The Development of the RSU U^2 Net+ Architecture for Brain Tumor Segmentation in 3D Images

Elvaret
Faculty of Computer Science
Esa Unggul University
Jakarta, Indonesia
Email: elvaretharefa [AT] gmail.com

Habibullah Akbar
Faculty of Computer Science
Esa Unggul University
Jakarta, Indonesia
Email: habibullah.akbar [AT] esaunggul.ac.id

***Abstract*—Segmenting brain tumors in medical images plays a crucial role in diagnosis and monitoring of medical conditions. However, the segmentation process is still performed manually, consuming time and exhibiting variability among assessors. This research aims to develop the RSU U^2-Net+ architecture for brain tumor multilabel segmentation in 3D images. The RSU U^2-Net+ architecture consists of 9 interconnected blocks, employing broader connectivity in each block. The architecture is reinforced with the use of Residual U-blocks (RSU) to enhance image understanding across various scales without significantly increasing computational load. Testing on data reveals that the RSU U^2-Net+ architecture performs well, as indicated by a dice coefficient score of 0.779, IoU of 0.6439, recall of 0.7541, and specificity of 0.9911. Evaluation is also conducted for each tumor label. Recall and specificity for edema are 0.8690 and 0.9851, for enhancing tumor are 0.7991 and 0.9956, and for non-enhancing tumor are 0.5942 and 0.9927. This research makes a significant contribution to the development of advanced medical image analysis technology. The achieved results have tangible benefits for medical practitioners and patients, with the potential to enhance the speed and consistency of brain tumor segmentation in 3D medical images.**

Keywords-component; Computer Vision, Deep Learning, RSU U^2 Net+, Brain Tumor

I. INTRODUCTION

Brain tumors are one of the most threatening diseases to a patient's health and quality of life. Brain tumors occur due to the growth of abnormal cells within the brain that can be malignant or benign. Malignant brain tumors, such as gliomas, can cause serious and life-threatening symptoms, such as severe headaches, seizures, visual disturbances, and behavioral changes [1]. According to the International Agency for cancer research at WHO in 2020 reported that there were 168,346 male patients and 139,756 female patients suffering from brain tumors worldwide [2].

The disease is often difficult to diagnose early, and when detected, treatment is complex [3]. Segmenting brain tumors from 3D medical images, such as MRI (Magnetic Resonance Imaging) and CT (Computed Tomography) images, is a key step in diagnosis, treatment planning, and monitoring disease progression.

In general, brain tumor segmentation is performed manually by radiologists, which is time-consuming and can have inter-rater variations [4]. Therefore, in recent years, the development of automated segmentation technology using neural networks has become very attractive [5]. Artificial neural networks, particularly in the form of Convolutional Neural Networks (CNNs), have produced promising results in medical image segmentation tasks, including brain images [6].

One CNN architecture that stands out in semantic segmentation is the U-Net architecture. U-Net is an artificial neural network architecture developed specifically for semantic segmentation tasks in image processing and has been successfully applied in 2D medical image segmentation tasks with impressive performance [7]. This architecture was first introduced by Olaf Ronneberger, Philipp Fischer, and Thomas Brox in their paper entitled "U-Net: Convolutional Networks for Biomedical Image Segmentation" in 2015 [8].

However, its use in the context of 3D medical images, which have additional dimensions and higher complexity, is still a challenge [9]. Brain tumor segmentation in 3D medical images requires the ability to identify and separate tumors from normal brain structures in three-dimensional space [10].

In this research, a RSU U^2 Net+ architecture is developed for brain tumor segmentation in 3D medical images that will make a significant contribution to the medical world [11]. This architecture will be trained on the BraTS public dataset which stands for "Brain Tumor Segmentation" with 3D brain tumor segmentation images [12]. RSU U^2 Net+ has a deeper architecture than the regular U-Net and applies residual blocks that allow the original information of the input to flow more freely through the neural network [13]. It also has better regularization techniques that help prevent overfitting and improve generalization to unseen data. The development of this architecture aims to improve accuracy and efficiency in the diagnosis process, assist in more targeted treatment planning, and enable better monitoring of disease progression [14]. In addition, by improving automation capabilities in brain tumor segmentation, it can reduce inter-rater variability and speed up the patient care process [15]. The results of this study will hopefully have a positive impact in the medical world, allowing medical professionals to better deal with the complexities of

brain tumors, and ultimately, improve the treatment and prognosis of patients suffering from this disease [16].

II. METHODS

A. Data

The dataset used comes from the BratTS 2017 3D MRI image dataset in Nifty format, which consists of 484 brain MRI scan images along with a segmentation mask that shows the border or area of brain tumor identification. This brain MRI scan image consists of several channels, namely flair, T1w, T1-Gd, and T2w. In this dataset, there is information about the brain tumor segmentation mask which is divided into several labels, namely edema, non-enhancing tumor, and enhancing tumor. Each 3D MRI image has a volume of $240 \times 240 \times 155 \times 4$ and mask $240 \times 240 \times 155$.

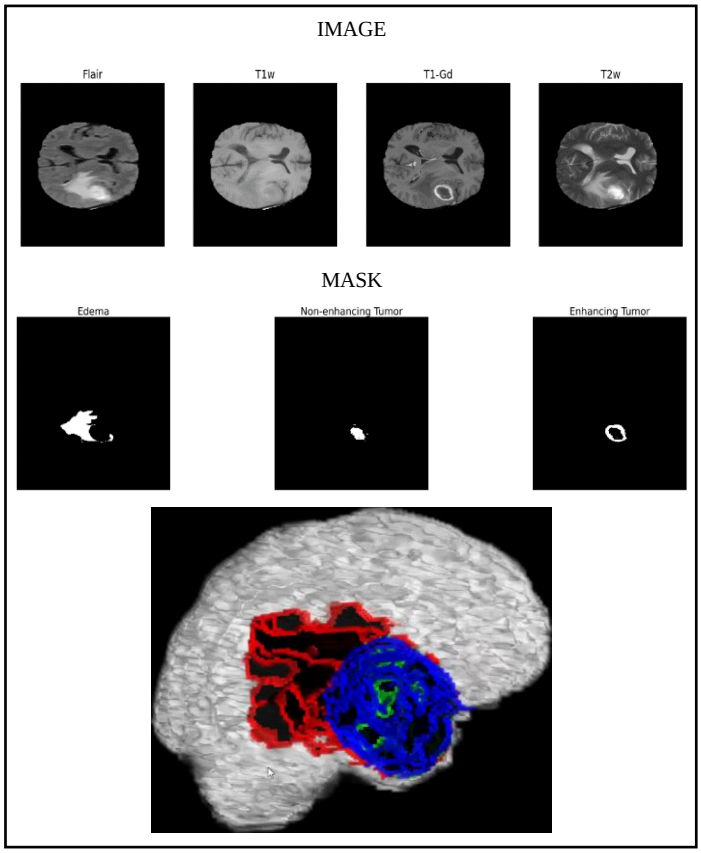

Figure 1. Image and Mask in 2D and 3D Preview

The researcher used one-hot encoding to provide binary representation for each mask, resulting in masks with their own vector dimensions of $240 \times 240 \times 155 \times 3$. After that, colors were assigned to the mask for each label: red for edema, green for non-enhancing tumor, and blue for enhancing tumor. The masks were then integrated into the 3D image.

B. Preprocessing

For data preprocessing, the researcher employed a patch approach to generate sub-volumes. This method randomly generates 5 sub-volumes for each 3D image and applies the criterion that only 10% of the tumor area will be selected. This approach is utilized to ensure that the model focuses on relevant pixels and guarantees that the data used in this research has a significant representation in important areas for analysis. Subsequently, this sub volume will be standardized to ensure uniformity pixel value that applied to the model. After generating sub volumes from 484 3D images, we will split the sub volume into 80% for training data, 10% for validation data, and 10% for testing data.

C. Proposed Architecture

The RSU U²-Net+ architecture consists of 6 interconnected U-Net+ blocks and 3 residual custom dilated blocks. Each U-Net+ block employs broader connectivity inspired by U-Net++. Additionally, the architecture is reinforced with the use of Residual U-blocks (RSU) to enhance image understanding across various scales without significantly increasing computational load. Furthermore, Residual Custom Dilated Blocks are utilized to examine information at greater depths or wider receptive fields up to dilation 8.

1) Phase 1: (RSU U-Net+ Block)

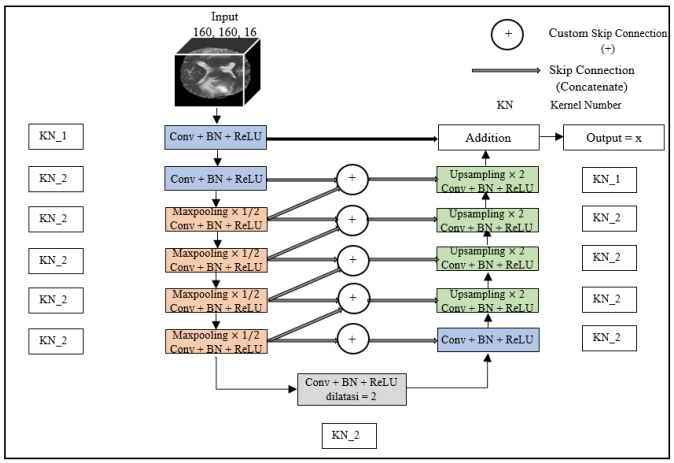

Figure 2. 1 Blok 3D RSU U-Net +

The first phase of the RSU U² Net+ architecture to be developed consists of 6 blocks with the same operational approach but varying numbers of layers per block. In the encoder section, a 3D image of a brain tumor is taken as input with dimensions of 160 x 160 x 16. Feature extraction is then performed using convolution layers (kernel size = 3×3×3), followed by Batch Normalization. Batch Normalization calculates the batch's mean and standard deviation, normalizing the data. The normalized data is then passed through ReLU activation. By using Batch Normalization before ReLU, it helps maintain the distribution of the convolution layer output,

making it more stable during training. Following that, there is a downsampling stage using maxpooling to reduce the spatial dimensions of the image data by half, aiding in extracting crucial features and allowing the U-Net network to learn increasingly complex features during subsequent convolution layers.

Subsequently, after the encoder, there is a bridge section. The bridge is a key component in U-Net that facilitates the flow of information between the encoder and decoder. This bridge utilizes convolution layers, Batch Normalization, ReLU activation, and dilation of 2 to merge features extracted by the encoder with features to be used in the decoding process. Dilation of 2 at this stage is employed to allow the CNN network to access information farther from points in the input image. This is valuable in tasks like image segmentation, where the network needs to understand a broader context. Next, there is the final part in the U-Net architecture, which is the decoder. After the bridge, the decoding process is performed using Upsampling or Conv3D Transpose operations to restore the image resolution to its original size. In this section, Concatenate operations are also conducted, merging information from the bridge with skip connections (contextual information from the encoder). This is crucial to preserve the necessary details and context in producing accurate segmentation or reconstruction results. The researcher also introduces a modification to this concatenation operation, using two convolution layers to create an additional skip connection. This modification is designed to allow an additional flow of information and help preserve higher-level details in the image segmentation task. This modification is inspired by the U-Net++ architecture introduced in the paper titled "Unet++: A Nested U-Net Architecture for Medical Image Segmentation" [17]. However, in this study, the researcher opts for a U-Net+ variation due to computational burdens and a significant increase in parameter count in 3D image segmentation tasks. The final operation in this block is Addition. The researcher adds the result of this decoding process to the original features extracted at the beginning. This aims to retain the original details of the image that may be lost during the convolution and pooling processes in the encoder.

2) *Residual Custom Dilated Block*

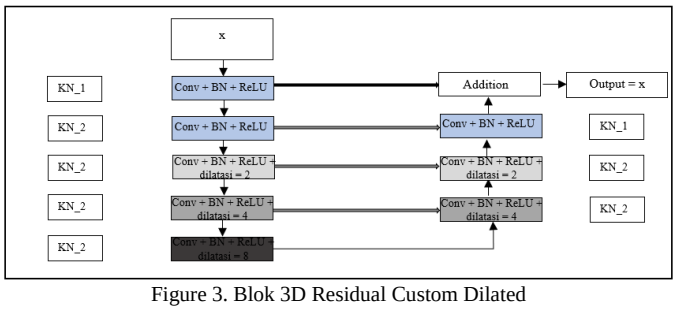

Figure 3. Blok 3D Residual Custom Dilated

The second phase of the RSU U^2 Net+ architecture consists of 3 blocks with the same operation and the same number of layers per block as the previous phase. These blocks utilize the same operations as in the previous phase. What sets this block apart is the use of dilation rates 4 and 8. This block is not a U-Net architecture consisting of encoder, bridge, and decoder, but rather a Residual Custom Dilated Block used to examine information further or a wider receptive field up to dilation 8. The use of larger dilation in these convolutions has several potential advantages such as a larger receptive field and expanding contextual understanding.

3) *Full RSU U^2 Net +*

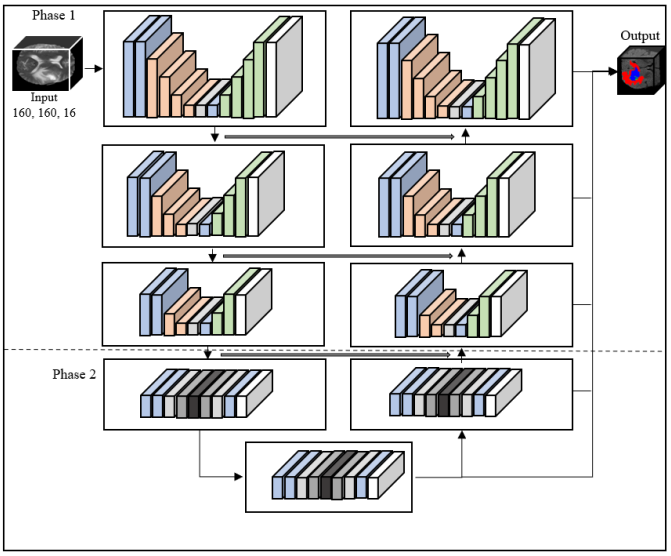

Figure 4. RSU U^2 Net + Architecture

The illustration above depicts the full architecture of the RSU U^2 Net+, consisting of 9 blocks until reaching the final output. Here are the numbers of convolutional kernels used in each block:

TABLE I. NUMBER OF KERNELS IN BLOCK 1 - 4

RSU U^2Net +	Blok 1	Blok 2	Blok 3	Blok 4
KN_1	32	64	128	256
KN_2	32	32	64	128

TABLE II. NUMBER OF KERNELS IN BLOCK 5 - 9

RSU U^2Net +	Blok 5	Blok 6	Blok 7	Blok 8	Blok 9
KN_1	512	256	128	64	32
KN_2	256	128	64	32	32

In this architecture as well, there are skip connections between blocks, allowing direct flow of information between different levels in the architecture, which helps maintain higher and lower-level feature information during the image segmentation process. For the final output, the researcher uses sigmoid activation to determine the regions that constitute

edema, non-enhancing tumor, and enhancing tumor. Sigmoid is used because brain tumor segmentation is a multilabel segmentation where certain areas can have multiple labels simultaneously, and each label is interpreted as an independent probability.

The RSU U^2 Net+ architecture is a robust model for image segmentation, especially in the context of object segmentation tasks that require high precision and accuracy. With hierarchical convolutional layers, dilation techniques, and information fusion from various resolution levels, this model can retain significant detail in the images. The output results from this architecture have high resolution and are suitable for applications in 3D brain tumor image segmentation.

D. Training

The model training process with RSU U^2 Net+ architecture will use training data and validation data. Validation data will be used to evaluate the model during the training process and monitor the model to measure how well the model work on the data that the model hadn't never seen. The researcher will use soft dice loss as loss function to handle the multilabel 3D segmentation because soft dice loss leads to smoother optimization landscapes. The continuous nature of the soft dice loss allows for more stable and gradual updates to the model weights during training. This can result in more reliable convergence during the optimization process. For the soft dice loss formula:

$$Soft\ Dice\ Loss = 1 - \frac{1}{3}\sum_{c=1}^{3}\frac{2\times\sum_{i=1}^{N} p_i \times g_i + \varepsilon}{\sum_{i=1}^{N} p_i^2 + \sum_{i=1}^{N} g_i^2 + \varepsilon} \quad (1)$$

Where:

p_i : the predicted value generated by the model for each pixel i
g_i : the ground truth value for each pixel i
N : the total number of pixels in the 3D image
ε : the epsilon value (0.00001) to avoid division by zero
c : class (1,2,3)

As the optimization algorithm, the researcher will use the Adam optimizer because this algorithm combines the advantages of two optimization methods, such as RMSprop and Momentum, making it very effective in finding the local minimum.

E. Evaluation

As the evaluation modeling we will use 4 evaluation matrix, such as soft dice score, jaccard index (IoU), sensitivity (recall), and specificity. The evaluation matrices of soft dice score and Jaccard index will be used to monitor the model training process and tune parameters. After the training process is completed, these four evaluation matrices will also be tested on the testing data.

1) Dice score coefficient (DSC): This matrix calculates the overlapping portion between the predicted and ground truth areas, multiplied by 2 to place greater emphasis on the intersecting regions. It is then divided by the total number of elements in both areas. This matrix provides an insight into how well the model predictions align with the location and shape of the actual areas in the identified 3D brain tumor segmentation data. For the dice score coefficient formula:

$$DSC = \frac{1}{3}\sum_{c=1}^{3}\frac{2\times\sum_{i=1}^{N} p_i \times g_i + \varepsilon}{\sum_{i=1}^{N} p_i + \sum_{i=1}^{N} g_i + \varepsilon} \quad (2)$$

Where:

p_i : the predicted value generated by the model for each pixel i
g_i : the ground truth value for each pixel i
N : the total number of pixels in the 3D image
ε : the epsilon value (0.00001) to avoid division by zero
c : class (1,2,3)

2) Intersection of Union (IOU): The calculation of IoU dividing the intersecting area by the union area of the prediction and ground truth. The higher the IoU value, the better the brain tumor segmentation results of the model in matching the volume with the 3D ground truth. For the IoU formula:

$$IoU = \frac{1}{3}\sum_{c=1}^{3}\frac{\sum_{i=1}^{N}(p_i \times g_i) + \varepsilon}{\sum_{i=1}^{N}(p_i + g_i - p_i \times g_i) + \varepsilon} \quad (3)$$

Where:

p_i : the predicted value generated by the model for each pixel i
g_i : the ground truth value for each pixel i
N : the total number of pixels in the 3D image
ε : the epsilon value (0.00001) to avoid division by zero
c : class (1,2,3)

3) Sensitivity (Recall): Sensitivity measures the model's ability to correctly detect true positive areas among all actual positive areas. This matrix indicates the extent to which a model identity and map actual tumor areas in medical imags. For sensitivity formula:

$$\text{Sensitivity} = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (4)$$

4) Specificity: Specificity measures the model's ability to correctly identitiy true negative areas among all actual negative areas. This matrix reflects how well the model can accurately distinguish non-tumor areas. For the specificity formula:

$$\text{Specificity} = \frac{True\ Negative}{True\ Negative + False\ Positive} \quad (5)$$

III. RESULT AND DISCUSSION

The data has dimensions of 240, 240, 155 with 4 channels including Flair, T1w, T1-Gd, and T2w, along with a mask sized 240, 240, 155 containing 3 labels: edema, non-enhancing tumor, and enhancing tumor. Following data processing, the researcher

generated 1000 sub-volumes from 200 3D images that meet the specified criteria. These sub-volumes were then split into 800 training data, 100 validation data, and 100 testing data. During the model training process using soft dice loss, the researcher utilized the training data for model training and the validation data for monitoring. The model's performance was assessed using testing data with evaluation metrics including dice score coefficient, IoU, recall, and specificity. The model training was conducted on Colab Pro using a paid version of GPU T4 with a 16 GB GPU memory.

A. Training Process

A sub-volume with dimensions (160, 160, 16) was processed using the RSU U^2 Net+ architecture for model training. The training process was conducted with 800 training data and 100 validation data to train the model and evaluate the performance of this architecture over 35 epochs, with a learning rate set to 0.00001 and a batch size of 2.

Figure 5. Training Process with RSU U^2 Net+ Architecture

This architecture is capable of learning patterns within the data and demonstrates consistent improvement from start to finish. The loss function values on the training set show a significant decrease from 0.4011 in the first epoch to 0.0814 in the last epoch (35 epochs), indicating effective learning and approaching convergence. Evaluation metrics such as mean Intersection over Union (mean IoU) and dice coefficient for both training and validation also exhibit consistent improvement throughout the training process. This suggests that the model becomes more proficient in understanding and generating object segmentation in previously unseen data. These changes indicate the model's success in recognizing more complex patterns and features over time, leading to overall performance improvement.

Figure 6. Training Process with U-Net Architecture

The U-Net architecture used as a comparison to the results of the training process with the RSU U^2 Net+ architecture also exhibits commendable performance. The loss function on the training set decreases from 0.4619 in the first epoch to 0.0993 in the last epoch (35 epochs), indicating model convergence and its ability to reduce prediction errors. The mean Intersection over Union (mean IoU) and dice coefficient values consistently improve during training, especially on the validation set.

TABLE III. LOSS AND VALIDATION LOSS DURING MODEL TRAINING

Architecture	Loss	Val loss
RSU U^{2}Net +	0.0814	0.1320
U-Net	0.0993	0.1357

TABLE IV. DSC AND IOU COMPARISON DURING MODEL TRAINING

Architecture	DSC	IoU	Val DSC	Val IoU
RSU U^{2}Net +	0.8497	0.7420	0.8007	0.6735
U-Net	0.8269	0.7105	0.8045	0.6790

TABLE V. COMPUTATIONAL PERFORMANCE

Architecture	GPU	Training Time (minutes)	Parameter
RSU U^{2}Net +	T4	397.7	51.490.695
U-Net	T4	433.6	16.318.307

Despite the RSU U² Net+ architecture having a larger number of parameters compared to U-Net, RSU U² Net+ manages to complete the training process in a shorter time. RSU U² Net+ requires approximately 397.7 minutes for training, while U-Net takes longer, around 433.6 minutes. Furthermore, RSU U² Net+ achieves competitive evaluation results, including Dice Coefficient, IoU, and loss on both training and validation data, when compared to U-Net. This indicates the efficiency and high performance of RSU U² Net+, making it a consideration for selecting an architecture for medical data segmentation tasks.

B. Model Evaluation

After the training process is completed, the model will undergo an evaluation process on 100 testing data using the Dice Score Coefficient, IoU, sensitivity (recall), and specificity.

TABLE VI. MODEL EVALUATION WITH DSC AND IOU

Architecture	Dice Score Coefficient	IoU
RSU U^{2}Net +	0.7779	0.6439
U-Net	0.7768	0.6424

1) DSC and IoU RSU U^2 Net+ Architecture

Based on the evaluation using 100 testing data, the RSU U² Net+ architecture achieves a Dice Coefficient of 0.7779 and an Intersection over Union (IoU) of 0.6439. These evaluation results demonstrate the model's capability in accurately segmenting objects in the test data, with a high level of precision and similarity between the predicted results and the ground truth.

2) DSC and IoU U-Net Architecture

Based on the evaluation using 100 testing data, the U-Net architecture also shows good results with a Dice Coefficient of 0.7768 and an IoU of 0.6424, although it is still surpassed by the RSU U² Net+ architecture.

TABLE VII. MODEL EVALUATION WITH SENSITIVITY AND SPECIFICITY

	RSU U²Net +		**U-Net**	
Label	**Sensitivity (Recall)**	**Specificity**	**Sensitivity (Recall)**	**Specificity**
Endema	0.8690	0.9851	0.8566	0.9841
Enhancing Tumor	0.7991	0.9956	0.7262	0.9970
Non-enhancing Tumor	0.5942	0.9927	0.5551	0.9951
Mean	0.7541	0.9911	0.7126	0.9921

3) Sensitivity and Specificity RSU U² Net+ Architecture

Based on the evaluation matrix results, the RSU U² Net+ architecture demonstrates excellent performance in classifying three labels: Edema, Enhancing Tumor, and Non-enhancing Tumor. The highest sensitivity (recall) is obtained for the Edema label at 0.8690, followed by Enhancing Tumor with a value of 0.7991, and Non-enhancing Tumor with a value of 0.5942. Overall, the average sensitivity reaches 0.7541, while specificity reaches 0.9911. These results indicate that RSU U² Net+ effectively recognizes the presence and location of the three types of lesions in medical images.

4) Sensitivity and Specificity U-Net Architecture

Based on the evaluation matrix results, the U-Net architecture also shows good results in classifying the three labels; however, the sensitivity values are lower compared to RSU U² Net+. The highest sensitivity is obtained for the Edema label at 0.8566, followed by Enhancing Tumor with a value of 0.7262, and Non-enhancing Tumor with a value of 0.5551. The average sensitivity of U-Net reaches 0.7126, while specificity reaches 0.9921. These results indicate that RSU U² Net+ is superior in recognizing the presence and location of the three types of lesions in medical images compared to U-Net.

C. Predict and Vizualization 1 Image 3D and 2D View

The trained and tested RSU U² Net+ model is used to observe the prediction results and visualizations for several brain tumor patients. In addition to displaying 3D predictions, the researchers also provide predictions and visualizations in the form of per-slice 2D with three different perspectives: sagittal (left-right), coronal (front-back), and transversal (top-bottom) to gain a deeper understanding.

1) Brain Tumor Patient 100

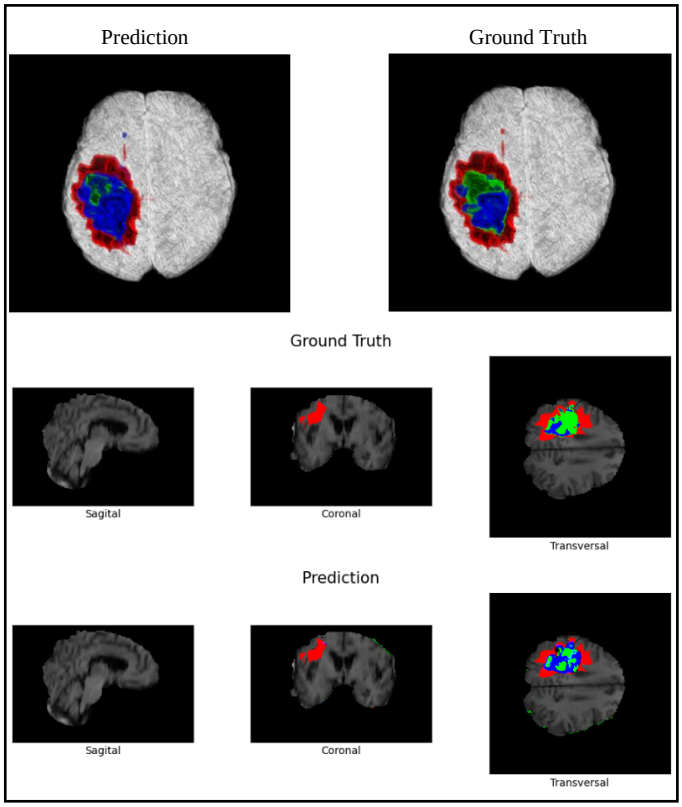

Figure 7. Prediction and Visualization 3D and 2D Brain Tumor Patient 100 Slice 110

TABLE VIII. 3D PREDICTION RESULTS EVALUATION FOR PATIENT 100 USING SENSITIVITY AND SPECIFICITY

	Endema	**Non-Enhancing Tumor**	**Enhancing Tumor**
Recall	0.8359	0.6508	0.9609
Specificity	0.9513	0.9915	0.9591

In the prediction and visualization of brain images for patient 100, the model provides good results with a high recall rate for enhancing tumors and edema. However, in the non-enhancing tumor category, the recall is slightly lower. The high specificity rate indicates that the model tends to minimize the number of false positives or errors in segmenting non-tumor areas as tumors. This is a positive indication regarding the model's ability to identify non-tumor areas in brain images. Despite the decrease in performance in the non-enhancing tumor category, overall, the model can be considered effective in identifying various types of brain tumors for patient 100.

2) Brain Tumor Patient 10

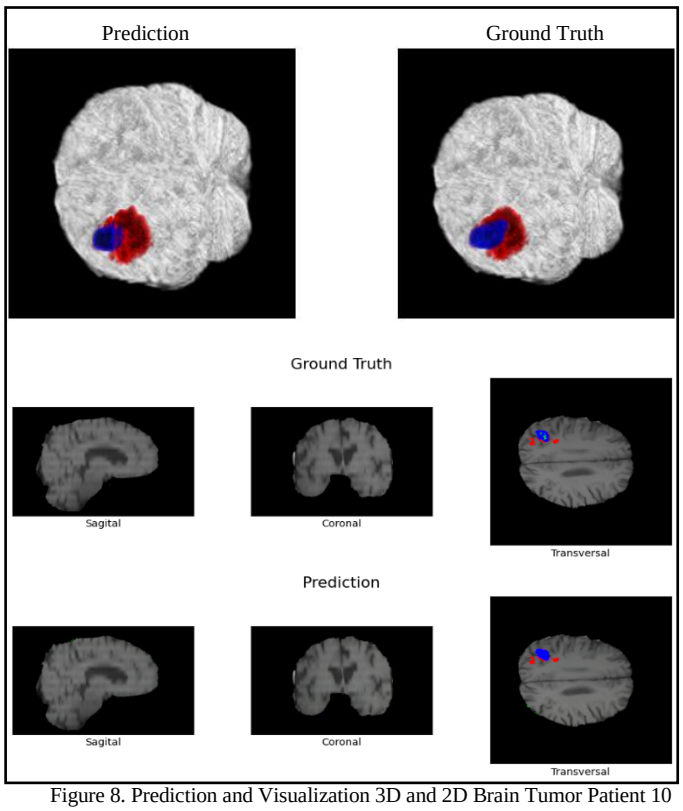

Figure 8. Prediction and Visualization 3D and 2D Brain Tumor Patient 10 Slice 100

TABLE IX. 3D PREDICTION RESULTS EVALUATION FOR PATIENT 10 USING SENSITIVITY AND SPECIFICITY

	Endema	**Non-Enhancing Tumor**	**Enhancing Tumor**
Recall	0.9265	0.0169	0.6067
Specificity	0.9639	0.9958	0.9593

In the prediction and 3D visualization for patient 10, the model shows quite impressive evaluation results, especially considering the relatively small size of the tumor (below 5% of the background) compared to the sizes of tumors in other patients. The model demonstrates a high ability to detect edema, as indicated by a recall of 92.65%, showcasing its capability to recognize areas genuinely affected by edema. However, in non-enhancing tumors, the recall is very low (1.69%) with a high specificity rate of 99.58%, indicating the model's tendency to minimize errors in classifying non-enhancing tumor areas as non-tumor. For enhancing tumor types, the model achieves a recall of 60.67% with a specificity of 95.93%. These results provide an understanding that the model is effective in detecting edema but faces challenges in recognizing non-enhancing tumors.

IV. CONCLUSION

Based on the design of the RSU U2 Net+ architecture and the conducted evaluations, several conclusions can be drawn. The RSU U^2 Net+ model successfully enhances efficiency and reduces segmentation variability among evaluators while achieving consistency in interpreting brain tumor images. The resulting architecture holds the potential for further development and integration into systems for automated segmentation of brain tumors. Optimized for effective segmentation of 3D medical images, particularly brain tumors, the RSU U^2 Net+ architecture demonstrates promising evaluation results, highlighting its suitability for applications in the field, as explored in the research titled 'The Application of the RSU U^2 Net+ Architecture for Brain Tumor Segmentation in 3D Images.

Although this study has made a positive contribution to the development of the RSU U^2 Net+ architecture for brain tumor segmentation in 3D images, there are some limitations that could be addressed to enhance the quality of future research. Particularly, future studies could benefit from comparing the RSU U^2 Net+ architecture with alternatives beyond U-Net. Recommendations for further research include obtaining a larger and more diverse dataset, implementing additional data augmentation techniques, and exploring variants of soft dice loss that are more sensitive to imbalanced labels in multilabel segmentation problems. These suggestions aim to refine and expand the capabilities of the RSU U^2 Net+ architecture in future investigations.

REFERENCES

[1] S. Hadidchi *et al.*, "Headache and Brain Tumor," *Neuroimaging Clin N Am*, vol. 29, no. 2, pp. 291–300, 2019, doi: 10.1016/j.nic.2019.01.008.

[2] WHO, *Brain, central nervous system*. 2020. [Online]. Available: https://gco.iarc.fr/today/data/factsheets/cancers/31-Brain-central-nervous-system-fact-sheet.pdf

[3] K. Dananjoyo, W. Nalendra Tama, R. Ghazali Malueka, A. Asmedi Staf Departemen Neurologi Fakultas Kedokteran, K. Masyarakat dan Keperawatan, and U. Gadjah Mada, "Nyeri kepala tumor otak pada dewasa Brain tumor headache in adult," 2019.

[4] M. Nadia Putri, I. Katili, A. Hariri, T. Asih Budiati, and G. Murti Wibowo, "94-2,5) Poltekkes Kemenkes Semarang, Indonesia 3,4) Pertamina Central Hospital Jakarta," *JImeD)*, vol. 7, 2021, [Online]. Available: http://ejournal.poltekkes-smg.ac.id/ojs/index.php/jimed/index

[5] M. S. Swetha and M. P. Devi, "NEURAL NETWORK SEGMENTATION IN MRIs IMAGES FOR BRAIN TUMORs SEGMENTATIONs," *International Journal of Scientific Research in Engineering and Management*, 2022, doi: 10.55041/IJSREM12703.

[6] M. B. Sudhan *et al.*, "Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model," *J Healthc Eng*, vol. 2022, 2022, doi: 10.1155/2022/1601354.

[7] I. B. L. M. Suta, M. Sudarma, and I. N. Satya Kumara, "Segmentasi Tumor Otak Berdasarkan Citra Magnetic Resonance Imaging Dengan Menggunakan Metode U-NET," *Majalah Ilmiah Teknologi Elektro*, vol. 19, no. 2, p. 151, Dec. 2020, doi: 10.24843/mite.2020.v19i02.p05.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," May 2015, [Online]. Available: http://arxiv.org/abs/1505.04597

[9] M. Kolařík, R. Burget, V. Uher, K. Říha, and M. K. Dutta, "Optimized high resolution 3D dense-U-Net network for brain and spine segmentation," *Applied Sciences (Switzerland)*, vol. 9, no. 3, Jan. 2019, doi: 10.3390/app9030404.

[10] Y. He, D. Yang, H. Roth, C. Zhao, and D. Xu, "DiNTS: Differentiable Neural Network Topology Search for 3D Medical Image Segmentation," Mar. 2021, [Online]. Available: http://arxiv.org/abs/2103.15954

[11] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, "U^2-Net: Going Deeper with Nested U-Structure for Salient Object Detection," May 2020, doi: 10.1016/j.patcog.2020.107404.

[12] M. Ghaffari, A. Sowmya, and R. Oliver, "Automated Brain Tumor Segmentation Using Multimodal Brain Scans: A Survey Based on Models Submitted to the BraTS 2012-2018 Challenges," *IEEE Rev Biomed Eng*, vol. 13, pp. 156–168, 2020, doi: 10.1109/RBME.2019.2946868.

[13] Z. Tabatabaei, A. Colomer, K. Engan, J. Oliver, and V. Naranjo, "Residual block Convolutional Auto Encoder in Content-Based Medical Image Retrieval," in *IVMSP 2022 - 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/IVMSP54334.2022.9816325.

[14] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," Jul. 2018, [Online]. Available: http://arxiv.org/abs/1807.10165

[15] X. Liu, L. Song, S. Liu, and Y. Zhang, "A review of deep-learning-based medical image segmentation methods," *Sustainability (Switzerland)*, vol. 13, no. 3, pp. 1–29, Feb. 2021, doi: 10.3390/su13031224.

[16] M. K. Abd-Ellah, A. I. Awad, A. A. M. Khalaf, and H. F. A. Hamed, "A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned," *Magnetic Resonance Imaging*, vol. 61. Elsevier Inc., pp. 300–318, Sep. 01, 2019. doi: 10.1016/j.mri.2019.05.028.

[17] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2018, pp. 3–11. doi: 10.1007/978-3-030-00889-5_1.