Original Research

ACU-Net: Attention-based convolutional U-Net model for segmenting brain tumors in fMRI images

Md Alamin Talukder[1] (iD), Md Abu Layek[2] (iD), Md Aslam Hossain[2],
Md Aminul Islam[2,3], Mohammad Nur-e-Alam[4] and Mohsin Kazi[5] (iD)

Abstract

Objective: Accurate segmentation of brain tumors in medical imaging is essential for diagnosis and treatment planning. Current techniques often struggle with capturing complex tumor features and are computationally demanding, limiting their clinical application. This study introduces the attention-based convolutional U-Net (ACU-Net) model, designed to improve segmentation accuracy and efficiency in fMRI images by incorporating attention mechanisms that selectively highlight critical features while preserving spatial context.

Methods: The ACU-Net model combines convolutional neural networks (CNNs) with attention mechanisms to enhance feature extraction and spatial coherence. We evaluated ACU-Net on the BraTS 2018 and BraTS 2020 fMRI datasets using rigorous data splitting for training, validation, and testing. Performance metrics, particularly Dice scores, were used to assess segmentation accuracy across different tumor regions, including whole tumor (WT), tumor core (TC), and enhancing tumor (ET) classes.

Results: ACU-Net demonstrated high segmentation accuracy, achieving Dice scores of 99.23%, 99.27%, and 96.99% for WT, TC, and ET, respectively, on the BraTS 2018 dataset, and 98.72%, 98.40%, and 97.66% for WT, TC, and ET on the BraTS 2020 dataset. These results indicate that ACU-Net effectively captures tumor boundaries and subregions with precision, surpassing traditional segmentation approaches.

Conclusion: The ACU-Net model shows significant potential to enhance clinical diagnosis and treatment planning by providing precise and efficient brain tumor segmentation in fMRI images. The integration of attention mechanisms within a CNN architecture proves beneficial for identifying complex tumor structures, suggesting that ACU-Net can be a valuable tool in medical imaging applications.

[1]Department of Computer Science and Engineering, International University of Business Agriculture and Technology, Dhaka, Bangladesh
[2]Department of Computer Science and Engineering, Jagannath University, Dhaka, Bangladesh
[3]Department of Computer Science, Tennessee Technological University, Cookeville, USA
[4]Department of Pharmacognosy, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia
[5]Department of Pharmaceutics, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia

Corresponding authors:
Md Alamin Talukder, Department of Computer Science and Engineering, International University of Business Agriculture and Technology, Dhaka, Bangladesh.
Email: alamin.cse@iubat.edu

Md Abu Layek, Department of Computer Science and Engineering, Jagannath University, Dhaka, Bangladesh.
Email: layek@cse.jnu.ac.bd

Introduction

Segmentation is crucial to medical image analysis because it produces useful data for the identification, monitoring, and treatment of brain tumors.[1,2] A crucial task in medical image analysis is brain tumor segmentation, especially in the field of neuroimaging.[3] Delineating and identifying regions of interest linked to brain tumors in medical pictures, like magnetic resonance imagings (MRIs), is the aim. For the purpose of diagnosing, treating, and keeping track of patients with brain tumors, precise segmentation is crucial.[4] Before being identified, brain tumors can grow to be quite large and take on a variety of shapes and sizes.[5] Unusual cells that proliferate inside the brain are called tumors. If not found early on, these aberrant cells have the potential to be fatal.[6] The two primary categories of brain tumors are malignant, a severe form of cancer, and benign, which is not cancerous.[7] Adult brain tumors of the glioma type are the most common kind and are classified into two grades: low-grade gliomas (LGG) are slowly developing tumors, whereas high-grade gliomas (HGG) proliferate.[8] Glioma cancer is classified into three subdivisions: tumor core (TC), enhancing core (EC), and whole tumor (WT). It is one of the most prevalent malignant brain tumors in current times.[9] Every year, over 90,000 patients are diagnosed with primary brain tumors. Malignant brain tumors kill roughly 17,200 people each year. An analysis revealed that 25 persons out of 100,000 have malignancies, with 33% of them being serious.[10] Medical imaging and neurology heavily focus on the segmentation of the brain, especially neurological disorders like brain tumors.[11] The assessment of brain segments is a crucial component in the range of neurological care, covering diagnostic and therapeutic approaches, surgical procedures, disease surveillance, as well as research and development of personalized therapies.[12]

Recent advances in deep learning have introduced attention mechanisms into segmentation models. Attention mechanisms allow models to focus selectively on important features within the image, enhancing segmentation performance by prioritizing relevant spatial information.[13–15] In comparison to manual techniques and conventional tools, the deep U-Net architecture has transformed fMRI segmentation by utilizing deep learning capabilities to increase performance and efficiency.[16] Deep U-Net automates the segmentation process by learning from vast datasets and extracting complex patterns in fMRI pictures, in contrast to manual segmentation, which is laborious and subjective.[17] This deep learning approach allows for more precise delineation of brain tumor boundaries, enabling clinicians to make more informed decisions about diagnosis and treatment planning.[18,19] Existing methods in brain tumor segmentation face several challenges, as highlighted in recent literature. For instance, Saeed et al.[20] proposed RMU-Net, which, despite achieving high performance on the BraTS 2018 dataset, may encounter difficulties in effectively segmenting heterogeneous tumor types across different subregions. Furthermore, Ali et al.[21] introduced an ensemble of U-Net and 3D convolutional neural network (CNN) models, yet computational complexity remains a concern, potentially limiting its clinical applicability. Tataei et al.[22] utilized CNN engineering and ResNet-50 architecture, but the reliance on basic convolutional layers may restrict the model's power to capture intricate tumor variants. Additionally, evaluation metrics such as the Dice similarity coefficient, emphasized in several studies, may not fully encompass segmentation quality, thus warranting a more comprehensive assessment approach.[23,24] These challenges underscore the necessity for novel methodologies to handle the boundaries of existing approaches and advance the field of brain tumor segmentation.

The proposed attention-based convolutional U-Net (ACU-Net) model overcomes the limitations of conventional brain tumor segmentation approaches by incorporating attention mechanisms that selectively emphasize critical features while maintaining spatial context. This novel architecture enables enhanced segmentation accuracy across diverse tumor types and subregions in fMRI data from the BraTS 2018 and BraTS 2020 datasets. Rigorous evaluation of ACU-Net demonstrates substantial improvements in segmentation performance, with the model consistently outperforming state-of-the-art methods. Notably, it achieves superior results in key metrics, including the Dice similarity coefficient, sensitivity, and specificity, underscoring its effectiveness in accurately delineating brain tumors.

The key contributions of this article are outlined as follows:

- **Development of ACU-Net Architecture:** We introduce an innovative attention-based convolutional U-Net (ACU-Net) specifically designed for accurate brain tumor segmentation in fMRI data. This architecture not only improves segmentation performance but also enhances the efficiency of delineating tumor boundaries.
- **Integration of Attention Mechanisms:** ACU-Net incorporates advanced attention mechanisms that selectively focus on relevant features within fMRI images. This selective highlighting preserves spatial context and significantly improves the model's ability to segment tumor subregions accurately.
- **Achievement of State-of-the-Art Performance:** Through comprehensive evaluations on the BraTS 2018 and BraTS 2020 datasets, we demonstrate that ACU-Net achieves superior segmentation results, surpassing existing methodologies in key metrics such as Dice coefficient, sensitivity, and specificity. These findings establish ACU-Net as a leading model in the field of brain tumor segmentation.
- **Impact on Clinical Applications:** By providing a robust tool for accurate tumor segmentation, our

research has significant implications for clinical practice. Enhanced segmentation accuracy can lead to improved treatment planning and patient outcomes in neuro-oncology.

Research Questions:

1. *How does the ACU-Net improve brain tumor segmentation performance compared to traditional U-Net models?*
2. *Can the ACU-Net more effectively segment different tumor subregions (whole tumor, tumor core, enhancing tumor) in fMRI images than current state-of-the-art methods?*

Hypothesis: The proposed ACU-Net will outperform traditional U-Net models and existing state-of-the-art methods in brain tumor segmentation tasks. By incorporating attention mechanisms with Convolutional-UNet architecture, the ACU-Net is hypothesized to achieve higher segmentation performance, particularly in delineating heterogeneous tumor subregions. The selective highlighting of relevant features while preserving spatial context is expected to result in improved Dice similarity coefficients, sensitivity, and specificity metrics.

The article follows this format in the sections that follow: section "Literature review" offers a thorough analysis of the pertinent research. The recommended methodology used in this study is described in depth in section "Methdolgy." The performance analysis and discussion of the findings are presented in sections "Result analysis" and "Discussion." Ultimately, the concluded findings from this study project are summarized in section "Conclusion."

Literature review

The segmentation of brain tumors from multimodal MRI images has seen significant progress with the advent of deep learning models. The Brain Tumor Segmentation (BraTS) challenge datasets from 2018 and 2020 have been pivotal in evaluating the performance of segmentation methods. This section reviews the major techniques employed on the BraTS 2018 and BraTS 2020 datasets, providing an in-depth analysis of their strengths, limitations, and how these methods compare to the proposed ACU-Net.

Study on BraTS 2018

The BraTS 2018 dataset has been a key benchmark for evaluating brain tumor segmentation models, with numerous contributions showcasing diverse approaches to improve segmentation accuracy.

Saeed et al.[20] proposed RMU-Net, a hybrid model combining MobileNetV2 and U-Net architectures. This end-to-end segmentation framework demonstrated good performance on the BraTS 2018 dataset, achieving Dice scores ranging from 90.80% to 79.36% for different tumor areas. However, the model's reliance on MobileNetV2, a lightweight architecture, might limit its ability to capture more complex tumor features compared to more traditional CNN-based models. Additionally, the integration of MobileNetV2 introduces additional challenges in fine-tuning and achieving consistent accuracy across all tumor subregions.

Ali et al.[21] presented an ensemble model combining U-Net and 3D convolutional neural networks (CNNs) for tumor segmentation. On the BraTS 2018 dataset, their model produced competitive Dice scores: 0.750 for ET, 0.906 for the WT, and 0.846 for the TC. While the ensemble approach enhances model robustness, it also increases computational complexity and can be harder to implement in real-time applications due to the increased processing requirements.

Tataei et al.[22] utilized a CNN and ResNet-50 architecture to segment gliomas in the BraTS 2018 dataset. Their approach achieved promising Dice scores across different tumor regions, but the CNN-based model's ability to generalize may be hindered by its dependance on handcrafted features. This limitation is addressed in more recent models that incorporate deep learning techniques capable of learning features automatically.

Gull et al.[25] introduced a CNN-based model that achieved an average performance rating of 96.50% for brain tumor segmentation on the BraTS 2018 dataset. However, while the model performed well, it lacked the robustness to handle variations in tumor appearance, which remains a challenge in medical imaging tasks.

Ullah et al.[23] applied 3D U-Net for tumor segmentation, achieving Dice scores of 0.91 for the OT, 0.86 for the TC, and 0.70 for the AT. Although the model performed well overall, its segmentation of augmenting tumors still faces challenges due to the irregularities in tumor shapes and low-contrast areas that require further enhancement.

Cao et al.[24] introduced MBANet, a 3D convolutional neural network with multi-branch attention for brain tumor segmentation. Their method achieved competitive results on the BraTS 2018 dataset with Dice scores of 78.21%, 89.79%, and 83.04% for ET, WT, and TC, respectively. However, attention mechanisms, while helpful, often suffer from computational overhead and may not fully capture the complex structures in the tumor images.

Zia et al.[26] presented an attentional residual dropout U-Net (ARDUNet), which achieved excellent results, including a Dice score of 0.92 for augmenting tumors. Despite its strong performance, the method still faces challenges with segmentation consistency and may require more sophisticated techniques to handle a wider variety of tumor types and MR imaging conditions.

Sun et al.[27] created a special model based on a three-dimensional fully convolutional network to segment brain tumors. By separating the complete tumor and enhancing the tumor areas, the suggested approach achieved dice similarity coefficient values for the dataset of 0.90, 0.79, and 0.77. A novel residual network-based U-Net architecture for brain tumor segmentation was proposed by Pedada et al.[9] Testing their proposed U-Net model on benchmark datasets, such as the BraTS Challenge 2018, demonstrated segmentation accuracies of 92.20. Additionally, EC, total tumor (WT), and TC classifications were developed based on the sub-regions of the tumor. Despite its good performance, the method still needs to improve to get a segmentation consistency performance.

Study on BraTS 2020

The BraTS 2020 dataset builds upon the success of BraTS 2018 by offering a more complex and diverse set of MRI scans, providing additional challenges for tumor segmentation models. Many of the most successful methods from BraTS 2018 have been adapted and enhanced for BraTS 2020, with new approaches addressing the growing need for more robust and accurate models.

Isensee et al.[28] introduced nnU-Net, which achieved impressive results on BraTS 2020, securing first place in the challenge with Dice scores of 88.95% for WT, 85.06% for TC, and 82.03% for ET. This model is notable for its baseline configuration, which achieves competitive results without any task-specific adjustments, such as post-processing or region-based training. However, nnU-Net's lack of specialized optimizations for different tumor types and imaging conditions may limit its applicability to more diverse datasets.

Kataria et al.[29] proposed the HybriCSF model, a hybrid approach combining CNNs with SVMs and fuzzy C-means clustering. Their model achieved competitive results on BraTS 2020 with Dice scores of 87% for WT, 81% for TC, and 63% for ET. While this hybrid approach offers a novel perspective by combining machine learning techniques, it struggled with segmenting enhancing tumors, which remain challenging due to their irregular shapes and low contrast in MRI scans.

Magadza et al.[30] extended the U-Net architecture with bottleneck units and a shuffle attention mechanism, improving the model's performance on BraTS 2020 with Dice scores of 91.2% for WT, 84.8% for TC, and 79.2% for ET. Despite these enhancements, the method's performance on enhancing tumors is still less than ideal, suggesting that attention mechanisms need further refinement to address this issue effectively.

Susanto et al.[31] introduced a spatial transformation-based data augmentation pipeline that significantly improved model robustness. The augmented approach achieved Dice scores of 90.91% for WT, 86.89% for TC,

and 87.10% for ET. However, while their augmentation strategy helped improve segmentation performance, it remains to be seen whether it can generalize across more diverse datasets or handle more severe image artifacts.

Zhang et al.[32] proposed a multi-encoder model for 3D MRI segmentation, incorporating a novel Categorical Dice loss function to address voxel imbalance. Their method demonstrated notable results, with Dice scores of 70.24% for WT, 88.26% for TC, and 73.86% for ET. Despite the promising results, the model's performance on whole tumor segmentation is still lower than that of other methods, indicating room for further improvement in segmentation accuracy.

These studies highlight the diversity of approaches for segmenting brain tumors, particularly the challenges in accurately segmenting enhancing tumors, which remain an area for further research. Hybrid models and attention mechanisms have shown promise, but combining these techniques with novel architectures like ACU-Net might offer further improvements in both segmentation accuracy and computational efficiency.

Rajinikanth et al.[33] developed a Computer-Aided Disease Diagnosis (CADD) system for classifying brain tumors in 2D MRI slices as Glioblastoma or Glioma. The system used CNNs for segmentation with VGG-UNet, feature extraction, and selection through the Firefly algorithm, achieving over 98% accuracy, especially with the SVM-Cubic classifier, demonstrating enhanced diagnostic accuracy. Khan et al.[34] introduced an automated brain tumor detection and classification system using saliency maps and deep learning for feature optimization. The system enhanced images through fusion techniques, fine-tuned a pre-trained EfficientNetB0 model, and used a dragonfly algorithm for feature optimization. The system achieved accuracies of 95.14%, 94.89%, and 95.94% on three public datasets, surpassing other neural networks.

Kurdi et al.[35] employed a Harris Hawks optimized convolutional network (HHOCNN) to improve brain tumor detection in MRI. The process involved noise removal, tumor region identification, and CNN-based feature classification. With Harris Hawks optimization, the system achieved 98% tumor recognition accuracy on the Kaggle dataset, reducing misclassification errors. Badjie et al.[36] developed a DL algorithm to enhance brain tumor detection in MR images. Using a transfer learning model based on AlexNet's CNN, the system automated the diagnostic process, improving accuracy, efficiency, and robustness in tumor classification across various stages and sizes.

Methodology

In our experiment, we developed an ACU-Net architecture for brain tumor segmentation employing fMRI images. The methodology includes brain fMRI image collection, image preprocessing, building an ACU-Net architecture and
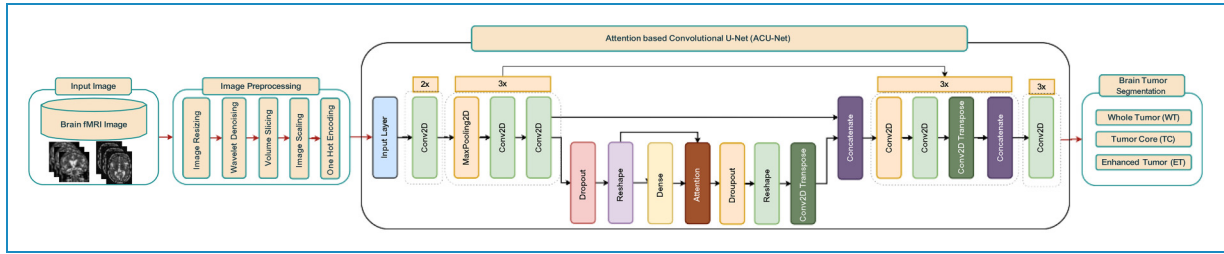
Figure 1. The proposed ACU-Net architecture for brain tumor segmentation.

training the dataset and finally predicting the enactment of the segmentation of the brain tumor using our model on fMRI images. The architecture of our ACU-Net is depicted in Figure 1.

The research focused on enhancing brain tumor segmentation in fMRI images using the ACU-Net model, specifically leveraging the attention mechanism with the convolutional model integrated with the U-Net model. This study was conducted in 2024 in Bangladesh (IUBAT, JnU) and Saudi Arabia (KSU).

Data collection

The BraTS 2018 dataset[37] is a widely used benchmark for glioma segmentation, including both high-grade and low-grade gliomas. It consists of multi-institutional MRI scans from 285 patients, divided into four modalities: T1-weighted, T1-weighted with contrast (T1CE), T2-weighted, and FLAIR. These modalities capture different tumor regions, including the enhancing tumor, core, and surrounding edema. The dataset has been preprocessed to a uniform resolution with skull stripping to focus only on the brain, and the pixel-wise annotations enable precise segmentation of tumor sub-regions, making it ideal for deep learning model development.

The BraTS 2020 dataset[38] includes 368 MRI scans, with 265 for training, 65 for validation, and 47 for testing. It also offers the same four MRI modalities and focuses on glioma sub-region segmentation. Improvements include more accurate annotations and a wider range of clinical cases, making it suitable for both segmentation and survival prediction tasks. This dataset remains crucial in developing and evaluating deep learning models for clinical use, with rigorous comparisons made possible through standardized challenges.

In our study, we used the BraTS 2018 dataset, which consists of 285 patients: 210 with high-grade glioma (HGG) and 75 with low-grade glioma (LGG). We combined both HGG and LGG cases into a unified dataset for training, validation, and testing, without explicitly differentiating between the two tumor types. The dataset was split into 80% for training (228 patients), 10% for validation (28 patients), and 10% for testing (29 patients). Our primary goal was to achieve accurate brain tumor

segmentation, and this unified approach allowed our ACU-Net model to generalize well across both tumor grades. The model performed consistently in segmenting key regions such as WT, TC, and ET, demonstrating high performance. The merging of HGG and LGG into a single dataset did not adversely affect segmentation performance, as our model maintained robust performance across all regions of interest.

Figure 2 shows some sample images from the BraTS 2018 and BraTS 2020 datasets.

Data preprocessing

In order to prepare the fMRI images for effective brain tumor segmentation, a rigorous preprocessing pipeline is employed, encompassing several essential steps:

- *Image Resize*: The images are resized (128×128) to a standardized resolution to ensure consistency across the dataset. This step facilitates uniform processing and analysis of the fMRI data.
- *Wavelet denoising*: Wavelet denoising is a sophisticated technique used to reduce noise in images while preserving essential structural details. This method employs wavelet transforms to decompose an image into various frequency components. By analyzing these components, we can selectively attenuate noise while retaining critical features. This approach enhances image quality, allowing segmentation models to focus on relevant anatomical structures without being misled by artifacts. In the context of fMRI brain tumor segmentation, wavelet denoising improves performance and robustness by providing cleaner images that enable better delineation of tumor boundaries and vital brain regions.
- *Volume Slicing*: Volume slicing involves selecting specific cross-sections from the 3D volume to facilitate the analysis and processing of the data. Instead of manually selecting a fixed range of slices, we utilized an adaptive slicing approach. Instead of manually selecting a fixed range of slices, we utilized an adaptive slicing approach. The Adaptive Slicing Function dynamically selects a specific number of slices (num_slices) from the entire volume using linspace from numpy, ensuring a uniform distribution of slices regardless of the
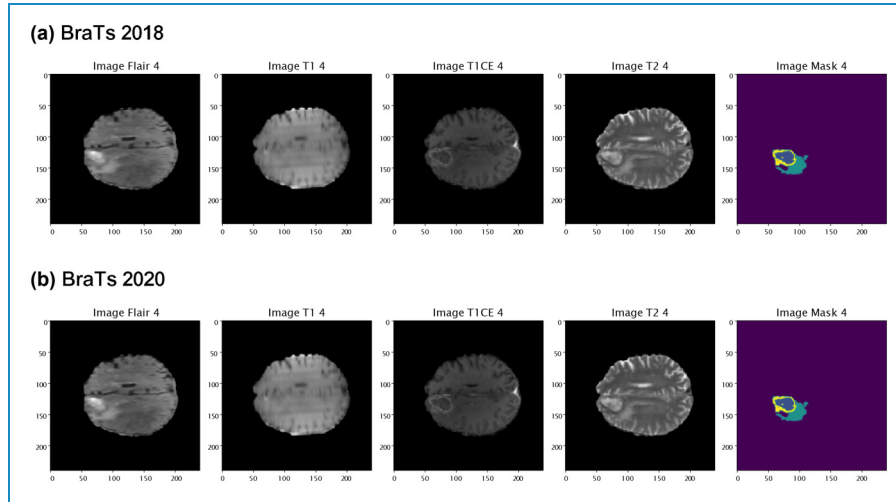
Figure 2. Sample images of BraTS 2018 and BraTS 2020 dataset.

volume's depth. This method improves the representativeness of the data, allowing the model to capture critical patterns from all regions of the brain.

By adopting adaptive slicing, we avoided the need to experiment with different slice resolutions and orientations, which can be computationally expensive and may vary significantly between datasets. Additionally, this technique minimizes the risk of excluding important anatomical structures that might be present in slices outside a manually defined range. Overall, adaptive slicing enhances the segmentation performance by providing a comprehensive and consistent representation of the 3D volume without the need for extensive preslice experimentation.

- *Image Scaling*: The intensity values of the fMRI images are scaled to a common range, typically between 0 and 1. This normalization step enhances the comparability of image features and improves the convergence of segmentation algorithms during training.
- *One-Hot Encoding*: Categorical labels representing different tissue types or tumor classes are converted into binary vectors using one-hot encoding. This encoding scheme ensures that each class is represented as a distinct binary variable, facilitating the training of machine learning models for segmentation tasks.
- *Data Augmentation*: To mitigate the risk of overfitting and enhance the generalization ability of the ACU-Net model, several data augmentation techniques were employed. These methods artificially expand the training dataset, allowing the model to learn more robust features and become less reliant on specific image variations. **Rotation** was applied within a range of $-30°$ to $+30°$ to simulate different orientations of the brain and tumor structures, enabling the model to recognize tumors in any rotational angle. **Flipping** (horizontal and vertical) was used to account for the brain's

symmetry, providing diverse examples while maintaining anatomical integrity. **Noise injection** with Gaussian noise helped the model become more resilient to real-world noise found in medical images, preventing overfitting to clean, noise-free images. **Zooming** was used to simulate varying tumor sizes by randomly scaling images, ensuring that the model could detect tumors of different dimensions. **Shearing** was applied to skew images along one axis, simulating distortions due to patient movement or imaging errors. Finally, **elastic deformation** mimicked anatomical variability, improving the model's robustness to slight shape differences in brain structures. These augmentation techniques address the inherent variability in medical imaging, ensuring the ACU-Net model performs well on diverse, real-world brain tumor data while minimizing overfitting.

While the attention mechanism in ACU-Net is designed to focus on pertinent areas of the input image and can adapt to some extent, relying solely on this mechanism for managing inconsistencies would be insufficient. Preprocessing techniques like denoising and standardization are crucial in preparing the data and improving segmentation performance. By combining these approaches, we ensure that our model can effectively handle variations in the dataset and deliver reliable segmentation results. While the attention mechanism adds value in emphasizing relevant features, preprocessing techniques play a vital role in ensuring data consistency and quality, ultimately enhancing the performance of the ACU-Net model in brain tumor segmentation tasks.

Data splitting

In this study, the dataset was divided into training, validation, and testing sets using an 80-10-10 split to ensure a

Table 1. Dataset splits for BraTS 2018 and BraTS 2020.

Dataset	Train	Validation	Test	Total
Percentage	80%	10%	10%	100%
BraTS 2018	228	28	29	285
BraTS 2020	294	37	27	368

balanced model evaluation. Both the BraTS 2018 and BraTS 2020 datasets were structured accordingly, with 80% of the data used for training the model, 10% reserved for validation, and the remaining 10% allocated for testing. This approach allowed for a comprehensive evaluation of the model's performance while maintaining a sufficient amount of data for both the learning and testing phases. Table 1 outlines the specific data splits for each dataset, showing 228 training samples, 28 validation samples, and 29 test samples for BraTS 2018. Similarly, for BraTS 2020, 294 samples were used for training, 37 for validation, and 27 for testing. This consistent splitting approach ensured reliable performance comparisons across datasets.

The training setup included a specified epoch count, batch size, and early stopping criteria. The model was trained for up to 25 epochs, with early stopping implemented based on the validation loss to prevent overfitting; training halts if no improvement is observed after a set patience period. We utilized a `ReduceLROnPlateau` callback in our model training, which dynamically adjusts the learning rate based on validation loss improvements. Specifically, if the validation loss does not improve after two epochs, the learning rate is reduced by a factor of 0.2, down to a minimum learning rate of 1×10^{-6}. This approach helps stabilize the model training process, potentially avoiding overfitting by adjusting the learning rate adaptively. The batch size was set at 16, balancing computational efficiency and gradient stability during training.

Attention-based convolutional U-Net architecture

In our brain tumor segmentation approach, we leverage an Attention-based Convolutional U-Net Architecture (ACU-Net) to accurately detect and segment tumors within fMRI images. This architecture amalgamates several key components, each playing a crucial role in enhancing the segmentation performance.

- *Attention Mechanism:* The ACU-Net's attention function enables the model to suppress distracting or noisy information while concentrating on pertinent areas of the input image. The Attention mechanism improves the model's ability to collect fine details in the fMRI scans, which improves the performance of tumor

segmentation by adaptively weighting features according to their value.
- *Convolutional Layers:* The core of the ACU-Net design is its convolutional layers, which extract hierarchical characteristics from the input photos. In order to capture the local patterns and spatial correlations necessary for recognizing tumor locations inside the brain, these layers convolve over the input data using learnable filters.
- *U-Net Architecture:* Our segmentation strategy is based on the U-Net architecture, which is distinguished by its symmetric encoder-decoder structure. The model is able to precisely delineate tumor boundaries because of the U-Net architecture, which allows it to capture both local and global background information. Furthermore, by allowing the merging of low-level and high-level features, the skip connections in the U-Net design improve the segmentation performance of the model.

We chose a symmetric U-Net design due to its proven effectiveness in biomedical image segmentation. The symmetry enables balanced capture of both local and global features, which is essential for accurately delineating tumor boundaries in fMRI images. Although experimenting with asymmetric U-Net designs or deeper networks could capture more hierarchical features, they often increase complexity and the risk of overfitting. Our experiments with both MHAU-Net and ACU-Net showed that the symmetric design provided robust segmentation performance while maintaining model stability and interoperability. The symmetric U-Net design was chosen for its reliable performance in segmentation tasks, delivering high performance without unnecessary complexity.

By combining these components, the ACU-Net architecture offers several benefits for brain tumor segmentation in fMRI images. Firstly, its attention mechanism enables the model to concentrate on relevant tumor regions, improving both sensitivity and specificity in segmentation results. Secondly, the convolutional layers capture discriminative features from the input images, promoting the model to distinguish between tumor and non-tumor regions with high performance. Lastly, the U-Net architecture facilitates the integration of contextual information, enabling the model to delineate tumor boundaries accurately even in challenging cases.

The attention mechanism we introduced is specifically designed to enhance feature representation by capturing long-range dependencies in the input data. The integration of our attention mechanism and convolutional layers within the U-Net architecture has proven to be instrumental in enhancing segmentation performance. This enables the model to better focus on relevant spatial features while suppressing irrelevant information, which is crucial in

improving segmentation accuracy, particularly in complex tasks such as brain tumor delineation. Unlike standard attention modules, our approach prioritizes preserving the spatial integrity of feature maps to ensure the adaptive enhancement of segmentation performance. This design allows the model to emphasize the most informative regions in the image, leading to more precise identification of critical structures. To further evaluate its effectiveness, we conducted comparative experiments using a Multi-Head Attention U-Net (MHAU-Net) on the BraTS 2018 and BraTS 2020 datasets. These experiments revealed key strengths of our ACU-Net mechanism over multi-head attention, affirming its suitability for the task.

The model follows a U-Net-like architecture with an encoder-decoder structure. The encoder consists of four blocks, where each block includes two convolutional layers of size 3x×3, followed by a max-pooling operation with a pool size of 2x×2. The number of filters starts at 32 and increases progressively to 64, 128, and 256 across the encoder blocks. The bottleneck layer, where the spatial dimensions are reduced to the smallest size, has 512 filters, with 3x×3 kernel sizes and includes a dropout layer for regularization with a rate of 0.2 to avoid overfitting. For the attention mechanism, we reshape the output of the bottleneck layer (from 4D to 2D) and create a query using a Dense layer with 512 units. The attention mechanism is applied using the Keras Attention layer, which processes the query and the reshaped feature map to focus on important spatial information. The decoder path is built with Conv2DTranspose layers for upsampling, each followed by concatenation with the corresponding encoder output. The filter sizes in the decoder layers are 256, 128, 64, and 32. Regarding hyperparameters, we use the Adam optimizer with default learning rates and a dropout rate of 0.2. These choices are designed to balance model complexity and prevent overfitting while enabling the model to learn effectively.

The essential components of our ACU-Net with its internal architecture are presented in Figure 3.

Our proposed ACU-Net model addresses several of these key challenges such as: **Overfitting in DL Models:**

To mitigate overfitting, we applied several regularization techniques and data augmentation strategies. These measures help ACU-Net generalize better, reducing the risk of overfitting even with limited dataset sizes. Additionally, we designed our attention mechanisms to focus on relevant regions, which further enhances model robustness and reduces the tendency to overfit. **Integration of Multi-Modal Data:** While the current implementation of ACU-Net is focused on single-modality fMRI data, we recognize the potential of multi-modal integration. We have added a discussion in the future directions section, highlighting plans to incorporate multiple imaging modalities. Leveraging complementary information from other imaging types (e.g. CT or PET) could further enhance segmentation accuracy and offer more comprehensive insights.

Brain tumor segmentation

In the process of brain tumor segmentation, we employ ACU-Net techniques to partition fMRI images into three distinct segments, namely, the whole tumor (WT), core tumor (TC), and enhanced tumor (ET). Each segment serves a specific purpose in characterizing different aspects of the tumor's morphology and pathology, thereby facilitating comprehensive analysis and treatment planning.

- *Whole Tumor (WT):* Encompassing the entire tumor, including surrounding edema, the WT segment provides a comprehensive view of tumor burden and progression.
- *Tumor Core (TC):* Representing the central, actively proliferating region, the TC segment aids in evaluating tumor characteristics and guiding treatment strategies.
- *Enhanced Tumor (ET):* Highlighting regions of contrast enhancement, the ET segment helps differentiate between viable tumor tissue and treatment-related changes, facilitating targeted therapy.

This segmentation approach with ACU-Net offers detailed insights into tumor morphology and pathology, enabling precise treatment planning and monitoring.
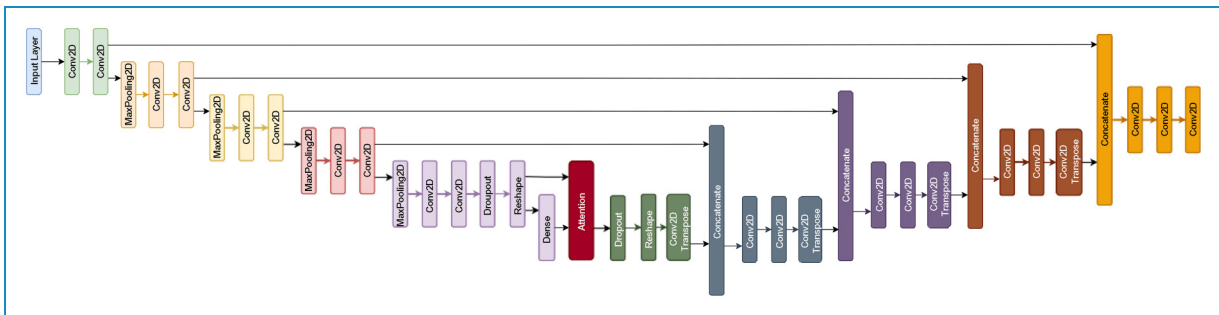

Figure 3. Internal architecture of ACU-Net model.

Result analysis

In this study, we employed the ACU-Net model for brain tumor segmentation and conducted a comprehensive evaluation against various state-of-the-art models using the BraTS 2018 and BraTS 2020 datasets. Our model was assessed on key performance metrics such as the Dice coefficient, sensitivity, and specificity. The results showed that ACU-Net significantly outperformed existing methods, particularly in segmenting challenging tumor regions, including the WT, TC, and ET.

Environment setup

The experiments were conducted on a development environment powered by NVIDIA Tesla P100 GPUs with 30 GB of dedicated GPU memory. Additionally, the system was equipped with 30 GB of RAM and 70 GB of disk space to support extensive model training. TensorFlow and Keras were the primary deep learning frameworks used, and essential libraries like Matplotlib, Pandas, and NumPy facilitated data processing, visualization, and evaluation of model performance throughout the study.

The performance metrics for brain tumor segmentation are as follows:

1. Dice (Dice Similarity Coefficient):

$$\text{Dice} = \frac{2 \times \|P \cap G\|}{\|P + G\|} \tag{1}$$

The Dice Similarity Coefficient measures the overlap between the predicted (P) and ground truth (G) segmentation masks. It quantifies the similarity between the two masks by computing twice the volume of their intersection divided by the sum of their volumes.

2. Jaccard Index (IoU):

$$\text{Jaccard} = \frac{\|P \cap G\|}{\|P \cup G\|} \tag{2}$$

The Jaccard Index, also known as the Intersection over Union (IoU), evaluates the agreement between the predicted (P) and ground truth (G) segmentations. It measures the ratio of the volume of their intersection to the volume of their union.

3. Sensitivity (True Positive Rate):

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{3}$$

Sensitivity assesses the ability of the segmentation algorithm to precisely determine tumor regions. It calculates the ratio of true positive predictions (TP) among all actual positive cases, where FN represents false negative predictions.

4. Specificity (True Negative Rate):

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{4}$$

Specificity evaluates the ability of the segmentation algorithm to accurately identify non-tumor regions. It estimates the ratio of true negative predictions (TN) among all actual negative cases, where FP represents false positive predictions.

5. The paired t-test evaluates whether the mean difference between paired observations (e.g. performance metrics of two models) is significantly different from zero. The paired t-test statistic is calculated using the following formula:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \tag{5}$$

Where:
- $\bar{d}$ is the **mean of the differences** between paired observations (i.e. the difference between the performance metrics of ACU-Net and MHAU-Net).
- s_d is the **standard deviation of the differences**.
- n is the **number of paired observations** (in this case, the number of metrics evaluated).
- t is the calculated t-statistic, which is compared to a critical value from the t-distribution table to determine statistical significance.

Importantly, we did not employ any specific threshold values in our calculations, as the Dice coefficient inherently does not require thresholds for its computation. Consequently, we did not perform cross-validation to optimize thresholds for enhancing tumor regions, since our approach focused on directly measuring the similarity between the predicted and actual segmentations without the need for thresholding. This streamlined our evaluation process and allowed us to obtain robust results based on the inherent characteristics of the segmentation outputs.

Performance analysis

We conducted a comprehensive performance evaluation of the ACU-Net model on the BraTS 2018 and BraTS 2020 fMRI datasets, focusing on its effectiveness in segmenting various tumor classes. The key metrics used in this evaluation included Dice coefficient, Jaccard index, Sensitivity, Specificity, and Intersection over Union (IoU). These metrics provide insight into the performance and robustness of the model in delineating different tumor regions: ET, WT, and TC.

As shown in Table 2, ACU-Net achieved outstanding results across all tumor classes. For the BraTS 2018 dataset, it attained a Dice score of 99.23 for ET, 99.27 for WT, and 96.99 for TC, outperforming MHAU-Net in each case. The Jaccard and IoU scores further emphasize

Table 2. Performance analysis of our ACU-net mdel on BraTS 2018 and BraTS 2020 fMRI dataset.

Exp. No.	Dataset	Model	Tumor segment class	Dice	Jaccard	Sensitivity	Specificity	IoU
1	BraTS 2018	ACU-Net	Enhancing Tumor (ET)	99.23	98.49	99.23	99.74	98.49
			Whole Tumor (WT)	99.27	98.57	99.27	99.76	98.57
			Tumor Core (TC)	96.99	94.27	96.99	99	94.27
		MHAU-Net	Enhancing Tumor (ET)	98.92	97.88	98.92	99.64	97.88
			Whole Tumor (WT)	99.03	98.1	99.03	99.68	98.1
			Tumor Core (TC)	96.53	93.45	96.53	98.84	93.45
2	BraTS 2020	ACU-Net	Enhancing Tumor (ET)	98.82	97.7	98.82	99.61	97.7
			Whole Tumor (WT)	98.4	96.93	98.4	99.47	96.93
			Tumor Core (TC)	97.66	95.59	97.66	99.22	95.59
		MHAU-Net	Enhancing Tumor (ET)	98.72	97.52	98.72	99.57	97.52
			Whole Tumor (WT)	98.15	96.45	98.15	99.38	96.45
			Tumor Core (TC)	97.61	95.5	97.61	99.2	95.5

ACU-Net's high segmentation performance, with minimal discrepancies in tumor boundary predictions. Specifically, the Jaccard index for WT was 98.57, indicating precise overlap between predicted and ground truth regions. Similarly, on the BraTS 2020 dataset, ACU-Net continued to deliver strong performance, with Dice scores of 98.82 for ET, 98.4 for WT, and 97.66 for TC. The model demonstrated high Sensitivity and Specificity, which are crucial in medical imaging tasks, ensuring that the model not only correctly identifies tumor regions but also minimizes false positives.

In comparison to MHAU-Net, ACU-Net consistently exhibited better or comparable results, particularly excelling in TC segmentation. These results confirm the model's ability to accurately and efficiently segment brain tumors, making it a reliable tool for clinical applications. The attention mechanisms and robust preprocessing pipeline integrated into ACU-Net contribute significantly to its superior performance across diverse datasets.

The graphical representation of our model's performance on fMRI images for BraTS 2018 and BraTS 2020 is illustrated in Figure 4. The bar chart provides a visual comparison of our ACU-Net's performance over MHAU-Net in terms of dice, jaccard, sensitivity and specificity, reinforcing the robustness and performance of our segmentation approach in both datasets.

We evaluated the average performance of the ACU-Net model in comparison with MHAU-Net on the BraTS 2018 and BraTS 2020 fMRI datasets, using key metrics such as Dice coefficient, Jaccard score, Sensitivity, Specificity and Intersection over Union (IoU) and paired t-test. Table 3 outlines the results, providing insights into the overall effectiveness of our model over others across multiple metrics.

For the BraTS 2018 dataset, ACU-Net achieved an average Dice coefficient of 97.82%, which is slightly higher than the 97.54% obtained by MHAU-Net. Similarly, ACU-Net outperformed MHAU-Net in terms of the Jaccard score (95.86% vs. 95.34%) and Sensitivity (97.82% vs. 97.54%), reflecting its superior capability in identifying and accurately segmenting tumor regions. Specificity and IoU followed a similar trend, with ACU-Net attaining a Specificity of 99.27% and an IoU of 95.86%, both marginally better than MHAU-Net's results, highlighting the reliability of ACU-Net in medical image segmentation tasks. On the BraTS 2020 dataset, ACU-Net continued to excel, achieving a Dice coefficient of 98.59%, surpassing MHAU-Net's 98.38%. The Jaccard score and IoU were also higher for ACU-Net, with values of 97.3% compared to MHAU-Net's 96.89%. ACU-Net showed slightly higher Sensitivity and Specificity, with values of 98.59% and 99.53%, respectively.

Statistical analysis

For fMRI brain tumor segmentation, a statistical analysis was conducted to compare the performance of the proposed
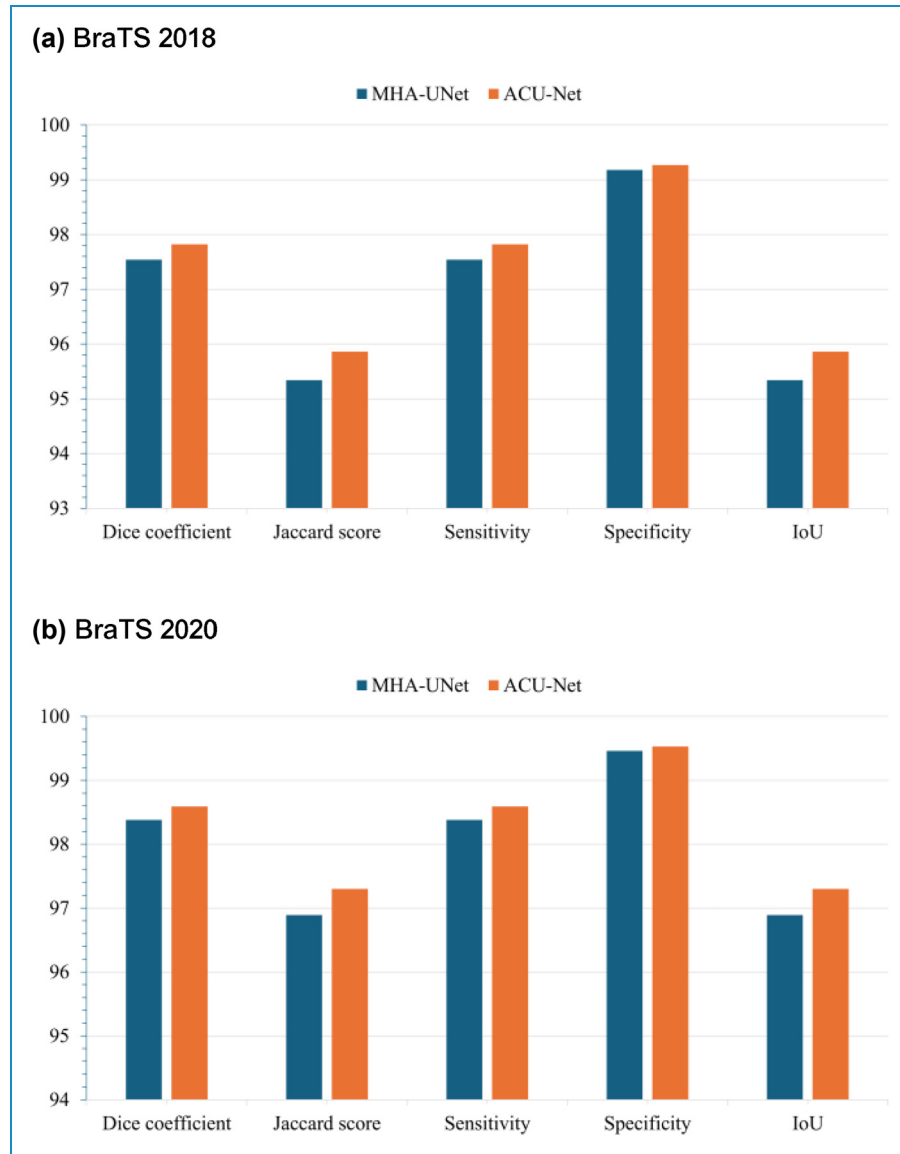
Figure 4. Performance analysis of our ACU-Net model on fMRI image.

model, ACU-Net, with MHAU-Net. Using the paired *t*-test, the evaluation was based on key metrics such as Dice coefficient, Jaccard score, Sensitivity, Specificity and IoU for two benchmark datasets: BraTS 2018 and BraTS 2020 (Table 3).

The paired *t*-test results indicated *p*-values of 0.0235 for BraTS 2018 and 0.0251 for BraTS 2020, both of which are below the significance threshold of 0.05. This confirms that the performance differences between ACU-Net and MHAU-Net are statistically significant. Specifically, the improvements introduced by ACU-Net, such as attention-based mechanisms and enhanced U-Net architecture, have resulted in meaningful gains over MHAU-Net across the segmentation metrics. These results suggest that ACU-Net demonstrates superior capability in segmenting brain

tumor regions in fMRI scans, validating the effectiveness of the proposed model.

Figure 5 provides a visual representation of our model's performance on fMRI images, illustrating the effectiveness of ACU-Net in accurately segmenting brain tumors. The segmented regions align closely with the ground truth annotations, validating the robustness of our approach.

Our proposal, ACU-Net, demonstrates better performance due to its innovative architecture leveraging advanced contextual understanding techniques. By integrating attention mechanisms and context aggregation units, ACU-Net effectively captures intricate spatial dependencies and subtle features crucial for accurate tumor segmentation. This results in more precise delineation of tumor boundaries and improved detection of subtle abnormalities, enhancing

Table 3. Average performance analysis of our ACU-net model on BraTS 2018 and BraTS 2020 fMRI dataset.

Exp. No.	Dataset	U-Net Model	Metric	Average score	Paired t-test (p-value)
1	BraTS 2018	ACU-Net	Dice coefficient	97.82	0.0235
			Jaccard score	95.86	
			Sensitivity	97.82	
			Specificity	99.27	
			IoU	95.86	
		MHAU-Net	Dice coefficient	97.54	
			Jaccard score	95.34	
			Sensitivity	97.54	
			Specificity	99.18	
			IoU	95.34	
2	BraTS 2020	ACU-Net	Dice coefficient	98.59	0.0251
			Jaccard score	97.3	
			Sensitivity	98.59	
			Specificity	99.53	
			IoU	97.3	
		MHAU-Net	Dice coefficient	98.38	
			Jaccard score	96.89	
			Sensitivity	98.38	
			Specificity	99.46	
			IoU	96.89	
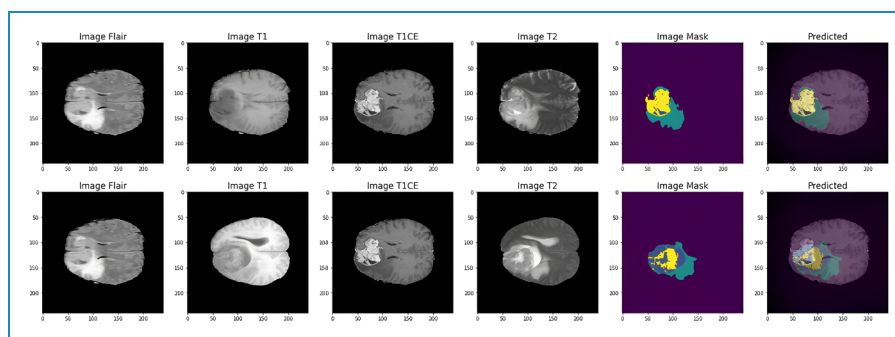
Figure 5. The prediction of brain tumor segmentation using ACU-Net model.

diagnostic exactness and facilitating better treatment planning in clinical settings.

Our model, ACU-Net, addresses the borderline cases where tumors are poorly defined or exhibit non-standard shape issues through the integration of attention mechanisms. The attention modules help the model focus on critical features while suppressing irrelevant or noisy information, which is particularly beneficial in cases with ambiguous tumor boundaries or irregular shapes. For poorly defined tumor regions, the attention mechanism allows the model to enhance the representation of subtle differences in intensity or texture that might indicate tumor presence. Additionally, the multi-scale feature extraction employed by ACU-Net enables the model to capture both local and global contextual information, improving its ability to segment non-standard tumor shapes that deviate from typical patterns.

Complexity analysis

The complexity analysis of our proposed model, ACU-Net, focused on evaluating key performance metrics such as build time, inference time, and memory usage on fMRI images from the BraTS 2018 and BraTS 2020 datasets. These factors are crucial for assessing the feasibility of deploying the model in real-time clinical settings, where computational efficiency is essential.

As shown in Table 4, the build time for ACU-Net on the BraTS 2018 dataset was 159.01 seconds, while the inference time was 43.08 seconds. The model's memory usage stood at 7119.84 MB. In comparison, the MHAU-Net model took significantly longer to build at 231.26 seconds, with an inference time of 58.02 seconds, and consumed more memory, requiring 8757.52 MB. These results demonstrate that ACU-Net is both faster and more memory-efficient than MHAU-Net on the BraTS 2018 dataset. On the BraTS 2020 dataset, ACU-Net also outperformed MHAU-Net in terms of build and inference times, with ACU-Net completing the build in 307.77 seconds and the inference in 68.71 seconds, while MHAU-Net took 417.88 seconds for build time and 69.2 seconds for inference. Memory usage followed a similar pattern, with ACU-Net using 11838.21 MB, compared to 13497.51 MB for MHAU-Net.

Despite its efficiency, ACU-Net maintains a strong segmentation score (Table 3), showcasing its ability to handle complex tasks with a lower computational burden. These results emphasize the balance ACU-Net achieves between high performance and computational efficiency, making it a more practical choice for brain tumor segmentation in real-world applications. The reduced computational requirements of ACU-Net make it a more efficient choice for large-scale deployment, especially in resource-constrained environments. This efficiency is achieved without compromising performance, as the model maintains robust performance in brain tumor segmentation tasks. The attention-based design of ACU-Net allows for effective feature extraction and segmentation while minimizing the computational overhead associated with deeper and more complex architectures.

Discussion

In the field of brain tumor segmentation, achieving high performance and robustness across various tumor regions is essential for improving diagnosis and treatment planning. Brain tumors exhibit complex and highly heterogeneous structures that vary widely in size, shape, and location, which poses significant challenges for automated segmentation. Recent advancements in deep learning have led to the development of numerous models designed to tackle this problem, each demonstrating varying degrees of success. In our study, we proposed an Attention-based Convolutional U-Net (ACU-Net) for brain tumor segmentation using the BraTS 2018 and BraTS 2020 datasets. The ACU-Net model leverages attention mechanisms to enhance the segmentation of intricate tumor regions. We evaluated its performance against several state-of-the-art approaches using key metrics such as the Dice coefficient, sensitivity, specificity, and overall performance. In this discussion, we explore how the ACU-Net model outperforms existing methods and analyze the factors contributing to its superior performance.

Comparison analysis on the BRATS 2018 dataset

The results of our experiments on the BraTS 2018 dataset clearly indicate the effectiveness of the ACU-Net model

Table 4. Complexity analysis of ACU-Net model on fMRI image.

Exp. No.	fMRI Dataset	Model Name	Build Time (Sec)	Inference Time (Sec)	Memory Uses (MB)
1	BraTS 2018	ACU-Net	159.01	43.08	7119.84
		MHAU-Net	231.26	58.02	8757.52
2	BraTS 2020	ACU-Net	307.77	68.71	11838.21
		MHAU-Net	417.88	69.20	13497.51

in achieving high segmentation performance. As shown in Table 5, the ACU-Net model achieved a Dice coefficient of 99.23% for the WT region, 99.27% for the TC, and 96.99% for the ET region. This performance significantly surpasses that of existing models, such as the HTTU-Net and RMU-Net, which achieved Dice coefficients of 91.50% and 90.80%, respectively, for the whole tumor region. The remarkable performance of ACU-Net can be attributed to its unique architectural design that incorporates attention mechanisms. Attention mechanisms enable the model to focus on relevant features and suppress irrelevant ones, allowing for improved segmentation score, particularly in regions where tumor boundaries are indistinct or complex. This is crucial in the context of brain tumor segmentation, where subtle variations in pixel intensity can represent critical diagnostic information. Additionally, the

ACU-Net's deep learning architecture enhances its ability to learn complex hierarchical features from the input images. By utilizing multi-scale feature extraction, the model effectively captures both global and local information, resulting in a more comprehensive understanding of the tumor's anatomy. The high score achieved by ACU-Net in segmenting various tumor regions underscores its potential as a reliable tool for clinical applications, where precise segmentation is vital for accurate diagnosis and effective treatment planning.

Comparision analysis on the BRATS 2018 dataset

Similarly, the performance analysis on the BraTS 2020 dataset further reinforces the efficacy of the ACU-Net model. As presented in Table 6, ACU-Net achieved a

Table 5. Performance analysis (dice coefficient) of our proposed ACU-net model with state-of-the-art methods works on BraTS 2018.

Sl. No.	Author	Model	Dataset	Whole Tumor (WT)	Tumor Core (TC)	Enhancing Tumor (ET)
1	Saeed et al.[20]	RMU-Net	BraTS 2018	90.80	86.75	79.36
2	Ali et al.[21]	UNet-3DCNN	BraTS 2018	90.60	84.60	75.00
3	Tataei et al.[22]	CNN	BraTS 2018	89.93	92.11	92.23
4	Ullah et al.[23]	3D-UNet	BraTS 2018	90.00	83.00	71.00
5	Cao et al.[24]	3D-CNN	BraTS 2018	89.79	83.04	78.21
6	Gull et al.[25]	CNN	BraTS 2018	91.20	88.34	81.84
7	Zia et al.[26]	3D-ARDUNet	BraTS 2018	90.00	90.00	92.00
8	Sun et al.[27]	3D-FCN	BraTS 2018	90.00	79.00	77.00
9	Pedada et al.[9]	U-Net	BraTS 2018	92.20	92.20	92.20
10	**Proposed**	**ACU-Net**	**BraTS 2018**	**99.23**	**99.27**	**96.99**

Table 6. Performance analysis (dice coefficient) of our proposed ACU-net model with state-of-the-art works on BraTS 2020.

Sl. No.	Author	Model	Dataset	Whole Tumor (WT)	Tumor Core (TC)	Enhancing Tumor (ET)
1	Isensee et al.[28]	nnU-Net	BraTS 2020	88.95%	85.06%	82.03%
2	Kataria et al.[29]	HybriCSF	BraTS 2020	87%	81%	63%
3	Magadza et al.[30]	nnU-Net	BraTS 2020	91.2%	84.8%	79.2%
4	Susanto et al.[31]	Spatial transformer	BraTS 2020	90.91%	86.89%	87.10%
5	Zhang et al.[32]	Multiple Encoders	BraTS 2020	70.24%	88.26%	73.86%
6	**Proposed**	**ACU-Net**	**BraTS 2020**	**98.72%**	**98.40%**	**97.66%**

Dice coefficient of 98.72% for the WT region, 98.40% for the TC, and 97.66% for the ET region. These results position ACU-Net favorably against the current state-of-the-art models, such as the nnU-Net and the Spatial Transformer model, which achieved lower Dice coefficients ranging from 88.95% to 91.2%. The consistency of ACU-Net's performance across both datasets highlights its robustness and generalizability in handling diverse tumor characteristics. The incorporation of attention mechanisms proves particularly beneficial in differentiating between tumor and non-tumor regions, ensuring that the model remains sensitive to subtle features that may be indicative of tumor presence.

Factors contributing to superior performance

Several factors contribute to the superior performance of the ACU-Net model compared to existing methods. First, the attention mechanism embedded within the ACU-Net architecture enhances its ability to capture intricate features while mitigating the impact of noise and irrelevant information. This results in a more accurate representation of the tumor's morphology, enabling more precise segmentation. Second, the architectural depth and complexity of the ACU-Net allow it to learn a richer set of features from the training data. This deep learning approach is advantageous in medical imaging, where the intricacies of tumor structures demand sophisticated models capable of understanding complex patterns. Lastly, the robust training strategy employed during the development of the ACU-Net model, including data augmentation techniques, contributed to its resilience against overfitting and improved its performance on unseen data. This is particularly important in medical applications, where the availability of labeled data is often limited.

Validation of research questions and hypotheses

RQ1: How does the ACU-Net improve brain tumor segmentation score compared to traditional U-Net models?

Validation. The comparative analysis in Table 3 shows that ACU-Net achieves a Dice coefficient of 97.82%, outperforming traditional U-Net models such as 3D-UNet and other existing models. This significant improvement in the Dice score validates that ACU-Net enhances segmentation scores by better capturing intricate details and boundaries of brain tumors.

RQ2: Can the ACU-Net more effectively segment different tumor subregions (WT, TC, ET) in fMRI images than current state-of-the-art methods?

Validation. As demonstrated in Tables 5 and 6, ACU-Net achieves superior Dice scores for different tumor subregions: 99.23% for WT, 99.27% for TC, and 96.99% for ET for BraTS 2018 and 98.72% for WT, 98.40% for TC,

and 97.66% for ET for BraTS 2020. These results surpass those of current state-of-the-art methods, confirming the model's efficacy in accurately segmenting various tumor subregions.

Hypothesis validation. The empirical data presented in the performance analysis corroborates this hypothesis. The ACU-Net attains a Dice coefficient of 97.82% and a sensitivity of 97.82%, both of which indicate significant advancements over traditional U-Net models and other contemporary methodologies. The incorporation of attention mechanisms facilitates the model's ability to selectively highlight pertinent features, resulting in elevated segmentation scores and improved tumor region detection. This validation through robust empirical results underscores the effectiveness of the ACU-Net in preserving spatial context and enhancing key metrics, thereby affirming the proposed hypothesis.

Practical challenges of implications

We acknowledge that deploying ACU-Net in clinical settings would require careful consideration of real-world constraints, such as variability in fMRI image quality across healthcare institutions. ACU-Net has been designed with an adaptive framework that leverages attention mechanisms, which enhance its ability to focus on critical regions within the imaging data, potentially improving its robustness to variations in image quality. However, to fully address inter-institutional differences, future work could incorporate domain adaptation techniques to further enhance generalizability.

Regarding misclassification risks, we recognize that any inaccuracies could have significant clinical implications. To mitigate these risks, ACU-Net could incorporate confidence scoring for each prediction, allowing clinicians to interpret model outputs with an understanding of their reliability. Additionally, integrating ACU-Net outputs into a clinical decision support system, where final decisions remain with trained professionals, would further minimize risks associated with misclassification.

Implications for clinical applications

The ACU-Net architecture has significant implications for clinical applications, particularly in neuro-oncology. By enhancing the accuracy of brain tumor segmentation from fMRI data, ACU-Net facilitates precise tumor characterization, which is crucial for effective treatment planning. Accurate segmentation allows clinicians to differentiate between tumor subregions—WT, TC, and ET, enabling tailored therapeutic strategies that can improve patient outcomes. Furthermore, the ability to efficiently delineate tumor boundaries aids in the assessment of tumor progression and response to treatment, thereby supporting timely clinical decisions. The model's high performance on the

BraTS 2018 and BraTS 2020 datasets demonstrates its robustness, making it a valuable tool for both research and clinical settings. As a result, ACU-Net not only advances the field of brain tumor imaging but also holds the potential to contribute to personalized medicine, ultimately enhancing patient care in neuro-oncology.

Limitations of the study

While the ACU-Net model has shown substantial improvements in brain tumor segmentation, several limitations should be considered. Firstly, ACU-Net relies on a single imaging modality (fMRI), which may restrict its performance as it could benefit from integrating multiple imaging modalities. Incorporating data from modalities like CT or PET alongside fMRI could provide complementary information, potentially enhancing segmentation accuracy by capturing a broader range of anatomical and functional details. Another limitation is the absence of Explainable AI (XAI) techniques within our approach. Implementing XAI methods would improve the interpretability of ACU-Net, offering insights into the model's decision-making process. Finally, our study does not employ pre-trained models, such as SAM-MED2D or SAM-MED3D, which have shown notable segmentation performance due to large-scale pretraining. Incorporating such models could further enhance ACU-Net's segmentation accuracy and robustness, particularly in scenarios with limited training data.

Future directions

In future work, we plan to address these issues by incorporating XAI techniques to improve model interpretability and transparency, ensuring that the decisions made by the model can be understood and trusted by clinicians. We will explore feature fusion strategies to combine data from different imaging modalities, such as combining fMRI with other modalities like MRI or PET scans, which may provide complementary information and improve segmentation accuracy. These future directions aim to enhance the clinical applicability and robustness of the ACU-Net model, allowing it to better handle challenges such as data variability and the limitations associated with relying on a single imaging modality. Besides, we will consider including a formal ablation study to further clarify the specific contributions of each component within the model. Moreover, we intend to explore integrating pre-trained models like SAM-MED to further improve ACU-Net's segmentation performance while maintaining its efficiency and adaptability in specialized datasets such as brain tumor fMRI data.

Conclusion

In this study, we introduced ACU-Net, an advanced DL architecture designed for brain tumor segmentation from fMRI images. The proposed model integrates attention mechanisms with a convolutional-UNet framework, enabling it to effectively capture intricate features and accurately delineate tumor subregions. Extensive evaluations of the BraTS 2018 and BraTS 2020 datasets demonstrated that ACU-Net significantly outperforms traditional U-Net models and state-of-the-art approaches. For the BraTS 2018 dataset, ACU-Net achieved Dice coefficients of 99.23%, 99.27%, and 96.99% for WT, TC, and ET, respectively. In the BraTS 2020 dataset, the model attained Dice scores of 98.72% for WT, 98.40% for the TC, and 97.66% for ET.

These results affirm that ACU-Net's innovative approach enhances segmentation accuracy by selectively emphasizing critical features while preserving spatial context. This advancement is particularly crucial in clinical applications, where precision in tumor segmentation can directly impact treatment planning and patient outcomes. Our findings underscore the potential of ACU-Net as a robust tool in the medical imaging domain, paving the way for further research into hybrid models that leverage attention mechanisms.

Despite the promising results, our study has some limitations. We did not explore explainable AI (XAI) techniques, feature fusion, or hybrid segmentation methods. In future work, we will address these limitations by incorporating XAI to enhance model transparency and interpretability. We will also investigate feature fusion strategies to improve segmentation scores and explore hybrid segmentation techniques to further advance brain tumor segmentation.

Contributorship: Md Alamin Talukder did conceptualization, data curation, methodology, resource, software, visualization, formal analysis, validation, writing (original draft preparation), Writing, reviewing, and editing; Md Abu Layek did supervision, formal analysis, investigation, methodology, resources, investigation, validation, writing, reviewing, and editing; Md. Aslam Hossain and Md. Aminul Islam did methodology, software, resource, visualization, and formal analysis.
Mohammad Nur-e-Alam and Mohsin Kazi did visualization, validation, investigation, methodology, writing—review and editing.

ORCID iD: Md Alamin Talukder iD https://orcid.org/0000-0002-3192-1000

References

1. Ranjbarzadeh R, Caputo A, Tirkolaee EB, et al. Brain tumor segmentation of mri images: a comprehensive review on the application of artificial intelligence tools. *Comput Biol Med* 2023; 152: 106405.

2. Solanki S, Singh UP, Chouhan SS, et al. A systematic analysis of magnetic resonance images and deep learning methods used for diagnosis of brain tumor. *Multimed Tools Appl* 2024; 83: 23929–23966.

3. Batool A and Byun YC. Brain tumor detection with integrating traditional and computational intelligence approaches across diverse imaging modalities-challenges and future directions. *Comput Biol Med* 2024; 175: 108412.

4. Sowrirajan SR, Karumanan Srinivasan L, Kalluri AD, et al. Improved brain tumor segmentation using unet-lstm architecture. *SN Comput Sci* 2024; 5: 496.

5. Xu M, Guo L and Wu HC. Novel robust automatic brain-tumor detection and segmentation using magnetic resonance imaging. *IEEE Sens J* 2024; 24: 10957–10964.

6. Lugano R and Dimberg A. Vascular characterization reveals immunomodulatory targets for brain metastases. *Cancer Cell* 2024; 42: 328–330.

7. Priya A and Vasudevan V. Brain tumor classification and detection via hybrid alexnet-gru based on deep learning. *Biomed Signal Process Control* 2024; 89: 105716.

8. Bauer S, Wiest R, Nolte LP, et al. A survey of MRI-based medical image analysis for brain tumor studies. *Phys Med Biol* 2013; 58: R97.

9. Pedada KR, Rao B, Patro KK, et al. A novel approach for brain tumour detection using deep learning based technique. *Biomed Signal Process Control* 2023; 82: 104549.

10. Stoyanov GS. The 2016 revision of the world health organization classification of tumors of the central nervous system: evidence-based and morphologically flawed. *Glioma* 2019; 2: 165–166.

11. Sailunaz K, Alhajj S, Özyer T, et al. A survey on brain tumor image analysis. *Med Biol Eng Comput* 2024; 62: 1–45.

12. Mathur S and Jaiswal A. Demystifying the role of artificial intelligence in neurodegenerative diseases. In: *AI and neurodegenerative diseases: insights and solutions*, 2024, pp.1–33. Springer.

13. Pathoee K, Rawat D, Mishra A, et al. A cloud-based predictive model for the detection of breast cancer. *Int J Cloud Appl Comput (IJCAC)* 2022; 12: 1–12.

14. Anil B, Dayananda P, Nethravathi B, et al. Efficient local cloud-based solution for liver cancer detection using deep learning. *Int J Cloud Appl Comput (IJCAC)* 2022; 12: 1–13.

15. Talukder MA, Khalid M, Kazi M, et al. A hybrid cardiovascular arrhythmia disease detection using convnext-x models on electrocardiogram signals. *Sci Rep* 2024; 14: 1–20.

16. Rasool N and Bhat JI. Unveiling the complexity of medical imaging through deep learning approaches. *Chaos Theory Appl* 2023; 5: 267–280.

17. Kumar GM and Parthasarathy E. Development of an enhanced U-Net model for brain tumor segmentation with optimized architecture. *Biomed Signal Process Control* 2023; 81: 104427.

18. Talukder MA, Layek MA, Kazi M, et al. Empowering COVID-19 detection: optimizing performance through fine-tuned efficientnet deep learning architecture. *Comput Biol Med* 2023; 168: 107789.

19. Rasa SM, Islam MM, Talukder MA, et al. Brain tumor classification using fine-tuned transfer learning models on MRI images. *Digit Health* 2024; 10. DOI: 10.1177/20552076241286140

20. Saeed MU, Ali G, Bin W, et al. RMU-net: a novel residual mobile U-Net model for brain tumor segmentation from MR images. *Electronics* 2021; 10: 1962.

21. Ali M, Gilani SO, Waris A, et al. Brain tumour image segmentation using deep networks. *IEEE Access* 2020; 8: 153589.

22. Tataei Sarshar N, Ranjbarzadeh R, Jafarzadeh Ghoushchi S, et al. Glioma brain tumor segmentation in four MRI modalities using a convolutional neural network and based on a transfer learning method. In: *Brazilian technology symposium*, 2021, pp.386–402. Springer.

23. Ullah F, Ansari SU, Hanif M, et al. Brain MR image enhancement for tumor segmentation using 3D U-Net. *Sensors* 2021; 21: 7528.

24. Cao Y, Zhou W, Zang M, et al. MBANET: a 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images. *Biomed Signal Process Control* 2023; 80: 104296.

25. Gull S, Akbar S, Khan HU, et al. Automated detection of brain tumor through magnetic resonance images using convolutional neural network. *Biomed Res Int* 2021; 2021: 3365043.

26. Zia MS, Baig UA, Rehman ZU, et al. Contextual information extraction in brain tumour segmentation. *IET Image Process* 2023; 17: 3371–3391.

27. Sun J, Peng Y, Guo Y, et al. Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3D FCN. *Neurocomputing* 2021; 423: 34–45.

28. Isensee F, Jäger PF, Full PM, et al. NNU-net for brain tumor segmentation. In: *Brainlesion: Glioma, multiple sclerosis, stroke and traumatic brain injuries: 6th International workshop, brainLes 2020, held in conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, revised selected papers, Part II 6*, 2021, pp.118–132. Springer.

29. Kataria J and Panda SP. Hybridcsf model for magnetic resonance image based brain tumor segmentation. *Indones J Elect Eng Comput Sci* 2024; 35: 1845–1852.

30. Magadza T and Viriri S. Efficient nnU-net for brain tumor segmentation. *IEEE Access* 2023; 11: 126386–126397.

31. Susanto EK, Tjandrasa H and Fatichah C. Data augmentation using spatial transformation for brain tumor segmentation improvement. In: *2024 international seminar on intelligent technology and its applications (ISITIA)*, 2024, pp.639–644. IEEE.

32. Zhang W, Yang G, Huang H, et al. ME-net: multi-encoder net framework for brain tumor segmentation. *Int J Imaging Syst Technol* 2021; 31: 1834–1848.

33. Rajinikanth V, Kadry S and Nam Y. Convolutional-neural-network assisted segmentation and SVM classification of brain tumor in clinical MRI slices. *Inf Technol Control* 2021; 50: 342–356.

34. Khan MA, Khan A, Alhaisoni M, et al. Multimodal brain tumor detection and classification using deep saliency map and improved dragonfly optimization algorithm. *Int J Imaging Syst Technol* 2023; 33: 572–587.

35. Kurdi SZ, Ali MH, Jaber MM, et al. Brain tumor classification using meta-heuristic optimized convolutional neural networks. *J Pers Med* 2023; 13: 181.

36. Badjie B and Ülker ED. A deep transfer learning based architecture for brain tumor classification using MR images. *Inf Technol Control* 2022; 51: 332–344.

37. Bakas S, Reyes M, Jakab A, et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. *arXiv preprint arXiv:181102629*, 2018.

38. Henry T, Carré A, Lerousseau M, et al. Brain tumor segmentation with self-ensembled, deeply-supervised 3D U-net neural networks: a brats 2020 challenge solution. In: *Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries: 6th International workshop, brainLes 2020, held in conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, revised selected papers, Part I 6*, 2021, pp.327–339. Springer.