Recursive Journal of Informatics
Vol. 4, No. 1, March 2026

p-ISSN 2963-5551 https://journal.unnes.ac.id/journals/index.php/rji e-ISSN 2986-6588

Optimization of Brain Tumor Segmentation on Magnetic Resonance Imaging (MRI) Using Attention Gate U-Net

Fariska Ratna Fauziah[1]*, Budi Prasetiyo[2]
[1,2]Computer Science Department, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Indonesia

Abstract. Brain tumor segmentation using Magnetic Resonance Imaging (MRI) plays a vital role in medical diagnosis, requiring high precision to support clinical decisions and reduce mortality rates.
Purpose: This research aims to enhance the segmentation process by implementing an Attention Gate into the U-Net model.
Methods/Study design/approach: In the segmentation stage, Attention Gate on U-Net is integrated to filter out relevant information from the extracted features, resulting in a more precise segmentation of the brain tumor to determine the location of the tumor.
Result/Findings: The performance of the model is assessed by calculating several evaluation metrics such as dice coefficient and intersection-over-union (IoU) for the segmentation process. The results showed that adding Attention Gate to the U-Net achieved a dice coefficient of 87.08% and IoU of 72.70%
Novelty/Originality/Value: The novelty of this study lies in the integration of the Attention Gate mechanism within the U-Net decoder stage to enhance focus on tumor regions. While U-Net is widely used in medical image segmentation, this specific attention-based enhancement significantly improves performance compared to conventional U-Net models without attention. This research contributes to advancing more accurate and efficient decision-support systems in the field of medical image analysis.

Keywords: Brain Tumor, Segmentation, U-Net, Attention Gate, Deep Learning
Received June 2025 / **Revised** March 2026 / **Accepted** March 2026

INTRODUCTION

The central nervous system, which consists of the brain and spinal cord, plays a vital role in transmitting sensory information and coordinating bodily responses [1], [2], [3]. Disorders of this system, such as brain tumors, can significantly threaten bodily functions. Brain tumors are caused by uncontrolled tissue growth and can be benign or malignant, depending on their size, location, and cellular characteristics [4]. The World Health Organization (WHO) classifies brain tumors into four grades, from I to IV, with varying degrees of malignancy and prognosis [5].

The prevalence of brain tumors is increasing globally each year. Brain tumors account for approximately 1.35% of all malignant neoplasms and 29.5% of tumor-related deaths, with survival rates ranging from 5–35% for malignant types and 90% for benign types [6]. Therefore, early detection and accurate monitoring are crucial [7]. One of the most commonly used diagnostic methods is Magnetic Resonance Imaging (MRI) [8]. However, due to the complexity of brain structure and the variation in tumor shape and size, the process of tumor segmentation and classification from MRI images poses a significant challenge [9].

Manual segmentation of brain tumors on MRI images is time-consuming and highly dependent on the expertise of medical personnel. Therefore, automated methods based on deep learning offer a promising solution to improve efficiency and consistency in diagnosis. However, traditional methods such as the Active Contour Model [10], Otsu's Thresholding [11], and Region Growing [12] are still widely used. However, these methods have several limitations, such as sensitivity to noise, dependence on initialization, and difficulty in handling complex tumor shapes. For example, Active Contour only

[1]*Corresponding author.
Email addresses: fariskafauziah07@students.unnes.ac.id (Fauziah)
DOI: 10.15294/rji.v4i1.27152

produces a dice score of 0.665, while Region Growing achieves 0.86, but is susceptible to intensity variations in MRI images.

As a more reliable alternative, deep learning models such as U-Net have been widely applied in brain tumor segmentation due to their ability to effectively capture spatial and contextual information. [13] demonstrated that U-Net can achieve a Dice score of up to 0.94 on the BRATS 2012 dataset, far surpassing traditional approaches. However, U-Net also has several limitations, such as low sensitivity to background details and complex ground truth [14], high computational requirements [15], and the risk of overfitting on data different from the training data [16].

To address these limitations, various studies have developed modifications to U-Net by adding an Attention Module. One example is Attention Res-UNet [17], which incorporates an Attention Gate (AG) to focus attention on tumor areas and ignore irrelevant parts, thereby improving segmentation performance on the BRATS 2019 dataset. The use of AG has also proven effective in other fields, such as remote sensing image segmentation to reduce noise and highlight important features [14]. This module works by assigning higher weights to relevant features and suppressing insignificant ones (Amisha & Adersh, 2023). The concept of AG was first introduced by [18] in the context of deep learning-based machine translation.

Several studies have utilized this approach to improve the accuracy of brain tumor segmentation. [19] developed MAG-Net, a U-Net-based model that integrates an attention mechanism and depthwise separable convolution. This model successfully reduced the number of parameters from 31 million to just 5.4 million without compromising performance, achieving a Dice coefficient of 0.74, IoU of 0.60, and accuracy of 99.52%. On the other hand, [20] proposed a combination of the Active Contour Model with Fuzzy C-Means to improve initial contour detection using the Radius Contraction and Expansion (RCE) technique, resulting in a Dice score of 0.665. Meanwhile, the Convolutional Autoencoder (CAN) approach by [21] simplifies the U-Net architecture without skip connections but still achieves a Dice score of 0.7287 with significantly fewer parameters.

Thus, the integration of CNN models like U-Net with attention mechanisms like Attention Gate presents a much more adaptive and efficient approach to brain tumor segmentation compared to conventional methods. This combination not only significantly improves accuracy and precision in identifying complex and varied tumor areas but also allows the model to automatically focus on the most relevant features, thereby reducing interference from noise and non-tumor structures. This innovation has the potential to transform the paradigm in medical diagnosis, providing a strong foundation for faster, more accurate, and personalized clinical decision-making, while opening opportunities for the development of advanced medical assistance systems that can be relied upon in various real-world conditions.

METHODS

After the brain tumor segmentation dataset was obtained, it was processed by resizing, grayscale conversion, bilateral filtering, and pixel normalization. The mask labels were encoded using one-hot encoding. The data was then split into training and testing datasets, which were used to train a U-Net model with Attention Gate. The process was concluded with testing and analysis of the results using the Dice and IoU metrics. The complete workflow is shown in Figure 1.

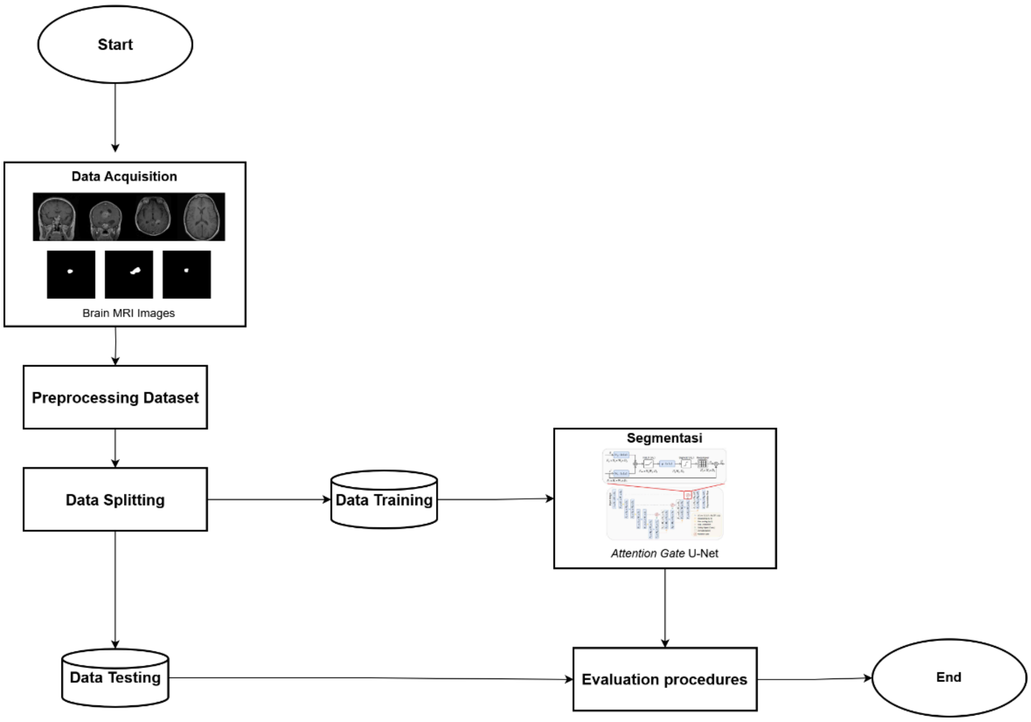

Figure 1. Flowchart Research

Data Acquisition

The data used in this study was obtained from the Figshare Dataset. For the segmentation task, this study used the CE-MRI Figshare Dataset obtained from Cheng's study [22]. This dataset consists of 3,064 brain MRI images from 233 patients diagnosed with brain tumors, collected between 2005 and 2010 at Tianjin University General Hospital and Nanfang Hospital in China. The MRI images in this dataset have three viewpoints: axial, coronal, and sagittal. The sample of the dataset is shown in Figure 2.

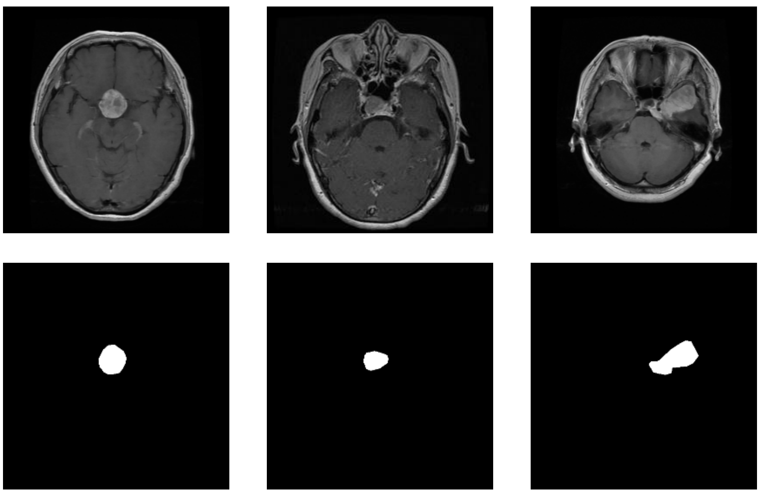

Figure 2. Image Samples and Masking for Segmentation

Data Preprocessing

Preprocessing of the brain tumor segmentation dataset was performed by adjusting the image structure and masking. The initial images with a resolution of 512×512 or 256×256 were resized to 224×224 pixels for training efficiency without compromising segmentation quality, especially in the U-Net architecture. All images are converted to grayscale format to maintain consistency and processing efficiency. To reduce noise while preserving important edge details, a bilateral filter is applied, followed by color mapping to convert grayscale intensity into a specific color representation. Normalization is also performed by scaling pixel values to the range [0, 1], supporting model stability and convergence acceleration. Finally, the masking labels are encoded using a Label Encoder and converted to one-hot encoding format to separate the two target classes: background and tumor. The results of the preprocessing steps performed on the segmentation dataset are shown in Figure 3.

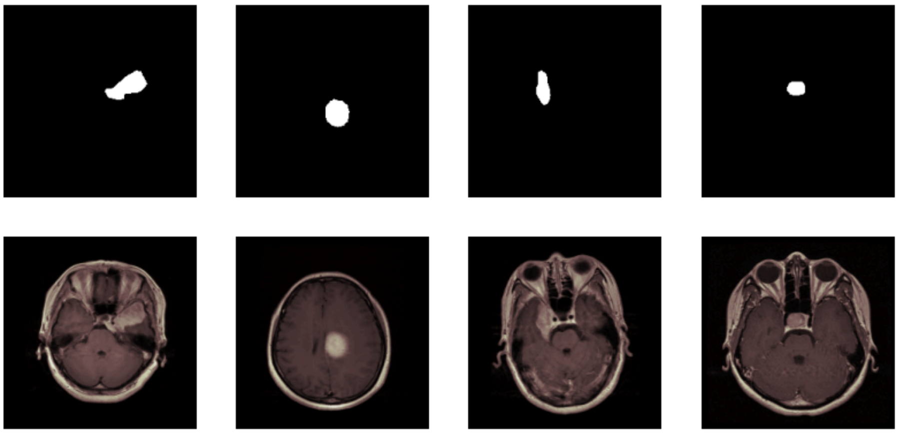

Figure 3. Sample of Preprocessing Results

Data Splitting

To ensure that the brain tumor segmentation model created does not experience overfitting and has optimal performance, the dataset consisting of 3064 images and their masks was divided into three subsets: training data, validation data, and testing data. In the first stage, 85% of the data (2604 images) was used as training data, while the remaining 15% (460 images) was allocated for testing data. Next, the training data was further divided into main training data comprising 80% (2,083 images) and validation data comprising 20% (521 images). This data division aims to ensure that the model is trained to its fullest potential on the training data, consistently validated on the validation data, and tested for performance on the testing data, thereby producing a more reliable model that avoids overfitting.

U-Net

The U-Net method was first introduced by Ronneberger in 2015 [23], initially applied to image segmentation and winning the ISBI 2015 Cell Tracking Challenge and Caries Detection Challenge. U-Net has inspired the development of many network structures and an increasing number of deep learning strategies expanded based on U-Net. The U-Net network structure consists of a contraction path to capture context and a symmetric expansion path for precise localization. U-Net can also be combined with data augmentation techniques to achieve end-to-end training with limited input data [24], [25], [26]. Due to its architecture resembling the letter "U," this method is named U-Net. The U-Net architecture can be seen in Figure 4.

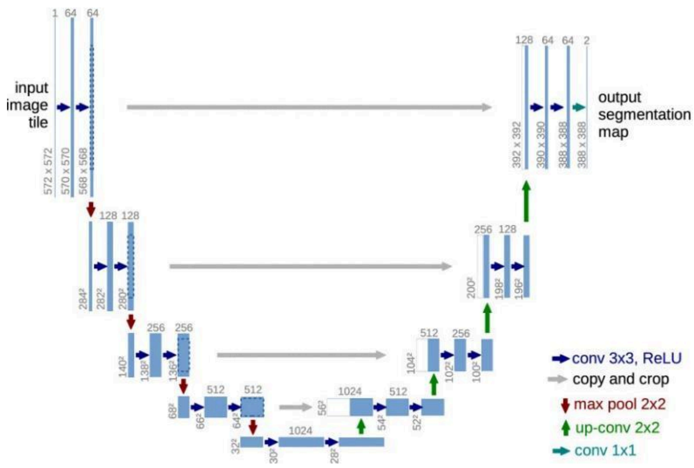

Figure 4. U-Net Architecture

Attention Gate Module

The Attention Gate (AG) in the U-Net model was first proposed by Oktay et al. (2018) [27]. This Attention module adaptively adjusts and automatically learns to focus on various shapes and sizes of target structures. Models enhanced with AG implicitly learn to highlight salient features useful for specific tasks while eliminating irrelevant areas in the input image. The AG module determines spatial regions by utilizing contextual information and activations obtained from the gating signal (g) at a coarser scale, as shown in Figure 5 below.

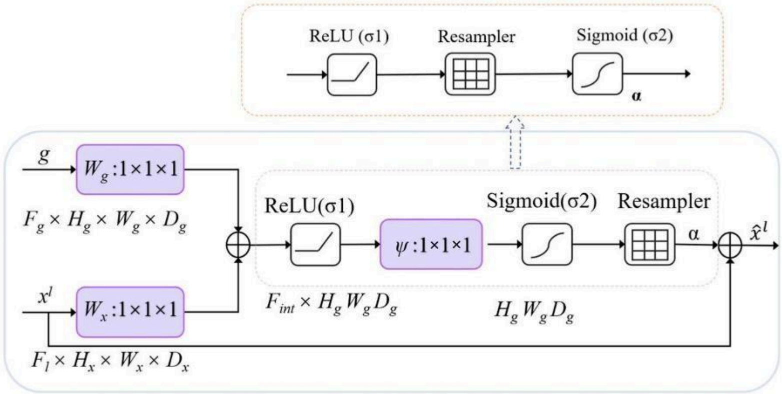

Figure 5. Attention Gate Module Structure

Metric Evaluation

Evaluation methods are crucial for assessing the performance of machine learning and deep learning models and ensuring reliability across various applications. These evaluation methods not only provide quantitative measures but also help in comparing different models, guiding improvements, and ensuring that the model can be generalized well to unseen data. The evaluation metrics used in this study are the Dice Coefficient and Intersection over Union (IoU). Both metrics have a scale from 0 to 1. A high value means that the model is more accurate and precise in predicting the area. IoU measures the similarity between the predicted mask and the ground truth by calculating the ratio of the intersection to the union. On the other hand, the Dice Coefficient metric is calculated as twice the area of overlap divided by the sum of the predicted area and the actual area.

RESULT AND DISCUSSION

After splitting the dataset into training, validation, and testing data, the next step is to train the Attention Gate U-Net model. Training is carried out using several different parameters to assess the best performance that can be produced by the model. Full details of the parameters used can be seen in Table 1.

Table 1. Trained Segmentation Model Parameters

Hyperparameter	Setting
Epochs	20
Batch Size	20
Optimizer	Adam, Adamax, AdamW, RMSProp
Learning Rate	0.001, 0.0001
Activation Function	ReLU, GELU
Loss Function	Dice Loss

In brain tumor classification experiments, the model training process is adjusted according to the combination of hyperparameters used to obtain optimal model performance.

Learning Rate and Optimizer Experiment

The first experiment was conducted by comparing the learning rate of several different optimizers, namely Adam, Adamax, AdamW, and RMSProp. The experiment was conducted using 20 epochs, 20 batch sizes, the ReLU activation function, and learning rates of 0.001 and 0.0001.

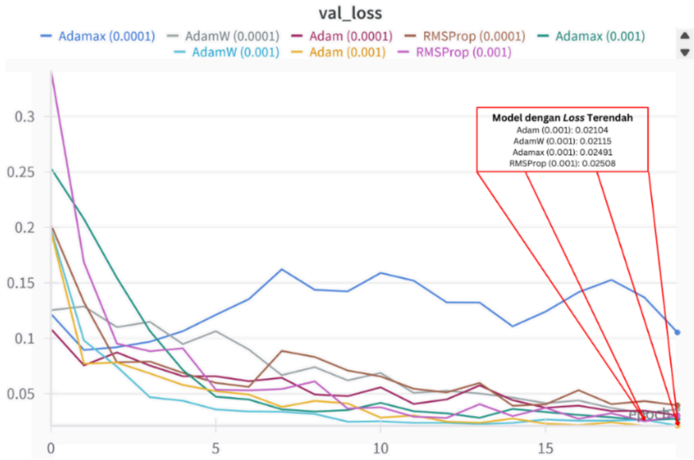

Figure 6. Testing Loss Segmentation against Learning Rate and Optimizer

The results showed in Figure 6 that at a learning rate of 0.001, Adam, Adamax, and AdamW provided faster and more stable validation loss reduction, indicating efficient convergence. Conversely, RMSProp showed the worst performance with the highest loss. Meanwhile, at a learning rate of 0.0001, the validation loss graph tended to be more volatile, especially for Adamax, RMSProp, and AdamW. Adam still showed relative stability, but the final results were still inferior to the configuration with a learning rate of 0.001.

The final evaluation shows that Adam (0.001) has the lowest validation loss value (0.0210), followed by AdamW (0.0211). Therefore, further experiments focus on the combination of Adam, AdamW, Adamax, and RMSProp optimizers with a learning rate of 0.001, as they demonstrate the best overall performance.

Activation Function Experiment

After determining that the optimal learning rate was 0.001 with Adam, AdamW, Adamax, and RMSProp optimizers, the study continued by evaluating the effect of two activation functions, ReLU and GELU, on model performance.

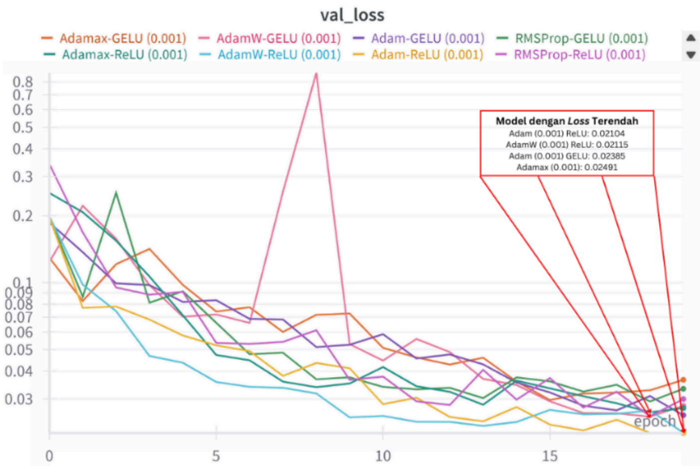

Figure 7. Testing Loss Segmentation on Activation Functions

The results of the experiment showed in Figure 7 that the combination with the ReLU activation function tended to produce a more stable validation loss graph and declined consistently across all optimizers. The Adam-ReLU and AdamW-ReLU combinations showed the best performance with low loss and a smooth training trend. Conversely, the use of GELU, especially with AdamW, caused significant fluctuations at the beginning of the epoch before eventually decreasing, indicating instability in the initial training process.

Nevertheless, GELU still shows fairly good results when used with Adam. The four configurations with the best results are Adam-ReLU (0.0210), AdamW-ReLU (0.0211), Adam-GELU (0.0253), and Adamax-ReLU (0.0272). Based on these results, the configurations to be used in subsequent experiments are combinations of ReLU with Adam, AdamW, and Adamax, as well as combinations of Adam with GELU, as they provide stable and near-optimal performance.

Accuracy Validation Comparison

Further experiments were conducted to evaluate the performance of the four best models using the Intersection over Union (IoU) metric and the Dice coefficient. Based on Figure 8, the model with the AdamW-ReLU combination showed the best performance with an IoU value of 0.6987, as well as a stable upward trend throughout training. Adamax-ReLU ranked second with a value of 0.6667, despite experiencing fluctuations in some epochs, but still showed a good upward trend. The Adam-ReLU model ranked third with an IoU of 0.6073 and a stable trend, though its improvement was not as high as the previous two models. Meanwhile, the Adam-GELU model had the lowest performance with an IoU of 0.5918 and significant fluctuations, indicating that this combination is less stable.

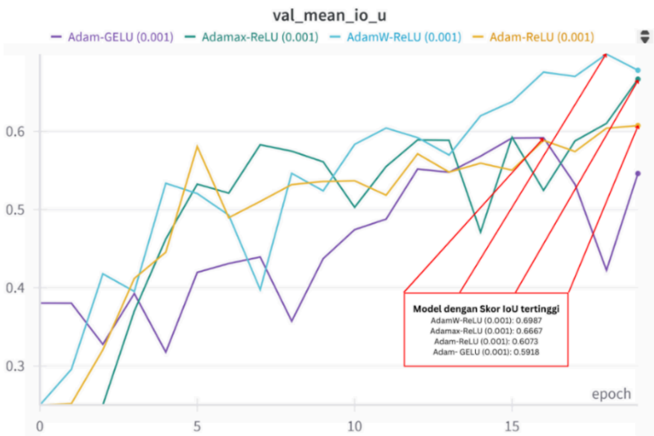

Figure 8. Comparison of IoU Scores on Segmentation Validation Data

The evaluation of the Dice coefficient in Figure 9 showed consistent results. AdamW-ReLU was again the best with a score of 0.8559, followed by Adam-ReLU with 0.8543. Although the values are nearly equivalent, Adam-ReLU tends to be more fluctuating. Adamax-ReLU records a score of 0.8330 with a relatively slow improvement, while Adam-GELU has the lowest score of 0.7974. The combination of Adam with GELU appears to be less optimal, possibly due to the complexity of the GELU function causing the learning process to be less stable.

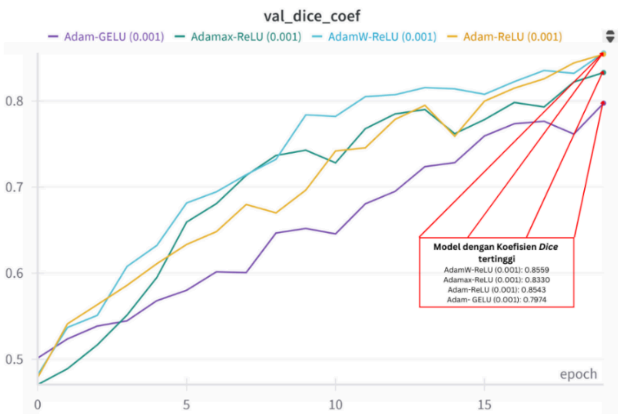

Figure 9. Comparison of Coefficient Dice Scores on Segmentation Validation Data

Overall, the combination of AdamW with ReLU proved to be the most consistent and effective, both in terms of IoU and Dice coefficient, indicating that this model is the most optimal for image segmentation tasks.

Comparison of Attention Gate U-Net with U-Net on Test Data

The final evaluation was conducted using 460 test images that had not been used during the training process, allowing for an objective measurement of the model's generalization. In this evaluation, four architectures were tested: U-Net and three AGU-Net variants using the Adam, AdamW, and Adamax optimizers. The test results showed that all AGU-Net variants outperformed U-Net in every evaluation metric, including the Dice coefficient and IoU value, as shown in Table 2.

Table 2. Comparison of Results Between Segmentation Methods

Model	Loss Value	Dice Coefficient	IoU Scores
U-Net	0.0200	0.8450	0.6071
AGU-Net Adam-ReLU 0.001	0.0195	0.8707	0.6519
AGU-Net Adamax-ReLU 0.001	0.0244	0.8503	0.7147
AGU-Net AdamW-ReLU 0.001	0.0193	0.8708	0.7270

Among all the variants tested, AGU-Net with the AdamW optimizer configuration, ReLU activation function, and learning rate of 0.001 proved to give the best results. This model recorded a significant improvement compared to the U-Net model with an 11.99% increase in IoU score and a 2.58% increase in the dice coefficient. This improvement indicates that the use of Attention Gate in the U-Net model, combined with AdamW, can enhance overall segmentation accuracy.

Table 3. Comparison of Segmentation Models with Previous Research

Researcher	Method	Dice Coefficient	IoU Score
Gupta et al. [19]	MAG-Net	74%	60%
Sheela & Suganthi [10]	Activate Contour Model dan Fuzzy C-Means	66.50%	-
Badža & Barjaktarović [21]	Convolutional Autoencoder	72.87%	-
Proposed Method	AGU-Net	87.08%	72.70%

Furthermore, the segmentation method, the Attention Gate U-Net (AGU-Net) model also showed competitive performance as shown in Table 3. When compared to the MAG-Net model by [19], which only achieved a Dice coefficient of 74% and an IoU of 60%, the segmentation methods Activate Contour Model and Fuzzy C-Means by [10] with a Dice coefficient of 66. 50%, or the Convolutional Autoencoder (CNA) method by [21] which achieved a Dice coefficient of 72.87%, AGU-Net shows a substantial improvement with a Dice coefficient of 87.98% and an IoU of 72.70%. This performance demonstrates that the Attention Gate enhances U-Net's ability to capture relevant areas of the image.

CONCLUSION

Based on the experimental results, the application of the attention mechanism in the AGU-Net model provides a significant performance improvement compared to the basic U-Net in the task of brain tumor segmentation. The Dice coefficient increased from 84.50% to 87.08%, while the Intersection over Union (IoU) value also rose from 60.71% to 72.70%. This improvement indicates that the integration of the attention mechanism not only enhances the model's ability to detect tumor types but also improves the accuracy of tumor localization in MRI images. Thus, the use of attention gates in the U-Net architecture highlights important features and ignores less relevant ones, resulting in more precise and reliable segmentation for medical applications. As a future development, this research can be expanded by testing the model on more diverse and complex datasets, including various types of brain tumors and other imaging modalities. Additionally, exploring the combination of attention mechanisms with data augmentation techniques or ensemble methods could be a potential step to further enhance model performance. Applying the model in real-time scenarios or integrating it with clinical diagnosis assistance systems also presents a promising area for further research, aiming to support faster and more accurate medical decision-making.

REFERENCES

[1] M. A. Khan *et al.*, "Brain tumor detection and classification: A framework of marker-based watershed algorithm and multilevel priority features selection," *Microsc Res Tech*, vol. 82, no. 6, 2019, doi: 10.1002/jemt.23238.

[2] M. Raza, M. Sharif, M. Yasmin, S. Masood, and S. Mohsin, "Brain image representation and rendering: A survey," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 4, no. 18, 2012.

[3] J. G. Park and C. Lee, "Skull stripping based on region growing for magnetic resonance brain images," *Neuroimage*, vol. 47, no. 4, 2009, doi: 10.1016/j.neuroimage.2009.04.047.

[4] M. Havaei *et al.*, "Brain tumor segmentation with deep neural networks," *Med Image Anal*, vol. 35, 2017, doi: 10.1016/j.media.2016.05.004.

[5] K. Abolanle AA *et al.*, "Brain tumor: An overview of the basic clinical manifestations and treatment," *Global Journal of Cancer Therapy*, 2020, doi: 10.17352/2581-5407.000034.

[6] M. Pichaivel, G. Anbumani, P. Theivendren, and M. Gopal, "An overview of brain tumor," *Brain Tumors*, pp. 1–10, 2022, [Online]. Available: www.intechopen.com

[7] S. Maqsood, R. Damasevicius, and F. M. Shah, "An efficient approach for the detection of brain tumor using fuzzy logic and U-Net CNN classification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2021. doi: 10.1007/978-3-030-86976-2_8.

[8] Q. Ke, J. Zhang, W. Wei, R. Damaševičius, and M. Woźniak, "Adaptive independent subspace analysis of brain magnetic resonance imaging data," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2893496.

[9] A. Wadhwa, A. Bhardwaj, and V. Singh Verma, "A review on brain tumor segmentation of MRI images," 2019. doi: 10.1016/j.mri.2019.05.043.

[10] C. J. J. Sheela and G. Suganthi, "Brain tumor segmentation with radius contraction and expansion based initial contour detection for active contour model," *Multimed Tools Appl*, vol. 79, no. 33–34, 2020, doi: 10.1007/s11042-020-09006-1.

[11] M. T. Nyo, F. Mebarek-Oudina, S. S. Hlaing, and N. A. Khan, "Otsu's thresholding technique for MRI image brain tumor segmentation," *Multimed Tools Appl*, vol. 81, no. 30, 2022, doi: 10.1007/s11042-022-13215-1.

[12] E. S. Biratu, F. Schwenker, T. G. Debelee, S. R. Kebede, W. G. Negera, and H. T. Molla, "Enhanced region growing for brain tumor MR image segmentation," *J Imaging*, vol. 7, no. 2, 2021, doi: 10.3390/jimaging7020022.

[13] A. Ilhan, B. Sekeroglu, and R. Abiyev, "Brain tumor segmentation in MRI images using nonparametric localization and enhancement methods with U-Net," *Int J Comput Assist Radiol Surg*, vol. 17, no. 3, 2022, doi: 10.1007/s11548-022-02566-7.

[14] M. Yu, X. Chen, W. Zhang, and Y. Liu, "AGs‑Unet: Building extraction model for high resolution remote sensing images based on attention gates u network," *Sensors*, vol. 22, no. 8, 2022, doi: 10.3390/s22082932.

[15] J. S. Suri *et al.*, "UNet deep learning architecture for segmentation of vascular and non-vascular images: A microscopic look at unet components buffered with pruning, explainable artificial intelligence, and bias," 2023. doi: 10.1109/ACCESS.2022.3232561.

[16] H. Xiao, L. Peng, S. Peng, and Y. Zhang, "Lung image segmentation based on involution UNet model," in *Proceedings - 2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering, AEMCSE 2022*, 2022. doi: 10.1109/AEMCSE55572.2022.00045.

[17] D. Maji, P. Sigedar, and M. Singh, "Attention Res-UNet with guided decoder for semantic segmentation of brain tumors," *Biomed Signal Process Control*, vol. 71, 2022, doi: 10.1016/j.bspc.2021.103077.

[18] D. Bahdanau, K. H. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[19] S. Gupta, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "MAG-Net: Multi-task attention guided network for brain tumor segmentation and classification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2021. doi: 10.1007/978-3-030-93620-4_1.

[20] C. J. J. Sheela and G. Suganthi, "Morphological edge detection and brain tumor segmentation in magnetic resonance (MR) images based on region growing and performance evaluation of modified Fuzzy C-means (FCM) algorithm," *Multimed Tools Appl*, vol. 79, no. 25–26, 2020, doi: 10.1007/s11042-020-08636-9.

[21] M. M. Badža and M. Barjaktarović, "Segmentation of brain tumors from MRI images using convolutional autoencoder," *Applied Sciences (Switzerland)*, vol. 11, no. 9, 2021, doi: 10.3390/app11094317.

[22] J. Cheng, "Brain Tumor Dataset," 2017, *figshare*. doi: 10.6084/m9.figshare.1512427.v8.

[23] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015. doi: 10.1007/978-3-319-24574-4_28.

[24] G. Du, X. Cao, J. Liang, X. Chen, and Y. Zhan, "Medical image segmentation based on U-Net: A review," *Journal of Imaging Science and Technology*, vol. 64, no. 2, 2020, doi: 10.2352/J.ImagingSci.Technol.2020.64.2.020508.

[25] N. Man, S. Guo, K. F. C. Yiu, and C. K. S. Leung, "Multi-layer segmentation of retina OCT images via advanced U-Net architecture," *Neurocomputing*, vol. 515, 2023, doi: 10.1016/j.neucom.2022.10.001.

[26] Z. Huang *et al.*, "A novel tongue segmentation method based on improved U-Net," *Neurocomputing*, vol. 500, 2022, doi: 10.1016/j.neucom.2022.05.023.

[27] O. Oktay *et al.*, "Attention U-Net: Learning where to look for the pancreas," Apr. 2018, [Online]. Available: http://arxiv.org/abs/1804.03999