INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY - BANGALORE

Medical Image Analysis (AIM-841)

Course Project Report - T1-24-25

on

Magnetic Resonance Image for Brain Tumor Detection, Classification and Segmentation

Submitted by:

Monish K - MS2024014

Garvit Singla - MS2024009

Anwesh Nayak - MS2024003

Submitted to:

Asst. Prof. Sushree S. Behera

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible, whose constant guidance and encouragement crowned our effort with success.

We wish to thank our Asst. Prof.**, Dr. Sushree Sangeeta Behera** for the support and encouragement for the presentation. We also thank her for the invaluable guidance provided which has helped in the creation of a better technical report.

Name & Roll No.:

Monish K - MS2024014

Garvit Singla - MS2024009

Anwesh Nayak - MS2024003

Date : 16th Dec 2024

ABSTRACT

Medical image analysis plays a critical role in advancing diagnostic and therapeutic approaches for neurological conditions. In this study, we address the challenges of brain tumor detection, classification, and segmentation using MRI scans from the BraTS20 dataset. Our project explores a comprehensive workflow that evaluates the performance of three distinct methodologies: handcrafted features, deep Convolutional Neural Networks (CNNs), and Transformer-based architectures.

For detection, we extract spatial and intensity-based handcrafted features, employing classical machine learning models to identify tumor presence. In parallel, CNNs are utilized for their ability to learn hierarchical spatial features directly from raw MRI data, while Transformers leverage their self-attention mechanisms to capture global dependencies across the image.

For classification and segmentation, we compare the effectiveness of these approaches in accurately identifying tumor types and delineating tumor regions. Quantitative metrics such as accuracy, Dice similarity coefficient, and F1-score are used to evaluate and benchmark the performance of each method. Our results demonstrate the strengths and limitations of each technique, highlighting scenarios where traditional methods outperform or complement modern deep learning models.

This comparative analysis provides insights into the suitability of different methodologies for brain tumor analysis, paving the way for improved diagnostic tools and personalized treatment strategies.

TABLE OF CONTENTS

1. INTRODUCTION

Brain tumors represent a significant medical challenge, with their accurate diagnosis and treatment heavily dependent on advancements in medical imaging technologies. Magnetic Resonance Imaging (MRI) has emerged as a cornerstone in neuroimaging due to its unparalleled ability to provide detailed soft-tissue contrast and high-resolution images of the brain.

Despite its diagnostic advantages, interpreting MRI data for brain tumor analysis remains a complex and labor-intensive task. Challenges such as the heterogeneity of tumor appearance, variability in patient anatomy, and the need for precise tumor boundary delineation necessitate the use of advanced computational methods. In recent years, medical image analysis has witnessed a paradigm shift with the advent of machine learning and deep learning techniques. These approaches have shown immense potential in automating and improving tasks such as tumor detection, classification, and segmentation.

1. Magnetic Resonance Imaging (MRI) and Its Role in Medical Imaging:

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique that has revolutionized the field of medical diagnostics. Developed in the early 1970s, MRI capitalizes on the magnetic properties of atomic nuclei, particularly hydrogen atoms abundant in water and fat molecules within the human body. The imaging process begins by placing the patient within a strong magnetic field, which aligns the hydrogen nuclei. When radiofrequency pulses are applied, these nuclei absorb energy and shift to a higher energy state. Upon cessation of the radiofrequency signal, the nuclei release this energy as they return to their equilibrium state, emitting signals that are detected and processed to form detailed cross-sectional images of the body.

MRI's primary strength lies in its ability to differentiate between various tissue types based on their water content and molecular composition, producing high-resolution images with excellent contrast. This makes MRI particularly well-suited for imaging soft tissues, such as the brain, where it is used extensively for diagnosing and monitoring neurological disorders, including brain tumors. By providing multi-modal imaging sequences (e.g., T1, T2, FLAIR), MRI offers complementary information about tissue structure, fluid content, and pathological changes, which is critical for comprehensive analysis.

2. The BraTS20 Dataset:

The Brain Tumor Segmentation (BraTS) 2020 dataset is a widely used benchmark in medical image analysis for brain tumor research. This dataset comprises preoperative MRI scans from patients diagnosed with gliomas, the most common and aggressive type of brain tumors. Gliomas are categorized into low-grade (Grade II) and high-grade (Grade III and IV) tumors, with high-grade glioblastomas (GBMs) being particularly challenging due to their rapid growth and invasive nature.

The BraTS20 dataset includes four distinct MRI modalities for each patient:

1. **T1-weighted (T1)**: Provides detailed anatomical information.
2. **T1-contrast enhanced (T1CE)**: Highlights enhancing tumor regions with contrast agents.
3. **T2-weighted (T2)**: Offers insights into fluid-filled regions and tumor edema.
4. **FLAIR (Fluid-Attenuated Inversion Recovery)**: Suppresses signals from cerebrospinal fluid to better visualize lesions.

Additionally, the dataset includes expert-annotated segmentation masks that delineate three key tumor regions:

1. **Enhancing Tumor Core (ET)**: Active and proliferative tumor regions.
2. **Tumor Core (TC)**: Includes necrotic and non-enhancing tumor core.
3. **Whole Tumor (WT)**: Combines the enhancing core, necrotic core, and peritumoral edema.

BraTS20 provides a standardized dataset for developing and benchmarking models for tumor detection, classification, and segmentation. Its extensive multi-modal data and expert annotations enable researchers to evaluate and compare the effectiveness of different computational approaches in a controlled setting.

2. BACKGROUND

Brain tumor analysis is a critical task in medical imaging, necessitating accurate detection, classification, and segmentation of tumors to guide clinical decision-making. Traditionally, this process has relied heavily on expert radiologists to interpret MRI scans, a labor-intensive and time-consuming approach prone to inter- and intra-observer variability. To address these challenges, computational methods have emerged as essential tools, leveraging both classical machine learning and modern deep learning techniques for automated tumor analysis.

1. Handcrafted Feature-Based Approaches

Handcrafted feature extraction represents one of the earliest methods in medical image analysis. This approach relies on extracting predefined features from MRI images, such as texture, intensity, shape, and gradient-based descriptors, to characterize tumor regions. While these methods are computationally efficient and interpretable, they are often limited by their reliance on domain expertise for feature selection and their inability to generalize across diverse datasets.

In this project, we implement handcrafted feature-based models as a baseline, focusing on spatial, intensity, and texture-related features. These features are computed using conventional image processing techniques, and their performance is benchmarked against modern deep learning approaches.

2. Deep Learning with Convolutional Neural Networks (CNNs)

Deep Convolutional Neural Networks (CNNs) have revolutionized the field of medical image analysis by automating the feature extraction process. CNNs learn hierarchical features directly from raw image data, capturing both low-level patterns (e.g., edges and textures) and high-level semantic information (e.g., tumor boundaries). Their ability to generalize across datasets and their superior performance in image classification and segmentation tasks have made them a popular choice for brain tumor analysis.

In this project, CNN architectures are used for tumor detection, classification, and segmentation tasks. Pretrained models such as VGG-16, ResNet and U-Net are fine-tuned using transfer

learning to adapt them to the BraTS20 dataset. The U-Net architecture, in particular, is well-suited for segmentation tasks due to its encoder-decoder structure, which preserves spatial information while learning feature representations.

3. Transformer-Based Architectures

Transformers, originally developed for natural language processing, have recently gained traction in computer vision tasks. By utilizing self-attention mechanisms, Transformers can capture global dependencies within an image, making them particularly effective for medical imaging, where context across large spatial regions is critical. Vision Transformers (ViTs) and Swin transformer are among the latest advancements in this domain, offering a complementary perspective to CNNs.

In this project, we explore the use of Vision Transformers, Swin-Untr for brain tumor classification and segmentation comparing their performance to CNNs and handcrafted feature-based models. The ability of Transformers to process entire image patches simultaneously enables them to capture intricate patterns and contextual information that may be missed by CNNs or traditional methods.

4. Transfer Learning and Pretraining

Transfer learning is employed in this study to leverage pretrained models on large-scale datasets such as ImageNet. Fine-tuning these models on the BraTS20 dataset significantly reduces the computational resources and training time required while improving performance on limited medical imaging datasets. Transfer learning is particularly valuable for CNNs and Transformers, as it enables the reuse of learned features that generalize well to medical image analysis tasks.

5. Evaluation and Comparison

To comprehensively evaluate the performance of these methodologies, we use standard metrics such as accuracy, Dice similarity coefficient (DSC), precision, recall, F1-score, IoU. These metrics allow us to assess the strengths and limitations of each approach across tasks such as tumor detection, classification into high and low-grade gliomas, and precise segmentation of tumor regions.

3. OBJECTIVES

The primary goal of this project is to explore and compare the performance of three different computational methodologies for brain tumor detection, classification, and segmentation using the BraTS20 dataset. These methodologies include:

1. **Handcrafted Feature-Based Models**: Traditional approaches leveraging predefined features such as texture, intensity, and shape to detect and classify tumors.
2. **Deep Convolutional Neural Networks (CNNs)**: Modern deep learning models that automatically learn hierarchical features from raw MRI data for accurate tumor analysis.
3. **Transfer Learning(CNNs based):** Use of Pre-existing classification, segmentation models to be applied on Brats20 dataset as they work better in data scarcity.
4. **Transformer Architectures**: Emerging deep learning models that utilize self-attention mechanisms to capture global relationships within MRI data.

By systematically comparing these approaches, the project aims to identify the strengths, weaknesses, and potential synergies between classical machine learning and state-of-the-art deep learning methods. The outcomes of this study will contribute to advancing automated tools for brain tumor diagnosis and treatment planning.

4. ANALYSIS

1. Detection

Tumor detection is the process of identifying whether a brain tumor exists in the MRI scan. This task is foundational as it marks the first step in the medical image analysis pipeline. We used both handcrafted features and deep learning models to approach the tumor detection task. The performance of these techniques was then evaluated using precision, recall, and F1-score to assess their ability to detect the presence of a tumor accurately.

1.1 Handcrafted Features for Tumor Detection

Handcrafted features are based on domain-specific knowledge and image statistics, manually derived from MRI scans. We used the following feature sets for tumor detection:

1. **Histogram of Oriented Gradients (HoG):** HoG features capture gradient-based edge information, helping in the detection of tumor boundaries. This method works well for detecting regions with sharp contrast, such as the boundaries of a tumor.
2. **Gray-Level Co-Occurrence Matrix (GLCM):** GLCM-based features focus on the texture of tumor regions, helping to distinguish different types of tissue based on intensity relationships.
3. **Local Binary Patterns (LBP):** LBP features describe the texture of the region surrounding the tumor, detecting subtle patterns and variations in intensity, which is often indicative of the presence of a tumor.

These features were then fed into machine learning classifiers such as Support Vector Machines (SVM) and Random Forests, which were trained to classify regions as either tumor or non-tumor.

1.2 Deep Learning for Tumor Detection

Deep learning models excel in automated feature extraction. We implemented several deep learning models for tumor detection:

1. **VGG-16 and ResNet-50 (Transfer Learning):** Both pre-trained models were fine-tuned on the BraTS20 dataset. Their layers were adapted to learn tumor-specific features from MRI scans, leveraging hierarchical feature representations.
2. **Vision Transformer (ViT):** ViTs use self-attention mechanisms to capture long-range dependencies within the image. This is particularly useful in detecting tumors in cases where their location and appearance might vary across different regions of the brain.

1.3 Performance Evaluation for Tumor Detection

The performance of both handcrafted and deep learning models was evaluated using:

1. **Precision:** Measures the proportion of true positive tumor detections out of all positive predictions.
2. **Recall:** Measures the proportion of true positive tumor detections out of all actual tumor instances in the MRI scan.
3. **F1-Score:** The harmonic mean of precision and recall, providing a balanced measure of model performance. A higher F1-score indicates better overall detection performance.

By evaluating these metrics, we assessed how well each model detects the presence of a tumor in MRI scans.

2. Classification

Classification refers to categorizing detected tumors into different types, such as high-grade or low-grade gliomas, which are crucial for treatment decisions. In this section, we compared traditional handcrafted features and deep learning models for tumor classification. The performance of each method was evaluated using precision, recall, and F1-score.

2.1 Handcrafted Features for Tumor Classification

Handcrafted features capture critical characteristics that distinguish different tumor grades. We used the following features:

1. **HoG:** Helps classify tumors based on the texture and boundary details, with high-grade tumors typically having irregular boundaries.

2. **GLCM:** Characterizes texture features, where high-grade gliomas are expected to show higher contrast and more inhomogeneous textures than low-grade gliomas.
3. **LBP:** Local texture patterns are used to identify fine variations between tumor types. High-grade tumors often exhibit more heterogeneous texture patterns than low-grade tumors.

These handcrafted features were used to train classification algorithms such as **Support Vector Machines (SVM)**, **Random Forests**, and **k-Nearest Neighbors (k-NN)**. The models were trained to distinguish between high-grade gliomas (HGG) and low-grade gliomas (LGG) based on the extracted features.

2.2 Deep Learning for Tumor Classification

Deep learning models, particularly CNNs and Transformers, were used to automatically extract features and classify tumors.

1. **VGG-16 and ResNet-50:** Fine-tuned for classification tasks, both models learn hierarchical features to classify the tumors as either high-grade or low-grade gliomas.
2. **3D U-Net:** Although primarily a segmentation model, we adapted 3D U-Net for classification by aggregating features from the entire 3D MRI scan. This allowed the model to capture the global context of tumor location and structure.
3. **Vision Transformer (ViT):** ViT's self-attention mechanism allows it to capture long-range dependencies, improving its ability to distinguish subtle differences between high- and low-grade tumors.

2.3 Performance Evaluation for Tumor Classification

The performance of the classification models was evaluated using the following metrics:

1. **Precision:** Proportion of true positives (correctly identified tumors) out of all predicted positive tumor classifications.
2. **Recall:** Proportion of actual positive cases (real tumors) correctly identified by the model.
3. **F1-Score:** The harmonic mean of precision and recall, which balances both false

positives and false negatives. A high F1-score indicates that the model is both accurate and complete in its tumor classification.

By analyzing these metrics, we assessed the ability of each model to classify tumors accurately, and understand the trade-offs between false positives and false negatives in the classification task.

3. Segmentation

Segmentation is the process of delineating the boundaries of the tumor from the surrounding brain tissue. Accurate segmentation is essential for precise treatment planning and monitoring tumor growth. In this section, we explore both traditional image processing methods and deep learning techniques for tumor segmentation. The performance of these techniques was evaluated using the Intersection over Union (IoU) metric, a standard measure for evaluating segmentation accuracy.

3.1 Traditional Segmentation Techniques

Traditional segmentation techniques are simple and fast but are often less accurate when dealing with complex structures like brain tumors. We applied several traditional methods for segmentation:

1. **Region Growing:** A seed-based technique that expands regions based on pixel similarity. While it can work well for homogeneous regions, it is sensitive to noise and often struggles with heterogeneous tumor types.
2. **Otsu Thresholding:** A global thresholding technique that separates foreground (tumor) and background based on pixel intensity distributions. This method works well when the tumor has a significantly different intensity from the background but may fail when intensity differences are subtle.
3. **Canny Edge Detection:** Edge detection is used to identify tumor boundaries. This method detects areas of high intensity change, which are typically associated with the boundary of a tumor. However, it requires post-processing to improve the segmentation quality.
4. **Clustering for Segmentation:** Clustering is a type of unsupervised classification that

groups pixels based on their distance from different cluster centers. The resulting clusters provide a concise summary of the image content.

3.2 Deep Learning for Tumor Segmentation

Deep learning methods have shown great promise in overcoming the limitations of traditional segmentation methods, especially when dealing with complex or irregularly shaped tumors. We implemented the following deep learning-based segmentation methods:

1. **3D U-Net:** A deep network designed for segmenting volumetric (3D) data, which is ideal for MRI scans. The 3D U-Net learns both local and global features, preserving the spatial context of the tumor in 3D space.
2. **Swin-UNETR:** A hybrid model that combines Swin Transformers for global feature extraction with U-Net's encoder-decoder structure for segmentation. The Swin Transformer's attention mechanism is particularly effective for capturing long-range dependencies, which is beneficial for segmenting complex and irregular tumors.

3.3 Performance Evaluation for Tumor Segmentation

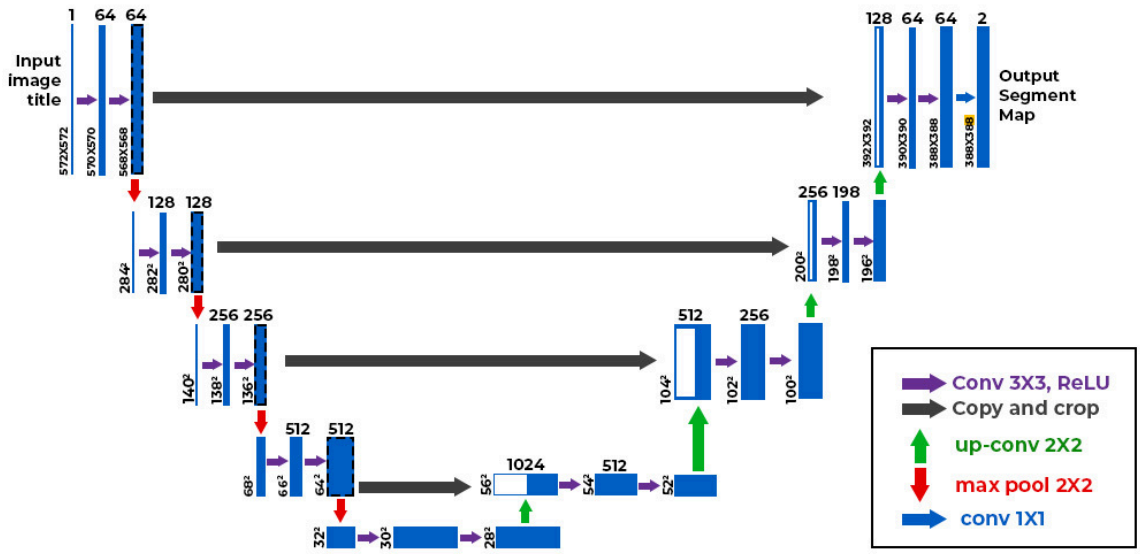
For segmentation, we used **Intersection over Union (IoU)** to evaluate the performance of the models. The IoU measures the overlap between the predicted segmentation mask and the ground truth mask, providing an indication of how accurately the model delineates tumor boundaries. Higher IoU values indicate better segmentation performance.

In addition to IoU, we also considered:

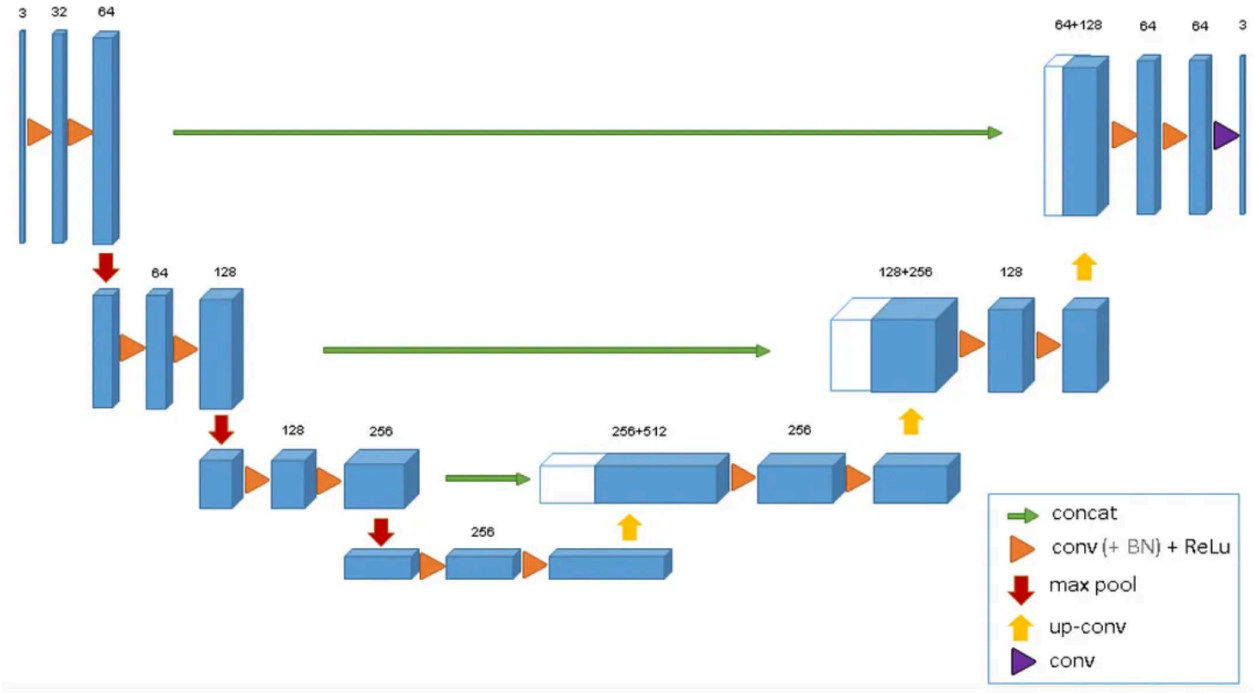
1. **Dice Similarity Coefficient (DSC):** Measures the overlap between the predicted and ground truth tumor masks.
2. **Focal Loss:** Measures the maximal distance between predicted and ground truth, providing an indication of how well the model follows the tumor region.

5. ARCHITECTURES

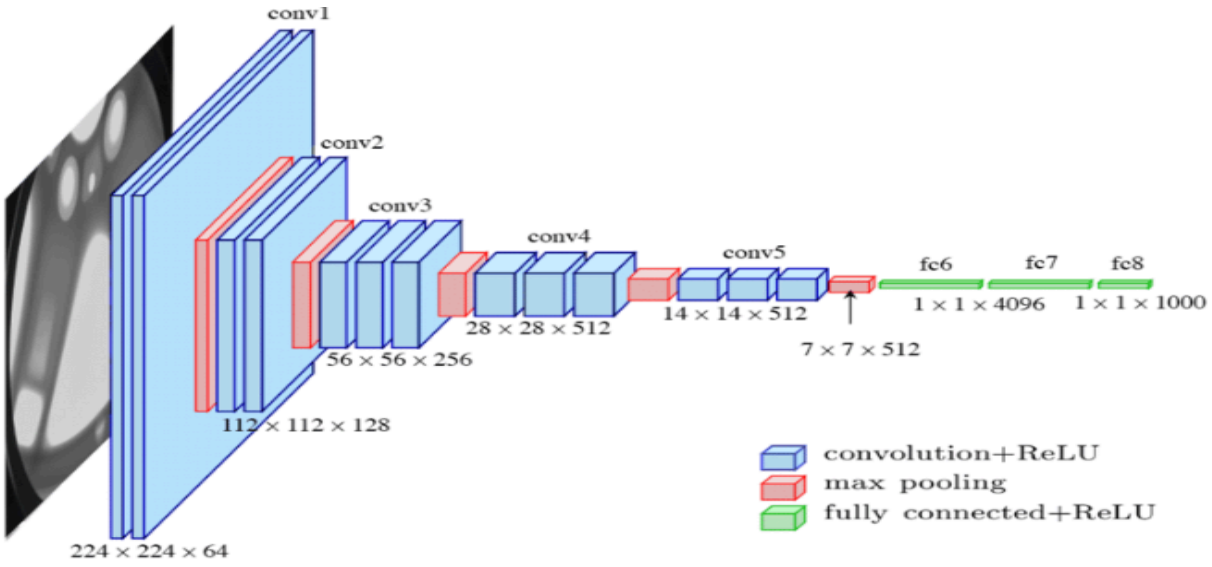
2D-Unet:



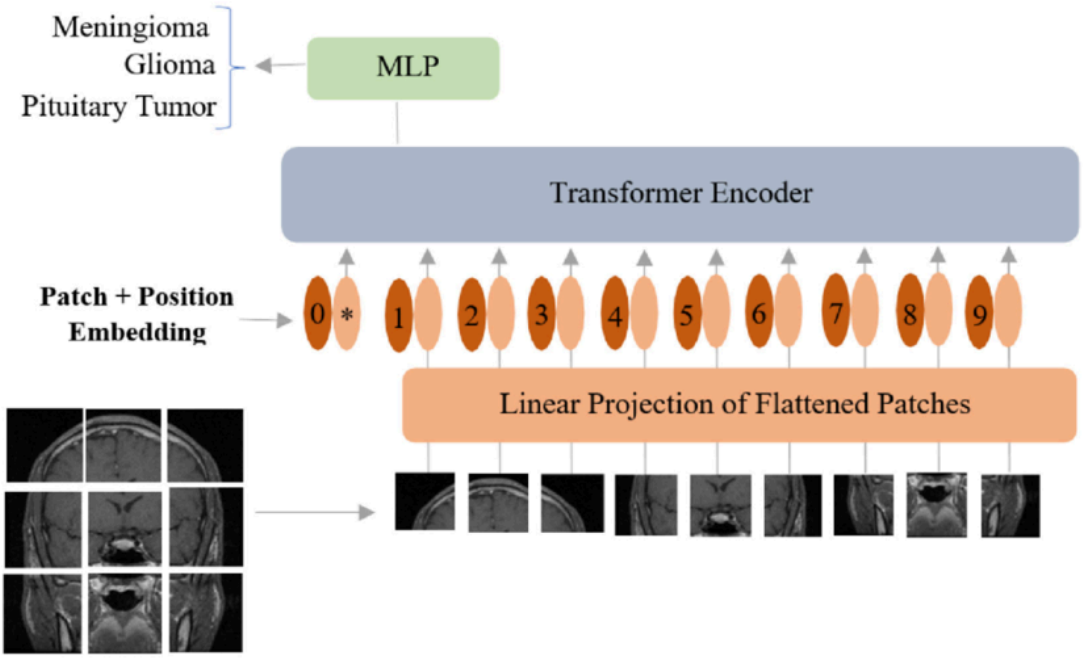
3d-Unet:



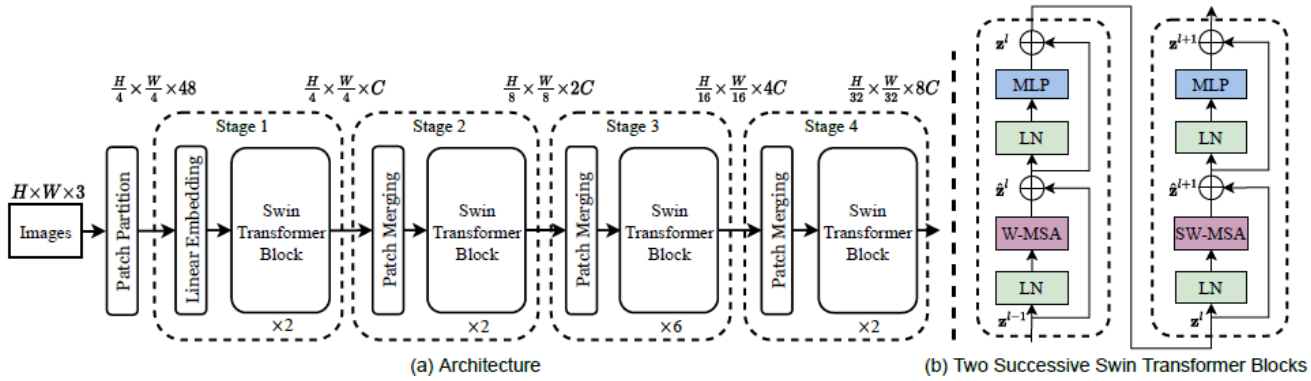
VGG-16:


conv1
conv2
conv3
conv4
conv5
fc6
fc7
fc8
224 × 224 × 64
112 × 112 × 128
56 × 56 × 256
28 × 28 × 512
14 × 14 × 512
7 × 7 × 512
1 × 1 × 4096
1 × 1 × 1000
convolution+ReLU
max pooling
fully connected+ReLU

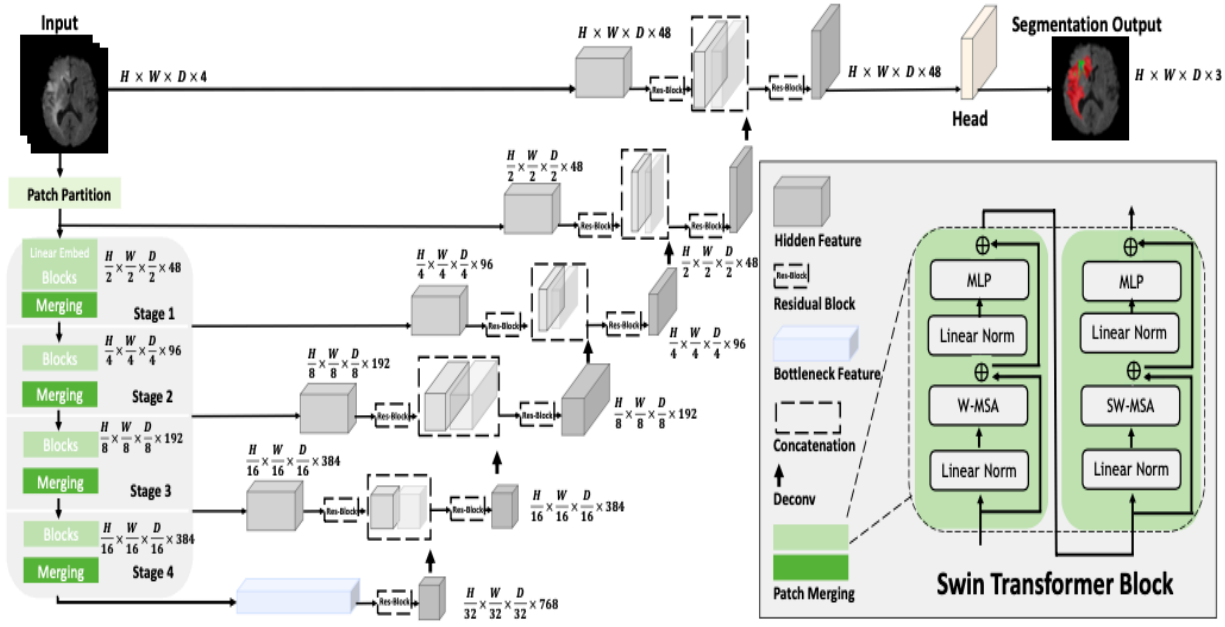
ViT:


Meningioma
Glioma
Pituitary Tumor
MLP
Transformer Encoder
Patch + Position Embedding
0
*
1
2
3
4
5
6
7
8
9
Linear Projection of Flattened Patches

Swin Transformer:



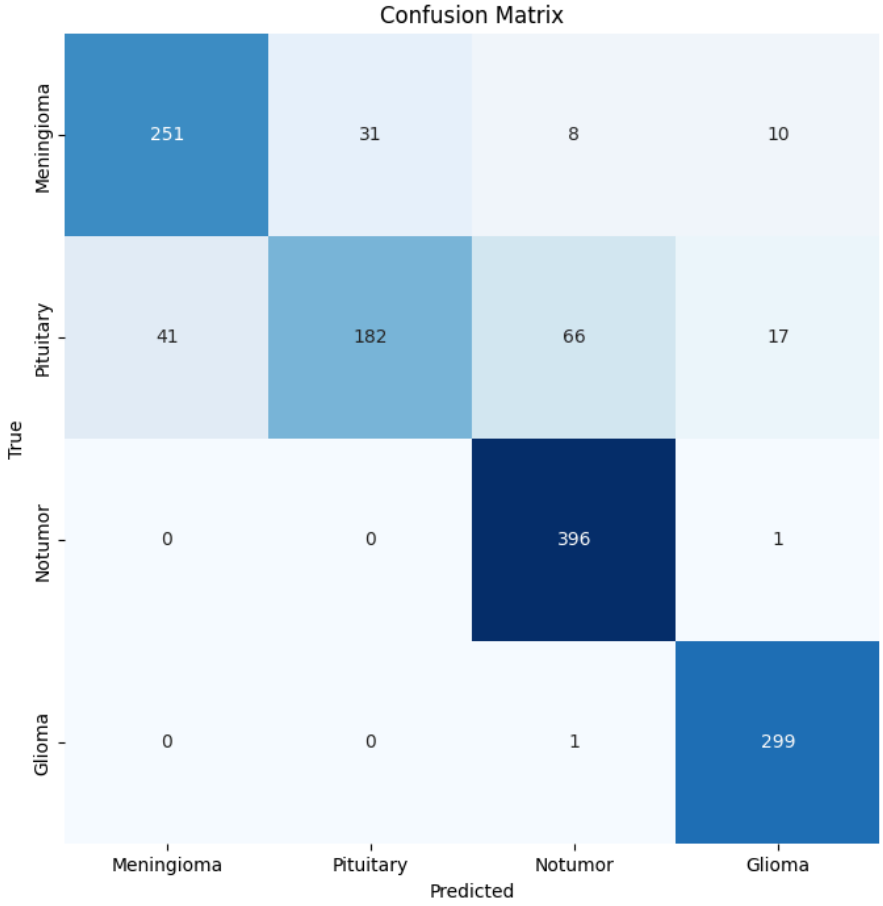
Swin Unetr:



6. RESULTS

Hand-Crafted feature based detection and classification:

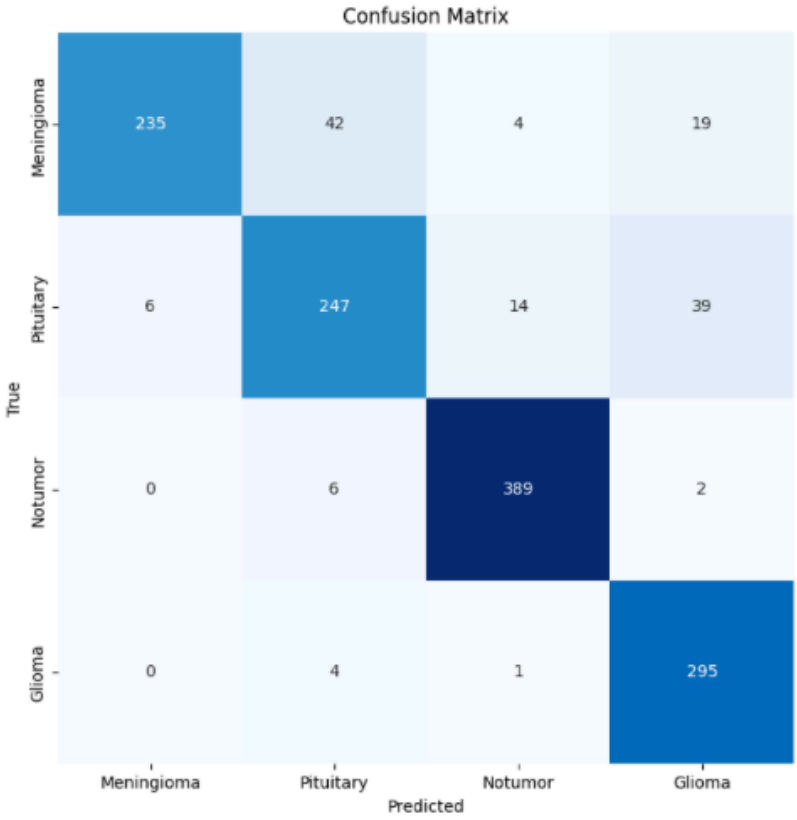
	precision	recall	f1-score
meningioma	0.86	0.84	0.85
pituitary	0.85	0.59	0.70
notumor	0.84	1.00	0.91
glioma	0.91	1.00	0.95



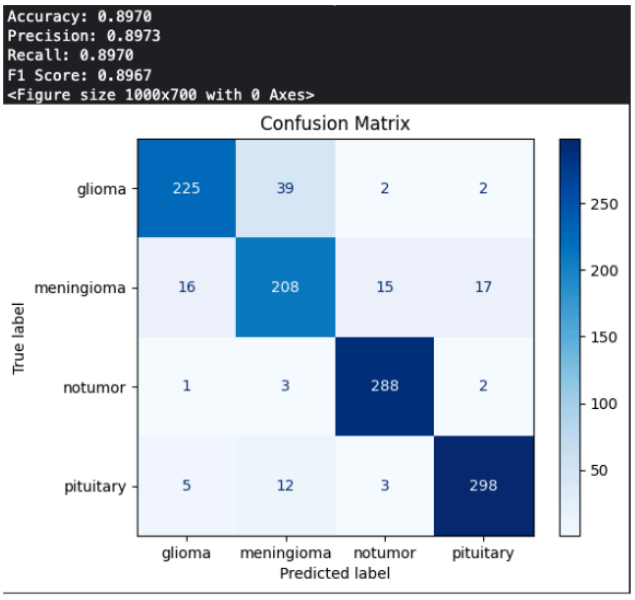
VGG-16 based classification:

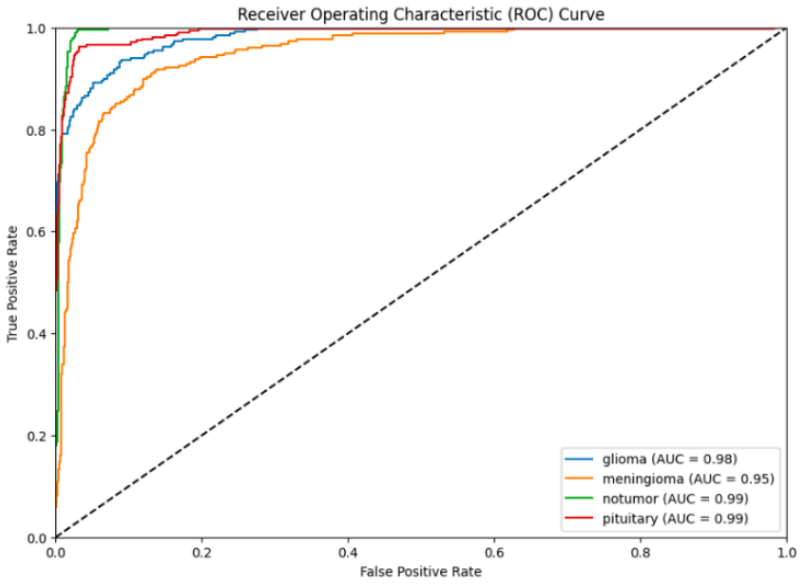
- Test Loss: 0.28074
- Test Accuracy: 0.89486

	precision	recall	f1-score
meningioma	0.98	0.78	0.87
pituitary	0.83	0.81	0.82
notumor	0.95	0.98	0.97
glioma	0.83	0.98	0.90

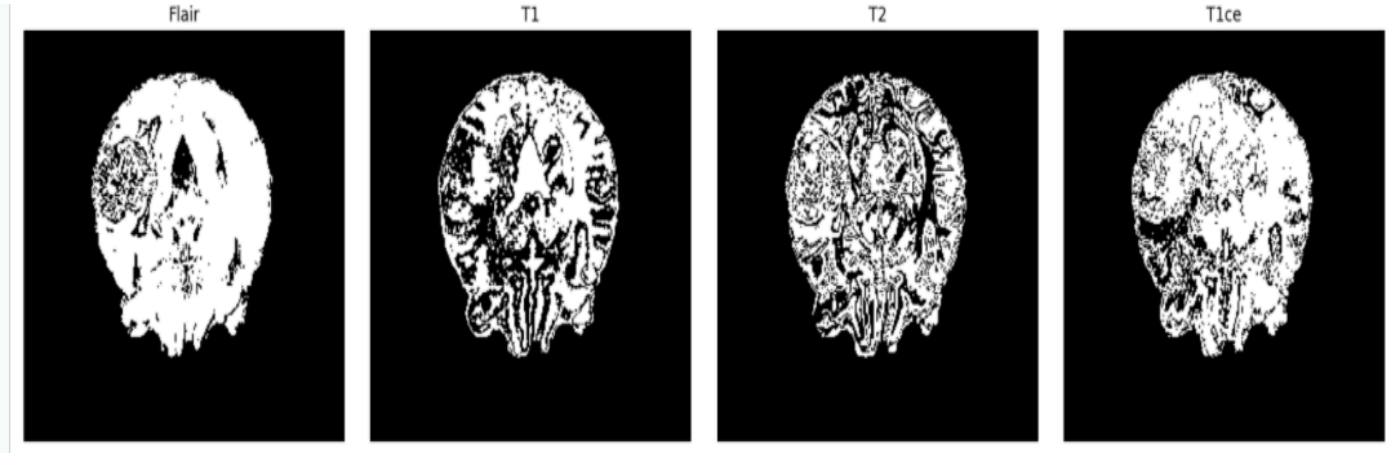


ViT:

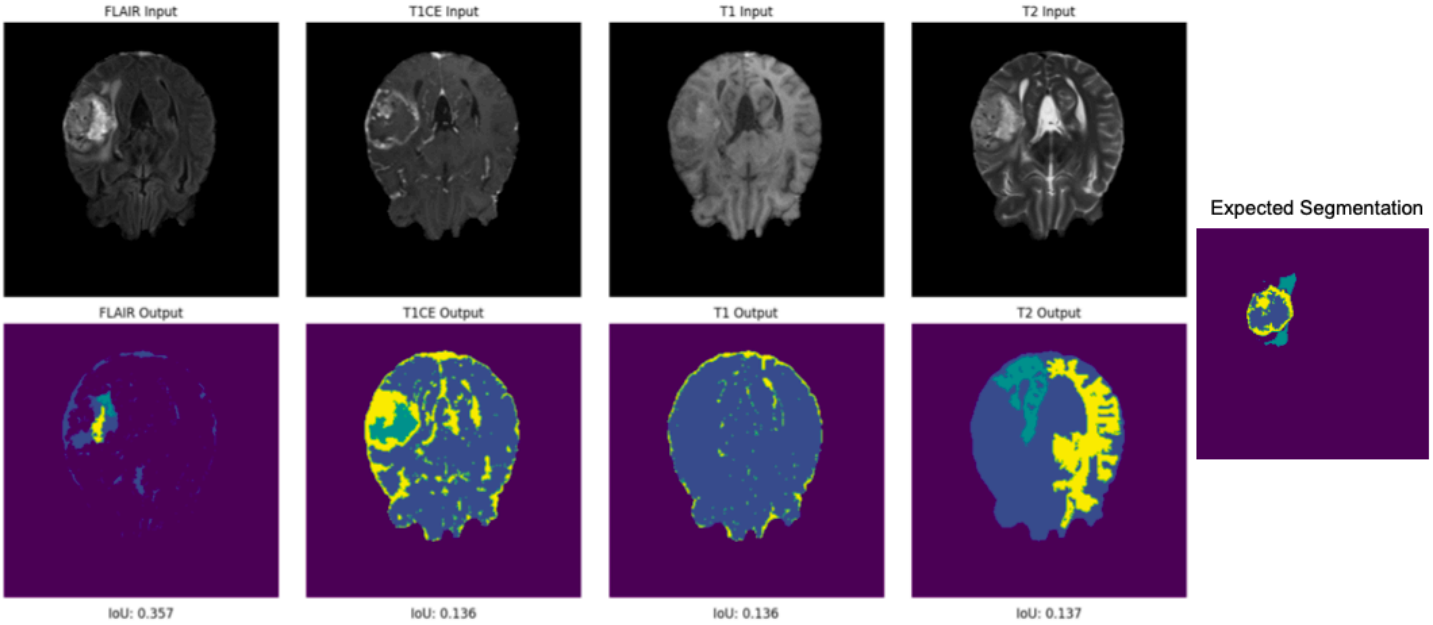




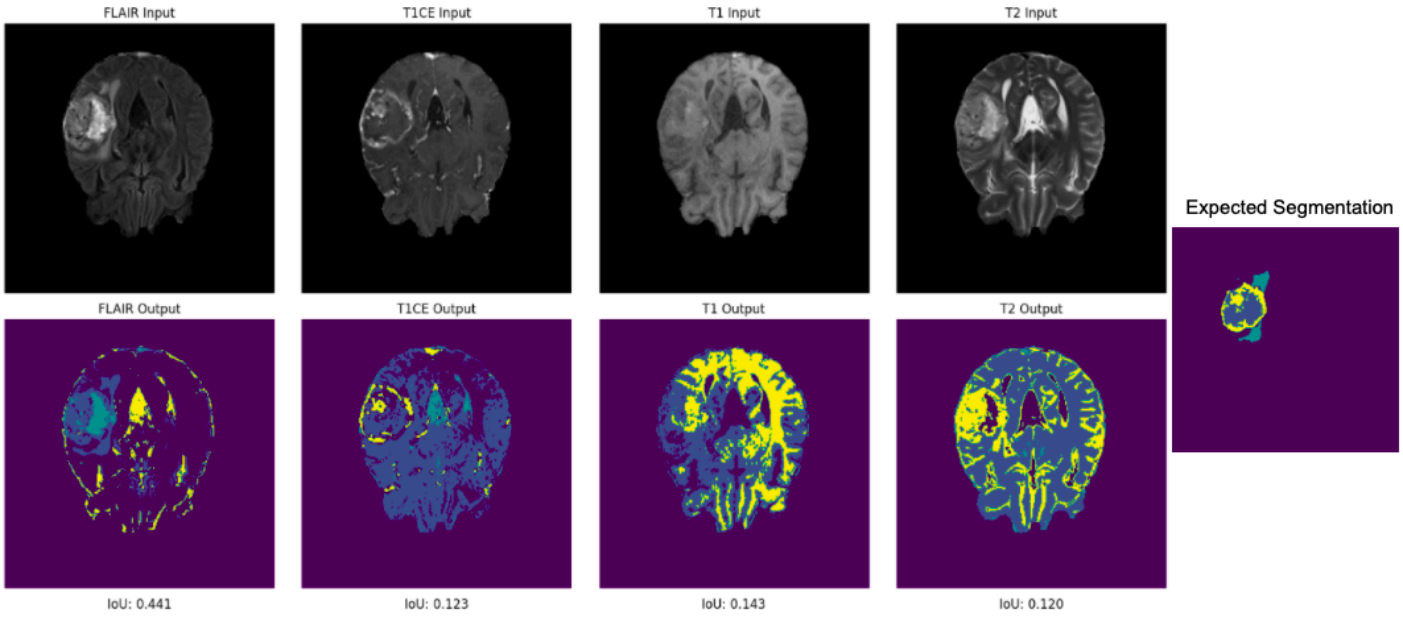
Otsu Thresholding:



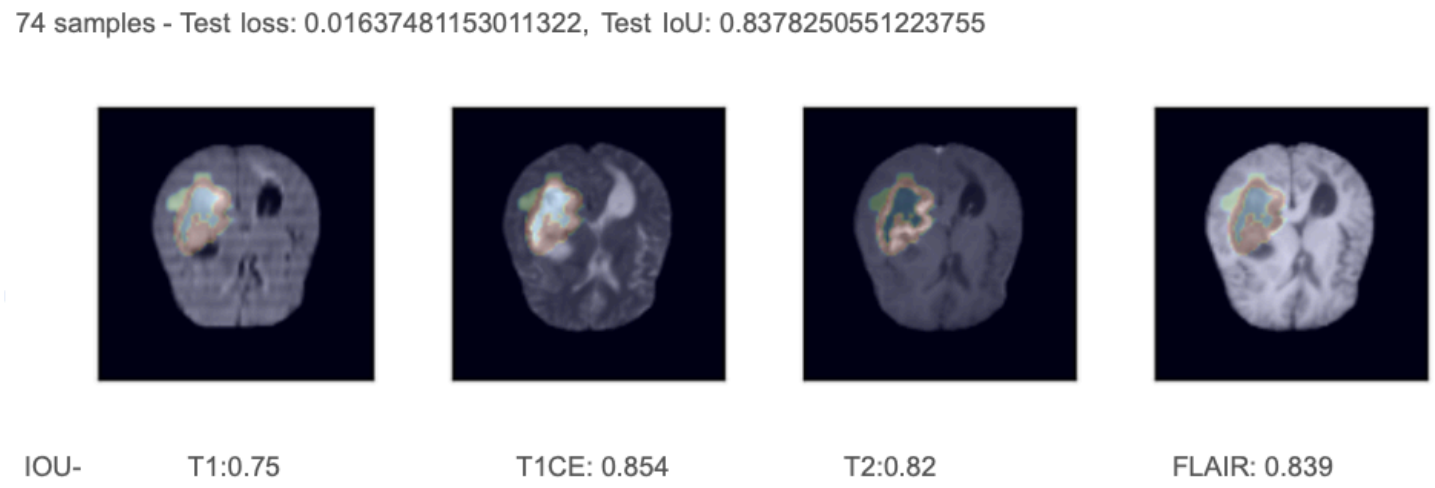
Region growing:



Clustering for segmentation:



Swin-Unetr:



7. REFERENCES

1. RSimonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the International Conference on Learning Representations (ICLR).
2. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
3. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Sander, C., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI).
4. Yu, L., & Liu, Y. (2015). Data Augmentation for Medical Image Segmentation with Convolutional Neural Networks. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI).
5. Tajbakhsh, N., et al. (2016). Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? IEEE Transactions on Medical Imaging.
6. Bakas, S., et al. (2018). Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progress, and Challenges. In Brain Tumor Image Analysis.
7. Dosovitskiy, A., et al. (2020). Vision Transformer for Image Recognition. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).
8. Cao, Z., Liu, L., Zhang, F., & Hu, Y. (2021). Swin-Unetr: Swin Transformer U-Net for Medical Image Segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).