Int. J. Advanced Networking and Applications
Volume: 17, Issue: 03, Pages : 6970-6981 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.2025.17310

Enhanced Brain Tumor Classification using Ensemble Machine Learning

Sugandha Singh
Ph.D. Scholar, Department of Computer Science, BBAU, Lucknow
Email: singhsugandha3@gmail.com
Vipin Saxena
Professor, Department of Computer Science, BBAU, Lucknow
Email: profvipinsaxena@gmail.com

ABSTRACT

Accurate classification of brain tumors is vital for timely diagnosis and effective treatment planning. This study investigates the performance of several machine learning algorithms—including Logistic Regression (LR), Random Forest (RF), Naive Bayes (NB), K-Nearest Neighbors (KNN), and Decision Tree (DT)—for the classification of meningioma and schwannoma brain tumors using MRI scan data. The experimental results reveal that KNN achieves the highest individual performance with 99% precision, recall, F1-score, and accuracy, while LR and RF follow closely. Naive Bayes, although less accurate, demonstrates utility in computationally constrained environments. The study further explores hybrid models by combining classifiers through ensemble techniques such as voting, bagging, and stacking. Among these, the voting classifier based on LR, KNN, and RF achieves an accuracy of 98%, indicating improved robustness and predictive capability. These findings underscore the potential of machine learning, particularly KNN and ensemble-based methods, in facilitating reliable brain tumor classification, thereby supporting clinical decision-making and improving patient outcomes.

Keywords **: brain tumor, K nearest neighbor, logistic regression, machine learning, magnetic resonance images, random forest**

Date of Submission: September 22, 2025 Date of Acceptance: October 20, 2025

1. Introduction

Brain tumor is a complex and heterogeneous group of neoplasms originating within the central nervous system, ranging from meningioma to schwannoma lesions. Among the frequently observed types, meningioma originates from the meninges—the protective layers surrounding the brain and spinal cord—and is generally non-cancerous but may exert pressure on nearby tissues. Schwannoma arises from Schwann cells responsible for forming the myelin sheath around peripheral nerves, most commonly affecting the vestibular nerve. Magnetic Resonance Imaging (MRI) serves as a vital diagnostic technique for detecting, categorizing, and assessing such tumors because of its excellent soft-tissue contrast and spatial resolution. Accurate classification [19] of tumors is critical for guiding treatment strategies, predicting patient prognosis, and optimizing therapeutic outcomes. Traditional diagnostic approaches, Which rely heavily on manual interpretation of medical imaging such as MRI [21], are time consuming and subject to variability inherent in human assessment. In recent years, Machine Learning (ML) techniques have emerged as promising tools for automating and enhancing brain tumor [20] classification process.

These algorithms are capable of learning complex patterns and relationships from large volumes of imaging data, enabling algorithms to make accurate predictions and classifications [22, 23].

Among the wide range of ML algorithms, logistic regression (LR) [1-2] random forest (RF) [3-4], naive bayes (NB) [5], k nearest neighbor (KNN) [6-7], and decision tree (DT) [9] have gained significant attention for their effectiveness in medical image analysis tasks.

In the present study, conduct a comparative analysis of KNN, LR, RF, NB, and DT for brain tumor classification using medical imaging data. The By evaluation is based on the key performance metrics, including accuracy, precision, recall, and F1 score, with the objective of highlighting the strengths and limitations of each algorithm. Furthermore, discuss the existing challenges and future directions in applying ML techniques to brain tumor classification, aiming to support the development of personalized, data-driven approaches in neuro-oncology.

2. Related Work

Numerous classifiers have been proposed for automatic brain tumor classification using machine learning techniques. These approaches generally follow a structured pipeline consisting of image pre-processing, feature extraction, and classification. While various methods have been explored, the majority of existing studies primarily emphasize the classification phase. Several notable approaches from the literature are outlined below.

Noreen et al. [8], proposed an automated system for classifying glioma, meningioma, and pituitary tumors using a combination of deep learning and machine learning techniques. The study utilized fine-tuned models such as Inception-v3 and Xception for feature extraction , followed by classifiers including Softmax, Random Forest, Support Vector Machine (SVM), KNN, and various ensemble methods. Among these, the Inception-v3 model achieved the highest test accuracy of 94.34%.

Similarly, Kaur et al. [9] developed an automated MRI-based brain tumor diagnostic system aimed at early detection to reduce cancer-related mortality. Their approach involved multiple image processing stages, including pre-processing, segmentation, feature extraction, and classification. A modified median filter and a multi-vector segmentation machine were employed to enhance tumor segmentation. The results demonstrated a 10% improvement in classification accuracy compared to systems that did not incorporate the modified filtering technique.

Jena et al. [4] investigated brain tumor classification and segmentation using supervised machine learning combined with image processing techniques. The study employed seven different texture feature extraction methods alongside several classifiers, including Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Boosted Decision Trees (BDT), Random Forest (RF), and ensemble approaches. Additionally, hybrid segmentation algorithms were evaluated for tumor area delineation. The results demonstrated high classification accuracies ranging from 94.25% to 97%, along with a mean Dice score of 90.16% for tumor segmentation, indicating effective performance in both tasks.

Ganesh and Swarnalatha [10] highlighted the transformative potential of machine learning in disease diagnosis, particularly in cancer detection. Their study proposed a Support Vector Machine (SVM) classifier for brain tumor segmentation and diagnosis, utilizing a dataset of 50 MRI images. The proposed method achieved a tumor diagnosis accuracy of 100% in distinguishing benign from malignant cases, with a kernel accuracy of 90% for each classification category.

Çınarer et al. [11] proposed a deep neural network (DNN) model enhanced by a synthetic minority over-sampling technique (SMOTE) for accurate classification of brain tumors, particularly multicentric gliomas and multiple lesions. Trained on the Visually Accessible Rembrandt Images (VARI) dataset, the model achieved an accuracy of 95.0%, with precision, recall, and F1-score values of 95.4%, 95.0%, and 94.9%, respectively.

Similarly, Kaur et al. [12] developed an ensemble learning approach for classifying benign and malignant brain tumors from MRI images. The method employed Otsu's segmentation to isolate tumors and extracted hybrid features for classification. DT, KNN, and Support Vector Machine (SVM) classifiers were combined using a stacking model, resulting in an average accuracy of 97.91%.

Jefferson and Shanmugasundaram [13] proposed a brain tumor classification method using 3D MRI data by leveraging radiomics features extracted through Radiomics and a 3D Convolutional Neural Network (3D-CNN). Their combinational model, which integrates Gray Level Co-occurrence Matrix (GLCM) features with 3D-CNN and employs a KNN classifier, achieved an accuracy of up to 96.7%.

Nayak and Kengeri [14] proposed a hybrid NB classifier for classifying MRI brain images, focusing on improving the accuracy of distinguishing between normal and abnormal cases. Their model integrates image pre-processing, feature extraction, and noise reduction techniques, resulting in enhanced classification performance.

Soni et al. [15] developed a hybrid framework combining Convolutional Neural Networks (CNN) and Random Forest Classifier (RFC) for brain tumor classification using MRI images. The approach consists of three stages: pre-processing with a Gaussian filter, automatic feature xtraction via CNN, and classification using multiple algorithms including RF, KNN, DT, SVM, and NB. Using the Kaggle dataset, their model achieved accuracies of 99.61% on training data, 92.16% on validation data, and 71.2% on test data.

Das et al. [16] proposed an automated framework for classifying pediatric brain tumor biopsy samples, specifically medulloblastoma, using texture features. The study evaluated 172 features extracted from GLCM, Gray Level Run Length (GRLN), Histogram of Oriented Gradients (HOG), Tamura, and Local Binary Patterns (LBP), tested with six different classifiers. The results showed that combining all texture features yielded an accuracy of 100% on a dataset of 80 biopsy images captured at 10x magnification.

Reddy and Adimoolan [17] compared the performance of KNN and RF algorithms for brain tumor detection in MRI images. Using a dataset of 400 MRI samples, KNN achieved a higher accuracy of 94.59% compared to 87.94% for RF. Additionally, KNN demonstrated superior sensitivity (96.61%) and specificity (93.84%) relative to RF. These results indicate that the KNN algorithm is a novel and more effective approach for brain tumor detection in MRI data.

3. Material and methods

This section describes the overall architecture for brain tumor classification, as illustrated in Fig. 1. First, the input MRI images undergo pre-processing to enhance quality and remove noise before being fed into the ML models. Next, the pre-processed images serve as input to several ML classifiers for initial classification. Based on their evaluation performance, the top three classifiers are selected. These selected classifiers are then combined in an ensemble module, where their features are concatenated and fed into another ML classifier to generate the final prediction.

3.1 Image pre-processing

To prepare the data for subsequent ML tasks, Principal Component Analysis (PCA) was first applied to reduce dimensionality while retaining 99% of the original variance. This approach effectively addresses the high dimensionality of the dataset, improving computational efficiency and potentially enhancing model performance by mitigating the curse of dimensionality.

3.2 Dataset

Experiments were conducted using publicly available brain MRI datasets for the task of brain tumor classification. Both datasets were downloaded from Kaggle [18] and collectively referred to as the MR-Small-BinaryClass dataset for simplicity. The MR-Small-BinaryClass dataset contains 1,309 images, with 864 images of meningioma and 445 images of schwannoma tumors. The dataset was split into training (80%) and testing (20%) subsets. Table I provides detailed statistics of the datasets used in the experiments. Sample images from the MR-Small-BinaryClass dataset are shown in Fig. 2.

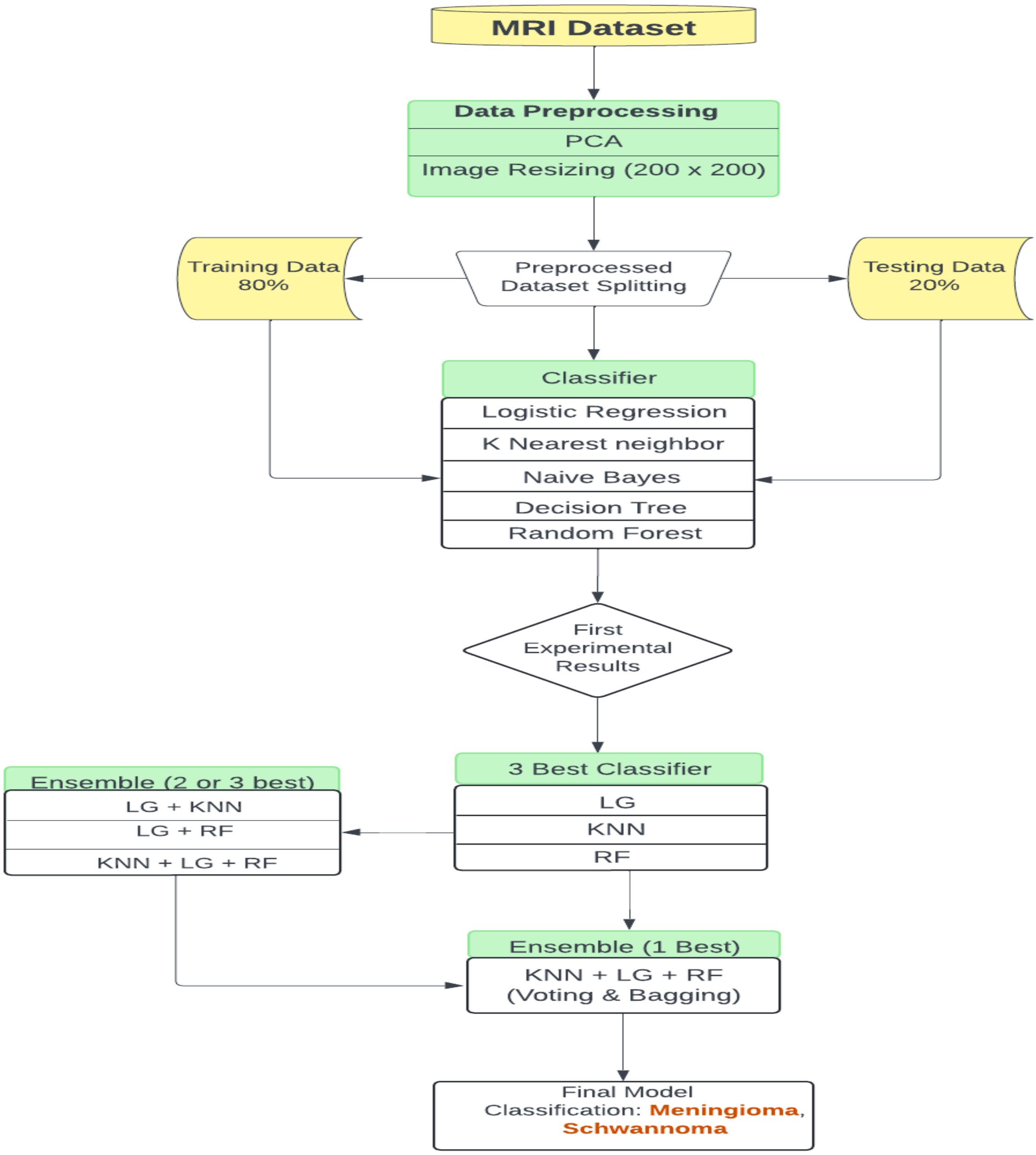
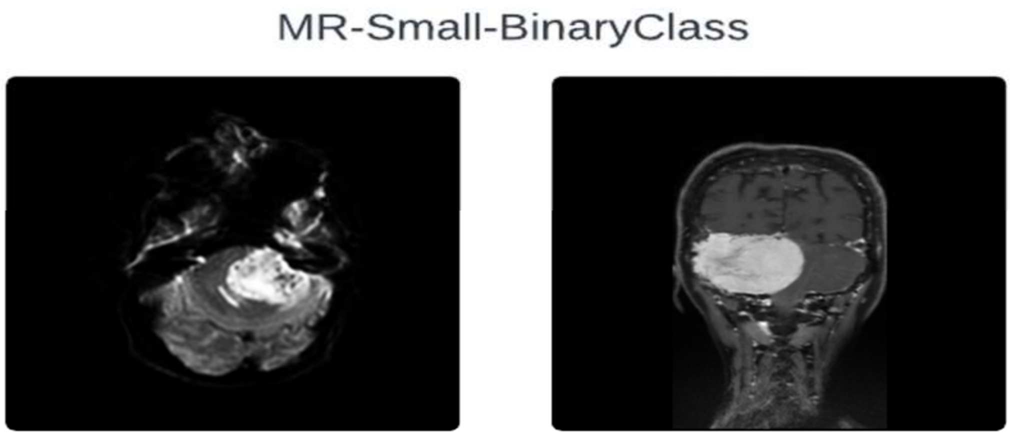


Fig.1 Architecture of proposed model



Int. J. Advanced Networking and Applications
Volume: 17, Issue: 03, Pages : 6970-6981 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.2025.17310

Fig. 2 Sample of brain MR images in MR-Small-BinaryClass dataset

TABLE 1.
DETAILS OF DATASET

Types	**Number of classes**	**Training set**	**Test set**
MR-Small-BinaryClass	2	864	225

3.3 Classification using ML Classifiers

ML algorithms are employed to classify MRI brain images as either meningioma or schwannoma. The study investigates and compares the performance of five machine learning algorithms: KNN, LR, RF, NB, and DT [19–23].

Naive Bayes (NB) classifiers, despite their conceptual simplicity, demonstrate competitive performance by assuming feature independence. This assumption makes them computationally efficient and well-suited for large-scale datasets. On the other hand, KNN is a simple yet powerful supervised machine learning algorithm used for both classification and regression tasks. As a non-parametric, instance-based learner, KNN does not make assumptions about the underlying data distribution and instead memorizes the training data without learning explicit model parameters.

Logistic Regression (LR) generally has a shorter inference time compared to ensemble methods because it relies on a single, simple mathematical model for prediction, making it computationally efficient.

Random Forest (RF) [24] is a versatile and widely used ensemble learning algorithm for both classification and regression tasks. It belongs to the family of decision tree-based methods and is known for its robustness, scalability, and capability to handle high-dimensional datasets with complex relationships.

Decision Tree (DT) is a non-parametric supervised learning method that partitions the feature space into regions based on input feature values. Each partition corresponds to a decision node in the tree, and the final prediction is made by traversing the tree from the root node down to a leaf node.

3.4 Evaluation Matrices

In this study, evaluate the performance of classifiers in distinguishing between meningioma and schwannoma brain tumors. The confusion matrix serves as a fundamental tool for this assessment, offering detailed insights into classifier accuracy as well as the types of classification errors. The confusion matrix for the classification results is presented below:

Table 2. Confusion matrix

		Predicted meningioma	**Predicted schwannoma**
	Class	**0**	**1**
Actual Meningioma	**0**	TP	FN
Actual Schwannoma	**1**	FP	TN

The confusion matrix, as shown in Fig. 3, provides several key metrics that indicate the classifier's performance. Accuracy, Precision, Recall, and F1-Score are the most important metrics and are calculated using the formulas provided in Equations (1)–(4).

$$\text{Accuracy} = (TN + TP) / (TN + TP + FP + FN) \quad (1)$$
$$\text{Precision} = TP / (FP + TP \quad (2)$$
$$\text{Recall} = TP / (FN + TP) \quad (3)$$
$$\text{F1-Score} = (2 \times \text{Recall} \times \text{Precision}) / (\text{Recall} + \text{Precision}) \quad (4)$$

where TN, TP, FN, and FP represent the number of True Negative, True Positive, False Negative, and False Positive instances, respectively.

4. Experimental Results

4.1 Experimental Settings

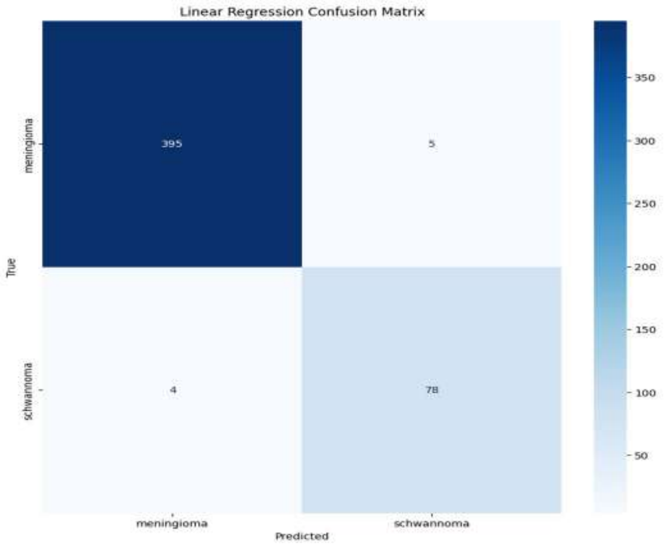
In the experiment, five different machine learning classifiers are used. Prior to training, the input images were pre-processed and resized to 200 × 200 pixels to meet the requirements of the classifiers. All experiments are conducted on Google Colab using a CPU hardware accelerator.

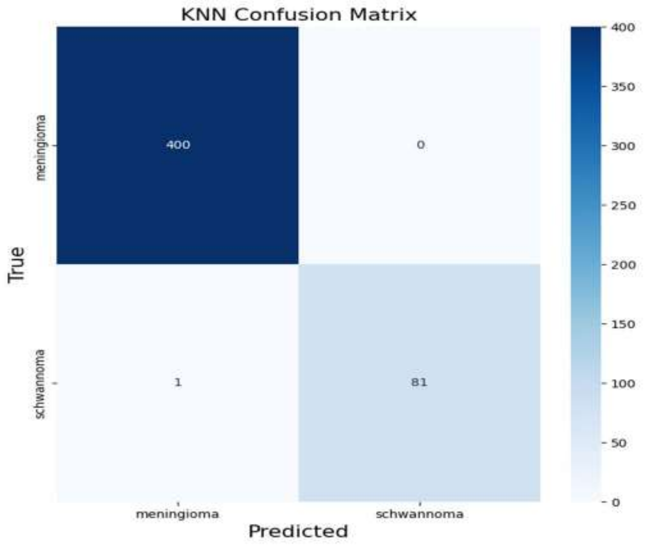
4.2 Results

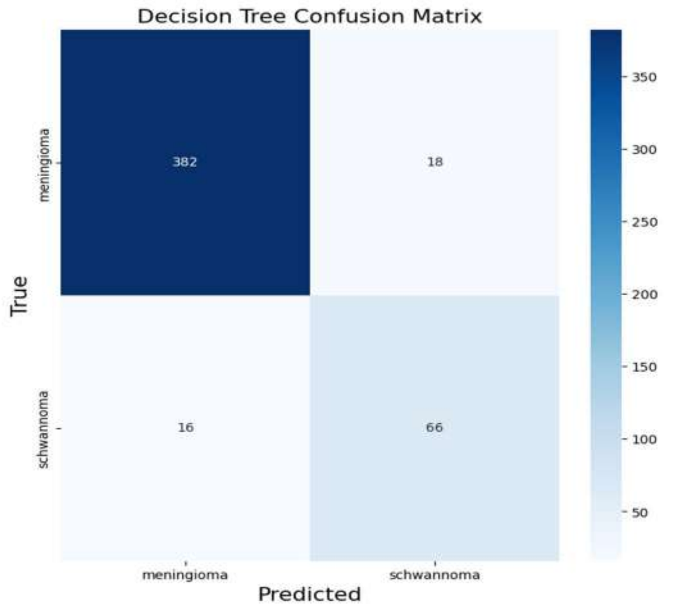
The first experiment is designed to compare the several machine learning classifiers for PCA dimensionality reduction. The second experiment aimed to demonstrate the effectiveness of the models built from the top 2 or 3 ML classifiers identified in the first experiment, using their selected features with various ML classifiers. The results of the first experiment on the small dataset are presented in Table III, where KNN, RF, and LR were identified as the top classifiers. As shown in Table IV and Fig. 4, the features from KNN, RF, and LR were selected to form the top three classifiers on the small dataset.

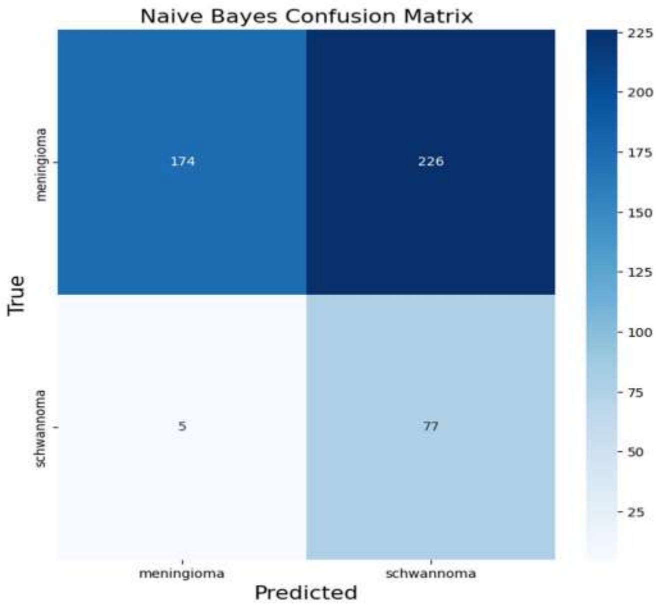
By analyzing both the AUC-ROC and AUC-PR curves, a more comprehensive evaluation of classification performance is achieved, enabling a balanced approach to model selection, especially in imbalanced learning scenarios, as shown in Fig. 4. Additionally, evaluating these metrics during both the training and testing phases helps in developing a robust and well-generalized classifier, as illustrated in Fig. 5.

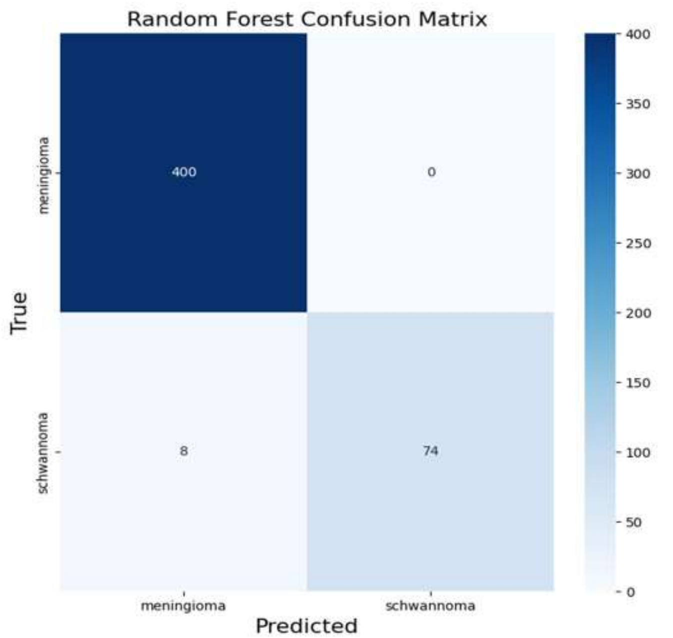
Int. J. Advanced Networking and Applications
Volume: 17, Issue: 03, Pages : 6970-6981 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.2025.17310

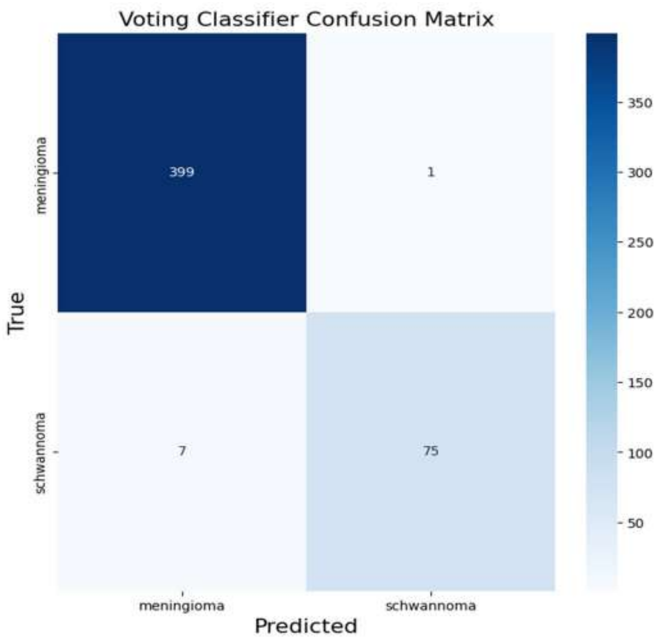












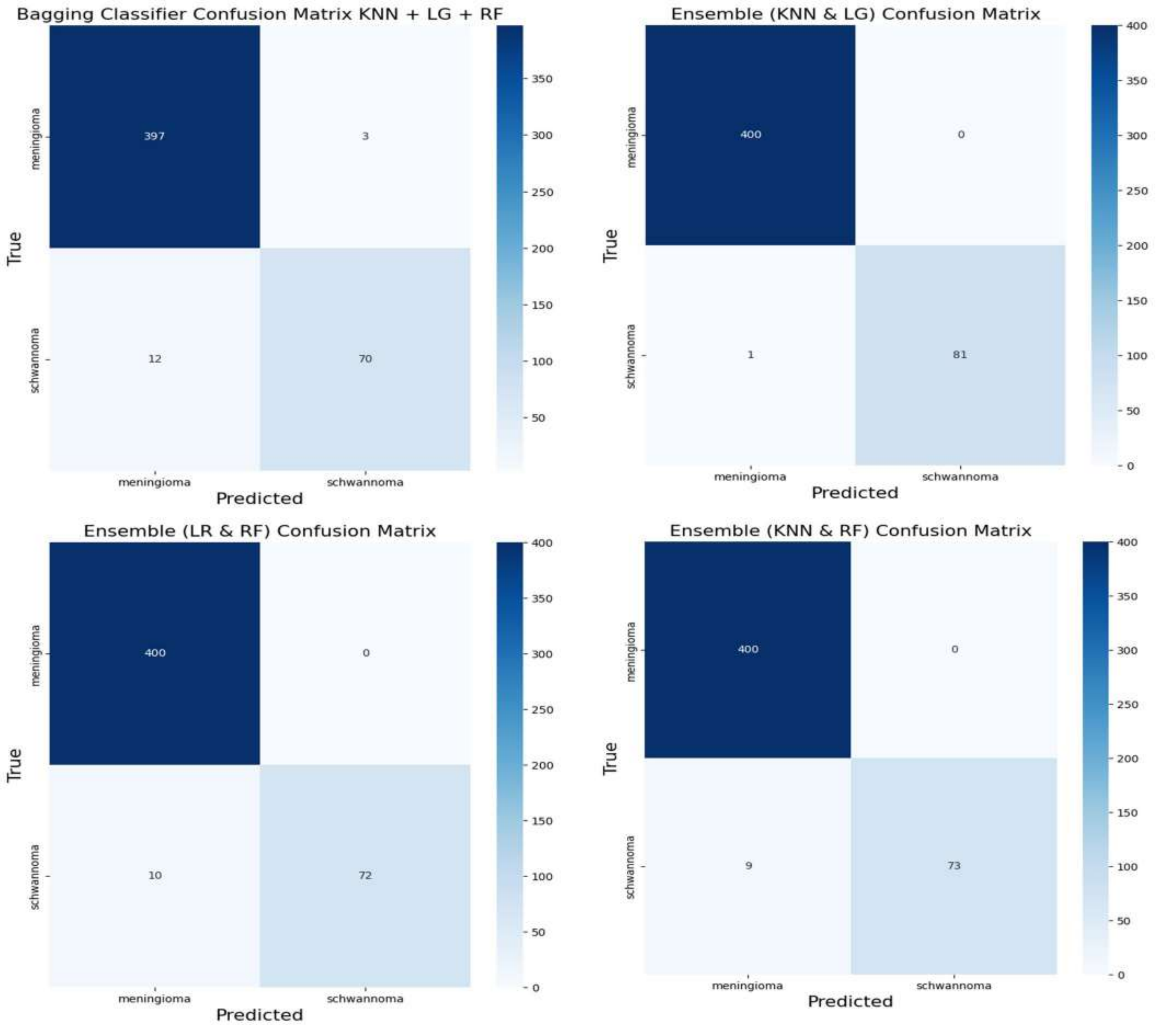
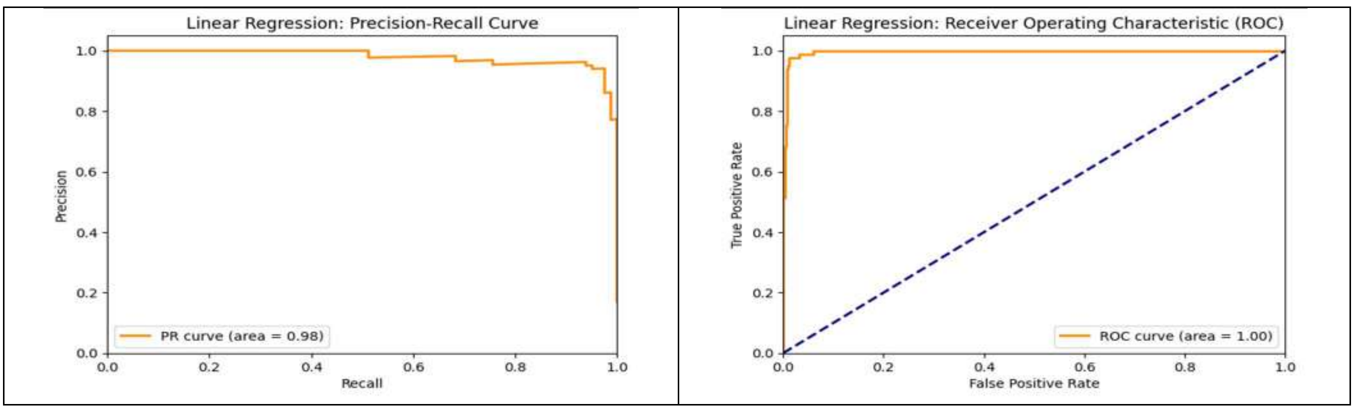


Fig. 3 Confusion matrices of KNN, LR, RF, NB, and DT for brain tumor classification



Int. J. Advanced Networking and Applications
Volume: 17, Issue: 03, Pages : 6970-6981 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.2025.17310

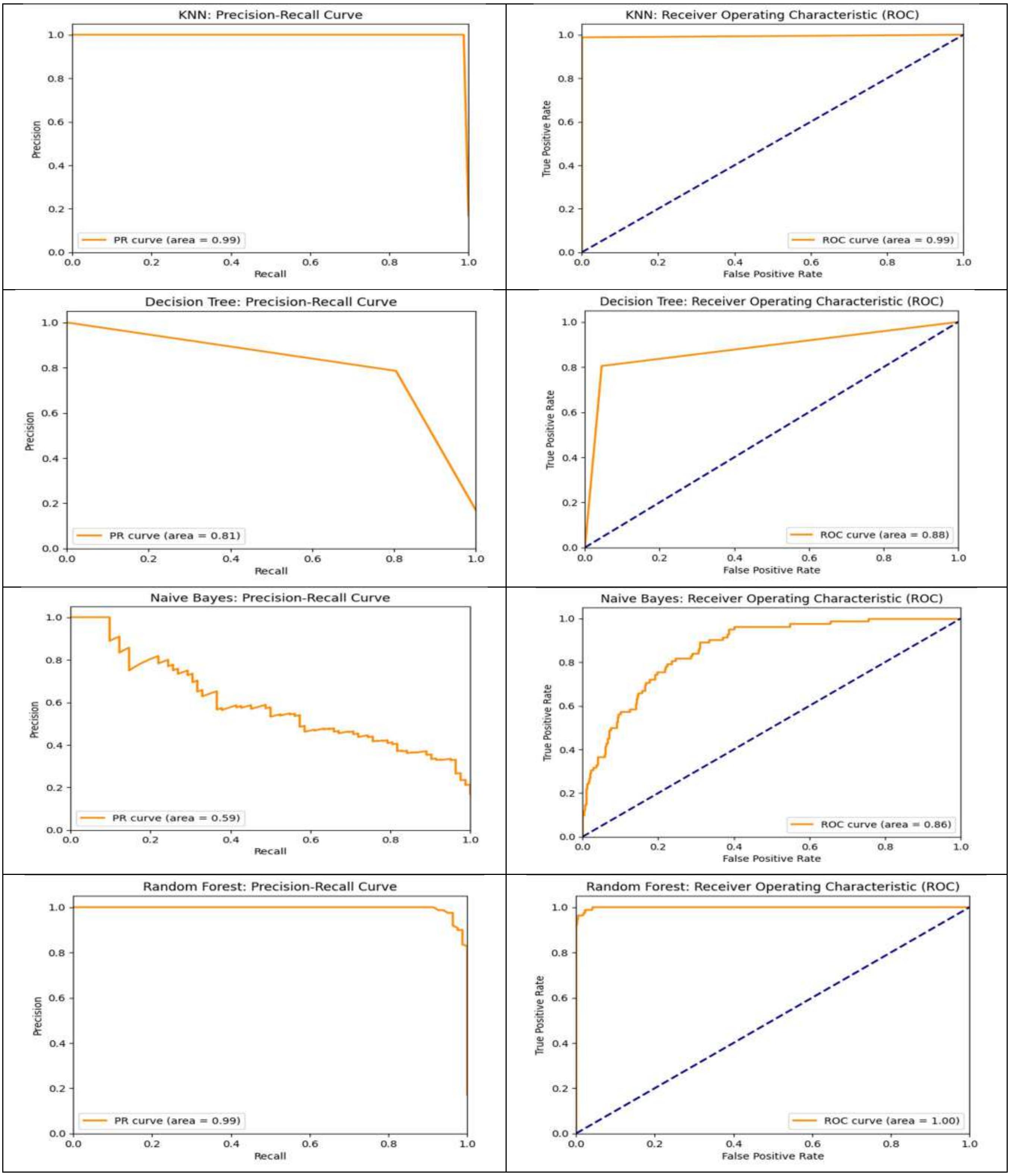


Fig. 4 Comparison of AUC-ROC and PR curves for evaluating model performance under class imbalance

```
Testing evaluation metrics: LG
              precision    recall  f1-score   support

           0       0.99      0.99      0.99       400
           1       0.94      0.95      0.95        82

    accuracy                           0.98       482
   macro avg       0.96      0.97      0.97       482
weighted avg       0.98      0.98      0.98       482

Accuracy: 0.9813278008298755
Precision: 0.9813278008298755
Recall: 0.9813278008298755
F1 score: 0.9813278008298755
Training evaluation metrics: LG
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1564
           1       1.00      1.00      1.00       363

    accuracy                           1.00      1927
   macro avg       1.00      1.00      1.00      1927
weighted avg       1.00      1.00      1.00      1927

Accuracy: 1.0
Precision: 1.0
Recall: 1.0
F1 score: 1.0
```

```
Testing evaluation metrics: KNN
              precision    recall  f1-score   support

           0       1.00      1.00      1.00       400
           1       1.00      0.99      0.99        82

    accuracy                           1.00       482
   macro avg       1.00      0.99      1.00       482
weighted avg       1.00      1.00      1.00       482

Accuracy: 0.9979253112033195
Precision: 0.9979253112033195
Recall: 0.9979253112033195
F1 score: 0.9979253112033195
Training evaluation metrics: KNN
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1564
           1       1.00      1.00      1.00       363

    accuracy                           1.00      1927
   macro avg       1.00      1.00      1.00      1927
weighted avg       1.00      1.00      1.00      1927

Accuracy: 1.0
Precision: 1.0
Recall: 1.0
F1 score: 1.0
```

```
Testing evaluation metrics:DT
              precision    recall  f1-score   support

           0       0.96      0.96      0.96       400
           1       0.81      0.80      0.81        82

    accuracy                           0.94       482
   macro avg       0.89      0.88      0.89       482
weighted avg       0.94      0.94      0.94       482

Accuracy: 0.9356846473029046
Precision: 0.9356846473029046
Recall: 0.9356846473029046
F1 score: 0.9356846473029046
Training evaluation metrics: DT
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1564
           1       1.00      1.00      1.00       363

    accuracy                           1.00      1927
   macro avg       1.00      1.00      1.00      1927
weighted avg       1.00      1.00      1.00      1927

Accuracy: 1.0
Precision: 1.0
Recall: 1.0
F1 score: 1.0
```

```
Testing evaluation metrics: NB
              precision    recall  f1-score   support

           0       0.93      0.82      0.87       400
           1       0.45      0.72      0.55        82

    accuracy                           0.80       482
   macro avg       0.69      0.77      0.71       482
weighted avg       0.85      0.80      0.82       482

Accuracy: 0.8029045643153527
Precision: 0.8029045643153527
Recall: 0.8029045643153527
F1 score: 0.8029045643153527
Training evaluation metrics: NB
              precision    recall  f1-score   support

           0       0.92      0.82      0.87      1564
           1       0.47      0.68      0.56       363

    accuracy                           0.80      1927
   macro avg       0.70      0.75      0.71      1927
weighted avg       0.83      0.80      0.81      1927

Accuracy: 0.796574987026466
Precision: 0.796574987026466
Recall: 0.796574987026466
F1 score: 0.796574987026466
```

```
RF: Testing evaluation metrics: RF
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       400
           1       1.00      0.89      0.94        82

    accuracy                           0.98       482
   macro avg       0.99      0.95      0.97       482
weighted avg       0.98      0.98      0.98       482

Accuracy: 0.9813278008298755
Precision: 0.9813278008298755
Recall: 0.9813278008298755
F1 score: 0.9813278008298755
RF: Training evaluation metrics:RF
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1564
           1       1.00      1.00      1.00       363

    accuracy                           1.00      1927
   macro avg       1.00      1.00      1.00      1927
weighted avg       1.00      1.00      1.00      1927

Accuracy: 1.0
Precision: 1.0
Recall: 1.0
F1 score: 1.0
```

```
Ensemble: Testing evaluation metrics KNN + LG
              precision    recall  f1-score   support

           0       1.00      1.00      1.00       400
           1       1.00      0.99      0.99        82

    accuracy                           1.00       482
   macro avg       1.00      0.99      1.00       482
weighted avg       1.00      1.00      1.00       482

Accuracy: 0.9979253112033195
Precision: 0.9979253112033195
Recall: 0.9979253112033195
F1 score: 0.9979253112033195
Ensemble: Training evaluation metrics KNN + LG
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1564
           1       1.00      1.00      1.00       363

    accuracy                           1.00      1927
   macro avg       1.00      1.00      1.00      1927
weighted avg       1.00      1.00      1.00      1927

Accuracy: 1.0
Precision: 1.0
Recall: 1.0
F1 score: 1.0
```

```
Ensemble (LR & RF): Testing Evaluation Metrics
              precision    recall  f1-score   support

           0       0.97      1.00      0.99       400
           1       1.00      0.87      0.93        82

    accuracy                           0.98       482
   macro avg       0.99      0.93      0.96       482
weighted avg       0.98      0.98      0.98       482

Accuracy: 0.9771784232365145
Precision: 0.9771784232365145
Recall: 0.9771784232365145
F1 score: 0.9771784232365145

Ensemble (LR & RF): Training Evaluation Metrics
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1564
           1       1.00      1.00      1.00       363

    accuracy                           1.00      1927
   macro avg       1.00      1.00      1.00      1927
weighted avg       1.00      1.00      1.00      1927

Accuracy: 1.0
Precision: 1.0
Recall: 1.0
F1 score: 1.0
```

```
Ensemble (KNN & RF): Testing Evaluation Metrics
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       400
           1       1.00      0.89      0.94        82

    accuracy                           0.98       482
   macro avg       0.99      0.95      0.97       482
weighted avg       0.98      0.98      0.98       482

Accuracy: 0.9813278008298755
Precision: 0.9813278008298755
Recall: 0.9813278008298755
F1 score: 0.9813278008298755

Ensemble (KNN & RF): Training Evaluation Metrics
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1564
           1       1.00      1.00      1.00       363

    accuracy                           1.00      1927
   macro avg       1.00      1.00      1.00      1927
weighted avg       1.00      1.00      1.00      1927

Accuracy: 1.0
Precision: 1.0
Recall: 1.0
F1 score: 1.0
```

```
Ensemble Accuracy: 0.983402489626556
Testing evaluation metrics: KNN + LG + RF (Voting)
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       400
           1       0.99      0.91      0.95        82

    accuracy                           0.98       482
   macro avg       0.98      0.96      0.97       482
weighted avg       0.98      0.98      0.98       482

Accuracy: 0.983402489626556
Precision: 0.983402489626556
Recall: 0.983402489626556
F1 score: 0.983402489626556
Training evaluation metrics: KNN + LG + RF (Voting)
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      1564
           1       1.00      0.99      1.00       363

    accuracy                           1.00      1927
   macro avg       1.00      1.00      1.00      1927
weighted avg       1.00      1.00      1.00      1927

Accuracy: 0.9984431759211209
Precision: 0.9984431759211209
Recall: 0.9984431759211209
F1 score: 0.9984431759211209
```

Fig. 5 Evaluation matrix for testing and training phases of the classifier

The results of the second experiment on the MR-Small-BinaryClass dataset are presented in Tables 4–6. The computational complexity of the ensemble models is compared based on their inference time on the test set of the MR-Small-BinaryClass dataset, as shown in Table 6. Based on these results, four key observations were made.

Investigation 1: KNN outperforms other machine learning classifiers.
Evaluation: Tables 3 and 4 demonstrate that KNN with PCA outperforms other classifiers on the MR-Small-BinaryClass dataset. This superior performance is attributed to KNN's non-parametric nature, simplicity, adaptability to local data structures, and its ability to effectively utilize all available data during prediction, compared to the other classifiers.

Investigation 2: Gaussian Naïve Bayes performs the worst among all evaluated classifiers.
Evaluation: As shown in Table 3, Gaussian Naïve Bayes (NB) exhibits the lowest performance compared to other classifiers on both datasets. This may be due to the combination of PCA and NB, where PCA reduces dimensionality by retaining components with the highest variance, which may not necessarily be the most informative for classification. As a result, important discriminative features could be lost, negatively impacting classification accuracy.

Table 3. Performance Parameters of ML Classifiers on MR-Small-BinaryClass

Classifier	Precision	Recall	F1-score	Accuracy
LG	98%	98%	98%	98%
KNN	99%	99%	99%	99%
DT	93%	93%	93%	93%
NB	80%	80%	80%	80%
RF	98%	98%	98%	98%

Investigation 3: Ensemble of features from two or three ML classifiers proves effective for improving performance across most classifiers on the dataset.

Evaluation: As demonstrated in Tables 4 and 5, models utilizing an ensemble of features from the top two or three classifiers generally achieve higher accuracy compared to individual classifiers—except for KNN, which still

Int. J. Advanced Networking and Applications
Volume: 17, Issue: 03, Pages : 6970-6981 (2025) ISSN: 0975-0290
https://doi.org/10.35444/IJANA.2025.17310

performs best individually. This improvement can be attributed to the ensemble model's ability to combine the strengths of multiple classifiers by concatenating their most informative features. These diverse feature representations capture complementary information, thereby enhancing classification performance. However, this approach shows limited effectiveness on smaller datasets due to insufficient training samples, which restrict the ensemble model's capacity to learn from a more complex feature space.

Investigation 4: Ensemble of three ML classifiers results in the highest inference time, while ensembles of two classifiers offer faster inference.

Evaluation: As presented in Table 6, ensembles involving three ML classifiers incur the highest inference time on the test set. This is primarily due to the increased computational complexity, as each classifier operates independently, resulting in cumulative processing overhead. In contrast, ensembles of two classifiers demonstrate shorter inference times. This relative efficiency can be attributed to fewer model evaluations and the potential for optimized parallel execution or more streamlined aggregation mechanisms, making them more suitable for time-sensitive applications.

Figures 6 and 7 demonstrate that the proposed model achieves state-of-the-art performance, delivering the highest classification accuracy compared to all other evaluated methods.

Table 4. Accuracy of Ensemble of Two ML Models for MR-Small-BinaryClass

Classifier	Precision	Recall	F1-score	Accuracy
LR+KNN	99%	99%	99%	99%
RF+LR	97%	97%	97%	97%
KNN+RF	98%	98%	98%	98%

Table 5. Accuracy of Ensemble of Three ML Models for MR-Small-BinaryClass

Classifier	Accuracy
LG+KNN+RF (Voting classifier)	**98**
LG+KNN+RF (Bagging classifier)	**96**
LG+KNN+RF (Stacking classifier)	95

Table 6. Computational Complexity of Ensemble of ML Classifiers on MR-Small-BinaryClass

Classifier	Total Time
KNN + LG	38s
LG + RF	43s
KNN + RF	28s
LG + KNN + RF (Voting)	60s
LG + KNN + RF (Bagging)	60s

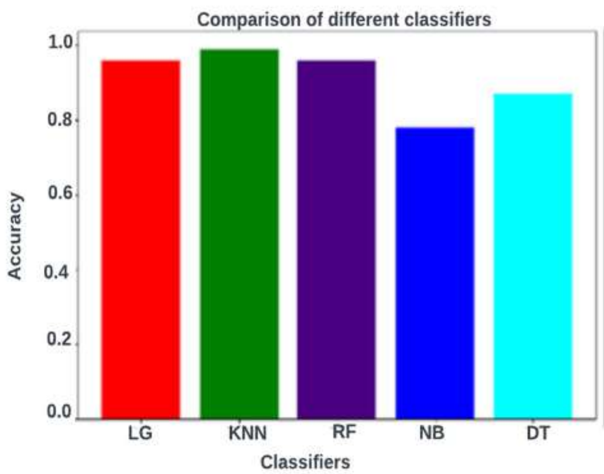


Fig. 6 Comparative accuracy of the ML classifiers

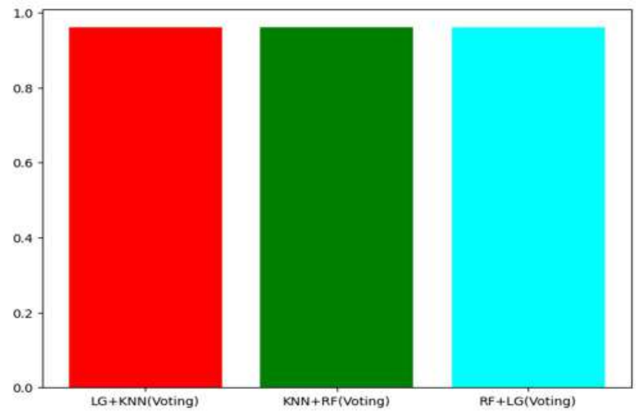


Fig. 7 Comparative accuracy by ensemble best two classifiers

5. Challenges and Future Directions

Despite significant advancements in ML based brain tumor classification, several challenges persist. Key issues include data imbalance, limited interpretability of complex models, and difficulties in generalizing across diverse patient populations. Addressing these challenges requires interdisciplinary collaboration and the adoption of innovative methodologies. In particular, integrating multimodal imaging data and developing explainable artificial intelligence (AI) frameworks are essential steps toward building more robust, transparent, and clinically applicable diagnostic systems.

6. Conclusions

This study underscores the efficacy of various machine learning classifiers in the identification and classification of brain tumors from MRI images. Among the evaluated models, KNN demonstrated superior performance, attributed to its non-parametric nature and adaptability to local data structures. In contrast, NB exhibited the lowest performance, likely due to its assumptions of feature independence and limitations when paired with dimensionality reduction techniques like PCA. Ensemble methods showed promising results, by effectively combining the strengths of multiple classifiers to enhance predictive accuracy. Overall, the results indicate that integrating machine learning approaches can substantially improve diagnostic precision and computational efficiency, thereby supporting clinicians in making timely and informed treatment decisions for brain tumor patients.

References

[1] S. Gajula, and V. Rajesh, “An MRI brain tumour detection using logistic regression-based machine learning model,” *Int J Syst Assur Eng Manag*, 15, 124–134, 2024. https://doi.org/10.1007/s13198-022-01680-8.

[2] H. Nankani, S. Gupta, S. Singh, and S. S. Ramesh, “Detection analysis of various types of cancer by logistic regression using machine learning,” *International Journal of Engineering and Advanced Technology (IJEAT)*, 9(1), 99-104, 2019.DOI:10.35940/ijeat.a1055.109119.

[3] J. Amin, M. Sharif, M. Raza, M. Yasmin, “Detection of brain tumor based on features fusion and machine learning.” *J Ambient Intell Human Comput*, 15, 983–999. 2024.DOI: 10.1007/s12652-018-1092-9.

[4] B. Jena, G. K. Nayak, and S. Saxena, “An empirical study of different machine learning techniques for brain tumor classification and subsequent segmentation using hybrid texture feature”, *Machine Vision and Applications*, 33, 6, 2022.https://doi.org/10.1007/s00138-021-01262-x.

[5] P. Nancy, G. Murugesan, A. S. Zamani, K. Kaliyaperumal, M. Jawarneh, S. K. Shukla, S. Ray, and A. Raghuvanshi, “Detection of brain tumour using machine learning based framework by classifying MRI images,” “International Journal of Nanotechnology”, 20, 880-896, 2023. DOI: 10.1504/IJNT.2023.134040.

[6] F. M. Refaat, M. M. Gouda, M. Omar, “Detection and Classification of Brain Tumor Using Machine Learning Algorithms”, *Biomed Pharmacol J*,;15, 4, 2381-2397, 2022.

[7] S. Saeed, A. Abdullah, N. Z. Jhanjhi, M. Naqvi, and A. Nayyar, “New techniques for efficiently k-NN algorithm for brain tumor detection,” *Multimed Tools Appl*, 81, 18595–18616, 2022. https://doi.org/10.1007/s11042-022-12271-x.

[8] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, and M. O. Alassafi, “Brain tumor classification based on fine-tuned models and the ensemble method,” *Computers, Materials and Continua*, 67(3), 3967-3982, 2021.https://doi.org/10.32604/cmc.2021.014158.

[9] P. Kaur, G. Singh, P. Kaur, “Classification and validation of MRI brain tumor using optimised machine learning approach”, ICDSMLA 2019. Lecture Notes in Electrical Engineering, vol 601. Springer, Singapore. DOI: 10.1007/978-981-15-1420-3_19.

[10] K. Ganesh, and R. Swarnalatha, “Improved brain tumor segmentation and diagnosis using an SVM-based classifier,” *Medical Imaging and Computer-Aided Diagnosis*, MICAD, Lecture Notes in Electrical Engineering, Springer, Singapore, 2020, DOI: 10.1007/978-981-15-5199-4_5.

[11] G. Çınarer, B. G. Emiroğlu, R. S. Arslan, and A. H. Yurttakal, “Brain tumor classification using deep neural network,” *Advances in Science, Technology*

and Engineering Systems Journal, 5(5), 765-769, 2020.DOI: 10.25046/aj050593.

[12] R. Kaur, A. Doegar, G. K. Upadhyaya, “An ensemble learning approach for brain tumor classification using MRI”, Soft Computing: Theories and Applications. Advances in Intelligent Systems and Computing, vol 1380. Springer, Singapore, 2022, DOI: https://doi.org/10.1007/978-981-16-1740-953.

[13] B. Jefferson, and R. S. Shanmugasundaram, “Brain tumor classification in 3D-MRI using features from radiomics and 3D-CNN combined with KNN classifier,” *International Journal of Electrical Engineering and Technology (IJEET)*, 12(2), 185-198, 2021. DOI: 10.34218/IJEET.12.2.2021.017.

[14] M. M. Nayak, S. D. Kengeri Anjanappa, “An efficient hybrid classifier for MRI brain images classification using machine learning based Naive Bayes algorithm,” *SN COMPUT. SCI.*, 4, 223, 2023.DOI: 10.1007/s42979-022-01614-y.

[15] A. K. Soni, A. Boobalan, E. Rajesh, V. Janakiraman, and V. Kesarwani, "Hybrid classification framework based on CNN and random forest method to predict brain tumor", IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT), Greater Noida, India, 2024.doi: 10.1109/IC2PCT60090.2024.10486772.

[16] D. Das, L. B. Mahanta, S. Ahmed, B. K. Baishya, and I. Haque, “Automated classification of childhood brain tumours based on texture feature,” *Songklanakarin Journal of Science and Technology*, 41(5), 1014-1020, 2019. DOI: 10.14456/sjst-psu.2019.128

[17] Y. Reddy, I. Adimoolam, and A. Sudha, “Comparison of accuracy on K-nearest neighbor and random forest algorithm based brain tumor detection in MRIhttps://doi.org/10.1063/5.0173997 images,” *AIP Conf. Proc.*, 14 November 2023; 2822 (1).

[18] F. Fernando, Brain Tumor MRI images 17 classes. https://www.kaggle.com/datasets/fernando2rad/brai n-tumor-mri-images -17-classes (accessed on 1 November 2023).

[19] S. Singh and V. Saxena. "Classification and segmentation of MRI images of brain tumors using deep learning and hybrid approach. International Journal of Electrical and Computer Engineering Systems. 15, 2,163-172, 2024, doi:https:// doi.org/10.32985/ijeces.15.2.5.

[20] S. Singh, and V. Saxena, "A Fine Tuned Pre-trained Model for Classification of Brain Tumor using Magnetic Resonance Imaging", GRENZE International Journal of Engineering and Technology, 10, 1,463-472, 2024. GrenzeID:01.GIJET.10.1.70_1.

[21] V. Saxena, S. Singh, KV, Singh. “MRI brain tumor prediction using Azure Streamlit framework and analysis of CNN activation functions”. Indian Journal of Science and Technology. 16(37):3129-3138. https://doi.org/10.17485/IJST/v16i37.427.

[22] S. Singh and V. Saxena, "Interpolation method for identification of brain tumor from magnetic resonance images", International Journal of Engineering and Manufacturing (IJEM), Vol.13, No.2, pp. 40-51, 2023. DOI:10.5815/ijem.2023.02.05.

[23] S. Singh, and V. Saxena, "A comparative study of segmentation techniques for identification of brain tumor from magnetic resonance images", BioGecko, A Journal for New Zealand Heptalogy, Vol. 12(2), 1800-1812, 2023.

[24] M. Thayumanavan, and A. Ramasamy, “An efficient approach for brain tumor detection and segmentation in MR brain images using random forest classifier,” *Concurrent Engineering*,29(3), 266-274, 2021.