

# Res-SANet: A Custom Deep Learning Model for Brain Tumor Segmentation

Nujhat Raisa Arpa  
2122152042

*Department of Electrical and Computer Engineering  
North South University  
Dhaka, Bangladesh  
nujhat.arpa@northsouth.edu*

Imroz Rahim  
2011292042

*Department of Electrical and Computer Engineering  
North South University  
Dhaka, Bangladesh  
imroz.rahim@northsouth.edu*

Mazharul Islam Nibir  
2121397042

*Department of Electrical and Computer Engineering  
North South University  
Dhaka, Bangladesh  
mazharul.nibir@northsouth.edu*

Shuvo Sarkar  
2121991042

*Department of Electrical and Computer Engineering  
North South University  
Dhaka, Bangladesh  
shuvo.sarkar@northsouth.edu*

**Abstract**—Around the world, brain tumors remain one of the leading causes of mortality. The lack of timely diagnosis is the leading cause of this problem, which is a world crisis. Moreover, brain tumor diagnosis is a critical process that depends on experts. Existing models do not provide efficient results related to radiologic diagnosis. So radiologists need a segmentation model to help them find brain tumors by segmenting the brain's medical images (MRI). Previous work in this field primarily relied on CNN-based models to segment regions of brain tumors, which typically do not provide satisfactory results to assist an expert. This research focused on comparing different segmentation models on a brain tumor dataset called BraTs23, which consists of 4 MRI modalities and includes three sub-regions: edema core, enhancing core, and necrotic core. Additionally, this study proposes a new model, ResSANet, which combines the UNet backbone SANet model with Residual Connections, a Dual-Attention mechanism, and creates a feature pyramid to capture multiscale contextual information. The model is assessed on the selected dataset containing high-quality brain MRI images and masks. Comparing the proposed model with other baseline models, the excellent outcome of this proposed model indicates that it is superior in terms of validation loss and dice coefficient for all individual sub-regions with a loss of 0.018, average dice score of 74%, dice score of edema of 82%, tumor enhancing dice score of 80% and necrotic dice score of 71% on BraTs23 dataset.

**Index Terms**—deep learning, brain tumor, image segmentation, BraTs 2023, preprocessing, ROI, 3D MRI images

## I. INTRODUCTION

The brain is the most advanced part of the human nervous system, and when the cells start to grow abruptly, voiding their natural growth pattern, a tumor originates [1]. A disruption in the growth cycle of brain cells leads to a rapid growth increase through abnormal cell division and turns into a solid mass of tissues that can be defined as tumors [2]. Depending on what tissue they arise from, the tumors vary. Primary brain tumors

arise from the brain, which stays within the brain. On the contrary, secondary brain tumors originate from somewhere else in the body and gradually spread to the brain [2]. Depending on the aggressiveness, brain tumors can be classified into benign tumors and malignant tumors. Malignant tumors can spread to neighboring healthy tissue and metastasize or form new tumors in distant body parts. In contrast, benign tumors can not invade nearby tissues and aren't capable of metastasizing. Once a benign tumor is removed, they do not regenerate; however, even benign tumors in vital organs such as the brain, if grown in size significantly, can pose a life-threatening risk to the patient [1], [2]. Radiologists primarily depend on MRIs for detecting brain tumors. Magnetic Resonance Imaging (MRI) is the conventional procedure for this. Radiologists pair T1-weighted (pre-contrast sequence) and T1Gd (post-contrast sequence) together to locate the tumor core. A combination of MRI modalities, such as T2-weighted and T2 Flair, defines the brain's water bodies [3]. Traditional treatments depend on the tumor sizes, locations, and stages of the tumors. Treatments include - surgical removal, radiotherapy, chemotherapy, immunotherapy, and targeted therapy [4]. However, to provide tumor patients with the necessary treatment, the doctor must locate the tumor as accurately as possible. According to a study, in 2019, approximately 347,992 cases of central nervous system cancers were reported globally. It was summarised using the Global Health Data Exchange query tool. In 2019, 347,992 (95% uncertainty interval) global cases of brain and CNS cancers were reported, which increased significantly from 1990 to 2019. The death toll in the same year was 246,253 due to the malignant tumor [5]. Although the tumor reports are low in low-income countries, the mortality rates are very high compared to those in high-income countries. This happens because of late diagnosis and poor healthcare systems [6].

In this paper, we employed five deep learning models, namely SAnet, Unet, Segnet, RAUnet, and Yolov8-seg, an efficient solution for segmentation tasks on brain tumor MRI images. Henceforth, the models will be denoted as SM5 models throughout the paper. The notable contributions of this paper are as follows:

- 1) Development of a dataset comprising 3D Brain MRI images: We worked on the BraTs competition dataset of 2023 named BraTS23, which consists of 3D MRI images, including masked images. We used this dataset to make a 2D dataset by making slices of those 3D MRI images.
- 2) Making an efficient preprocessed dataset: We focused on making the dataset more effective by using ROI (Region of Interest), which targets the most probable region of an MRI image to contain the tumor. We also worked on normalizing and denoising the photos so that ROI can give better results.
- 3) Tumor segmentation on the preprocessed dataset: In our analysis, only a few related studies have implemented ROI extraction to create an entirely new dataset for segmentation-related tasks, with nearly no implementations observed for BraTS2023. Moreover, our incentive for the study was to create a proposed model that beats the SM5 benchmark models.
- 4) Comparative analysis using different techniques: Using different preprocessing techniques or without techniques, this study provides a detailed analysis of the SAnet model to come up with the best combination of preprocessed techniques to work on.
- 5) Evaluation metrics to compare results: To assess the performance of the chosen models on the BraTS2023 dataset, we used several evaluation metrics such as accuracy, precision, specificity, sensitivity, dice score, mean-Iou, and loss. This study demonstrates dice coefficient scores for each category of tumor named edema, enhancing, and necrotic tumors. These metrics provide us with a way to compare the performance of the SM5 models.
- 6) Finally, this study provides a comprehensive knowledge of which technique and model combining gives the best results on our segmentation task.

The rest of the paper is organized as follows: Motivation is addressed in section (II), which mentions why we should work in this field. The literature review is discussed in section (III), outlines other works in this field, and gives an idea of how to improve our study by comparing all works. Section (IV) presents the methodology, which covers data acquisition, data preprocessing, all the deep learning models used, and the performance analysis methods that are used. Implementation & Evaluation are described in section (V), where implementation details and evaluation metrics are defined. Results & discussions are addressed in Section (VI), where all the results of the models are discussed by comparing each other and discussing how the experiments helped us fulfill the objectives of our

study. Finally, we conclude the work and discuss the future work in section (VII).

## II. MOTIVATION

Brain tumor is still one of the most prominent areas of research, and correct detection of this fatal disease is the most critical factor for research. It can affect people of all ages. The survival rate of this depends on the type, location, and stages of the tumor. It is estimated that about 23,820 adults and 5270 children were diagnosed with brain tumors in the United States alone in 2020 [7]. Around 250,000 people are diagnosed with brain tumors globally [8]. In aggressive brain tumors like-HGG, the survival rate is less than 10% [9]. These tumors affect other parts of the body, leading to severe morbidity and reducing the quality of life. Accurate segmentation of brain tumors using MRI images alone is a complex task because of the structure and appearance of the tumors. Tumors often have fuzzy borders, and they may spread into nearby regions of the brain, making it difficult to differentiate between healthy and metastatic tissue. Precise segmentation of gliomas from surrounding tissues is crucial for early detection as it can significantly improve the chances of survival. However, this remains a formidable task because predicting tumor location and grade is challenging. Automated segmentation tools would solve these problems and provide fast and accurate diagnosis.

## III. LITERATURE REVIEW

Jabbar et al. [1] proposed a hybrid model named Caps-VggNet that combined the CapsNet model with the VGGNet model by adding the layers of VGGNet. The authors used a combination of 2 datasets, the Brats Challenge 2019 dataset for training and the Brats Challenge 2020 dataset for testing, which includes MRI images of High-Grade Glioma (HGG) patients and Low-Grade Glioma (LGG) patients. Each MRI includes 4 modalities T1, T1ce (post contrast T1), T2 and T2 FLAIR. Annotation is done by experts and divided into three sub-regions (Whole Tumor, Tumor Core, and Enhanced Tumor). The authors used multi-grade segmentation to distinguish and label different tumor regions based on the types. First, The authors performed Data Preprocessing that includes methods like Z score normalization, skull stripping, median and bilateral filtering, removing noise, image cropping and resizing all images into 224x224 pixels and augmentation by performing rotations, scaling, flipping, and transformations. Next, they used GLCM (Grey level concurrence matrix) to extract critical information from the preprocessed images. Then, they implemented the hybrid learning model Caps-VGGNet for brain tumor classification, and they also used the transfer learning paradigm to train the model on the Brats 2020 dataset. Finally, they used the hybrid Caps-VGGNet model to classify the MRI images into High-Grade Glioma(HGG) and Low-Grade Glioma(LGG). The proposed model outperformed most traditional models like VGG16 and ResNet and achieved an accuracy of 99%, specificity of 99%, and sensitivity of 98% on the Brats 2020 dataset. This study required a decent level

of computer resources such as an Intel core i7 processor with RAM size of 32GB and GPU.

Roy et al. [10] introduced two CNN-based models, S-Net and SA-Net, expanding on the U-Net architecture for brain tumor segmentation. Their models incorporated 'Merge Blocks' for contextual integration and 'Attention Blocks' for region-specific focus using both hard and soft attention mechanisms. Preprocessing involved converting 3D to 2D images, Numpy array conversion, and threshold-based segmentation refinement. Evaluation metrics included accuracy, sensitivity, specificity, F-score, and Dice similarity coefficient.

Metlek and Çetiner [12] proposed a hybrid model, ResUNet+, which uses UNet as its backbone. On top of the UNet model, modifications like residual blocks and nodes were added between the encoder-decoder structures to prevent the vanishing of gradients and semantic gaps that often occur with other CNN models. Unlike traditional CNN models, they applied convolution on the tumor region only by extracting the region of interest (ROI). They achieved this by applying several different pre-processing techniques such as normalization, morphological operations, and transformations like highHat and lowHat, etc. They then subtracted the healthy tissue region by removing the background. This approach makes tumor segmentation more accurate but also uses significantly less computing power. They tested ResUNet+ on data from the BraTS competitions in 2018, 2019, and 2020, and it got promising scores: 92.80% for whole tumors, 93.10% for enhanced tumors, and 91.90% for tumor nuclei. These results show that ResUNet+ is just as good as, or even better than, a lot of top models out there.

J. Lin et al. [13] developed a hybrid model with multimodal cross-attention for brain tumor segmentation called "CKD-TransBTS: Clinical Knowledge-Driven Hybrid Transformer." The authors used the Brats Challenge 2021 private dataset, which includes 1251 3D MRI samples; each MRI includes four modalities. The mask annotation was split into three sub-categories: Enhancing Tumor (ET), Edema (ED), and the Necrotic Tumor Core (TC). First, the authors performed data preprocessing by skull stripping and resampling, image grouping, and data augmentations by random zooming, flipping, adding Gaussian noise, applying denoising with Gaussian blurring, and adjusting the pixel intensity levels with random contrast. Afterward, they implemented the CKD-TransBTS model with a double-branching hybrid encoder and a calibration-focused feature decoder to segment the brain tumors into specific sub-regions based on the annotations provided. A Modality-Correlated Cross-Attention (MCCA) block was used in the encoder to combine and extract the multimodal features. A Trans&CNN Feature Calibration (TCFC) block was used in the decoder to bridge the gap between the features of the transformer and CNN pipelines. Finally, the performance of the proposed model was evaluated using Dice Score, 95% Hausdorff distance, and sensitivity. The authors obtained a 0.9066 Dice Score for the overall segmentation performance of the proposed model across all tumor sub-regions. The authors

also compared the proposed model with several top-of-the-line models, including six CNN-based and six transformer-based models. Through extensive experiments, the authors showed that the proposed model was able to beat several baseline models.

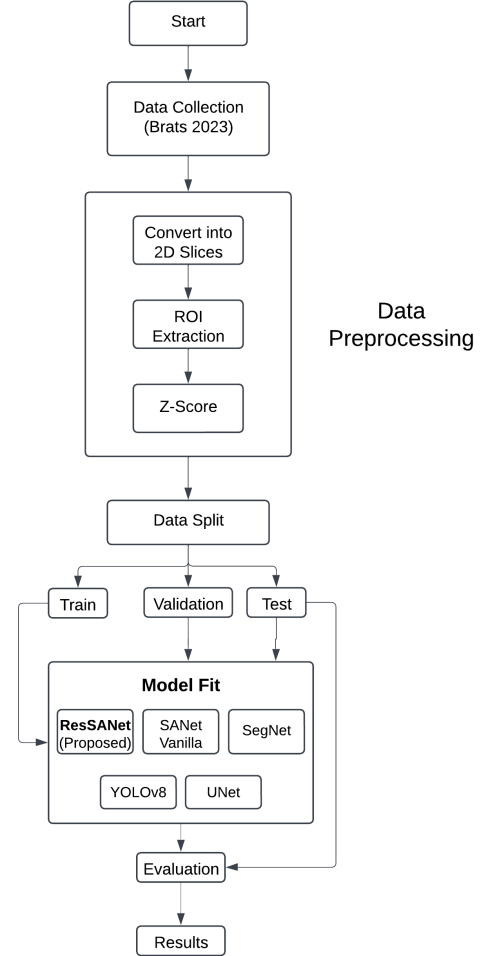


Fig. 1. Brain Tumor Segmentation Research Flow Diagram.

#### IV. METHODOLOGY

Fig. 1 represents the key steps of our research study. It starts with a thorough review of some relevant research papers from recent years; following that, the data collection process is done, which then goes through some preprocessing steps and selection. Finally, selected SM5 models are trained on the preprocessed data and evaluated using standard parameters by following these steps:

##### A. Dataset Acquisition

The BraTS Challenge 2023 dataset was used for this study. Table I summarizes the dataset, showing a total of 1251 3D brain MRI images. Each MRI sample has four modalities, which include T1, T1ce, T2, and Flair. It also includes their corresponding Ground Truth masks are shown in [Fig. 2].

Metric	Original (BraTS 2023)	Custom (ROI Applied)
Total Patient Samples	1251	1098
Modalities	4	4
Tumor Sub-Regions	3	3
Total Images	6255	5490
Segmentation Masks	1251	1098
T1 Contrast Images	1251	1098
T1 Native Images	1251	1098
T2 Flair Images	1251	1098
T2 Weighted Images	1251	1098

TABLE I  
COMPARISON OF ORIGINAL AND CUSTOM MEDICAL IMAGING DATASETS

The dataset comprises annotations for different tumor sub-regions, including Gd-enhancing tumor (ET), Necrotic Tumor Core (TC), and Edema (ED), which radiologists have identified based on VASARI features. We used two modalities out of the 4 - T2Flair along with T1ce, and we segmented them using the masks. We decided to take 80 out of 155 2D slices of the images within a range of 60-140 as this range contained the most prominent tumor images.

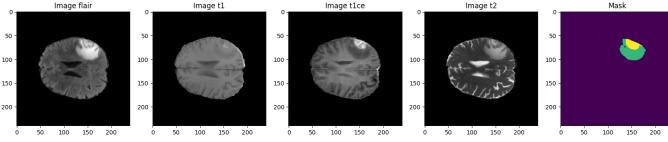


Fig. 2. Four Modalities & Corresponding mask

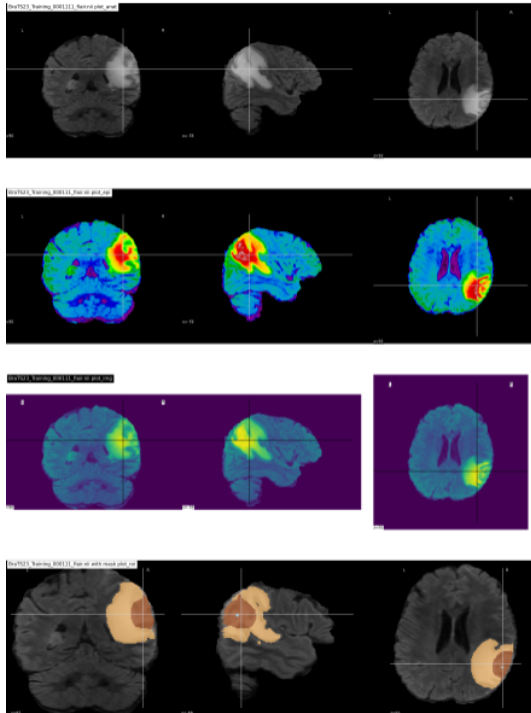


Fig. 3. Data visualization

### B. Data Preprocessing

One of the biggest challenges when segmenting tumors from MRI images is the artifacts caused by thermal noise and magnetic field interference. This noise can significantly reduce the performance of CNN-based approaches. For this reason, we employed several preprocessing methods to normalize the noisy images and boost their visibility. An exhaustive search was conducted to find the different combinations of preprocessing techniques that will generate more accurate results.

*a) Region Of Interest(ROI) extraction:* This was the first task as preprocessing was performed. In the dataset, not every brain MRI had tumor cells visible. As the segmentation was performed only on the brain tumor, the tumor part of the brain MRIs was the region of interest. ROI extraction technique was applied only on the brain MRIs that had tumors in them. The rest of them were kept as they were. The segmentation masks were in grayscale format. It was converted into a binary format so that the tumor boundaries were visible correctly. Morphological closing was also applied. The bounding box is expanded by 10 pixels so that no information about the tumor gets rejected. After refining the masks, the largest region was selected for tumor representation. This region in the corresponding T2 Flair and T1ce slices was detected. After detection, the rest of the area was isolated with the background(true black). A new custom dataset was created using this ROI extraction consisting of 746 patient samples. Original MRI image vs ROI image are shown in [Fig. 4]

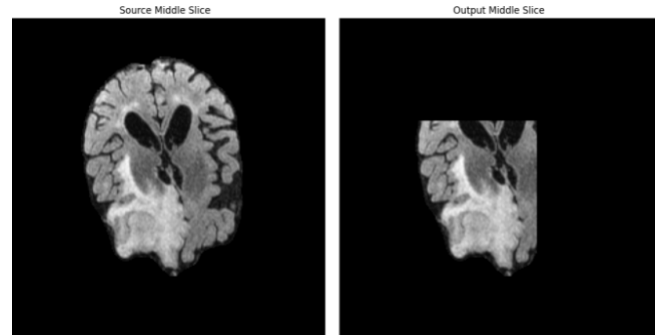


Fig. 4. Original MRI image vs ROI





Fig. 5. ROI vs mask

*b) Other Techniques:* Alongside ROI extraction, the other preprocessing methods involved Z-Score Normalization. We also experimented with CLAHE (Contrast Limited Adaptive Histogram Equalization) [Fig. 6], Histogram Equalization, Sharpening filters, and Denoising with Gaussian Blur. It was detected that if CLAHE and sharpening filters were used, it would affect the overall segmentation process. The boundary of the tumor region gets distorted, resulting in comparatively less accurate tumor segmentation. After a thorough experiment on the preprocessing techniques, it was seen that using only ROI extraction and Z-score normalization gives the best results. Besides these, resizing and cropping were also performed. Augmentation was not implemented as the dataset was already large and versatile enough.

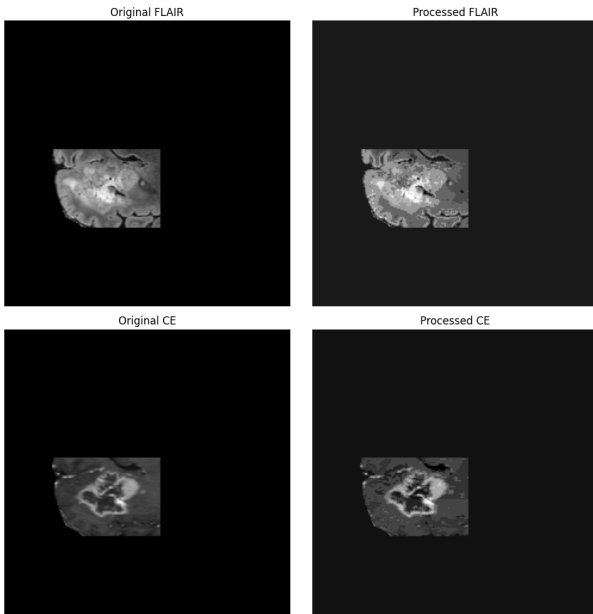


Fig. 6. Preprocessing technique CLAHE (CE) applied

### C. Model Architecture

**a) SANet:** The proposed model Res-SANet uses SANet as its backbone, which has some additional features. SANet is designed to conduct segmentation tasks. It uses an encoder-

decoder structure and a single attention mechanism to enhance feature extraction in the decoder phase.

- **Encoder Path:** This part of the model pipeline extracts features from the input through convolutional and pooling layers. Four convolutional blocks, each block with two convolutional layers, followed by max-pooling layers, were used for downsampling. Dropout was also applied after each block to prevent overfitting. This first convolutional layer uses 32 filters, which doubles and goes up to 256 filters in the final layer.
- **Bottleneck Layer:** It serves as a middle layer to connect the encoder and decoder path. The bottleneck layer uses a convolutional block with 512 filters and dropout to prevent overfitting. It serves as a connection between the contracting path and the expanding path, capturing the high-level features from the input.
- **Attention Mechanism:** It is used to enhance the feature selection by focusing on spatially important features. A single attention mechanism was used for the decoding phase.
- **Decoder Path:** Decoder reconstructs the resolution with upsampling followed by a convolutional block and then an attention block. In each upsampling layer, the dimension was doubled. Up to the final layer, the filter decreases from 256 to 32.
- **Output Layer:** This generates the final map. The output produces a multiclass segmentation.

**b) Res-SANet:** Our proposed model used the U-Net-like SANet model as its backbone. ResSANet adds the following modifications to the standard SANet architecture [Fig. 7],

- **Residual Connection:** Each convolution block integrates a residual connection, which acts as a shortcut path that prevents gradient vanishing and ensures efficient feature propagation to the deeper layers of the network.
- **Batch Normalization:** Batch Normalization was used after every convolution to stabilize the learning.
- **Feature Pyramid Network:** Inspired by our results from the YOLOv8 results, high-resolution features from early encoder layers were fused with upsampled deep features to enhance multiscale representation learning.
- **Attention Mechanism:** Instead of just using single attention, the ResSANet uses a dual-attention mechanism.
  - **Channel Attention:** It extracts global context using average pooling and max pooling.
  - **Spatial Attention:** It incorporates gating signals and spatial attention maps to refine feature localization.
- **Dropout Layers:** Dropout layers were used throughout the entire network to prevent overfitting and to improve generalization.

## V. IMPLEMENTATION & EVALUATION

### A. Implementation Details

This study was done using TensorFlow with the help of Kaggle utilizing an NVIDIA Tesla P100 GPU. BraTS2023 was

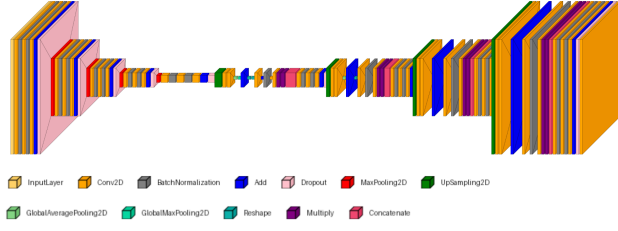


Fig. 7. Res-SAnet Model Architecture

used for the training of the proposed and other selected SM5 models. Training all of the models were initialized at epoch 0 with random weights without using any pre-trained weights. The data was split into two parts, 80% for training and 20% for validation, and then the 80% training data was again split into two parts, leaving a final distribution of 68% training, 20% validation and 12% testing [Fig. 8]. The images of the modalities were 240x240x155 in terms of dimensions. The 3rd dimension signifies the number of slices. While segmenting the brain tumors, using only one modality is not enough since the boundary between the Whole Tumor, Enhancing Tumor, Tumor Core( also known as Necrotic Tissue), and Edema regions along with healthy tissue are not clear. ROI extraction and other preprocessing techniques, especially, Z-Score normalization was employed to identify the tumor-affected areas from all the modalities, and then the T2Flair and T1ce were used as the FLAIR modality performed excellently at differentiating edema regions from other parts of the tumor and T1ce was used to make the borders more visible, effectively making use of multi-modal data. After obtaining these ROI images and having applied the preprocessing, the model was then ready to be trained. To achieve the best results, we trained the model using several hyperparameters. The initial learning rate was dynamically adjusted using ReduceLROnPlateau, reducing by a factor of 0.5 when validation loss plateaus with a threshold of 1e-6 to stabilize the convergence. The training progress was monitored using the Validation Dice Coefficient value.

### B. Evaluation Metrics

Brain Tumor Segmentation often focuses only on segmenting the whole tumor region. This study aims to segment several parts of the tumor region, including the Enhancing Tumor, Necrotic Tissue, and Edema regions. We have obtained Dice scores for each tumor region separately. The Dice Coefficient represents the spatial overlap between the segmented and Ground Truth results.

For each of our tests in this study, we used several accuracy metrics, named accuracy, precision, specificity, sensitivity, dice coefficient, and mean-iou (Jaccard index), to evaluate

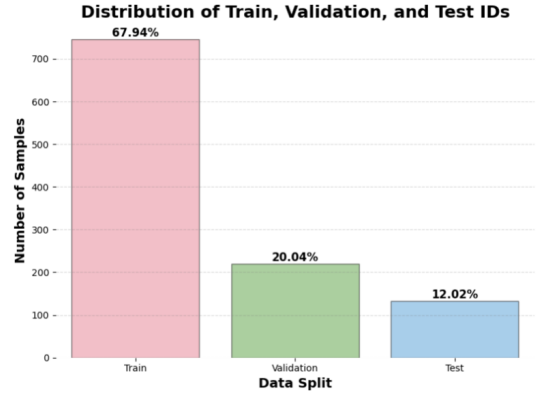


Fig. 8. Data distribution of the custom dataset

our models for segmentation tasks. The mentioned metrics, especially dice score and Jaccard index, are commonly used for segmentation tasks in computer vision and deep learning.

- **Accuracy:** Accuracy refers to the ratio of correct predictions to the number of predictions. Which can be expressed as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

- **Precision:** Precision can be described as the ratio of actual predicted positive cases to total positive cases. Also known as PPV or positive predictive values. Which can be expressed as:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

- **Sensitivity:** Sensitivity is defined as the proportion of the true positive cases that are correctly identified. Which can be expressed as:

$$Sensitivity = \frac{TP}{TP + FN} \quad (3)$$

- **Specificity:** Specificity is the actual negative rate that is defined as the proportion of true negative cases that are predicted correctly as negative. Which can be expressed as:

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

- **Dice Coefficient:** The Dice Coefficient quantifies the similarity between the predicted and actual positive cases. Which can be expressed as:

$$Dice = \frac{2TP}{FP + 2TP + FN} \quad (5)$$

Class-Specific Dice Coefficients are used to evaluate the model's performance in specific classes.

- **Mean Jaccard Index:** Mean-iou (Jaccard Index) measures the level of overlap between the predicted positive and actual positives. It compares the common positive

TABLE II  
MODEL PERFORMANCE METRICS

Models	Accuracy	Precision	Loss	Dice-Coeff	Dice Coefficient Metrics			Sensitivity	Specificity
					Edema	Enhancing	Necrotic		
Res-SAnet	99.46%	99.42%	<b>0.018</b>	<b>0.74</b>	<b>0.82</b>	<b>0.80</b>	<b>0.71</b>	99.36%	99.81%
SAnet	99.44%	99.47%	0.05	0.68	0.81	0.78	0.63	99.29%	99.82%
RAUnet	99.45%	99.41%	0.02	0.73	0.79	0.79	0.70	99.38%	99.80%
Segnet	98.98%	99.67%	0.03	0.56	0.71	0.53	0.45	98.85%	99.67%
Unet	98.98%	99.12%	0.03	0.53	0.64	0.64	0.43	98.74%	99.70%
Yolov8-seg	99.27%	99.25%	0.02	0.67	0.77	0.69	0.56	99.19%	99.75%

instances to all other predicted and actual positives. Which can be expressed as:

$$Jaccard = \frac{TP}{TP + FP + FN} \quad (6)$$

## VI. RESULTS & DISCUSSION

### A. Results

This section discusses the result of the proposed study. This study benchmarked the proposed model against several other models, such as - SAnet, RAUnet, SegNet, Unet, and YOLOv8-seg (SM5). We used the models that use the UNet as the backbone because it gave the best results in the field of medical imaging segmentation tasks. In Table II, the models were compared using the performance metrics such as - Accuracy, Loss, Dice-Coefficient, Dice-Coefficient-Edema, Dice-Coefficient-Enhancing, Dice-Coefficient-Necrotic, Precision, Specificity, and Sensitivity. The conclusion was measured following the observations based on the results of Table II. Dice-Coefficient was the parameter whose score was observed to highlight the model's behavior. The Dice-Coefficient measures the segmentation performance of the selected (SM5) models respectively - 0.68, 0.73, 0.56, 0.00, and 0.67. Among these, the proposed model gives comparatively higher values in almost all the parameters. Res-SAnet achieved 0.74 as a Dice-Coefficient score, which exceeds the score of all other models that were compared [Fig. 10]. Res-SAnet exceeds SAnet, RAUnet, Segnet, Unet, and YOLOv8-seg, respectively, by - 8.4%, 1.36%, 27.6%, 0%, and 9.9%.

### B. Discussion

Res-SAnet performs better because of some key additional changes it has in its model architecture. It uses the residual connection, which ensures efficient feature propagation. The model had better stabilization as Batch Normalization was used. A feature pyramid network helped to upsample the encoder layer. The prominent addition was the dual-attention mechanism, which extracts both global features and local features. Moreover, in SAnet's paper, the authors ran the model over 100 epochs. On the other hand, Res-SAnet outperformed their score only after 50 epochs.

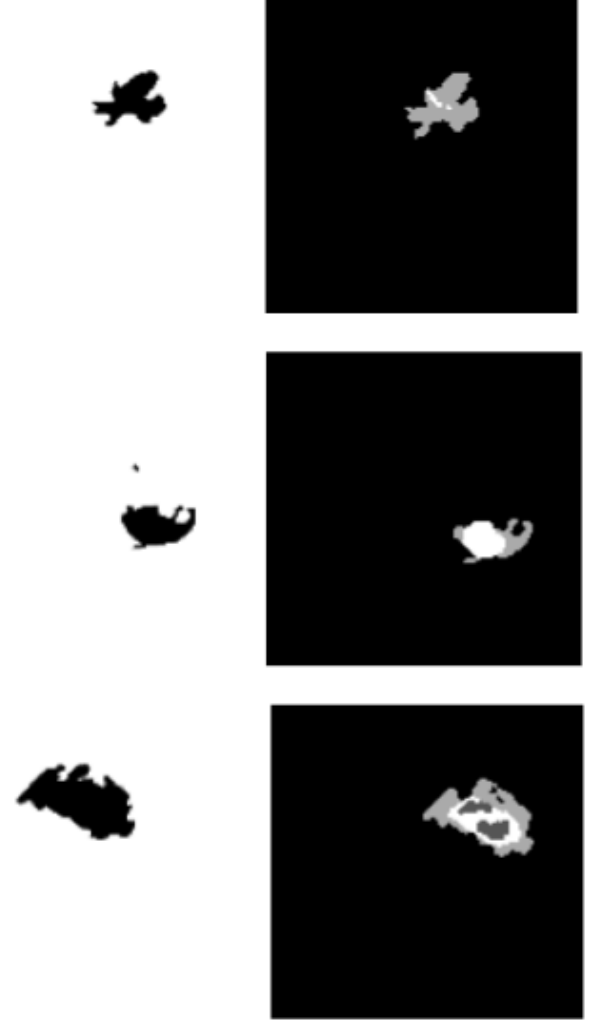
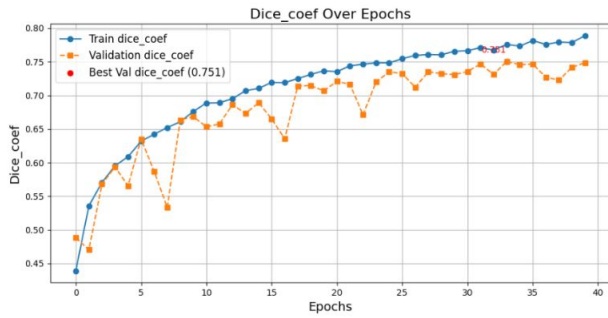


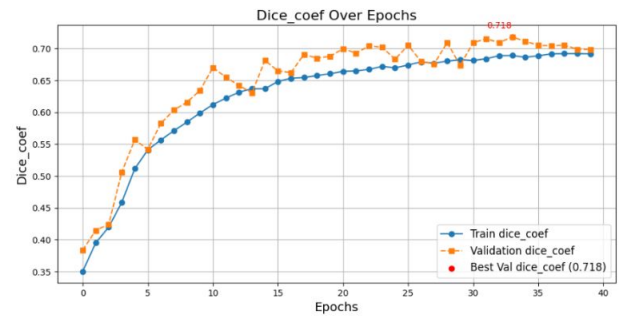
Fig. 9. Binary formatted original Mask VS generated mask by the Res-SAnet model

## VII. CONCLUSION & FUTURE WORK

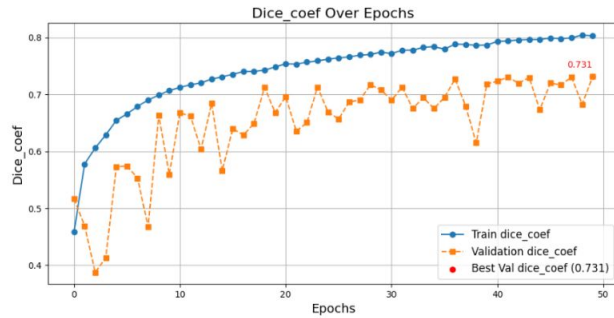
In this study, we demonstrated the capability and efficiency of our proposed model, ResSAnet, for brain tumor segmentation, achieving promising results and beating the other benchmark models in all Dice score categories. The concepts of Residual Connections and Dual-Attention (Channel and Spatial) were introduced to achieve even higher performance than



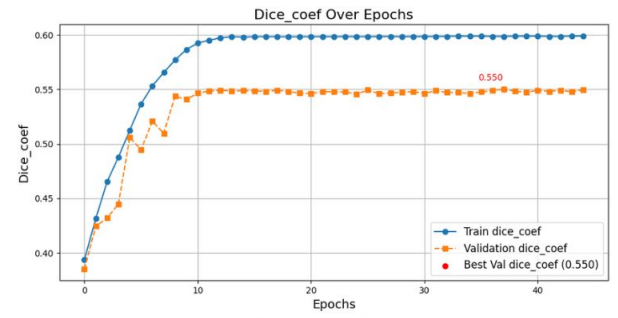
**Res-SANet**



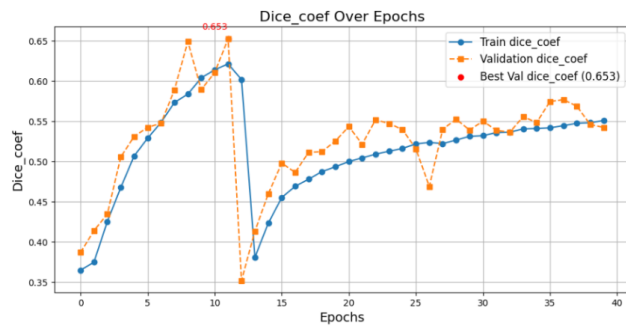
**SANet**



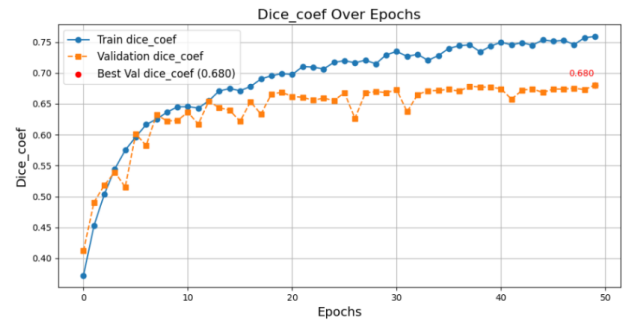
**RAUnet**



**SEGnet**



**Unet**



**YOLOv8-seg**

Fig. 10. Performance across all models

the Baseline SANet architecture. In our study, we faced several challenges, such as the limited availability of computational resources; hence, the experiments conducted were limited to 50 epochs. In the future, we intend to do more exhaustive training by increasing the number of epochs. We plan to experiment with augmentation strategies and further refine the model architecture to better capture complex tumor shapes. We can also fine-tune hyperparameters to achieve even better results. We also intend to extend this study by incorporating all four modalities instead of 2. To ensure a comprehensive evaluation, we expand the evaluation metrics by incorporating the Hausdorff Distance 95th Percentile (HD95). Also, we can compare the time performance of all the models we will run. Furthermore, future experiments will be conducted on the BraTS2024 dataset to evaluate the adaptability and robustness

of the model. Additionally, we intend to explore 3D-based networks. We will also incorporate XAI (Explainable AI) techniques to provide a better understanding of the model's decision-making.

## REFERENCES

- [1] A. Jabbar, S. Naseem, T. Mahmood, T. Saba, F. S. Alamri and A. Rehman, "Brain Tumor Detection and Multi-Grade Segmentation Through Hybrid Caps-VGGNet Model," *IEEE Access*, vol. 11, pp. 72518-72536, 2023.
- [2] M. F. Almufareh, M. Imran, A. Khan, M. Humayun, and M. Asim, "Automated Brain Tumor Segmentation and Classification in MRI Using YOLO-Based Deep Learning," *IEEE Access*, vol. 12, pp. 16189-16207, 2024.
- [3] Y. Zhang, Y. Liu, Y. Wang, Y. Zhang, and J. Liu, "CKD-TransBTS: Clinical Knowledge-Driven Hybrid Transformer With Modality-Correlated



Cross-Attention for Brain Tumor Segmentation,” IEEE Transactions on Medical Imaging, vol. 42, no. 8, pp. 2452-2463, 2023.

- [4] A. Rodríguez-Camacho, J.G. Flores-Vázquez, J. Moscardini-Martelli, J.A. Torres-Ríos, A. Olmos-Guzmán, C.S. Ortiz-Arce, D.R. Cid-Sánchez, S.R. Pérez, M.D.S. Macías-González, L.C. Hernández-Sánchez, et al., “The Landscape of Novel Therapeutics and Challenges in Glioblastoma Multiforme: Contemporary State and Future Directions,” *International Journal of Molecular Sciences*, vol. 23, no. 7207, 2022.
- [5] Y. Fan, X. Zhang, C. Gao, S. Jiang, H. Wu, Z. Liu, and T. Dou, “Burden and trends of brain and central nervous system cancer from 1990 to 2019 at the global, regional, and country levels,” *Archives of Public Health*, vol. 80, no. 1, pp. 1–14, 2022.
- [6] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, “Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality,” *CA: A Cancer Journal for Clinicians*, vol. 74, no. 6, pp. 400–421, Nov. 2024.
- [7] D. Daimary, M. B. Bora, K. Amitab, D. Kandar, “Brain Tumor Segmentation from MRI Images using Hybrid Convolutional Neural Networks,” *Procedia Computer Science*, Volume 167, 2020
- [8] P. Jain and S. Santhanalakshmi, “Early Detection of Brain Tumor and Survival Prediction Using Deep Learning and An Ensemble Learning from Radiomics Images,” *IEEE 3rd Global Conference for Advancement in Technology (GCAT)*, Bangalore, India, 2022
- [9] S. Solanki, U. P. Singh, S. S. Chouhan and S. Jain, “Brain Tumor Detection and Classification Using Intelligence Techniques: An Overview,” *IEEE Access*, vol. 11, pp. 12870-12886, 2023
- [10] S. Roy, R. Saha, S. Sarkar, R. Mehera, R. K. Pal and S. K. Bandyopadhyay, “Brain Tumour Segmentation Using S-Net and SA-Net” *IEEE Access*, vol. 11, pp. 28658-28679, 2023
- [11] H. Çetiner and S. Metlek, “DenseUNet+: A novel hybrid segmentation approach based on multi-modality images for brain tumor segmentation,” *Journal of King Saud University - Computer and Information Sciences*, vol. 35, 2023
- [12] S. Metlek and H. Çetiner, “ResUNet+: A New Convolutional and Attention Block-Based Approach for Brain Tumor Segmentation,” *IEEE Access*, vol. 11, pp. 69884-69902, 2023.
- [13] J. Lin et al., “CKD-TransBTS: Clinical Knowledge-Driven Hybrid Transformer With Modality-Correlated Cross-Attention for Brain Tumor Segmentation,” in *IEEE Transactions on Medical Imaging*, vol. 42, no. 8, pp. 2451-2461, Aug. 2023