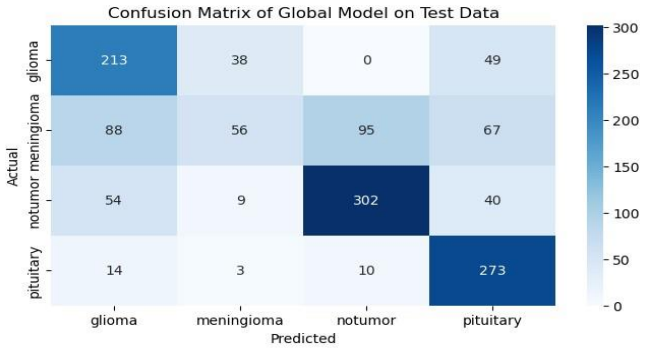

Figure 14: Confusion matrix of FL+DP model over 20 rounds

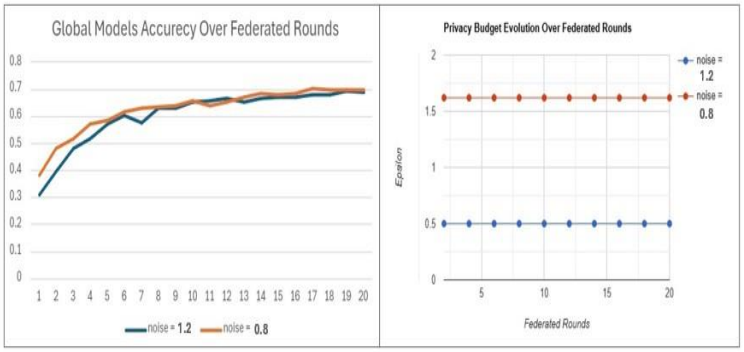

Figure 15: Models' accuracy and epsilon at different noise amounts

V. Conclusion

Federated Learning (FL) allows multiple institutions to collaborate on training a model without exchanging raw data, thereby protecting privacy and utilizing diverse datasets. This project evaluated FL as an alternative to centralized Deep Learning (DL) for brain tumor classification using MRI images and integrated Differential Privacy (DP) to enhance data protection. DP-SGD was applied on the client side to ensure that individual data points could not be inferred from shared gradients.

The experiments demonstrated that FL achieved performances comparable to centralized DL, with both reaching around 97% accuracy, while maintaining data locality. However, when DP was integrated, accuracy dropped to 64%, reflecting the trade-off between privacy and model performance. Despite this decline, FL+DP successfully improved privacy protection, confirming the potential of combining these methods for sensitive medical applications.

Several challenges were encountered, including limited computational resources and smaller local datasets when distributing data among multiple clients, which affected generalization. However, the findings revealed that having fewer clients with balanced data led to improved performance.

Future research should concentrate on enhancing dataset diversity, investigating more sophisticated federated learning algorithms like FedProx, and fine-tuning differential privacy parameters to achieve a better balance of accuracy and privacy. Overall, the study confirms that FL—especially when supported by DP—holds strong promise for secure and privacy-preserving medical image analysis.

VI. Acknowledgment

We would like to express our gratitude to the members of the faculty of computer science and information technology at King Abdulaziz University for their support and guidance. We also extend our sincere thanks to Dr. Yousef Al-Sinani for his valuable advice regarding the methodology of preprocessing data, which greatly contributed to the quality and direction of this study.

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

VII. References

[1] Guslovesmath, "Brain Tumor MRI Dataset," Kaggle, 2024. [Online]. Available: https://www.kaggle.com/code/guslovesmath/cnn-brain-tumor-classification-99-accuracy

[2] Q. Mastoi, S. Latif, M. Shahid, R. Ullah, "Explainable AI in Medical Imaging: An Interpretable and Collaborative Federated Learning Model for Brain Tumor Classification," Frontiers in Oncology, vol. 15, Art. 1535478, 2025. doi:10.3389/fonc.2025.1535478.

[3] H. Guan and M. Liu, "Federated Learning for Medical Image Analysis: A Survey," arXiv preprint arXiv:2306.05980, 2023. doi: 10.48550/arXiv.2306.05980.

[4] 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, p. 1953, 2022.

[5] Central Board for Accreditation of Healthcare Institutions (CBAHI), "CBAHI National Standards for Healthcare Institutions," Saudi Ministry of Health, 2025. [Online]. Available: https://portal.cbahi.gov.sa/english/cbahi-standards

[6] W. Li, F. Milletari, D. Xu, N. Rieke, H. Roth, S. Albarqouni, S. Bakas, M. Cardoso, A. Feng, and Z. Xu, "Privacy-Preserving Federated Brain Tumour Segmentation," in *Machine Learning in Medical Imaging*, LNCS, vol. 11861, pp. 133-141, Springer, Cham, 2019. doi:10.1007/978-3-030-32692-0_16.

[7] N. Sivakumar and A. Raza Khan, "A Hybrid Brain Tumor Classification Using Federated Learning with FedAvg and FedProx for Privacy and Robustness Across Heterogeneous Data Sources," IEEE Access, 2025. doi:10.1109/ACCESS.2025.3549440.

[8] M. E. Yahiaoui, M. Derdour, and M. A. Sarem, "Federated Learning with Privacy Preserving for Multi-Institutional 3D Brain Tumor Segmentation," 2024. [Prerprint]

[9] A. E. Ouadrhiri and A. Abdelhadi, "Differential Privacy for Deep and Federated Learning: A Survey," IEEE Access, vol. 10, pp. 22359–22373, 2022. doi:10.1109/ACCESS.2022.3151670.

[10] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in Proc. 20th International Conf. on Artificial Intelligence and Statistics (AISTATS), vol. 54, pp. 1273-1282, Fort Lauderdale, FL, USA, Apr. 2017.

[11] J. D. Qammar and H. Ning, "Federated Learning Attack Surface: Taxonomy, Cyber Defences, Challenges, and Future Directions," Artificial Intelligence Review, vol. 55, pp. 3569–3606, 2022. doi:10.1007/s10462-021-10098-w.

[12] Y. Zhao, C. Jing, and W. Zhang, "Differential Privacy Preservation in Deep Learning: Challenges, Opportunities and Solutions," IEEE Access, vol. 7, pp. 48901–48934, 2019. doi:10.1109/ACCESS.2019.2921417.

[13] F. Mar'i, A. Supianto, and F. Bachtiar, "Comparison of Federated and Centralized Learning for Image Classification," *PIKSEL: Penelitian Ilmu Komputer Sistem Embedded and Logic*, vol. 11, no. 2, pp. 393–400, 2023. doi:10.33558/piksel.v11i2.7367.

[14] A. Sadilek, L. Liu, D. Nguyen, M. Kamruzzaman, S. Serghiou, B. Rader, A. Ingerman, S. Mellem, P. Kairouz, E. Nsoesie, J. MacFarlane, A. Vullikanti, M. Marathe, P. Eastham, J. Brownstein, B. Arcas, M. Howell, and J. Hernandez, "Privacy-First Health Research with Federated Learning," npj Digital Medicine, vol. 4, p. 132, 2021. doi:10.1038/s41746-021-00489-2.

[15] P. N. Srinivasu, G. J. Lakshmi, and M. F. Ijaz, "Enhancing Medical Image Classification via Federated Learning and Pre-Trained Model," *Egyptian Informatics Journal*, vol. 27, no. 47, 2024. doi:10.1016/S1110-8665(24)00108-7.

[16] E. Albalawi, M. Mahesh, G. S. Basha, and A. Almusharraf, "Integrated Approach of Federated Learning with Transfer Learning for Classification and Diagnosis of Brain Tumor," *BMC Medical Imaging*, 2024. doi:10.1186/s12880-024-01261-0.

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[17] F. Ullah, M. Nadeem, M. A. Faisal, and S. Khan, “Enhancing Brain Tumor Segmentation Accuracy through Scalable Federated Learning with Advanced Data Privacy and Security Measures,” *Mathematics*, vol. 11, no. 4189, 2023. doi:10.3390/math11194189.

[18] R. Camajori Tedeschini, S. Savazzi, L. Serio, R. Sorrentino, and M. Nicoli, “Decentralized Federated Learning for Healthcare Networks: A Case Study on Tumor Segmentation,” IEEE Access, vol. 10, pp. 8693–8706, 2022. doi:10.1109/ACCESS.2022.3141913.

[19] A. K. V. Md. Ashik Khan, “Comparative Analysis of Resource-Efficient CNN Architectures for Brain Tumor Classification,” 2024. [Preprint]

Received: August 14, 2025