ILKOM Jurnal Ilmiah Vol. 17, No. 2, August 2025, pp. 86-97
Accredited 2nd by KEMDIKTISAINTEK No. 10/C/C3/DT.05.00/2025; E-ISSN 2548-7779 | P-ISSN 2087-1716

Performance Comparison of Ensemble Learning Models for Brain Tumor Detection on Augmented MRI Datasets

Gilbert Valentino Titaley[a,1]; Nurul Rismayanti[a,2]; Anik Nur Handayani[a,2,*]; Jevri Tri Ardiansah[a,4]
[a] *Universitas Negeri Malang, Jl. Semarang Gading Kasri, Kecamatan Klojen, Kota Malang, 65115, Indonesia*
[1] *gilberth.valentino.2405348@students.um.ac.id;* [2] *nurul.rismayanti.2305348@students.um.ac.id;* [3] *aniknur.ft@um.ac.id;*
[4] *jevri.ardiansah.ft@um.ac.id*
** Corresponding author*

Article history: Received May 09, 2024; Revised July 08, 2024; Accepted August 14, 2025; Available online August 19, 2025

Abstract

Brain tumors are highly fatal diseases, making early detection a critical factor in improving patient survival rates. Magnetic Resonance Imaging (MRI) has become a primary tool in brain tumor diagnosis; however, manual analysis processes are often time-consuming and prone to subjective errors. This study employs a machine learning-based classification model to detect four categories of brain tumors—glioma, meningioma, pituitary, and healthy—with high accuracy. The methods include image segmentation using the U-Net model, which excels in medical image analysis due to its encoder-decoder architecture with skip connections, allowing efficient integration of spatial and contextual information. Features are extracted using HuMoments, known for their invariance to rotation, translation, and scale, ensuring robust spatial pattern representation. Data normalization is conducted using Robust Scaling and L2 Normalization to address outliers and harmonize feature scales, enhancing model performance. The MRI dataset, originally comprising 7,023 images, was augmented to 8,000 images using techniques such as rotation, flipping, and contrast adjustments to improve class balance and minimize overfitting. Three ensemble algorithms—Random Forest, XGBoost, and Stacking—were employed to train the models, with performance evaluation based on accuracy, ROC-AUC, F1-score, and confusion matrix. The results demonstrate that Random Forest achieved the best performance with an accuracy of 72% and an ROC-AUC of 0.91. This study illustrates the potential of machine learning approaches for automated brain tumor diagnosis, with further improvement possible through model optimization and the use of more diverse datasets.

Keywords: Brain Tumor; Classification Model; Ensemble Algorithms; Machine Learning; Magnetic Resonance Imaging.

Introduction

Brain tumors are among the diseases with a high fatality rate, making early detection a critical aspect in improving patient survival rates. Magnetic Resonance Imaging (MRI) has long been used as a medical imaging technique capable of providing high-resolution structural details of the brain, making it a primary tool in brain tumor diagnosis [1], [2]. However, manual analysis of MRI results requires a high level of expertise, is time-consuming, and is often prone to subjective errors. These limitations drive the need for more accurate and efficient automation solutions, particularly through machine learning approaches. In recent years, machine learning methods have demonstrated rapid progress in the automatic diagnosis of brain tumors, offering greater speed and consistency compared to manual analysis [3]–[5]. Nevertheless, key challenges remain in enhancing classification accuracy and the ability of models to handle the diversity of complex imaging data. Factors such as image segmentation quality, feature representation, and the selection of suitable algorithms play a crucial role in determining the success of these approaches.

This study proposes a machine learning-based approach leveraging ensemble algorithms such as Random Forest, XGBoost, and Stacking [6]–[9]. The U-Net model is frequently employed for MRI image segmentation because of its encoder-decoder architecture with skip connections, which integrates high-resolution spatial details from the encoder with contextual information from the decoder. This design enhances the accuracy of segmentation tasks and has proven effective in various medical imaging applications, including organ and tumor segmentation [10], [11]. Feature extraction is performed using HuMoments, which are invariant to rotation, translation, and scale, enabling more effective representation of spatial patterns in the images [12], [13]. This method is particularly advantageous in this research due to its robustness in capturing the shape and geometric features of brain tumor regions, which often exhibit variations in orientation and size across different MRI samples. By providing a compact yet informative numerical representation, HuMoments ensure that the extracted features are consistent and reliable for classification tasks, even

when dealing with diverse imaging conditions. Additionally, data normalization is conducted using Robust Scaling and L2 Normalization to ensure uniform data scaling [14], which is essential for machine learning and deep learning algorithms. Robust Scaling is employed to mitigate the influence of outliers commonly found in MRI image features, ensuring that the model focuses on meaningful patterns rather than noise. Meanwhile, L2 Normalization is used to harmonize feature magnitudes across the dataset, facilitating better convergence during training and improving the stability of machine learning algorithms.

This research builds upon various prior studies that have made significant contributions to the field of medical data analysis. For instance research, [15] employed the Naive Bayes algorithm to detect brain tumors from MRI images, although the achieved accuracy was limited to approximately 60%. Research [16] developed an approach that combined segmentation techniques and texture-based feature extraction, such as the Gray-Level Co-occurrence Matrix (GLCM), with algorithms like KNN, FCM, and K-Means, which proved more efficient for medical image segmentation tasks. Meanwhile study, [17] compared the performance of various machine learning algorithms, such as Random Forest, XGBoost, and Logistic Regression, for cardiovascular disease classification. Their research highlighted the superiority of Random Forest in handling high variability in data, making it a reliable choice for similar applications. This study aims to address the primary challenges in machine learning-based brain tumor diagnosis by integrating U-Net-based segmentation techniques, optimal feature representation through HuMoments, and ensemble algorithms to improve classification accuracy. Unlike prior studies, such as [15], which relied on simpler algorithms like Naive Bayes for brain tumor detection and achieved limited accuracy (approximately 60%), this research combines U-Net for precise segmentation and HuMoments for robust feature extraction. The ensemble learning methods, including Random Forest and XGBoost, enhance the classification process, addressing limitations in handling complex patterns and diverse datasets. Additionally, previous studies struggled to manage larger and more diverse datasets, as well as overlapping features, which often led to unreliable classifications [15]. By incorporating advanced augmentation techniques to balance the dataset and ensemble models for better generalization, this study bridges these gaps, achieving higher accuracy and robustness for practical clinical applications. The main contribution of this research lies in the development of a more reliable and efficient approach to brain tumor diagnosis, which can be broadly applied in technology-driven medical analysis systems.

Method

A. Research Design

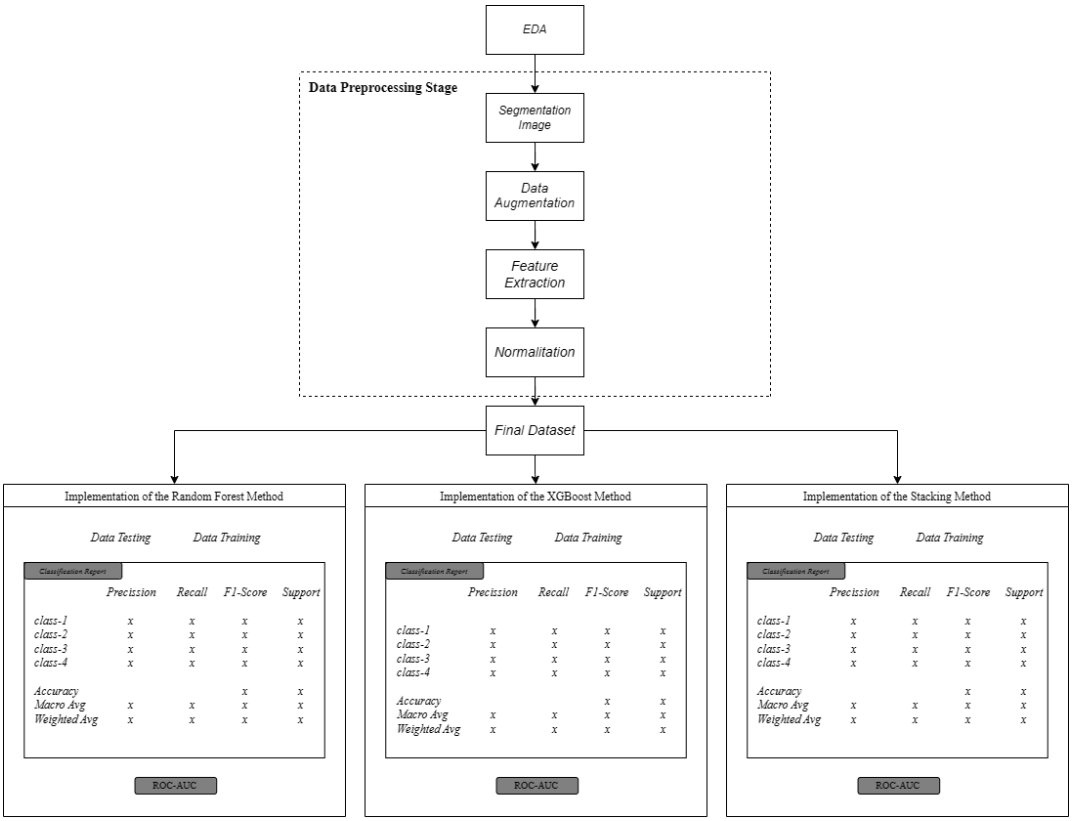

Figure 1. Research Design

This study employs a machine learning approach to classify brain MRI images into four categories: glioma, healthy, meningioma, and pituitary. The research stages include image segmentation, feature extraction, data normalization, model training, and performance evaluation. These processes are designed to ensure accurate results by utilizing ensemble algorithms such as Random Forest, XGBoost, and Stacking. The U-Net segmentation model is used to isolate relevant brain regions from the MRI image background [11], [18]. The segmentation results are then utilized for further analysis using machine learning techniques.

B. Data Collection

1. Dataset

The dataset used in this study was obtained from the public platform Kaggle and consists of 7,023 brain MRI images divided into four classes: glioma (1,621 images), healthy (2,000 images), meningioma (1,645 images), and pituitary (1,757 images). The images in this dataset have consistent resolution, facilitating the segmentation and analysis processes. With a relatively balanced distribution across classes, the dataset enables machine learning and deep learning models to learn unique patterns for each class without significant bias risks.

The selection of this dataset is based on the variety of tumor types available (glioma, meningioma, pituitary) as well as the normal class (healthy), providing a realistic representation for clinical analysis. Additionally, the dataset supports research reproducibility as it is publicly accessible through Kaggle. To provide a visual overview of the dataset, examples of images from each class glioma, healthy, meningioma, and pituitary are presented. These visualizations offer initial insights into the unique visual patterns in each class and assist in validating both the segmentation process and further analysis.

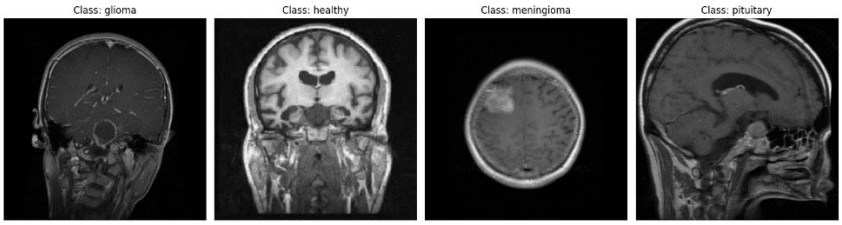

Figure 2. Sample of Dataset

2. Image segmentation

Image segmentation is a crucial initial step to isolate relevant brain areas from MRI images, particularly for detecting tumors with high precision. In this study, the U-Net model was chosen for segmentation due to its superior performance in semantic segmentation tasks compared to other methods, such as Fully Convolutional Networks (FCN) [19]. The symmetrical encoder-decoder structure of U-Net allows for the integration of high-resolution spatial information from the encoder with contextual information from the decoder, resulting in more accurate segmentation [11], [18], [20], [21]. The choice of U-Net was further motivated by its proven success in similar tasks, such as organ segmentation in medical imaging.

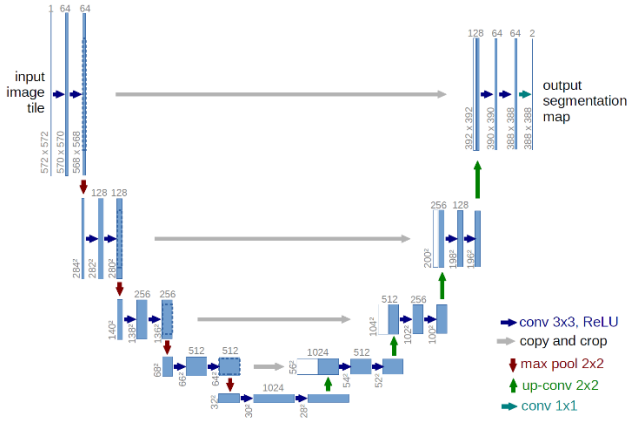

Figure 3. U-Net Architecture [18]

To optimize the segmentation process, the dice loss function was employed, given its capability to handle pixel imbalances between the background and the target object, such as tumors, which are typically small in proportion to the entire image. The primary goal of segmentation is to separate the brain regions with potential

tumors from the image background, thereby enhancing the focus of subsequent analyses. The optimization of the U-Net model is performed using the dice loss function, which is formulated as follows.

$$Dice\ loss = 1 - \frac{2 \cdot \sum(P \cdot G)}{\sum P + \sum G} \quad (1)$$

where P represents the predicted pixel probability, and G denotes the ground truth. The following are the original images and their segmentation results, demonstrating the effectiveness of U-Net in isolating significant regions from the background.

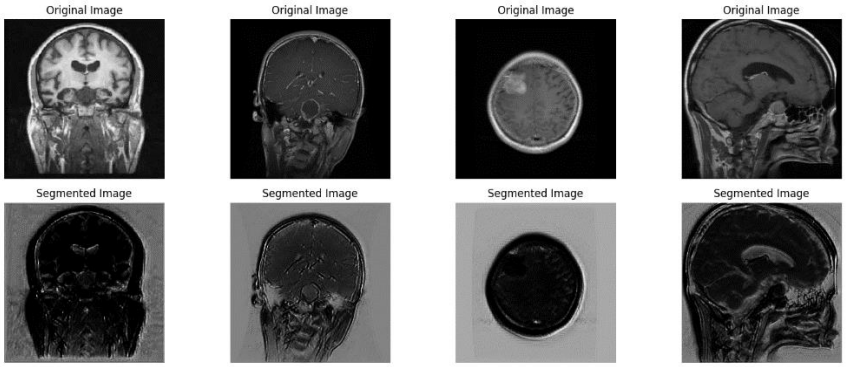

Figure 3. Image segmentation

3. Data Augmentation

Data augmentation was applied to enhance the diversity of the dataset while minimizing the risk of overfitting during model training. The augmentation process involved various transformations, including rotation, horizontal and vertical flipping, as well as contrast adjustments, to generate more varied data samples [22], [23]. As a result, the number of samples in each class was balanced to 2,000, ensuring an even distribution for both training and testing datasets. Data augmentation plays a critical role in this study as it helps mitigate the risk of overfitting by providing a more diverse dataset, enabling the model to generalize better to unseen data. Without augmentation, the imbalance in the dataset—such as fewer samples in certain classes—could lead to biased learning, where the model performs well on dominant classes but poorly on underrepresented ones. By generating additional samples through techniques such as rotation, flipping, and contrast adjustment, data augmentation ensures that each class is adequately represented during training and testing, which is crucial for improving classification performance and robustness, particularly in challenging tasks like brain tumor detection.

4. Feature Extraction

After segmentation, features were extracted using HuMoments, a technique that numerically represents the shape and distribution of objects within an image [24]–[26]. HuMoments were utilized for their ability to capture spatial patterns of shapes and distributions in MRI images. Additionally, their invariance to rotation, translation, and scale makes them ideal for MRI data, which often exhibit variations in orientation and size across subjects. The basic formula for geometric moments M_{pq} in Equation 2:

$$M_{pq} = \sum_{x} \sum_{y} x^p y^q I(x, y) \quad (2)$$

Where $I(x, y)$ represents the pixel intensity at the coordinate x, y. Based on these moments, HuMoments are computed using non-linear combinations to produce features that are invariant to transformations. This feature extraction process enables a compact and informative numerical representation, which is crucial for machine learning applications.

Table 1. Extraction Results

No	H1	H2	H3	H4	H5	H6	H7	Label
1	0.061531	0.00011	4.33E-05	7.02E-05	1.47E-09	-4.65E-07	3.58E-09	1
2	0.10655	0.005565	0.00019	6.17E-05	-6.66E-09	-3.56E-06	5.09E-10	1
…	…	…	…	…	…	…	…	…
7998	0.031556	1.41E-05	9.61E-07	3.34E-06	1.56E-12	-1.25E-08	5.78E-12	3
7999	0.023381	6.03E-05	8.06E-07	1.19E-06	-1.07E-12	-6.87E-09	-4.71E-13	3

No	H1	H2	H3	H4	H5	H6	H7	Label
8000	0.062603	0.000192	7.55E-07	4.22E-05	6.55E-11	1.70E-07	2.29E-10	3

5. Normalization

Normalization is performed to align the feature scales and reduce the impact of outliers. Two methods are employed: Robust Scaling and L2 Normalization.

- Robust Scaling: Robust Scaling is chosen for its ability to handle outliers commonly found in imaging data features. The formula is Equation 3:

$$X_{scaled} = \frac{X - Median(X)}{IQR(X)} \tag{3}$$

- L2 Normalization: L2 Normalization is applied to harmonize the scale across features, enabling algorithms such as XGBoost and Random Forest to effectively learn data patterns. The formula is given by Equation 4:

$$X_{norm} = \frac{X}{||X||_2}, ||X||_2 = \sqrt{\sum_{i=1}^{n} X_i^2} \tag{4}$$

The results of the normalization process are visualized using scatter plots. Scatter plots serve as an effective tool for understanding data distribution and relationships between features, facilitating the analysis of relevant patterns in the dataset. This visualization helps evaluate the success of normalization in aligning feature scales and mitigating the influence of outliers.

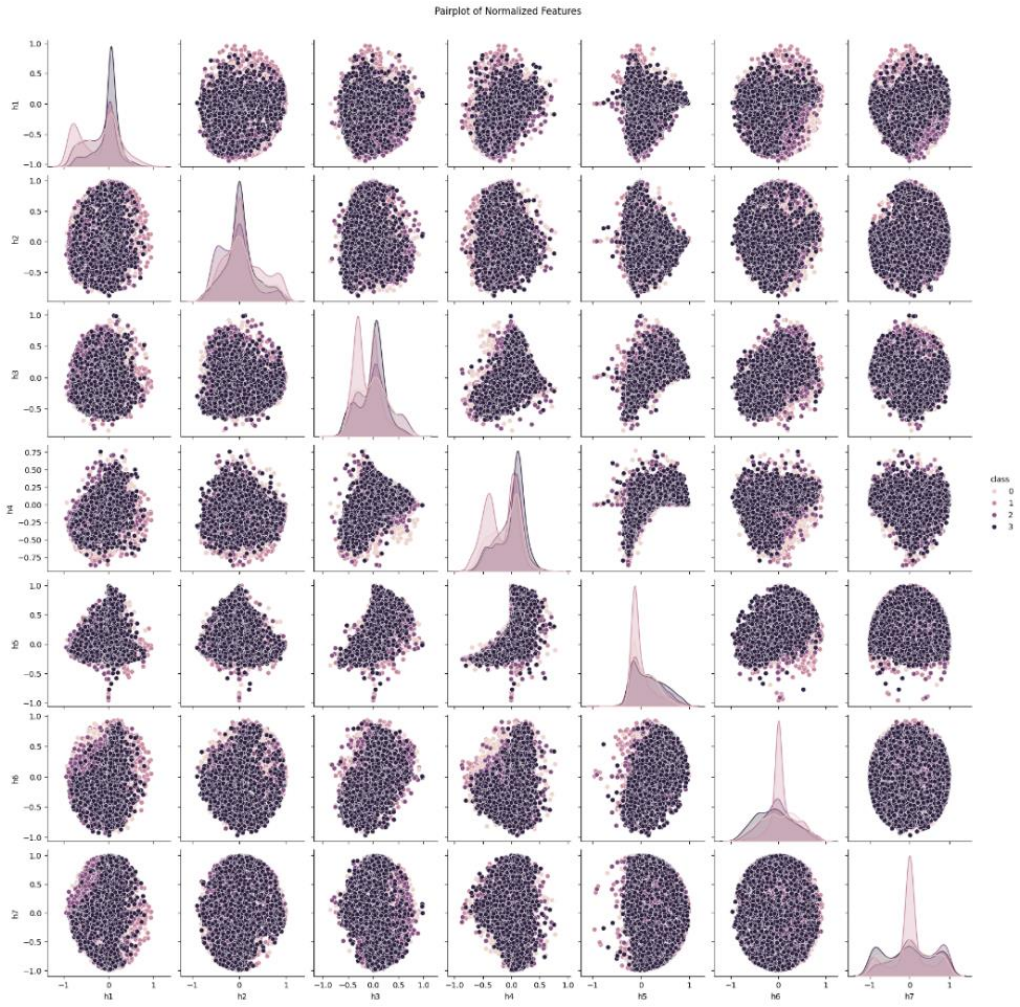

Figure 4. Scatter Plot

In the scatter plot visualization below, each point represents an MRI data sample, with colors indicating its class (glioma, meningioma, pituitary, or healthy). This scatter plot aids in evaluating the extent to which data

normalization successfully aligns feature scales and mitigates the influence of outliers. Through this visualization, the distribution patterns across classes and potential overlaps between features can be better understood, which is a crucial step in improving the accuracy of machine learning models.

C. Data Analysis

The data analysis process in this study involves training machine learning models, validating their performance, and evaluating their classification capabilities using various metrics. The following steps outline the analysis process:

1. Splitting Data
   The dataset was divided into two subsets: 80% for training and 20% for testing, with the division performed randomly. This split ensures balanced class distributions in both subsets, enabling the evaluation of the model's generalization capability. The training data is used to train the models, while the testing data provides an independent evaluation to assess the models' real-world performance.

2. Training the Models
   Three ensemble algorithms were employed for model training: Random Forest, XGBoost, and Stacking. Each algorithm was chosen for its unique strengths in handling complex datasets and its ability to improve classification performance:
   - Random Forest: This algorithm leverages multiple decision trees to enhance accuracy and stability [27]–[29]. It predicts the output class based on the average probabilities from all trees in the forest. Random Forest excels in handling high-dimensional data and reducing the risk of overfitting by using a combination of multiple decision trees, each trained on random subsets of data and features. This ensemble approach not only improves accuracy but also enhances model robustness by averaging the predictions, making it less sensitive to noise in the dataset. Moreover, its ability to measure feature importance provides valuable insights for understanding which variables contribute most to the classification, defined as Equation 5:

$$P(y = c|X) = \frac{1}{N}\sum_{i=1}^{N} P_i(y = c|X) \quad (5)$$

     Where p_i represents the probability predicted by the i-th tree.

   - XGBoost: A gradient boosting algorithm that iteratively optimizes a loss function with regularization to control model complexity. XGBoost is particularly advantageous for its efficiency and scalability, utilizing techniques like parallel processing and tree pruning to reduce computational cost. Its regularization features, such as L1 and L2 penalties, prevent overfitting, making it suitable for datasets with high variability. Furthermore, XGBoost employs a weighted quantile sketch algorithm for better handling of sparse data and uneven distributions, ensuring robust performance in real-world applications. The optimization is expressed as Equation 6:

$$L = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (6)$$

     Where l is the loss function, and $\Omega(f_k)$ is the regularization term.

   - Stacking: This technique combines predictions from multiple base models (Random Forest, Gradient Boosting, and Support Vector Classifier) into a meta-model. The meta-model uses the probability outputs from the base models as input features, enabling it to learn patterns and relationships across the individual model predictions. The strength of Stacking lies in its ability to exploit the complementary nature of the base models; for instance, Random Forest excels at reducing variance through ensemble averaging, Gradient Boosting captures complex patterns by iteratively minimizing errors, and Support Vector Classifier effectively handles high-dimensional feature spaces. By integrating these diverse capabilities, Stacking improves both accuracy and generalization, particularly in complex datasets. This method is particularly beneficial when the individual base models exhibit varied strengths and weaknesses, as the meta-model synthesizes these insights to produce more reliable and robust predictions.

3. Evaluation Model Performance
   The trained models were evaluated using three primary metrics:
   - Classification Report: This provides detailed metrics for each class, including precision, recall, F1-score, and overall accuracy [30]. These metrics offer insights into the model's performance in predicting each class [31].
   - Confusion Matrix: This presents a detailed view of prediction errors by comparing the actual and predicted classes, highlighting areas where the model underperforms.

- ROC-AUC: The Receiver Operating Characteristic Area Under the Curve (ROC-AUC) measures the model's ability to distinguish between positive and negative classes. It is calculated as the area under the ROC curve by Equation 7:

$$AUC = \int_0^1 TPR(FPR)\, d(FPR) \tag{7}$$

where TPR (True Positive Rate) and FPR (False Positive Rate) are defined in Equation 8:

$$TPR = \frac{TP}{TP + FN}, FPR = \frac{FP}{FP + FN} \tag{8}$$

This structured approach ensures that the models' performance is rigorously assessed, and their strengths and limitations are thoroughly analyzed. The evaluation metrics provide comprehensive insights into the classification results, aiding in the refinement of the models for improved accuracy and reliability.

Results and Discussion

A. Results

This study evaluates three machine learning models, namely Random Forest, XGBoost, and Stacking, to classify brain MRI images into four categories: glioma, meningioma, pituitary, and healthy. The evaluation was conducted using metrics such as accuracy, ROC-AUC, F1-score, and confusion matrix to comprehensively assess the performance of each model. The performance comparison results are shown in **Figure 5**.

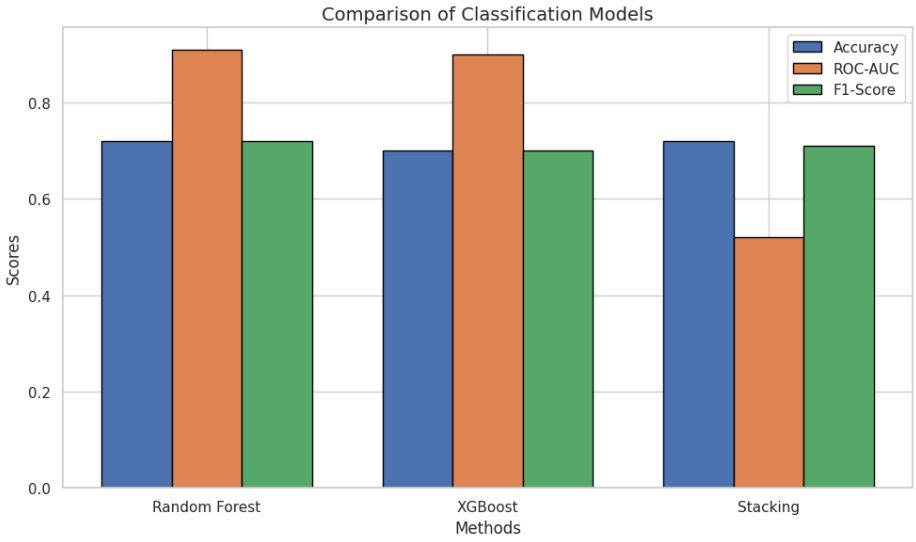

Figure 5. Comparison of Classification Models

1. Random Forest

Table 2. Classification Report of the Random Forest Model

	Precision	Recall	F1-score	Support
Glioma	0.70	0.67	0.69	397
Healthy	0.91	0.87	0.89	400
Meningioma	0.60	0.59	0.6	399
Pituitary	0.67	0.72	0.69	399
Accuracy			0.72	1595
Macro avg	0.72	0.72	0.72	1595
Weighted avg	0.72	0.72	0.72	1595

The Random Forest model demonstrated the best performance with an accuracy of 72% and the highest ROC-AUC of 0.91, indicating its superior ability to distinguish between categories. The average precision, recall, and F1-score for this model are each 0.72.

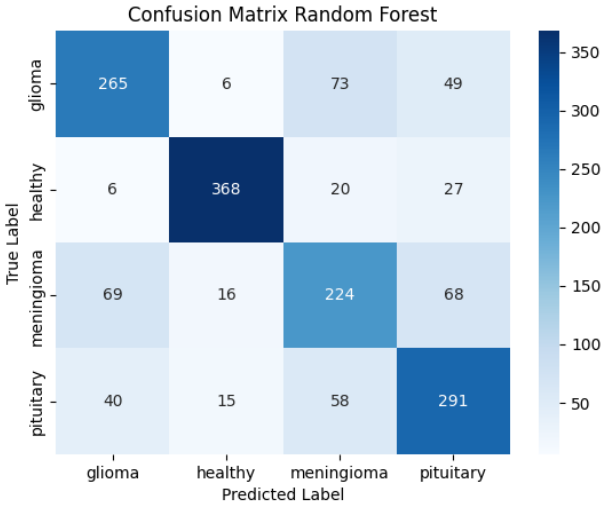

Figure 6. Confusion Matrix for the Random Forest Model

Confusion matrix indicates that Random Forest has the best prediction distribution, with fewer errors compared to the other two models. However, there is a significant misclassification of glioma as meningioma (73 cases), suggesting that the features of these two classes share visual similarities.

2. XGBoost

Table 3. Classification Report of the XGBoost Model

	Precision	**Recall**	**F1-score**	**Support**
Glioma	0.67	0.66	0.67	397
Healthy	0.83	0.85	0.84	400
Meningioma	0.62	0.59	0.6	399
Pituitary	0.66	0.70	0.68	399
Accuracy			0.70	1595
Macro avg	0.70	0.70	0.70	1595
Weighted avg	0.70	0.70	0.70	1595

The XGBoost model achieved an overall accuracy of 70% on the test dataset, demonstrating competitive performance across the four classes. For the glioma class, the model recorded a precision of 0.67, a recall of 0.66, and an F1-score of 0.67, indicating moderate performance in correctly identifying this category.

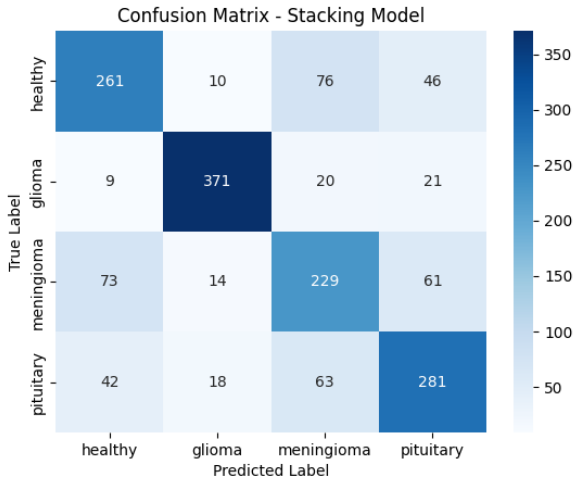

Figure 7. Confusion Matrix for the XGBoost Model

The confusion matrix for XGBoost reveals slightly weaker performance, particularly in the meningioma class, where many predictions were incorrectly classified as pituitary. This could be attributed to XGBoost's limitations in capturing feature patterns without further parameter optimization.

3. Stacking

Table 4. Classification Report of the Stacking Model

	Precision	Recall	F1-score	Support
Glioma	0.90	0.88	0.89	397
Healthy	0.68	0.66	0.67	400
Meningioma	0.59	0.61	0.60	399
Pituitary	0.69	0.70	0.69	399
Accuracy			0.72	1595
Macro avg	0.71	0.71	0.71	1595
Weighted avg	0.72	0.72	0.72	1595

The Stacking technique achieved an accuracy comparable to Random Forest (72%), but its significantly lower ROC-AUC score of 0.52 highlights the model's limitation in effectively distinguishing between classes. Nevertheless, the average precision, recall, and F1-score of 0.71 indicate that the model remains competitive, particularly in reducing misclassifications for the glioma and healthy categories.

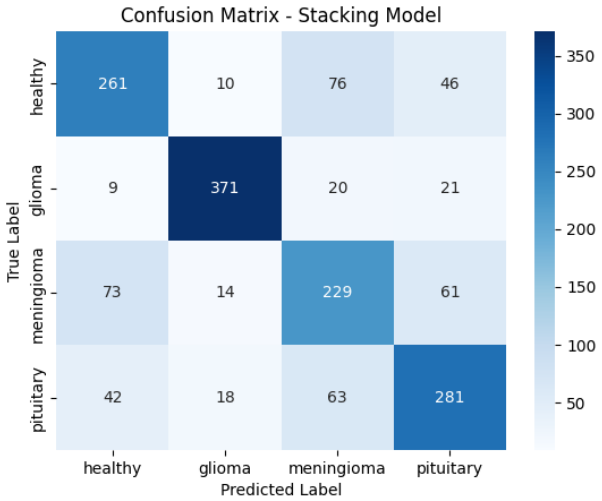

Figure 8. Confusion Matrix for the XGBoost Model

The confusion matrix for Stacking, despite achieving accuracy comparable to Random Forest, reveals weaknesses in distinguishing classes with similar feature distributions. Specifically, in this research, the Stacking model shows significant misclassifications in the meningioma class, where many instances are predicted as glioma (73 cases) or pituitary (61 cases). This indicates that the model struggles to differentiate meningioma from these two classes, likely due to overlapping texture and intensity patterns. Additionally, the healthy class has moderate misclassifications, with some instances being predicted as meningioma (76 cases).

B. Discussion

The results of this study indicate that the Random Forest model achieved the best overall performance, particularly evidenced by its ROC-AUC score of 0.91, demonstrating a strong ability to distinguish between positive and negative classes. The superiority of Random Forest can be attributed to its ensemble nature, which combines predictions from multiple decision trees, thereby enhancing the model's stability and accuracy.

The XGBoost model, although competitive with an accuracy of 70%, requires further optimization of parameters such as learning rate and tree depth to maximize its ability to capture inter-class patterns. Optimizing these parameters could enable XGBoost to better capture subtle distinctions between overlapping classes, reducing misclassifications. For example, a deeper tree depth allows the model to learn more complex hierarchical relationships within the data,

while a finely tuned learning rate balances convergence speed and generalization. Research [32] demonstrated that XGBoost can outperform traditional ensemble methods when parameter tuning is conducted effectively, which highlights a limitation in this study as exhaustive parameter optimization was not performed. The Stacking technique, while achieving an accuracy of 72%, exhibited weaknesses with its low ROC-AUC score of 0.52. A low ROC-AUC score indicates that the model struggles to distinguish between positive and negative classes, which can lead to inconsistent predictions, particularly for classes with overlapping features such as meningioma and glioma. This limitation may reduce its reliability for high-stakes applications such as clinical diagnoses, where clear class separation is critical.

This suggests that the meta-model structure was not effective in combining outputs from the base models. The ineffectiveness likely stems from insufficient diversity among the base models, as they may have captured similar feature representations, leading to redundant input for the meta-model. Moreover, the limited amount of training data for the meta-model might have constrained its ability to generalize effectively. To enhance the performance of the Stacking model, future research should consider integrating a more diverse set of base learners, including fundamentally different models such as neural networks or probabilistic classifiers. Additionally, optimizing the meta-model with advanced techniques like regularization or weighted voting schemes could improve its ability to synthesize the strengths of the base models.

Conclusion

This study successfully demonstrated that a machine learning-based approach combining U-Net image segmentation, HuMoments feature extraction, and data normalization can produce a reliable brain tumor classification model. The results showed that the Random Forest model delivered the best performance with an accuracy of 72%, an ROC-AUC score of 0.91, and an F1-score of 72%. This model excelled in distinguishing various brain tumor categories, particularly the healthy category. XGBoost followed with an accuracy of 70% and an ROC-AUC score of 0.90, while Stacking achieved an accuracy comparable to Random Forest but was limited by an ROC-AUC score of 0.52.

These findings support the initial statement in the introduction that integrating machine learning methods can improve the efficiency and accuracy of brain tumor diagnosis compared to manual approaches. However, the results also indicate that certain categories, such as glioma and meningioma, still frequently experience misclassification, highlighting the need for improvements in feature extraction and model parameter adjustments. Future research could be developed by incorporating texture- or morphology-based features and using larger and more diverse datasets to enhance the model's generalization capability. Additionally, hyperparameter optimization techniques for XGBoost and Stacking have the potential to further improve overall performance. With this approach, the resulting model could contribute to the development of more accurate and efficient machine learning-based medical diagnostic systems for clinical applications.

References

[1] M. Martucci *et al.*, "Magnetic Resonance Imaging of Primary Adult Brain Tumors: State of the Art and Future Perspectives," *Biomedicines*, vol. 11, no. 2, p. 364, Jan. 2023, doi: 10.3390/biomedicines11020364.

[2] R. Ranjbarzadeh, A. Caputo, E. B. Tirkolaee, S. Jafarzadeh Ghoushchi, and M. Bendechache, "Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools," *Comput. Biol. Med.*, vol. 152, p. 106405, Jan. 2023, doi: 10.1016/j.compbiomed.2022.106405.

[3] R. Vallée, "Machine learning decision tree models for multiclass classification of common malignant brain tumors using perfusion and spectroscopy MRI data," *Front. Oncol.*, vol. 13, 2023, doi: 10.3389/fonc.2023.1089998.

[4] A. A. Asiri, "Machine Learning-Based Models for Magnetic Resonance Imaging (MRI)-Based Brain Tumor Classification," *Intell. Autom. Soft Comput.*, vol. 36, no. 1, pp. 299–312, 2023, doi: 10.32604/iasc.2023.032426.

[5] K. Prabhakar, "Implementation of Machine Learning Algorithms for Brain Tumor Classification using MRI and Numerical Dataset," *2022 IEEE 3rd Global Conference for Advancement in Technology, GCAT 2022*. 2022, doi: 10.1109/GCAT55367.2022.9971904.

[6] C. J. Tseng, "An optimized XGBoost technique for accurate brain tumor detection using feature selection and image segmentation," *Healthc. Anal.*, vol. 4, 2023, doi: 10.1016/j.health.2023.100217.

[7] O. A. M. F. Alnaggar, "MRI Brain Tumor Detection Using Boosted Crossbred Random Forests and Chimp Optimization Algorithm Based Convolutional Neural Networks," *Int. J. Intell. Eng. Syst.*, vol. 15, no. 2, pp. 36–46, 2022, doi: 10.22266/ijies2022.0430.04.

[8] M. Sandeep, "Brain Tumor Detection Using Random Forest Algorithm in Comparison With K-nearest Neighbors Algorithm to Measure the Accuracy, Precision and Recall," *AIP Conference Proceedings*, vol. 2821, no. 1. 2023, doi: 10.1063/5.0158397.

[9] G. Pallavi, "Brain tumor detection with high accuracy using random forest and comparing with thresholding method," *AIP Conference Proceedings*, vol. 2853, no. 1. 2024, doi: 10.1063/5.0198189.

[10] A. Ilhan, "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, pp. 589–600, 2022, doi: 10.1007/s11548-022-02566-7.

[11] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021, doi: 10.1109/ACCESS.2021.3086020.

[12] Y. Jusman, "Machine Learnings of Dental Caries Images based on Hu Moment Invariants Features," *Proc. - 2021 Int. Semin. Appl. Technol. Inf. Commun. IT Oppor. Creat. Digit. Innov. Commun. within Glob. Pandemic, iSemantic 2021*, pp. 296–299, 2021, doi: 10.1109/iSemantic52711.2021.9573208.

[13] D. V Kondusov, "Comparison of 3D Models Using Hu Moment Invariants," *Russ. Eng. Res.*, vol. 40, no. 7, pp. 570–574, 2020, doi: 10.3103/S1068798X20070199.

[14] C. Lin, "Harmonic Beltrami Signature: A Novel 2D Shape Representation for Object Classification," *SIAM J. Imaging Sci.*, vol. 15, no. 4, pp. 1851–1893, 2022, doi: 10.1137/22M1470852.

[15] D. Divyamary, S. Gopika, S. Pradeeba, and M. Bhuvaneswari, "Brain Tumor Detection from MRI Images using Naive Classifier," in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Mar. 2020, pp. 620–622, doi: 10.1109/ICACCS48705.2020.9074213.

[16] Preetika, "MRI Image based Brain Tumour Segmentation using Machine Learning Classifiers," *2021 International Conference on Computer Communication and Informatics, ICCCI 2021*. 2021, doi: 10.1109/ICCCI50826.2021.9402508.

[17] Y. Boer, "Classification of Heart Disease: Comparative Analysis using KNN, Random Forest, Gaussian Naive Bayes, XGBoost, SVM, Decision Tree, and Logistic Regression," *2023 5th International Conference on Cybernetics and Intelligent Systems, ICORIS 2023*. 2023, doi: 10.1109/ICORIS60118.2023.10352195.

[18] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Med. Image Comput. Comput. Interv. – MICCAI 2015 (MICCAI 2015)*, vol. 9351, no. Cvd, pp. 12–20, 2015, doi: 10.1007/978-3-319-24574-4.

[19] O. Ozturk, B. Sarıtürk, and D. Z. Seker, "Comparison of Fully Convolutional Networks (FCN) and U-Net for Road Segmentation from High Resolution Imageries," *Int. J. Environ. Geoinformatics*, vol. 7, no. 3, pp. 272–279, Dec. 2020, doi: 10.30897/ijegeo.737993.

[20] 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.

[21] J.-S. Long, G.-Z. Ma, E.-M. Song, and R.-C. Jin, "Learning U-Net Based Multi-Scale Features in Encoding-Decoding for MR Image Brain Tissue Segmentation," *Sensors*, vol. 21, no. 9, p. 3232, May 2021, doi: 10.3390/s21093232.

[22] D. R. Nayak, "Brain Tumor Classification Using Dense Efficient-Net," *Axioms*, vol. 11, no. 1, 2022, doi: 10.3390/axioms11010034.

[23] A. B. Abdusalomov, "Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging," *Cancers (Basel).*, vol. 15, no. 16, 2023, doi: 10.3390/cancers15164172.

[24] I. N. A. Nastase, "Image Moment-Based Features for Mass Detection in Breast US Images via Machine Learning and Neural Network Classification Models," *Inventions*, vol. 7, no. 2, 2022, doi: 10.3390/inventions7020042.

[25] N. Elazab, "Brain Cancer Diagnosis Based on Histopathological Images Using Handcrafted Features," *18th International Computer Engineering Conference, ICENCO 2022*. pp. 108–113, 2022, doi: 10.1109/ICENCO55801.2022.10032507.

[26] A. Abisha, "Feature Extraction from Plant Leaves and Classification of Plant Health Using Machine Learning," *Lecture Notes in Electrical Engineering*, vol. 858. pp. 867–876, 2022, doi: 10.1007/978-981-19-0840-8_67.

[27] P. Nagaraj, "Ensemble Machine Learning (Grid Search & Random Forest) based Enhanced Medical Expert Recommendation System for Diabetes Mellitus Prediction," *3rd International Conference on Electronics and Sustainable Communication Systems, ICESC 2022 - Proceedings*. pp. 757–765, 2022, doi: 10.1109/ICESC54411.2022.9885312.

[28] M. Salem, "Random Forest modelling and evaluation of the performance of a full-scale subsurface constructed wetland plant in Egypt," *Ain Shams Eng. J.*, vol. 13, no. 6, 2022, doi: 10.1016/j.asej.2022.101778.

[29] X. Yu, "Random forest algorithm-based classification model of pesticide aquatic toxicity to fishes," *Aquat. Toxicol.*, vol. 251, 2022, doi: 10.1016/j.aquatox.2022.106265.

[30] H. Azis, L. Syafie, F. Fattah, and ..., "Unveiling Algorithm Classification Excellence: Exploring Calendula and Coreopsis Flower Datasets with Varied Segmentation Techniques," *2024 18th Int. ...*, 2024, doi: 10.1109/IMCOM60618.2024.10418246.

[31] N. Rismayanti and A. P. Utami, "Improving Multi-Class Classification on 5-Celebrity-Faces Dataset using Ensemble Classification Methods," *Indones. J. Data ...*, 2023, doi: 10.56705/ijodas.v4i2.78

[32] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.