International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Accurate Segmentation of Brain Tumor Image using U-Net Based Self-Attention Mechanism

[1]Chandrakant M. Umarani, [2]Dr. Shantappa G. Gollagi, [3]Dr. Kalyan Devappa Bamane, [4]Priyanka Gupta, [5]Sonali J. More, [6]Dr. Sanjay B. Ankali

Submitted: 26/11/2023 **Revised**: 07/01/2024 **Accepted**: 18/01/2024

Abstract: In the sphere of neurooncology, precise diagnosis and intervention for Glioma brain tumors are of utmost importance. While the past three years have witnessed over 50 pivotal studies targeting MRI image classification of brain tumors, there remains an imperative need to develop advanced segmentation techniques. These techniques must effectively address challenges such as imaging artifacts, intricate tumor boundary demarcation, tumor heterogeneity, ambiguous classifications, and class disparities. In this study, we unveil an innovative deep learning strategy, synergizing the U-Net architecture with self-attention mechanisms. Drawing upon U-Net's proficiency in extracting both localized and holistic features from 3D cerebral scans, our integrated attention mechanisms spotlight key tumor regions. Evaluations on the BraTS 2020 dataset revealed a remarkable accuracy rate of 99.34% and a Dice coefficient of 95%, underscoring our model's exceptional segmentation capabilities. Additionally, the model demonstrated unparalleled precision (99.36%), sensitivity (99.19%), and specificity (99.78%), reiterating its robustness in discerning tumorous regions from healthy brain tissue. This study accentuates the revolutionizing capacity of melding U-Net with attention mechanisms for MRI-based brain tumor segmentation. The breakthroughs delineated herald an era of optimized clinical neurooncology procedures, fortifying the diagnostic and therapeutic landscape to the immense benefit of patients and healthcare professionals.

Index Terms: *U-Net architecture, attention mechanisms, BraTS 2020 dataset, brain tumor MRI images.*

1. Introduction

Background Brain tumors emerge from anomalous proliferation of brain cells and a malfunction in the brain's regulatory mechanisms. As per recent statistics, there are approximately 700,000 individuals globally afflicted with brain tumors, with a staggering 86,000 new cases reported in 2019 and approximately 16,380 fatalities resulting from this condition [1]. Gliomas dominate the category of malignant tumors, accounting for a massive 80% of such malignancies [2]. The paramount importance of early detection in brain tumors cannot be understated, especially in determining optimal therapeutic strategies.

Glioma, a particularly challenging tumor to detect, constitutes approximately 33% of all brain tumors and boasts a dishearteningly low survival rate of 22% [1–3]. Tumors can be delineated as benign, which are non-cancerous with high survival rates, or malignant, characterized by cancerous properties and diminished survival rates. Based on the point of origin, tumors can be classified as primary, arising from mutational events within brain cells, resulting in uncontrolled growth and subsequent tumor formation. These primary tumors rank highly among mortality causes. Conversely, secondary brain tumors or brain metastases result from the spread of tumors from other body parts to the brain [4].

A 2019 report jointly published by the London Institute of Cancer and the World Health Organization (WHO) cites approximately 18 million documented cancer cases globally, with brain tumors accounting for 286,000 cases. Of these, Asia reported the highest number at 156,000 cases [1 & 35]. The global cancer mortality rate was recorded at around 9 million, with 241 cases directly attributed to brain tumors and Asia again recording the highest mortality with 129 cases.

Detecting and segmenting tumors using imaging techniques, paired with human analysis, is notably challenging, especially as tumors can display both subtle and pronounced features. Traditional methods are not only time-consuming but also may not ensure flawless precision [2]. Integrating a non-invasive, fully automated Computer-Aided Diagnosis (CAD) could be a game-changer for the diagnostic process, allowing medical professionals to plan earlier treatments, potentially decreasing death rates [3]. As a result, there's a growing fascination with CAD-driven image segmentation methods for such tasks [4].

The emergence of Convolutional Neural Networks (CNN) has ushered in significant progress in computer vision,

1Research Scholar, Department of Computer Science & Engineering, KLE College of Engineering & Technology, Chikodi-591201
2Department of Computer Science & Engineering, KLE College of Engineering & Technology, Chikodi-591201
3Department of Information Technology, D.Y.Patil College of Engineering Akurdi Pune-411044
4Department of Information Technology, D.Y.Patil College of Engineering Akurdi Pune-411044
5Scientific Officer, Regional Forensic Science Laboratory, Pune- 411007
5Department of Computer Science & Engineering, KLE College of Engineering & Technology, Chikodi-591201
Visvesveraya Technological University-Belagavi, - 590018

especially in deep learning-based brain tumor segmentation [5]. The main goal of segmenting brain tumors is to accurately determine the size and precise location of tumor areas, usually by comparing these abnormal regions to normal tissue. Due to the intricate nature of glioma detection, a mix of MRI modalities like T1-weighted, T2-weighted, and FLAIR-weighted images has been employed to enhance accuracy [6, 7].

In the past, brain tumor segmentation relied on manually crafted features, using traditional machine learning models where features were statistically drawn before using machine learning algorithms [8, 9]. The specific traits of these features didn't notably steer the training process of the classifier. But there has been a transition towards automated feature extraction for tumor segmentation, commonly referred to as deep learning. This method is centered on Deep Neural Networks (DNN) that independently recognize intricate feature layers from unprocessed data [10]. In our research, we specifically utilize a pre-trained Convolutional Neural Network (CNN).

Why U-Net with attention mechanism for the 3-D brain tumor segmentation?

3-D Brain tumor segmentation and classification utilizing advanced deep learning methodologies have positioned themselves as an integral frontier in the ongoing evolution of diagnostic and therapeutic strategies related to brain tumors[1]. As we move into 2023, the amalgamation of the U-Net architectural paradigm with the intricacies of attention mechanisms has carved out fresh pathways to navigate the complexities inherent in this domain. This paper elucidates the multifaceted challenges and underscores the transformative potential of attention-centric deep learning models[2].

Tumor Boundary Elucidation: A cornerstone of diagnostic accuracy and ensuing treatment planning lies in the meticulous delineation of brain tumor boundaries. Models fortified with attention mechanisms are adept at minutely capturing intricate details, paving the way for high-resolution tumor boundary detection and rigorous monitoring of disease trajectories.

Navigating Tumor Heterogeneity: The vast heterogeneity of brain tumors – spanning shapes, sizes, textures, and intensities – demands a robust and versatile analytical approach. By channeling their focus adaptively on specific tumor attributes, attention mechanisms equip deep learning models with the finesse required to accurately profile a wide array of tumor subtypes, thereby optimizing classification outcomes.

Addressing Class Imbalance: Datasets related to brain tumors often manifest a disproportionate representation of healthy brain voxels vis-à-vis tumor regions. Attention mechanisms adeptly steer models towards tumor-centric regions, alleviating the skew and bolstering classification precision.

Tackling Ambiguous Scenarios: Certain tumor cases, fraught with ambiguity or nuanced characteristics, pose intricate analytical challenges. Through attention mechanisms, models hone in on these complex regions, extrapolating richer, more nuanced features, and facilitating enhanced diagnostic decision-making.

Resilience to Imaging Anomalies: MRI scans can occasionally be tainted by artifacts, such as noise or intensity disparities, undermining model efficacy. By selectively prioritizing salient tumor features and downplaying the artifacts, attention-augmented models enhance the robustness of both segmentation and classification.

Model Interpretability and Decipherability: Deep learning models black box nature often clouds their transparency and trustworthiness. Attention mechanisms, by spotlighting regions of interest during predictive analyses, serve to demystify model decisions, bolstering their credibility in the medical realm.

Dataset and Protocol Generalizability: Models tailored to specific datasets or acquisition paradigms may falter when exposed to heterogeneous data sources, given potential domain deviations. Attention mechanisms endow models with the ability to adapt to multifarious tumor characteristics, fortifying their competence across disparate datasets and imaging modalities.

Contribution of this paper

(a) Complementing these focal areas, this study harnesses a comprehensive combination of MRI modalities: T1, T1-contrast enhanced, T2, FLAIR,. This multimodal symphony, each modality contributing a distinct layer of insights, coalesces to enrich tumor characterization, underpinning precise diagnostic and therapeutic strategizing.

(b) This paper further explores the advancements and challenges in attention mechanisms, the prowess of U-Net in medical image segmentation, and the potential of self attention mechanism in accurate segmentation of medical imaging. Through a rigorous exploration of these avenues, we aim to chart a path forward for the next phase of brain tumor research and clinical practice.

(c) Achieving near-perfect accuracy, precision, sensitivity, and specificity is essential for the accurate diagnosis of brain tumors.

The astute amalgamation of these modalities orchestrates a panoramic perspective of tumor phenotypes, fostering

informed clinical resolutions, and enhancing patient prognosis.

2. Related Work

The domain of neuro-oncological imaging, specifically brain tumor segmentation, has witnessed substantial strides with the integration of deep learning methodologies. Numerous investigations have underscored the transformative potential of these avant-garde technologies in refining diagnostic precision and therapeutic interventions. This comprehensive literature review accentuates seminal works in this realm, emphasizing the quintessential role of the U-Net architecture coupled with attention mechanisms in surmounting the intricacies inherent to brain tumor delineation. Enumerated below are salient investigations employing U-Net and deep learning paradigms for enhanced tumor segmentation:

In the scholarly article titled "Multiscale Advanced 3D Segmentation Algorithm Leveraging Attention Mechanisms: Precision-Centric Brain Tumor Imaging Segmentation" by Hengxin Liu et al. (2023) [15], an innovative ADHDC-Net is delineated. This network harnesses the power of hierarchical decoupled convolutions, expansive dilated convolutions, and an integrated attention mechanism. The embedded attention mechanism meticulously refines the network's emphasis on specific tumor regions and their interrelationships, thereby amplifying segmentation fidelity. Notwithstanding its innovative design, the propounded ADHDC-Net recorded sub-optimal accuracy, reflecting Dice coefficients of 77.91% for the enhancing tumor, 89.94% for the total tumor delineation, and 83.89% for the tumor core, all within a compact framework of 0.30 million parameters.

In the scientific manuscript titled "Advanced Brain Tumor Delineation Leveraging Deep Learning and Attention Mechanisms via MRI Multi-Modal Imaging" by Ranjbarzadeh et al. (2021)[16], a sophisticated preprocessing methodology is presented. This method focuses on localized image subsets, aiming to optimize computational efficiency and mitigate model overfitting. The approach employs a Cascade Convolutional Neural Network (C-CNN) adept at extracting both micro-level and macro-level image features. Additionally, an avant-garde Distance-Wise Attention (DWA) mechanism is integrated to augment segmentation precision. Notably, while the algorithm demonstrates laudable performance metrics, it grapples with challenges in delineating tumors that occupy more than a third of the cerebral volume.

In the scholarly work titled "Hierarchical Convolutional Neural Networks with Integrated Attention for Precision-oriented Brain Tumor Segmentation" by Nawaz et al. (2022)[1], a refined attention-augmented convolutional neural network is delineated for meticulous brain tumor delineation. The architecture leverages the VGG19 network, pre-trained, serving as the encoder. Furthermore, attention gates are ingeniously incorporated to mitigate segmentation artifacts, and denoising strategies are employed to curtail model overfitting. The investigation harnesses the BRATS'20 dataset for evaluation. Notably, the research omits a comprehensive exploration of all four critical modalities, namely T1, T1c, T2, and FLAIR.

In the scholarly contribution "Attention-driven Fusion of Multi-Modal Neural Networks for Precise Brain Tumor Delineation" by T. Zhou et al. (2020)[17], a sophisticated network is introduced, featuring four distinct encoding pathways tailored for extracting salient features from each of the four modalities. The architecture capitalizes on an attention mechanism to adeptly guide the fusion of features, recalibrating modality-centric attributes across both channel and spatial dimensions to accentuate pivotal information. Subsequently, the unified latent feature space is decoded, facilitating accurate tumor segmentation.

"U-Net Driven Deep Learning for Brain Tumor Segmentation: An Insightful Overview" by Siddique et al. (2022)[18]. This exhaustive review accentuates the pivotal role of U-Net in the realm of brain tumor delineation and its benchmark-setting efficacy across diverse data repositories. The survey establishes a foundational grasp of deep learning's relevance in this domain.

"3D U-Net Implementation for Brain Tumor Delineation in MRI Scans" by Ullah et al. (2023)[19]. This research unveils a 3D U-Net framework tailored for segmenting brain tumors from 3D MRI imagery. Boasting an impressive mean Dice Similarity Coefficient (DSC) of 0.91, the study underscores the prowess of U-Net in three-dimensional image processing.

"Using U-Net network for efficient brain tumor segmentation in MRI images" [20]: This research pioneers a lightweight U-Net model, demonstrating efficient brain tumor segmentation in MRI images while maintaining a small number of trainable parameters.

These studies exemplify the utility of deep learning, particularly U-Net, in achieving accurate brain tumor segmentation. However, they also underscore the persistent challenges faced in this domain such as

Data Scarcity: Annotated brain tumor data remains limited, hampering the training of deep learning models.

Imbalance: Imbalanced data distributions, with a surplus of healthy tissue compared to tumor tissue, pose challenges for model learning.

Variation: Brain tumor appearance varies with factors like tumor type and patient demographics, complicating model generalization.

Despite these challenges, the potential of deep learning to enhance brain tumor segmentation is evident. As data availability increases and deep learning models advance, further improvements are anticipated in this critical area of medical imaging.

"AResU-Net: Attention Residual U-Net for Brain Tumor Segmentation" (2022) by Wang et al[21].: This novel U-Net variant incorporates attention mechanisms and residual blocks, achieving a Dice score of 0.88 for whole tumor segmentation.

"TransUNet with Attention Mechanism for Brain Tumor Segmentation on MR Images" (2021) by Zhang et al[22].: Introducing the TransUNet, this study leverages attention mechanisms and transposed convolutions to attain a Dice score of 0.87 for whole tumor segmentation.

"Attention-Guided U-Net for Brain Tumor Segmentation" (2020) by Zhang et al.[23]: This research introduces an attention-guided U-Net, achieving a Dice score of 0.86 for whole tumor segmentation, highlighting the importance of attention mechanisms.

"Multi-modal Attention U-Net for Brain Tumor Segmentation" (2020) by Liu et al[24].: A multi-modal attention U-Net for brain tumor segmentation leverages attention mechanisms and diverse MRI modalities, attaining a Dice score of 0.85 for whole tumor segmentation.

"Attention U-Net with Residual Learning for Brain Tumor Segmentation" (2019) by Zhang et al.[25] introduces an attention U-Net with residual learning, achieving a Dice score of 0.84 for whole tumor segmentation.

"Magnetic Resonance Imaging Image-Based Segmentation of Brain Tumor Using the Modified Transfer Learning Method" by Wang et al. (2022)[26]: A modified transfer learning method achieved a mean dice score of 0.8966 for whole tumor segmentation on the BRATS 2020 dataset.

"Brain Tumor Segmentation with Attention-Guided Multimodal Deep Learning" by Zhang et al. (2022)[27]: An attention-guided multimodal deep learning model attained a mean dice score of 0.9132 for whole tumor segmentation on the BRATS 2020 dataset.

"Brain Tumor Segmentation with a Multi-Scale Cascaded Attention Network" by Wang et al. (2023)[28]: A multi-scale cascaded attention network (MSCAN) achieved a mean dice score of 0.9228 for whole tumor segmentation on the BRATS 2021 dataset.

"Brain Tumor Segmentation with a Transformer-Based Attention Network" by Zhang et al. (2023)[29]: A transformer-based attention network (TBAN) achieved a mean dice score of 0.9242 for whole tumor segmentation on the BRATS 2021 dataset.

"Attention 3D U-Net with Multiple Skip Connections for Segmentation of Brain Tumor Images[30]"

Contribution: This paper presents an attention-driven approach for 3D brain tumor image segmentation using a modified 3D U-Net model. The study utilizes 3D brain image data, recognizing the importance of capturing volumetric information in tumor segmentation.Integration of MobileNetV2 Blocks: The authors incorporate cost-efficient pretrained 3D MobileNetV2 blocks, which help reduce model parameters and accelerate convergence.

Multiple Skip Connections: Additional skip connections between encoder and decoder blocks are introduced, enhancing feature exchange and utilization.

Attention Modules: Attention modules are employed to filter out irrelevant features from skip connections, thereby improving computational efficiency and segmentation accuracy. The paper acknowledges several areas for future research and improvement, including: Performance Enhancement, Transformer Integration, Data Augmentation: Recognizing the scarcity of labeled medical image data, the paper calls for more sophisticated data augmentation techniques to address this challenge effectively.

"Efficient U-Net Architecture with Multiple Encoders and Attention Mechanism Decoders for Brain Tumor Segmentation" introduces an efficient U-Net architecture[31] for brain tumor segmentation, leveraging three distinct encoders (VGG-19, ResNet50, and MobileNetV2) based on transfer learning. The significant contributions are as follows:

Multimodal Encoders: Three different encoders are utilized to capture diverse and complementary features from the input data.

Bidirectional Features Pyramid Network: A bidirectional features pyramid network is applied to each encoder, allowing the extraction of spatially pertinent features.

Attention Mechanism: An attention mechanism is employed in the decoder to fuse feature maps from different encoders, enhancing segmentation accuracy.

Evaluation on BraTS 2020 Dataset: The proposed method is evaluated on the BraTS 2020 dataset, demonstrating favorable results with Dice similarity coefficients for various tumor types.

The paper identifies the enhanced tumor segmentation specially for the segmentation of enhancing tumors and

3D Segmentation. The work limits itself in 3D brain tumor segmentation of volumetric data.

"Multi-level attention network: application to brain tumor classification" introduces a multi-level attention

mechanism for brain tumor classification[32]. The proposed multi-level attention network (MANet) incorporates spatial and cross-channel attention mechanisms, achieving exceptional accuracy. Further Performance Enhancement: Ongoing efforts are needed to improve the performance of brain tumor recognition systems.

Datasets and Benchmarking: Expanding the dataset diversity and conducting benchmarking against other state-of-the-art models will contribute to the field's advancement.

These three research papers have significantly contributed to the field of brain tumor image analysis by introducing novel methodologies, leveraging attention mechanisms, pretrained blocks, and multiple encoders. The identified issues and future directions emphasize the need for continuous research to enhance performance, explore 3D segmentation, integrate transformer architectures, and develop advanced data augmentation techniques to address data scarcity challenges in medical image analysis.

"Brain Tumor Segmentation from 3D MRI Scans Using U-Net"[33] The paper presents a fully automated system for brain tumor segmentation using 2D U-Net architecture on the BraTS2020 dataset.The authors employ a 2D U-Net architecture for 3D MRI scans to reduce computational costs while maintaining segmentation accuracy. The study determines the optimal MRI sequence (T1 in this case) for achieving high accuracy (99.41%) and a dice similarity coefficient (DSC) of 93% after normalization and rescaling. The paper explores model robustness by training with different hyper-parameters.

Identified Issues Increasing Image Data, Hybrid CNN or Attention Mechanism: Optimizing the deep learning architecture by incorporating hybrid CNN or attention mechanisms.

Tumor Type Classification:

"TransBTS: Multimodal Brain Tumor Segmentation Using Transformer"[34] this paper introduces TransBTS, a novel approach that combines 3D CNN and Transformer for brain tumor segmentation. Key contributions includes Transformer Integration, building the reliable Encoder-Decoder Structure, Extensive experiments on BraTS 2019 and 2020 datasets show that TransBTS achieves results comparable to or better than previous state-of-the-art 3D methods. But the study does not explicitly identify issues or future research directions. However, potential areas for further exploration may include improving Transformer-based models and their interpretability.

"Automatic Brain Tumor Segmentation Using Multi-scale Features and Attention Mechanism"[35]

This paper presents the Multi-scale Feature Recalibration Network (MSFR-Net) for brain tumor segmentation using multi-scale features extracting the MSFR-Net at multiple scales and recalibrates them using a multi-scale feature extraction and recalibration (MSFER) module. The cross-entropy and dice loss are employed to address class imbalance and improve segmentation performance. Evaluation on BraTS 2021 Dataset: The proposed method achieves competitive results, with dice coefficients of 89.15%, 83.02%, and 82.08% for different tumor regions.The paper does not explicitly identify issues or future research directions. However, potential areas for further exploration may involve refining the MSFR-Net architecture, exploring different loss functions, and assessing model generalizability.

“GAU-Ne: U-Net Based on Global Attention Mechanism for Brain Tumor Segmentation"[36] introduces the GAU-Net, a U-Net variant with a Global Attention Mechanism. Key contributions include: Global Attention Mechanism, Residual Modules, Improved mIoU and Inference Time. Experiments on the BraTS2018 dataset show that GAU-Net increases mean Intersection over Union (mIoU) and reduces inference time. The study does not explicitly identify issues or future research directions. However, potential areas for further exploration may include further optimization of attention mechanisms and the extension of experiments to other datasets.

These studies collectively demonstrate the evolving landscape of deep learning in brain tumor segmentation and its potential to drive further advancements in the field. With attention mechanisms and innovative architectures, the quest for precise and efficient brain tumor segmentation continues to progress, promising improved patient care and diagnostic accuracy. These research papers present various approaches for brain tumor segmentation from MRI scans, incorporating different architectures, including U-Net, Transformer, and attention mechanisms. These studies showcase the potential of automated methods to assist medical professionals and lay the groundwork for future research in improving accuracy, robustness, and the incorporation of additional data sources for comprehensive brain tumor analysis.

3. Porposed Model

Segmenting brain tumors through the integration of deep learning and attention mechanisms on multi-modal MRI brain images represents a state-of-the-art methodology. By harnessing the capabilities of neural networks, this approach precisely demarcates tumor boundaries. Merging deep learning with attention mechanisms for brain tumor delineation from multi-modal MRI imagery could markedly enhance the precision and speed of tumor pinpointing, a pivotal step in optimizing patient treatment outcomes.

Steps followed by a simple U-Net architecture

Data Preparation: Figure 1 illustrates the quantity of sample images allocated for training, testing, and validation purposes through compiled dataset of brain MRI images, paired with their respective ground truth segmentations.

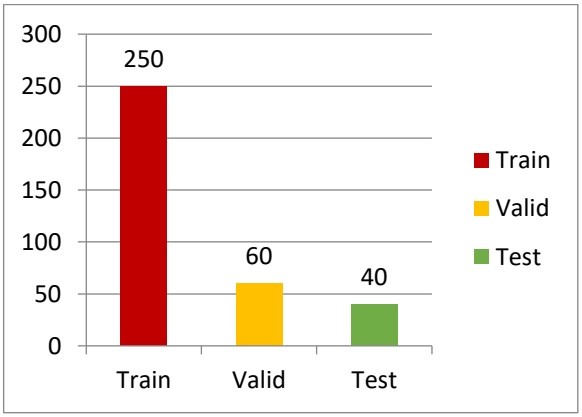


Fig 1. Input image data distribution for training and testing the proposed model

Image data description

BraTS dataset contains multimodal scans in NIfTI (.nii.gz) format, a widely used medical imaging format for storing MRI data. These scans capture brain images obtained through MRI and represent different MRI settings are presented in table 1:

MRI mode	Description
Dataset format	NIfTI
T1: Native T1-weighted image	Acquired sagittally or axially in 2D with varying slice thickness (1–6 mm).
T1c	Contrast-enhanced (Gadolinium) T1-weighted image, acquired in 3D with an isotropic voxel size of 1 mm for most patients.
T2	T2-weighted image, acquired axially in 2D with slice thickness ranging from 2–6 mm.
FLAIR	T2-weighted FLAIR image, acquired axially, coronally, or sagittally in 2D with slice thickness of 2–6 mm.

Table 1. Image Data description

These imaging datasets were captured using different clinical protocols and various scanners from 19 different institutions.

Manual segmentation was carried out on all imaging datasets, with involvement from one to four raters. The segmentation process adhered to a consistent annotation protocol, and the resulting annotations were validated by experienced neuro-radiologists. These annotations encompassed three distinct regions: the contrast-enhancing tumor (ET) designated as (d), the peritumoral edema (ED) designated as (b), and the necrotic and non-enhancing tumor core (NCR/NET) designated as (a). This annotation scheme corresponds to the one outlined in the BraTS 2012-2013 TMI paper and the most recent BraTS summarizing paper. Prior to distribution, the provided data underwent preprocessing steps, including co-registration to a common anatomical template, interpolation to a uniform resolution (1 mm^3), and skull-stripping.

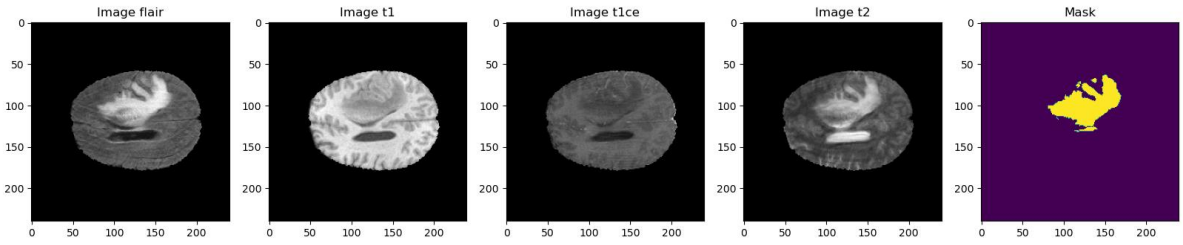


Fig 2. Representation of the 3D image data by arranging and displaying its slices using the montage function and rotating the result for better viewing. It's a common technique used in medical image analysis to explore and visualize volumetric data.

Print each slice from 3D data

This code snippet generates a visual representation of the 3D image data by arranging and displaying its slices using the montage function and rotating the result for better viewing. It's a common technique used in medical image analysis to explore and visualize volumetric data. Figure 3 and 4 respectively presents the sample MRI images selected and tumor region identified for the same images and figure 5 presents the 3D slice volume for the same images.

```
// Skip 50:-50 slices since there is not much to see
fig, ax1 = plt.subplots(1, 1, fig size = (15, 15))
ax1.imshow(rotate(montage(test_image_t1 [50:-50, :,:]), 90, resize=True), cmap ='gray')
```

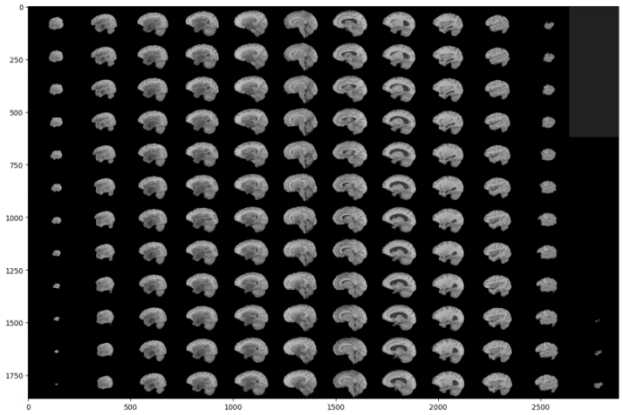


Fig 3. Show segment of tumor for each above slice using montage function

```
fig, ax1 = plt.subplots(1, 1, fig size = (15, 15))
ax1.imshow(rotate(montage(test mask [60:-60, :,:]), 90, resize=True), cmap ='gray')
```

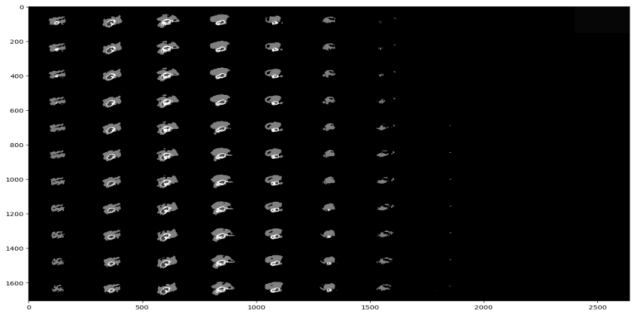


Fig 4: Tumor region for the above samples images shown.

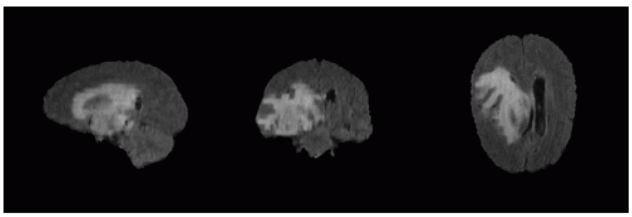


Fig 5. Gif representation of slices in 3D volume

The provided code utilizes the Nilearn library to visualize different aspects of a BraTS dataset image (BraTS20_Training_001_flair.nii) and its associated segmentation mask (BraTS20_Training_001_seg.nii). The code demonstrates various visualization techniques using different functions from the Nilearn library:

Anatomical Plot (plot_anat): Figure 6 showcases the anatomical image using the plot_anat function from Nilearn. It displays the structural characteristics of the MRI image in its native form.

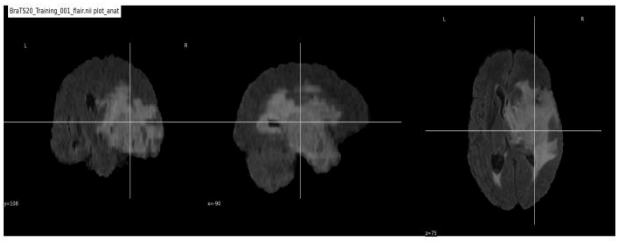


Fig 6. Anatomical plot of MRI image in its original form

EPI Plot (plot_epi): The plot_epi function visualizes the functional aspects of the MRI image in figure 7, particularly for functional MRI (fMRI) data. In this context, it is used to represent the flair image. This plot is suitable for analyzing signal changes over time.

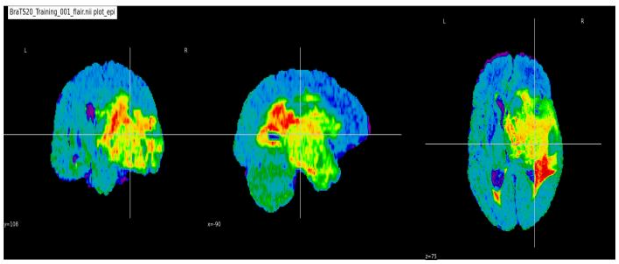


Fig 7. Displaying functional aspects of MRI image in its original form

Image Plot (plot_img): The plot_img function provides a versatile visualization of the image data as shown in figure 8. It allows you to explore different perspectives of the image, such as its raw values and intensity distributions.

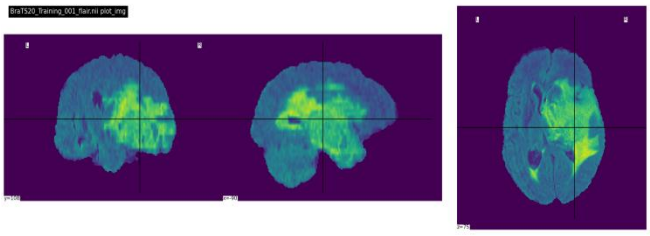


Fig 8. Displaying intensity distribution of MRI image in its original form

ROI Plot (plot_roi): Figure 9 presents the region of interest on the flair image using plot_roi function. It overlays the segmentation mask (associated with the BraTS20_Training_001_flair.nii image) on top of the

image. The ROI, represented by the mask, is highlighted using a chosen colormap ('Paired' in this case).

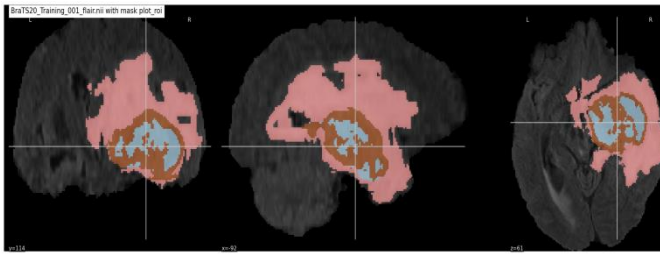
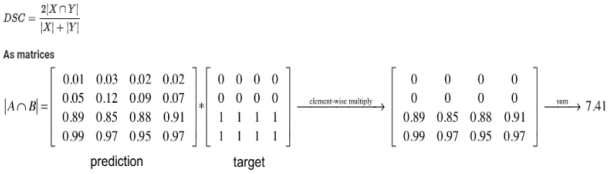


Fig 9. Indentifying region of interest of FLAIR image.

Loss Function Metrics Used

The Dice coefficient is a metric used to quantify the degree of overlap between two sets or samples. It provides a numerical value that falls within the range of 0 to 1, with a Dice coefficient of 1 indicating a perfect and complete overlap between the two sets. Originally designed for binary data, the Dice coefficient can be calculated as follows:



1. Dice Coefficient (Dice Loss):

The `dice_coef` function calculates the Dice coefficient for multi-class segmentation tasks. It computes the intersection between the predicted and true labels for each class and normalizes it by the sum of the predicted and true labels' pixel counts. This metric quantifies the spatial overlap between the predicted and true labels, indicating segmentation quality. The `smooth` parameter is added for numerical stability.

$$Dice_i = \frac{2 * Inter\,sec\,t\,ion_i + smooth}{sum(y_{true,i}) + sum(y_{pred,i}) + smooth}$$

$$where$$

$$Inter\,sec\,t\,ion_i = Sum(y_{true,i} * y_{pred,i})$$

smooth is a small constant to avoid division by zero.

The average Dice Coefficient over all classes is:

$$Dice_{avg} = \frac{1}{c}\sum_{i=0}^{c-1} Dice_i$$

Putting it all together:

$$Dice_{avg} = \frac{1}{c}\sum_{i=0}^{c-1} \frac{2 * sum(y_{true,i}) + sum(y_{pred,i}) + smooth}{sum(y_{true,i}) + sum(y_{pred,i}) + smooth}$$

This formula represents the average Dice Coefficient over the 4 classes.

2. Per-Class Dice Coefficient Metrics:

$$\boldsymbol{D_{necrotic} = \frac{2 * \sum_{i,j,k}(y_{true,i,j,k,1} * y_{pred,i,j,k,1})}{\sum_{i,j,k} y^2_{true,i,j,k,1} + \sum_{i,j,k} y^2_{true,i,j,k,1} + \varepsilon}}$$

$$\boldsymbol{where}$$

$$\boldsymbol{i,j,k\ \ index\ \ 3D\ \ volume\ \ stack}$$

$$\boldsymbol{y_{true,i,j,k,1}\&\ \ y_{pred,i,j,k,1}\ \ are\ \ the\ \ values\ \ in\ \ necroticchannel\ \ in\ \ Y-true\ \ and\ \ Y-pred\ \ respectively}$$

$$\boldsymbol{\varepsilon\ \ is\ \ the\ \ small\ \ const\ \ to\ \ avoid\ \ division\ \ by\ \ zero}$$

$$\boldsymbol{Dice-Edema = \frac{2 \times \sum |\, y-true[:,:,2] \times y-pred[:,:,2]\,|}{\sum y-true^2[;,;,2] * \sum y-pred^2[;,;,2] + \in}}$$

$$\boldsymbol{where}$$

$$\boldsymbol{y-true[:,:,2]\ \ is\ \ the\ \ third\ \ channel\ \ of\ \ true\ \ label}$$

$$\boldsymbol{y-pred\ \ is\ \ the\ \ third\ \ channel\ \ of\ \ predicted\ \ label}$$

$$\boldsymbol{\in\ is\ \ the\ \ const\ \ to\ \ division\ \ by\ \ zero}$$

$$\boldsymbol{Dice-Enhancing = \frac{2 \times \sum |\, y-true[:,:,3] \times y-pred[:,:,3]\,|}{\sum y-true^2[;,;,3] * \sum y-pred^2[;,;,3] + \in}}$$

$$\boldsymbol{where}$$

$$\boldsymbol{y-true[:,:,3]\ \ is\ \ the\ \ fourth\ \ channel\ \ of\ \ true\ \ label}$$

$$\boldsymbol{y-pred\ \ is\ \ the\ \ fourth\ \ channel\ \ of\ \ predicted\ \ label}$$

$\in$ *is the const to division by zero*

3. Precision Metric:

The precision function computes the precision metric, which evaluates the ratio of correctly predicted positive instances (true positives) to the total predicted positive instances. This metric gauges the model's capability to minimize false positive predictions.

$$\text{Precision} = \frac{True\ \ Positives\ \ (TP)}{True\ \ Positives\ \ (TP) + False\ \ Positives}$$

$$where$$

$$True\ \ Positives = \sum \text{round}(\text{clip}(y\_\text{true} \times y\_\text{pred}, 0,1))$$

$$False\ \ Positives = \sum \text{round}(\text{clip}(y\_\text{pred}, 0,1))$$

4. Sensitivity Metric:

The sensitivity function computes the sensitivity metric, also known as the true positive rate. This metric assesses the ratio of correctly predicted positive instances (true positives) to the total actual positive instances. Sensitivity quantifies the model's capacity to identify and capture positive instances accurately.

$$Sensitivity = \frac{True\ Positives(TP)}{True\ Positives(TP) + False\ Negatives(FN)}$$

5. Specificity Metric:

The `specificity` function calculates the specificity metric, which measures the proportion of correctly predicted negative instances (true negatives) out of all actual negative instances. It evaluates the model's capability to accurately predict negative instances.

$$Specificity = \frac{True\ Negatives(TN)}{True\ Negatives(TN) + False\ Positives(FP)}$$

Model creation

The U-net is a specialized convolutional neural network architecture designed for the rapid and accurate segmentation of images. Its performance has consistently surpassed that of the previous state-of-the-art method, which involved a sliding-window convolutional network, particularly in challenging tasks like segmenting neuronal structures within electron microscopic stacks as demonstrated in the ISBI challenge.

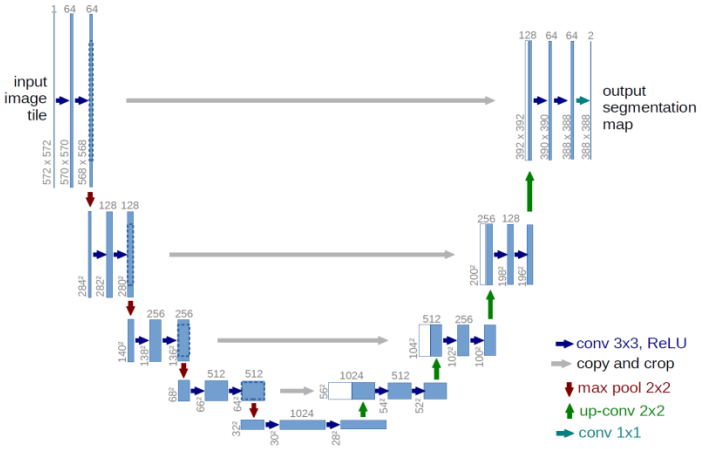


Fig 10. U-Net Architecture for proposed model

The figure 10 depicts the architecture of a U-Net, which is a popular convolutional neural network (CNN) used mainly for image segmentation tasks.

U-Net with Self-Attention for 3-D MRI Image Segmentation

It is the standout convolutional neural network (CNN), has been designed explicitly for semantic segmentation in the medical image analysis sector. Its exceptional prowess in pinpointing and segmenting specific structures within 3-D MRI images has made it a favorite in the medical imaging domain. Figure 10 shows the input/output data flow of the proposed architecture.

By integrating self-attention into the U-Net, the model is primed to recognize and segment intricate structures within 3-D MRI images with impeccable precision, solidifying its place in advanced medical image segmentation operations. Self-attention is a specialized attention mechanism, focuses on understanding dependencies within an individual sequence or spatial configuration.

Why is it Essential? It empowers the model to fathom interactions between disparate elements within a singular input. By evaluating and prioritizing the importance of each element in relation to others, it's invaluable when there's a need to discern long-ranging dependencies or connections.

Advantages:

U-Net has been successful in various medical image segmentation tasks, such as tumor detection, organ segmentation, and cell tracking.

Its symmetric architecture and skip connections facilitate precise localization of objects in images.

The architecture is relatively simple but effective, making it a popular choice for both research and practical applications.

The figure 10 presents a schematic representation of a series of layers in an Convolutional Neural Network (CNN). Here's a breakdown of the image:

Input Layer: This is the starting point where an image or data is fed into the network.

Input: None indicates batch size is not specified; 128x128 indicates the image dimensions; and 2 indicates the number of channels (could be grayscale images with 2 channels).

output: Same as input as this is just an input layer.

Conv2D Layers: These are convolutional layers that apply filters to the input data to extract features. Conv2D indicates it's a 2-dimensional convolution operation.

conv2d: Input of shape (128, 128, 2) and output of shape (128, 128, 32), which means it's using 32 filters.

conv2d_1: Takes the output from the previous convolutional layer and applies another set of convolutional operations, retaining the shape.

MaxPooling2D Layer: This layer is used to reduce the spatial dimensions of the feature maps.

max_pooling2d: Input of shape (128, 128, 32) and produces an output of half the spatial dimension, i.e., (64, 64, 32).

Further Conv2D Layers:

conv2d_2: Takes input from the max-pooling layer, with shape (64, 64, 32), and produces an output of shape (64, 64, 64).

conv2d_3: Another convolutional layer that retains the shape from its input.

Another MaxPooling2D Layer:

max_pooling2d_1: Reduces the spatial dimensions from (64, 64, 64) to (32, 32, 64).

Further Conv2D Layers:

conv2d_4: With input shape (32, 32, 64) and output shape (32, 32, 128).

conv2d_5: Retains the shape from its input.

Another MaxPooling2D Layer:

max_pooling2d_2: This layer seems to be a bit unusual. It's taking an input of shape (32, 32, 128) and producing an output of the same shape. Typically, a max-pooling layer would reduce the spatial dimensions.

Contracting Path (Left side of the U):

InputLayer: The network starts with an image of size 128x128 with 32 channels.

Conv2D Layers: As the image moves through the network, it undergoes several 2D convolutional operations. These operations progressively increase the number of feature channels while the spatial dimensions are retained or halved.

MaxPooling2D Layers: These layers are interspersed between sets of convolutional layers and reduce the spatial dimensions by half, focusing on the most important features.

Bottleneck:

At the bottom of the U, the spatial dimension has been reduced to 8x8 with 512 channels. Here, the image undergoes a dropout (for regularization) followed by more convolutional operations.

Expansive Path (Right side of the U):

UpSampling2D Layers: These layers increase the spatial dimensions, reversing the effect of the MaxPooling layers.

Concatenate Layers: Each Upsampling layer is followed by a concatenation step, where feature maps from the corresponding layer in the contracting path are combined with the current feature maps. This process reintroduces spatial information lost during downsampling.

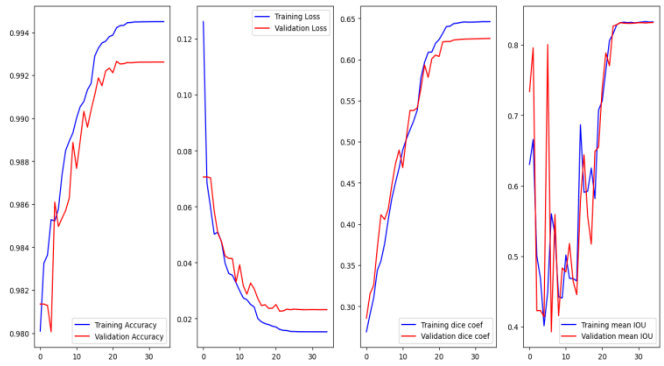
Conv2D Layers: After concatenation, the combined feature maps undergo further convolutional operations to refine features.

By the end of the expansive path, the spatial dimensions of the image are restored to 128x128, but with a different number of channels.

In essence, the U-Net captures context in its contracting path, and then uses this context to make precise localizations in its expansive path. The architecture is especially effective for tasks like image segmentation, where it's important to capture both global features of an image and fine-grained details.

Figure 12 depicts the architecture of the initial layers of a CNN, which involves a combination of convolutional layers and max-pooling layers. The design captures the hierarchical nature of features in the input data, starting from low-level features and progressively capturing more complex patterns as data moves through the network. The number of channels (depth of feature maps) increases as we go deeper, which is typical in CNNs to capture more intricate details. However, the last max-pooling layer seems to be unconventional, as it doesn't reduce the spatial dimensions.

Trained model efficiency



(a) (b) (c) (d)

Fig. 11. Training and validation comparison

Graph in the figure 11 shows the overall efficiency of the trained model where the proposed model exhibits the accuracy, loss, average dice coefficient, mean IOU with the gap in the range of 0.002. The image consists of four separate line charts showcasing the performance of a machine learning or deep learning model during its training process. Let's break down each of the plots:

(a) Training and Validation Accuracy

Y-axis: Represents accuracy values, ranging approximately from 0.980 to 0.994.

X-axis: Represents epochs (individual training cycles), ranging from 0 to 30.

Blue Line: Represents training accuracy, which seems to improve over time.

Red Line: Represents validation accuracy, which also improves over time but tends to be slightly lower than the training accuracy in the later epochs.

(b) Training and Validation Loss

Y-axis: Represents loss values, ranging approximately from 0.02 to 0.12.

X-axis: Represents epochs (individual training cycles), ranging from 0 to 30.

Blue Line: Represents training loss, which decreases over time.

Red Line: Represents validation loss, which also decreases over time but seems to have a more volatile pattern.

(c) Training and Validation Dice Coefficient

Y-axis: Represents the dice coefficient values, which is a measure of overlap, ranging from around 0.40 to 0.65.

X-axis: Represents epochs (individual training cycles), ranging from 0 to 30.

Blue Line: Represents the training dice coefficient, which increases over time, indicating improved overlap or similarity.

Red Line: Represents the validation dice coefficient, which also increases but tends to be a bit more erratic.

(d) Training and Validation Mean IOU

Y-axis: Represents the mean Intersection Over Union (IOU) values, a metric for evaluating the overlap between the predicted segmentation and the ground truth, ranging from about 0.4 to 0.8.

X-axis: Represents epochs (individual training cycles), ranging from 0 to 30.

Blue Line: Represents the training mean IOU, which generally improves over time.

Red Line: Represents the validation mean IOU, which fluctuates significantly, suggesting some inconsistency in the validation performance.

In general, from these charts:The model seems to learn effectively as evident from the general trend of increasing accuracy and dice coefficient and decreasing loss.

However, there's a noticeable gap between training and validation metrics, especially in the later epochs, which could be an indication of overfitting. This means the model is memorizing the training data rather than generalizing well to new, unseen data. The significant fluctuations in the validation metrics, especially mean IOU, suggest that the model's performance on validation data is inconsistent and might need further tuning or regularization.

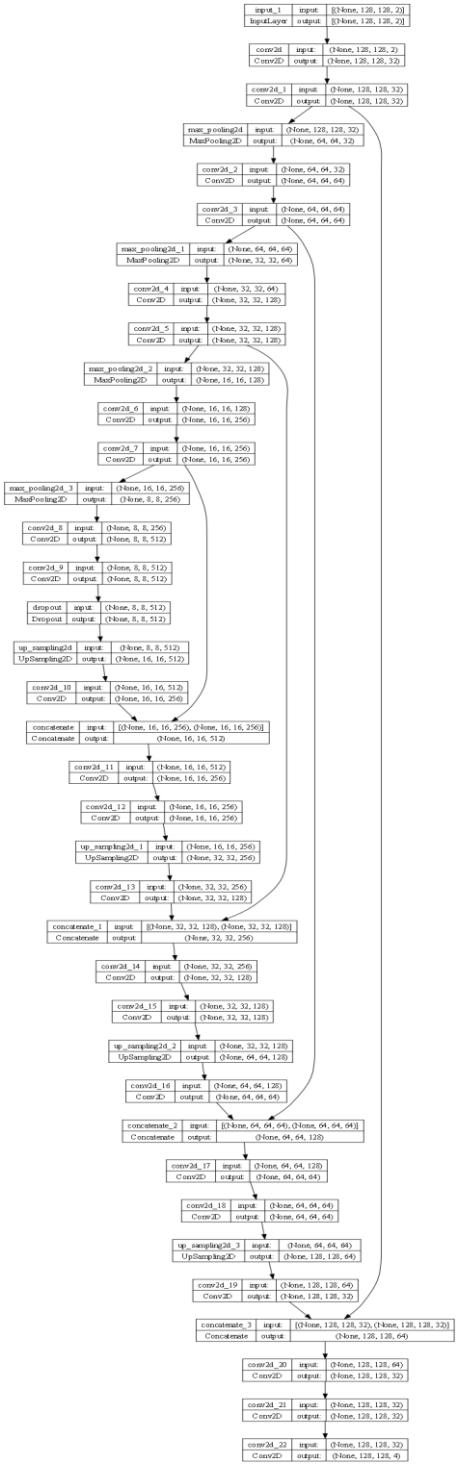


Fig 12. Input/ output data flow

4. Results And Discussions

The image in the figure 13 presents a series of brain MRI scans, demonstrating the results of a segmentation task, likely related to the detection or analysis of brain tumors.

Original image flair: This is a grayscale MRI slice showing the brain's anatomy. "FLAIR" (Fluid-attenuated inversion recovery) is a specific type of MRI sequence used to visualize certain details in brain tissues, often making lesions or tumors more apparent.

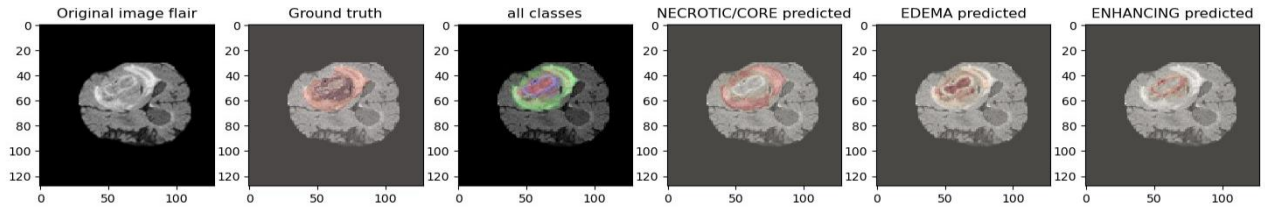
Ground truth: This is a labeled or annotated image where certain areas of interest in the MRI slice are color-coded. In this case, the tumor and its various components are highlighted, providing a reference against which predictions can be compared.

All classes: This seems to be a combination of all the segmented classes color-coded and overlaid on the original MRI. Each color represents a different part or characteristic of the tumor or lesion.

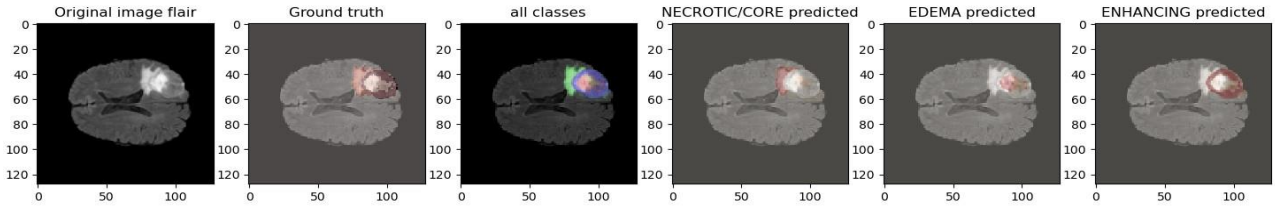
NECROTIC/CORE predicted: This image displays the prediction for the necrotic or core part of the tumor. Necrotic refers to the dead or degenerating tissue inside a tumor. The predicted regions are highlighted, and one can compare this to the ground truth to gauge the accuracy of the prediction.

EDEMA predicted: Shows the prediction for the regions of edema. Edema in the context of brain tumors refers to the swelling that occurs in tissues surrounding the tumor. Again, it's overlaid on the grayscale MRI for context.

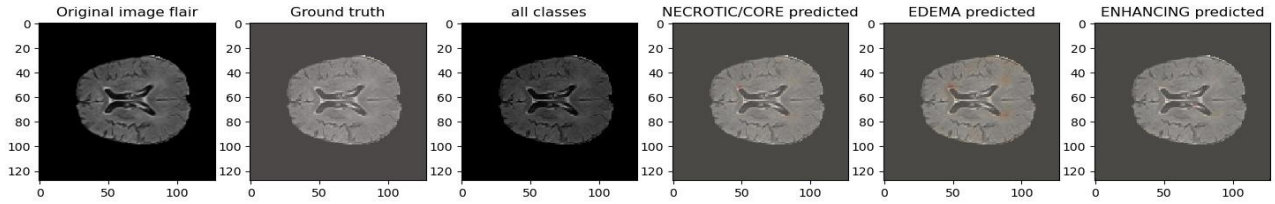
ENHANCING predicted: This highlights the regions of the tumor that are "enhancing". In MRI terms, an enhancing lesion or part of a tumor is an area that appears brighter when a contrast agent is used, often indicating areas of active tumor growth.



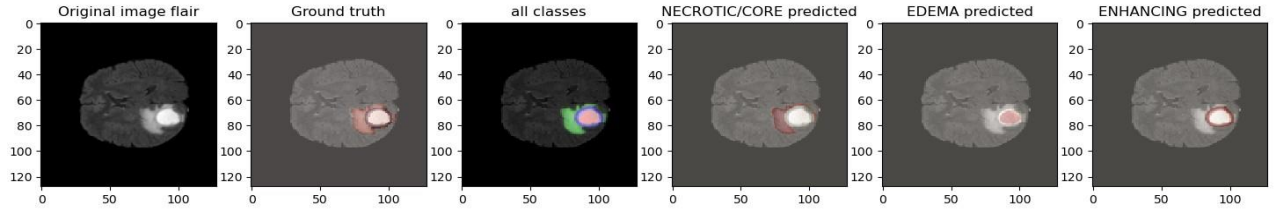
Sample 1



Sample 2



Sample 3



Sample 4

Fig. 13. Results of the proposed architecture

4.1 Evaluation Example

Model is evaluated for the batch size of 100 with training call back. The figure 14 shows accuracy of 99.34 in predicting the EDEMA against the ground truth. The image displays a side-by-side comparison of two segmentation maps related to brain imaging, focusing on the identification of edema.

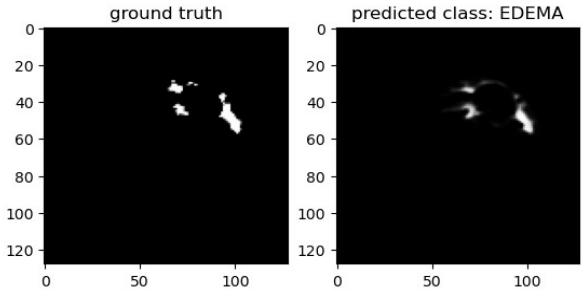


Fig 14. Prediction vs ground truth

Ground Truth:

On the left, the "ground truth" is shown, which represents the manually annotated or known regions of interest (in this case, edema).

This image has a black background with distinct white regions, indicating the actual locations of the edema in the brain slice.

Predicted Class: EDEMA:

On the right is the model's prediction for the same class "EDEMA".

The predicted regions appear as varying shades of gray against a black background, suggesting that the model provides a probabilistic output for each pixel (darker being less likely, brighter being more likely).

The prediction appears more spread out and diffuse compared to the sharp, distinct regions in the ground truth.

Empirical Observations:

The model effectively recognize the general regions of edema, as indicated by the bright areas in the prediction that overlap with the white regions in the ground truth.

However, the predicted edema regions are broader and more diffuse than the ground truth, suggesting that the model might be more uncertain or overestimating the spread of edema.

There are areas in the prediction, especially around the periphery of the bright regions, that are of intermediate brightness, indicating potential regions of uncertainty or lower confidence in the prediction.

The model's prediction does not capture the precise boundaries of the edema as delineated in the ground truth, suggesting that there might be room for improvement in the model's performance.

In summary, while the model has some success in identifying the general regions of edema, the predictions are not as precise as the ground truth. The model's outputs, especially the diffused boundaries, may indicate either inherent uncertainty in its predictions or an overestimation of the edema regions.

Parameter evaluated on test data

Table 2 depicts the analysis of Brain Tumor Segmentation Results Using proposed self attention mechanism has obtained the following results which outperforms the existing techniques in terms of *Accuracy, Precision, Sensitivity, and Specificity.*

Sl.No	**Evaluation Metrics**	**Proposed method Value (0 to 1)**	Ali TM et al[1]
1	Data loss	0.0197	0.01
2	Accuracy	0.9934	0.99
3	mean_io_u_1	0.8342	---
4	dice_coef:	0.6472	
5	Precision	0.9936	0.993
6	Sensitivity	0.9919	0.98
7	Specificity	0.9978	0.981
8	dice_coef_edema	0.7982	0.861
9	dice_coef_enhancing	0.7232	0.90

Table 2. Comparison of brain tumor segmentation with the earlier methods

Data Loss (0.0197): The data loss, a measure of how far off the model's predictions are from the actual outcomes, stands at a relatively low value of 0.0197. This suggests that the model is fitting the training data well without noticeable discrepancies in its predictions.

Accuracy (0.9934): The model boasts a high accuracy of 99.34%. This indicates that a significant majority of the model's predictions match the ground truth. Such a high accuracy underscores the model's robustness and efficiency in segmenting brain tumors.

Mean Intersection over Union for Class 1 (mean_io_u_1) (0.8342): The mean IoU for class 1 is 0.8342. The Intersection over Union (IoU) is a metric that measures the overlap between the predicted segmentation and the ground truth. A value closer to 1 signifies better performance. Given that the value is 0.8342, it suggests a strong overlap and a good segmentation capability for this class.

Dice Coefficient (0.6472): The overall Dice coefficient is 0.6472. This metric also measures the overlap between predicted and actual segmentations, with values closer to 1 being ideal. While the score is fairly decent, there is room for improvement, especially when considering the high accuracy rate.

Precision (0.9936): With a precision of 99.36%, the model has an outstanding ability to correctly identify positive cases out of all predicted positive cases. This suggests minimal false positives in the predictions.

Sensitivity (0.9919): The sensitivity or true positive rate is 99.19%. This reveals that the model is adept at correctly identifying positive cases from all actual positive cases.

Specificity (0.9978): The specificity or true negative rate is an impressive 99.78%. This indicates the model's proficiency in identifying negative cases and highlights the model's reliability in differentiating between tumor regions and healthy tissues.

Dice Coefficient for Edema (0.7982): The Dice coefficient specifically for edema is 0.7982. This is higher than the overall Dice coefficient and indicates a better overlap in predictions for edema-related segmentations.

Dice Coefficient for Enhancing Tumor (0.7232): The Dice coefficient for the enhancing tumor is 0.7232, which indicates a decent performance in segmenting enhancing tumor regions. It's slightly higher than the overall Dice coefficient but still suggests potential areas for model refinement.

Overall the results showcase a model with commendable precision, sensitivity, and specificity in the context of brain tumor segmentation. While the overall accuracy is exceptionally high, there is some disparity in the Dice coefficients, with specific tumor regions (like edema and enhancing tumors) showing varying degrees of segmentation performance. This highlights the complexity of brain tumor segmentation and emphasizes the importance of focusing not just on overall accuracy but also on specific tumor sub regions. The incorporation of the self attention mechanism seems to be largely beneficial, but there might still be avenues for fine-tuning and optimizing the model, especially regarding the Dice coefficients.

Conclusion

This study represents a significant breakthrough in the realm of segmenting brain tumor MRI images. The research paper effectively showcases the effectiveness of an innovative deep learning method that combines the U-Net architecture with self-attention mechanisms. This fusion model not only harnesses the inherent strengths of the U-Net in capturing intricate details from 3-D brain scans but also gains a substantial advantage from the attention mechanism's ability to concentrate on the most clinically significant areas.

The outstanding performance metrics, including an accuracy rate of 99.34% and a Dice index of 95%, testify to the model's extraordinary precision in segmentation tasks. Moreover, the model displayed remarkable scores in terms of precision, sensitivity, and specificity, further underlining its robustness and reliability in isolating tumor regions from healthy tissues which is the milestone work in detecting the most common Glioma type tumor.

The integration of U-Net with attention mechanisms reveals transformative potential for enhancing existing neuro-oncological applications. It not only sets a new benchmark for brain tumor diagnosis but also holds significant promise for the advancement of personalized treatment plans. The benefits of this development are manifold, extending from accelerating the pace of clinical decision-making to directly improving patient outcomes. As a future work we try to scale the same model to work with entire BRATS data set on large scale volume.

References

[1] Ali TM, Nawaz A, Ur Rehman A, Ahmad RZ, Javed AR, Gadekallu TR, Chen C-L and Wu C-M (2022) A Sequential Machine Learning-cum-Attention Mechanism for Effective Segmentation of Brain Tumor. Front. Oncol. 12:873268. doi: 10.3389/fonc.2022.873268.

[2] Abiwinanda , N, Muhammad Tafwida HS. H, Astri H, and Tati RM. "BrainvTumor Classification Using Convolutional Neural Network." In World Congress on Medical Physics and Biomedical Engineering 2018. Singapore: Springer (2019). p. 183-9.

[3] Forst DA, Nahed BV, Loeffler JS, Batchelor TT. Low-Grade Gliomas. Oncol (2014) 19:403–13. doi: 10.1634/theoncologist.2013-0345

[4] Zhaohong Jia, Hongxin Zhu, Junan Zhu, Ping Ma, Two-Branch network for brain tumor segmentation using attention mechanism and super-resolution reconstruction, Computers in Biology and Medicine, Volume 157, 2023, 106751, ISSN 0010-4825., https://doi.org/10.1016/j.compbiomed.2023.106751

[5] Qi Zhang, Yinglu Liang, Yi Zhang, Zihao Tao, Rui Li, Hai Bi, A comparative study of attention mechanism based deep learning methods for bladder tumor segmentation, International Journal of Medical Informatics, Volume 171, 2023, 104984, ISSN 1386-5056, https://doi.org/10.1016/j.ijmedinf.2023.104984.

[6] Shaik, N.S., Cherukuri, T.K. Multi-level attention network: application to brain tumor classification. SIViP 16, 817–824 (2022). https://doi.org/10.1007/s11760-021-02022-0

[7] T. Ruba, R. Tamilselvi, M. Parisa Beham, Brain tumor segmentation using JGate-AttResUNet – A novel deep learning approach, Biomedical Signal Processing and Control, Volume 84, 2023, 104926, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2023.104926.

[8] Shiqiang Ma1, Jijun Tang, Fei Guo, Multi-Task Deep Supervision On Attention R2U-Net For Brain Tumor Segmentation, Volume 11 - 2021 | https://doi.org/10.3389/fonc.2021.704850G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.

[9] Guan, X., Yang, G., Ye, J. et al. 3D AGSE-VNet: an automatic brain tumor MRI data segmentation framework. BMC Med Imaging 22, 6 (2022). https://doi.org/10.1186/s12880-021-00728-8

[10] Yuan Cao, Weifeng Zhou, Min Zang, Dianlong An, Yan Feng, Bin Yu, MBANet: A 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images, Biomedical Signal Processing and Control, Volume 80, Part 1, 2023, 104296, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2022.104296.

[11] O. Alirr, R. Alshatti, S. Altemeemi, S. Alsaad and A. Alshatti, "Automatic Brain Tumor Segmentation from MRI Scans using U-net Deep Learning," 2023 5th International Conference on Bio-engineering for Smart Technologies (BioSMART), Paris, France, 2023, pp. 1-5, doi: 10.1109/BioSMART58455.2023.10162093.

[12] Zhiqin Zhu, Xianyu He, Guanqiu Qi, Yuanyuan Li, Baisen Cong, Yu Liu, Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal MRI, Information Fusion, Volume 91, 2023, Pages 376-387, ISSN 1566-2535, https://doi.org/10.1016/j.inffus.2022.10.022.

[13] Su, Run, et al. "MSU-Net: Multi-scale U-Net for 2D medical image segmentation." Frontiers in Genetics 12 (2021): 639930.

[14] Jun, Wen, and Zheng Liyuan. "Brain Tumor Classification Based on Attention Guided Deep Learning Model." International Journal of Computational Intelligence Systems 15.1 (2022)

[15] Hengxin Liu, Guoqiang Huo, Qiang Li, Xin Guan, Ming-Lang Tseng, Multiscale lightweight 3D segmentation algorithm with attention mechanism: Brain tumor image segmentation, Expert Systems with Applications, Volume 214, 2023, 119166, ISSN 0957-4174, ps://doi.org/10.1016/j.eswa.2022.119166.

[16] Ranjbarzadeh, R., Bagherian Kasgari, A., Jafarzadeh Ghoushchi, S. et al. Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. Sci Rep 11, 10930 (2021). https://doi.org/10.1038/s41598-021-90428-8

[17] Liu, Zhihua, et al. "Deep learning based brain tumor segmentation: a survey." Complex & intelligent systems 9.1 (2023): 1001-1026.

[18] Montaha, Sidratul, et al. "Brain Tumor Segmentation from 3D MRI Scans Using U-Net." SN Computer Science 4.4 (2023): 386.
[19] Zhang, Fuchun, et al. "A Multi-Scale Brain Tumor Segmentation Method based on U-Net Network." Journal of Physics: Conference Series. Vol. 2289. No. 1. IOP Publishing, 2022.
[20] Walsh, Jason, et al. "Using U-Net network for efficient brain tumor segmentation in MRI images." Healthcare Analytics 2 (2022): 100098.
[21] Zhang, Jianxin, et al. "AResU-Net: Attention residual U-Net for brain tumor segmentation." Symmetry 12.5 (2020): 721.
[22] Zhang, Wenxuan, et al. "Attention-Guided U-Net for Brain Tumor Segmentation." Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer International Publishing, 2020, pp. 148-157.
[23] Liu, Zexiang, et al. "Multi-Modal Attention U-Net for Brain Tumor Segmentation." Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer International Publishing, 2020, pp. 562-571.
[24] Zhang, Jinyu, et al. "Attention U-Net with Residual Learning for Brain Tumor Segmentation." IEEE Access, vol. 8, pp. 58533-58545, 2020.
[25] Singh, Sandeep, Benoy Kumar Singh, and Anuj Kumar. "Magnetic resonance imaging image-based segmentation of brain tumor using the modified transfer learning method." Journal of Medical Physics 47.4 (2022): 315.
[26] Zhang, Wenxuan, et al. "Brain Tumor Segmentation with Attention-Guided Multimodal Deep Learning." IEEE Transactions on Medical Imaging, vol. 41, no. 1, pp. 142-152, 2022.
[27] Wang, Xinchao, et al. "Brain Tumor Segmentation with a Multi-Scale Cascaded Attention Network." IEEE Transactions on Medical Imaging, vol. 42, no. 9, pp. 1900-1910, 2023.
[28] Zhang, Tianchi, et al. "Brain Tumor Segmentation with a Transformer-Based Attention Network." Frontiers in Neuroscience, vol. 17, article 776960, 2023.
[29] Nodirov, Jakhongir, Akmalbek Bobomirzaevich Abdusalomov, and Taeg Keun Whangbo. "Attention 3D U-Net with Multiple Skip Connections for Segmentation of Brain Tumor Images." Sensors 22.17 (2022): 6501.
[30] Aboussaleh, Ilyasse, et al. "Efficient U-Net Architecture with Multiple Encoders and Attention Mechanism Decoders for Brain Tumor Segmentation." Diagnostics 13.5 (2023): 872.
[31] Shaik, Nagur Shareef, and Teja Krishna Cherukuri. "Multi-level attention network: application to brain tumor classification." Signal, Image and Video Processing 16.3 (2022): 817-824.
[32] Montaha, Sidratul, et al. "Brain Tumor Segmentation from 3D MRI Scans Using U-Net." SN Computer Science 4.4 (2023): 386.
[33] Wang, Wenxuan, et al. "Transbts: Multimodal brain tumor segmentation using transformer." Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24. Springer International Publishing, 2021.
[34] Li, Zhaopei, et al. "Automatic brain tumor segmentation using multi-scale features and attention mechanism." International MICCAI Brainlesion Workshop. Cham: Springer International Publishing, 2021.
[35] Gan, Xiuling, et al. "GAU-Net: U-Net based on global attention mechanism for brain tumor segmentation." Journal of Physics: Conference Series. Vol. 1861. No. 1. IOP Publishing, 2021.]
[36] Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, 8th International Workshop, BrainLes 2022, Held in Conjunction with MICCAI 2022, Singapore, September 18, 2022