Deep Learning-Based Multi-Compartment Brain Glioma Segmentation:Using U-Net, and Residual U-Net Architectures

Team-5

Bala krishna Ragannagari

Sai Avinash

Bhagavath Sai Darapureddy

Ujjawal Dwivedi

Contents

Introduction

Gliomas are one of the most common and aggressive brain tumors, posing significant challenges in diagnosis and treatment due to their invasive nature. Post-treatment MRI scans are critical for monitoring tumor progression, but accurate segmentation of tumor sub-regions remains challenging due to the tumor's complex structure.

This project focuses on segmenting multi-compartment gliomas from post-treatment MRI scans, aiming to identify regions such as:

- **Enhancing tissue (ET)**: Active tumor regions that show contrast enhancement.
- **Non-enhancing tumor core (NETC)**: Necrotic and cystic tumor areas.
- **Surrounding non-enhancing FLAIR hyperintensity (SNFH)**: Tumor infiltration and edema.
- **Resection cavity (RC)**: Space left after surgical removal of tumor tissue.

In this project, we explore and compare three deep learning models— **U-Net**, and **Residual U-Net**—for this segmentation task using the **BraTS 2024 dataset**. We aim to evaluate the models based on accuracy metrics like **Dice coefficient**, with the goal of contributing to the development of automated tools for glioma assessment and treatment planning.

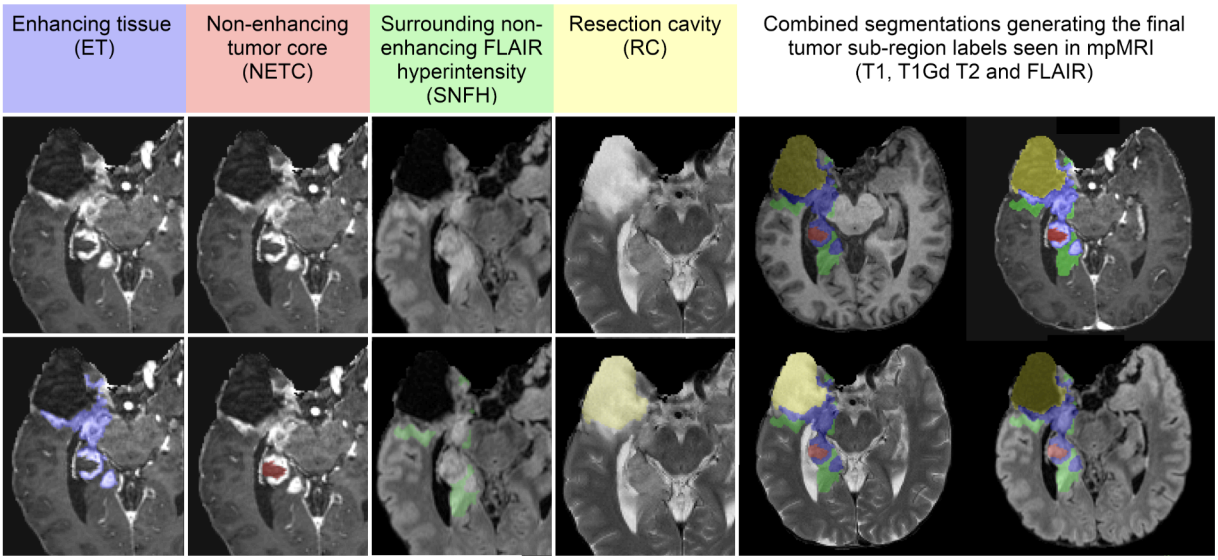


Fig 1.1 Multi-Compartment Glioma Segmentation in Post-Treatment MRI Scans

The following image illustrates the segmentation of these tumor sub-regions in post-treatment MRI scans. Each compartment is visually highlighted, with Enhancing Tissue (ET) shown in blue,

Non-enhancing Tumor Core (NETC) in red, Surrounding Non-enhancing FLAIR Hyperintensity (SNFH) in green, and Resection Cavity (RC) in yellow. The final column demonstrates the combined segmentation labels of these regions in multi-parametric MRI (mpMRI) images, including T1, T1Gd, T2, and FLAIR sequences.

Dataset

The **BraTS 2024 dataset** is a multi-institutional collection of **post-treatment MRI scans** of patients with glioma, made available for research purposes in the field of medical image segmentation. The dataset is specifically designed for the segmentation of brain tumor sub-regions, providing crucial information for understanding tumor progression and guiding treatment planning.This dataset includes data from 1350 patients, each having four types of MRI scans and corresponding segmented images for comprehensive analysis. The MRI scan types are:

- **T1c or T1-Gd (T1-weighted post-contrast)**: Primarily used to detect Enhancing Tissue (ET), which appears bright and highlights active tumor cells and areas where the blood-brain barrier is disrupted.
- **T1w or T1n (T1-weighted or native T1)**: Crucial for identifying the Non-enhancing Tumor Core (NETC), which appears dark on scans, indicating necrosis or cysts, with no contrast enhancement.
- **T2w (T2-weighted)**: Helps identify Surrounding Non-enhancing FLAIR Hyperintensity (SNFH), which appears hyperintense and includes edema, infiltrating tumor growth, and treatment-related changes, while excluding chronic ischemic damage.
- **T2f (T2-weighted FLAIR)**: Also highlights SNFH, offering additional contrast to differentiate pathological features.

Image Specifications

Each MRI scan is stored in NIfTI (.nii) format with the following dimensions:

- **First Index (182)**: 182 slices in the sagittal plane, showing a side view of the brain.
- **Second Index (216)**: 216 slices in the coronal plane, providing a front view.
- **Third Index (182)**: 182 slices in the axial plane, offering a top-down view of the brain.

Tissue Classification in MRI Scans

- **Enhancing Tissue (ET)**: Bright on T1-Gd images, indicating active tumor regions.
- **Non-enhancing Tumor Core (NETC)**: Appears dark on T1 and T1-Gd images but brighter on native T1 images, indicating necrotic or cystic regions.

- **Surrounding Non-enhancing FLAIR Hyperintensity (SNFH)**: Appears hyperintense on T2 and T2f images, representing edema and infiltrating tumor growth.
- **Resection Cavity (RC)**: The intensity varies with the cavity's age. Chronic cavities appear similar to cerebrospinal fluid, while recent ones may contain air, blood, or proteinaceous materials, showing varied signal characteristics.

Data formation for the preprocessing and Training :

We started by setting up our data directory and using random seed for reproducibility. We then split our dataset into training (80%) and validation (20%) sets. This split helps us train our models and validate their performance separately.

For the training data, we created two main folders: 'train_images' and 'masks'. We developed a process to handle the BraTS 2024 dataset efficiently. This process goes through each patient's directory in the training set, checking for the necessary image files (t1c, t1n, t2f, t2w) and the corresponding segmentation file. If all files are present, it saves the image files in the 'train_images' folder and the segmentation file in the 'masks' folder.Similarly, for the validation data, we created 'val_images' and 'val_masks' folders. This organization helps us keep our data neat and accessible for our U-Net and Residual U-Net models. It also makes it easier to load the data during the training and validation phases of our project.

Preprocessing Techniques

To enhance the performance and manage computational efficiency of our model on the Brats 2024 glioma MRI dataset, we implemented the following preprocessing steps:

Image Resizing: The MRI images were resized to (132, 132, 116) to simplify data handling and reduce computational load, while preserving essential features.

Z-Score Normalization: This technique standardized the pixel intensity values across all scans, setting the mean to zero and standard deviation to one. This step is critical for improving model convergence and reducing sensitivity to input variations.

Data Stacking: Considering the 3D nature of MRI data, we combined all four MRI scans (T1c, T1n, T2w, T2f) for each patient into a single comprehensive dataset. This approach allows the

model to learn from complex spatial relationships and enhances its ability to accurately classify tumor-related features.

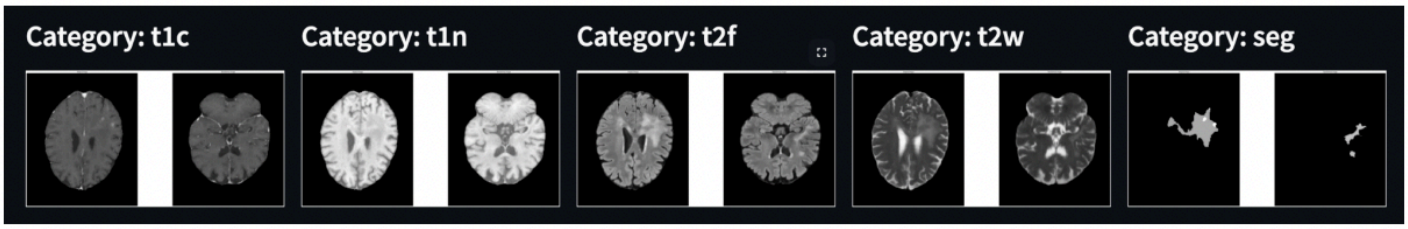
Data Augmentation Techniques

To augment our dataset and increase the robustness of our model, we employed several data augmentation techniques:

Zooming: Applied a 20% zoom to emphasize critical features within the MRI scans, particularly focusing on tumor areas.

Cropping: Following zooming, 80% of each image was cropped to maintain focus on the central and most informative parts of the scans.

Rotation: Each scan was rotated by 45 degrees to introduce variability and better simulate different imaging conditions encountered in clinical settings.



The image above demonstrates the use of rotation as an augmentation technique on various MRI images. In each category, the original scans are shown on the left, while the augmented versions after rotation are displayed on the right. Upon closer inspection, noticeable differences can be observed between the original and rotated images, particularly in the segmented regions.

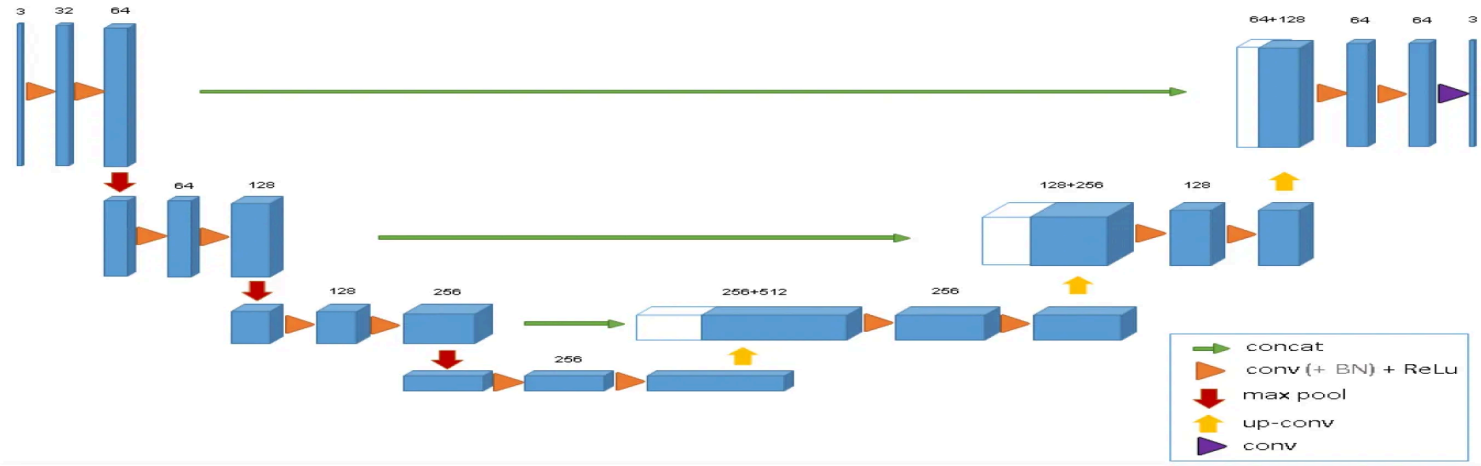
Model Architecture

U-Net (3D):

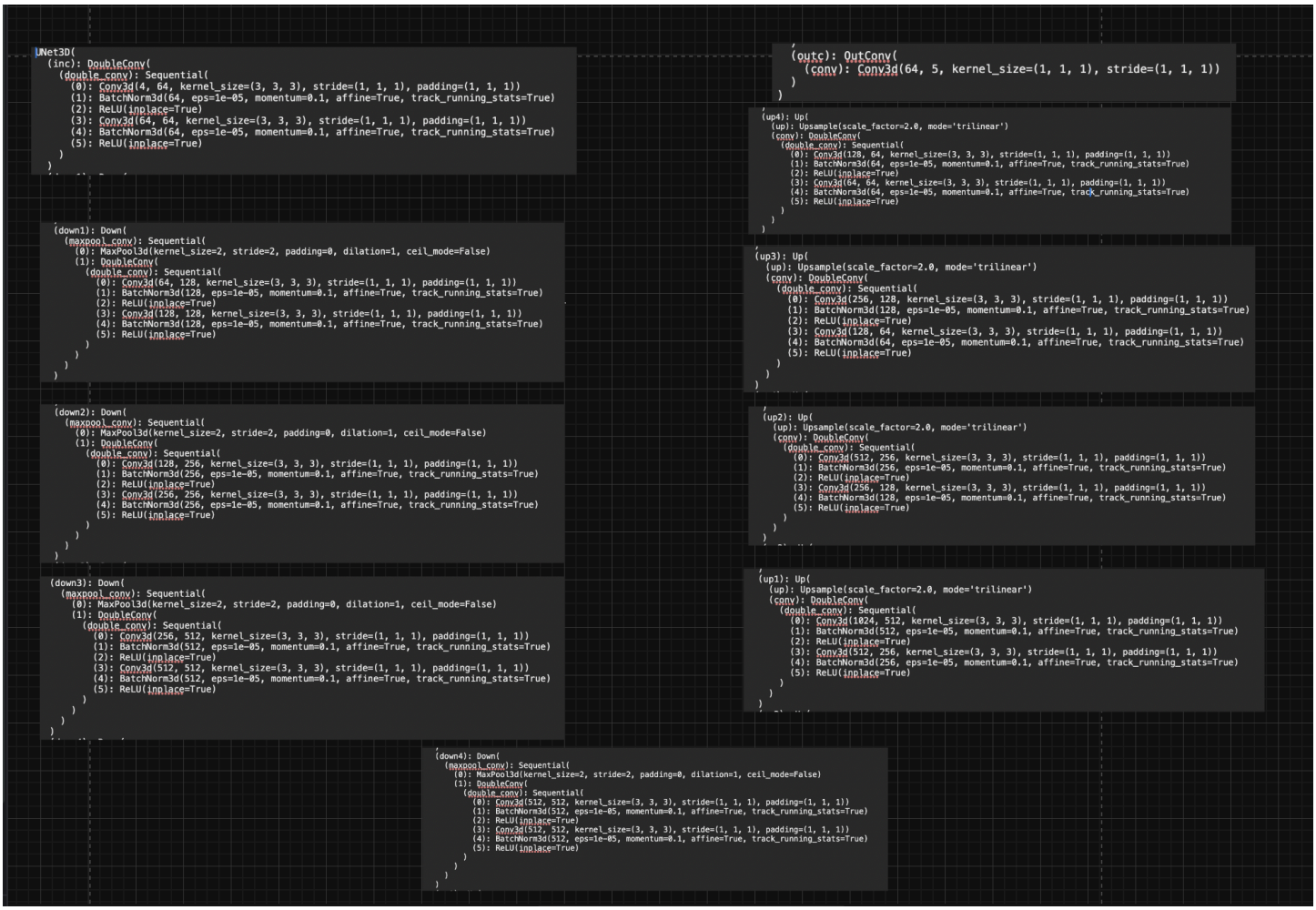
The U-Net architecture is specifically designed for medical image segmentation tasks, including glioma segmentation. It utilizes an encoder-decoder structure, with skip connections to retain spatial information for precise localization. Below is the 3D U-Net model architecture used in this project.

Model Structure:

- **Contracting Path (Encoder):** A series of convolutions, followed by max pooling, which reduces the spatial dimensions of the image while increasing feature complexity.
- **Bottleneck:** The deepest layer, where the most abstract features are learned.
- **Expansive Path (Decoder):** Upsampling the feature maps through transposed convolutions, with concatenation of features from the encoder to preserve spatial resolution.
- **Final Output Layer:** A 1x1 convolution to produce the segmentation maps for the tumor sub-regions.



STANDARD U NET ARCHITECTURE



Model architecture summary

3D U-Net Model Architecture

The two primary paths of the 3D U-Net architecture—the analysis path and the synthesis path—are tailored for volumetric segmentation and are derived from the classic U-Net architecture

Input Specifications: The model processes inputs of dimension (4, 166, 132, 132), suitable for volumetric medical imaging data.

.

The analysis path (encoder): consists of a sequence of convolutional layers, with two 3×3×3 convolutions in each layer. To add non-linearity, a ReLU activation function is then applied. To decrease the spatial dimensions and boost the representational capacity, a 2×2×2 max pooling operation with strides of two is conducted in each dimension after the convolutions.

Synthesis Path (Decoder): The synthesis or decoding path features layers of up-convolutions with a 2×2×2 kernel and strides of two in each dimension. Each up-convolution is followed by two 3×3×3 convolutions, each paired with a ReLU activation. This path gradually restores the spatial dimensions to match the original input size.

Shortcut Connections: To enhance feature propagation and reduce information loss, shortcut connections are employed from corresponding layers of equal resolution in the analysis path to the synthesis path. These connections help in preserving important spatial hierarchies necessary for accurate segmentation.

Results (U NET)

Loss and Dice Score Plots:

CrossEntropyLoss

We used CrossEntropyLoss for our segmentation task.

$$\text{Loss} = -\sum_{c=1}^{C} y_c \cdot \log(\hat{y}_c)$$

C: Total number of classes

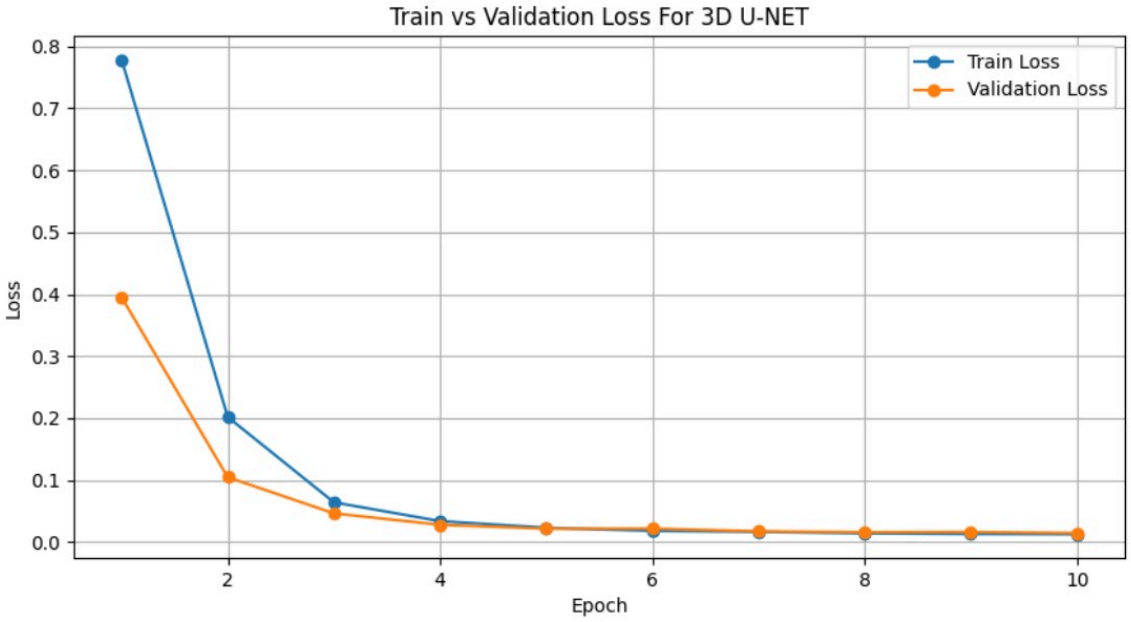
y_c: Ground truth label

$y\hat{}_c$: Predicted probability

CrossEntropyLoss is used as the criterion for training the 3D U-Net model in this project. It measures the difference between the predicted probability distribution and the true class labels, making it suitable for multi-class segmentation tasks like brain tumor segmentation

Training and Validation Loss:

Below is the **line plot** of the **training and validation loss**. It illustrates how the loss decreased over time during training, with both the training and validation losses showing similar trends.The graph shows how the training and validation loss change over 10 epochs for a 3D U-Net model. At the beginning, the training loss drops quickly, meaning the model is learning well from the training data. The validation loss also decreases at first, but more slowly, and then stays steady. This suggests the model is doing a good job on new, unseen data without overfitting, as both losses become quite stable by the end.



Dice coefficient:

For our segmentation task we implemented the Dice coefficient as a performance metric. The Dice coefficient measures the overlap between the predicted segmentation and the ground truth, providing a value between 0 and 1, where 1 indicates perfect overlap.

Our Dice coefficient formula is calculated as follows:

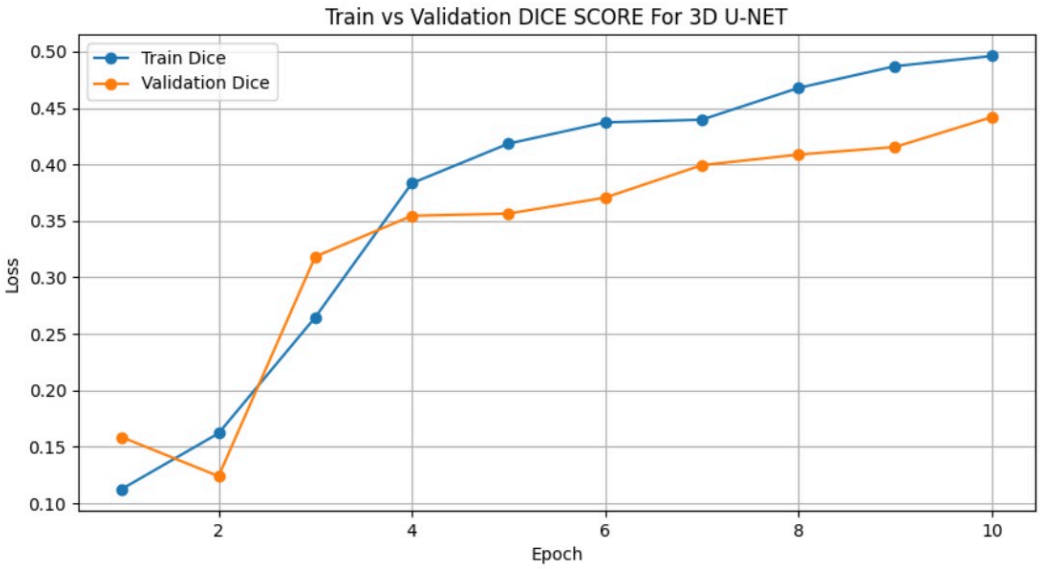
$$\text{Dice} = \frac{1}{C-1} \sum_{c=1}^{C-1} \frac{2|\text{Pred}_c \cap \text{Target}_c|}{|\text{Pred}_c| + |\text{Target}_c| + \epsilon}$$

- C is the number of classes
- Pred_c is the predicted segmentation for class c
- Target_c is the ground truth segmentation for class c
- ε is a small constant (1e-6) added for numerical stability

We calculate the Dice coefficient for each class separately (excluding the background class) and then average the results.

Training and Validation Dice Score:

This plot shows the **Dice coefficient** for both **train** and **validation**. It indicates the model's ability to segment the tumor sub-regions accurately. You can observe the improvement in the **train Dice score** while the **validation Dice score** shows some fluctuations, suggesting room for improvement in generalization.



Validation Results Summary:

metric	value
Validation Loss	0.0164
Dice Scores	
Enhancing Tumor (ET)	0.095
Non-enhancing Tumor Core (NETC)	0.297
Surrounding Non-enhancing FLAIR Hyperintensity (SNFH)	0.693
Resection Cavity (RC)	0.524

After training our model, we assessed its performance on the validation dataset, revealing a validation loss of 0.0164. The Dice scores for different tumor regions showed variability: the model identified the enhancing tumor with a score of 0.095, the non-enhancing tumor core with a score of 0.297, the surrounding non-enhancing FLAIR hyperintensity with a score of 0.693, and the resection cavity with a score of 0.524. These results highlight that while the model performed reasonably well in some areas, particularly in detecting surrounding hyperintensities, there is still significant need for improvement, especially in accurately identifying the enhancing tumor regions.

Model Performance Visualization

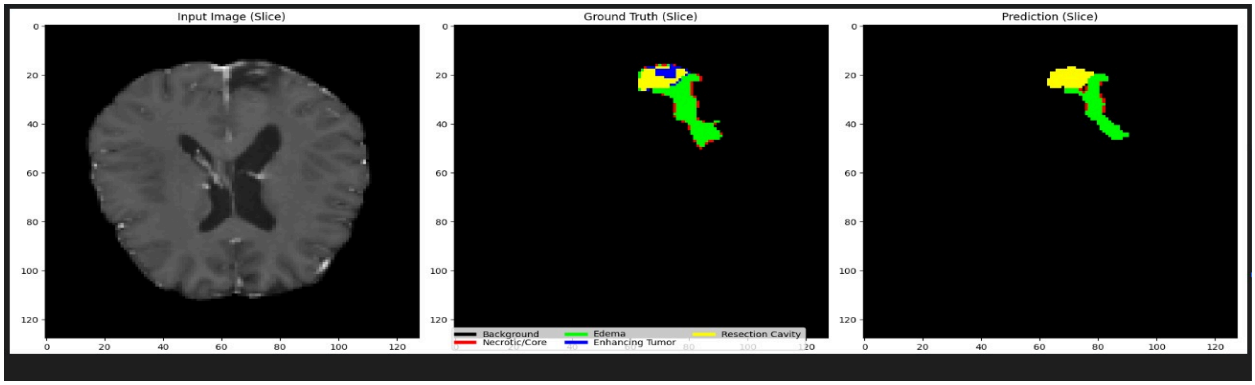
The following images display an example of model performance for segmentation:

- **Input Image**: The original MRI scan of the brain.
- **Ground Truth**: The manually annotated segmentation for the tumor sub-regions, including **Enhancing Tumor (blue)**, **Necrotic/Core (red)**, **Edema (Green)**, and **Resection Cavity (Yellow)**.
- **Prediction**: The segmentation predicted by the model.

Example: Tumor Segmentation

Ground Truth and Prediction Comparison:

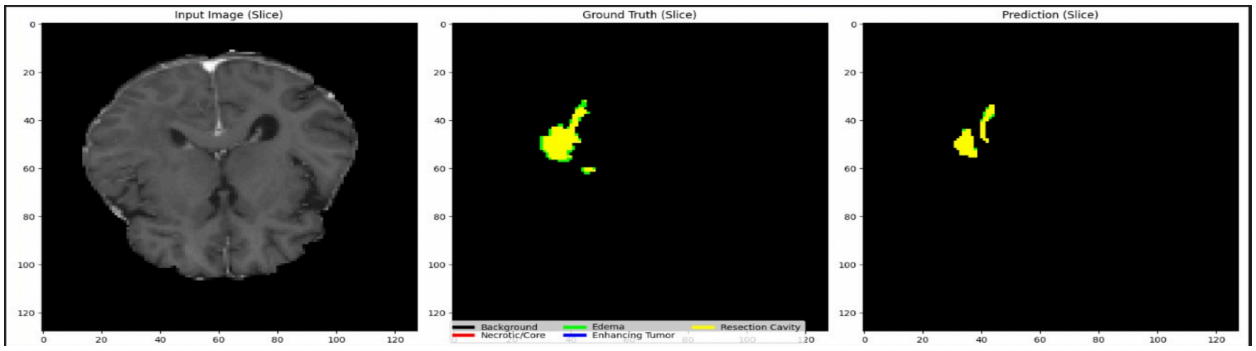
- This image shows a **slice of the MRI input**, with the corresponding **ground truth** (middle) and **predicted segmentation** (right).
- The **model successfully predicts the Edema (Green)** and **Resection Cavity (Yellow)** sub-regions, which appear clearly in the prediction. However, it struggles with the **Enhancing Tumor (Blue)** and **Necrotic/Core (Red)** areas, where the predicted segmentation is less accurate.



Example 2: Tumor Segmentation

Ground Truth and Prediction Comparison:

- This image shows a **slice of the MRI input**, with the corresponding **ground truth** (middle) and **predicted segmentation** (right).
- The **model predicts the Resection Cavity (Yellow)** well, but there is an incomplete prediction of the **Edema (Green)**, which is partly missing.



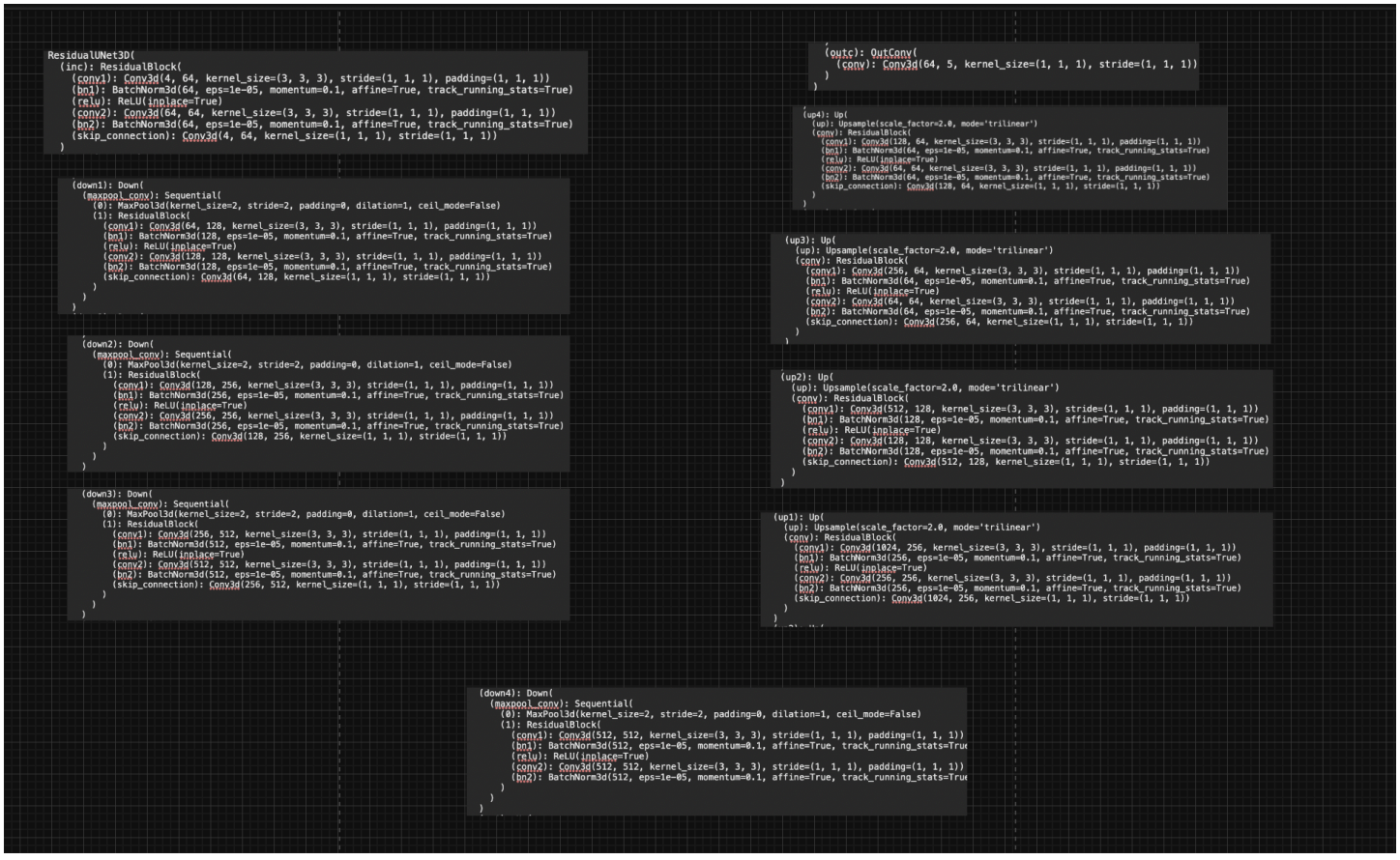
Residual 3D U-Net Architecture

The Residual U-Net architecture is a modified version of the original U-Net, incorporating residual connections to improve learning by allowing gradients to flow more easily through the network. This modification helps mitigate the vanishing gradient problem, which can be common in deep networks. The residual connections allow the model to learn both the identity and complex features simultaneously, improving both training speed and accuracy.

The Residual 3D U-Net is a powerful model for volumetric image segmentation, as it incorporates 3D convolutions, making it suitable for tasks involving 3D medical images, such as brain tumor segmentation.

Key features:

- **Residual Blocks**: The model includes residual blocks that add skip connections, improving gradient flow.
- **Downsampling and Upsampling:** The architecture includes downsampling layers (max-pooling) and upsampling layers (transposed convolutions).
- **Skip Connections:** These connections help preserve spatial resolution during upsampling by adding low-level features from earlier layers.



For the Residual U-Net model, we enhanced the traditional U-Net architecture by incorporating residual connections to improve feature learning and gradient flow. Here's what we did:

Residual Blocks:

- We replaced the standard convolutional blocks in U-Net with residual blocks. Each residual block consists of two 3D convolutional layers, batch normalization, and ReLU activation.

- A skip connection was added within each block, which either directly passes the input or uses a 1x1 convolution to match dimensions if the input and output channels differ. This helps mitigate vanishing gradient issues and allows the network to learn identity mappings more effectively.

Network Architecture:

- The encoder path uses a series of downsampling modules (Down), where each module applies max pooling followed by a residual block.
- The decoder path uses upsampling modules (Up), where features from the encoder are concatenated with upsampled features from the decoder, followed by a residual block.
- The final layer uses a 1x1x1 convolution to map the features to the desired number of output classes.

Input and Output:

- The model was designed to handle 4 input channels (corresponding to different MRI modalities) and produce 5 output classes (background and tumor subregions).

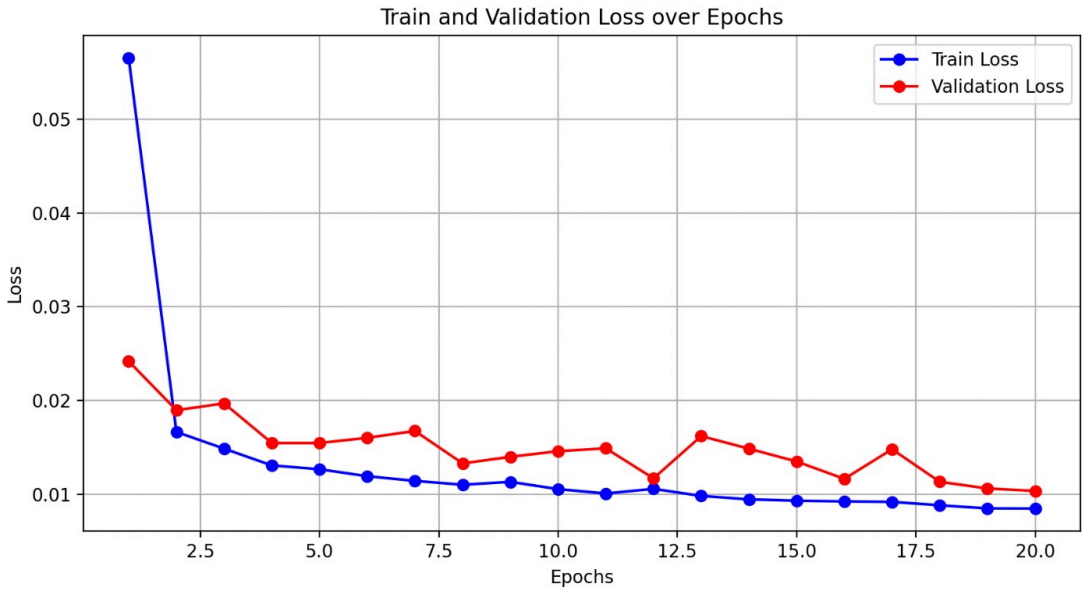
By integrating residual connections, this architecture improves gradient flow during backpropagation, enabling deeper networks to be trained effectively. The model captures both low-level and high-level features better, which is crucial for segmenting complex structures like brain tumors.This Residual U-Net model builds on the strengths of U-Net while addressing its limitations, making it more robust for our brain tumor segmentation task.

Results :(Residual U net)

Loss and Dice Plots:

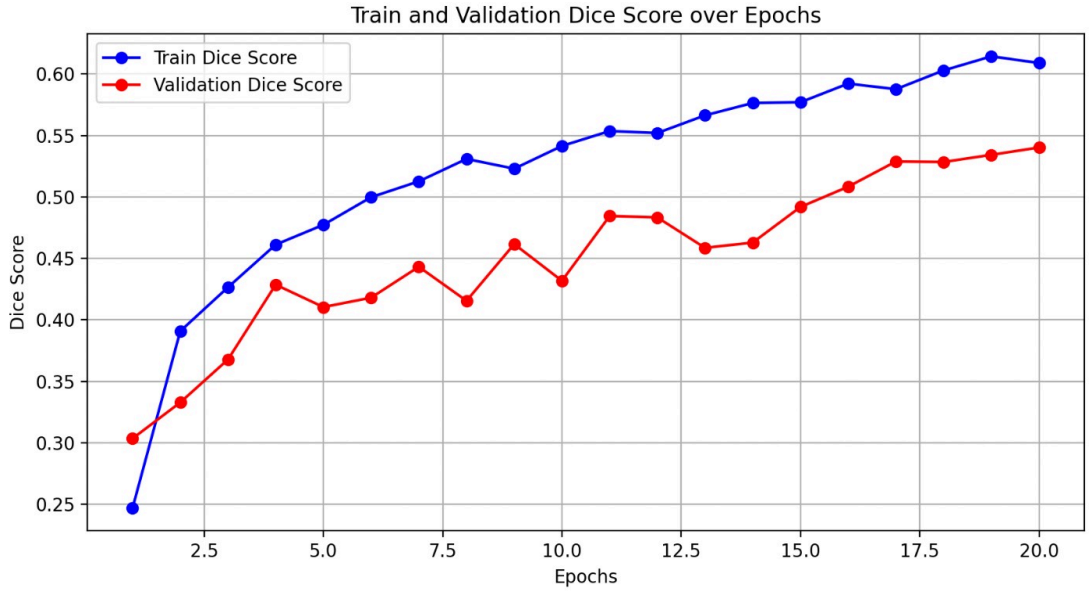
Train and Validation Loss:

- The blue curve represents the training loss, which decreases steadily over the epochs.
- The red curve represents the validation loss, showing a similar decreasing trend, with some fluctuations around the middle epochs, which is common during training.
- Both losses appear to converge towards the lower end, indicating good model performance and generalization.



Train and Validation Dice Score:

- The blue curve (train Dice score) steadily increases, indicating improving performance on the training set.
- The red curve (validation Dice score) also increases but with more fluctuation compared to the training score. This suggests that while the model is improving, some adjustments or regularization might be needed to stabilize the validation performance.



Validation Results Summary-3d residual Unet:

metric	value
Validation Loss	0.0104
Dice Scores	
Enhancing Tumor (ET)	0.325
Non-enhancing Tumor Core (NETC)	0.422
Surrounding Non-enhancing FLAIR Hyperintensity (SNFH)	0.797
Resection Cavity (RC)	0.61

The validation loss was 0.0104, indicating a low overall error in predictions. The Dice scores for individual tumor regions showed that the model performed well in segmenting the surrounding non-enhancing FLAIR hyperintensity with a score of 0.797 and the resection cavity with a score of 0.61. The non-enhancing tumor core achieved a score of 0.422, while the enhancing tumor had a

lower score of 0.325. These results show that the model is effective in segmenting certain tumor regions compared with U net.

Model Performance Visualization:

The following images display an example of model performance for segmentation:

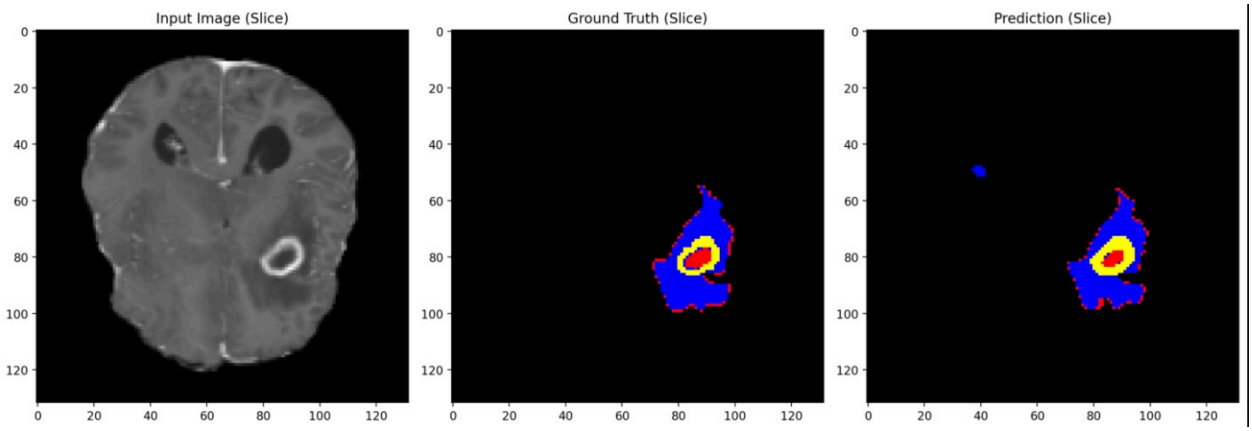
- **Input Image**: The original MRI scan of the brain.
- **Ground Truth**: The manually annotated segmentation for the tumor sub-regions, including **Enhancing Tumor (blue)**, **Necrotic/Core (red)**, **Edema (Green)**, and **Resection Cavity (Yellow)**.
- **Prediction**: The segmentation predicted by the model.

Example 1: Tumor Segmentation

- **Ground Truth and Prediction Comparison**:

  This image shows a **slice of the MRI input**, with the corresponding **ground truth** (middle) and **predicted segmentation** (right).

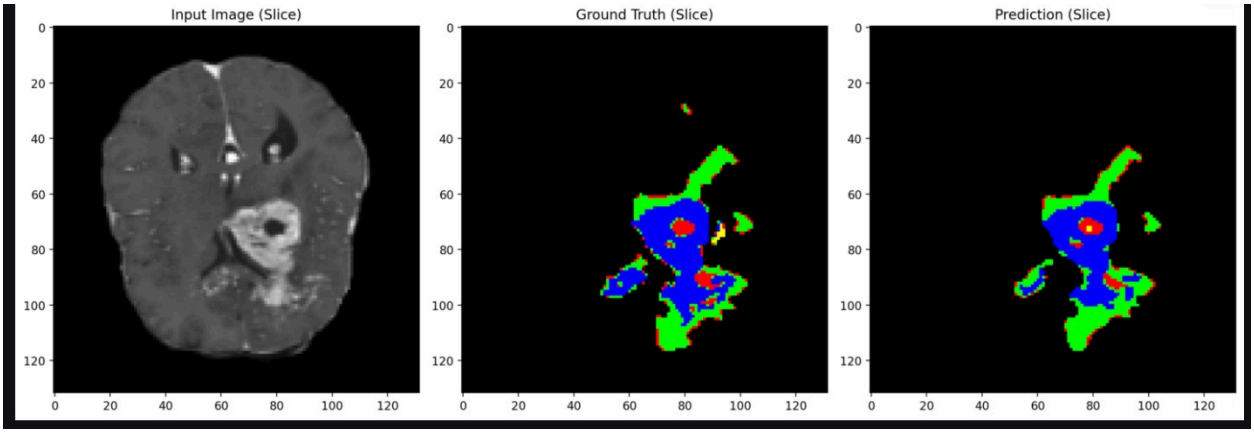
  The **model successfully predicts the Enhancing Tumor (blue)**and **Resection Cavity (Yellow)** ,**Necrotic/Core (Red)**sub-regions, which appear almost clear.



Example 2: Tumor Segmentation

Ground Truth and Prediction Comparison:

- This image shows a **slice of the MRI input**, with the corresponding **ground truth** (middle) and **predicted segmentation** (right).
- The **model successfully predicts the almost Enhancing Tumor (blue)**and edema (green) ,sub-regions, which appear almost clear. There is partial prediction **Necrotic/Core (Red),Resection Cavity (Yellow)** .



Conclusion :

In our project, we focused on improving the segmentation of brain tumors in MRI scans from the BraTS 2024 glioma dataset. We started by organizing the data into training and validation sets, which helped us manage the model development process effectively.

To prepare the images for analysis, we applied several preprocessing techniques. We resized the MRI images to a uniform size of (132, 132, 116) to make them easier to handle and reduce computational demands. We also used Z-score normalization to standardize pixel intensity values, which is crucial for helping our model learn effectively. Additionally, we stacked multiple MRI modalities (T1c, T1n, T2f, T2w) into a single dataset to capture complex spatial relationships in the images. To further enhance our dataset, we employed data augmentation techniques like zooming, cropping, and rotating the images to create variations that help the model generalize better.

We implemented two models: U-Net and Residual U-Net. The U-Net model showed decent performance but had some limitations in accurately identifying certain tumor regions. In contrast, the Residual U-Net, which included local skip connections to improve learning, significantly outperformed the U-Net across all tumor types. For example, it achieved better Dice scores for enhancing tumors and surrounding hyperintensities.

Overall, our project demonstrated that using advanced deep learning techniques like U-Net and Residual U-Net can greatly enhance brain tumor segmentation accuracy in MRI scans. These improvements can potentially lead to better diagnosis and treatment planning for patients with gliomas.

Reference :

- https://arxiv.org/pdf/1606.06650
- https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10635180&tag=1
- https://www.sciencedirect.com/science/article/pii/S1361841524002056