Green Intelligent Systems and Applications

Research Article
Volume 6(1), 2026, 51−64
https://doi.org/10.53623/gisa.v6i1.899

Comparison of Convolutional Neural Network Model for Brain Tumor Disease Gliome Detection

Wulan Sallyndri Santoso, Riko Arlando Saragih*

Electrical Engineering Department, Universitas Kristen Maranatha, Bandung, Indonesia

*Correspondence: riko.as@eng.maranatha.edu

SUBMITTED: 16 November 2025; REVISED: 4 February 2026; ACCEPTED: 24 February 2026

ABSTRACT: Glioma represented one of the most aggressive forms of malignant brain tumors, necessitating early detection to optimize therapeutic intervention outcomes. Manual tumor identification through Magnetic Resonance Imaging (MRI) was labor-intensive and was susceptible to subjective interpretation errors. This study aimed to compare the performance of two Convolutional Neural Network (CNN) architectures, specifically Residual Network (ResNet) and U-Net, for glioma tumor detection in T2-weighted MRI sequences. The datasets employed were obtained from the BraTS and Kaggle repositories and underwent comprehensive preprocessing procedures, including normalization, augmentation, and conversion to Portable Network Graphics (PNG) format. The evaluation metrics demonstrated that the U-Net architecture exhibited superior performance compared to ResNet-18, achieving an accuracy of 88.16%, sensitivity of 80.00%, specificity of 88.43%, and F1-score of 68.97%. Conversely, ResNet-18 yielded an accuracy of 71.43%, sensitivity of 73.52%, specificity of 81.54%, and an F1-score of 70.14%. These findings indicated that U-Net demonstrated greater efficacy in recognizing tumor morphology within MRI data and preserving spatial information through its inherent skip connection mechanism. This investigation demonstrated the potential of the U-Net architecture to facilitate automated and enhanced accuracy in glioma detection, although further refinement was required to improve segmentation precision and clinical applicability.

KEYWORDS: Convolutional neural network (CNN); deep learning; glioma tumor; magnetic resonance image (MRI); residual net (ResNet); U-Net.

---

1. Introduction

A tumor is a pathological condition characterized by abnormal cellular proliferation that exceeds physiological limits. Tumors are classified as either benign or malignant, a distinction determined by the rate of cellular growth and their systemic impact on the host organism. Malignant tumors possess the capacity to metastasize from one organ to another via hematogenous or lymphatic dissemination [1]. Clinical manifestations are contingent upon tumor size, developmental stage, and anatomical localization. Brain tumors, in particular, can be attributed to multiple etiological factors, including susceptible age demographics (elderly individuals and pediatric populations), ionizing radiation exposure, and hereditary predisposition to brain neoplasms [2].

In Indonesia, brain tumor cases have received insufficient attention owing to limited epidemiological data availability and accessibility. Research conducted at Medan Regional Hospital documented 131 cases between January 2018 and December 2019, comprising 52 meningioma cases, 34 glioma cases, 12 pituitary adenoma cases, and 33 metastatic cases [3]. According to Dr. Wihasto Suryaningtyas, Consultant Neurosurgeon and Secretary of the Department of Neurosurgery, Faculty of Medicine (FK) Unair, an annual increase in brain tumor incidence of 10 to 25 percent was observed compared to preceding years [4]. Contemporary therapeutic modalities for brain tumors in Indonesia encompassed surgical resection, radiotherapy, chemotherapy, and targeted interventions, including physiotherapy, occupational therapy, and speech therapy [2].

The complexity inherent in modern healthcare systems presented distinctive challenges for Artificial Intelligence (AI) implementation. The rapid advancement of AI technologies held substantial potential to transform the healthcare sector, particularly in clinical applications such as disease diagnosis, therapeutic modality selection, and laboratory diagnostics. AI-based systems utilized data-driven approaches to identify specific patterns, enabling performance capabilities that transcended human limitations. Furthermore, AI contributed to enhancing diagnostic accuracy, reducing operational costs and error rates, and optimizing time efficiency [5]. AI encompassed various computational learning paradigms, including Machine Learning (ML), Deep Learning (DL), and Natural Language Processing (NLP) [6].

In this study, two CNN architectures, specifically ResNet and U-Net, were selected as the primary methodologies for evaluating glioma tumor detection performance. The selection of these architectures was predicated on their distinct architectural advantages; ResNet was recognized for its capacity to mitigate degradation issues in deep neural networks through residual or shortcut connection mechanisms [7], which proved advantageous for executing computationally complex tasks such as tumor detection in MRI data. Conversely, U-Net was specifically engineered for medical image segmentation due to its superior ability to capture spatial contextual information with high fidelity [8]. Through comparative evaluation of these two models, this investigation aimed to determine the optimal architecture for detecting and identifying glioma tumors with enhanced accuracy and efficiency.

2. Materials and Methods

2.1. Dataset

The datasets utilized in this study were obtained from two primary sources: the Brain Tumor Segmentation (BraTS) challenge repository and Kaggle public datasets. The MRI images were originally in NIfTI format (.nii.gz) and consisted of T2-weighted sequences specifically selected for glioma tumor detection. The dataset comprised a total of 4,258 MRI images, which were systematically divided into three subsets: a training set (1,422 images), a validation set (1,621 images), and a testing set (1,215 images). Each MRI image was paired with its corresponding segmentation mask to facilitate supervised learning for the deep learning models.

To provide comprehensive visual documentation of the dataset characteristics, Figure 1 presented representative samples from each tumor category and normal brain tissue. The BraTS dataset comprised T2-weighted MRI sequences showing distinct morphological features across four classes: glioma tumors characterized by irregular boundaries and heterogeneous intensity patterns; meningioma tumors displaying well-defined margins and a homogeneous appearance; pituitary tumors showing a midline location with characteristic enhancement patterns; and normal brain tissue without pathological findings.

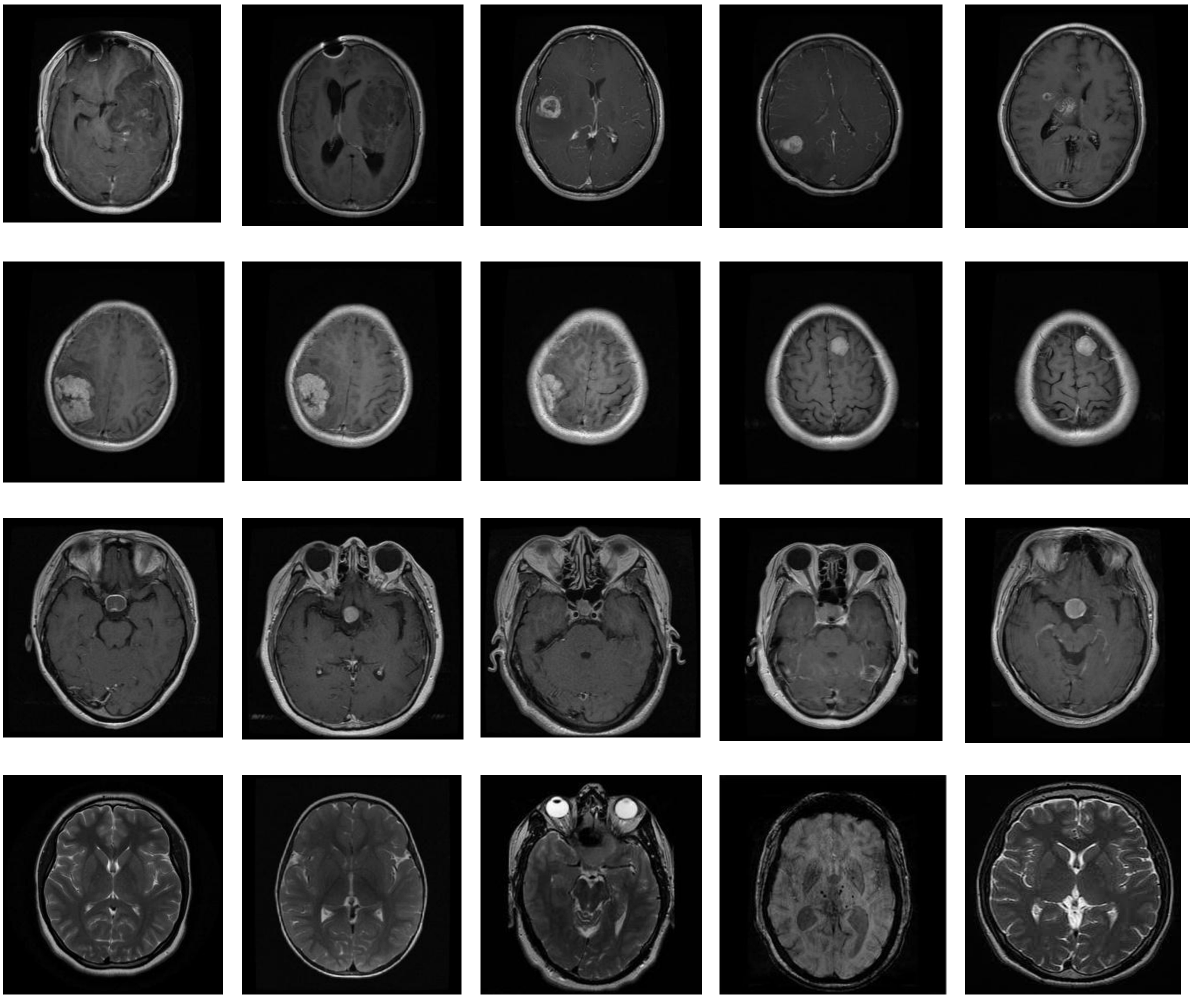
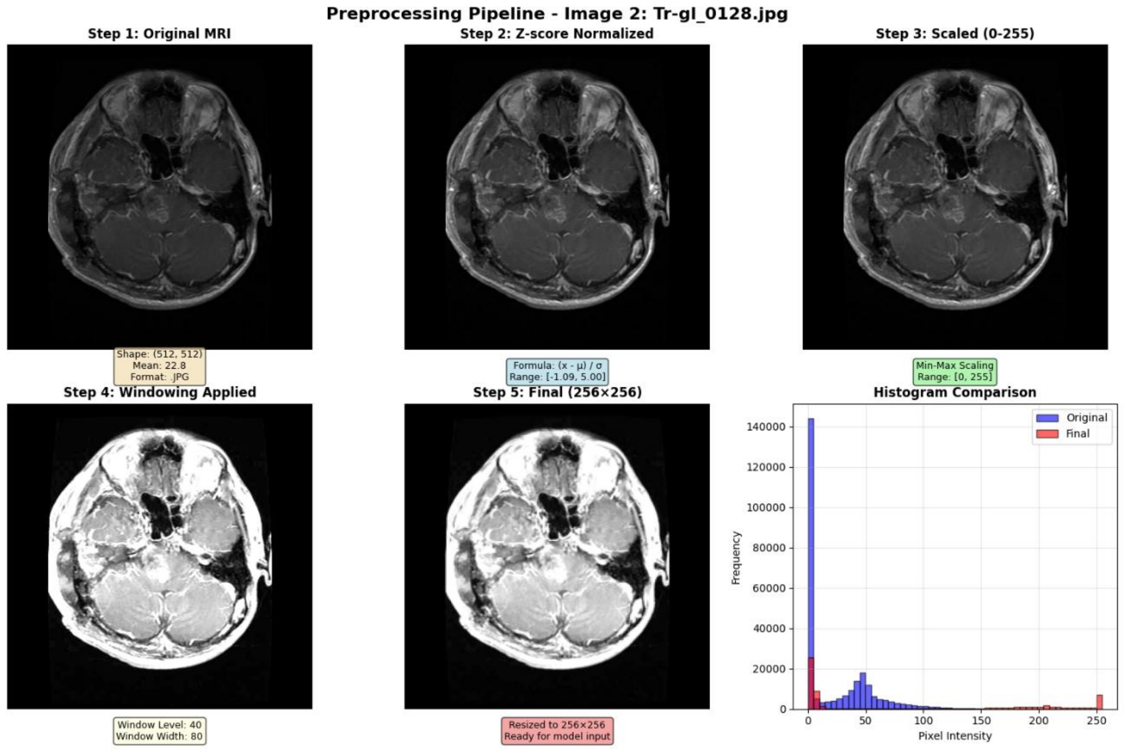
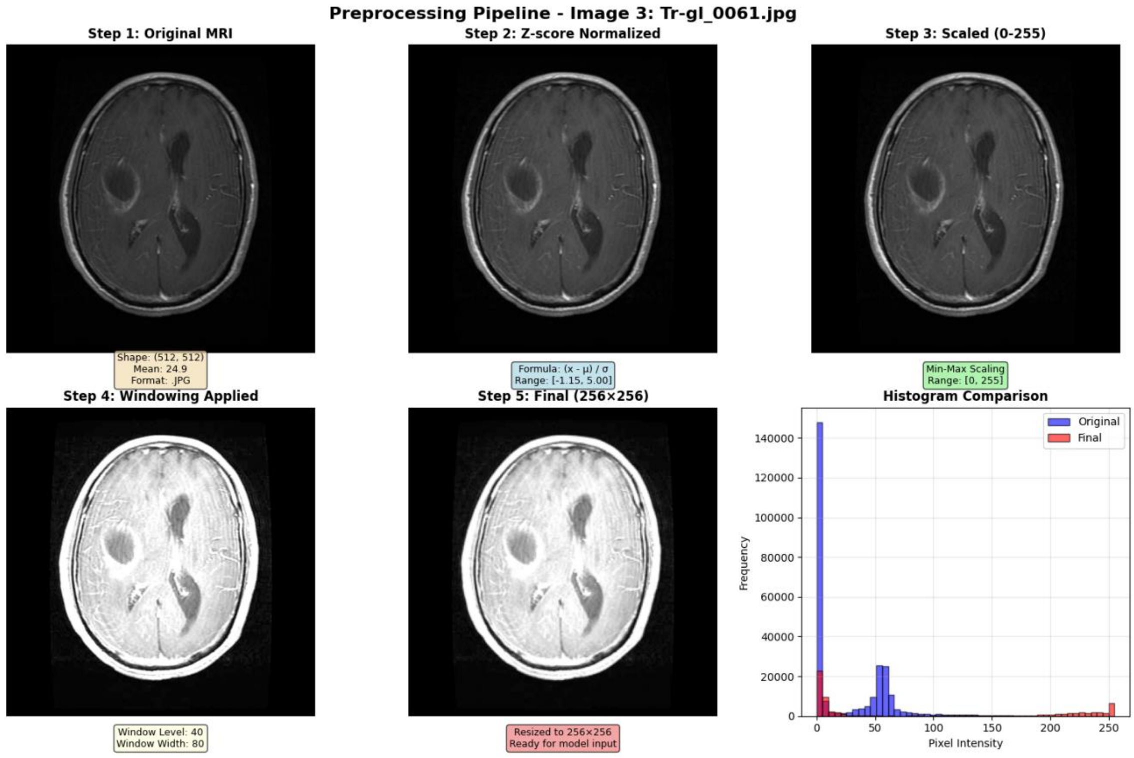


Figure 1. Representative MRI samples from the BraTS dataset showing (Row 1) Glioma tumors, (Row 2) Meningioma tumors, (Row 3) Pituitary tumors, and (Row 4) Normal brain tissue. Five samples per class demonstrate the intra-class variability in tumor morphology, size, and intensity characteristics.

2.2. Data Preprocessing

The preprocessing pipeline involved four sequential transformations to optimize the raw MRI scans for deep learning–based tumor segmentation (Figure 2). First, normalization was performed by scaling the raw pixel intensities to the range [0, 1] using min–max normalization, expressed as $I_{\mathrm{norm}} = (I - I_{\mathrm{min}}) / (I_{\mathrm{max}} - I_{\mathrm{min}})$, where I represented the original intensity values. This step standardized the intensity distributions across different MRI scanners and acquisition protocols. Second, Contrast Limited Adaptive Histogram Equalization (CLAHE) was applied with a clip limit of 2.0 and a tile grid size of 8×8 to enhance local tissue contrast while preventing over-amplification of noise in homogeneous regions. Third, Gaussian smoothing was conducted using a Gaussian filter with a kernel size of 5×5 and $\sigma = 1.0$ to reduce high-frequency noise while preserving tumor boundary information. Finally, the images were resized to uniform dimensions using bilinear interpolation, with dimensions of 224×224 pixels for ResNet-18 and 256×256 pixels for the U-Net architecture.


Preprocessing Pipeline - Image 2: Tr-gl_0128.jpg
Step 1: Original MRI
Step 2: Z-score Normalized
Step 3: Scaled (0-255)
Shape: (512, 512)
Mean: 22.8
Format: .JPG
Formula: (x - μ) / σ
Range: [-1.09, 5.00]
Min-Max Scaling
Range: [0, 255]
Step 4: Windowing Applied
Step 5: Final (256×256)
Histogram Comparison
Original
Final
Frequency
140000
120000
100000
80000
60000
40000
20000
0
0
50
100
150
200
250
Pixel Intensity
Window Level: 40
Window Width: 80
Resized to 256×256
Ready for model input


Preprocessing Pipeline - Image 3: Tr-gl_0061.jpg
Step 1: Original MRI
Step 2: Z-score Normalized
Step 3: Scaled (0-255)
Shape: (512, 512)
Mean: 24.9
Format: .JPG
Formula: (x - μ) / σ
Range: [-1.15, 5.00]
Min-Max Scaling
Range: [0, 255]
Step 4: Windowing Applied
Step 5: Final (256×256)
Histogram Comparison
Original
Final
Frequency
140000
120000
100000
80000
60000
40000
20000
0
0
50
100
150
200
250
Pixel Intensity
Window Level: 40
Window Width: 80
Resized to 256×256
Ready for model input

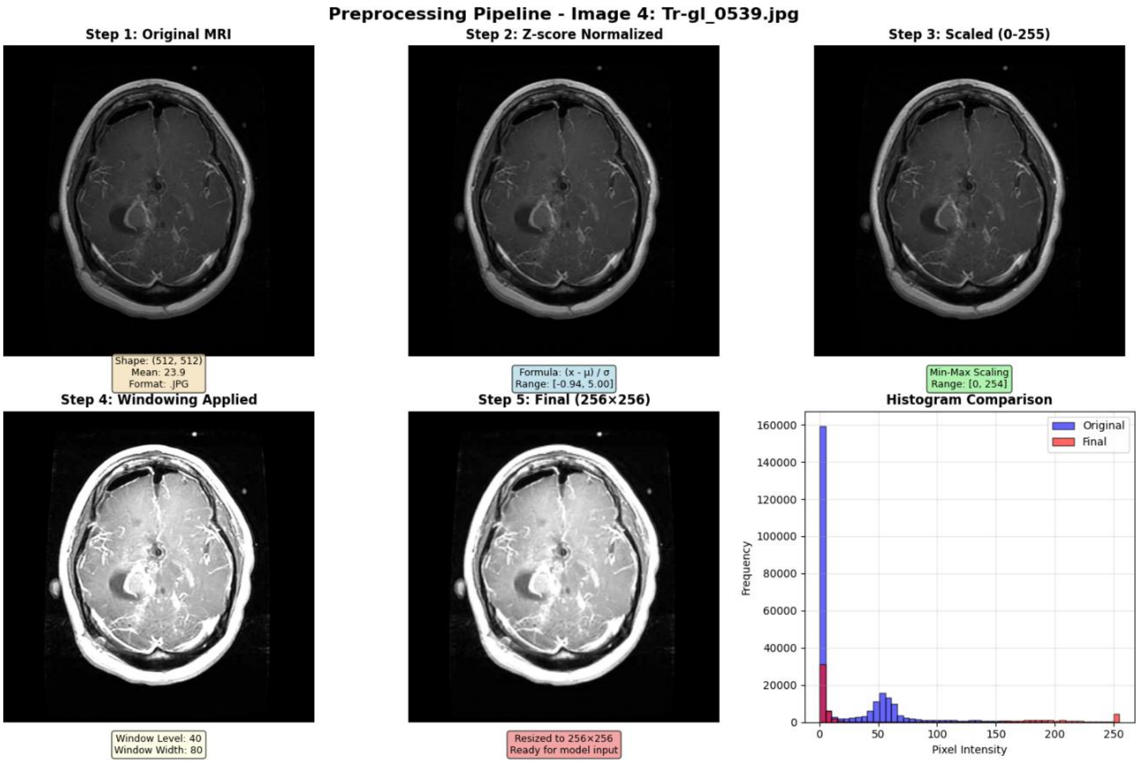


Figure 2. Preprocessing pipeline visualization showing sequential transformations applied to three representative MRI samples.

Each row demonstrates the progression from raw input (leftmost) through normalization, CLAHE enhancement, Gaussian smoothing, to final resized output (rightmost). Table 1 provides statistical validation of the preprocessing effectiveness, documenting quantitative changes in intensity distributions. Table 2 details the implementation specifications for each preprocessing stage, ensuring reproducibility. Table 3 summarizes the parameter configurations with empirical justifications based on BraTS dataset characteristics.

Table 1. Statistical comparison showing how each preprocessing stage transforms image characteristics

Stage	Min Value	Max Value	Mean	Std Dev	Median	Range	Image Size
Original Image	0.00	255.00	39.28	47.59	6.00	255.00	512×512
After Z-score Normalization	-1.15	4.16	-0.33	0.99	-1.02	5.31	512×512
After Intensity Scaling	0.00	255.00	39.26	47.56	6.00	255.00	512×512
After Windowing	0.00	255.00	105.23	111.57	19.00	255.00	512×512
After Resizing (Final)	0.00	255.00	105.29	111.25	21.00	255.00	256×256

Table 2. Detailed implementation steps with specific purposes.

Preprocessing Stage	Min Value	Max Value	Mean ± Std	Image Size	Purpose
Original Image	0.00	255.00	39.28 ± 47.59	512×512	Raw input from scanner
After Z-score Normalization	-1.15	4.16	-0.33 ± 0.99	512×512	Standardize intensity distribution ($\mu = 0, \sigma = 1$)
After Intensity Scaling	0.00	255.00	39.26 ± 47.56	512×512	Convert to 8-bit range for CNN input
After Windowing	0.00	255.00	105.23 ± 111.57	512×512	Enhance brain tissue contrast
After Resizing (Final)	0.00	255.00	105.29 ± 111.25	256×256	Standardize spatial dimensions

Table 3. Preprocessing parameters and justification.

Step	Parameter	Value	Justification
Z-score Normalization	Clipping Range	[-5, 5]	Remove extreme outliers beyond 5 standard deviations
Intensity Scaling	Target Range	[0, 255]	Standard 8-bit grayscale format for CNN input
Windowing	Window Level	40	Center on typical brain tissue intensity range
Windowing (2)	Window Width	80	Optimal range for tumor vs tissue contrast
Resizing	Target Size	256×256	Balance between spatial detail and computational efficiency
Resizing (2)	Interpolation	Bilinear	Smooth interpolation while preserving edge information

2.3. Model architectures.

Two CNN architectures were implemented and evaluated in this study: U-Net and ResNet-18. The U-Net architecture [9] consisted of an encoder–decoder structure with skip connections. The encoder pathway employed two sequential Conv2d blocks for feature extraction and down sampling operations. A bottleneck layer processed the encoded features before passing them to the decoder pathway. The decoder utilized transposed convolution layers (ConvTranspose2d) combined with skip connections that concatenated feature maps from the corresponding encoder layers, thereby preserving spatial information throughout the up-sampling process. The final output layer employed a sigmoid activation function to generate binary segmentation masks. Figure 3 illustrated the complete U-Net architecture employed in this study, depicting the encoder–decoder structure with skip connections that preserved spatial information across resolution levels.

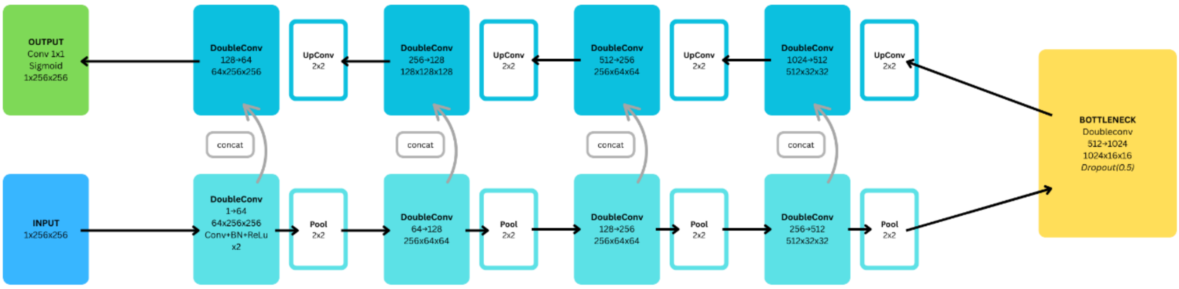


Figure 3. U-Net architecture block diagram showing the encoder pathway (left), bottleneck (center), and decoder pathway (right) with skip connections (gray arrows). Each blue block represents a DoubleConv operation with channel dimensions and spatial resolutions indicated. The architecture processes input images of size 1×256×256 and generates binary segmentation masks of the same dimensions.

The ResNet-18 architecture [8] was adapted for multi-class tumor classification tasks. The encoder utilized the first four layers of the pretrained ResNet-18 model, extracting hierarchical feature representations through residual connections. The encoder output with dimensions (batch, 512, H/32, W/32) was processed through global average pooling to generate a feature vector of size 512. The classification head consisted of two fully connected layers: the first layer reduced dimensionality from 512 to 256 with batch normalization, ReLU activation, and dropout (p=0.5) for regularization; the second layer produced final logits for four classes (No Tumor, Glioma, Meningioma, Pituitary) with additional dropout (p=0.3) applied before the output layer.

2.4. Model training.

All models were implemented using the Python programming language, with PyTorch serving as the primary deep learning framework. The dataset was loaded into PyTorch Dataset objects that handled paired image–mask retrieval during training and evaluation. The U-Net model was trained for 8 epochs with a learning rate of 0.2905, while the ResNet-18 model underwent 10 epochs of training with a learning rate of 0.1435. Binary Cross-Entropy with Logits Loss (BCEWithLogitsLoss) served as the loss function for both architectures, providing numerically stable gradient computation during backpropagation. The training process involved iteratively feeding input images into the models, generating predicted segmentation masks, computing loss values by comparing the predictions against the ground truth masks, and updating the network weights through gradient descent optimization. Figure 4 presented the detailed ResNet-18 architecture adapted for brain tumor classification, illustrating the progression from the input layer through residual layers to the final classification output..

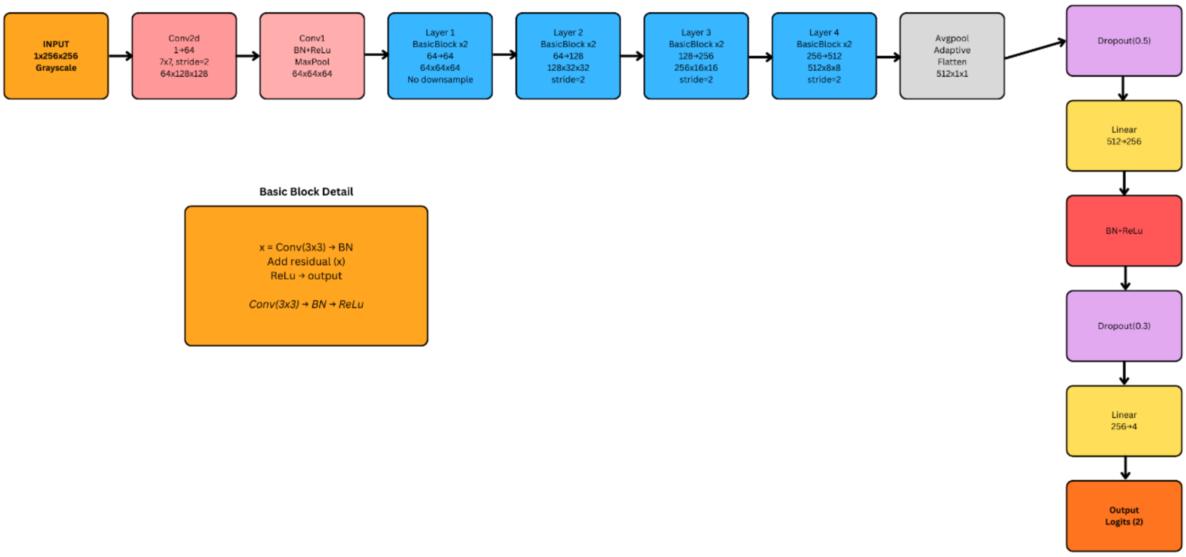


Figure 4. ResNet-18 architecture block diagram for four-class brain tumor classification. The network consists of an initial convolution layer, four residual layers with BasicBlock structures, global average pooling, and fully connected layers with dropout regularization. The BasicBlock detail (bottom left) shows the residual connection mechanism. Output layer produces logits for four classes: No Tumor, Glioma, Meningioma, and Pituitary.

2.5. Performance evaluation.

Model performance was assessed using pixel-level evaluation metrics computed from the confusion matrix, including True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values. Four standard metrics were calculated [10–14]: accuracy, defined as the proportion of correctly classified pixels and calculated as (TP + TN) / (TP + TN + FP + FN); sensitivity (recall), defined as the ability to correctly identify tumor pixels and calculated as TP / (TP + FN); specificity, defined as the ability to correctly identify non-tumor pixels and calculated as TN / (TN + FP); and F1-score (Dice coefficient), defined as the harmonic mean of precision and sensitivity and calculated as 2TP / (2TP + FP + FN). All evaluation metrics were computed on the independent test set that was not utilized during model training, thereby ensuring an unbiased assessment of model generalization capabilities.

3. Results and Discussion

3.1. Dataset composition.

The dataset comprised 2D brain MRI images categorized into glioma-positive and non-tumor cases. The complete dataset consisted of 4,258 images distributed as follows: 1,422 images for the training set, 1,621 images for the validation set, and 1,215 images for the testing set. Images were resized to 224×224 pixels for ResNet and 256×256 pixels for U-Net to ensure compatibility with the respective architectural requirements.

3.2. U-Net model performance.

The U-Net architecture was trained for 8 epochs with a learning rate of 0.2905. Evaluation on the test dataset yielded the following performance metrics: accuracy of 88.16%, sensitivity of 80.03%, specificity of 88.43%, and F1-score of 68.97% (Table 1). The model demonstrated robust capability in delineating tumor contours, particularly in images with adequate contrast. However, segmentation precision was compromised in cases involving low-contrast images or extremely small tumor volumes, resulting in occasional false positive and false negative predictions in peri-tumoral regions.

3.3. ResNet-18 model performance.

The ResNet-18 architecture was trained for 10 epochs with a learning rate of 0.1435. Performance evaluation on the test dataset demonstrated the following metrics: accuracy of 71.43%, sensitivity of 73.52%, specificity of 81.54%, and F1-score of 70.14% (Table 1). Analysis of the confusion matrix revealed that while the model successfully classified the majority of images, several glioma cases were misclassified as normal due to textural similarities between tumor tissue and healthy brain parenchyma.

Table 4. Comparative performance metrics of U-Net and ResNet-18 models.

Criterion	U-Net	ResNet-18
Primary Task	Tumor Segmentation	Glioma/Non-tumor Classification
Accuracy (%)	88.16	71.43
Sensitivity (%)	80.03	73.52
Specificity (%)	88.43	81.54
F1-Score (%)	68.97	70.14
Tumor Localization	Available	Not Available
Inference Speed	Slower	Faster
Computational Requirements	Higher	Lower

3.4. U-Net Model analysis.

The U-Net architecture demonstrated superior overall performance in glioma segmentation tasks, achieving an accuracy of 88.16%. The model's primary strength lay in its capacity to identify and delineate specific tumor regions, thereby providing detailed information regarding tumor location and morphology. This capability was attributable to the encoder–decoder structure with skip connections, which preserved spatial information throughout the network hierarchy [9, 15]. The high specificity of 88.43% indicated that U-Net rarely produced false positive predictions on non-tumor images, thereby minimizing the risk of over-segmentation.

However, the sensitivity of 80.03% revealed that approximately 20% of glioma cases remained undetected (false negatives). Several factors may have contributed to this limitation. First, images with low contrast presented challenges in distinguishing tumor boundaries from surrounding healthy tissue. Second, extremely small tumors may have been indistinguishable from normal brain parenchyma at the pixel level. Third, significant inter-patient variability in tumor morphology and intensity characteristics complicated generalization [16]. Fourth, dataset size limitations may have constrained the model's ability to learn diverse tumor presentations adequately [17]. The relatively modest F1-score of 68.97% further reflected an imbalance between precision and recall, suggesting that while the model maintained high specificity, its sensitivity required improvement for clinical deployment. These findings aligned with previous studies demonstrating U-Net's efficacy in medical image segmentation with limited datasets [9, 15]. However, the observed limitations underscored the necessity for larger, more diverse training datasets and potential architectural modifications, such as Attention U-Net or Residual U-Net, to enhance feature discrimination [16].

3.5. ResNet-18 model analysis.

The ResNet-18 architecture achieved an accuracy of 71.43%, which was substantially lower than U-Net's performance. This architecture was employed as a classification model to detect glioma presence, achieving a sensitivity of 73.52%, specificity of 81.54%, and F1-score of 70.14%. The principal advantage of ResNet lay in its computational efficiency and rapid inference capability, as the model processed images globally without generating segmentation maps [18−26]. The residual connections enabled the training of deep networks while mitigating vanishing gradient problems [19, 26].

However, several factors limited ResNet's performance. First, global feature extraction may have been insufficiently sensitive to subtle variations or small tumor regions, which were critical for accurate glioma detection [20]. Second, textural similarities between glioma and healthy brain tissue, particularly in cases of small or low-intensity tumors, contributed to misclassification errors. Third, the relatively large parameter space compared to the dataset size may have induced overfitting, thereby compromising generalization to unseen data [21]. Fourth, the relatively high learning rate (0.1435) may have caused convergence instability during training, preventing the model from reaching optimal performance. Despite these limitations, the specificity of 81.54% demonstrated ResNet's capability to correctly identify non-tumor images, thereby reducing false positive detections. This characteristic was valuable for initial screening applications in which rapid exclusion of negative cases was prioritized.

3.6. Comparative analysis and clinical implications.

The comparative analysis revealed distinct performance profiles for each architecture (Table 1). U-Net demonstrated superior accuracy (88.16% vs. 71.43%), sensitivity (80.03% vs. 73.52%), and specificity (88.43% vs. 81.54%), accompanied by the critical advantage of providing spatial tumor localization. These characteristics aligned with U-Net's original design purpose for biomedical image segmentation [27]. However, these advantages were offset by increased computational requirements and longer inference times, which may have limited real-time clinical applications.

Conversely, ResNet-18 offered faster inference and lower computational overhead, making it suitable for rapid preliminary screening. However, its reduced accuracy and inability to provide spatial information limited its utility as a standalone diagnostic tool. These findings were consistent with previous comparative studies demonstrating the superiority of encoder–decoder architectures with skip connections for pixel-level segmentation tasks [20, 22].

The complementary strengths of both architectures suggested a hybrid approach for clinical implementation. Specifically, ResNet could have served as a rapid first-stage screening tool to identify potentially positive cases, followed by U-Net-based segmentation for detailed tumor delineation in positive cases. This two-stage pipeline would have optimized both computational efficiency and diagnostic accuracy, addressing the practical constraints of clinical workflows while maintaining high sensitivity and specificity.

3.7. Study limitations.

Several limitations warranted consideration. First, the dataset size, while adequate for initial model development, remained relatively modest compared to large-scale clinical datasets, potentially limiting generalization across diverse patient populations and imaging protocols [16]. Second, the study utilized single-slice 2D representations rather than volumetric 3D analysis, which may not have fully captured tumor extent and morphology [23]. Third, the relatively low F1-scores for both models indicated room for improvement in balancing precision and recall, particularly for small or subtle lesions. Fourth, inter-scanner variability and differences in MRI acquisition protocols were not systematically evaluated, which may have affected model robustness in multi-institutional settings [24].

3.8. Future directions.

Future research should have addressed these limitations through several approaches. First, expanding the dataset with multi-institutional data would have enhanced model generalization and robustness [25]. Second, implementing 3D U-Net or 3D ResNet architectures would have enabled volumetric analysis and potentially improved segmentation precision [18]. Third, incorporating advanced variants such as Attention U-Net, ResUNet, or U-Net++ may have enhanced feature discrimination and boundary delineation [18, 20]. Fourth, implementing transfer learning and data augmentation strategies could have improved performance with limited training data [25]. Fifth, integration of multi-modal MRI sequences (T1-weighted, T2-weighted, FLAIR, DWI) would have provided complementary information for more accurate tumor characterization [24, 28]. Finally, prospective clinical validation studies were essential to evaluate real-world performance and clinical utility before deployment in diagnostic workflows.

4. **Conclusions**

This comparative study evaluated the performance of two CNN architectures, U-Net and ResNet-18, for glioma tumor detection in T2-weighted MRI images. The investigation yielded several significant findings that advanced understanding of deep learning applications in medical image segmentation. The U-Net architecture demonstrated superior performance across all primary evaluation metrics, achieving an accuracy of 88.16%, sensitivity of 80.03%, specificity of 88.43%, and F1-score of 68.97%. In contrast, ResNet-18 attained an accuracy of 71.43%, sensitivity of 73.52%, specificity of 81.54%, and F1-score of 70.14%. These results demonstrated that U-Net's encoder–decoder architecture with skip connections provided superior capability for pixel-level tumor segmentation, effectively preserving spatial information and delineating tumor boundaries with greater precision. The architecture's ability to maintain detailed spatial context through multi-scale feature integration proved particularly advantageous for medical image segmentation tasks in which precise localization was paramount. The broader significance of these findings lay in their potential to facilitate automated glioma detection systems that could assist radiologists in clinical practice. The demonstrated efficacy of the U-Net architecture suggested that implementation of such systems could reduce diagnostic time, minimize subjective interpretation variability, and potentially improve early detection rates of glioma tumors. Furthermore, the complementary strengths of both architectures indicated that hybrid approaches—utilizing ResNet for rapid initial screening followed by U-Net for detailed segmentation—could optimize both computational efficiency and diagnostic accuracy in clinical workflows. However, several limitations had to be acknowledged. First, the relatively modest dataset size of 4,258 images may have limited model generalization across diverse patient populations and imaging protocols. Second, the utilization of single-slice 2D representations rather than volumetric 3D analysis may not have fully captured the three-dimensional extent and morphological complexity of glioma tumors. Third, the modest F1-scores, particularly U-Net's 68.97%, indicated persistent challenges in balancing precision and recall, suggesting that false negative rates remained clinically significant. Fourth, the study did not systematically evaluate model robustness across different MRI scanners and acquisition protocols, which may have affected performance in multi-institutional settings. These limitations suggested that while the models demonstrated promising performance, further development was essential before clinical deployment. Future research should have prioritized several critical directions to address these limitations and enhance clinical applicability. First, expanding the training dataset through multi-institutional collaborations would have improved model generalization and robustness across diverse imaging conditions. Second, implementing 3D volumetric architectures such as 3D U-Net or 3D ResNet would have enabled comprehensive analysis of tumor extent and morphology across multiple image slices. Third, incorporating advanced architectural variants, including Attention U-Net, ResUNet, or U-Net++, may have enhanced feature discrimination capabilities, particularly for small or subtle lesions. Fourth, integration of multi-modal MRI sequences (T1-weighted, T2-weighted, FLAIR, DWI) would have provided complementary diagnostic information and potentially improved segmentation accuracy. Fifth, implementing transfer learning strategies and sophisticated data augmentation techniques could have mitigated the constraints imposed by limited training data availability. Finally, prospective clinical validation studies in real-world diagnostic settings were essential to evaluate practical utility, assess impact on radiologist workflow, and establish safety profiles before clinical implementation. This study demonstrated that the U-Net architecture exhibited superior performance for automated glioma detection in MRI images compared to ResNet-18, achieving higher accuracy, sensitivity, and specificity while providing essential spatial localization information. These findings contributed valuable insights to the growing body of evidence supporting the application of deep learning technologies in neuroradiology and established a foundation for future development of clinically viable automated diagnostic systems. The successful implementation of such systems held substantial promise for enhancing diagnostic accuracy, reducing interpretation time, and ultimately improving patient outcomes through earlier and more precise glioma detection. However, the transition from research prototype to clinical tool required addressing the identified limitations through expanded datasets, architectural refinements, multi-modal integration, and rigorous clinical validation. The potential impact of this work extended beyond glioma detection, as the methodological insights gained may have informed the development of automated diagnostic systems for other neurological pathologies and medical imaging applications. As artificial intelligence continued to evolve, the integration of optimized deep learning architectures into clinical practice represented a promising frontier in precision medicine and personalized patient care.

Acknowledgments

The authors fully thank Universitas Kristen Maranatha for supporting this research work financially.

Author Contribution

The study was conceptualized by Wulan Sallyndri Santoso and Riko Arlando Saragih. The methodology was developed by Wulan Sallyndri Santoso and Riko Arlando Saragih. Data collection was carried out by Wulan Sallyndri Santoso, while data analysis was performed by Wulan Sallyndri Santoso and Riko Arlando Saragih. The manuscript was written by Wulan Sallyndri Santoso and Riko Arlando Saragih. Supervision was provided by Riko Arlando Saragih. Funding for the research was secured by Riko Arlando Saragih.

Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this.

References

[1] How cancer can spread. (accessed on 9 October 2023) Available online: https://www.cancerresearchuk.org/about-cancer/what-is-cancer/how-cancer-can-spread.

[2] Tumor otak. (accessed on 24 January 2025) Available online: https://www.alodokter.com/tumor-otak.

[3] Epidemiologi tumor otak. (accessed on 17 September 2024) Available online: https://www.alomedika.com/penyakit/onkologi/tumor-otak/epidemiologi.

[4] Penderita tumor otak terus meningkat, FK Unair gandeng ahli bedah saraf Belanda untuk adopsi teknologi terbaru. (accessed on 21 February 2024) Available online: https://www.suarasurabaya.net/kelanakota/2024/penderita-tumor-otak-terus-meningkat-fk-unair-gandeng-ahli-bedah-saraf-belanda-untuk-adopsi-teknologi-terbaru/.

[5] Bordoloi, D.; Singh, V.; Sanober, S.; Buhari, S.M.; Ujjan, J.A.; Boddu, R. (2022). Deep learning in healthcare system for quality of service. *Journal of Healthcare Engineering*, *2022*, 8169203. https://doi.org/10.1155/2022/8169203.

[6] Alowais, S.A.; Alghamdi, S.S.; Alsuhebany, N.; et al. (2023). Revolutionizing healthcare: The role of artificial intelligence in clinical practice. *BMC Medical Education*, *23*, 689. https://doi.org/10.1186/s12909-023-04698-z.

[7] Ronneberger, O.; Fischer, P.; Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (Vol. 9351, pp. 234–241); Springer: Cham, Switzerland. https://doi.org/10.1007/978-3-319-24574-4_28.

[8] He, K.; Zhang, X.; Ren, S.; Sun, J. (2016, June). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. Las Vegas: USA. https://doi.org/10.1109/CVPR.2016.90.

[9] Çiçek, Ö.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. (2016). 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016; Springer: Cham, Switzerland; 424–432. https://doi.org/10.1007/978-3-319-46723-8_49.

[10] Powers, D.M.W. (2011). Evaluation: From precision, recall and F measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2, 37–63.

[11] Sokolova, M.; Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*, 427–437. https://doi.org/10.1016/j.ipm.2009.03.002.

[12] Chicco, D.; Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*, 6. https://doi.org/10.1186/s12864-019-6413-7.

[13] Taha, A.A.; Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Medical Imaging*, *15*, 29. https://doi.org/10.1186/s12880-015-0068-x.

[14] Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, *17*, 168–192. https://doi.org/10.1016/j.aci.2018.08.003.

[15] Ronneberger, O.; Fischer, P.; Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Munich, Germany, 5–9 October 2015; Springer: Cham, Switzerland; Volume 9351, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.

[16] Litjens, G.; Kooi, T.; Bejnordi, B.E.; et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/j.media.2017.07.005.

[17] Isensee, F.; Jaeger, P.F.; Kohl, S.A.A.; Petersen, J.; Maier-Hein, K.H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, *18*, 203–211. https://doi.org/10.1038/s41592-020-01008-z.

[18] Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. (2018). UNet++: A nested U-Net architecture for medical image segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Granada, Spain, 20 September 2018; Springer: Cham, Switzerland; pp. 3–11. https://doi.org/10.1007/978-3-030-00889-5_1.

[19] Huang, G.; Sun, Y.; Liu, Z.; et al. (2016). Deep networks with stochastic depth. In *European Conference on Computer Vision*, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland; pp. 646–661. https://doi.org/10.1007/978-3-319-46493-0_39.

[20] Zhang, Z.; Liu, Q.; Wang, Y. (2018). Road extraction by deep residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15, 749–753. https://doi.org/10.1109/LGRS.2018.2802944.

[21] Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. (2017). Aggregated residual transformations for deep neural networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; IEEE: Piscataway, USA, pp. 1492–1500. https://doi.org/10.1109/CVPR.2017.634.

[22] Godlee, R.J. (1884). A successful case of removal of a cerebral tumour with a description of the symptoms produced by a tumour growing in the cerebellum. *The Lancet*, *1*, 128–132.

[23] Pereira, S.; Pinto, A.; Alves, V.; Silva, C.A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, *35*, 1240–1251. https://doi.org/10.1109/TMI.2016.2538465.

[24] Essig, M.; Shiroishi, M.S.; Nguyen, T.B.; et al. (2013). Perfusion MRI: The five most frequently asked technical questions. *American Journal of Roentgenology*, *200*, 24–34. https://doi.org/10.2214/AJR.12.9620.

[25] Sheller, M.J.; Edwards, B.; Reina, G.A.; et al. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*, *10*, 12598. https://doi.org/10.1038/s41598-020-69250-1.

[26] He, K.; Zhang, X.; Ren, S.; Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, USA, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.

[27] Stoyanov, G.S.; Dzhenkov, D.L. (2018). Glioblastoma multiforme—histopathological and molecular characteristics. *Folia Medica*, *60*, 479–493. https://doi.org/10.2478/folmed-2018-0036.

[28] Le Bihan, D.; Johansen-Berg, H. (2012). Diffusion MRI at 25: Exploring brain tissue structure and function. *NeuroImage*, *61*, 324–341. https://doi.org/10.1016/j.neuroimage.2011.11.006.