IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 June 2025, accepted 11 July 2025, date of publication 17 July 2025, date of current version 25 July 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3590291

RESEARCH ARTICLE

BrainAR: Automated Brain Tumor Diagnosis With Deep Learning and 3D Augmented Reality Visualization

MERIEM KHEDIR[1], KAHINA AMARA[2], NASSIMA DIF[3], OUSSAMA KERDJIDJ[2], SHADI ATALLA[4], (Senior Member, IEEE), AND NAEEM RAMZAN[5], (Senior Member, IEEE)

[1]École Supérieure en Informatique, Sidi Bel Abbès 22000, Algeria
[2]Centre of Development of Advanced Techniques, Algiers 16303, Algeria
[3]École Supérieure en Informatique, LabRI-SBA Laboratory, Sidi Bel Abbès 22000, Algeria
[4]College of Engineering and Information Technology, University of Dubai, Dubai, United Arab Emirates
[5]School of Computing, Engineering and Physical Sciences, University of the West of Scotland, PA1 2BE Paisley, U.K.

Corresponding author: Kahina Amara (kahina.amara88@gmail.com)

ABSTRACT Augmented Reality (AR) technology offers promising applications in healthcare by enabling interactive 3D visualization of anatomical structures. However, current AR implementations often lack patient-specific detail, limiting their effectiveness in clinical settings. In this paper, we present BrainAR, an innovative mobile AR-based application designed for the automatic segmentation, 3D visualization, localization, and interaction with brain tumors using multiparametric 3D Magnetic Resonance Imaging (MRI) data. Our method leverages a 3D Residual U-Net, trained on the BraTS2021 dataset, achieving a mean Dice score of 0.886 for accurate tumor segmentation. The segmentation outputs are integrated into a real-time 3D engine to enable precise and dynamic visualization of brain tumors. Key contributions of our work include: 1) a server-side deployment of the segmentation model for online, patient-specific inference; 2) seamless AR integration enabling interactive exploration through hand gestures and voice commands; and 3) a mobile-based platform aimed at enhancing accessibility and usability in clinical environments. The proposed solution facilitates early detection and diagnosis by providing clinicians with an intuitive, immersive, and patient-specific tool for enhanced medical imaging interaction.

INDEX TERMS Brain tumor segmentation, computer-aided diagnosis, patient-specific 3D visualization, tumor localization, interaction, 3D U-Net, model deployment, MRI, augmented reality.

I. INTRODUCTION

Despite rapid advancements in medical technology, brain cancer continues to pose a significant challenge [1]. According to the World Health Organization's International Agency for Research on Cancer, brain cancer is one of the leading causes of cancer-related deaths worldwide. With over 300,000 cases reported annually, brain tumors continue to raise concerns within the international medical community [2]. This highlights the critical need for innovative diagnostic and treatment methodologies to facilitate early detection, prevent advanced cases of brain cancer, and improve surgical success rates. Brain tumor detection is traditionally a manual process where MRI scans are visually examined, which can be time-consuming and prone to errors. Despite modern artificial intelligence (AI) advancements for automating brain tumor diagnosis through deep learning, challenges remain in achieving sufficient tumor visualization, thereby constraining the interpretation of potential segmentation results from these methods.

Brain tumor diagnosis involves a multi-step approach that begins with a neurological examination assessing vision, hearing, reflexes, and coordination to localize potential abnormalities. If clinical findings suggest a tumor, advanced imaging techniques are employed, with contrast-enhanced MRI serving as the standard due to its superior soft-tissue

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Sharif.

resolution for delineating tumor boundaries and surrounding structures. Specialized MRI sequences including functional MRI (fMRI) for mapping eloquent cortex, perfusion MRI for assessing tumor vascularity, and magnetic resonance spectroscopy (MRS) for analyzing metabolic profiles provide critical data for surgical planning and differential diagnosis. When imaging indicates a mass, histopathological analysis via stereotactic biopsy or surgical resection is performed to determine tumor type and grade according to WHO criteria [2], integrating molecular profiling for precise classification. Complementary tools such as PET-CT (evaluating metabolic activity) and lumbar puncture further characterize tumor behavior and guide therapeutic decisions. This integrated diagnostic paradigm enables risk stratification and personalized treatment strategies. Existing solutions inadequately address critical challenges in brain tumor diagnosis and surgical planning: (1) Human error mitigation remains unresolved, as existing AI tools focus on desktop-based analysis without spatially anchored guidance to reduce errors during diagnosis or surgery; (2) Workload reduction is hindered by manual processes dominating tumor localization and resource planning, with no unified system to automate segmentation, 3D visualization, and intraoperative guidance; and (3) Clinical integration is compromised by AR system requiring unmanageable hardware or disjointed workflows, exacerbating workload inefficiencies.

Advanced medical imaging is essential for brain tumor detection, segmentation, and treatment, as it provides detailed and precise views of the brain's internal structures, including blood vessels, tumors [3], and other critical areas [4]. Studies demonstrate that deep learning shows significant promise in medical imaging, providing improved functionalities for the classification and segmentation of brain tumors. However, in AR systems for brain tumor diagnosis, the integration of AI-driven insights remains a major limitation. While AI is capable of processing medical imaging data and extracting significant diagnostic information, existing AR platforms face significant challenges in integrating these capabilities into patient-specific clinical environments. Deep learning models for complex tasks such as tumor segmentation and classification require significant computational resources, making online and patient-specific integration challenging. Using AR for brain tumor diagnosis faces key obstacles, particularly in the integration of high-quality medical imaging data into the AR environment. This paper proposes BrainAR, an approach to address these gaps. It uses AI-driven 3D tumor segmentation and AR overlay to automate tumor boundary detection and resource allocation, aiming to reduce workload and analysis time. It also supports collaboration through a mobile AR system designed to integrate into clinical workflows, allowing clinicians to visualize tumors in 3D while interacting with patients and medical teams.

The remainder of this paper is organized as follows: Section II reviews deep learning, augmented reality (AR), and mobile AR methods relevant to brain tumor diagnosis. Section III outlines the proposed methodology for integrating brain tumor segmentation models with augmented reality, including data collection, model training, deployment, and AR application development. Section IV presents a detailed discussion of the results. Finally, Section V concludes the paper and highlights potential directions for future research.

II. RELATED WORKS

A. DEEP LEARNING FOR BRAIN TUMOR SEGMENTATION

Artificial intelligence-driven image analysis can potentially improvee accuracy and efficacy of tumor segmentation [12], [13], [14]. Deep learning has emerged as a widely adopted approach in medical imaging [15], [16], [17], owing to its capacity to learn and extract complex features from large datasets [18]. Due to their high accuracy and adaptability, U-Net based architectures [19] are widely employed models for medical image segmentation tasks [20], [21], [22], particularly in brain tumor segmentation. Several studies have demonstrated the superior precision of U-Net-based models in the domain of brain tumor segmentation [5], [7], [9], [10], [11], [23], as outlined in Table 1. The paper, presented in [5], employed a 3D Dense U-Net to implement a triple network architecture, where each network predicts a specific tumor region (whole tumor, tumor core, or enhancing tumor) in the context of a binary segmentation task. The study highlighted the robustness of this method, particularly its minimal preprocessing requirements, efficient multi-volume fusion method to reduce false positives, and patch-based approach to manage memory constraints, which achieved high accuracy across multiple datasets. Another investigation, [6], proposed to train two ensembles of U-Net models and to combine their generated maps. The proposed approach was validated on the BraTS_2020 training dataset. In addition, the authors in [7] introduced a Binary U-Net model designed to identify tumor boundaries and classify tumor types, demonstrating high accuracy in heterogeneous tissue images.

Gupta et al. [9] conducted a comparative analysis of several U-Net models, including the original U-Net, Attention U-Net, Residual U-Net, ResUNet++, and Recurrent Residual U-Net, on brain tumor MRI scans. Their findings indicate that the Recurrent Residual U-Net model achieved the highest performance, with a Mean Intersection Over Union 'IoU' of 0.8665. Another study [8] introduced a Deep Convolutional Neural Network (DCNN) that attained a validation accuracy of 98% on the BraTS dataset through robust preprocessing and hyperparameter tuning.

A recent study [10] improved U-Net models by integrating tumor edges as learning targets, resulting in enhanced segmentation accuracy. Finally, the study [11] proposed a novel 3DUV-NetR+ architecture, which uses the encoding features of 3D U-Net and V-Net, fused at each depth, and enhanced by Transformers to capture complex spatial relationships and dependencies across 3D multimodal MRI data. This hybrid approach combines local and global contextual information, enabling more accurate segmentation of brain tumor sub-regions. It achieved high performance on the BraTS_2020 dataset, with Dice Similarity Coefficients

TABLE 1. Deep learning models and techniques used in the state of the art for brain MRI segmentation.

Reference	Model	Dataset	Optimizer	Loss	Metric
[5]	3D Dense U-Net	Brats2018	Adam	-	Dice for whole tumor = 0.90, Hausdorff distance for whole tumor = 6.0.
[6]	Ensemble of 3D U-Net	Brats2020	Adam	Dice Loss	Dice of 0.79, 0.89 and 0.84
[7]	Binary U-Net	LGG segmentation dataset	–	Dice loss	Dice=0.9289, Recall=0.9368, Precision=0.9712.
[8]	DCNN	BraTs2021	Adam	Categorical Cross-Entropy loss	Dice=0.90, Mean-IOU=0.91, Sensitivity=0.96, Specificity=0.99.
[9]	U-Net, Attention U-Net, Residual U-Net, ResUnet++, Recurrent Residual U-Net	The brain T1-weighted CE-MRI dataset	Adam	Binary Cross-Entropy loss	Precesion, Recall [9], F1-score [9], Mean-IOU [9].
[10]	U-Net, VNET, Attention U-Net, U-Net 3-plus, Swin-Unet, Hybrid MRU-Net.	BraTs2020	Adam	Focal loss	Dice=0.894, bi-direction Hausdorff95 distance = 6.79.
[11]	3DUV-NetR+	Brats2020	Adam	combinaison of Categorical Focal loss and Dice loss	Dice for whole tumor = 0.9195, Hausdorff distance for whole tumor = 4.9.

of 91.95% for Whole Tumor (WT), 82.80% for Tumor Core (TC), and 81.70% for Enhancing Tumor (ET).

B. AUGMENTED REALITY IN MEDICAL PRACTICE

Augmented reality interactive technology has emerged as a transformative technology in the medical field, it provides an intuitive method for visualizing human anatomy, enabling users to explore and navigate internal anatomical structures from multiple viewpoints [24], [25], [26]. In surgical applications, AR can significantly aid surgeons in tumor localization, delineating dissection planes, and reducing the risk of injuring hidden structures [27].

The advent of imaging equipment that provides basic anatomical and pathological information beyond just diagnosis has led to significant advancements in the treatment of brain tumors in recent years [27]. In the same context, AR provides substantial advantages, particularly when integrated with advanced imaging modalities such as MRI. The implementation of AR facilitates patient-specific representation of the brain and tumor in a three-dimensional format derived from MRI data, affording clinicians a more comprehensive understanding of the tumor's location and its spatial relationship with surrounding brain tissue. Although deep learning models excel in automating the segmentation of brain tumors, the interpretability of their outputs can be challenging without enhanced visualization techniques. AR mitigates this limitation by transforming segmented data into interactive and detailed 3D visualizations, thereby improving the interpretation of results generated by artificial intelligence.

The capability to manipulate and interact with 3D models, such as rotation, zooming, and switching between various views of the tumor and adjacent brain tissue, enhances the clarity and depth of insights gained from these visualizations. AR systems have been investigated in prior research to enhance intraoperative tumor visualization and improve segmentation accuracy during brain surgery. This research underscores AR's potential to enhance spatial awareness, alleviate cognitive strain, and streamline clinical workflows through 3D data integration.

C. MOBILE AUGMENTED REALITY FOR BRAIN TUMOR DIAGNOSIS ASSISTANCE

Reference [24] introduced a mobile augmented reality system for the visualization of 3D brain tumors utilizing facial feature tracking. The system calibrates the camera by estimating its orientation and position through a set of 3D points and their corresponding 2D projections. Upon calibration, a 3D reconstructed model of the brain tumor is presented at the estimated pose. In [28], an accessible augmented reality application for brain tumor diagnosis was introduced. The proposed workflow employs the Gradient Vector Flow (GVF) Snakes model for tumor segmentation, facilitating automatic segmentation, augmented reality visualization, and three-dimensional interaction with brain and tumor models. The authors evaluated their augmented reality application using a phantom head in a controlled environment, reporting average errors ranging from 2 mm to 4.5 mm on smartphone displays. Expanding on these, [29] presented a handheld augmented reality navigation (ARN) system that displayed

3D virtual structures overlaid on an image of the surgical field via a tablet PC monitor. The system was tested on 10 patients undergoing various procedures, and the surgeons evaluated its usefulness at each stage of surgery. The ARN system was beneficial for superficial tumor resections but less effective for deep-seated tumor resections, except when using transcortical or interhemispheric approaches. The study concluded that the tablet-based ARN system could be practical during craniotomy, dura and skin incisions, and superficial tumor resections, as well as for transcortical and interhemispheric approaches for deep-seated tumors.

The study presented in [30] examined the impact of augmented reality on brain tumor surgery through intraoperative visualization. They evaluated the usefulness of AR by considering factors such as surgical resection time, the duration/type/mode of AR, control usability, displayed objects, pointer-based navigation checks, and quality indicators in fifty-five intracranial lesion cases. These procedures were performed either with an AR-navigated microscope or conventional neuronavigation. The main advantage of AR visualization in brain tumor surgery is its integrated continuous display, enabling pointer-less navigation. The navigation view provides the highest usability while obstructing the operative field less frequently. Visualization quality could further improve with enhancements in registration accuracy and depth perception.

D. SHORTCOMINGS OF EXISTING APPROACHES

Mobile augmented reality faces notable hardware limitations, such as small screen size and resolution, which hinder the clear display of fine anatomical details on smartphones. Even high-resolution displays often require frequent zooming to view intricate structures, disrupting workflow and increasing the risk of misinterpretation. Additionally, mobile devices' narrow field of view (FoV) forces users to manually reposition themselves to examine large-scale models, fragmenting spatial understanding. In contrast, head-mounted AR systems offer a wider FoV that better supports comprehensive spatial visualization.

Patient-specific rendering of high-fidelity 3D medical models also places heavy demands on mobile GPUs, leading to overheating and rapid battery depletion, serious concerns in time-sensitive clinical settings. Clinicians may hesitate to adopt mobile AR due to unfamiliar interfaces or doubts about its reliability compared to established tools like radiology workstations. Integration and security challenges further hinder the adoption of mobile AR in healthcare. Many applications lack seamless interoperability with hospital systems, limiting access to patient-specific data during visualization.

Moreover, transmitting sensitive medical information to mobile devices introduces potential compliance risks unless strong encryption and offline functionality are implemented to ensure data privacy. This study aims to address the limitations of traditional methods by leveraging AI-driven techniques and augmented reality for brain tumor segmentation, paving the way for more effective and interactive diagnostic tools.

The selection of augmented reality over virtual reality was driven by the specific clinical and practical requirements of our brain tumor diagnosis framework. AR offers the critical advantage of preserving real-world context by overlaying 3D tumor models and diagnostic data directly onto the clinician's view, maintaining situational awareness in environments where interaction with physical elements such as patient scans, surgical tools, or team members is essential. In contrast, virtual reality (VR) immerses users in a fully virtual environment, isolating them from their surroundings and potentially hindering collaboration and decision-making during medical procedures. Additionally, AR allows for seamless integration into clinical workflows, enabling clinicians to use mobile devices or headsets without disrupting established practices. For instance, surgeons can visualize tumor projections on MRI scans while simultaneously consulting with colleagues. VR, however, necessitates disconnection from the physical workspace, which can reduce efficiency and limit its practicality in dynamic, fast-paced clinical settings. BrainAR integrates a mobile AR, edge-optimized deep learning, and ergonomic design to reduce error-proneness and workload, with a focus on deep learning-driven MRI segmentation:

-MRI Segmentation via Deep Learning: Traditional manual or semi-automated tumor segmentation in MRI is time-consuming and prone to intra-/inter-observer variability, especially for heterogeneous tumors with irregular boundaries. Existing automated tools often struggle with low-contrast regions or diffuse tumor margins. BrainAR introduces a deep learning model for heterogeneous tumor segmentation, leveraging a hybrid architecture, 3D U-Net with residual blocks, to prioritize ambiguous regions. Unlike conventional threshold-based or atlas-driven methods, our model dynamically adapts to tumor heterogeneity and achieves an 88.6% Dice score on challenging glioma cases. The integration of deep learning for MRI segmentation may reduce clinician workload by automating a traditionally manual, hours-long process into one completed in minutes. At the same time, attention mechanisms enhance accuracy by resolving boundary ambiguities in low-contrast regions. This contribution ensures reproducible, standardized tumor volumetry, directly supporting precise treatment planning and minimizing variability in clinical workflows.

-Server-Backed Deployment: Most AI-based segmentation tools are desktop-bound, limiting accessibility in resource-constrained settings and requiring high-end GPUs. Real-time 3D rendering of segmented tumors for AR visualization is computationally prohibitive on mobile devices. BrainAR adopts a server-backed edge computing pipeline, where MRI segmentation and 3D mesh generation occur on a centralized server, while lightweight AR clients (e.g., tablets, HoloLens) handle user interaction. This hybrid approach reduces latency by preprocessing MRI data on the server, and enables real-time AR visualization on mobile

devices without compromising segmentation quality. The server-backed deployment ensures accessibility by enabling clinicians to utilize AI tools on low-power devices (e.g., tablets in operating rooms), enhances efficiency by offloading GPU-intensive tasks like 3D rendering to centralized servers to avoid hardware bottlenecks, and streamlines clinical workflows through seamless interoperability with hospital PACS/RIS systems via DICOM standards.

-Spatial Contextualization for Error Reduction: AI segmentation outputs are often treated as opaque systems, leading to diagnostic uncertainties when predictions conflict with anatomical context. Existing AR tools lack seamless integration with AI outputs, forcing clinicians to mentally map 2D predictions to 3D anatomy. BrainAR bridges this gap by converting deep learning outputs into AR-compatible 3D models spatially registered to the patient's MRI coordinate system, introducing two key innovations: hybrid human-AI validation, where clinicians interactively adjust AR overlays using anatomical landmarks to resolve ambiguities, and real-time feedback, where discrepancies between AI predictions and AR-rendered anatomy trigger model recalibration, iteratively refining accuracy and ensuring alignment with clinical observations. By enabling clinicians to verify AI outputs against spatial context. This transparent visualization fosters clinician trust in AI tools by demystifying decision-making processes, while AR-guided intraoperative navigation ensures alignment with preoperative plans, minimizing surgical drift and enhancing precision during tumor resection.

III. PROPOSED APPROACH

This paper introduces the BrainAR application, a novel approach designed to address the challenges associated with manual segmentation of brain tumors. The proposed application combines artificial intelligence and augmented reality. The AI component involves training a deep learning model for automated tumor segmentation in 3D brain MRI data. The segmentation results are then used in an augmented reality application to produce a patient-specific interactive 3D visualization. The proposed framework is illustrated in Figure 1, this diagram illustrates the end-to-end processing pipeline from medical imaging input to augmented reality visualization. Raw multiparametric MRI scans (T1, T1ce, T2, FLAIR) undergo preprocessing including skull-stripping, N4 bias correction, and intensity normalization. The core AI component employs a 3D U-Net architecture trained on BraTS2021 data to segment tumor subregions. The model that demonstrated the best performance was then deployed on a local server using a Flask application, where post-processing techniques were integrated to enable online 3D rendering. An AR platform was subsequently developed using Unity 3D [31] and ARCore [32], handling the visualization and interaction with the segmentation results through a mobile interface. ARCore provides robust motion tracking capabilities through its feature point detection and simultaneous localization and mapping (SLAM) techniques. ARCore enables the application to track the device's position and orientation without external markers, ensuring a stable overlay of virtual content onto the physical environment. This markerless approach ensures persistent alignment between virtual content and the physical world without requiring predefined markers or external sensors.

BrainAR aims to enhance brain tumor diagnosis and clinical decision making by projecting interactive 3D tumor reconstructions from preoperative MRI, directly onto the patient's anatomy via augmented reality, enabling surgeons to visualize tumor boundaries, functional areas, and critical structures in real-time spatial context. This reduces diagnostic uncertainty, improving tumor margin delineation accuracy versus conventional imaging alone, while providing intraoperative guidance that boosts resection completeness, and shortens surgery time. By offering an intuitive AR overlay, BrainAR aims to support critical decisions like preserving eloquent brain zones during resection and differentiating tumor subtypes, ultimately serving as a precision navigation tool that translates imaging data into actionable clinical insights for improved surgical outcomes.

For practical operating room adoption, comprehensive clinical evaluation (e.g., multi-center trials measuring surgical outcomes) and end-user studies assessing surgeon usability, workflow integration, and cognitive load in live surgical settings will be essential to validate its real-world impact as a precision navigation tool that translates imaging data into actionable clinical insights.

A. DEEP LEARNING FOR BRAIN TUMOR SEGMENTATION

1) DATA COLLECTION

Training a deep learning model for brain tumor MRI image segmentation requires a high-quality dataset. A comparative study demonstrated that the BraTS (Brain Tumor Segmentation) dataset [33] was identified as the most widely used and suitable, particularly for 3D visualization in the AR application, due to its inclusion of 3D MRI scans that are ideal for automatic 3D rendering. In the proposed approach, the BraTS2021 version of the BraTs dataset was used [33], which includes 1251 segmented 3D MRI images. The BraTS2021 dataset contains multiparametric MRI (mpMRI) scans, including: T1 native, T1 post-contrast (T1CE), T2-weighted (T2), and T2 Fluid Attenuated Inversion Recovery (T2-FLAIR). Each mpMRI image is associated with a segmentation mask with four class labels: 0 for background, 1 for necrotic tumor core, 2 for peritumoral edematous/invaded tissue, and 4 for Gd-enhancing tumor. Each modality and its corresponding mask have dimensions of $240 \times 240 \times 155$.

2) DATA PREPROCESSING

To effectively segment brain tumors for 3D visualization and interaction, preprocessing is crucial for preparing the dataset for model training. In the proposed approach, first, cropping was applied to each MRI modality and its corresponding

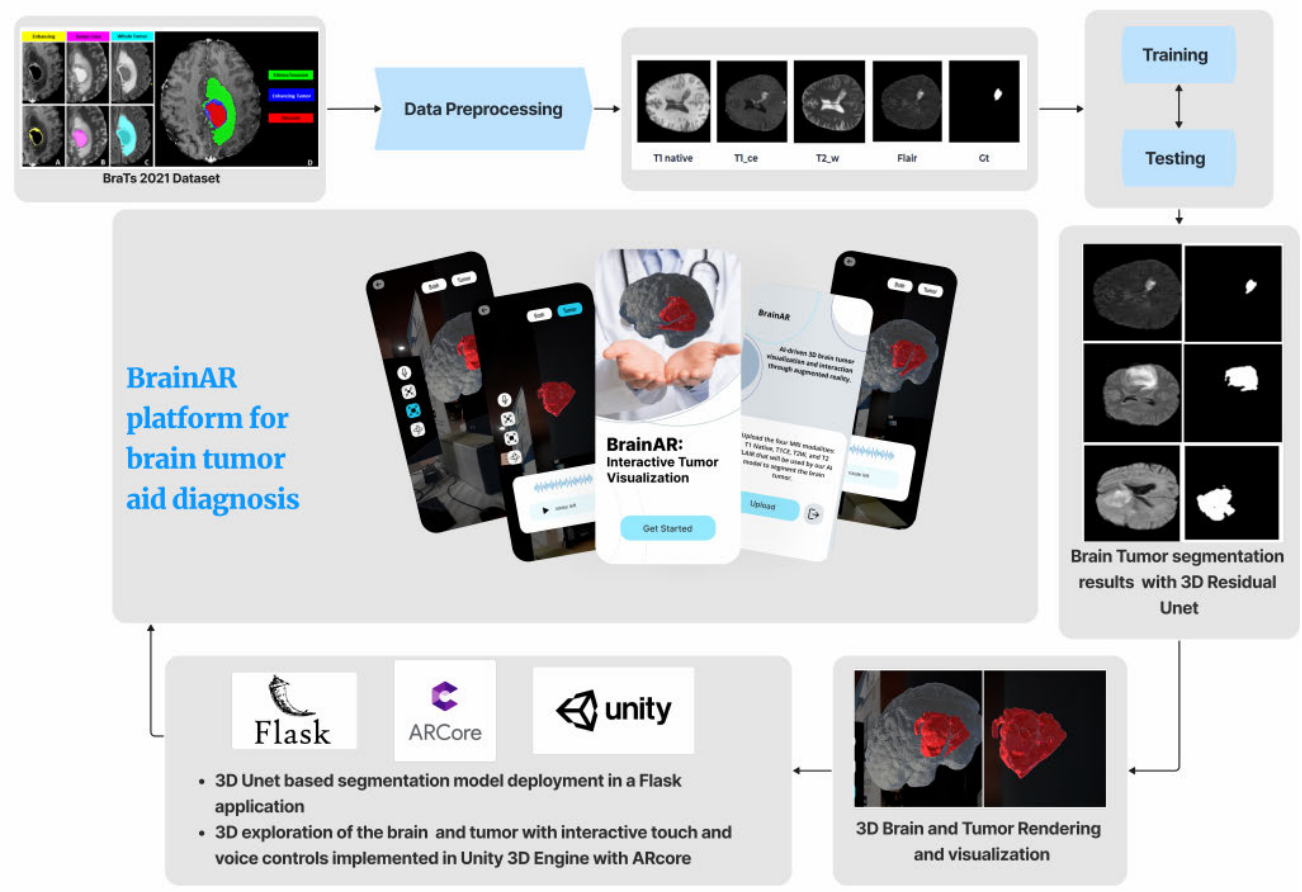


FIGURE 1. The workflow of the proposed framework.

segmentation mask to convert them to a smaller size of 192 × 192 × 128. Cropping was applied to reduce computational complexity and training time. The 192 × 192 × 128 dimension represents the smallest size that fully contains the brain tissue, with only the background being removed. Subsequently, all images were normalized by dividing each pixel value by 255, scaling them to a range of 0 to 1 to ensure faster and more stable convergence during model training. Each MRI modality (T1 native, T1CE, T2-weighted, T2-FLAIR) was standardized using the z-score technique to ensure consistency across modalities, with a mean of 0 and a standard deviation of 1. All modalities were then stacked into a single 4D image in the order of T1 native, T1CE, T2-weighted, and T2-FLAIR, resulting in images of size 4 × 192 × 192 × 128.

This fusion leverages complementary anatomical and pathological information across modalities, enhancing the model's ability to learn robust and discriminative features. For the segmentation masks, label transformation was applied to combine the tumor sub-regions (1, 2, 4) into a single label (1), representing the tumor, while label (0) was retained to represent the background class, resulting in binary masks. This simplification was intended to reduce model complexity, while also aligning with the primary goal of producing intuitive, whole-tumor visualizations in the AR application. To mitigate limitations from the modest original dataset size, we implemented rigorous data augmentation techniques exclusively during training. These included domain-specific transformations: random Gibbs noise (simulating MRI acquisition artifacts), 3D elastic deformations, multi-axis rotations (along x/y/z axes), scaling variations (zoom-in/out), and strategic flipping (axial/sagittal planes). Critically, each augmentation was applied to mutually exclusive image subsets to prevent compounding artifacts and preserve anatomical integrity. This protocol expanded the training set by 1,547 synthetically generated volumes, increasing the effective training pool from 1,251 to 2,547 samples while maintaining patient distribution fidelity.

3) MODEL TRAINING

In recent years, U-Net architectures [19] have become the most popular deep learning models for medical data segmentation. This is mostly because they are accurate and good at processing different kinds of medical scans, as explained in the literature [7], [9], and [10]. This study focuses on selecting U-Net variants that offer high performance while maintaining a reduced number of parameters. The analysis identified the Residual U-Net [34] and the Attention U-Net [35] as

promising architectures for evaluation on the pre-processed dataset.

The 3D Residual U-Net [34] employed in this study is an enhanced version of the U-Net architecture, it has residual units that include a mechanism for dimension matching, as described in [36]. Each layer of the network consists of an encoding and decoding path with a skip connection. Unlike traditional U-Net implementations, the encoding path achieves downsampling through strided convolutions, while the decoding path performs upsampling using strided transpose convolutions at the beginning of each block. This Residual 3D U-Net is composed approximately of 4.8 million parameters and five layers, each with a stride of 2 and a kernel size of 3. On the other hand, the Attention U-Net [35] used in this study is composed of approximately 5.9 million parameters across five layers, each with a stride of 2 and a kernel size of 3. A final sigmoid activation layer was added to both networks to output probabilities between 0 and 1, representing the two classes: tumor or background.

Both the 3D residual U-Net and 3D attention U-Net models were trained using the Adam optimizer with an initial learning rate of 0.001. A StepLR learning rate scheduler was employed to decay the learning rate of each parameter group by gamma = 0.1 every five epochs. Early stopping was applied to prevent overfitting by stopping the training if the validation loss did not improve for three consecutive epochs. The choice of hyperparameters was informed by a review of state-of-the-art literature, where the Adam optimizer and a learning rate of 0.001 are commonly adopted for medical image segmentation tasks. Due to the computational demands of processing high-dimensional 3D MRI data with four input channels, conducting an extensive hyperparameter search was not feasible.

During the training of the 3D Residual U-Net, two loss functions were evaluated. The first (Equation 1) combined Binary Cross-Entropy (BCE) loss with Dice loss, while the second (Equation 2) combined Dice loss with Focal loss. The batch size was limited by GPU memory, with a maximum value of 4. Using the loss 1, the model converged in just 7 epochs, whereas training with the loss 2 required 25 epochs to converge.

For the 3D Attention U-Net, the second loss function 2 was selected based on its superior performance in the Residual U-Net. Due to memory constraints, the batch size was set to 2. The model converged after 21 epochs, with early stopping triggered when no improvement in validation loss was observed.

$$L1 = 0.5 \cdot Dice + 0.5 \cdot BinaryCrossEntropy \tag{1}$$

$$L2 = 0.5 \cdot Dice + 0.5 \cdot Focal \tag{2}$$

4) EVALUATION METRICS

The performance of the models was evaluated using several metrics: Dice score, 95th percentile Hausdorff distance (HD95), Intersection Over Union (IOU), Specificity, and Sensitivity. As described in equation 3, the Dice score computes the overlap between the predicted segmentation and the ground truth.

$$Dice = \frac{2 \cdot |P \cap G|}{|P| + |G|} \tag{3}$$

where:

- P (Predicted): The predicted segmentation region.
- G (Ground Truth): The actual (true) segmentation region.
- $|P|$ and $|G|$: The areas of the predicted and ground truth regions, respectively.
- $|P \cap G|$: The overlap area between the predicted and ground truth segmentation.

The 95th percentile Hausdorff Distance (HD95) (equation 4) indicates the spatial distance between the boundaries of the predicted and actual segmentation, using the 95th percentile of the boundary distances rather than the maximum, making it less sensitive to outliers.

$$\mathrm{HD}_{95}(A, B) = \max\Big\{\mathrm{percentile}_{95}\big(\{\min_{b \in B} d(a, b) : a \in A\}\big), \mathrm{percentile}_{95}\big(\{\min_{a \in A} d(b, a) : b \in B\}\big)\Big\} \tag{4}$$

where:

- $d(a, b)$: The distance between points a and b in the metric space.

The Intersection over Union (IoU) measures the intersection ratio to the union of the predicted and true regions, as illustrated in equation 5.

$$IoU = \frac{|P \cap G|}{|P \cup G|} \tag{5}$$

where:

- P (Predicted): The predicted segmentation region.
- G (Ground Truth): The actual (true) segmentation region.
- $|P \cap G|$: The area of intersection between the predicted and ground truth regions.
- $|P \cup G|$: The area of the union of the predicted and ground truth regions.

Specificity quantifies the proportion of correctly identified negative samples (equation 6), and sensitivity reflects the model's ability to correctly identify true positives, as illustrated in equation 7.

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

where:

- TN (True Negatives): The number of correctly predicted negative instances.
- FP (False Positives): The number of negative instances that were incorrectly classified as positive.

$$Sensitivity = \frac{TP}{TP + FN} \tag{7}$$

where:

- *TP* (True Positives): The number of instances that were correctly predicted as positive.
- *FN* (False Negatives): The number of positive instances that were incorrectly predicted as negative.

Before applying these metrics, the segmentation results were binarized to obtain binary masks comparable to the ground truth, using two different threshold values: 0.45 and 0.5. The threshold of 0.5 was retained, as it yielded better results in most cases on both the validation and test sets. For the IoU and 95th percentile Hausdorff Distance metrics, one-hot encoding was applied before computation.

B. MODEL DEPLOYMENT

For deployment, we selected the best-performing deep learning model and integrated it into a Flask[1] application hosted on a server to handle the computational tasks involved in MRI analysis efficiently. As illustrated in Figure 2, the distributed system architecture separates compute-intensive processing from real-time visualization tasks. The client application on a smartphone manages AR tracking, rendering, and user interactions. Communication occurs through HTTPS, this design allows model inference while maintaining responsive AR visualization, crucial for intraoperative use where computational resources are constrained. This setup allows users to send MRI scans from the mobile application to the server via a POST request. Upon receiving the MRI data, the API initiates a series of pre-processing steps, including cropping, normalization, standardization, and image fusion. The preprocessed data is then passed to the model, which generates a segmentation mask of the tumor. Post-processing is then applied to refine the output, including binarizing the mask by applying a threshold of 0.5. Padding is then used to restore the mask from the deep learning model's output size of $192 \times 192 \times 128$ back to the original dimensions of MRI scans, ensuring proper spatial alignment. The binary mask is then overlaid onto the corresponding brain image, with the mask visualized in red and the brain in grayscale for accurate localization. Finally, the result is converted into OBJ format and encoded in Base64 to enable transfer within the POST response. The processed result is packaged into JSON format and sent back to the client via a POST response. The client application extracts the Base64-encoded string from the JSON and decodes it to retrieve the OBJ format, which can be easily handled by AR libraries. This data is then used to automatically generate an interactive 3D visualization of the tumor and the brain, enabling users to examine and interact with the image in an AR environment. The deployment process was initially successful when tested on a local machine, which served as the backend server. It was equipped with an NVIDIA T550 GPU (12GB VRAM), a 12th Gen Intel® Core™ i7-1260P CPU, 16GB of RAM, and ran Python 3.10.13 with PyTorch 2.2.0 and CUDA 11.8.

In designing the ARBrain mobile app, ergonomic considerations were carefully integrated to ensure an efficient, comfortable, and user-friendly experience for medical professionals. The application features real-time 3D visualization of the segmented brain tumour within an anatomically accurate 3D brain model, enabling clinicians to explore and understand tumour location, size, and structure from multiple perspectives. The AR interface is optimized for handheld devices, minimizing the need for excessive arm movement and reducing user fatigue. To further enhance usability, the app supports multimodal interaction, including intuitive hand gesture control for manipulating the 3D model (e.g., rotate, zoom, highlight regions) and vocal commands for hands-free operation, which is particularly useful in sterile or multitasking environments. The interface design prioritizes clarity, with high-contrast visuals, accessible button placement, and support for both right- and left-handed use. These ergonomic and interaction design choices ensure that ARBrain is suitable for extended clinical use in diverse lighting and working conditions.

This system offers significant advantages by managing all heavy computational tasks on a remote server, thereby maintaining a lightweight and responsive user interface. The server GPU was utilized to accelerate the inference process, resulting in faster response times. This approach enables access to the application from various devices, promoting broader usability. Furthermore, it enhances overall performance, facilitating a streamlined and efficient workflow, which contributes to an improved user experience.

C. AUGMENTED REALITY MOBILE PLATFORM

The use of mobile devices and smartphones to assist physicians and healthcare professionals is rapidly increasing [37]. Given their portability, accessibility, and ease of use, developing an AR mobile platform for medical assistance would be beneficial [37]. However, several factors need to be considered, including creating a user-friendly interface, enabling intuitive interactions, and incorporating 3D visualization. These aspects are crucial to ensuring that the application is easy to use and lightweight for end users.

As a first step, Figma[2] was used to create a prototype for the application, starting with a minimum viable product (MVP) that includes only the essential buttons and screens. As shown in figure 3, the initial screen (screen a) displayed upon launching the app contains a ''Get Started'' button, which leads to the second screen (screen b) for uploading MRI scans. This screen features an ''Upload'' button to submit the scans, an ''Exit'' button to close the app, and a brief description of the application. Once the user uploads the MRI scans via screen (c), they are redirected to the augmented reality environment (figure 6), where spatial repositioning uses a three-finger drag gesture tracked via smartphone' articulated hand tracking and 3D models of the brain and tumor can be visualized. The switch buttons

[1] https://flask.palletsprojects.com/

[2] https://www.figma.com/fr-fr/

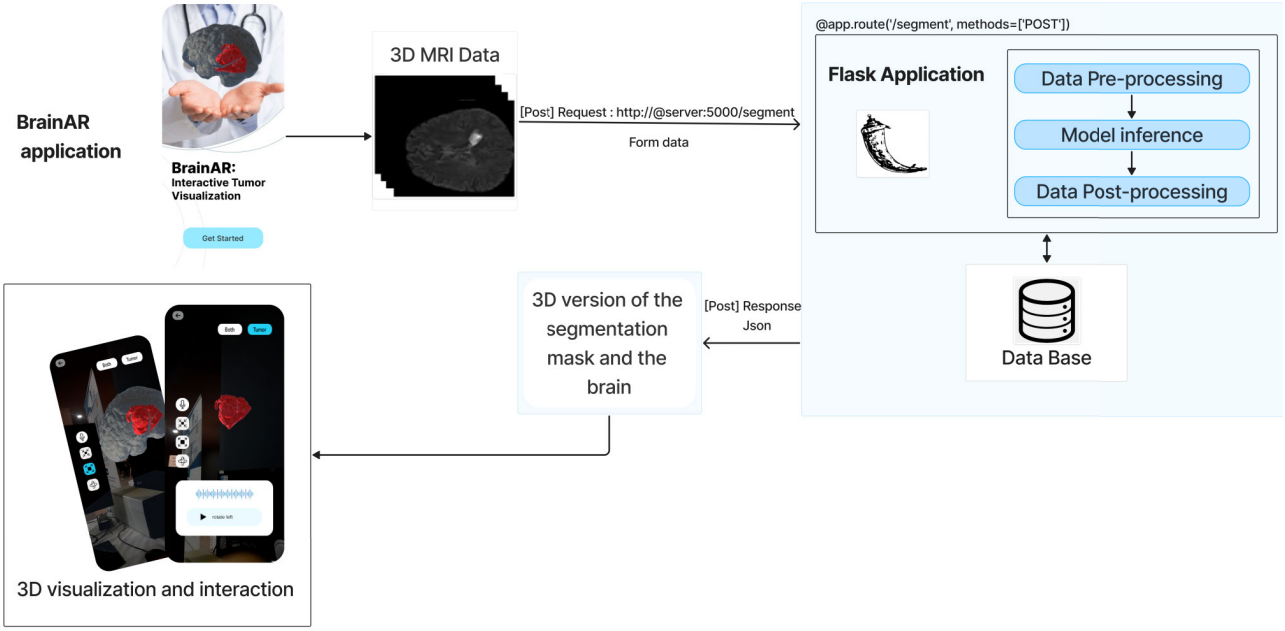
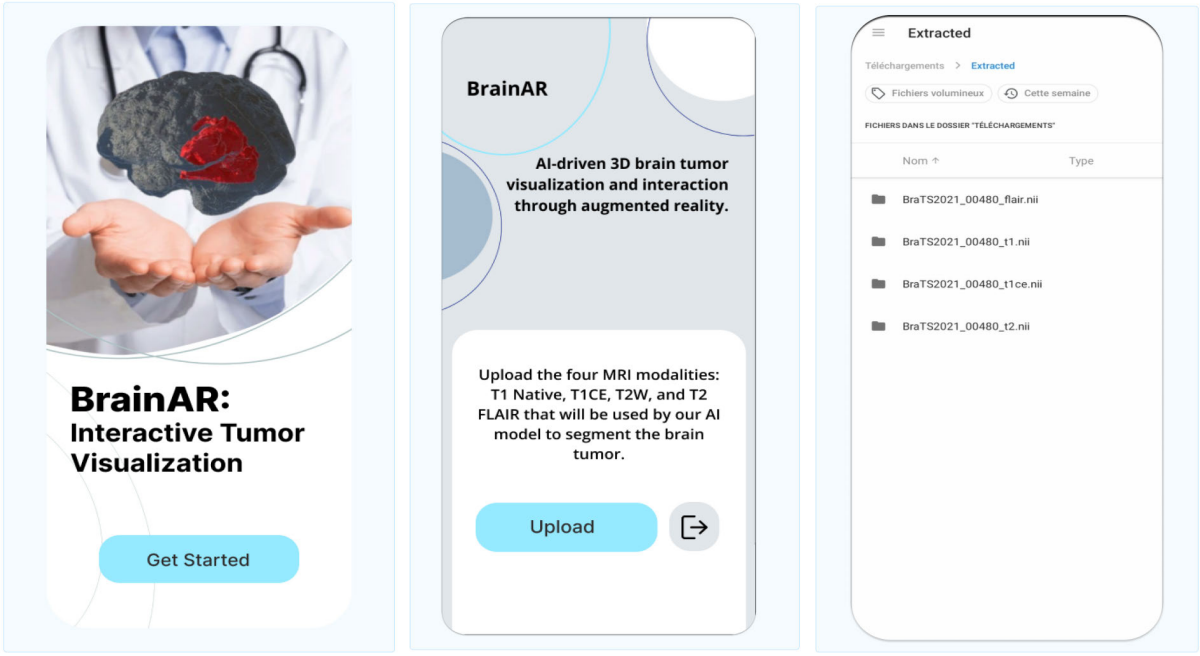


FIGURE 2. The Client-Server architecture utilized in the BrainAR system.



(a) (b) (c)

FIGURE 3. (a) The Get-Started screen of the BrainAR application, (b) The Upload screen of the BrainAR application, (c) Uploading MRI scans from local storage.

highlighted in Figure 4 allow users to toggle between the tumor and the brain 5, volumetric rendering demonstrates clinical value and employs depth-enhanced rendering techniques. The brain model maintains transparency to preserve situational awareness. While the interaction buttons facilitate user engagement. The microphone icon represents voice interaction, as shown in figures 9, while the other buttons control the rotation 8, zooming 7, and translation of the 3D model 6. Additionally, a ''Back'' button is included to return to the upload screen for the option to try different MRI scans if needed.

The application was developed using Unity 3D in conjunction with the ARCore library. Initially, the MRI scans are accessed from local files and subsequently sent via POST request to the server for processing. Upon receiving the response, the application utilizes the OBJ file format

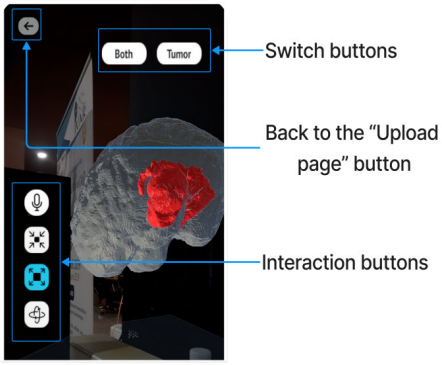


FIGURE 4. AR environment with different visualization and interaction buttons.

to prepare the 3D models of the brain and tumor for visualization. The interaction logic, which includes manipulation of the 3D models, is implemented in C# within the Unity 3D environment. Gesture interactions, such as rotating, zooming in and out, and repositioning the model were implemented using ARCore's hand tracking capabilities, enabling intuitive actions like tap, swipe, and pinch to select, rotate, or scale the 3D object. In parallel, voice commands were integrated via the SpeechToText[3] API, which converts spoken words into text and automatically maps them to the corresponding transformations of the 3D model. These voice commands were carefully selected based on usability testing and user-centered design principles to ensure they are intuitive, contextually relevant, and effective in supporting hands-free interaction within the AR environment.

IV. EXPERIMENTAL STUDY

First, the dataset, composed of 1251 3D multi-parametric MRI images, was divided into three subsets: 80% for training, 10% for validation, and 10% for testing, resulting in 1000 images for training, 125 for validation, and 126 for testing. Table 2 highlights the results obtained on the test set. When the 3D Residual U-Net was trained using a combination of Dice and binary cross-entropy (loss 1), it converged quickly within just 7 epochs. However, this rapid convergence limited the model's ability to fully capture the underlying data features, resulting in suboptimal performance, as reflected in the evaluation metrics. In contrast, training with a combination of Dice and focal loss (loss 2), which is more effective at handling class imbalance, led to improved metric scores, particularly in terms of the Dice coefficient. The 3D Attention U-Net was subsequently trained using the loss2 combination, which had shown superior results in the previous training. This model converged after 21 epochs. Despite this, the best overall performance was achieved by the 3D Residual U-Net. These findings highlight the effectiveness of the 3D Residual U-Net architecture compared to the 3D Attention U-Net when trained on the original dataset.

[3]https://github.com/yasirkula/UnitySpeechToText

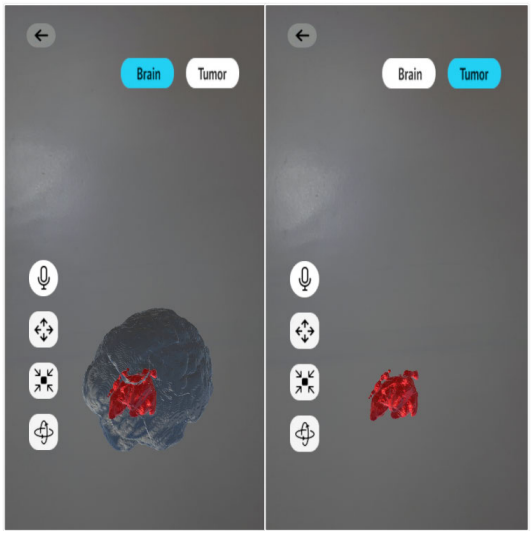


FIGURE 5. Brain and tumor visualizations.

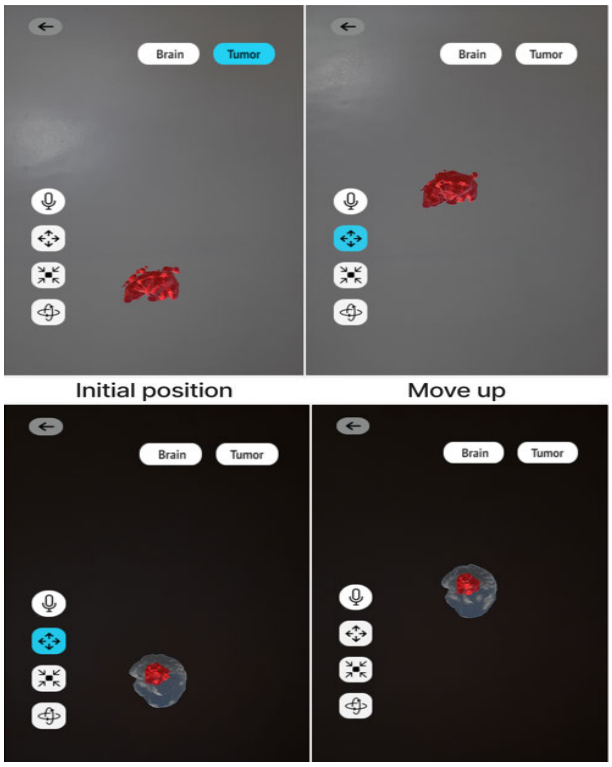


FIGURE 6. Translation interaction.

To enhance the results, data augmentation was implemented to increase the variability of the training set. Due to the extended training time, it was impractical to retrain all models. Therefore, the 3D Residual U-Net which had initially shown the best performance was retrained on the augmented dataset using loss2. As shown in Table 2, the Dice score improved by 0.005, the Hausdorff distance by 0.809, the IoU by 0.004, and the sensitivity by 0.002. Although these improvements are modest, even after we applied data augmentation techniques, they indicate that

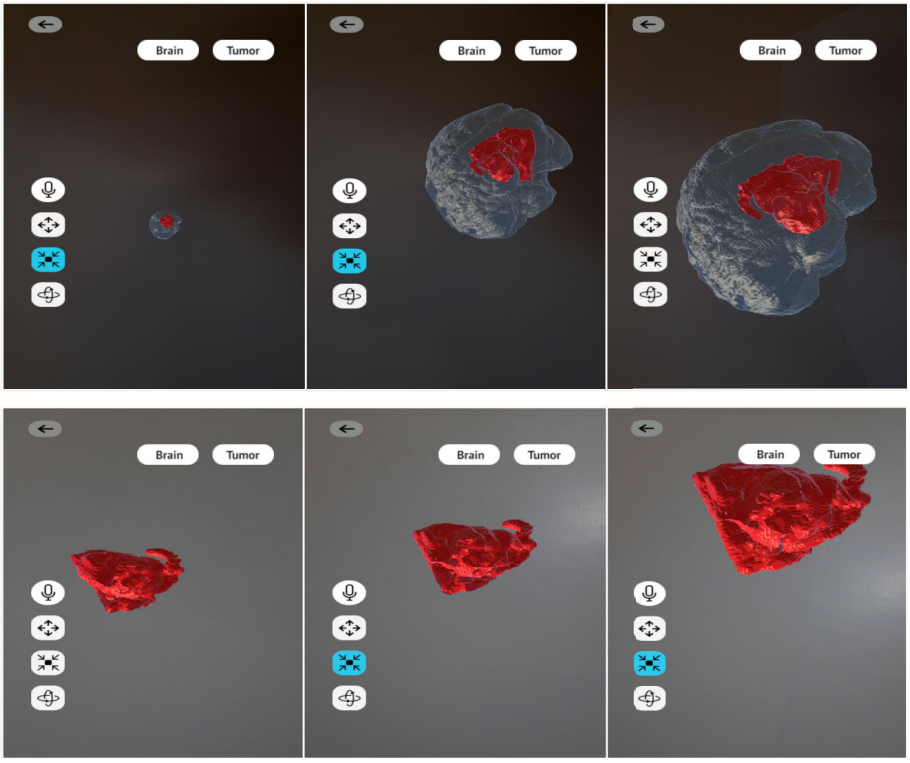


FIGURE 7. Zoom interaction.

TABLE 2. Obtained results on the test set.

Model	Dice	95% Hausdorff distance (HD)	IOU	Sensitivity	Specificity
3D Residual U-Net with loss (equation 1)	0.799	17.671	0.844	0.874	0.996
3D Residual U-Net with loss (equation 2)	0.881	10.272	0.898	0.889	**0.998**
3D Attention U-Net with loss (equation 2)	0.874	10.919	0.893	0.874	**0.998**
3D Residual U-Net with loss (equation 2) and data augmentation	**0.886**	**9.463**	**0.902**	**0.891**	**0.998**

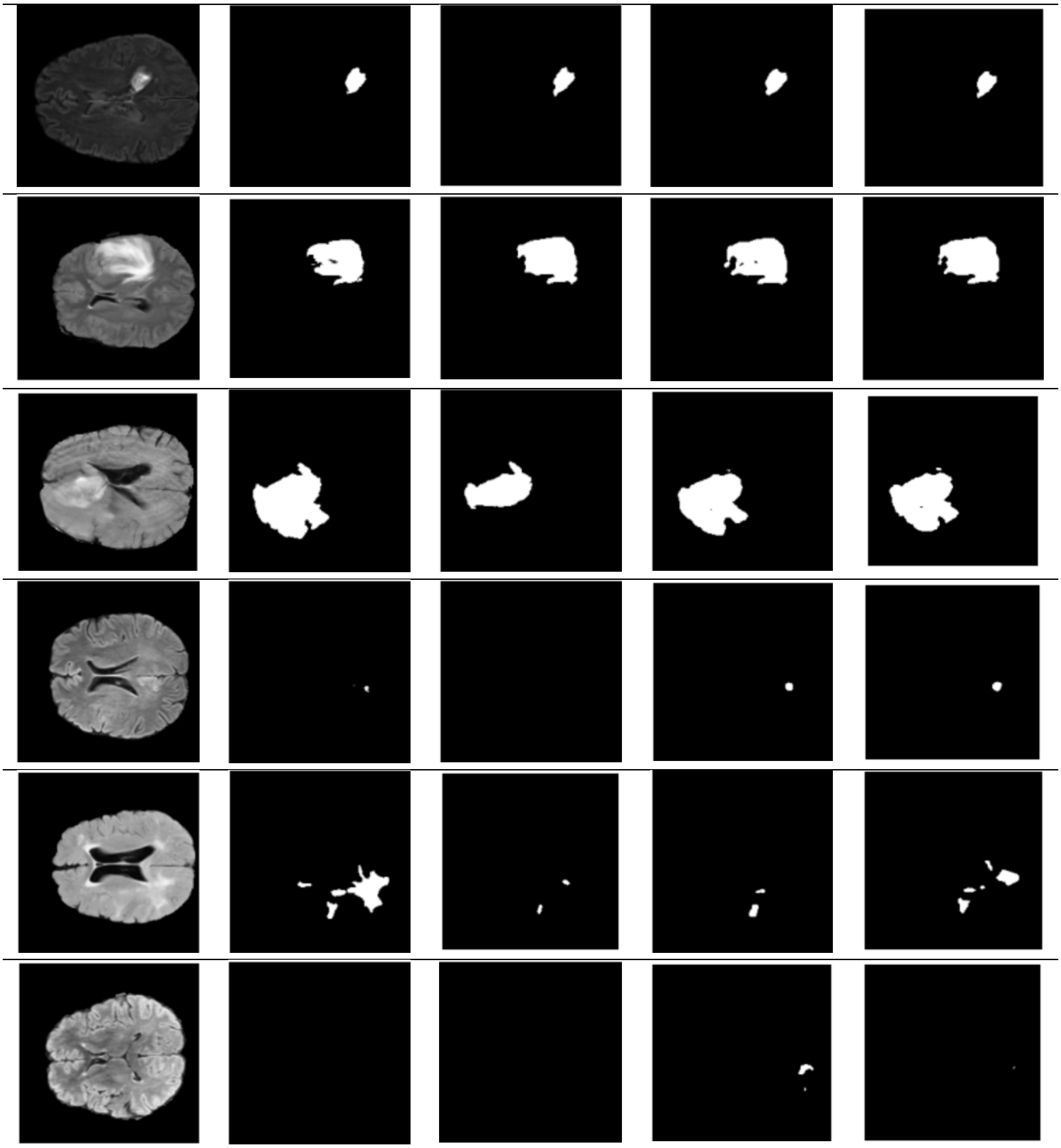
our model is highly sensitive to class imbalance. This is because the background class is significantly larger than the tumor region, and the data augmentation may not have been sufficient to address this issue. Across all models, a notable difference between sensitivity and specificity was observed, indicating that the models often missed positive cases (false negatives) while accurately identifying most negative cases (true negatives). These results also demonstrate the effectiveness of using focal loss, which has led to better performance compared to binary cross-entropy, as focal loss handles class imbalance more effectively. Table 3 presents sample prediction results from the trained models on the test set. These images were obtained from the 64th slice along the depth dimension of the 3D mask.

As demonstrated in both the results in Table 2 and Table 3, the 3D residual U-Net model, trained on the augmented dataset, was selected for deployment due to its superior performance.

The BrainAR application demonstrated strong performance in terms of timing. The deployment of the model on the server equipped with a GPU significantly reduced the computational load on the mobile device, thereby minimizing response times. Benchmarking over five test cases showed an average end-to-end prediction time of 61.57 seconds (standard deviation: 21.74 seconds), with a 95th percentile of 93.80 seconds. These metrics reflect the entire processing pipeline from image upload and preprocessing to segmentation and 3D rendering, supporting responsive, on-demand AR visualization for clinical use. As demonstrated in Figure 7, the 3D visualization of the tumor produced by the system achieves a rendering quality that is visually comparable to manual reconstruction methods, while being fully automated

TABLE 3. Prediction results on the test set using the 3D Attention U-Net and the 3D Residual U-Net trained with loss 2, both with and without data augmentation.

Flair	Ground truth mask	3D Attention U-Net	3D Residual U-Net before data augmentation	3D Residual U-Net after data augmentation



and requiring no manual intervention for either reconstruction or rendering.

Unlike non-mobile AR systems such as the HoloLens, which require expensive hardware, this platform is significantly more cost-effective and portable, making it ideal for use in resource-constrained environments. The system provides an accessible and portable solution, reducing the time required for diagnosis through AI-driven segmentation while offering AR-based support for surgical planning. This combination has the potential to impact patient outcomes by enabling early detection and improving surgical preparation, which may contribute to reduced mortality rates and higher treatment success. However, the precision of segmentation remains a key limitation that requires further improvement in the future.

Another limitation of the interaction experience stems from ARCore's requirement for surface detection,[4] which can make object manipulation, like zooming or moving

[4]https://developers.google.com/ar/design/content/content-placement

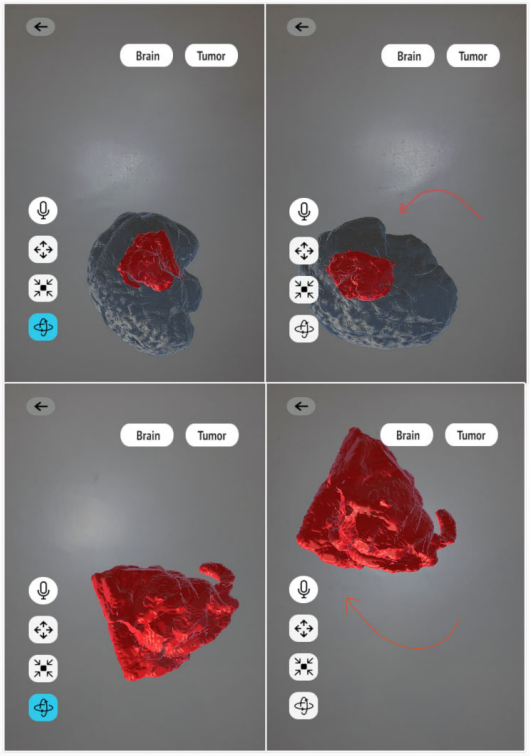


FIGURE 8. Rotation interaction.

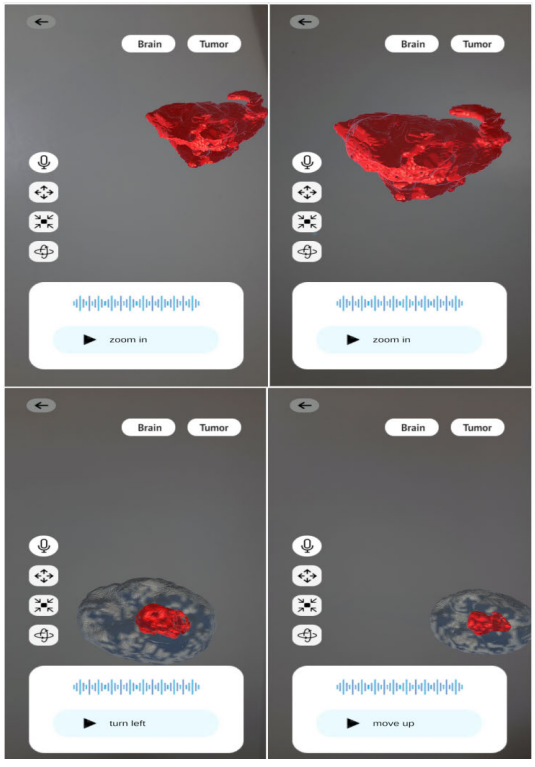


FIGURE 9. Voice interaction.

less seamless, sometimes requiring users to reposition their device to reorient the model. While this does not affect the functionality, it may impact the fluidity of interaction in some cases. Although this prototype represents the first iteration of a novel system, it can potentially substantially impact the healthcare sector. The developed approach could be applied to other types of cancer or medical conditions, provided that the deep learning model is adapted to the specific task. To rigorously validate the application's ease of use, we plan to conduct a formal evaluation study involving a diverse cohort of users, ranging from novice users to domain experts.

V. CONCLUSION

This paper presents a system that combines deep learning with augmented reality to enhance brain cancer diagnosis and treatment by providing patient-specific 3D localization, augmented reality visualization, and interactive capabilities for segmented brain tumors. Leveraging a 3D residual U-Net architecture, we achieved automatic tumor segmentation from 3D multiparametric MRI data. The model was integrated into a Flask-based backend and deployed on a local server to manage computationally intensive tasks efficiently. This setup was then interfaced with an augmented reality application, BrainAR, enabling online, interactive 3D visualizations of the segmented tumors. The developped mobile application enhances accessibility and portability, allowing healthcare professionals to use the system on various devices, including smartphones and tablets, regardless of location. Furthermore, the server-side deployment offloads computationally intensive tasks, ensuring that even resource-constrained devices can access and utilize the system, making it broadly accessible to a wide range of medical professionals without the need for specialized hardware. However, this work represents an initial version, requiring several improvements.

Future enhancements will focus on three key directions to advance both technical robustness and clinical utility: model capabilities, user interaction, and segmentation performance. First, we plan to extend the framework to multi-label segmentation, enabling simultaneous delineation of tumors, edema, and healthy tissues to provide clinicians with a holistic, anatomically informed view that enhances diagnostic efficiency. Additionally, we aim to integrate tumor classification and staging by coupling segmentation outputs with lightweight prognostic modules, enabling end-to-end analysis of anatomical and prognostic features. Second, user interaction will be refined through dynamic 3D visualization tools such as adjustable resolution rendering, region-of-interest exploration, and voice command integration. We also aim to overcome current limitations of ARCore, particularly in object placement and tracking stability, to improve usability and provide a more seamless and responsive AR experience in clinical contexts. Third, segmentation accuracy will be improved by addressing class imbalance via hybrid strategies: resampling minority classes, class-weighted loss functions, and a dual-model architecture where one network detects tumors and another refines their boundaries. To strengthen model generalization and ensure clinical applicability, we will conduct cross-domain validation using external datasets acquired with varied MRI

protocols. Additionally, efforts will be directed toward strengthening data security and user confidentiality, including the implementation of authentication systems and data encryption mechanisms to protect sensitive medical information and ensure compliance with privacy standards in clinical environments. These technical upgrades will be paired with longitudinal patient-specific statistics to support personalized clinical decision-making. Collectively, these efforts aim to bridge granular computational analysis with real-world diagnostic workflows, ensuring both precision and usability in clinical deployments.

REFERENCES

[1] T. T. T. Nguyen, L. A. Greene, H. Mnatsakanyan, and C. E. Badr, "Revolutionizing brain tumor care: Emerging technologies and strategies," *Biomedicines*, vol. 12, no. 6, p. 1376, Jun. 2024. [Online]. Available: https://www.mdpi.com/2227-9059/12/6/1376

[2] A. Cohen-Gadol. *72 Must-Know Brain Tumor Statistics (2024)*. Accessed: Oct. 2024. [Online]. Available: https://www.aaroncohen-gadol.com/en/patients/brain-tumor/types/statistics

[3] K. Amara, M. A. Guerroudji, O. Kerdjidj, N. Zenati, S. Atalla, and N. Ramzan, "Enhancing arrhythmia diagnosis through ECG deep learning classification deploying and augmented reality 3D heart visualization and interaction," *IEEE Access*, vol. 13, pp. 103198–103219, 2025.

[4] K. Amara, H. Kennouche, A. Aouf, O. Kerdjidj, N. Zenati, O. Djekoune, and M. A. Guerroudji, "Augmented reality for medical practice: A comparative study of deep learning models for Ct-scan segmentation," in *Proc. Int. Conf. Adv. Technol. Electron. Electr. Eng. (ICATEEE)*, Nov. 2022, pp. 1–6.

[5] C. G. B. Yogananda, B. R. Shah, M. Vejdani-Jahromi, S. S. Nalawade, G. K. Murugesan, F. F. Yu, M. C. Pinho, B. C. Wagner, K. E. Emblem, A. Bjørnerud, B. Fei, A. J. Madhuranthakam, and J. A. Maldjian, "A fully automated deep learning network for brain tumor segmentation," *Tomography*, vol. 6, no. 2, pp. 186–193, Jun. 2020. [Online]. Available: https://www.mdpi.com/2379-139X/6/2/186

[6] T. Henry, A. Carré, M. Lerousseau, T. Estienne, C. Robert, N. Paragios, and É. Deutsch, "Brain tumor segmentation with self-ensembled, deeply-supervised 3D U-Net neural networks: A BraTS 2020 challenge solution," in *Proc. Int. MICCAI Brainlesion Workshop*, Lima, Peru. Cham, Switzerland: Springer, Oct. 2020, pp. 327–339.

[7] N. Rao, D. L. S. Reddy, and H. Gujja. (Sep. 2022). *Brain MRI Segmentation Binary U-Net 875 Based Architecture Using Deep Learning Algorithm*. [Online]. Available: https://www.researchsquare.com/article/rs-1916275/v1

[8] A. M. Mostafa, M. Zakariah, and E. A. Aldakheel, "Brain tumor segmentation using deep learning on MRI images," *Diagnostics*, vol. 13, no. 9, p. 1562, Apr. 2023. [Online]. Available: https://www.mdpi.com/2075-4418/13/9/1562

[9] A. Gupta, M. Dixit, V. K. Mishra, A. Singh, and A. Dayal, "Brain tumor segmentation from MRI images using deep learning techniques," in *Advanced Computing* (Communications in Computer and Information Science), vol. 1781, D. Garg, V. A. Narayana, P. N. Suganthan, J. Anguera, V. K. Koppula, and S. K. Gupta, Eds., Cham, Switzerland: Springer, 2023, pp. 434–448. [Online]. Available: https://link.springer.com/10.1007/978-3-031-35641-436

[10] S. Sahayam and U. Jayaraman, "Integrating edges into U-Net models with explainable activation maps for brain tumor segmentation using MR images," 2024, *arXiv:2401.01303*.

[11] I. Aboussaleh, J. Riffi, K. E. Fazazy, A. M. Mahraz, and H. Tairi, "3DUV-NetR+: A 3D hybrid semantic architecture using transformers for brain tumor segmentation with MultiModal MR images," *Results Eng.*, vol. 21, Mar. 2024, Art. no. 101892. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2590123024001452

[12] R. Ranjbarzadeh, A. Caputo, E. B. Tirkolaee, S. Jafarzadeh Ghoushchi, and M. Bendechache, "Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools," *Comput. Biol. Med.*, vol. 152, Jan. 2023, Art. no. 106405. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482522011131

[13] M. A. Guerroudji, K. Amara, D. Aouam, N. Zenati, O. Djekoune, and S. Benbelkacem, "Segmentation of the breast masses in mammograms using active contour for medical practice: AR based surgery," in *Artificial Intelligence and Its Applications*, B. Lejdel, E. Clementini, and L. Alarab, Eds., Cham, Switzerland: Springer, 2022, pp. 447–457.

[14] Y. Zhang, J. Chu, L. Leng, and J. Miao, "Mask-refined R-CNN: A network for refining object details in instance segmentation," *Sensors*, vol. 20, no. 4, p. 1010, Feb. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/4/1010

[15] S. Dayarathna, K. T. Islam, S. Uribe, G. Yang, M. Hayat, and Z. Chen, "Deep learning based synthesis of MRI, CT and PET: Review and analysis," *Med. Image Anal.*, vol. 92, Feb. 2024, Art. no. 103046.

[16] R. Kumar, P. Kumbharkar, S. Vanam, and S. Sharma, "Medical images classification using deep learning: A survey," *Multimedia Tools Appl.*, vol. 83, no. 7, pp. 19683–19728, Jul. 2023.

[17] S. Rai, J. S. Bhatt, and S. K. Patra, "Deep learning in medical image analysis: Recent models and explainability," in *Explainable AI in Healthcare*. Boca Raton, FL, USA: CRC Press, 2024, pp. 23–49.

[18] S. K. Zhou, H. Greenspan, C. Davatzikos, J. S. Duncan, B. Van Ginneken, A. Madabhushi, J. L. Prince, D. Rueckert, and R. M. Summers, "A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises," *Proc. IEEE*, vol. 109, no. 5, pp. 820–838, May 2021.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," 2015, *arXiv:1505.04597*.

[20] R. Azad, E. K. Aghdam, A. Rauland, Y. Jia, A. H. Avval, A. Bozorgpour, S. Karimijafarbigloo, J. Cohen, E. Adeli, and D. Merhof, "Medical image segmentation review: The success of U-Net," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Jan. 2022.

[21] J. Chen, J. Mei, X. Li, Y. Lu, Q. Yu, Q. Wei, X. Luo, Y. Xie, E. Adeli, Y. Wang, M. P. Lungren, S. Zhang, L. Xing, L. Lu, A. Yuille, and Y. Zhou, "TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers," *Med. Image Anal.*, vol. 97, Oct. 2024, Art. no. 103280.

[22] L. Qian, C. Wen, Y. Li, Z. Hu, X. Zhou, X. Xia, and S.-H. Kim, "Multi-scale context UNet-like network with redesigned skip connections for medical image segmentation," *Comput. Methods Programs Biomed.*, vol. 243, Jan. 2024, Art. no. 107885.

[23] S. Ali, R. Khurram, K. U. Rehman, A. Yasin, Z. Shaukat, Z. Sakhawat, and G. Mujtaba, "An improved 3D U-Net-based deep learning system for brain tumor segmentation using multi-modal MRI," *Multimedia Tools Appl.*, vol. 83, no. 37, pp. 85027–85046, May 2024.

[24] Q. Shan, T. E. Doyle, R. Samavi, and M. Al-Rei, "Augmented reality based brain tumor 3D visualization," *Proc. Comput. Sci.*, vol. 113, pp. 400–407, Jan. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050917317660

[25] K. Amara, O. Kerdjidj, M. A. Guerroudji, N. Zenati, and O. Djekoune, "Augmented reality visualization and interaction for COVID-19 CT-scan NN automated segmentation: A validation study," *IEEE Sensors J.*, vol. 23, no. 11, pp. 12114–12123, Jun. 2023.

[26] K. Amara, A. Aouf, H. Kennouche, A. O. Djekoune, N. Zenati, O. Kerdjidj, and F. Ferguene, "COVIR: A virtual rendering of a novel NN architecture O-Net for COVID-19 ct-scan automatic lung lesions segmentation," *Comput. Graph.*, vol. 104, pp. 11–23, May 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0097849322000358

[27] N. Montemurro, S. Condino, M. Carbone, N. Cattari, R. D'Amato, F. Cutolo, and V. Ferrari, "Brain tumor and augmented reality: New technologies for the future," *Int. J. Environ. Res. Public Health*, vol. 19, no. 10, p. 6347, May 2022. [Online]. Available: https://www.mdpi.com/1660-4601/19/10/6347

[28] M. A. Guerroudji, K. Amara, and N. Zenati, "HoloBrain: 3D low-cost mobile augmented reality rendering of brain tumour using the GVF snake model segmentation," *Comput. Methods Biomechanics Biomed. Eng., Imag. Visualizat.*, vol. 11, no. 7, pp. 1–16, Jan. 2024, doi: 10.1080/21681163.2024.2302051.

[29] M. Satoh, T. Nakajima, T. Yamaguchi, E. Watanabe, and K. Kawai, "Evaluation of augmented-reality based navigation for brain tumor surgery," *J. Clin. Neurosci.*, vol. 94, pp. 305–314, Dec. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967586821005403

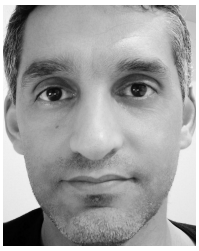
[30] A. L. Roethe, J. Rösler, M. Misch, P. Vajkoczy, and T. Picht, "Augmented reality visualization in brain lesions: A prospective randomized controlled evaluation of its potential and current limitations in navigated microneurosurgery," *Acta Neurochirurgica*, vol. 164, no. 1, pp. 3–14, Jan. 2022.

[31] Unity Technologies. (2024). *Unity 3D*. Accessed: Jan. 2024. [Online]. Available: https://unity.com/
[32] Google. (2023). *Arcore*. [Online]. Available: https://developers.google.com/ar
[33] U. Baid et al., "The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification," 2021, *arXiv:2107.02314*.
[34] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual U-Net," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, May 2018.
[35] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*.
[36] E. Kerfoot, J. Clough, I. Oksuz, J. Lee, A. P. King, and J. A. Schnabel, "Left-ventricle quantification using residual U-Net," in *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*, M. Pop, M. Sermesant, J. Zhao, S. Li, K. McLeod, A. Young, K. Rhode, and T. Mansi, Eds., Cham, Switzerland: Springer, 2019, pp. 371–380.
[37] M. Lee, A. B. S. B. Mahmood, E. S. Lee, H. E. Smith, and L. T. Car, "Smartphone and mobile app use among physicians in clinical practice: Scoping review," *JMIR mHealth uHealth*, vol. 11, Mar. 2023, Art. no. e44765. [Online]. Available: https://mhealth.jmir.org/2023/1/e44765

MERIEM KHEDIR received the M.Sc. degree in computer science from École Supérieure en Informatique, Sidi Bel Abbès, in 2024. Her main research interests include biomedical image analysis, computer vision, and artificial intelligence in healthcare.

KAHINA AMARA received the master's degree in control and robotics and the Ph.D. degree from the University of Science and Technology Houari Boumediene (USTHB), Algiers, Algeria, in 2011 and 2018, respectively. She is currently a Permanent Researcher with CDTA, IRVA Team. Her main research interests include augmented and virtual reality, 3D interaction, affective computing, computer vision, emotion recognition, and healthcare.

NASSIMA DIF received the B.Sc., M.Sc., and Ph.D. degrees in computer science from the University of Djillali Liabes, in 2015, 2017, and 2020, respectively. She is currently a Teacher and a Researcher with École Supérieure en Informatique, Sidi Bel Abbès. Her research interests include machine learning, deep learning, feature selection, and biomedical image analysis.



OUSSAMA KERDJIDJ received the Ph.D. degree from the University of Laghouat, Algeria, in 2019. His leading research interests include hardware and software implementation and artificial intelligence applied to healthcare applications.

SHADI ATALLA (Senior Member, IEEE) is currently an Associate Professor with the College of IT, University of Dubai. He has more than ten years of teaching and research experience at the university level. He was involved in several European research projects, where he played the role of a Research and Development Engineer. He is also involved with designing and implementing ICT industrial-related projects. He was a Research Fellow with Pervasive Technologies Research Group, Istituto Superiore Mario Boella, Turin, Italy, from 2012 to 2015. He has published a number of papers in international scientific journals and at international conferences. His main research interests include high-speed switching architectures, optical packet switching, architectures and trade-offs, wireless networks, and the Internet of Things.

NAEEM RAMZAN (Senior Member, IEEE) received the M.Sc. degree in telecommunication from the University of Brest, France, in 2004, and the Ph.D. degree in electronics engineering from the Queen Mary University of London, London, U.K., in 2008. He is currently a Full Professor of computing engineering and the Chair of the Affective and Human Computing for Smart Environment Research Centre and the Co-Lead of AVCN Institute, University of the West of Scotland (UWS). Before joining UWS, he was a Senior Research Fellow and a Lecturer with the Queen Mary University of London, from 2008 to 2012. He has published more than 200 papers in journals, conferences, and book chapters, including standard contributions. He is focused on leading high quality interdisciplinary research and teaching in the areas of video processing, analysis and communications, multimedia search and retrieval, video quality evaluation, brain-inspired multi-modal cognitive technology, multimodal human–computer interfaces, DNA computing, fall detection, big data analytics, affective computing, the IoT/smart environments, natural multimodal human–computer interaction, and eHealth/connected health. He is a fellow of the Royal Society of Edinburgh (FRSE), a Senior Fellow of the Higher Education Academy (SFHEA), the Co-Chair of the MPEG HEVC Verification (AHG5) Group, and a Voting Member of the British Standards Institution (BSI). In addition, he holds key roles with the Video Quality Expert Group (VQEG), such as the Co-Chair of the Psycho-Physiological Quality Assessment (PsyPhyQA). He has been awarded the Scottish Knowledge Exchange Champion Award, in 2023 and 2020. He was awarded the Staff Appreciation and Recognition Scheme (STARS) Award for the "Outstanding Research and Knowledge Exchange" (University of the West of Scotland), in 2014 and 2016, and was awarded the Contribution Reward Scheme for the "Outstanding Research and Teaching Activities" (Queen Mary University), in 2011 and 2009.