Brain Tumor Segmentation with W-net and Deep Neural Network Strategies

M. JASMINE PEMEENA PRIYADARSINI
School of Electronics Engineering
Vellore Institute of Technology
Vellore, India
jasmin@vit.ac.in

Chetan Sai Borra
School of Electronics Engineering
Vellore Institute of Technology
Vellore, India
borrachetan.sai2020@vitstudent.ac.in

Jaswanth Gurujala
School of Electronics Engineering
Vellore Institute of Technology
Vellore, India
gurujala.jaswanth2020@vitstudent.ac.in

Gopireddy Harsha Vardhan Reddy
School of Electronics Engineering
Vellore Institute of Technology
Vellore, India
gopi.harshavardhan2020@vitstudent.ac.in

R. PREETHA
School of Electronics Engineering
Vellore Institute of Technology
Vellore, India
email address or ORCID

***Abstract*—Early detection of brain tumors is crucial for patient health and survival, given the life-threatening risks posed by higher grades. Magnetic Resonance Imaging (MRI) plays a vital role in this process, yet segmenting healthy brain cells from tumors proves challenging due to the vast MRI data and variations in tumor characteristics. Therefore, an accurate and automated segmentation method is essential for effective treatment planning and diagnosis. This study evaluates various convolutional neural networks and deep neural network architectures such as Enhanced U-Net, VGG, and W-Net on the BRATS 2020 Dataset, aiming to enhance the accuracy of tumor segmentation from MRI scans.**

***Index Terms*—Brain tumor, Convolutional neural networks, Deep neural networks, Automated segmentation, U-Net, W-Net.**

I. Introduction

Brain tumors pose a major health risk due to their complexity, diverse manifestations, and potentially fatal consequences. Detecting them early and planning accurate treatments are vital for better patient outcomes. A critical approach to diagnosing and evaluating brain tumors is through medical imaging, especially Magnetic Resonance Imaging (MRI). However, extracting brain tumor segments from MRI data is complex because of differences in tumor shape, size, location, and type. This is where sophisticated deep learning models, like U-Net and W-Net, play a pivotal role in effective tumor segmentation.

The Brain Tumor Segmentation (BRATS) Challenge provides a benchmark dataset for evaluating different segmentation approaches. The BRATS 2020 dataset contains multimodal MRI scans with annotations for various tumor types, making it a useful resource resource for research in the field. It includes four MRI modalities: T1, T1ce (contrast-enhanced), T2, and FLAIR, allowing for comprehensive analysis and segmentation.

U-Net, introduced by Ronneberger et al. [3], is a popular convolutional neural network architecture designed for biomedical image segmentation. Its distinctive structure, featuring a contracting path (encoder) and an expansive path (decoder), along with skip connections, allows for detailed segmentation with high accuracy. The U-Net architecture has been widely used in medical imaging tasks due to its efficiency and flexibility[3].

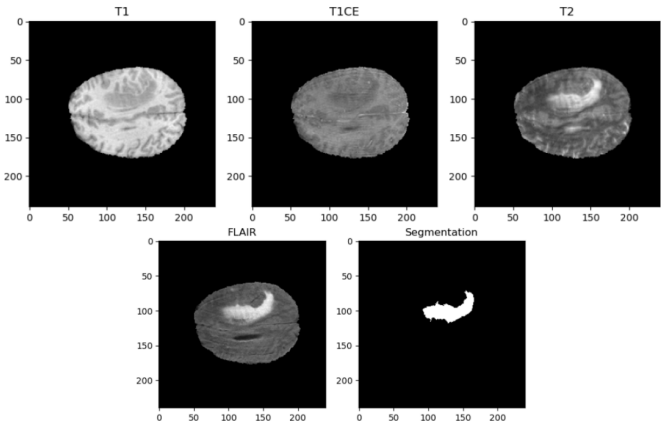

Fig. 1. These are the different modalities in the BRATS 2020

W-Net, on the other hand, is an extension of the U-Net architecture with a double U-shaped structure, providing additional capacity for feature extraction and detailed segmentation. W-Net can process 2D images, predict brain tumor segmentation, and then recombine the results into a single output image with high precision.

Other researchers have also explored variations of the U-Net architecture to enhance segmentation accuracy. For example, Li et al. [9] introduced a dense U-Net structure that utilizes dense connections to improve feature reuse, and Isensee et al. [10] proposed nnU-Net, which automates U-Net adaptation to different segmentation tasks. These developments indicate the ongoing evolution of neural network architectures for medical image segmentation.

II. Exploring Dataset

A. BraTS2020

BraTS 2020 multimodal scans are accessible as NIfTI files (.nii.gz) and include descriptions of a) standard T1-weighted images, b) post-contrast T1-weighted (T1Gd) images, c) T2-weighted images (T2), and d) T2 Fluid Attenuated Inversion Recovery (T2-FLAIR) images. Each imaging dataset underwent manual segmentation by one to four raters, adhering to a consistent annotation protocol. Experienced board-certified neuro-radiologists approved the annotations, which include delineations of the GD-enhancing tumor (ET — labeled 4), peritumoral edematous/invaded tissue (ED — labeled 2), and necrotic tumor core (NCR — labeled 1)as detailed in [1] , [2]and [11].

The dataset includes distinct directories for Training and Validation. In the Training folder, there are 369 individual brain scans, each containing four modalities (T1, T1ce, T2, Flair) along with a segmentation mask in the .nii file format. Each scan comprises images of the brain with dimensions of 240*240*155, and it offers three distinct views: Axial (240*240), Coronal (240*155), and Sagittal View (240*155).

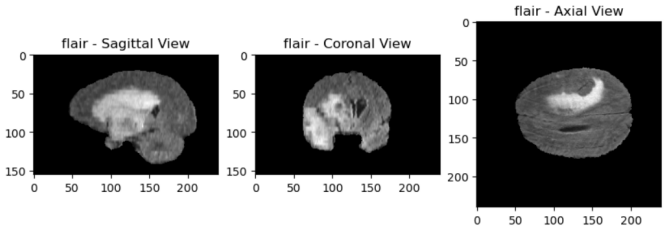

Fig. 2. An example of the brain scan depicted from three distinct perspectives.

For our CNN models, we utilize 75 images from each modality, spanning from the 60th to the 135th slice. This selection amounts to a total of 27,675 images per modality in the axial view. These modality images will be resized to 128*128 for training, validation, and testing purposes across different models. Additionally, the segmentation masks are resized to match the 128*128 dimensions for training.

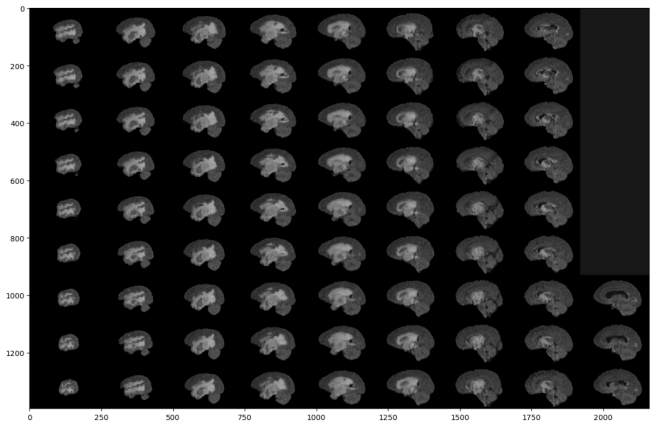

Fig. 3. The images of the brain scans from 60 to 135 slices in Flair.

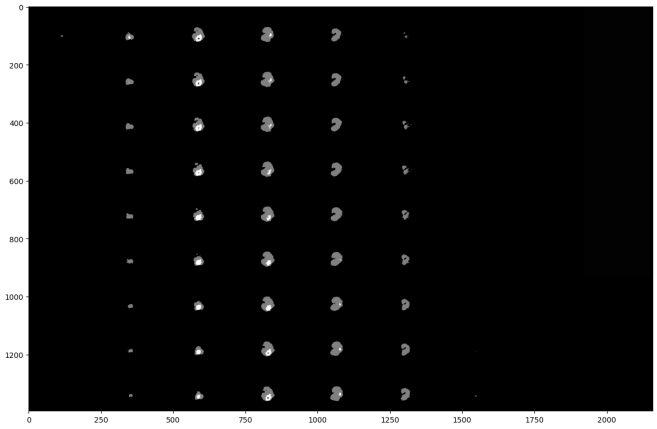

Fig. 4. The images of masks from 60 to 135 slices.

III. Related Work

The U-Net was first introduced by Olaf Ronneberger, Philipp Fischer, and Thomas Brox in 2015, as detailed in their paper titled "U-Net: Convolutional Networks for Biomedical Image Segmentation"[3]. The U-Net architecture [in figure-5] is distinctive for its structure comprising a contracting path and an expansive path. In the contracting path, encoder layers are employed to gather contextual details and diminish the spatial resolution of the input. Conversely, the expansive path integrates decoder layers to reconstruct the encoded data, leveraging information from the contracting path through skip connections to produce a segmentation map.

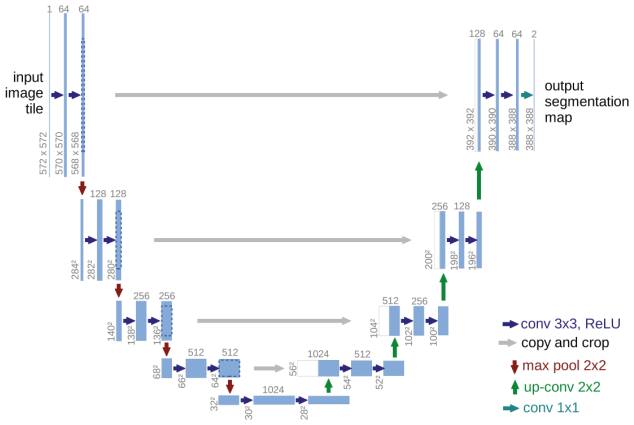

Fig. 5. Basic U-Net block diagram [3]

Diakogiannis and colleagues [4] introduced ResUNet, a groundbreaking deep learning framework that integrates principles from various advanced computer vision modules utilized in semantic segmentation tasks. They investigated the effectiveness of several versions of the generalized dice loss for semantic segmentation and developed a specialized variation loss function tailored for object semantic segmentation. This function demonstrates improved convergence characteristics and maintains strong performance, even in scenarios where classes are heavily imbalanced.

Z. Sobhaninia and his team have explored a variety of neural network architectures for brain tumor segmentation and related tasks. In one study, they employed LinkNet, a lightweight neural network architecture, to generate segmentation masks [7]. In another work [5], they utilized U-Net, a well-known architecture for biomedical image segmentation, with a modified algorithm to improve segmentation results. Additionally, they incorporated a Modified Self-Organizing Feature Map (SOFM) to predict patient survival outcomes.

Furthermore, in a separate study [8], Minhazur Rahman's team experimented with the SC-SE U-Net architecture, which integrates an Attention Mechanism with Squeeze-and-Excitation Networks. This approach was tested on the TCGA-LGG dataset, a publicly available dataset that contains lower-grade glioma data. The combination of Attention Mechanisms with Squeeze-and-Excitation helps the network focus on significant features during the segmentation process, potentially improving the accuracy and precision of the results.

IV. Proposed Methodologies

This study introduces a new architecture inspired by W-Net [6], utilizing a feature extraction channel. The design of the proposed network is depicted in Figure-5. The network comprises 18 modules, including encoder and decoder blocks. Each module contains two 2D convolutional layers. The encoder modules are followed by max pooling to reduce the image size, while the decoder modules are followed by an upsampling operation. The encoder layers start with 32 filters, doubling at each step until reaching 1024 filters in the bottommost layer.

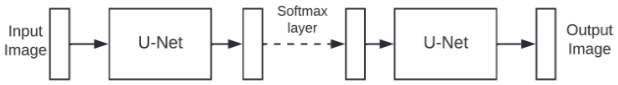

Fig. 6. Basic W-Net block diagram [6]

The dataset is divided into training samples, validation samples, and testing samples with a 68:12:20 ratio, leading to 250 samples for training, 74 for validation, and 45 for testing. The training was conducted over 70 to 100 epochs, enabling the loss to reach its minimum and allowing various parameters to be calculated.

A. Evaluation Measures and parameters.

The "Sørensen–Dice coefficient," developed by Lee R. Dice and Thorvald Sørensen in 1945, has gained traction due to its simplicity and effectiveness. The Dice coefficient measures the similarity between two sets, A and B, with values ranging from 0 to 1. A score of 0 means there's no overlap between the sets, while a score of 1 indicates that the sets are identical. The count of elements common to both sets is denoted by $|A \cap B|$.

$$Dice\ coefficient = \frac{2 * |A \cap B|}{(|A| + |B|)}$$

Mean IOU: When comparing data samples in machine learning, one of the most basic techniques is the intersection over union (IOU) method. In statistics, the Jaccard Index is often referred to as IoU. IoU is a technique used for evaluating computer vision tasks such as object tracking, object recognition, and semantic segmentation. In essence, IoU assesses the degree of overlap between two sets of elements. By employing bounding box regression, where A denotes the ground truth box and B represents the prediction box, IoU is calculated based on the coordinates, width, and height of A's bounding box.

$$Mean\ IOU = \frac{Areaof Overlap}{Areaof Union}$$

V. Results Observations

Initially, we started by training and assessing the U-net architecture described in [3], employing various combinations of modalities like Flair, Flair combined with T1ce, and Flair combined with T2. Among these combinations, the one utilizing Flair and T1ce stood out, showcasing impressive accuracy of 98.29, a dice loss of 50.84, and a Mean IOU of 73.28 compared to others as shown in table [1]. As a result of these promising results, we opted to utilize Flair and T1ce as the input modalities for subsequent models in our study.

TABLE I
Results of U-net

Metrics	Input Modalities		
	Flair	***Flair + T1ce***	***Flair + T2***
Accuracy	0.9868	0.9829	0.9693
MeanIOU	0.8276	0.7328	0.3761
Dice coefficient	0.4592	0.5084	0.3745

After training and testing the 2D ResNet-50 backbone algorithm with the BRATS 2020 dataset for approximately 100 epochs, we achieved an accuracy of 99% with a loss of 36% Ȧdditionally, the mean IOU and Dice coefficient were measured at 57% and 66 % respectively.

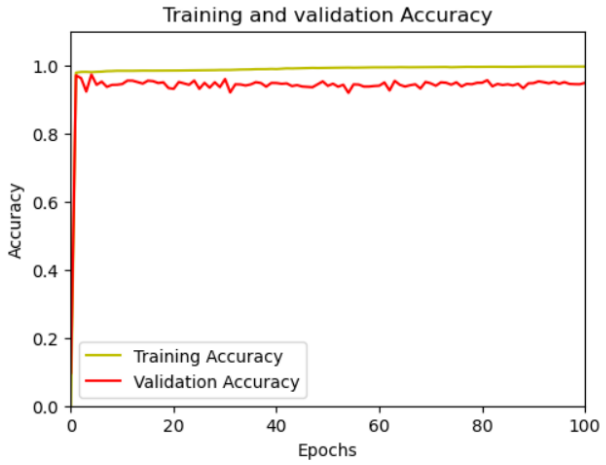

Fig. 7. the graph shows the accuracy of the resnet model.

The results indicate that the proposed W-Net approach has been rigorously evaluated using the publicly available BRATS

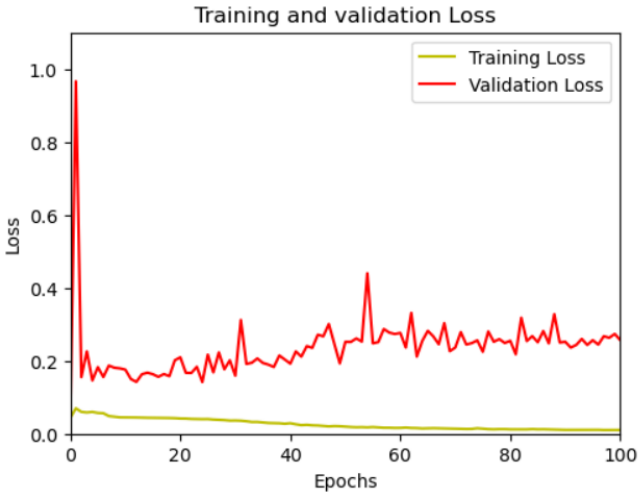

Fig. 8. the graph shows the loss of the resnet model.

2020 dataset, a well-known standard dataset with varying dimensions. The proposed model operates by processing a 2D image, predicting the brain tumor, then recombining the results into a single image of the original size, and finally saving the predicted output. The following figure illustrates the predicted segmentation mask alongside its corresponding ground truth mask.

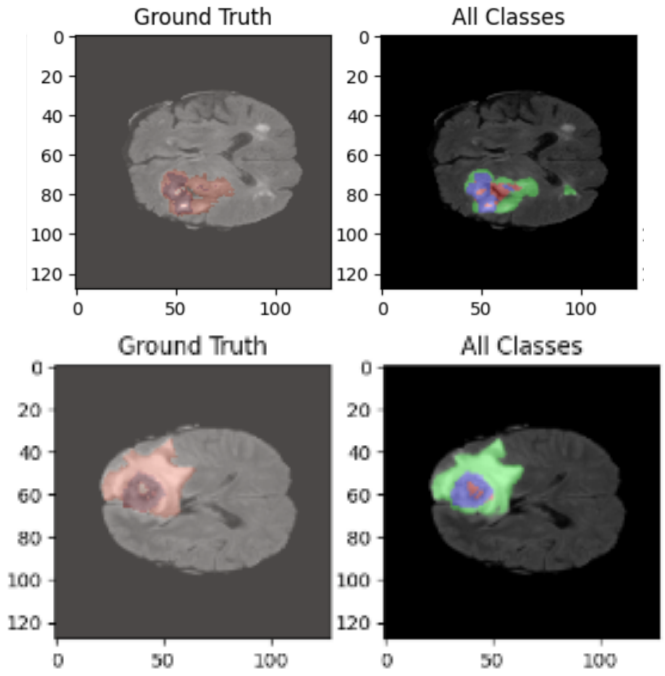

Fig. 9. The above figure shows that ground truth and predicted segmentation mask

A comparison table [table-2] that showcases the performance of various existing models. This table is designed to offer a clear overview of the differences in metrics such as accuracy, dice coefficient, mean IOU across different architectures. It provides a quick reference for understanding how each model performs under similar conditions. This comparison can serve as a valuable tool for both assessing current methodologies and guiding future development in this field.

TABLE II
COMPARISON OF METRICS ON DIFFERENT MODELS.

Models	Parameters		
	MeanIOU	***Dice coefficient***	***Accuracy***
U-Net [3]	0.7328	0.5084	0.9829
U-Net Modeied [5]	0.62	-	0.98
resUnet [4]	0.5787	0.6643	0.9909
LinkNet [7]	-	0.73	-
SC-SE U-Net architecture [8]	84.22	0.91	-
Proposed model W-Net	**0.9371**	**0.7697**	**0.9919**

VI. Conclusion and Future Work

In this paper, we introduced a new brain tumor segmentation model based on the W-Net architecture. Our experiments with the BRATS 2020 dataset demonstrated this model's effectiveness, achieving a mean IOU of 93, an accuracy of 99, and a dice coefficient of 76. These results underscore the model's ability to segment brain tumors with precision, a critical aspect for supporting clinical diagnosis and treatment planning.

However, there's still room for improvement. To further enhance segmentation accuracy, several modifications could be considered for the proposed model. Incorporating residual blocks can boost feature extraction by allowing deeper layers to reference earlier ones, reducing information loss during training. Attention mechanisms, which help the model focus on the most significant parts of the image, might also improve segmentation precision, potentially increasing the dice coefficient and mean IOU.

Overall, the W-Net-based model provides a solid foundation for brain tumor segmentation. The possibility of integrating additional advanced blocks, such as residual and attention mechanisms, indicates that this model could be further refined to improve performance. Future studies could focus on these enhancements and assess their impact on segmentation outcomes, ultimately contributing to more precise and reliable tools for brain tumor diagnosis and treatment.

Acknowledgment

References

[1] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015)

[2] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J.S. Kirby, et al., "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", Nature Scientific Data, 4:170117 (2017)

[3] Ronneberger O, Fischer P, Brox T (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". arXiv:1505.04597

[4] Diakogiannis, F. I., Waldner, F., Caccetta, P., and Wu, C. (2019). ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS Journal of Photogrammetry and Remote Sensing..

[5] D. S. VINOD, S. P. SHIVA PRAKASH, HUSSAIN ALSALMAN, ABDULLAH Y. MUAAD, MD. BELAL BIN HEYAT, "Ensemble Technique for Brain Tumor Patient Survival Prediction", in IEEE Access, 2024.

[6] Alok Patel, Kumar Prateek, Soumyadev Maity "A W-Net Based Architecture with Residual Block for Liver Segmentation," International Conference on Inventive Research in Computing Applications, 2022.

[7] Z. Sobhaninia, S. Rezaei, A. Noroozi, M. Ahmadi, H. Zarrabi, N. Karimi, A. Emami, and S. Samavi, "Brain tumor segmentation using deep learning by type specific sorting of images," arXiv preprint arXiv:1809.07786, 2018

[8] A. F. M. Minhazur Rahman, and Md. Ali Hossain, "Attention-refined U-Net with Skip Connections for Effective Brain Tumor Segmentation from MRI Images", 26th International Conference on Computer and Information Technology, 2023

[9] Li, X., Chen, H., Qi, X., Dou, Q., and Heng, P. A. (2018). H-DenseUNet: Hybrid densely connected U-Net for liver and tumor segmentation from CT volumes. IEEE Transactions on Medical Imaging.

[10] Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., and Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. Nature Methods.

[11] S. Bakas, M. Reyes, A. Jakab, S. Bauer, M. Rempfler, A. Crimi, et al., "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge", arXiv preprint arXiv:1811.02629 (2018)