Brain Tumor Segmentation Using Deep Learning

B-tech VIth semester, Indian Institute of Information Technology Allahabad

Harsh Garg IIT2020082 Ramavath Vasanth Naik IIT2020096 Tarun Harishchandra Pal IIT2020098 Mohit Kumar Mina IIT2020100

***Abstract*—The paper discusses deep learning models for segmenting MRI images, specifically the UNET model for Brain Tumor Segmentation. Glioma, the most prevalent and dangerous type of brain tumor, can be life-threatening when its grade is high. Early detection of these tumors can improve and save the patient's life. Automatic segmentation of brain tumors from MRI scans plays a vital role in treatment planning and timely diagnosis. In simple UNET local features are lost during encoding or downsampling, resulting in constant learning as the model goes deeper. The use of spatial attention helps preserve local features, and the modified architecture involves additional modules in the encoding path which are concatenated during upsampling or decoding. Our architecture achieved results with the average dice score for the tumor core (TC), whole tumor (WT), and enhancing tumor (ET) on the BraTS 2020 dataset of 0.9027, 0.8868, and 0.9067, respectively. To demonstrate the robustness of the proposed model in real-world clinical settings, validation of the trained model on an external cohort is performed on the BraTS 2021 benchmark dataset. The achieved dice scores on the external cohort are 0.8039, 0.7640, and 0.6689 for TC, WT, and ET, respectively.**

I. Introduction

Brain tumors include the most threatening types of tumors around the world [1]. A brain tumor is an abnormal growth of cells within the brain or the surrounding tissues that can cause a range of neurological symptoms and potentially life-threatening complications.

Glioma, the most common primary brain tumor, occurs due to the carcinogenesis of glial cells in the spinal cord and brain. Brain tumors can be benign or malignant, and they can originate from different types of brain cells or from other organs that have spread to the brain . A brain tumor is an abnormal growth of cells within the brain or the surrounding tissues that can cause a range of neurological symptoms and potentially life-threatening complications. Glioma, the most common primary brain tumor, occurs due to the carcinogenesis of glial cells in the spinal cord and brain. Brain tumors can be benign or malignant, and they can originate from different types of brain cells or from other organs that have spread to the brain .

However, the manual segmentation and analysis of structural MRI images of brain tumors is an arduous and time-consuming task that, thus far, can only be accomplished by professional neuroradiologists [2] . Therefore, an automatic and robust brain tumor segmentation will have a significant impact on brain tumor diagnosis and treatment. There are many different deep learning methods to segment MRI images but all methods didn't result in the most accurate segmentation.

In this report we implemented 3D-UNET with Spatial Attention which gives better results than the normal 3D-UNET and Modified Unet 3D, The problem with simple UNET is that during encoding or downsampling as we perform different operations the local features got lost and we go deep in more layers the learning also becomes constant.So to not get lost these local features we use spatial attention [3] , it also has the same architecture as Res-UNET 3D but with some modifications. The encoding path involves some extra modules (min,max pooling) at each level , which helps in saving local features which of these modules are concatenated to the corresponding level during upsampling or decoding.

II. Related Work and LR

The Previous research included the SAResU-Net, TransBTS, the hypercolumn technique-based model, the Spatial Attention-based Efficiently Features Fusion Network, and the Deep Neural Network-Based Novel Mathematical Model. The findings suggest that a combination of different techniques, such as deep learning models, attention mechanisms, and multi-modal imaging, have shown promising results in improving brain tumor segmentation performance. The SAResU-Net combined shuffle attention and residual modules with a basic 3D U-Net and leveraged a self-ensemble module for improved performance.

TransBTS used a transformer-based architecture to process multimodal medical images and generate tumor segmentation masks. The hypercolumn technique based model employed attention modules and residual blocks for improved feature extraction and classification. The Spatial Attention-based Efficiently Features Fusion Network used a series of dilated multi-fiber units and spatial attention mechanisms for feature refinement and fusion. The Deep Neural Network-Based Novel Mathematical Model utilized a deep neural network for 3D brain tumor segmentation. The findings of the literature review have implications for future research and practice in the field of brain tumor segmentation.

The current state of the field suggests that there is still room for improvement, particularly in terms of increasing the accuracy and robustness of the models. Future research may focus on incorporating more advanced techniques, such as graph convolutional networks, adversarial training, and transfer learning, to further enhance the performance of brain tumor segmentation models. Additionally, the use of large, diverse, and well-annotated medical imaging datasets will be crucial for training and validating these models.

TABLE I
LITERATURE REVIEW

S.No.	Title	Method	Achieved Performance	Dataset	Disadvantages
1	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.	SAResUNet 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 can't interpret the modules. complex modules are used cause overfitting in data.	2019 and 2020 Brats Dataset
2	MBANet: A 3D convolutional neural network with multi-branch attention for brain tumor segmentation from MRI images. Biomedical Signal Processing and Control, 80, 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
3	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

4.	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 UNet model to handle the issue of vanishing gradient. The proposed model is designed to take advantage of 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
5.	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

III. Material Methods

We have a Brats2020 dataset which will be trained on a U-Net model with attention mechanisms for the segmentation of brain tumors in multi-modalities MRI images

- U-Net is one of the most popular architectures used for segmentation. It was designed for image segmentation in the biomedical field [9].
- In contracting path feature maps get spatially smaller, whereas in an expanding path, the feature maps are expanded back to their original size.
- It consists of a contraction path and an expansion path with skip connections.
- The encoder(left) part obtains the context information of the input image, and the decoder(right) part performs accurate segmentation.
- The Encoded part takes the 3D volume as input of size 128x128x128x3 .There are a total 4 layers including the input layer.
- Channel-wise attention mechanisms are applied at different levels of the model during encoding and decoding which help to not get lost features.

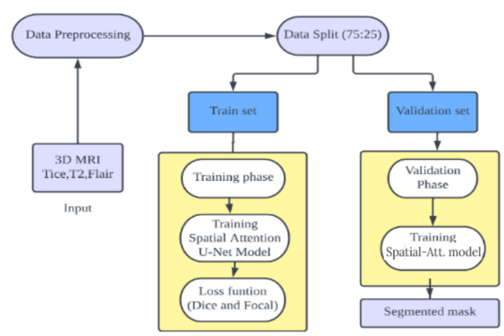

Fig. 1. The Proposed Methodology's General Flow [10]

- From the top to the bottom of the encoder, the number of feature maps increases and the size of feature maps reduces.
- The dimension of the feature map is doubled after each upsampling . The top of the decoder is the output layer consisting of a convolution layer and a sigmoid activation function.
- The testing was done on the Brats 2021 dataset.

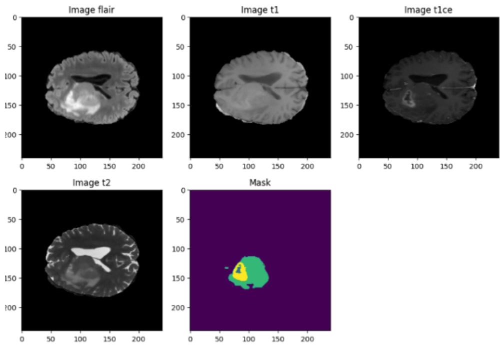

Fig. 2. Dataset of MultiModality MRI scans of Brain Dataset 2020

IV. Pre-processing

Custom generator is generated and preprocessing is done for the BraTS2020 dataset. The BraTS2020 dataset consists of MRI scans of brain tumors and their corresponding segmentation masks. Steps:

- Define a MinMaxScaler object which scales the intensity values of the MRI scans.
- The scaler is used to normalize the data to a range of 0 to 1 to improve the convergence of the deep learning model.
- The MRI scans are stored in separate lists for each image modality, i.e., T2, T1ce, and FLAIR, while the masks are stored in a separate list.
- Load the MRI scans and their corresponding masks using the nibabel library.
- Scale the intensity values of the MRI scans using the MinMaxScaler object
- Stack the three MRI scans into a single 3D image volume.
- Crop the image volume and its corresponding mask to remove any empty borders.
- Convert the mask to a one-hot encoded representation using the to_categorical function from tensorflowkerasutils.
- Save the preprocessed image and mask as NumPy arrays.
- Then the preprocessed data is split into training and validation sets with a ratio of 75:25.

V. Model Architecture

The U-Net model is a popular deep-learning architecture that has been widely used for semantic segmentation tasks, particularly in medical imaging. The architecture is designed to handle input data of varying sizes and has a contracting and expanding path that helps to capture both low and high-level features of the input data.

The contracting path of the U-Net model used in this report consists of four 3D convolutional layers followed by max-pooling layers. The purpose of the contracting path is to reduce the spatial resolution of the input while increasing the number of channels. Each block in the contracting path consists of two 3D convolutional layers that apply a set of learnable filters to the input feature map. The filters are applied to small patches of the input volume, and the resulting output is passed through a batch normalization layer to reduce internal covariate shifts. The activation function used in the model is the Rectified Linear Unit (ReLU) which introduces non-linearity into the model and helps to capture non-linear relationships in the input data.

The max-pooling layers in the contracting path help to reduce the spatial resolution of the feature maps by a factor of 2. This helps to reduce the computational complexity of the model while increasing the receptive field of the convolutional filters. By reducing the spatial resolution of the feature maps, the model is able to capture more global information about the input volume. After each max-pooling layer, a channel-wise attention mechanism is applied to improve the model's performance. The attention mechanism computes the mean and max pooling along the channel dimension and concatenates them. This concatenated feature map is then passed through a 3D convolutional layer with a softmax activation function to compute the channel attention weights. The channel attention weights are then applied to the input feature map using element-wise multiplication.

The expansion path of the U-Net model consists of four up-convolutional layers followed by concatenation with the corresponding layer in the contracting path. The purpose of the expansion path is to recover the spatial resolution of the feature maps while decreasing the number of channels. Each block in the expansion path consists of two 3D convolutional layers followed by batch normalization and ReLU activation function.

The first up-convolutional layer is followed by concatenation with the fourth layer in the contracting path, and so on. The concatenation operation helps to preserve the high-level features captured in the contracting path while recovering the spatial resolution of the feature maps.

Finally, a 1x1x1 3D convolutional layer with a sigmoid activation function is applied to obtain the segmentation mask. The sigmoid activation function maps the output of the convolutional layer to a probability distribution over the two classes (lesion and non-lesion). The segmentation mask is then thresholded to produce the final binary segmentation.

The model is trained using the Adam optimizer with a learning rate of 0.0001. The loss function used in the model is binary cross-entropy. The binary cross-entropy loss is a commonly used loss function for binary classification tasks. It measures the difference between the predicted probability distribution and the true probability distribution. The goal of the model during training is to minimize the binary cross-entropy loss to improve its accuracy in predicting the segmentation mask.

In summary, the U-Net model architecture used in this project is a powerful tool for performing semantic segmentation tasks on 3D medical imaging data. The model uses a spatial wise attention mechanism to improve accuracy.

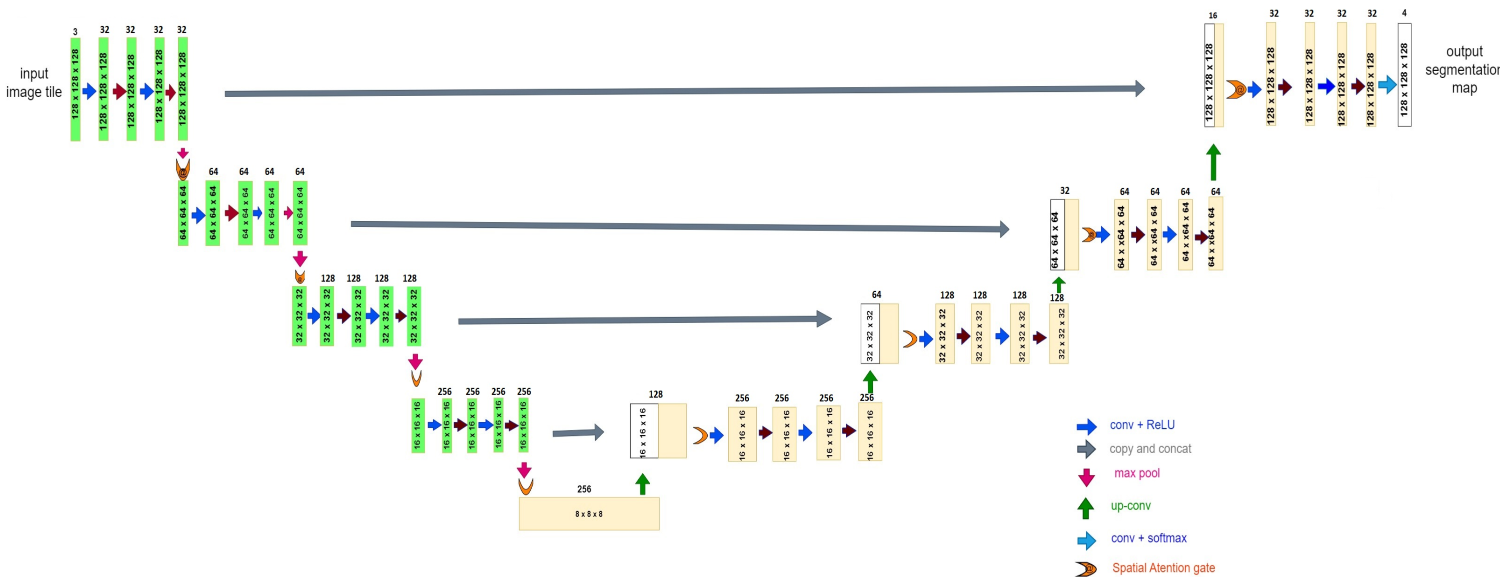

Fig. 3. Spatial attention Based Architecture

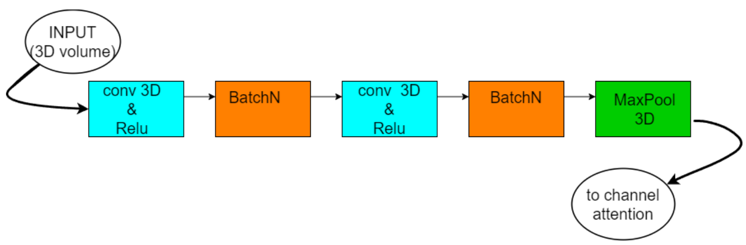

Fig. 4. Operations of Each layer of Contracting path

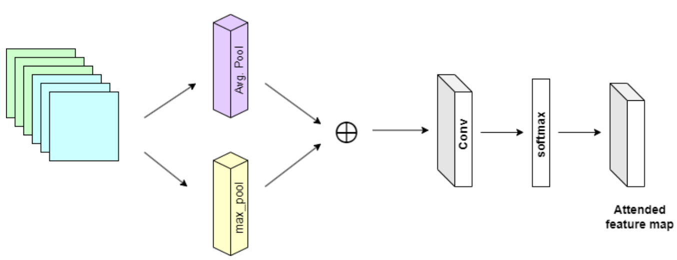

Fig. 5. Spatial Attention Module Overview

The attention mechanism in the model helps to improve its performance by highlighting the important features in the input. This helps the model to focus on relevant parts of the image and suppress irrelevant ones. The attention weights computed by the model are applied to the input feature maps using element-wise multiplication. This operation allows the model to assign higher weights to important features and lower weights to less important. In the upsampling path the batch normalization layer helps to improve the stability and convergence of the model during training by normalizing the input to the activation function. Finally, a 1x1x1 3D convolutional layer with a sigmoid activation fuction is applied to obtain the segmentation mask. The sigmoid activation function helps to ensure that the output values are btw 0 and 1,

VI. Loss Function,Metrics and Optimization

Loss Function:

– Dice Loss Function: Dice loss originates from Sørensen–Dice coefficient, which is a statistic developed in the 1940s to gauge the similarity between two samples [11]. It was brought to the computer vision community by Milletari et al. in 2016 for 3D medical image segmentation.

$$DiceLoss = 1 - (2 * TP)/(2 * TP + FP + FN) \quad (1)$$

where TP (True Positive) is the number of pixels that are correctly classified as positive, FP (False Positive) is the number of pixels that are incorrectly classified as positive, and FN (False Negative) is the number of pixels that are incorrectly classified as negative.

In segmentation tasks, TP represents the number of pixels that are correctly classified as part of the object of interest, and FN represents the number of pixels that are incorrectly classified as not part of the object of interest. FP represents the number of pixels that are incorrectly classified as part of the object of interest when they are not.

The Dice Loss is used to optimize the neural network model for segmentation tasks, where the goal is to maximize the overlap between the predicted segmentation mask and the ground truth mask. The Dice Loss is a differentiable and continuous function that can be used as the loss function during the training process of the neural network model.

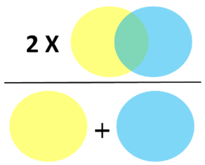

Fig. 6. Dice Loss Function Venn diagram [10]

In boundary detection tasks, real boundary pixels and predicted boundary pixels can be considered as two sets. By taking advantage of the loss of the dice, the two sets are trained to overlap little by little. As shown in the diagram above, the denominator considers the total number of boundary pixels globally, while the numerator considers the overlap between two local sets. Therefore, Dice loss takes into account both local and global loss information, which is essential for high accuracy.

Categorical Focal Loss: A Focal Loss function addresses class imbalance during training in tasks like object detection. Focal loss applies a modulating term to the cross entropy loss in order to focus learning on hard misclassified examples. It is a dynamically scaled cross-entropy loss, where the scaling factor decays to zero as confidence in the correct class increases. Intuitively, this scaling factor can automatically down-weight the contribution of easy examples during training and rapidly focus the model on hard examples.[12]

$$FocalLoss = -\alpha(1-p)^{\gamma}\log(p) \tag{2}$$

where is the balancing parameter, p is the predicted probability, and is the focusing parameter. The balancing parameter is used to balance the contribution of each class to the loss function, while the focusing parameter is used to control the rate at which the loss function focuses on hard-to-classify samples.

Metrics:
The total loss function is a combination of the Dice Loss and the Categorical Focal Loss. The weights of the two loss functions are adjusted using a hyperparameter to control the trade-off between them during the training process. In the stated paper,, the weight of the Categorical Focal Loss is set to 1.
Two evaluation metrics are defined - accuracy and IOU (Intersection over Union) Score. The IOU Score measures the overlap between predicted and ground truth masks and is a popular metric for segmentation tasks. The threshold parameter is set to 0.5, which means that any value above 0.5 is considered a positive prediction. The formula for Intersection over Union (IoU) Score is as follows:

$$IoUScore = TP/(TP+FP+FN) \tag{3}$$

where TP (True Positive) is the number of pixels that are correctly classified as part of the object of interest, FP (False Positive) is the number of pixels that are incorrectly classified as part of the object of interest, and FN (False Negative) is the number of pixels that are incorrectly classified as not part of the object of interest.

Optimiser

– Adam Optimiser :

The Adam optimizer is a popular optimization algorithm for training neural networks. It is a stochastic gradient descent method that uses adaptive learning rates to update the weights of the neural network. The algorithm combines the advantages of two other popular optimization algorithms, Adagrad and RMSprop. The Adam optimizer uses two momentum parameters, beta1 and beta2, to control the decay rate of the moving averages of the gradient and the squared gradient, respectively. The update rule for the weights is given by [13]:

$$w(t+1) = w(t) - \eta\frac{m(t)}{\sqrt{v(t)}+\epsilon} \tag{4}$$

where w(t) is the weight at time step t, is the learning rate, m(t) is the moving average of the gradient, v(t) is the moving average of the squared gradient, and epsilon is a small value added to the denominator to prevent division by zero. The Adam optimizer provides several advantages over other optimization algorithms. It is computationally efficient, requires minimal memory, and is well-suited for large-scale deep-learning applications. It is also robust to noisy gradients and can converge quickly to a good solution.

VII. Result and Analysis

In this section, the details about the evaluation measures employed to assess the performance of the proposed model, implementation details, and results achieved from the proposed method are discussed.

Evaluation Measure

- Dice coefficient score (DSC), specificity, and sensitivity are used to assess the proposed approach

$$Dice = 2TP/TP+FP+FN \tag{5}$$

- A binary cross-entropy loss function has been used.

VIII. Implementation Detail

- The proposed model (Modified-Net with spatial attention) was implemented using Python programming language, Keras library, and TensorFlow as the backend. The proposed model was implemented using Python programming language, Keras library, and TensorFlow as the backend.
- For the experimental purpose, an ADAM optimizer with a learning rate of 0.0001 was used. The activation function

ReLU with batch normalization was employed. Batch normalization normally increases the stability of the model and normalizes the network at each layer.
- The model was trained for 100 epochs on a batch size of 2 due to the limited computational resources. The experiments were conducted on the BraTS 2020 benchmark dataset, from which 75% of the data was used for training, and 25% data for validation.
- The testing was done on the Brats 2021 benchmark dataset, it was tested for the batch size of 6.

TABLE II
HYPERPAREMETERS OF PROPOSED SPATIAL ATTENTION MODEL

Hyperparameters	**values**
Input size	256 × 256 × 256 ×3
Learning rate	0.0001
batch Size	2
The hidden layer activation function	ReLU
Optimizer	ADAM
Loss function	binary cross-entropy loss function
No. of epochs	100
Dropout	0.1 - 0.2
Output size	128 × 128 × 128 × 3
Output layer activation function	Softmax

IX. ANALYSIS OF THE RESULTS

Training and Validation Details

The training loss per epoch and validation loss per epoch is shown with The Accuracy Per epoch of training and validation.

- The Training and validation was done on the Brats2020 dataset which consists of 369 images
- After preprocessing we got 344 useful images.
- These 344 further split into 75:25 ratio for training and validation that is 258 for training and 86 for validation

Testing details

- The Testing was done on the Brats2021 dataset which consists of 1250 images

TABLE III
RESULTS OF PROPOSED MODELS

*Models	Accuracy(2020)		IOU(2020)		IOU(Test)-2021
	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

TABLE IV
QUANTITATIVE RESULTS OF TRAIN AND TEST SET OF THE PROPOSED ARCHITECTURE

Models	Dataset	Tumor Core (TC)	Whole Tumor (WT)	Enhancing Tumor (ET)
Simple-Unet	Train-2020	0.8930	0.8684	0.8910
	Test-2021	0.7966	0.7495	0.6857
Modify-Unet	Train-2020	0.7331	0.7751	0.8700
	Test-2021	0.8206	0.7682	0.6630
Spatial-Unet	Train-2020	0.9027	0.8868	0.9067
	Test-2021	0.8039	0.7640	0.6689

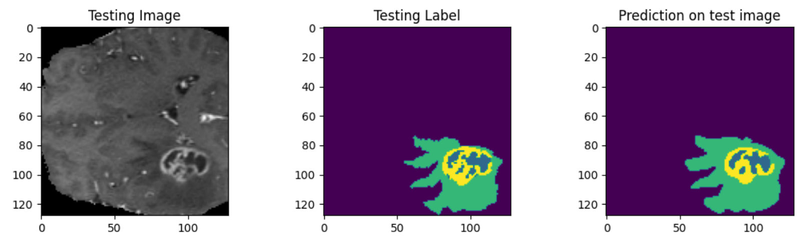

Fig. 7. Simple UNet Architecture Prediction Image

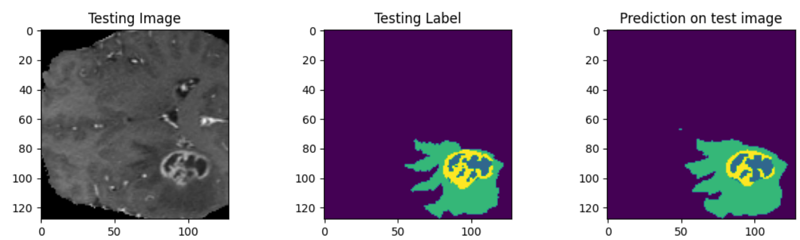

Fig. 8. Modified UNet Architecture Prediction Image

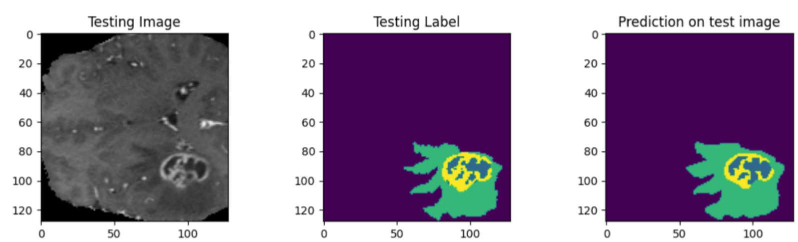

Fig. 9. Spatial Attention UNet Architecture Prediction Image

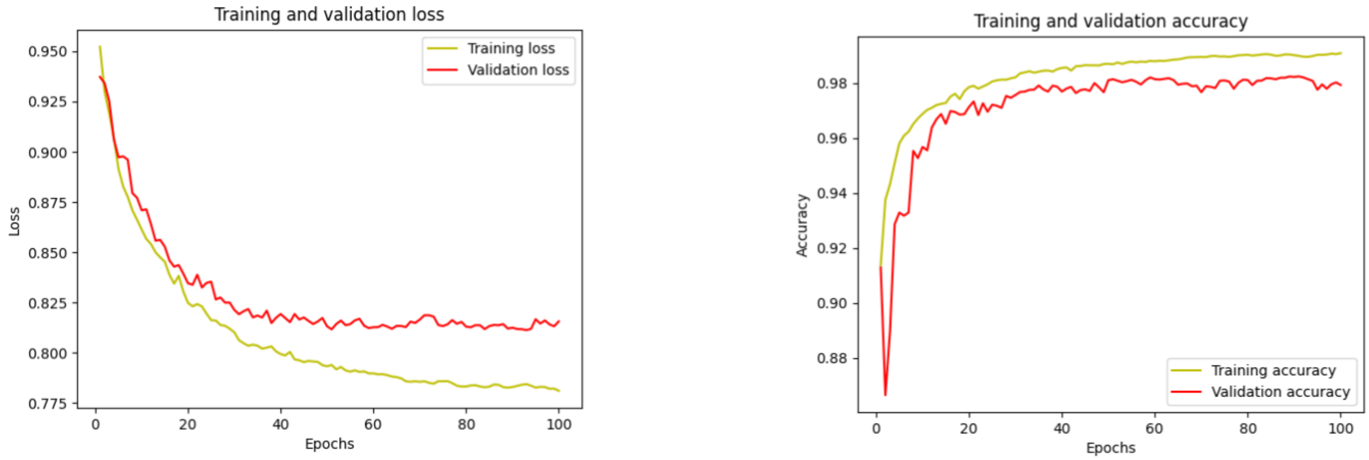

Fig. 10. Simple Unet Training And Validation Loss with Training and Validation accuracy

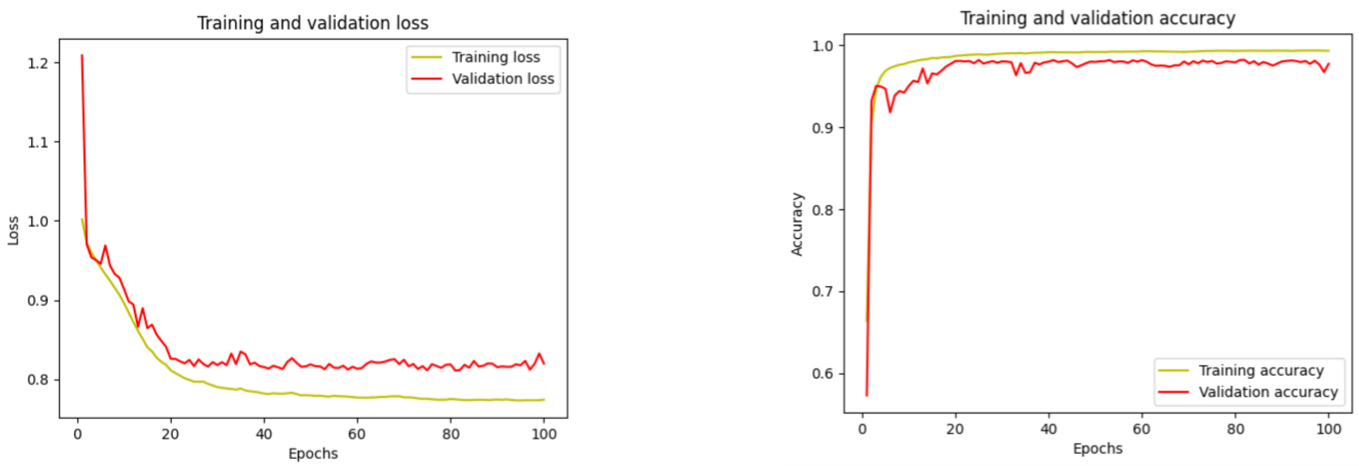

Fig. 11. Modified Unet Training And Validation Loss with Training and Validation accuracy

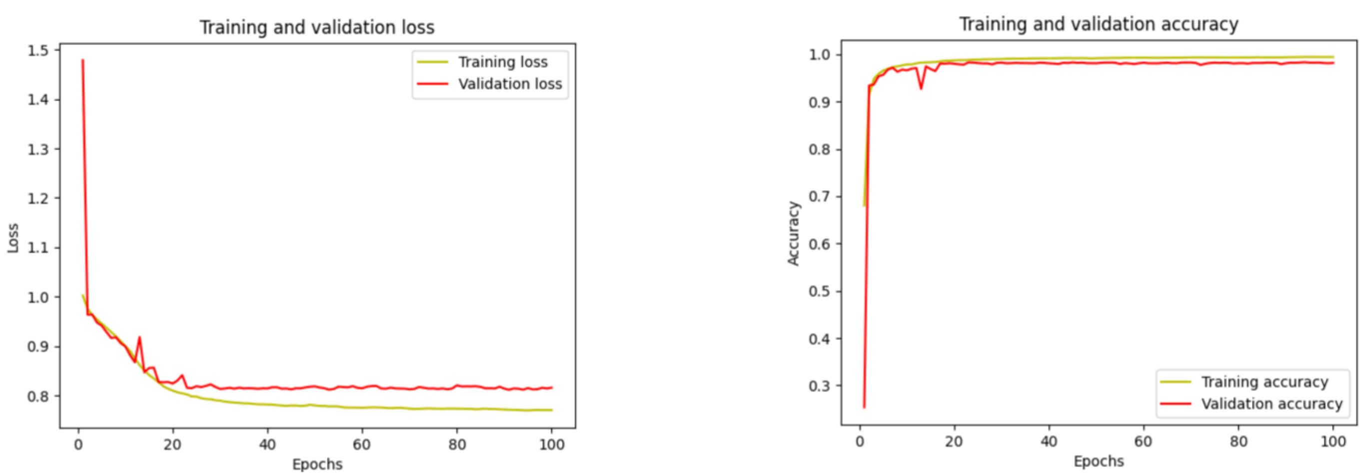

Fig. 12. Spatial attention Unet Training And Validation Loss with Training and Validation accuracy

X. 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. PMID: 11723374.

[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] 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.

[7] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv. https://doi.org/10.48550/arXiv.1505.04597

[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] Tie, Juhong; Peng, Hui; Zhou, Jiliu(2021).MRI Brain Tumor Segmentation Using 3D U-Net with Dense Encoder Blocks and Residual Decoder Blocks

[10]https://en.wikipedia.org/wiki/S%C3%B8rensen%E2 %80%93Dice$_coefficient$

[11]https://paperswithcode.com/method/focal-loss: :text=Focal%20loss%20applies%20a %20modulating,in%20the%20correct %20class%20increases.

[12]https://towardsdatascience.com/how-to-implement -an-adam-optimizer-from-scratch-76e7b217f1cc