Deep Learning Project

Topic - Brain Tumor Segmentation

Group Members

1. Sameer Ahmed(IIT2020053)
2. Harsh Garg (IIT2020082)
3. Anuj Chaturvedi (IIT2020019)
4. Siddhant Agarwal (IIT2020228)

Index

Introduction & Motivation

- Brain and nervous system tumors have a high death rate and are the leading cause of cancer death among youth. Gliomas have the highest incidence rate among such malignancies.[1]
- Many deep learning methods have been used for MRI image segmentation but not all produce accurate results, this paper uses 3D-UNET with Spatial Attention which outperforms other methods.[3]
- UNET can lose local features during encoding, resulting in reduced accuracy.Spatial attention is used to preserve local features during encoding and decoding. The encoding path includes additional modules at each level to preserve local features.

Problem Statement

Develop a Brain Tumor Segmentation model to improve Diagnosis and Treatment Planning from MRI scans

Literature Review

Paper title /Year	Method/ approach used	Achieved Performance	Disadvantages	Dataset
Raza, Rehan, et al. "dResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI." Biomedical Signal Processing and Control 79 (2023): 103861.	The proposed model is a hybrid of the deep residual network and U-Net model (dResU-Net). The residual network is used as an encoder in the proposed architecture with the decoder of the U□Net model to handle the issue of vanishing gradient. The proposed model is designed to take advantage from low□level and high-level features simultaneously for making the prediction. In addition, shortcut connections are employed between residual network to preserve low-level features at each level.	Dice score for the tumor core (TC) - 0.8357, whole tumor (WT) -0.8660, and enhancing tumor (ET)-0.8004, on the BraTS 2020 dataset	dResU-Net 3D is computationally expensive to train and execute and This model requires a large amount of medical images to train effectively .dResU-Net 3D has shown relatively poor performance in segmenting heterogeneous tumors, which are tumors with varying shapes, sizes, and intensities.	The model was trained and validated on the BraTS 2020 training dataset.
Allah, Ahmed M. Gab, Amany M. Sarhan, and Nada M. Elshennawy. "Edge U-Net: Brain tumor segmentation using MRI based on deep U-Net model with boundary information." Expert Systems with Applications 213 (2023): 118833.	The Edge U-Net model is a modified version of the U-Net model that incorporates boundary information to improve the segmentation of brain tumors in MRI images. The model uses a combination of convolutional and deconvolutional layers, skip connections, and edge maps to identify and segment tumor regions.	The Edge U-Net model achieved a Dice similarity coefficient (DSC) of 0.88 and a Jaccard index (JI) of 0.81 on the BraTS 2018 validation dataset, which outperforms several state-of-the-art methods.	The main disadvantage of the Edge U-Net model is that it requires additional computational resources and time compared to the original U-Net model due to the incorporation of boundary information. Additionally, the model may be sensitive to the quality of the edge maps generated during training.	The Edge U-Net model was evaluated on the BraTS 2018 dataset, which includes multimodal MRI images of brain tumors from multiple institutions. The dataset contains 285 training images and 66 validation images

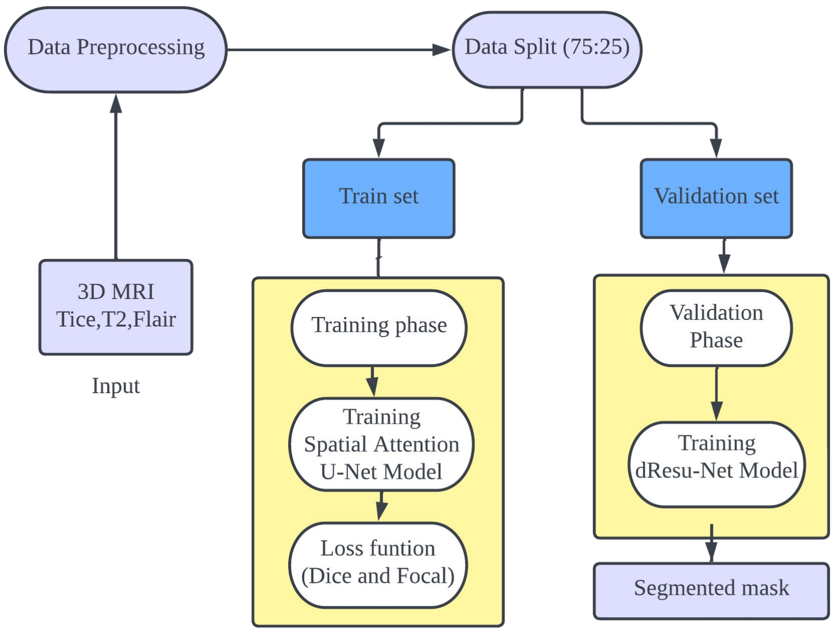
Paper title /Year	Method/ approach used	Achieved Performance	Disadvantages	Dataset
"SAResU-Net: Shuffle attention residual U-Net for brain tumor segmentation," 2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Beijing, China, 2022, pp. 1-6, doi: 10.1109/CISP-BMEI56279.2022.9979978.	SAResU□Net combines several shuffle attention (SA) blocks and residual modules with a basic 3D U-Net, where SA blocks are added to skip connection positions to capture the local spatial and channel information. In addition, a self-ensemble module is leveraged to further boost the model performance.	DSC values of 77.74%, 90.40% an 83.58%, 79.17%, 90.02% and 82.00% for the enhancing tumor (ET), the whole tumor (WT), and tumor core(TC) on the BraTS 2019 and 2020 validation dataset	We cant intrepret the modules that are used and also more complex modules, using these can cause overfitting in data.	2019 and 2020 Brats Dataset
MBANet: A 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images. Biomedical Signal Processing and Control, 80, 104296. https://doi.org/10.1016/j.bspc.2022.104296	The 3D convolutional neural network with 3D multi-branch attention called MBANet is proposed in the paper. First, the basic unit (BU) module of MBANet is constructed using the optimised shuffle unit. The BU module uses channel shuffle to jumble the convolutional channels after fusion and group convolution to execute convolution after the input channel has been split. The attention layer of the encoder is then used by MBANet's novel multi-branch 3D Shuffle Attention (SA) module.	Experiments on the BraTS 2018 and BraTS 2019 show that the dice of ET, WT and TC reach 80.18%, 89.80%, 85.47% and 78.21%, 89.79%, 83.04%, respectively.	MBANet solely uses publicly available multimodal datasets for its experiments.	BraTS 2018 and BraTS 2019 validation datasets
DPAFNet: A Residual Dual-Path Attention-Fusion Convolutional Neural Network for Multimodal Brain Tumor Segmentation, Biomedical Signal Processing and Control, Volume 79, Part 1, 2023, 104037, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2022.104037. (https://www.sciencedirect.com/science/article/pii/S1746809422005146)	A new 3D model, DPAFNet (Dual-Path and Multi-scale Attention Fusion Network), is proposed for efficient brain tumor segmentation. The model consists of a 3D feature extraction block and a 3D IDCM (Integrated Dense Context Module) block. The feature extraction block uses a Dual-Path module and a Multi-scale Attention Fusion module to capture both spatial and contextual information and aggregate global and local information. The 3D IDCM block merges feature maps with different receptive fields for dense pixel-level prediction. The proposed model is expected to improve accuracy and speed of brain tumor segmentation.	results on BraTS2020 training set is DPAFNet 78.1 89.4 83.2 and on BraTS2018 training set is DPAFNet Yes 78.9 89.5 79.9.	The DPAFNet model has some limitations, including: complexity, computational demands, limitations in handling non-uniform tumors, and limitations in handling low contrast tumors. These limitations should be considered when evaluating the suitability of the model for a specific use case.	BraTS2018, BraTS2019 and BraTS2020

Methods

1.Preprocessing

- Loaded MRI scans and masks.
- Normalized MRI scans intensity values to a range of 0 to 1.
- Applied data augmentation techniques to increase training dataset variability: random rotations, translations, shearing, scaling, and horizontal flipping of input data.
- Stacked three MRI scans into a single 3D image volume.
- Cropped image volume and its corresponding mask to remove empty borders.
- Converted mask to a one-hot encoded representation.
- Saved preprocessed image and mask as NumPy arrays.
- Data is split into training and validation sets with a ratio of 75:25.

Proposed Methodology



General Flow

Dataset

The Dataset selected is Brain Tumor Segmentation(BraTS2020) for training and BraTS 2021 for testing.

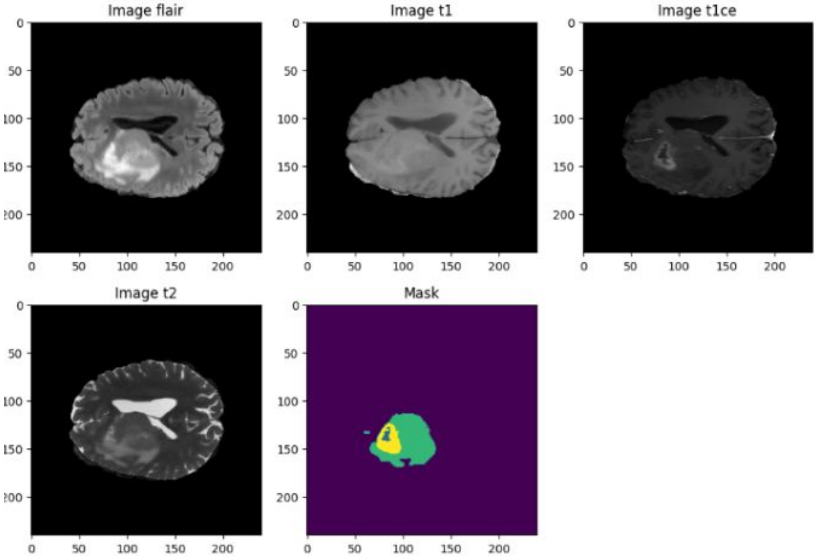
Tools and Libraries used

python 3.7, pytorch 1.6.0

torchvision 0.7.0

Pickle, Keras, glob

Nibabel, splitfolders



Sample MRI images and their ground truth (green for Edema, blue for Necrosis, and yellow for Enhancing tumor)

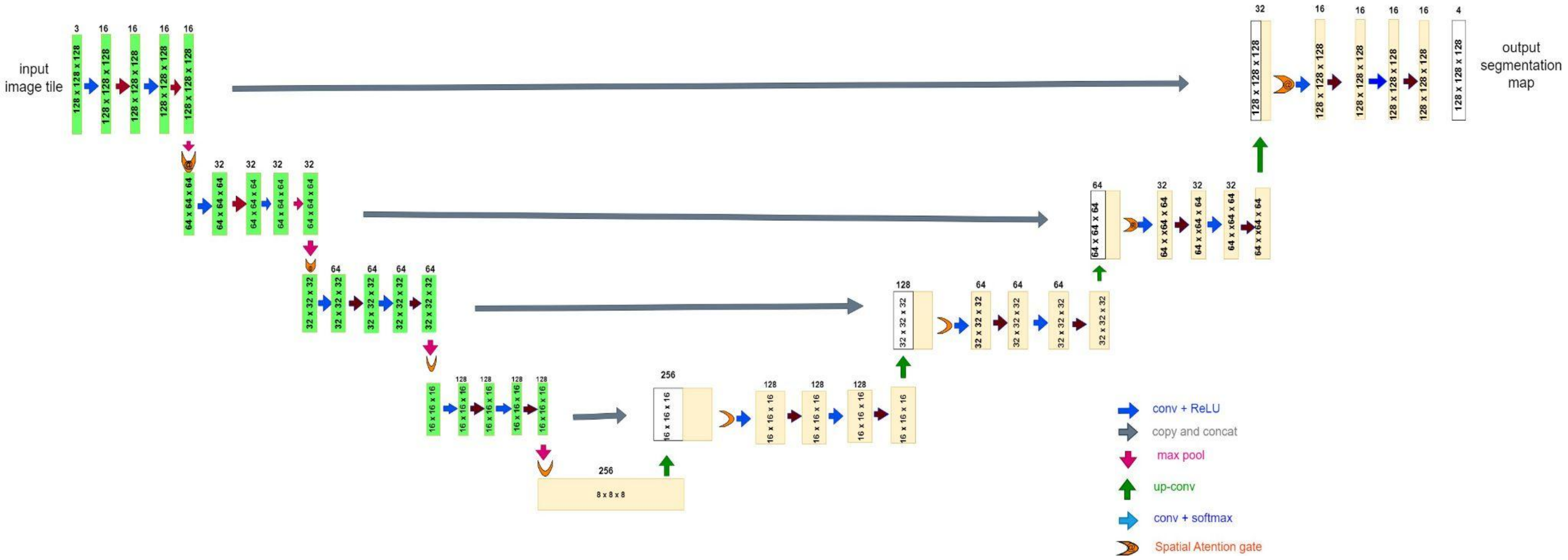
2. Model Architecture - Basic U-Net Model

The U-Net model is commonly used for semantic segmentation in medical imaging. The model consists of a contraction path, which downsamples the input volume and captures context information, and an expansion path, which upsamples the feature maps and combines them with the corresponding feature maps from the contraction path to produce the segmentation mask.

- The contraction path consists of several 3D convolutional layers with batch normalization and ReLU activation, followed by max pooling to reduce the spatial dimensions of the feature maps.
- The expansion path consists of several 3D transposed convolutional layers, also known as deconvolutional layers, that upsample the feature maps and concatenate them with the corresponding feature maps from the contraction path.
- The upsampled feature maps are then passed through a series of 3D convolutional layers with batch normalization and ReLU activation.

Overall, the U-Net architecture is a powerful model for accurately segmenting brain tumors in MRI.

Spatial Attention U-Net Model



Spatial Attention U-Net Model

The model architecture consists of an encoder and a decoder path, which are connected by skip connections.

The encoder path consists of a series of convolutional and max pooling layers with batch normalization and activation functions.

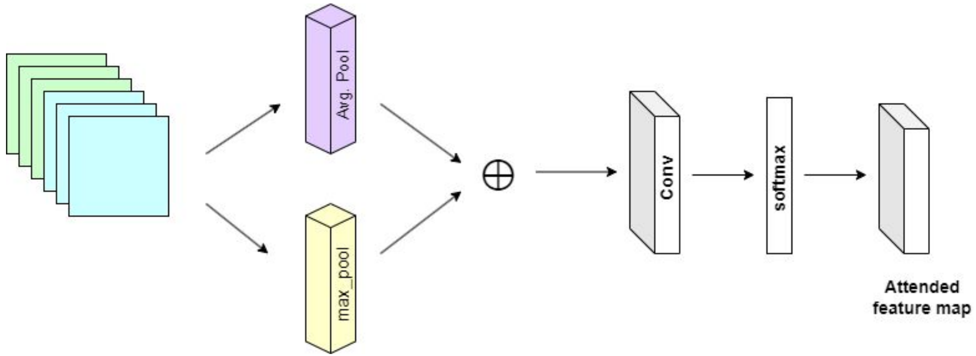
- The attention mechanism is applied to the output of the max pooling layers.

The decoder path consists of a series of transposed convolutional layers with skip connections from the corresponding encoder layers.

- The attention mechanism is applied to the output of the concatenation of the transposed convolutional layer and the corresponding encoder layer.

Spatial Attention Module

The attention mechanism is implemented using the channel-wise attention approach, which is a type of self-attention mechanism that focuses on the importance of individual channels in the feature maps.



The channel_attention() function computes the mean and max pooling of the feature map along the channel dimension and concatenates them. Then, a 3D convolution is applied to the concatenated feature map to obtain channel attention weights. Finally, the feature map is multiplied with the attention weights to obtain the attended feature map.

3.Loss Function and Optimization

The Adam optimizer is an optimization algorithm that combines the advantages of AdaGrad and RMSProp. It maintains an adaptive learning rate for each parameter and adjusts the learning rate based on the first and second moments of the gradients. The algorithm uses these moments to estimate the variance and bias, respectively, and corrects them to provide a more accurate estimate of the gradients.

The update rule for Adam optimizer is as follows:

m_t = $\boldsymbol{\beta}$1 * m_t-1 + (1 - $\boldsymbol{\beta}$1) * g_t)

v_t = $\boldsymbol{\beta}$2 * v_t-1 + (1 - $\boldsymbol{\beta}$2) * g_t^2

m_t_hat = m_t / (1 -$\boldsymbol{\beta}$1t)

v_t_hat = v_t / (1 - $\boldsymbol{\beta}$2t)

$\boldsymbol{\Theta}$_t = $\boldsymbol{\Theta}$_t-1 - $\boldsymbol{\alpha}$* m_t_hat / (sqrt(v_t_hat) + $\boldsymbol{\varepsilon}$)

where m and v are the first and second moments, beta1 and beta2 are the exponential decay rates for the moment estimates, g is the gradient, t is the time step, theta is the parameter to be optimized, alpha is the learning rate, and epsilon is a small constant to prevent division by zero.

Loss Functions-

- Dice Loss-

Dice loss is a loss function used in image segmentation problems, which is based on the Dice coefficient. The Dice coefficient is a similarity metric that measures the overlap between two sets or images, where a higher value indicates a better overlap. In image segmentation, the Dice coefficient is used to measure the similarity between the predicted segmentation mask and the ground truth mask.

The Dice loss is defined as:

$$Dice\_loss = 1 - 2 * (|X \cap Y|) / (|X| + |Y|)$$

where X and Y are the predicted and ground truth segmentation masks, respectively, and |.| denotes the cardinality or number of elements in the set.

- Focal Loss-

In binary classification, the objective is to predict the correct label for each input, which can be either 0 or 1. However, in many real-world scenarios, the number of examples in one class can be significantly larger than the other class, leading to class imbalance. This can result in the model performing poorly on the minority class.

Focal loss addresses this issue by down-weighting the loss for well-classified examples, which are easy to classify, and focusing more on hard examples, which are incorrectly classified. The loss function is defined as:

FL(p,y) = -(1-p)^gamma * log(p) if y=1 -(p)^gamma * log(1-p) if y=0

where p is the predicted probability, y is the true label (either 0 or 1), and gamma is a tunable parameter that controls the degree of down-weighting for easy examples. When gamma=0, focal loss reduces to cross-entropy loss.

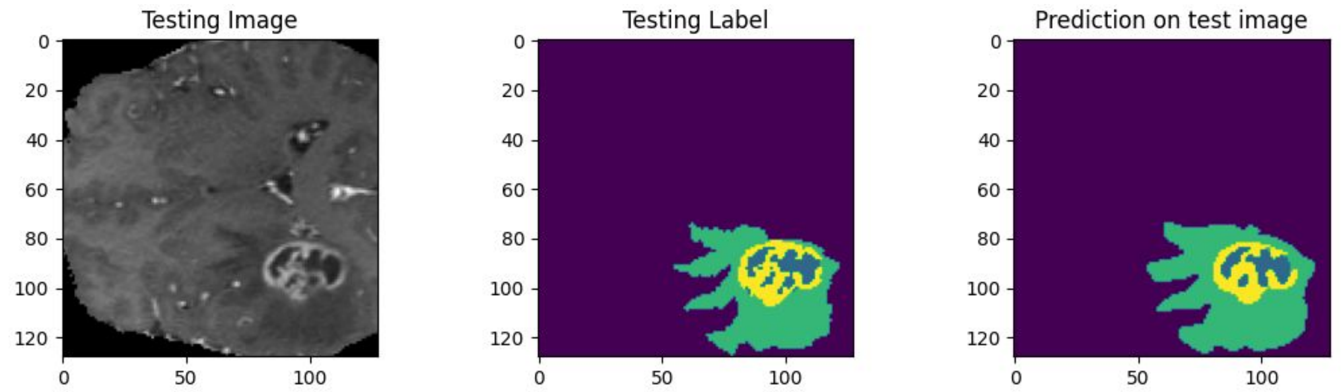
- IOU Test-

To calculate the IOU, we first find the intersection between the predicted and ground truth areas, which is the overlapping region. Then we divide this area by the union of the predicted and ground truth areas, which is the total area covered by both.

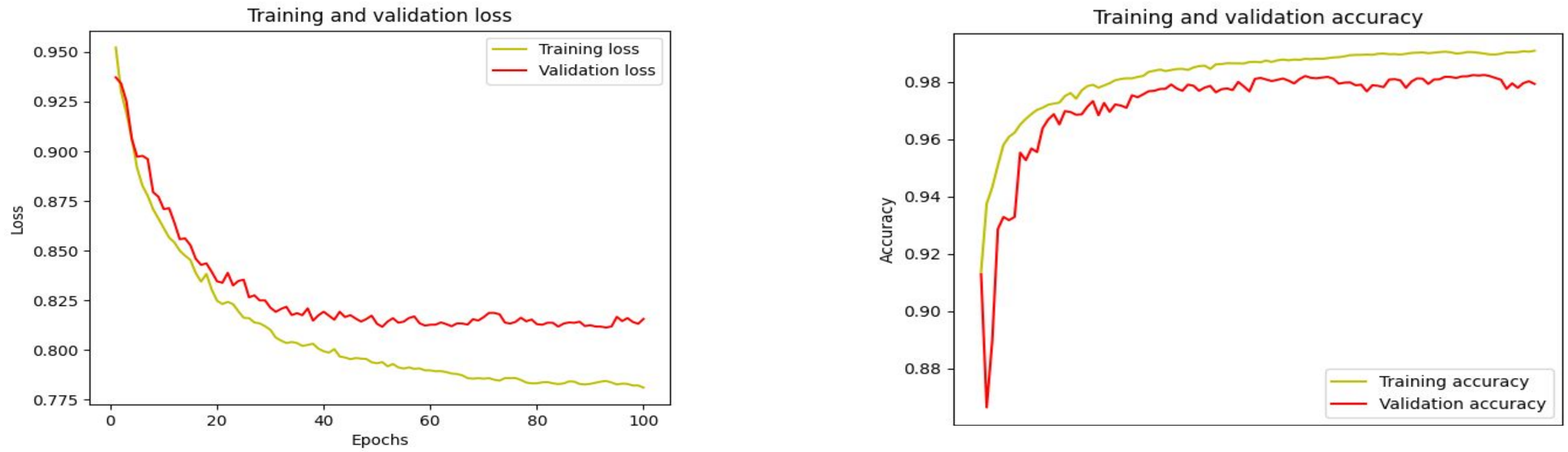
The IOU is a value between 0 and 1, where 0 indicates no overlap and 1 indicates perfect overlap. A higher IOU value indicates better performance of the model in predicting the target object.

Result Analysis

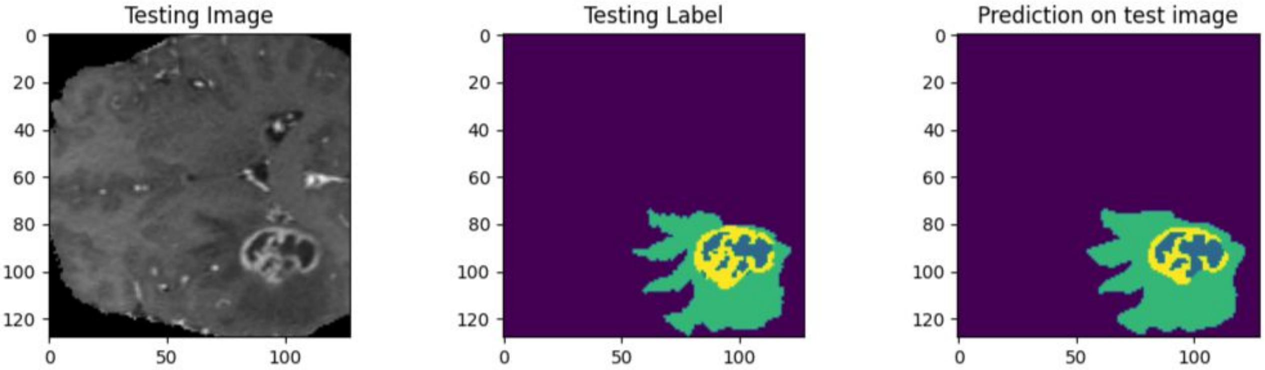
- **Implementation Details**
  - The proposed model was implemented using Python, Keras library, and TensorFlow as backend.
  - ADAM optimizer with a learning rate of 0.0001 was used.
  - ReLU activation function with batch normalization was employed for increased stability and normalization.
  - The model was trained for 100 epochs with a batch size of 2 due to limited computational resources.
  - Experiments were conducted on the BraTS 2020 benchmark dataset with 75% data used for training and 25% for validation, and BraTS 2021 dataset was used for testing.



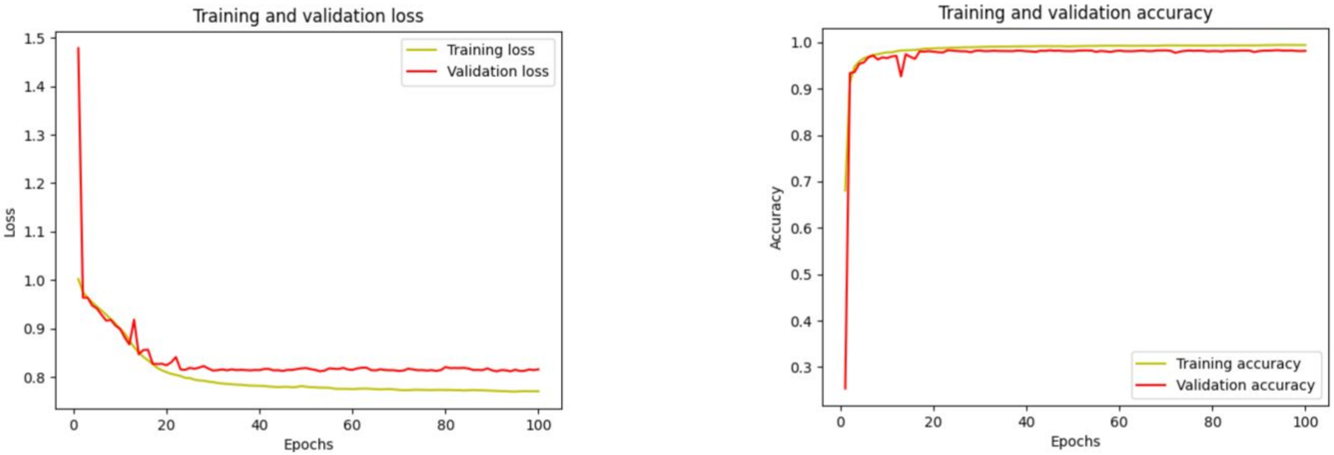
Simple UNet Architecture Prediction Image[6]



Simple Unet Training And Validation Loss with Training and Validation accuracy [6]



Spatial Attention UNet Architecture Prediction Image



Spatial attention Unet Training And Validation Loss with Training and Validation accuracy

- For Testing - BraTS 2020 Dataset
- For Training - BraTS 2021 Dataset

*Models	Accuracy		IOU		IOU(Test)
	train	val	train	val	
Simple Unet	0.9908	0.9792	0.8123	0.6753	0.6965
Optimized-unet	0.9932	0.9774	0.8502	0.6675	0.7025
Spatial Attention Unet.	0.9942	0.9818	0.8698	0.6773	0.6962

Simple UNet[6], Optimized Unet[12]

Conclusion

- Spatial U-Net proposed to improve brain tumor segmentation on 3D MRI images.
- The attention mechanism is implemented using the channel-wise attention approach, which is placed in the encoder network layer and skip connection of network
- It can effectively focus on the brain tumor region in the feature maps and extract more effective image features for more accurate segmentation
- Future improvements can be made through data augmentation, exploring other 3D-based architectures, and utilizing post-processing techniques.
- Proposed method can be applied to other segmentation problems in medical imaging.

References

[1] Miller, KD, Ostrom, QT, Kruchko, C, Patil, N, Tihan, T, Cioffi, G, Fuchs, HE, Waite, KA, Jemal, A, Siegel, RL, Barnholtz-Sloan, JS. Brain and other central nervous system tumor statistics, 2021. CA Cancer J Clin. 2021. https://doi.org/10.3322/caac.21693

[2] Holland EC. Progenitor cells and glioma formation. Curr Opin Neurol. 2001 Dec;14(6):683-8. doi: 10.1097/00019052-200112000-00002.

[3] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P., & Larochelle, H. (2017). Brain tumor segmentation with Deep Neural Networks. Medical Image Analysis, 35, 18-31. https://doi.org/10.1016/j.media.2016.05.004

[4] Cao, Y., Zhou, W., Zang, M., An, D., Feng, Y., & Yu, B. (2023). MBANet: A 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images. Biomedical Signal Processing and Control, 80, 104296. https://doi.org/10.1016/j.bspc.2022.104296

[5] Raza, R., Ijaz Bajwa, U., Mehmood, Y., Waqas Anwar, M., & Hassan Jamal, M. (2023). DResU-Net: 3D deep residual U-Net based brain tumor segmentation from multimodal MRI. Biomedical Signal Processing and Control, 79, 103861. https://doi.org/10.1016/j.bspc.2022.103861

[6]Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv. /abs/1505.04597
[7] Jiangyun Li, Hong Yu, Chen Chen, Meng Ding, Sen Zha Category Guided Attention Network for Brain Tumor Segmentation in MRI https://arxiv.org/abs/2203.15383
[8] Chang, Y., Zheng, Z., Sun, Y., Zhao, M., Lu, Y., & Zhang, Y. (2023). DPAFNet: A Residual Dual-Path Attention-Fusion Convolutional Neural Network for Multimodal Brain Tumor Segmentation. Biomedical Signal Processing and Control, 79, 104037. https://doi.org/10.1016/j.bspc.2022.104037
[9] Bruno Machado Pacheco , Guilherme de Souza e Cassia b, Danilo Silva c Towards fully automated deep-learning-based brain tumor segmentation: Is brain extraction still necessary? https://www.sciencedirect.com/science/article/pii/S1746809422009685
[10] Allah, Ahmed M. Gab, Amany M. Sarhan, and Nada M. Elshennawy. "Edge U-Net: Brain tumor segmentation using MRI based on deep U-Net model with boundary information." Expert Systems with Applications 213 (2023): 118833.
[11] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv. https://doi.org/10.48550/arXiv.1505.04597
[12] Futrega, M., Milesi, A., Marcinkiewicz, M., & Ribalta, P. (2021). Optimized U-Net for Brain Tumor Segmentation. ArXiv. /abs/2110.03352