RESEARCH ARTICLE

Decentralized and Explainable Brain Tumor Segmentation: A Blockchain-Secured Federated Learning Approach

Aiza Mukhtar[1], **Aamir Ali**[1,*], **Wajiha Farooq**[1], **Muhammad Aqib Fareed**[2], **Sadia Sohail**[3] and **Azka Ahsan**[2]

[1] Department of Computer Science, COMSATS University Islamabad, Sahiwal Campus, Sahiwal 57000, Pakistan
[2] Department of Computer Science, Government College University Faisalabad, Sahiwal Campus, Sahiwal 57000, Pakistan
[3] Department of Computer Science, University of Okara, Okara 56300, Pakistan

Abstract

Brain tumor segmentation plays an important role in clinical diagnosis, treatment planning, and neuro-oncology assessment using multimodal Magnetic Resonance Imaging (MRI) data. However, conventional centralized deep learning systems often face limitations associated with patient data privacy, secure inter-institutional collaboration, and limited model interpretability. This study presents a decentralized and privacy-preserving brain tumor segmentation framework that integrates Federated Learning (FL), a blockchain-inspired audit and coordination mechanism, and Explainable Artificial Intelligence (XAI) within a collaborative medical imaging environment. A 3D U-Net architecture was trained on the BraTS 2020 dataset under a simulated multi-institutional federated setting using non-IID MRI data distributions. Federated Averaging (FedAvg) was employed for global model aggregation, while a blockchain-inspired permissioned ledger mechanism was used to record model update hashes and aggregation metadata for auditability across communication rounds. Grad CAM and LIME were incorporated to provide interpretable visualization of tumor related regions contributing to segmentation predictions. Experimental evaluation demonstrated stable convergence behavior with an overall voxel accuracy of 98.52%, a weighted F1 score of 98.3%, and Dice coefficient improvement from 0.752 to 0.830 during federated training. The generated segmentation outputs showed strong agreement with manually annotated tumor regions while preserving decentralized data privacy. The proposed framework contributes a secure, interpretable, and collaborative segmentation pipeline for trustworthy brain tumor analysis across distributed healthcare environments.

Keywords: brain tumor segmentation, federated learning, blockchain-inspired audit logging, explainable AI, privacy

Submitted: 26 August 2025
Revised: 01 June 2026
Accepted: 04 June 2026
Published: 17 August 2026

Vol. 2, **No.** 3, 2026.
10.62762/JIAP.2026.354772

***Corresponding author:**
Aamir Ali
amirali4436823@gmail.com

Citation
Mukhtar, A., Ali, A., Farooq, W., Fareed, M. A., Sohail, S., & Ahsan, A. (2026). Decentralized and Explainable Brain Tumor Segmentation: A Blockchain-Secured Federated Learning Approach. *ICCK Journal of Image Analysis and Processing*, 2(3), 168–189.

preserving AI, medical diagnostics.

1 Introduction

One of the most severe and complex forms of neurological disorders is brain tumors, which are responsible for a considerable amount of global disability and mortality. For a better quality of life and more efficient planning of surgical interventions, early and accurate definition of the border and subdivisions of a brain tumor is needed. Traditional methods of analysis have several issues. For example, reliance on centralized data stores which could present some privacy concerns, and the use of "black box" AI algorithms which are difficult to decipher, leading to lower trust and adoption by clinicians [1, 2]. Figure 1 illustrates the anatomical structure of the human brain and highlights the representative location of a brain tumour within the cerebral region.

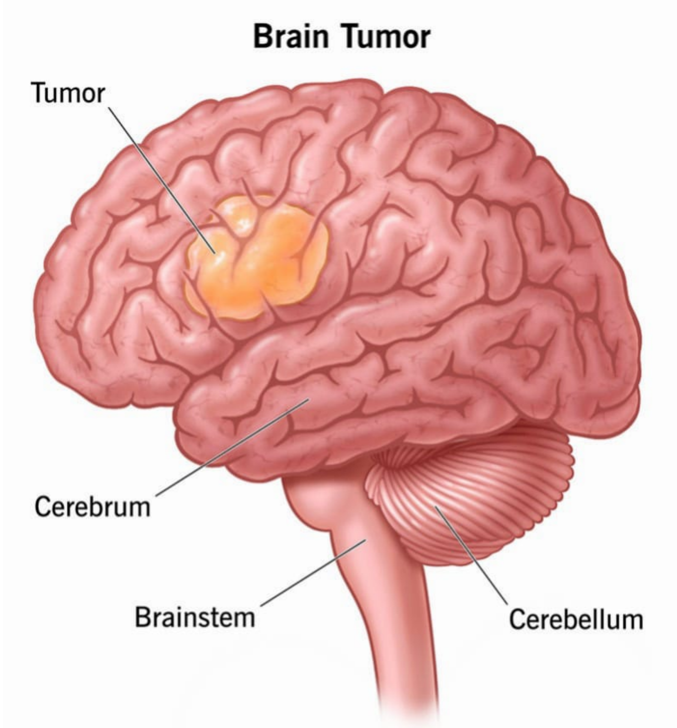


Figure 1. Anatomical illustration of the human brain showing the location of a brain tumour and major brain regions including the cerebrum, cerebellum, and brainstem.

The application of cutting-edge deep learning techniques to medical images — specifically Magnetic Resonance Imaging (MRI) scans of brain tumors — has produced promising results [3]. However, most existing AI solutions are built under the assumption that patient data can be uploaded to centralized servers, which conflicts with stringent privacy regulations like Health Insurance Portability and Accountability Act (HIPAA) and General Data Protection Regulation (GDPR). Moreover, these systems are vulnerable to impaired diagnostic security threats that include model poisoning and data breaches.

There is a troubling 'lack of trust' paradigm with deployed conventional AI models; particularly with deep neural networks. Underpinned by security and explainability of models, such trust loss impedes their widespread adoption without reservations. AI models legally and clinically require centralized datasets for training, meaning that such systems pose serious risks to data privacy and may lead to unauthorized exposure of sensitive patient information [4].

Federated Learning (FL) is a decentralized machine learning technique which trains models across multiple data sources without the exchange or centralization of the raw underlying data, thereby retaining data privacy. Constructed around an iterative centralization process, a central server provides a global model to client institutions (e.g., hospitals). Each client intrinsically trains this model privately on their data. Rather than data, model changes (e.g., weight gradients) are returned. These are consolidated by the server to constitute an updated global model. This process is cyclical, the model learning from all participants, while the locally sensitive data remains locked [5].

Blockchain concepts can be incorporated into federated learning environments to improve auditability and traceability of model update records [6, 7]. In healthcare settings, blockchain-inspired ledgers may provide immutable logging of aggregation events and model update histories, supporting transparency and accountability during collaborative AI development. However, practical deployment often introduces challenges related to scalability, consensus management, and infrastructure overhead [6].

Effectively and accurately segmenting brain tumors on MRI scans is still an unresolved area in medicine and remains an important challenge in clinical practice. Most existing AI systems use centralized systems which greatly understate or ignore the patient privacy and legal obligation concerns. In attempting decentralized systems, the processes of 'federation' or 'aggregating' still pose security challenges such as model poisoning and malicious tampering which are diagnostic and clinically devastating. The unique opaque character of deep learning models results in an interpretability problem, which is another barrier to clinical adoption, as clinicians will not use systems they do not trust. In addition to the issues stated, the problems of computational burden and data diversity across institutions will further limit the scalability and robustness of solutions. Each of these challenges underscores the urgent need to address privacy, security, and explainability in brain tumor segmentation frameworks in order to facilitate their integration into everyday clinical practice.

The combination of privacy concerns and the need for a centralized approach to AI leads to problems related to HIPAA and GDPR in the context of MRI datasets. The proposed framework combines federated learning with a blockchain-inspired audit and coordination mechanism to improve traceability during collaborative model training. Federated learning enables multiple hospitals to train a shared AI model without centralized data pooling, while the audit layer records model update information and aggregation metadata to support transparency throughout the training process [8]. With federated learning, hospitals train local models on their own data and transfer only model updates, keeping patient MRI data in their possession. In this context, a blockchain-inspired permissioned ledger is used to record model update hashes and aggregation metadata generated during each federated learning round.

Although federated learning preserves raw data privacy, aggregation records and model update histories still require transparent tracking mechanisms. To support traceability, the proposed framework employs a blockchain-inspired permissioned ledger that records hashed model updates, aggregation metadata, performance metrics, and encrypted checkpoints generated during federated training. These records provide an auditable history of model evolution throughout collaborative learning.

At the conclusion of each federated round; pivotal information including the hashed values of client contributions, the aggregated global model, performance metrics, and encrypted checkpoints, is captured and sealed within a block. This provides a traceable record of model lineage and aggregation history throughout collaborative training. This level of transparency fosters trust among the collaborating partners, a fundamental requirement within the medical domain, where model performance is of utmost importance to patient safety.

Explainable AI (XAI) helps make advanced AI tools clinically applicable [2, 9]. While deep learning algorithms for tumor segmentation are accurate, they are often opaque, causing distrust among clinicians [1, 2]. To address this interpretability gap, Grad-CAM [10] was employed to provide visual explanations of model predictions by generating heatmaps that highlight the MRI regions most relevant to segmentation decisions. Federated learning has been widely explored as a privacy-preserving training paradigm for distributed healthcare applications [11]. Based on the identified research gaps and objectives of this study, the following research questions were formulated.

- How can inter-institutional collaboration be established for the development of an AI model for brain tumor segmentation while maintaining data privacy regulations?
- How can we leverage the principles of blockchain to safeguard and verify the security of model updates that are shared within the network?
- How can we utilize explainable AI to demonstrate to medical practitioners the underlying rationale of the AI's outcomes in segmentation analysis?
- What are the relative accuracy, security, efficiency, and overall integration of our system in comparison to existing alternatives?

This study proposes a decentralized and interpretable brain tumor segmentation framework that combines Federated Learning, blockchain assisted secure coordination, and Explainable Artificial Intelligence within a unified medical imaging environment. Unlike conventional centralized segmentation systems, the proposed framework enables collaborative multi-institutional model training without direct sharing of sensitive MRI data. The methodological contribution of this work lies in the integration of secure federated segmentation, blockchain-based auditability, and explainable prediction analysis within a single privacy-preserving architecture for brain tumor analysis.

The proposed framework introduces a blockchain-inspired audit and coordination mechanism that records model update hashes and aggregation metadata to improve traceability during collaborative federated training. In addition, Grad CAM and LIME-based explainability modules are incorporated to improve transparency of segmentation predictions and support clinical interpretability of tumor related regions across multi-modal MRI scans. The framework further employs a 3D U-Net segmentation architecture within a decentralized non-IID federated setting to support stable collaborative learning across distributed healthcare institutions. The major contributions of this study are summarized as follows:

1. Development of a privacy-preserving federated brain tumor segmentation framework using distributed multimodal MRI datasets.
2. Integration of a blockchain-inspired audit logging mechanism for improved traceability and trusted inter-institutional collaboration.
3. Incorporation of Grad CAM and LIME-based explainable AI techniques for interpretable tumor segmentation analysis.
4. Evaluation of decentralized 3D U-Net segmentation performance under non-IID federated learning conditions using the BraTS 2020 dataset.
5. Establishment of a secure and explainable collaborative medical imaging pipeline for trustworthy AI-assisted neuro-oncology analysis.

The remainder of this paper is organized as follows. Section 2 reviews related work on federated learning, blockchain, and explainable artificial intelligence for brain tumour segmentation. Section 3 presents the proposed materials and methods, including dataset preparation, federated learning, blockchain coordination, and explainability integration. Section 4 discusses the experimental results and performance evaluation. Section 5 outlines the limitations of the study. Finally, Section 6 concludes the paper and summarizes the major findings.

2 Related Work

The integration of explainable AI, federated learning, and blockchain into medical image analysis — particularly for brain tumor segmentation — represents a promising and rapidly evolving research direction. This section summarizes the findings of the primary works, highlights the existing gaps, and proposes a unique integration of the three technologies to solve the problems of privacy, trust, and transparency.

Kumar et al. [12] integrate blockchain with federated learning for privacy-preserving brain tumor segmentation from medical images and identify the primary challenge. Participating hospitals train models in isolation on their sensitive Magnetic Resonance Imaging (MRI) datasets without transmitting sensitive data for tumor segmentation and model training. Blockchain technology secures the federated learning process and makes it transparent as hospitals' local model updates (weights) are amalgamated to construct a segmentation model. Some major limitations of the work include the high computational requirements and intricate nature of federated coordination and blockchain technology, along with the potential communication inefficiencies and model divergence created by differing data distributions across various medical institutions.

Yahiaoui et al. [13] suggest a federated learning model specifically tailored for multi-institutional collaboration on 3D brain tumor segmentation (and protecting patient data privacy). This method allows many hospitals to train robust segmentation models like 3D Convolutional Neural Networks on their local MRI datasets without sharing the volumetric MRI slices. The most significant limitation concerns the communication overhead associated with scheduling federated training across geographically distributed sites, which is particularly pronounced when processing large 3D volumetric MRI datasets.

Giri et al. [18] have designed a new system that combines a Federated Learning approach to brain tumor segmentation with a customized hybrid neural network, ResUHybridNet, and variable neural architecture. It helps maintain data privacy across various institutions. The principal limitations involve the computational and communication costs of training sophisticated hybrid models in a federated context, possible model drift due to data heterogeneity across various medical centers, and model personalization challenges in a decentralized training structure. Raheem et al. [19] put forward a new framework called Improved Privacy-Preserving Collaborative Convolutional Neural Network (IPC-CNN) that offers precise brain tumor segmentation while also increasing data privacy during multi-institutional collaborations. Albalawi et al. [20] propose an integrated framework that jointly utilizes Federated Learning (FL) and Transfer Learning (TL) to improve the classification and diagnosis of brain tumors while safeguarding data privacy between various organizations.

The main obstacles are the federated fine-tuning of complicated models, the impact of disparate (non-IID) distribution models on AI performance, and the

Table 1. Comparative analysis of recent federated learning, blockchain, and artificial intelligence-based approaches for brain tumour segmentation and medical imaging applications.

Ref	Year	Idea	Dataset	Technology	Shortcomings
[14]	2024	Large-scale multi-institutional benchmark for evaluating FL in brain tumor segmentation.	FeTS 2022 / BraTS	FedAvg, personalization, clustered FL	Benchmarking-focused; lacks novel algorithmic innovation; needs real hospital validation.
[15]	2025	Federated and transfer learning for multi-institutional cancer detection based on image analysis.	Public datasets (TCGA, METABRIC, BraTS)	Federated Learning, Transfer Learning, CNN	Domain shift across cancer types; limited generalization to 3D volumetric segmentation tasks.
[16]	2024	Enhancing general medical image classification via federated learning and pre-trained model integration.	General medical imaging datasets (publicly available)	Federated Learning, Pre-trained Models, Deep Learning	Limited to classification tasks; not evaluated on volumetric brain tumor segmentation; dataset heterogeneity across institutions.
[17]	2024	Brain tumor Detection, blockchain	Kaggle Brain Tumor MRI Dataset.	Blockchain IPFS, CNN	Expandability, Model Generalizability
[13]	2024	FL with PPTS for brain Tumor segmentation	BraTS 2020 dataset	3D U-Net. Federated Learning.	Privacy and poor Performance
[18]	2024	Federated trained via FL	MRI scans from 5 hospitals	ResUHybridNet;FL	Heterogeneity: Computation Integration:
[19]	2024	IPC-CNN FL with a custom CNN for brain Tumor segmentation	Figshare Dataset: BRATS 2018: Multi-modal MRI scans	IPC-CNN Architecture: Federated Learning:	Data Heterogeneity: GAN Dependency:
[20]	2024	Combining FL and TL to classify Brain Tumors	Combined Datasets figshare, SARTAJ, and Br35H.	FL, TL, CNN	Data Biases, Scalability, and Complexity
[21]	2024	FL framework with InceptionV3 for glioma detection in MRI images	3,242 MRI images (2,594 for training, 648 for testing	Federated Learning (FL), InceptionV3	Computational Cost Drop, Limited Aggregation Methods:
[22]	2024	BL integrated with IPFS	MRI images	DL, Blockchain, IPFS	Limited scope, Security complexity
[23]	2025	Blockchain and (PETs) integrated with FL for privacy in smart healthcare systems (SHSs)	MNIST, CIFAR-10, and healthcare-specific datasets.	Blockchain, DL	communication inefficiencies, and scalability challenges
[12]	2024	blockchain-based FL for brain tumor segmentation	BraTS 2020 dataset: Data distribution:	FL, Blockchain: 3D U-Net.	Computational Overhead,Complexity.

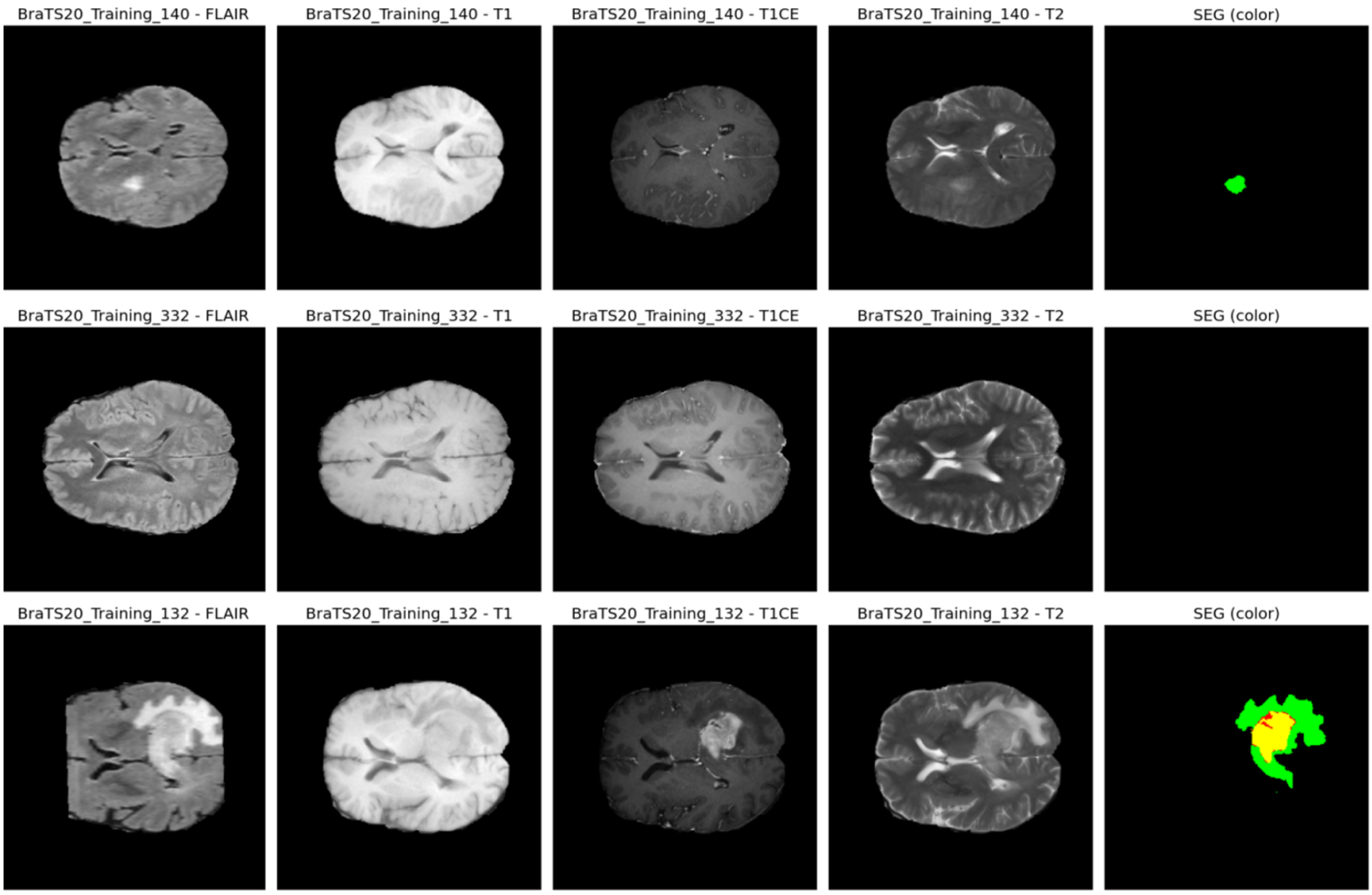


Figure 2. Representative samples from the training dataset with multi-modal MRI sequences and segmentation masks.

extent to which the pre-trained model is relevant to the medical imaging domain. Sharma et al. [21] suggest a new privacy-preserving federated learning framework built for cross-silo collaboration to enhance glioma detection using distributed datasets on brain tumors. The main concerns on the complex and time-consuming privacy-preserving safeguards, the impact of non-IID data from different clinical silos, variability in data annotation, and the dependency on annotated data so the model can be reliable are discussed in detail. Table 1 summarizes recent studies related to federated learning, blockchain integration, transfer learning, and deep learning-based approaches for brain tumour segmentation and medical imaging applications.

Manthe et al. [14] conducted an extensive federated benchmark evaluation for brain tumor segmentation across multiple institutions, demonstrating the generalization potential of federated models but noting the absence of novel algorithmic contributions and the need for real clinical validation. Recent studies have demonstrated the potential of Federated Learning, blockchain integration, and explainable AI for privacy-preserving medical image analysis. However, most existing approaches focus primarily on isolated aspects such as decentralized training, secure coordination, or segmentation accuracy, while limited attention has been given to the unified integration of privacy preservation, security verification, and interpretability within a single collaborative framework. Several prior studies also reported challenges associated with non-IID data distributions, communication overhead, lack of transparent model auditing, limited explainability evaluation, and reduced segmentation consistency across distributed institutions. In contrast, the proposed framework integrates a 3D U-Net-based federated segmentation architecture with blockchain-assisted model verification and Grad-CAM/LIME-based interpretability mechanisms within a decentralized healthcare environment. The blockchain-inspired audit layer provides traceable logging and verification of model update records generated during federated training. Furthermore, the proposed framework demonstrates stable federated convergence and consistent segmentation performance using distributed BraTS 2020 MRI datasets without requiring centralized patient data sharing. This integrated design addresses limitations of prior studies by jointly supporting secure collaboration, decentralized learning, and interpretable brain tumor segmentation within a unified framework.

3 Materials and methods

3.1 Dataset collection

The dataset used for this research is BraTS 2020 (Brain Tumor Segmentation), which is a widely used benchmark dataset for brain tumor segmentation that contains multi-modal MRI scans [24]. This dataset comprises T1-weighted (T1), contrast-enhanced T1 (T1ce), T2-weighted (T2), and Fluid-Attenuated Inversion Recovery (FLAIR) MRI sequences. Each MRI volume of this dataset is annotated by clinicians. For each annotated volume, they provide voxel-level ground truth masks for the sub-regions of the tumor: enhancing tumor (ET), tumor core (TC), and whole tumor (WT) depicting high precision. This dataset is particularly useful for deep learning segmentation algorithms since each MRI provides volumetric data of $240 \times 240 \times 155$ voxels, which is of sufficient quality for deep learning-based segmentation. In this study, the central 40 axial slices were extracted from each volume during preprocessing, resulting in an effective input dimension of $240 \times 240 \times 40$ per MRI scan.

In order to simulate a federated learning setting that involves multiple institutions, the BraTS 2020 dataset is split and distributed to multiple client nodes, each acting as an independent hospital or medical center. Each client node is allocated a subset of patient MRI scans which is non-overlapping, to make certain that institutions are not trading any raw data. This data follows a non-independently and identically distributed (non-IID) configuration to portray real-world heterogeneity, such as the diversity in scanner types used by hospitals, imaging protocols and patient population.

Each subset of the dataset was constructed to ensure balance and diversity: 70% of the data was used for training, while the remaining 30% was divided evenly for validation and testing. This configuration enables the global model to learn from various training samples while keeping independent datasets for performance assessments. In the non-IID setup, model training under the federated architecture must generalize to disparate domains of data for effective cross-learning.

Figure 2 illustrates representative samples from the training dataset used in this study. Each sample contains four MRI modalities, namely FLAIR, T1, T1CE, and T2, together with the associated segmentation mask. The figure demonstrates the complementary information provided by multi-modal MRI imaging and the diversity of tumor appearances within the dataset.

3.2 Dataset pre-processing

All MRI scans undergo the same preprocessing steps of intensity normalization, skull stripping, and resizing to a specified voxel space prior to distribution. This ensures clients receive uniform input. In keeping with the federated learning paradigm, preprocessing steps were performed by each institution independently, and the data was kept local.

3.2.1 Image Normalization and Augmentation

The inconsistency in brightness and contrast of MRI images can be attributed to the use of different scanning machines and different acquisition settings. Each MRI volume undergoes normalization to reduce this variation. Intensity normalization techniques are also performed to rescale each of the voxel intensity in the volume to within some range, e.g., [0, 1]. With normalization, the training of the scanner can be done without any bias relative to the distribution of intensity [25].

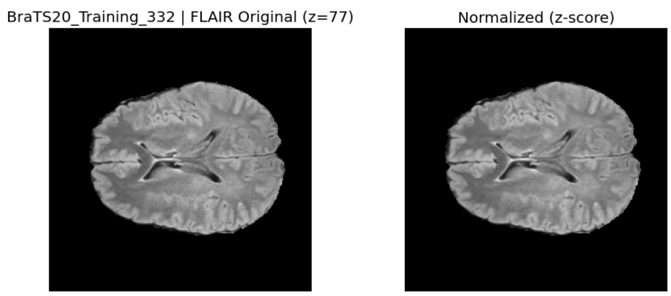


Figure 3. Representative samples from the BraTS20 dataset before and after normalization.

Figure 3 presents an example FLAIR MRI slice from the BraTS20 dataset in its original form and after z-score normalization. The normalization process standardizes image intensity values, improving consistency across MRI scans and facilitating stable deep learning model training.

Considering that MRI data consists of three-dimensional volumetric images, analyzing all slices can become costly in terms of computation. In this study, the approach is further refined to the central forty slices from each MRI volume, as these slices have the most pertinent information related to the tumor. To ensure equal input size across all clients, all MRI slices are resized to a fixed dimension, for example, 128 × 128 pixels. To improve generalization and reduce overfitting, each client employs local data augmentation techniques, such as random rotations, horizontal and vertical flips, and minor changes to the intensity [26].

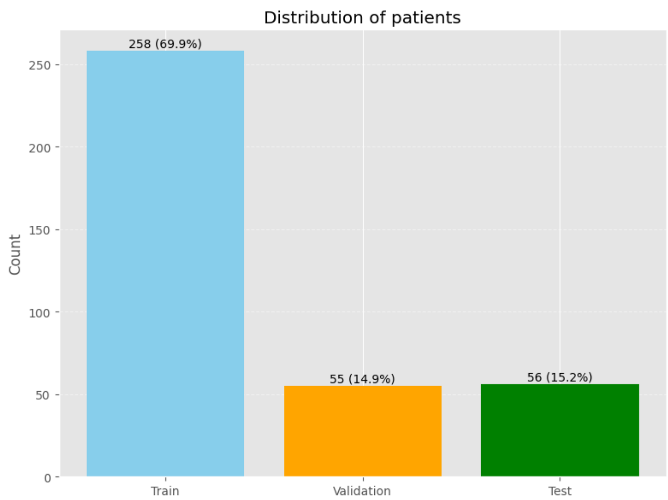


Figure 4. Distribution of patients across training, validation, and testing datasets.

To partition the data into training, validation, and testing subsets, 258 patient scans were allocated for training, 55 for validation, and 56 for testing, ensuring a distribution that meets the requirements of the federated learning evaluation protocol. Figure 4 illustrates the allocation of patient samples into the training, validation, and testing subsets. A total of 258 patients (69.9%) were assigned to the training set, while 55 patients (14.9%) and 56 patients (15.2%) were allocated to the validation and testing sets, respectively.

3.3 Experimental Configuration

The proposed federated brain tumor segmentation framework was implemented using the PyTorch deep learning library with MONAI medical imaging utilities for volumetric MRI preprocessing and segmentation model development. All experiments were conducted on a workstation equipped with an NVIDIA RTX 4090 GPU with 24 GB memory, 64 GB RAM, and an Intel Core i9 processor running Ubuntu Linux. The decentralized federated learning environment was simulated using four participating client nodes, where each client represented an independent healthcare institution with locally stored BraTS 2020 MRI datasets. Non-IID data distribution was maintained across client nodes to emulate realistic inter-institutional variability in imaging characteristics and patient populations.

The proposed segmentation framework employed a 3D U-Net architecture [27, 28] trained using the Adam optimizer with an initial learning rate of 0.0001. A batch size of 2 was used during local client training due to the computational requirements of volumetric MRI segmentation. The Dice Loss combined with Cross Entropy Loss was utilized as the optimization objective to improve segmentation overlap performance across tumor subregions [29].

To analyze convergence behavior, training metrics were periodically recorded at intermediate monitoring checkpoints throughout the optimization process. These checkpoints were used solely for performance monitoring and did not alter the federated aggregation procedure.

Model evaluation was conducted after each aggregation cycle using validation Dice score, Intersection over Union (IoU), voxel accuracy, and Hausdorff Distance metrics [30]. For auditability and coordination, a blockchain-inspired permissioned ledger mechanism was implemented to record encrypted model update hashes, aggregation metadata, and client participation logs during each federated communication round. The ledger was simulated within the experimental environment. However, it was not deployed as a fully decentralized blockchain network with consensus mechanisms. Its primary purpose was to evaluate traceability and audit logging of model aggregation events within the federated learning workflow.

3.4 Secure Federated Coordination with a Blockchain-Inspired Audit Layer

Following preprocessing, decentralized training commences through the deployment of a Federated Learning system augmented with a blockchain-inspired audit mechanism for the secure orchestration of model upgrades and validation [5, 7]. A federated learning system consists of three primary modules, namely, local training, secure model updates, and global aggregation.

3.4.1 *Federated Learning workflow*

Each collaborating client in the suggested framework stands in for a separate hospital or healthcare facility. Each client uses the chosen 3D U-Net framework to train models locally and keeps their own local dataset of brain MRI scans. Iterative training rounds make up the federated learning process. All participating clients receive the current global model from the central server at the start of each round. Only the

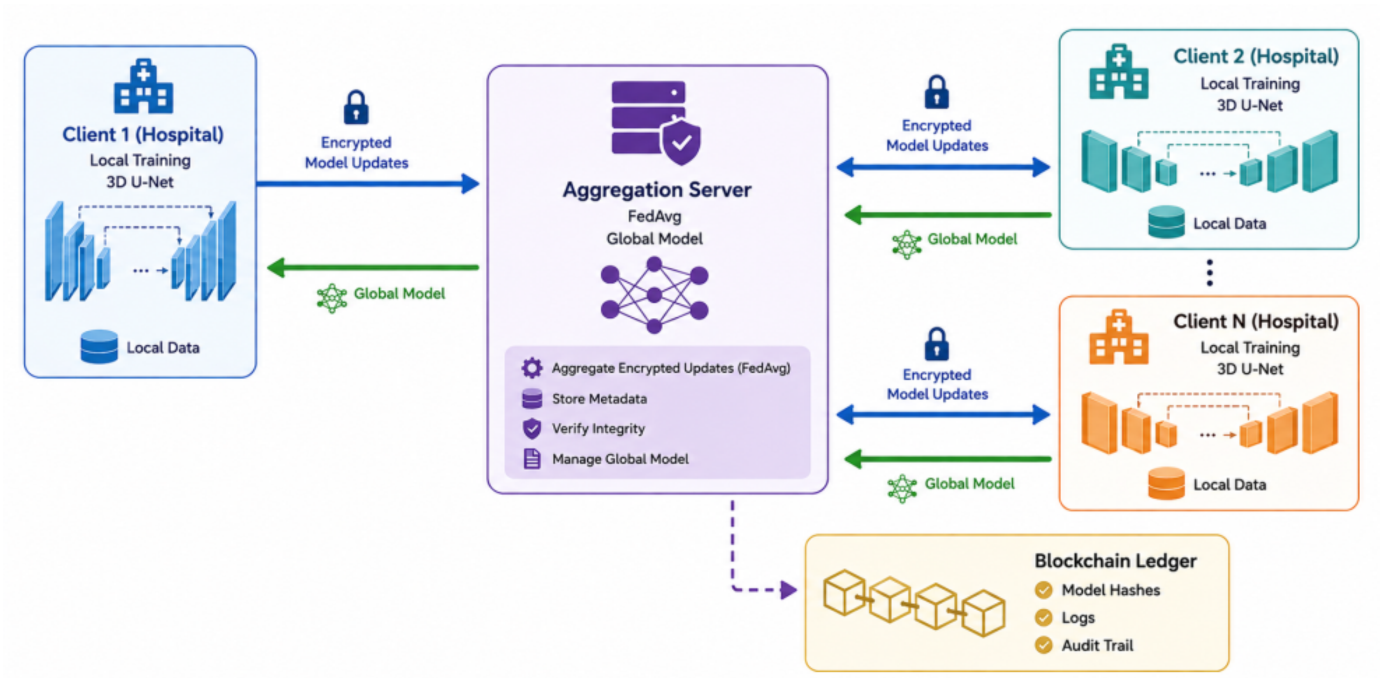


Figure 5. Conceptual architecture of the blockchain-inspired federated learning framework for collaborative brain tumor segmentation.

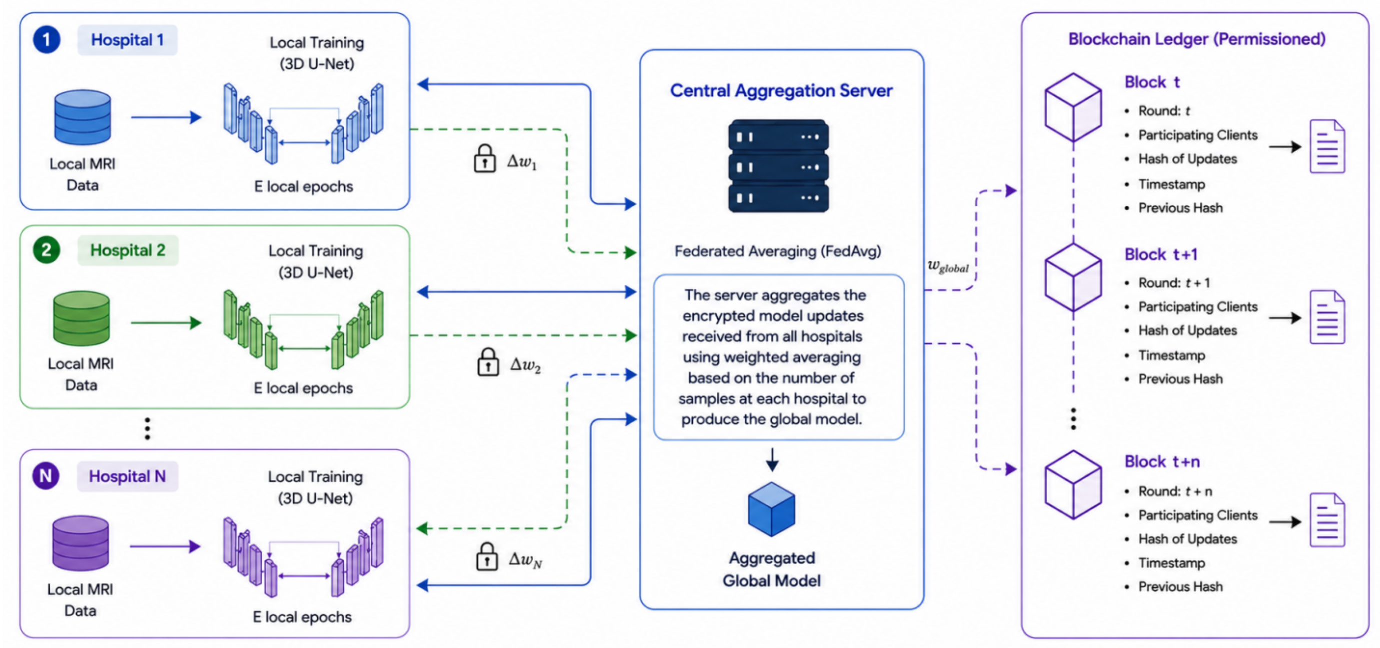


Figure 6. Detailed operational workflow of federated model training, aggregation, and audit logging during communication rounds.

changed gradients or model parameters are sent back to the server upon completion. The Federated Averaging (FedAvg) algorithm [31] was employed as the aggregation strategy. During each communication round, participating clients independently trained the current global model using their local MRI datasets. Upon completion of local training, the updated model parameters were transmitted to the aggregation server. The server then generated a new global model by computing a weighted average of the client updates, where the contribution of each client was proportional to the size of its local training dataset.

The federated learning pipeline incorporates a blockchain layer to improve security and efficiency.

A blockchain-inspired audit ledger is used to record cryptographic hashes of model updates, participation records, and aggregation metadata [7]. Figure 5 illustrates the high-level conceptual architecture of the proposed framework, highlighting the interaction among participating hospitals, the federated aggregation server, and the blockchain-inspired audit layer.

Figure 6 illustrates the detailed operational workflow executed during federated training, including local model updates, global aggregation, communication rounds, and audit record generation. Each participating hospital locally trains a 3D U-Net segmentation model using private MRI datasets and shares only encrypted model updates with the central aggregation server. Federated Averaging (FedAvg) is used to generate the updated global model, while a blockchain-inspired permissioned ledger records aggregation metadata and update hashes to support traceability and auditability across institutions. Here, w_1 denotes the trained model weights from Hospital 1, w_2 denotes the trained model weights from Hospital 2, and w_n denotes the trained model weights from Hospital N.

In the blockchain section of Figure 6, t represents the communication round or training iteration number in the federated learning process. Block t corresponds to the blockchain record generated during the current federated learning round. Block $t + 1$ represents the next training and aggregation round. Block $t + n$ denotes later communication rounds as training continues. The notation shows that blockchain entries are created sequentially after every federated aggregation cycle, enabling traceability and auditability throughout training.

Each client (hospital) performs local training on private MRI data using a 3D U-Net model and shares only encrypted model updates. The aggregation server applies Federated Averaging (FedAvg) to construct a global model, while a blockchain ledger ensures immutable logging, verification, and auditability of each training round.

3.4.2 Secure Federated Training and Aggregation

Every hospital trains a 3D U-Net model locally on its own pre-processed MRI scans. For the local training, hospitals do not communicate the raw MRI data. Only the learned model parameters or gradients are sent to a central coordinator. Every local model independently learns tumor characteristics to augment the collective learning.

After completing local training, hospitals send encrypted updates of the model instead of the raw data. The aggregation process followed the standard FedAvg protocol. Model updates received from participating clients were combined according to the relative size of each client's local dataset, allowing institutions with larger training datasets to contribute proportionally more to the global model update. This procedure was repeated across successive communication rounds until completion of federated training.

The global model was updated using the Federated Averaging (FedAvg) aggregation rule, in which local model updates from participating clients were combined according to the relative size of their training datasets. The updated global model was computed using the standard Federated Averaging (FedAvg) formulation shown in Equation (1).

$$w^{(t+1)} = \sum_{k=1}^{K} \frac{n_k}{N} w_k^{(t+1)} \tag{1}$$

where $w^{(t+1)}$ denotes the updated global model parameters after aggregation round t, $w_k^{(t+1)}$ represents the local model parameters submitted by client k, n_k is the number of training samples available at client k, N is the total number of training samples across all participating clients, and K is the total number of participating clients.

These model updates are recorded and monitored through a blockchain-inspired permissioned ledger that stores aggregation metadata, update hashes, timestamps, and participation records. The ledger supports traceability and auditability of federated training events while maintaining a lightweight implementation suitable for the experimental environment. The audit logging mechanism provides a verifiable record of model aggregation activities and client participation throughout federated training. All verified updates are synergized to produce an enhanced global model. Redistributing the globally aggregated model securely allows each client to receive the model. Each hospital updates the global model and engages in further local training on its own data. This new global knowledge is incorporated into the local updates, resulting in the generation of a generalized global segmentation model capable of learning from diverse institutional datasets. This process does not compromise the privacy of any participating institution's local data.

3.4.3 Blockchain-Inspired Coordination and Audit Mechanism

To improve auditability and traceability during decentralized training, the proposed framework incorporates a blockchain-inspired permissioned ledger mechanism within the federated learning environment. Rather than implementing a fully decentralized blockchain network, the ledger records cryptographic hashes of model updates, aggregation metadata, encrypted checkpoints, and client participation logs generated during each communication round.

During federated training, participating clients locally train the 3D U-Net model and transmit encrypted parameter updates to the aggregation server. Corresponding hash values are recorded within the blockchain ledger before global aggregation. Each block stores update hashes, timestamps, participating client identifiers, aggregation metadata, and the previous block hash to maintain traceability across communication rounds. The permissioned ledger allows authorized participants to verify aggregation records and model update histories, thereby supporting transparency and traceability during collaborative learning. Lightweight validation mechanisms were used to verify update integrity while minimizing computational overhead.

The blockchain-inspired ledger improves auditability by maintaining traceable records of aggregation events and encrypted model update hashes. Although the framework does not implement decentralized consensus or Byzantine fault tolerant validation, the ledger provides a structured mechanism for recording model lineage and aggregation history throughout federated training.

3.5 Explainable Artificial Intelligence Integration

To improve understanding, trustworthiness, and trust within the clinic, the proposed structure pairs Explainable Artificial Intelligence (XAI) with the deployed global federated model. Such interpretability mechanisms automate model prediction while human prediction closes the loop and boosts trust in the AI-diagnosis in the clinic. Having included XAI in the federated structure, it provides insight into the decision-making process of the model by predicting what, why, and how. These features help clinicians comprehend, interpret, and validate model attention areas fostering effective collaborations between humans and AI.

Figure 7 demonstrates the explainability mechanism integrated within the proposed segmentation framework. Multi-modal MRI scans are processed using the 3D U-Net model to generate tumor segmentation maps, which are subsequently interpreted using Grad-CAM heatmaps.

3.6 Model Evaluation and Validation

In the last stage of the methodology, the trained global segmentation model goes through rigorous evaluation on an independent test set. To assess the model's performance, we consider the accuracy, precision, recall, F1 score and the domain-specific performance metrics such as Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD). DSC indicates the degree of overlap between the predicted segmentation of the tumor and the segmentation ground truth provided by an expert. The higher the Dice score, the better the segmentation accuracy. While HD measures the degree of spatial similarity between the predicted and ground truth tumor regions and their boundaries. The lower the HD, the better the boundary delineation.

Table 2. Primary quantitative performance metrics used for evaluation of the proposed federated brain tumor segmentation framework.

Performance Measure	Description
Dice Score (DSC)	Overlap between prediction and ground truth.
IoU	Intersection over union of masks.
HD95	Boundary distance accuracy.
Precision	Correct tumor predictions.
Recall	Detected true tumor regions.
F1 Score	Balance of precision and recall.
Accuracy	Correctly classified voxels.

Table 2 summarizes the primary quantitative performance metrics used to evaluate the proposed federated brain tumor segmentation framework. These metrics assess segmentation overlap accuracy, boundary agreement, classification performance, and overall prediction reliability across multi modal MRI datasets.

The global model's performance is compared with non-federated baselines to show the effectiveness of the federated and blockchain-integrated configuration. The integration of Federated Learning, Blockchain,

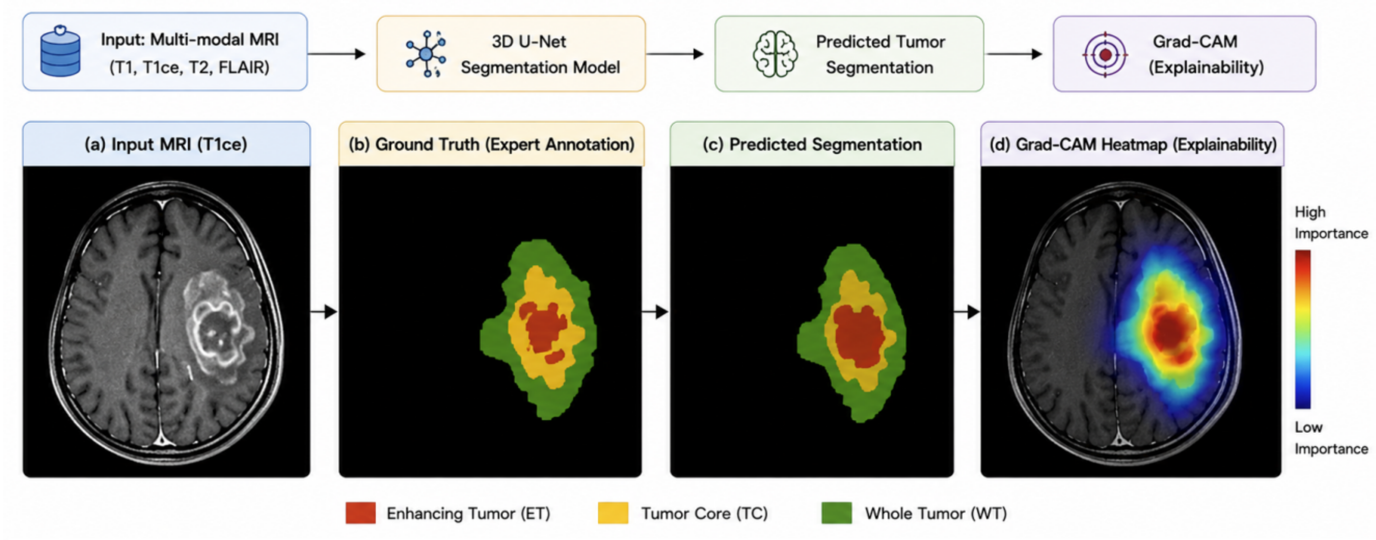


Figure 7. Explainable brain tumor segmentation using the proposed 3D U-Net and Grad-CAM visualization framework.

and Explainable AI enables hospitals to work together without the sharing of data, and clinicians gain the trust of the model because they can comprehend its rationale.

Figure 8 presents the overall architecture of the proposed framework for privacy-preserving and interpretable brain tumor segmentation. Multi-institutional MRI datasets are locally pre-processed and used to train 3D U-Net models at individual hospitals without sharing raw patient data. Encrypted model updates are securely aggregated using Federated Averaging (FedAvg) and recorded through a permissioned blockchain to ensure transparency and secure collaboration. The aggregated global model generates tumor segmentation outputs, while Grad-CAM-based explainability provides interpretable visualizations to support clinical decision-making and trustworthy AI-assisted diagnosis.

4 Results and Discussion

4.1 Segmentation Performance Evaluation

The proposed federated framework demonstrated consistent segmentation performance across the training rounds and evaluation stages. Multi-institutional MRI datasets were processed through the decentralized 3D U-Net architecture, while aggregation traceability was maintained through the blockchain-inspired audit coordination mechanism. The obtained results indicate stable convergence behavior with progressively improved segmentation quality during training.

Table 3 presents the overall voxel-wise classification performance of the proposed framework. The model achieved an overall voxel accuracy of 0.985, indicating strong agreement between predicted and reference segmentation labels. Weighted average precision, recall, and F1 score values exceeded 0.98, demonstrating stable segmentation performance across the evaluated MRI volumes. Comparatively lower macro average scores were observed due to class imbalance and the increased complexity associated with smaller tumor subregions such as necrosis.

Table 3. Overall voxel-wise classification performance of the proposed federated segmentation framework.

Metric	Value
Overall Voxel Accuracy	0.985
Macro Average Precision	0.749
Macro Average Recall	0.599
Macro Average F1 Score	0.648
Weighted Average Precision	0.983
Weighted Average Recall	0.985
Weighted Average F1 Score	0.983

The convergence trend further supports the stability of the proposed federated learning configuration. Training progression remained smooth with no abrupt fluctuations in loss or segmentation metrics, suggesting that the decentralized aggregation strategy successfully maintained model synchronization among participating clients.

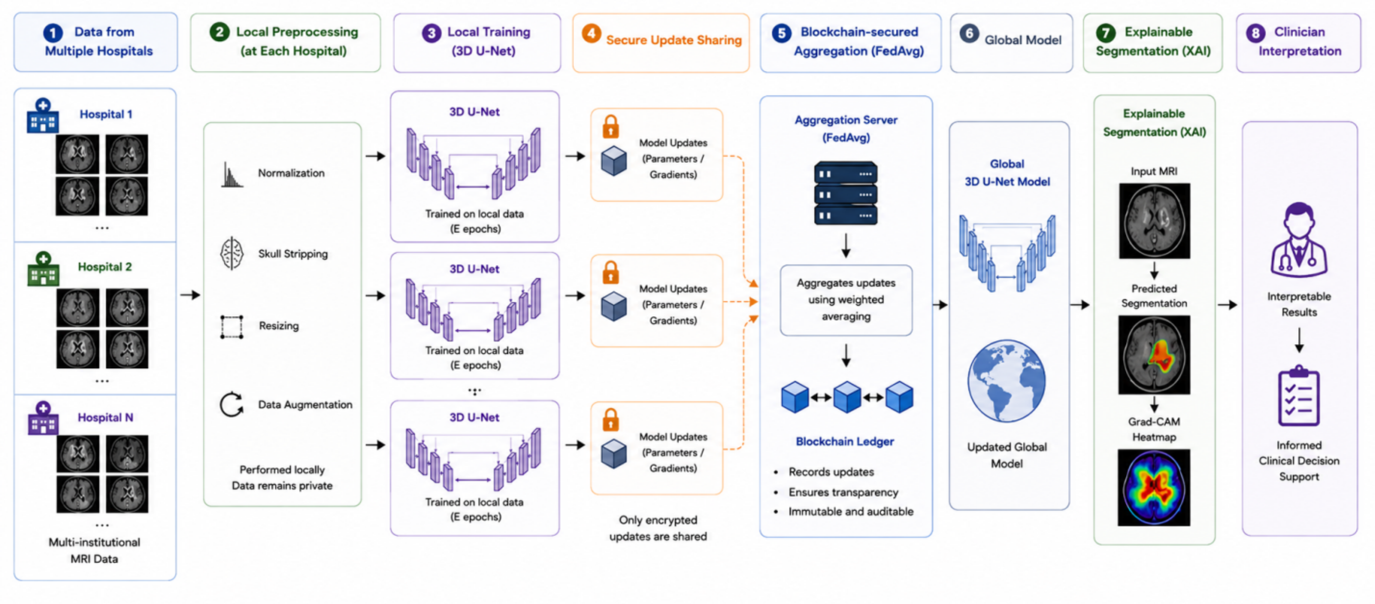


Figure 8. End-to-end framework of the proposed decentralized, blockchain-inspired, and explainable brain tumor segmentation system.

4.2 Epoch-wise Training Analysis

The quantitative training results recorded at successive monitoring checkpoints during federated optimization are presented in Table 4. A gradual reduction in both training loss and validation loss was observed as the epochs progressed. Training loss decreased from 0.150 at Checkpoint 1.0 to 0.065 at Checkpoint 5.0, while validation loss decreased from 0.165 to 0.092 during the same interval. These findings indicate that the optimization process converged steadily throughout training.

The Dice coefficient demonstrated substantial and consistent improvement throughout training. Initially recorded at 0.752, the Dice score progressively increased to 0.830 at Checkpoint 5.0. This steady upward trend indicates significantly improved overlap between predicted tumor regions and expert annotations, confirming that the segmentation model successfully refined tumor boundary localization capabilities throughout the federated learning process.

Table 4 summarizes the progression of training accuracy, training loss, validation loss, and Dice coefficient across successive monitoring checkpoints during federated segmentation training.

To monitor convergence behavior in greater detail, performance metrics were recorded at intermediate checkpoints during training rather than only after the completion of full training epochs. The values reported at 1.5, 2.5, 3.5, and 4.5 represent these intermediate monitoring checkpoints used to track optimization progress between consecutive training stages, providing a finer-grained view of model evolution.

Table 4. Checkpoint-wise training and validation performance of the proposed federated segmentation framework.

Checkpoint	Accuracy	Train Loss	Validation Loss	Dice Coefficient
1.0	0.9950	0.150	0.165	0.752
1.5	0.9925	0.125	0.142	0.768
2.0	0.9900	0.105	0.128	0.781
2.5	0.9875	0.092	0.115	0.793
3.0	0.9850	0.085	0.108	0.802
3.5	0.9825	0.078	0.102	0.812
4.0	0.9800	0.072	0.098	0.819
4.5	0.9775	0.068	0.095	0.825
5.0	0.9750	0.065	0.092	0.830

The observed discrepancy between voxel accuracy and Dice coefficient is a well-recognized characteristic of medical image segmentation tasks involving severe class imbalance. In the BraTS dataset, background voxels represent the overwhelming majority of the image volume, whereas tumor subregions occupy relatively small spatial areas. During early training stages, the high global voxel accuracy is largely a reflection of the model's ability to classify the dominant background class. As training progresses, the model increasingly shifts its optimization focus toward delineating complex, smaller tumor structures

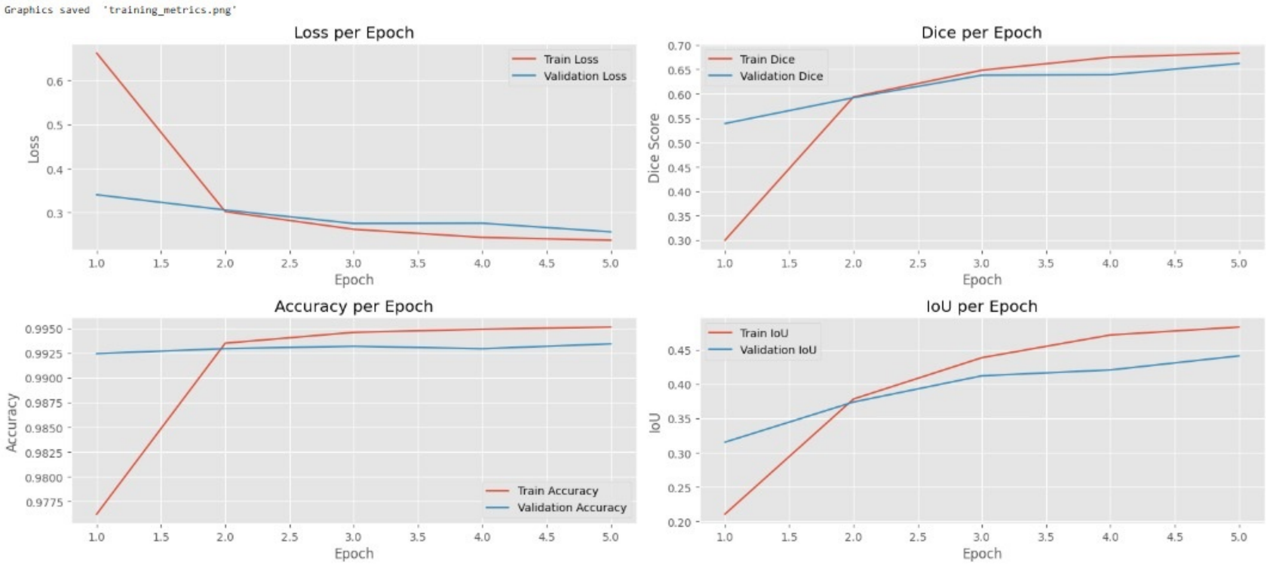


Figure 9. Training and validation performance curves across monitoring checkpoints for the proposed federated brain tumor segmentation framework.

and refining boundary regions. This strategic shift inherently leads to a modest reduction in overall voxel accuracy (from 0.9950 to 0.9750) while simultaneously driving a substantial improvement in clinically relevant segmentation metrics. Consequently, the significant rise in the Dice coefficient from 0.752 to 0.830 serves as a far more informative and trustworthy indicator of true model performance than the global accuracy.

The graphical performance trends further support these quantitative observations. Loss curves demonstrated stable convergence behavior with progressively decreasing validation loss values. Similarly, Dice and IoU curves displayed continuous upward progression without major oscillations, indicating that the federated training process remained stable throughout successive communication rounds.

Figure 9 illustrates the progression of training and validation metrics across successive monitoring checkpoints during federated training. Training and validation losses decreased steadily, while Dice coefficient and IoU values improved progressively during training. Although overall voxel accuracy exhibited a slight decline, segmentation overlap performance continued to improve, indicating enhanced tumor region delineation. The smooth convergence behavior confirms stable optimization and effective learning throughout the federated segmentation process.

4.3 Segmentation Quality Assessment

Visual segmentation analysis demonstrated close agreement between the predicted tumor masks and the manually annotated ground truth regions. The segmented outputs preserved the overall tumor structure and spatial localization across the evaluated MRI slices. In several cases, the predicted tumor regions closely matched the annotated enhancing tumor boundaries, indicating that the model successfully learned clinically relevant spatial features from the distributed datasets.

The qualitative evaluation also showed that the proposed framework maintained segmentation continuity while suppressing most background noise regions. Minor discrepancies were observed around smaller tumor boundaries and low contrast regions, particularly within necrotic tissue areas. However, the major tumor structures remained accurately localized in most evaluated slices.

The prediction results further demonstrated that the federated learning strategy did not substantially compromise segmentation performance despite decentralized data distribution. The generated segmentation masks remained visually consistent with the reference annotations while preserving data privacy across participating institutions.

Table 5 summarizes the inference configuration and segmentation output information for a representative MRI sample evaluated using the proposed framework. The model processed a four channel volumetric MRI input with dimensions of 240 × 240 × 40 and

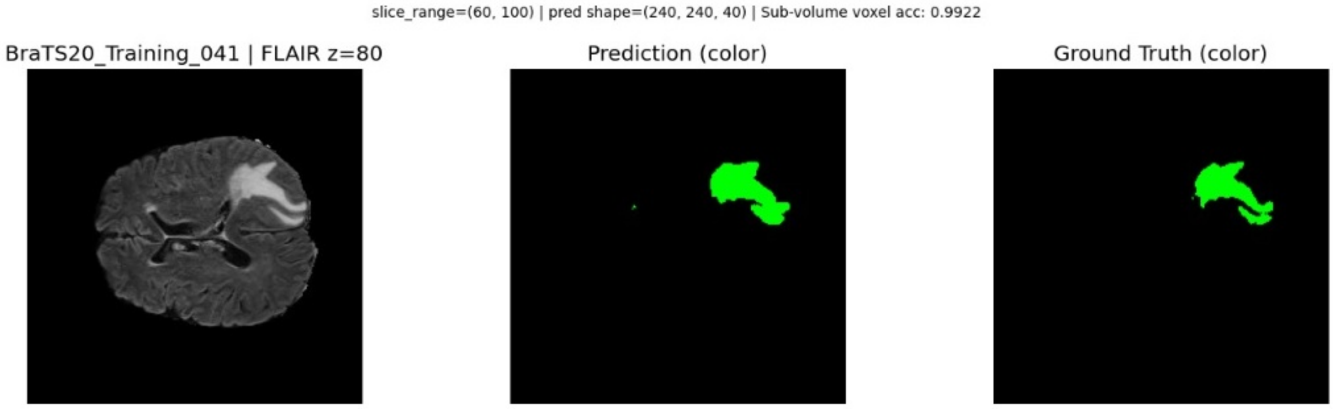


Figure 10. Qualitative comparison of MRI input slices, predicted tumor segmentation masks, and ground truth annotations generated by the proposed framework.

Table 5. Inference and prediction details for a representative brain tumor segmentation sample.

Parameter	Value
Patient ID	BraTS20_Training_041
Loaded Checkpoint	global_model_best.pth
Input Volume Shape	(1, 4, 240, 240, 40)
Ground Truth Availability	TRUE
Predicted Volume Shape	(240, 240, 40)
Saved Prediction Format	NIfTI (.nii.gz)
Slice Range Evaluated	(60, 100)
Sub-volume Voxel Accuracy	0.9922

generated the corresponding segmentation output volume successfully. The obtained sub volume voxel accuracy of 0.9922 indicates strong voxel wise agreement between the predicted segmentation and the annotated ground truth regions.

Figure 10 presents representative segmentation outputs produced by the proposed federated learning framework. The predicted tumor regions show strong visual correspondence with the manually annotated ground truth masks, indicating accurate localization and delineation of tumor structures across the evaluated MRI slices.

4.4 Classification and Segmentation Metrics

Table 6 presents the class-wise precision, recall, F1 score, and support values obtained during segmentation evaluation.

The classification report demonstrated high voxel-level prediction performance for the background class, with precision, recall, and F1 score values exceeding 0.99. Segmentation performance for tumor-related classes

Table 6. Class-wise voxel-level classification performance of the proposed segmentation framework.

Class	Precision	Recall	F1-Score	Support
Background	0.9914	0.9985	0.995	125,597,453
Necrosis	0.5773	0.2051	0.3026	919,940
Edema	0.6598	0.6095	0.6337	1,792,243
Active Tumor	0.7662	0.5822	0.6616	714,364

showed comparatively lower values; particularly, for necrosis regions, where the Dice coefficient reached 0.3026. This behavior may be associated with class imbalance and the relatively smaller spatial representation of necrotic tumor regions within the MRI volumes.

Edema and active tumor classes achieved Dice scores of 0.6337 and 0.6616, respectively. These findings indicate comparatively stronger segmentation capability for larger and visually distinguishable tumor structures. IoU values followed similar trends, with active tumor segmentation achieving an IoU value of 0.4943.

Table 7 summarizes the segmentation performance of individual tumor classes using Dice coefficient, IoU, HD95, precision, recall, and F1 score metrics. The segmentation performance varied across tumor subregions. Background segmentation achieved the highest Dice and IoU values, while edema and active tumor classes demonstrated comparatively stronger localization capability than necrosis regions. HD95 analysis further indicated acceptable spatial boundary agreement across most evaluated tumor structures.

Hausdorff Distance analysis demonstrated comparatively lower boundary deviation for active tumor segmentation compared with edema regions. The HD95 values suggest that the proposed framework

Table 7. Class-wise segmentation performance metrics of the proposed federated learning framework.

Class	Dice	IoU	Accuracy	HD95	Precision	Recall	F1
Background	0.995	0.99	0.9902	N/A	0.9914	0.9985	0.995
Necrosis	0.3026	0.1783	0.9933	14.542	0.5773	0.2051	0.3026
Edema	0.6337	0.4638	0.9902	18.704	0.6598	0.6095	0.6337
Active Tumor	0.6616	0.4943	0.9967	13.765	0.7662	0.5822	0.6616

preserved acceptable spatial boundary agreement across the evaluated tumor subregions. HD95 is reported only for tumor classes (necrosis, edema, and active tumor), as boundary distance metrics are not applicable to the background class. Figure 11 illustrates the class-wise segmentation performance of the proposed framework using Dice score, IoU, accuracy, HD95, precision, recall, and F1 score metrics. Background segmentation achieved the strongest performance across all evaluation metrics, while edema and active tumor regions demonstrated comparatively balanced segmentation capability. Lower Dice and recall values observed for necrosis regions indicate increased segmentation difficulty associated with smaller and heterogeneous tumor structures.

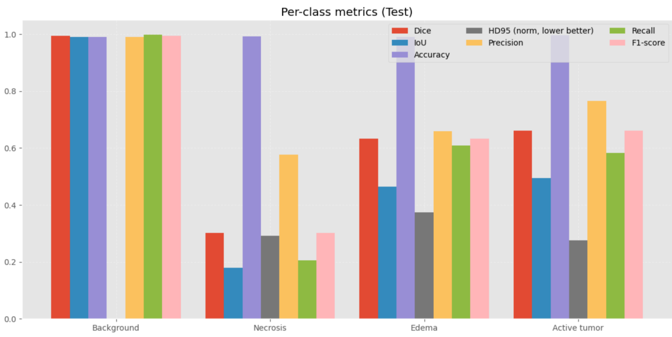


Figure 11. Per class segmentation performance analysis of the proposed federated learning framework.

4.5 Federated Learning Stability Analysis

The round wise federated learning analysis demonstrated stable collaborative learning behavior across all participating client nodes. Training loss continuously decreased during communication rounds, while validation Dice scores improved gradually with each aggregation cycle. This trend indicates that the Federated Averaging mechanism effectively consolidated distributed knowledge from multiple hospitals without destabilizing the global segmentation model.

The learning rate remained constant during training, which contributed to smooth optimization progression and prevented abrupt convergence instability. Similarly, validation accuracy values remained comparatively stable throughout the federated rounds, indicating that the decentralized training strategy maintained generalization capability despite non-IID data distributions among clients.

Figure 12 illustrates the batch-wise testing performance of the proposed segmentation framework using loss, Dice score, accuracy, and IoU metrics. Although minor fluctuations were observed across batches, the overall testing accuracy remained consistently high, while Dice and IoU values demonstrated acceptable segmentation consistency throughout evaluation. The aggregated global model also demonstrated consistent learning progression during successive rounds of encrypted model sharing and blockchain-recorded coordination. These findings suggest that the audit logging mechanism introduced negligible impact on segmentation convergence and overall training behavior.

The figure presents the federated learning performance trends across communication rounds. Training loss decreased progressively, while Dice and validation accuracy values improved steadily throughout the aggregation cycles. The stable metric progression demonstrates consistent collaborative learning and effective global model synchronization among distributed client nodes.

Table 8 presents the overall testing behavior of the proposed framework across batch-level evaluations. Test accuracy remained consistently high throughout most batches, indicating stable classification capability and preserved model generalization. Although fluctuations were observed in Dice score, IoU, and test loss values, the overall trends suggest acceptable segmentation consistency across varying tumor structures and imaging conditions. The observed oscillations may be associated with class imbalance and differences in tumor complexity among evaluated MRI samples.

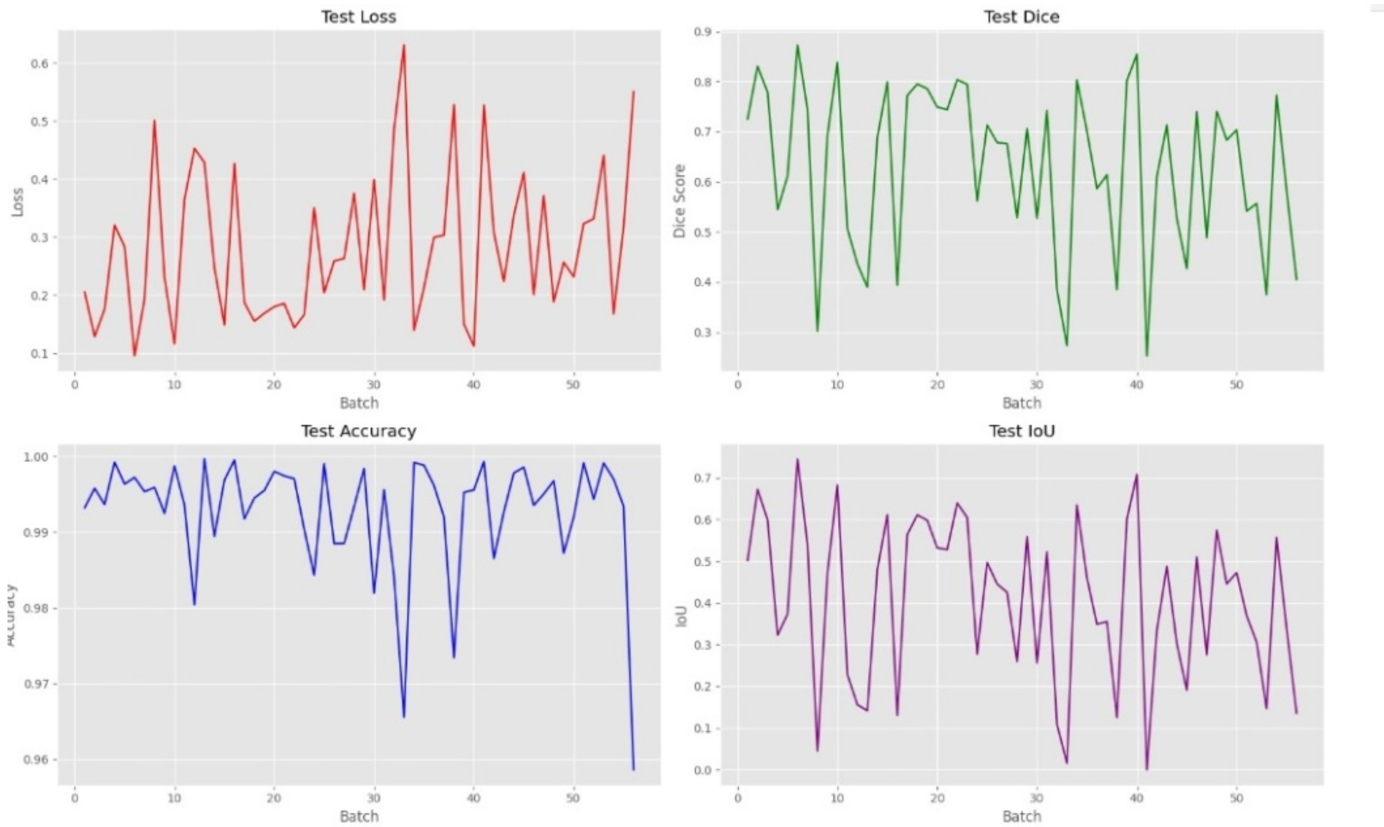


Figure 12. Batch-wise testing performance analysis of the proposed federated brain tumor segmentation framework.

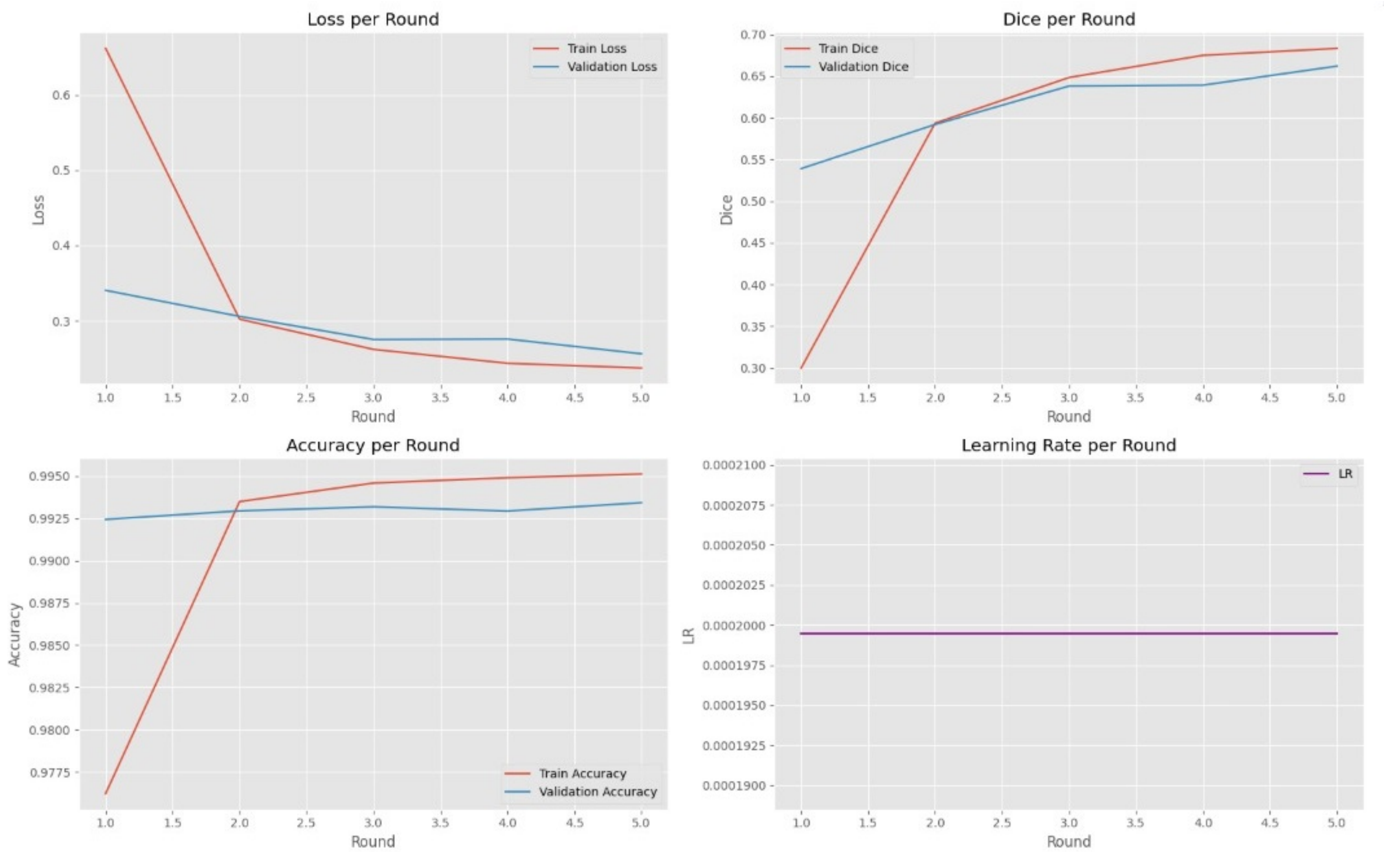


Figure 13. Round-wise federated learning performance analysis of the proposed segmentation framework.

Table 8. Summary of batch-wise testing performance metrics of the proposed segmentation framework.

Metric	Approx. Range	Mean / Stable Value	Trend / Observation	Interpretation
Test Loss	0.1 – 0.6	~0.25	Highly fluctuating but stays below 0.3 most of the time	Indicates good optimization; fluctuations suggest sensitivity to batch variation
Test Dice Score	0.3 – 0.85	~0.7	Alternates frequently with moderate oscillations	Shows effective segmentation capability though stability could improve
Test Accuracy	0.97 – 0.99	~0.985	Generally stable, minor dips around batches 30 and 55	Reflects excellent classification performance and generalization
Test IoU	0.1 – 0.7	~0.45	Fluctuates between batches, peaks around 0.7	Demonstrates moderate but reliable segmentation precision

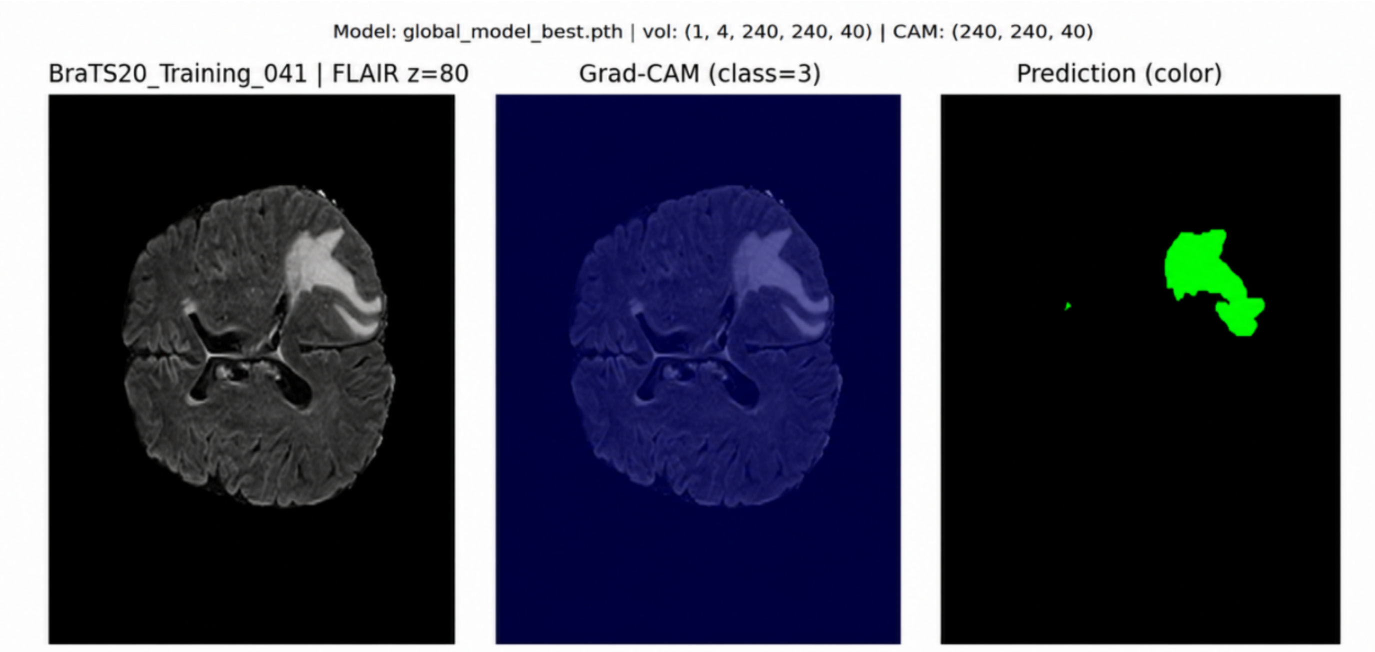


Figure 14. Grad CAM-based visualization of tumor attention regions generated by the proposed federated segmentation framework.

Figure 13 illustrates the round wise federated learning performance of the proposed brain tumour segmentation framework using training loss, validation loss, Dice coefficient, accuracy, and learning rate metrics. Training and validation losses decreased progressively across communication rounds, while Dice scores and accuracy values demonstrated steady improvement throughout training. The stable convergence behavior and consistent metric progression indicate effective global model aggregation and collaborative learning among distributed client nodes.

4.6 Explainable Artificial Intelligence for Model Interpretability

Explainable Artificial Intelligence techniques were incorporated into the proposed federated segmentation framework to improve interpretability and clinical transparency during brain tumor analysis. Gradient weighted Class Activation Mapping (Grad-CAM) [10] and Local Interpretable Model-agnostic Explanations (LIME) [32] were utilized to visualize the image regions contributing to segmentation predictions across multi-modal MRI scans.

Grad CAM analysis demonstrated that the segmentation model primarily focused on tumor-associated regions within the MRI slices. The generated activation maps showed close correspondence between highlighted regions and the predicted tumor areas. This observation suggests that the model learned relevant spatial and anatomical information during segmentation training. The activation patterns were mainly concentrated around hyperintense tumor regions visible in the FLAIR modality.

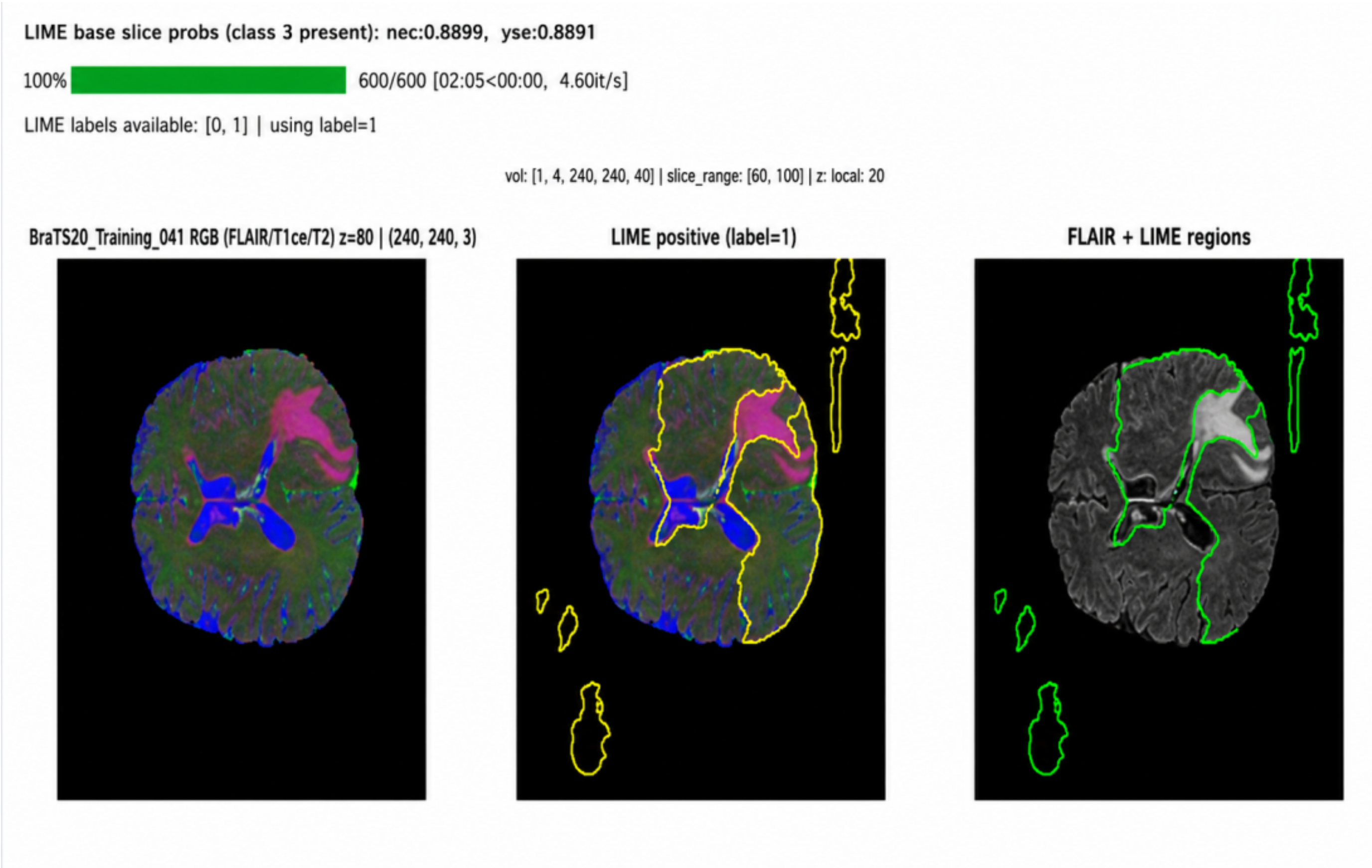


Figure 15. LIME-based local interpretability analysis of tumor-related regions in multi-modal MRI slices using the proposed framework.

Figure 14 presents representative Grad CAM visualization results generated using the proposed framework. The highlighted activation regions indicate the MRI areas that contributed most strongly to tumor segmentation predictions.

LIME-based analysis was further conducted to provide localized interpretability for individual segmentation predictions. The generated explanations identified influential superpixel regions associated with tumor classification and segmentation outcomes. These visualizations demonstrated that the framework preserved interpretable decision patterns while operating within a decentralized federated learning environment.

Figure 15 illustrates representative LIME based interpretability analysis obtained from multi modal MRI slices. The highlighted regions represent localized image areas that contributed significantly to tumor prediction and segmentation decisions.

The experimental findings indicate that the proposed framework successfully combined federated learning, blockchain-assisted coordination, and explainable medical image segmentation within a decentralized environment. Stable convergence patterns, progressively improving Dice scores, and visually consistent segmentation outputs were achieved throughout training and evaluation. The framework maintained high overall accuracy while preserving collaborative learning across distributed datasets. In addition, segmentation outputs remained aligned with manually annotated tumor regions, supporting the applicability of the proposed approach for privacy-preserving brain tumor segmentation tasks in multi-institutional healthcare environments.

5 Limitations

The proposed federated framework may also be affected by challenges associated with non-IID data distributions, communication overhead, and client heterogeneity. Variations in computational resources and imaging distributions across participating client nodes can influence training consistency and aggregation efficiency. In addition, repeated communication of encrypted model updates may increase latency and computational cost during large-scale federated deployment.

A further limitation is that the blockchain component was implemented as a simulated permissioned ledger rather than a fully decentralized blockchain network. Consequently, properties such as distributed consensus, Byzantine fault tolerance, and blockchain scalability were not experimentally evaluated. Future work should investigate deployment using real permissioned blockchain platforms such as Hyperledger Fabric to assess security, latency, storage overhead, and operational scalability under realistic healthcare environments.

Although the proposed framework demonstrated promising segmentation performance, several limitations remain. First, the experiments were conducted using the BraTS 2020 dataset within a simulated federated learning environment rather than a real multi-institutional clinical deployment. Consequently, variations associated with real hospital infrastructure, network instability, and institutional workflow integration were not fully evaluated.

Additionally, the presence of non-IID data distributions across client nodes may still affect global model consistency and segmentation performance, particularly for smaller tumor subregions such as necrosis. In addition, the integration of blockchain-based coordination introduced additional computational and communication overhead that may limit scalability in resource-constrained clinical environments.

Furthermore, the explainability analysis using Grad CAM and LIME was qualitatively assessed and was not validated through formal clinician-centered evaluation studies. The framework was also evaluated primarily on brain tumor segmentation tasks, which may limit direct generalization to other medical imaging applications and disease categories.

Future research may focus on large-scale real-world federated deployment, lightweight communication strategies, improved handling of heterogeneous client distributions, and clinically validated explainability assessment frameworks.

6 Conclusion

This study presents a federated learning-based framework for privacy-preserving brain tumor segmentation using multimodal MRI data. The proposed framework integrated decentralized learning, blockchain-assisted coordination, and explainable artificial intelligence techniques to support secure and interpretable medical image analysis across distributed institutional environments. Experimental evaluation demonstrated stable segmentation performance with an overall voxel accuracy of 98.5%. The framework achieved effective segmentation performance across tumor subregions while preserving patient data privacy through distributed model training. In addition, the Grad CAM and LIME-based interpretability analysis provided visual explanations of tumor-related regions contributing to segmentation predictions, supporting transparency during model evaluation. The findings indicate that federated learning can support collaborative medical image analysis without direct sharing of sensitive clinical data. Although challenges associated with non-IID data distributions, computational overhead, and real-world deployment remain, the proposed framework demonstrated consistent learning behavior and stable segmentation performance throughout training and evaluation. This work contributes toward the development of privacy-preserving and interpretable AI systems for brain tumor segmentation within decentralized healthcare environments. Future research may focus on large-scale clinical validation, lightweight federated optimization strategies, and improved handling of heterogeneous medical imaging data across institutions.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

This study did not require ethical approval or informed consent as it involved secondary analysis of the publicly available and fully anonymized BraTS 2020 dataset. The dataset was originally collected and de-identified by the respective data providers under their institutional review board approvals, and no new patient data were collected or accessed in this research.

References

[1] Kumar, K., & Jyoti, K. (2025, January 30). Enhancing transparency and trust in brain tumor diagnosis: An in-depth analysis of deep learning and explainable AI techniques. [CrossRef]

[2] Van der Velden, B. H., Kuijf, H. J., Gilhuijs, K. G., & Viergever, M. A. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical image analysis, 79*, 102470. [CrossRef]

[3] Dorfner, F. J., Patel, J. B., Kalpathy-Cramer, J., Gerstner, E. R., & Bridge, C. P. (2025). A review of deep learning for brain tumor analysis in MRI. *NPJ Precision Oncology, 9*(1), 2. [CrossRef]

[4] Kaissis, G. A., Makowski, M. R., Rückert, D., & Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nature machine intelligence, 2*(6), 305-311. [CrossRef]

[5] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ digital medicine, 3*(1), 119. [CrossRef]

[6] Nguyen, D. C., Ding, M., Pham, Q. V., Pathirana, P. N., Le, L. B., Seneviratne, A., ... & Poor, H. V. (2021). Federated learning meets blockchain in edge computing: Opportunities and challenges. *IEEE Internet of Things Journal, 8*(16), 12806-12825. [CrossRef]

[7] Zhu, L., Zhao, B., & Peng, R. (2026). Blockchain-enabled federated learning: A survey on system design, key challenges, and future directions. *Electronics, 15*(8), 1572. [CrossRef]

[8] de Aguiar, E. J., Traina Jr, C., & Traina, A. J. (2023). Security and privacy in machine learning for health systems: strategies and challenges. *Yearbook of medical informatics, 32*(01), 269-281. [CrossRef]

[9] Tjoa, E., & Guan, C. (2020). A survey on explainable artificial intelligence (xai): Toward medical xai. *IEEE transactions on neural networks and learning systems, 32*(11), 4793-4813. [CrossRef]

[10] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision* (*ICCV*) (pp. 618-626). IEEE. [CrossRef]

[11] Dhade, P., & Shirke, P. (2024). Federated learning for healthcare: a comprehensive review. *Engineering Proceedings, 59*(1), 230. [CrossRef]

[12] Kumar, R., Bernard, C. M., Ullah, A., Khan, R. U., Kumar, J., Kulevome, D. K., ... & Zeng, S. (2024). Privacy-preserving blockchain-based federated learning for brain tumor segmentation. *Computers in Biology and Medicine, 177*, 108646. [CrossRef]

[13] Yahiaoui, M. E., Derdour, M., Abdulghafor, R., Turaev, S., Gasmi, M., Bennour, A., ... & Sarem, M. A. (2024). Federated learning with privacy preserving for multi-institutional three-dimensional brain tumor segmentation. *Diagnostics, 14*(24), 2891. [CrossRef]

[14] Manthe, M., Duffner, S., & Lartizien, C. (2024). Federated brain tumor segmentation: An extensive benchmark. *Medical Image Analysis, 97*, 103270. [CrossRef]

[15] Bechar, A., Medjoudj, R., Elmir, Y., Himeur, Y., & Amira, A. (2025). Federated and transfer learning for cancer detection based on image analysis. *Neural Computing and Applications, 37*(4), 2239-2284. [CrossRef]

[16] Srinivasu, P. N., Lakshmi, G. J., Narahari, S. C., Shafi, J., Choi, J., & Ijaz, M. F. (2024). Enhancing medical image classification via federated learning and pre-trained model. *Egyptian Informatics Journal, 27*, 100530. [CrossRef]

[17] Sami, S. A., & Habib, M. A. (2024). Firefly Algorithm-Based Enhancement of MRI Images for Improved Brain Tumor Detection and Blockchain-Ensured Security. *Available at SSRN 4963964*. [CrossRef]

[18] Giri, A., Thapa, P., Banu, J. S., Poudyal, S., Rijal, B., & Karki, S. (2024). Harnessing ResUHybridNet with Federated Learning: A New Paradigm in Brain Tumour Segmentation. *Revue d'Intelligence Artificielle, 38*(3), 765. [CrossRef]

[19] Raheem, A., Yang, Z., Yu, H., Yaqub, M., Sabah, F., Ahmed, S., ... & Chuhan, I. S. (2024). IPC-CNN: A Robust Solution for Precise Brain Tumor Segmentation Using Improved Privacy-Preserving Collaborative Convolutional Neural Network. *KSII Transactions on Internet & Information Systems, 18*(9), 2589. [CrossRef]

[20] Albalawi, E., TR, M., Thakur, A., Kumar, V. V., Gupta, M., Khan, S. B., & Almusharraf, A. (2024). Integrated approach of federated learning with transfer learning for classification and diagnosis of brain tumor. *BMC medical imaging, 24*(1), 110. [CrossRef]

[21] Sharma, S., Guleria, K., Dogra, A., & Agrawal, S. K. (2024). A privacy-protected federated learning with cross-silo brain tumour dataset for glioma detection. *SN Computer Science, 5*(8), 1144. [CrossRef]

[22] Patel, Y., Gupta, R., Kakkar, R., Tanwar, S., & Guizani, M. (2024, May). Blockchain and dl-based brain tumour prediction scheme for telesurgery systems in healthcare 4.0. In *2024 International Wireless Communications and Mobile Computing* (*IWCMC*) (pp. 67-72). IEEE. [CrossRef]

[23] Ngoupayou Limbepe, Z., Gai, K., & Yu, J. (2025). Blockchain-based privacy-enhancing federated learning in smart healthcare: a survey. *Blockchains, 3*(1), 1. [CrossRef]

[24] Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., ... & Van Leemput, K. (2014). The multimodal brain tumor image segmentation

benchmark (BRATS). *IEEE transactions on medical imaging, 34*(10), 1993-2024. [CrossRef]

[25] Nyúl, L. G., Udupa, J. K., & Zhang, X. (2000). New variants of a method of MRI scale standardization. *IEEE transactions on medical imaging, 19*(2), 143-150. [CrossRef]

[26] Bhardwaj, N., Sood, M., & Gill, S. S. (2023, April). Data Pre-processing Techniques for Brain Tumor Classification. In *International Conference on Women Researchers in Electronics and Computing* (pp. 195-204). Singapore: Springer Nature Singapore. [CrossRef]

[27] Koc, A. B., & Akgun, D. (2026). A Review of U-Net Based Deep Learning Frameworks for MRI-Based Brain Tumor Segmentation. *Diagnostics, 16*(4), 506. [CrossRef]

[28] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016, October). 3D U-Net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention* (pp. 424-432). Cham: Springer International Publishing. [CrossRef]

[29] Milletari, F., Navab, N., & Ahmadi, S. A. (2016, October). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)* (pp. 565-571). IEEE. [CrossRef]

[30] Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC medical imaging, 15*(1), 29. [CrossRef]

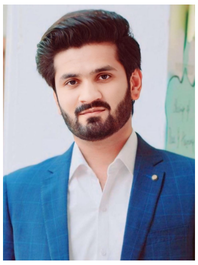
[31] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), PMLR 54* (pp. 1273-1282).

[32] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135-1144). ACM. [CrossRef]

Aiza Mukhtar is a researcher focused on AI, healthcare analytics, and IoT, aiming to develop smart tools for early disease detection and better healthcare through machine learning and deep learning. Her work also addresses cybersecurity challenges in medical systems. She is committed to advancing innovative, data-driven solutions for real-world clinical applications. (Email: misaizarao1122@gmail.com)

Aamir Ali is a dedicated researcher in artificial intelligence, healthcare analytics, and IoT environments. His expertise includes machine and deep learning applications in medical diagnosis, speech recognition, and facial emotion recognition. He has also worked on IoT-focused projects such as malware classification, anomaly detection, and cyberattack prediction. Aamir is passionate about developing AI-driven solutions for early disease detection and improving patient care. He actively contributes to interdisciplinary research across clinical and image data domains. (Email: amirali4436823@gmail.com)

Wajiha Farooq is a lecturer at COMSATS University Islamabad, Sahiwal Campus, in the Department of Computer Science. Her research interests include underwater wireless sensor networks (UWSNs), machine learning, deep learning, and edge computing. She is currently exploring AI-based solutions for real-world sensing and classification problems. (Email: wajihafarooq980@gmail.com)



Muhammad Aqib Fareed is a Lecturer in Computer Science and a dedicated researcher specializing in Artificial Intelligence (AI), healthcare analytics, and IoT-based systems. His research focuses on applying machine learning and deep learning techniques to medical diagnostics, intelligent healthcare solutions, and cybersecurity applications. With a strong passion for innovation and emerging technologies, he aims to develop smart and reliable systems for early disease detection, healthcare decision support, and secure digital environments. (Email: aqibfareed284@gmail.com)

Sadia Sohail is a researcher focused on AI, healthcare analytics, and IoT, aiming to develop smart tools for early disease detection and better healthcare through machine learning and deep learning. Her work also addresses cybersecurity challenges in medical systems. She is committed to advancing innovative, data-driven solutions for real-world clinical applications. (Email: amnas1071@gmail.com)

Azka Ahsan is a researcher focused on AI, healthcare analytics, and IoT, aiming to develop smart tools for early disease detection and better healthcare through machine learning and deep learning. Her work also addresses cybersecurity challenges in medical systems. She is committed to advancing innovative, data-driven solutions for real-world clinical applications. (Email: azkaahsan915@gmail.com)