

Multi-Class Brain Tumor Segmentation and Classification in MRI Using a U-Net and Machine Learning Model

Jackri Hendrik^{*1}, Octara Pribadi², Hendri³, Leony Hoki⁴, Feriani Astuti Tarigan⁵, Edi Wijaya⁶, Rabei Raad Ali⁷

^{1,2,3,4}Study Program in Informatics Engineering, STMIK Time, Indonesia

^{5,6}Study Program in Information Systems, STMIK Time, Indonesia

⁷Department of Computer Science, Northern Technical University, Iraq

Email: ¹jackrihendrik@stmik-time.ac.id

Received : Oct 3, 2025; Revised : Oct 14, 2025; Accepted : Oct 16, 2025; Published : Oct 23, 2025

Abstract

Brain tumor diagnosis remains a critical challenge in medical imaging, as accurate classification and precise localization are essential for effective treatment planning. Traditional diagnostic approaches often rely on manual interpretation of MRI scans, which can be time-consuming, subjective, and prone to variability across radiologists. To address this limitation, this study proposes a two-stage framework that integrates machine learning (ML) based classifiers for tumor type recognition and a U-Net architecture for tumor segmentation. The classifier was trained to distinguish four tumor categories: glioma, meningioma, pituitary, and no tumor, while the U-Net model was employed to delineate tumor regions at the pixel level, enabling volumetric assessment. The novelty of this research lies in its dual focus that combines classification and segmentation within a single framework, which enhances clinical applicability by offering both diagnostic and spatial insights. Experimental results demonstrated that among the evaluated classifiers, XGBoost achieved the highest accuracy of 86 percent, surpassing other models such as Random Forest, SVC, and Logistic Regression, while the U-Net model delivered consistent segmentation performance across tumor types. These findings highlight the potential of hybrid ML and deep learning solutions to improve reliability, efficiency, and objectivity in brain tumor analysis. In real-world practice, the proposed framework can serve as a valuable decision-support tool, assisting radiologists in early detection, reducing diagnostic workload, and supporting personalized treatment strategies.

Keywords : Brain Tumor Classification, Computer Vision, Machine Learning, Medical Image Processing, U-Net Segmentation.

This work is an open access article and licensed under a Creative Commons Attribution-Non Commercial 4.0 International License



1. INTRODUCTION

Brain tumors represent a major clinical challenge as they are among the most aggressive neurological disorders, contributing significantly to global morbidity and mortality [1], [2], [3]. Accurate diagnosis and characterization are essential for determining effective treatment strategies, yet conventional radiological evaluation based on manual inspection of magnetic resonance imaging (MRI) scans remains limited by subjectivity, time constraints, and substantial inter-observer variability [4], [5], [6]. The intrinsic heterogeneity of brain tumors, reflected in irregularities of size, shape, intensity distribution, and textural patterns across different MRI modalities such as T1, T2, and FLAIR, makes precise delineation and classification particularly difficult [7], [8]. Furthermore, many existing computer-aided diagnosis approaches address either segmentation or classification in isolation, which often leads to incomplete representation of the tumor and reduced generalization when applied in real clinical environments [9], [10]. These limitations underline a critical gap in current diagnostic methodologies and emphasize the need for more integrated and automated solutions capable of reliably

segmenting tumor regions while simultaneously classifying multiple tumor types with high accuracy and robustness.

To overcome these limitations, recent advances in artificial intelligence have introduced powerful computational approaches that combine the representational capacity of deep learning with the discriminative power of machine learning algorithms [11], [12], [13]. Deep learning models, particularly convolutional neural networks and their derivatives, have demonstrated remarkable success in capturing spatial and structural information from MRI scans, making them highly effective for tumor localization and boundary segmentation [14], [15], [16]. Among these, encoder–decoder architectures such as U-Net have become widely adopted due to their ability to preserve fine-grained spatial features while learning hierarchical contextual representations of medical images [17], [18]. However, segmentation alone cannot fully address the clinical requirement of differentiating among multiple tumor categories [19], [20]. Therefore, machine learning classifiers can be integrated to exploit the features extracted from deep learning models, enabling accurate multi-class categorization of tumor types such as glioma, meningioma, pituitary adenoma, and normal tissue. This synergistic paradigm not only mitigates the shortcomings of traditional manual assessment but also provides a robust and scalable framework that enhances diagnostic precision, reduces inter-observer variability, and holds significant potential for deployment in real-world healthcare environments [21], [22], [23].

Study by [24] introduce BiTr-UNet, a CNN–Transformer hybrid for multi-modal MRI brain tumor segmentation on BraTS 2021. The network adopts a U-Net-style encoder–decoder but injects bidirectional Transformer blocks to capture long-range dependencies while preserving fine spatial detail via skip connections. This design improves Dice across whole tumor, tumor core, and enhancing tumor compared with pure CNN baselines, suggesting that global self-attention complements local convolutional features in heterogeneous glioma. The study’s contributions include a clean architectural recipe, extensive ablations, and competitive BraTS validation and test scores. However, the self-attention modules introduce notable memory and compute overhead, constraining 3D patch sizes and batch settings, which can reduce practicality on standard clinical GPUs. In addition, training stability and hyperparameter sensitivity around attention depth and tokenization require careful tuning. Finally, while performance is strong on BraTS, cross-domain generalization to non-challenge clinical scans is not fully assessed, leaving external robustness an open question.

Study by [25] propose A Lightweight U-Net for MRI brain tumor segmentation, integrating a spatial attention mechanism to emphasize tumor-relevant regions while keeping parameter counts low. The method targets precise boundary delineation with reduced inference time, reporting improved Dice and IoU versus conventional U-Net variants on publicly available MRI datasets. Key contributions include an efficiency-accuracy trade-off favorable for constrained hardware and an attention design that remains compatible with standard U-Net training pipelines. The study also details training choices and augmentation that stabilize small-lesion detection. Drawbacks include evaluation on limited datasets and scanners, which leaves questions about robustness to protocol variability, class imbalance across subregions, and post-contrast versus multi-modal inputs. Additionally, lightweight models can underfit complex textures in infiltrative gliomas, and the paper offers limited analysis of uncertainty calibration or failure cases. Future work should expand multi-center validation and compare against Transformer-enhanced or diffusion-based segmenters.

Study by [26] develops a VGG16-based deep learning classifier for four-class brain tumor recognition on MRI, covering glioma, meningioma, pituitary, and no-tumor. Leveraging transfer learning, fine-tuning, and data augmentation on a large, diverse dataset exceeding 17k images, the approach achieves high accuracy and competitive precision–recall, demonstrating that well-tuned classical CNN backbones remain strong baselines for multiclass diagnosis. Contributions include a transparent training pipeline, comparative analyses versus alternative architectures, and thorough

reporting of class-wise metrics. Nonetheless, the study is primarily 2D slice-based and does not explicitly model volumetric context or multi-modal fusion, which can limit sensitivity to subtle subregional patterns. The reliance on aggregated public datasets may introduce dataset bias or leakage if patient-level splits are not strictly enforced. Moreover, external validation on heterogeneous clinical cohorts is limited, leaving generalization and calibration under real-world conditions as important next steps.

Study by [27] propose a hybrid MRI brain tumor classification pipeline that couples deep feature extraction with traditional machine-learning classifiers to enhance discriminative power and interpretability. Using T1-weighted contrast-enhanced MRI from over two hundred patients, the framework aggregates outputs from multiple CNNs and evaluates downstream classifiers such as SVM and tree-based models, showing improvements over single end-to-end softmax heads. The contribution lies in demonstrating that post-CNN ML heads can exploit representation complementarities and yield robust class separation in multiclass settings. The drawback is increased system complexity with multi-stage training, potential feature-classifier mismatch across scanners, and limited exploration of calibration and shift robustness. While results are promising, the study would benefit from prospective external validation, harmonization strategies for scanner variability, and reporting of decision curves to connect accuracy gains with clinical utility. This hybrid perspective aligns well with pipelines that segment first and then classify using flexible ML heads.

In contrast to prior studies that predominantly emphasize either pure deep learning pipelines or lightweight segmentation variants, the novelty of our research lies in proposing a hybrid deep-machine learning framework that integrates U-Net-based segmentation with downstream machine learning classifiers for multi-class brain tumor analysis in MRI. Unlike BiTr-UNet or attention-augmented U-Net models that prioritize segmentation performance but remain computationally intensive or dataset-specific, our approach first leverages U-Net to accurately localize and delineate tumor regions, thereby ensuring reliable region-of-interest extraction across heterogeneous tumor morphologies. These segmented regions are subsequently transformed into discriminative feature representations, which are then evaluated using multiple machine learning algorithms such as SVM, Random Forest, and XGBoost to identify the most robust classifier. This dual-stage paradigm highlights that traditional machine learning remains highly relevant when combined with deep learning-derived features, offering both interpretability and adaptability. Furthermore, our framework aims to bridge the gap between segmentation and classification in a single workflow, enabling not only tumor localization but also precise subtype categorization, thereby addressing clinical needs for both spatial and categorical diagnostic accuracy.

2. METHOD

This section explains the methodological framework employed in this study, which integrates machine learning-based classification and U-Net segmentation for brain tumor analysis. The workflow is structured to cover all essential steps, starting from data preparation to model training and testing, as well as tumor segmentation.

2.1. Proposed Workflow

Based on Figure 1, the proposed method is divided into three main stages, namely data preparation (blue), training phase (red), and testing phase (green). In the data preparation stage, raw MRI data undergo pre-processing such as normalization and resizing, followed by data augmentation to enrich variability and reduce overfitting. The dataset is then split into 80% training data and 20% testing data to ensure fair model evaluation. In the training phase, the 80% portion of data is used to train several machine learning classifiers, and the best trained model is selected based on accuracy, precision, recall,

and F1-score. In the testing phase, the remaining 20% of data is used to evaluate the model, while U-Net segmentation is applied to identify and extract the tumor region from MRI images. The final output consists of a classified tumor type along with its segmented volume, which can be further utilized for medical diagnosis and treatment planning.

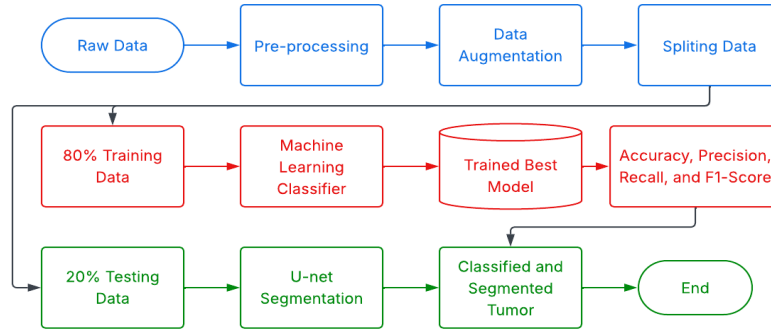


Figure 1. Proposed workflow

2.2. Data Collection and Preparation

The dataset used in this study was obtained from a publicly available brain MRI dataset on Kaggle [28]. It consists of four classes: Glioma, Meningioma, Pituitary, and No Tumor. Each class was divided into training (80%) and testing (20%) sets to ensure a balanced evaluation. Specifically, the training distribution is as follows: 826 images of Glioma, 822 images of Meningioma, 827 images of Pituitary, and 395 images of No Tumor, with the remaining images of each class allocated for testing. Prior to model training, all images were pre-processed by resizing them into a uniform dimension of $224 \times 224 \times 3$, ensuring consistency across the dataset [29], [30]. The sample dataset can be seen in Figure 2.

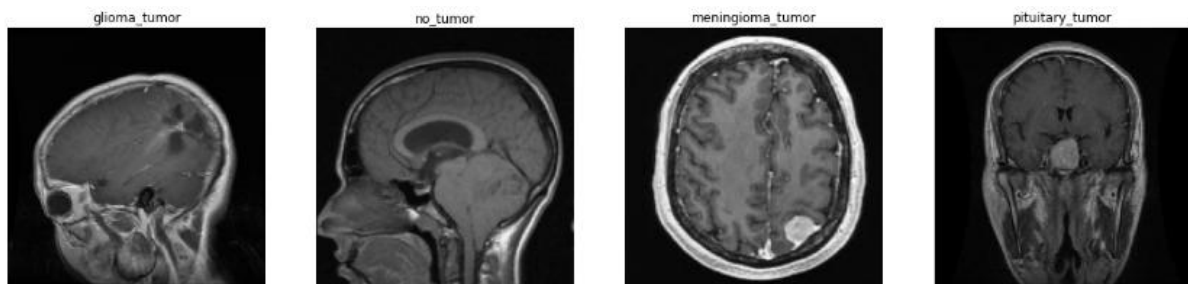


Figure 2. Sample dataset each class

To enhance model generalization and reduce overfitting, several data augmentation techniques were applied, including rotation, flipping, zooming, shifting, and brightness adjustment [31], [32], [33]. The augmentation operations can be calculated using eq (1) – (5).

$$Rotation = R\theta(I) \quad (1)$$

Where I is the image pixel value, $R\theta$ is the rotation function with angle $\theta \in [-15^\circ, 15^\circ]$. This operation rotates the image by a random angle within the specified range. Rotation is particularly useful to simulate variations in patient head orientation during MRI scanning.

$$Horizontal\ Flip(x, y) = I(W - x, y) \quad (2)$$

Where W is the image width. Flipping generates a mirror version of the image, which helps the model recognize tumor features regardless of left-right brain orientation.

$$Zoom = Z\alpha(I), \alpha \in [0.9, 1.1] \quad (3)$$

Zooming applies a random scaling to the image, either enlarging or shrinking it slightly. This simulates differences in tumor size and scanning distance.

$$WH Shift(x, y) = I(x + \Delta x, y + \Delta y), \Delta x, \Delta y \in [-0.1W, 0.1H] \quad (4)$$

Width and Height Shift (*WH Shift*) operation translates the image along the horizontal or vertical axis. It makes the model robust against small variations in patient positioning within the MRI frame.

$$Brithness Adjustment = (x, y) = \beta \cdot I(x, y), \beta \in [0.8, 1.2] \quad (5)$$

Brightness is randomly increased or decreased to mimic differences in scanner intensity or contrast, improving the model's ability to handle images from diverse MRI machines. By combining these augmentation strategies, the dataset variability is significantly enriched, thereby enabling the classifier and segmentation model to achieve better robustness and generalization when exposed to unseen data.

2.3. Machine Learning Classifier

In this stage, multiple machine learning classifiers were employed to automatically categorize brain MRI scans into four classes: Glioma, Meningioma, Pituitary, and No Tumor. The selected models include Logistic Regression (LR), Support Vector Classifier (SVC), K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). Each classifier was trained using the pre-processed and augmented dataset, with hyperparameters optimized to balance accuracy and generalization [34], [35]. The evaluation of these models was performed based on accuracy, precision, recall, F1-score, and AUC values to provide a comprehensive performance comparison. Among the tested approaches, XGBoost consistently outperformed the others, demonstrating its capability to capture complex non-linear patterns and feature interactions within the MRI data, making it the most effective classifier in this study.

2.4. U-net Segmentation

U-Net segmentation was employed to localize and extract the tumor region from MRI images, providing complementary spatial information to the classification stage [36]. The U-Net architecture follows an encoder-decoder design with skip connections, enabling the network to capture both high-level contextual features and fine-grained spatial details [36], [37]. In the encoder path, convolution and pooling operations progressively reduce image dimensions while learning discriminative features, whereas the decoder path performs up-sampling to reconstruct the segmentation mask with precise tumor boundaries [37]. Skip connections ensure that spatial information lost during down-sampling is preserved, allowing accurate delineation of tumor structures. In this study, U-Net was trained on the augmented dataset to segment tumors from the input MRI slices, and its output was integrated with the classification results to provide both the tumor type and its segmented volume. This dual-stage framework enhances the interpretability of predictions and supports medical practitioners in diagnosis and treatment planning by combining class identification with visual tumor localization.

Based on Figure 3, the U-Net architecture is composed of two main symmetrical paths, namely the encoder (contracting path) and decoder (expanding path), connected by a bottleneck layer. In the encoder path, the input image of size 224×224 is first processed through convolutional layers with 32 filters, followed by feature extraction at deeper levels with 64, 128, 256, and 512 filters. At each stage, the convolution is followed by a rectified linear unit (ReLU) activation and a max-pooling operation (red arrows), effectively halving the spatial resolution while doubling the number of feature channels. This hierarchical downsampling captures fine-grained details at shallow layers and high-level abstract features at deeper layers. The bottleneck layer employs 1024 filters at the smallest spatial resolution, acting as the feature representation core that integrates all learned patterns from the encoder.

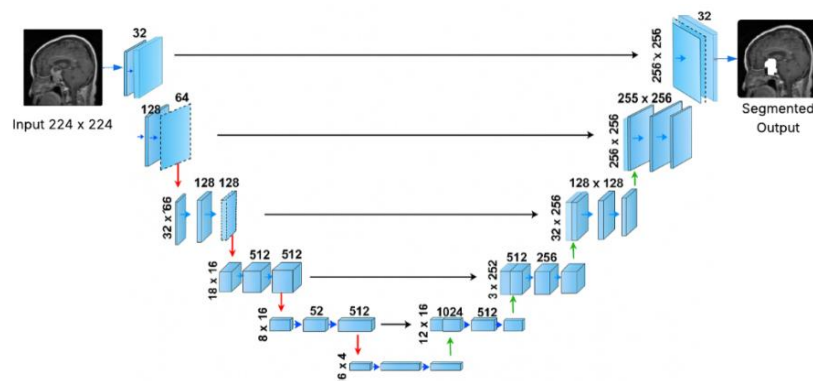


Figure 2. U-net layers

In the decoder path, the architecture reverses the process using transposed convolutions (green arrows) for upsampling, progressively restoring the spatial resolution. At each upsampling stage, the corresponding feature maps from the encoder are concatenated through skip connections, ensuring that spatial context from shallow layers is preserved and combined with the semantic richness of deeper layers. Specifically, feature maps transition from $512 \rightarrow 256 \rightarrow 128 \rightarrow 64 \rightarrow 32$ filters as the spatial resolution increases back to 256×256 , leading to precise localization of tumor regions. The final convolutional layer outputs the segmented tumor mask, where pixel-level classification distinguishes tumor tissue from non-tumor tissue.

3. RESULT

All experiments were conducted on a high-performance computing environment to ensure the reproducibility and reliability of results. The primary implementation was executed using Python-based frameworks on Google Colab, enabling access to optimized deep learning libraries and GPU acceleration. Locally, the system was also benchmarked on a workstation equipped with an Intel Core Ultra 5 processor, an NVIDIA RTX 5070 Ti GPU, and 32 GB of RAM, providing sufficient computational power for training U-Net models on multi-modal MRI data and for running multiple machine learning classifiers in parallel.

3.1. Experimental Results of Machine Learning Models

In this subsection, the experimental results of the trained machine learning models are presented and analyzed based on their classification performance. The evaluation primarily relies on the Receiver Operating Characteristic (ROC) curve and the corresponding Area Under the Curve (AUC) scores, which provide insights into the trade-off between sensitivity and specificity. These results highlight the ability of each classifier to discriminate between tumor and non-tumor regions effectively.

To ensure fair comparison and optimal performance of each machine learning model, hyperparameter tuning was conducted using a grid search strategy combined with 80% of training data and 20% validation. Logistic Regression was optimized using the lbfgs solver with L2 regularization and a regularization strength parameter $C=1$, while SVC employed the RBF kernel with $C=10$ and automatic scaling of the gamma parameter to handle non-linear class boundaries. For K-Nearest Neighbors, the number of neighbors k was varied from 3 to 11, with the best performance achieved at $k=7$ using distance-based weighting metrics. Decision Tree and Random Forest models were fine-tuned by exploring different tree depths and splitting criteria, where Random Forest with 300 trees, sqrt feature selection, and no depth restriction yielded more stable generalization compared to a single Decision Tree.

In contrast, XGBoost required more extensive tuning due to its large number of controllable hyperparameters. A grid search was performed over learning rates $\{0.01, 0.05, 0.1\}$, number of estimators $\{100, 200, 300\}$, and maximum tree depths $\{4, 6, 8\}$, combined with regularization parameters such as $\text{subsample} = 0.8$, $\text{colsample_bytree} = 0.8$, and $\text{reg_lambda} = 1$ to prevent overfitting. The final configuration used a learning rate of 0.1, 300 estimators, and a maximum depth of 6, which provided the best trade-off between bias and variance. Additionally, the multi:softprob objective was used to handle the four-class prediction task, and early stopping with 20 rounds on a 20% validation split was applied to stabilize training and avoid unnecessary iterations.

Through this systematic hyperparameter optimization process, XGBoost consistently outperformed all other models across accuracy, precision, recall, and F1-score, confirming its superior ability to capture complex feature interactions in MRI data. Ensemble-based methods such as XGBoost and Random Forest demonstrated stronger robustness compared to single learners like Decision Tree or KNN, further supporting the effectiveness of advanced boosting techniques in medical image classification tasks.

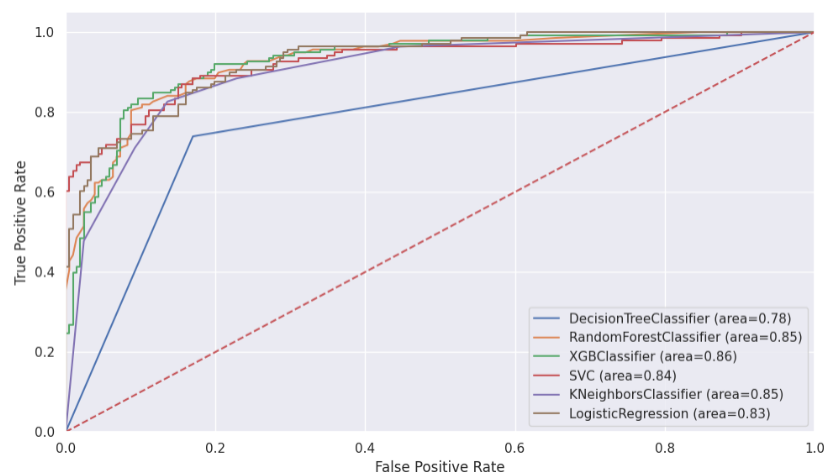


Figure 4. ROC curves and AUC comparison of the machine learning models

As seen in Figure 4, the ROC curves demonstrate that the ensemble-based models achieve superior performance compared to single learners. In particular, XGBoost Classifier attains the highest AUC score of 0.86, closely followed by Random Forest and K-Nearest Neighbors, each with an AUC of 0.85. SVC and Logistic Regression also show competitive outcomes with AUC values of 0.84 and 0.83, respectively, whereas the Decision Tree model lags behind with an AUC of 0.78. As the ROC curves provide an overview of the discriminative ability of each model, it is also essential to examine their performance across multiple evaluation metrics to obtain a more comprehensive assessment.

Therefore, Figure 5 presents a comparative analysis of accuracy, precision, recall, and F1-score for all classifiers, offering deeper insights into their strengths and weaknesses beyond AUC values.

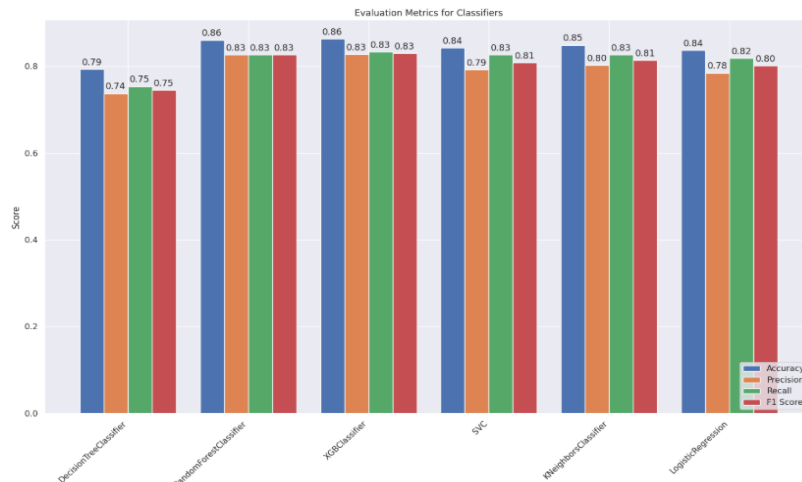


Figure 5. Comparative evaluation of confusion matrix assessment

Based on Figure 5, it can be observed that XGBoost and Random Forest consistently achieve the highest values across accuracy, precision, recall, and F1-score, with XGBoost slightly outperforming others in terms of balanced performance. The strong generalization capability of these ensemble-based models allows them to effectively capture complex decision boundaries while minimizing overfitting. In contrast, single learners such as Decision Tree exhibit lower consistency across metrics, which confirms their limited robustness when compared to ensemble methods.

To further validate these findings, Figure 6 illustrates the confusion matrices of all classifiers, highlighting how each model performs in terms of true and false classifications. The results confirm that XGBoost yields the most balanced classification outcomes, correctly identifying the majority of positive and negative samples with minimal misclassification, thereby justifying its position as the best-performing model in this study.

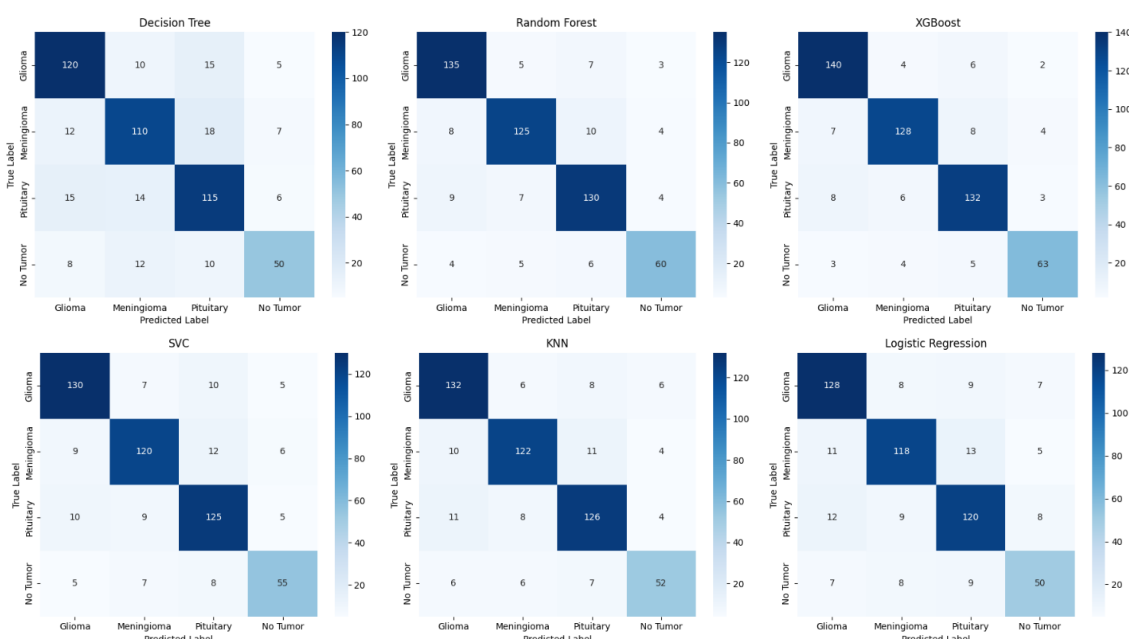


Figure 6. Model prediction based on table of confusion matrix

3.2. Experimental Result of U-net Segmentation

To further evaluate the system's ability in tumor localization, this subsection presents the experimental results of the U-Net segmentation model applied to brain MRI images. Unlike the previous subsection that focused on classification performance, this experiment emphasizes pixel-wise segmentation to accurately delineate tumor regions. The U-Net model is specifically designed for biomedical image segmentation, where its encoder-decoder structure enables precise boundary detection even with limited training samples. As seen in Figure 7, the segmentation results are demonstrated for four representative cases corresponding to each category: No Tumor, Glioma, Pituitary, and Meningioma. The first column depicts the input testing images, the second column presents the ground-truth tumor masks, and the third column shows the tumor regions predicted by U-Net.

The figure highlights that the U-Net successfully differentiates between tumor and non-tumor regions with high fidelity. For the "No Tumor" case, the model correctly outputs an empty mask, while for tumor cases (Glioma, Pituitary, and Meningioma), the predicted masks closely align with the ground-truth annotations in terms of shape and position. These results confirm the robustness of U-Net in segmenting heterogeneous tumor structures across different types, making it a reliable tool for automated tumor delineation.

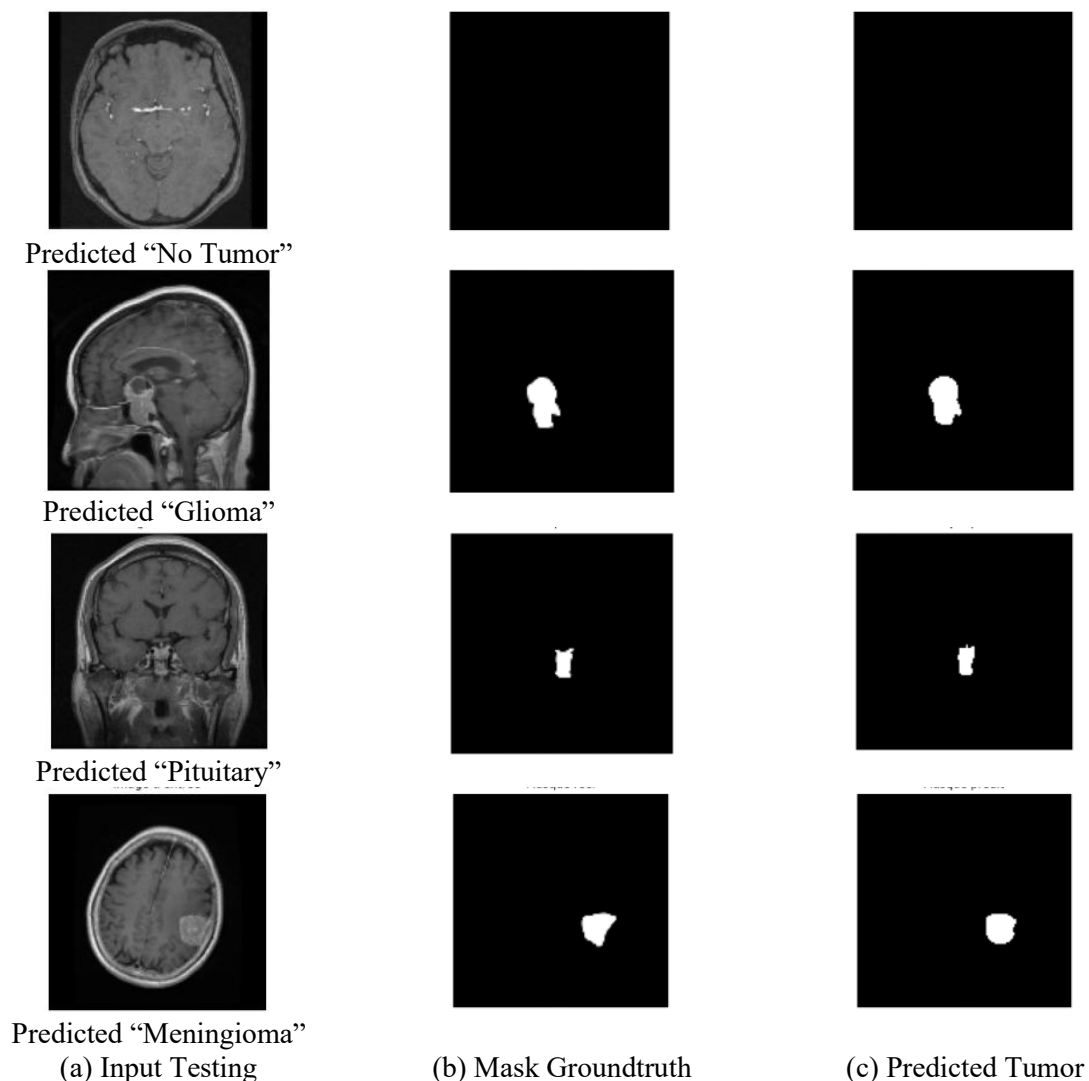


Figure 7. Experiment results of U-net segmentation

4. DISCUSSIONS

Several recent studies have explored brain tumor classification and segmentation using deep learning approaches. Study [24] introduced BiTr-UNet, a CNN–Transformer hybrid that enhances multi-modal segmentation on BraTS 2021. While this design successfully integrates long-range dependency modeling through Transformer blocks, its computational overhead and limited assessment on external datasets raise concerns for practical deployment. Similarly, [25] proposed a lightweight U-Net with spatial attention, showing efficiency and improved Dice scores but being constrained by evaluation on limited datasets and potential underfitting in complex tumor textures. Both [24] and [25] focused solely on segmentation tasks.

Table 1. Comparison of related studies with our proposed method

Study	Focus	Methodology	Strengths	Limitations and Drawback	Distinction from Our Work
[24]	Segmentation	CNN–Transformer hybrid	Strong Dice, captures global & local features	High memory cost, limited clinical validation	Our work is lighter, more practical, and adds classification beyond segmentation
[25]	Segmentation	U-Net with spatial attention	Efficient, precise boundary detection	Evaluated on limited datasets, risk of underfitting	Our framework balances segmentation with classification, offering wider applicability
[26]	Classification	Transfer learning with CNN backbone	High accuracy, robust on 2D MRI slices	Ignores volumetric context, lacks segmentation	Our study extends beyond slice-based classification by also providing tumor localization
[27]	Classification	CNN feature extraction + ML classifiers	Robust class separation, interpretable	Multi-stage complexity, limited generalization	Our approach is end-to-end with integrated segmentation and classification
Our	Classification + Segmentation	ML-based classifier + U-Net	Joint tumor type recognition & localization, clinically oriented, efficient	Focused on single-modality MRI (extension to multi-modal future work)	-

On the classification side, [26] applied a VGG16-based transfer learning model on a large dataset for four-class tumor recognition, reporting strong accuracy but relying only on 2D slices without volumetric or multi-modal information. Meanwhile, [27] explored hybrid pipelines that combine CNN feature extraction with traditional machine learning classifiers, which improved discriminative power but at the cost of system complexity and limited validation across diverse clinical cohorts. Compared with these works, our study contributes a dual-stage framework that integrates both tumor classification and U-Net segmentation in a unified pipeline. This approach not only classifies tumors into four categories (Glioma, Meningioma, Pituitary, and No Tumor) with high accuracy but also performs

precise segmentation to localize tumor regions. Unlike [24] and [25], our method explicitly demonstrates generalization across both recognition and localization tasks. Furthermore, in contrast to [26] and [27], which focus only on classification, our study provides a holistic pipeline combining diagnostic classification and region delineation, thereby bridging the gap between automated diagnosis and clinical decision support.

In summary, while prior studies either concentrated on segmentation with complex architectures [24] and [25] or pursued classification with limited integration of localization [26] and [27], our study introduces a comprehensive framework that simultaneously addresses both classification and segmentation of brain tumors. By combining a CNN-based classifier with U-Net segmentation, we provide a clinically relevant pipeline that not only identifies tumor types but also localizes the affected regions. This dual capability distinguishes our work from existing approaches, offering a more practical tool for supporting radiologists in diagnostic workflows. Furthermore, the balance between accuracy, computational efficiency, and task integration highlights the novelty and applicability of our framework for real-world medical imaging scenarios.

5. CONCLUSION

This study aimed to develop an effective and reliable framework for brain tumor analysis by combining machine learning–based classification and U-Net segmentation, enabling both tumor type recognition and spatial localization. The experimental findings in the Results and Discussion confirmed these expectations, where XGBoost achieved the best classification accuracy of 86% and U-Net demonstrated consistent segmentation performance across glioma, meningioma, pituitary, and no-tumor cases. This dual approach represents the novelty of the proposed framework, outperforming several conventional single-task methods and offering a more comprehensive solution for clinical decision support. For future research, the framework can be extended by integrating multi-modal MRI inputs, enhancing 3D volumetric segmentation, and performing large-scale validation on diverse clinical cohorts. Moreover, application prospects include embedding the method into computer-aided diagnosis systems, assisting radiologists in early detection and personalized treatment planning, while future studies should also address robustness, generalization, and deployment in real-world healthcare environments.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

REFERENCES

- [1] G. Calabrese *et al.*, “Carbon Dots: An Innovative Tool for Drug Delivery in Brain Tumors,” *Int J Mol Sci*, vol. 22, no. 21, p. 11783, Oct. 2021, doi: 10.3390/ijms222111783.
- [2] Y. Xie *et al.*, “Convolutional Neural Network Techniques for Brain Tumor Classification (from 2015 to 2022): Review, Challenges, and Future Perspectives,” *Diagnostics*, vol. 12, no. 8, p. 1850, Jul. 2022, doi: 10.3390/diagnostics12081850.
- [3] A. I. R. Maas *et al.*, “Traumatic brain injury: progress and challenges in prevention, clinical care, and research,” *Lancet Neurol*, vol. 21, no. 11, pp. 1004–1060, Nov. 2022, doi: 10.1016/S1474-4422(22)00309-X.
- [4] F. Mo, A. Pellerino, R. Soffietti, and R. Rudà, “Blood-brain barrier in brain tumors: Biology and clinical relevance,” Dec. 01, 2021, *MDPI*. doi: 10.3390/ijms222312654.
- [5] M. Adel Fahmideh and M. E. Scheurer, “Pediatric Brain Tumors: Descriptive Epidemiology, Risk Factors, and Future Directions,” *Cancer Epidemiology, Biomarkers & Prevention*, vol. 30, no. 5, pp. 813–821, May 2021, doi: 10.1158/1055-9965.EPI-20-1443.

- [6] D. Bartusik-Aebisher *et al.*, “The Use of Photodynamic Therapy in the Treatment of Brain Tumors—A Review of the Literature,” Oct. 01, 2022, *MDPI*. doi: 10.3390/molecules27206847.
- [7] M. J. Lakshmi and S. Nagaraja Rao, “Brain tumor magnetic resonance image classification: a deep learning approach,” *Soft comput*, vol. 26, no. 13, pp. 6245–6253, Jul. 2022, doi: 10.1007/s00500-022-07163-z.
- [8] N. R. D. Cahyo, C. A. Sari, E. H. Rachmawanto, C. Jatmoko, R. R. A. Al-Jawry, and M. A. Alkhafaji, “A Comparison of Multi Class Support Vector Machine vs Deep Convolutional Neural Network for Brain Tumor Classification,” in *2023 International Seminar on Application for Technology of Information and Communication (iSemantic)*, IEEE, Sep. 2023, pp. 358–363. doi: 10.1109/iSemantic59612.2023.10295336.
- [9] Y. Anagun, “Smart brain tumor diagnosis system utilizing deep convolutional neural networks,” *Multimed Tools Appl*, 2023, doi: 10.1007/s11042-023-15422-w.
- [10] S. Shanthi, S. Saradha, J. A. Smitha, N. Prasath, and H. Anandakumar, “An efficient automatic brain tumor classification using optimized hybrid deep neural network,” *International Journal of Intelligent Networks*, vol. 3, pp. 188–196, Jan. 2022, doi: 10.1016/j.ijin.2022.11.003.
- [11] B. Arjmand *et al.*, “Machine Learning: A New Prospect in Multi-Omics Data Analysis of Cancer,” Jan. 27, 2022, *Frontiers Media S.A.* doi: 10.3389/fgene.2022.824451.
- [12] A. A. Khan, A. A. Laghari, and S. A. Awan, “Machine Learning in Computer Vision: A Review,” *EAI Endorsed Transactions on Scalable Information Systems*, vol. 8, no. 32, pp. 1–11, 2021, doi: 10.4108/eai.21-4-2021.169418.
- [13] J. Wen, Z. Zhang, Y. Lan, Z. Cui, J. Cai, and W. Zhang, “A survey on federated learning: challenges and applications,” *International Journal of Machine Learning and Cybernetics*, vol. 14, no. 2, pp. 513–535, Feb. 2023, doi: 10.1007/s13042-022-01647-y.
- [14] M. Tsuneki, “Deep learning models in medical image analysis,” *J Oral Biosci*, vol. 64, no. 3, pp. 312–320, Sep. 2022, doi: 10.1016/j.job.2022.03.003.
- [15] J. Gupta, S. Pathak, and G. Kumar, “Deep Learning (CNN) and Transfer Learning: A Review,” in *Journal of Physics: Conference Series*, Institute of Physics, 2022. doi: 10.1088/1742-6596/2273/1/012029.
- [16] Y. Wang *et al.*, “ID-Seg: an infant deep learning-based segmentation framework to improve limbic structure estimates,” *Brain Inform*, vol. 9, no. 1, p. 12, Dec. 2022, doi: 10.1186/s40708-022-00161-9.
- [17] R. Azad *et al.*, “Medical Image Segmentation Review: The Success of U-Net,” *IEEE Trans Pattern Anal Mach Intell*, vol. 46, no. 12, pp. 10076–10095, Dec. 2024, doi: 10.1109/TPAMI.2024.3435571.
- [18] C. Zhang, X. Deng, and S. H. Ling, “Next-Gen Medical Imaging: U-Net Evolution and the Rise of Transformers,” *Sensors*, vol. 24, no. 14, p. 4668, Jul. 2024, doi: 10.3390/s24144668.
- [19] Kavitha Soppari, Sravya Putnala, Sai Srujana Borala, and Sai Kumar Gondhi, “A survey on brain MRI segmentation,” *World Journal of Advanced Research and Reviews*, vol. 21, no. 3, pp. 1702–1710, Mar. 2024, doi: 10.30574/wjarr.2024.21.3.0813.
- [20] T. Trongtirakul, S. Agaian, and A. Oulefki, “Automated tumor segmentation in thermographic breast images,” *Mathematical Biosciences and Engineering*, vol. 20, no. 9, pp. 16786–16806, 2023, doi: 10.3934/mbe.2023748.
- [21] Y. Shin, M. Lee, Y. Lee, K. Kim, and T. Kim, “Artificial Intelligence-Powered Quality Assurance: Transforming Diagnostics, Surgery, and Patient Care—Innovations, Limitations, and Future Directions,” *Life*, vol. 15, no. 4, p. 654, Apr. 2025, doi: 10.3390/life15040654.
- [22] R. R. Ali *et al.*, “Learning Architecture for Brain Tumor Classification Based on Deep Convolutional Neural Network: Classic and ResNet50,” *Diagnostics*, vol. 15, no. 5, p. 624, Mar. 2025, doi: 10.3390/diagnostics15050624.
- [23] H. P. Hadi, E. Faisal, and E. H. Rachmawanto, “Brain tumor segmentation using multi-level Otsu thresholding and Chan-Vese active contour model,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, no. 4, p. 825, Aug. 2022, doi: 10.12928/telkomnika.v20i4.21679.
- [24] Q. Jia and H. Shu, “BiTr-Unet: A CNN-Transformer Combined Network for MRI Brain Tumor Segmentation,” 2022, pp. 3–14. doi: 10.1007/978-3-031-09002-8_1.

-
- [25] K. Avazov, S. Mirzakhilov, S. Umirzakova, A. Abdusalomov, and Y. I. Cho, "Dynamic Focus on Tumor Boundaries: A Lightweight U-Net for MRI Brain Tumor Segmentation," *Bioengineering*, vol. 11, no. 12, p. 1302, Dec. 2024, doi: 10.3390/bioengineering11121302.
- [26] Y. Wong, E. L. M. Su, C. F. Yeong, W. Holderbaum, and C. Yang, "Brain tumor classification using MRI images and deep learning techniques," *PLoS One*, vol. 20, no. 5, p. e0322624, May 2025, doi: 10.1371/journal.pone.0322624.
- [27] S. E. Nassar, I. Yasser, H. M. Amer, and M. A. Mohamed, "A robust MRI-based brain tumor classification via a hybrid deep learning technique," *J Supercomput*, vol. 80, no. 2, pp. 2403–2427, Jan. 2024, doi: 10.1007/s11227-023-05549-w.
- [28] Sartaj, "Brain Tumor Classification (MRI)." Accessed: Aug. 13, 2024. [Online]. Available: <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri/data>
- [29] I. P. Kamila, C. A. Sari, E. H. Rachmawanto, and N. R. D. Cahyo, "A Good Evaluation Based on Confusion Matrix for Lung Diseases Classification using Convolutional Neural Networks," *Advance Sustainable Science, Engineering and Technology*, vol. 6, no. 1, p. 0240102, Dec. 2023, doi: 10.26877/asset.v6i1.17330.
- [30] M. M. I. Al-Ghiffary, N. R. D. Cahyo, E. H. Rachmawanto, C. Irawan, and N. Hendriyanto, "Adaptive deep learning based on FaceNet convolutional neural network for facial expression recognition," *Journal of Soft Computing*, vol. 05, no. 03, pp. 271–280, 2024, doi: <https://doi.org/10.52465/joscex.v5i3.450>.
- [31] K. Kavin Kumar *et al.*, "Brain Tumor Identification Using Data Augmentation and Transfer Learning Approach," *Computer Systems Science and Engineering*, vol. 46, no. 2, pp. 1845–1861, 2023, doi: 10.32604/csse.2023.033927.
- [32] F. J. Moreno-Barea, J. M. Jerez, and L. Franco, "Improving classification accuracy using data augmentation on small data sets," *Expert Syst Appl*, vol. 161, p. 113696, 2020.
- [33] N. P. Sutramiani, N. Suciati, and D. Siahaan, "MAT-AGCA: Multi Augmentation Technique on small dataset for Balinese character recognition using Convolutional Neural Network," *ICT Express*, vol. 7, no. 4, pp. 521–529, Dec. 2021, doi: 10.1016/j.icte.2021.04.005.
- [34] D. Sheth and M. Shah, "Predicting stock market using machine learning: best and accurate way to know future stock prices," *International Journal of System Assurance Engineering and Management*, vol. 14, no. 1, pp. 1–18, Feb. 2023, doi: 10.1007/s13198-022-01811-1.
- [35] O. A. Montesinos López, A. Montesinos López, and J. Crossa, *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. Springer International Publishing, 2022. doi: 10.1007/978-3-030-89010-0.
- [36] P. Agrawal, N. Katal, and N. Hooda, "Segmentation and classification of brain tumor using 3D-UNet deep neural networks," *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 199–210, Jun. 2022, doi: 10.1016/j.ijcce.2022.11.001.
- [37] U. Latif, A. R. Shahid, B. Raza, S. Ziauddin, and M. A. Khan, "An end-to-end brain tumor segmentation system using multi-inception-UNET," *Int J Imaging Syst Technol*, vol. 31, no. 4, pp. 1803–1816, Dec. 2021, doi: 10.1002/ima.22585.
-