

Doi:10.32604/dedt.2025.058943

ARTICLE





# Advancing Brain Tumor Classification: Evaluating the Efficacy of Machine Learning Models Using Magnetic Resonance Imaging

Khalid Jamil<sup>1</sup>, Wahab Khan<sup>1</sup>, Bilal Khan<sup>2</sup> and Sarwar Shah Khan<sup>2,\*</sup>

<sup>1</sup>Department of Computer Science, City University of Science and Information Technology, Peshawar, 25000, Pakistan
 <sup>2</sup>Department of Computer Science, University of Engineering and Technology, Mardan, 23200, Pakistan
 \*Corresponding Author: Sarwar Shah Khan. Email: sarwarshah@uetmardan.edu.pk

Received: 24 September 2024; Accepted: 22 January 2025; Published: 28 February 2025

**ABSTRACT:** Brain tumors are one of the deadliest cancers, partly because they're often difficult to detect early or with precision. Standard Magnetic Resonance Imaging (MRI) imaging, though essential, has limitations, it can miss subtle or early-stage tumors, which delays diagnosis and affects patient outcomes. This study aims to tackle these challenges by exploring how machine learning (ML) can improve the accuracy of brain tumor identification from MRI scans. Motivated by the potential for artificial intillegence (AI) to boost diagnostic accuracy where traditional methods fall short, we tested several ML models, with a focus on the K-Nearest Neighbors (KNN) algorithm. Our results showed that KNN achieved impressive accuracy, reaching 94.09%, and outperformed other models in key metrics like the F1-Score and Matthews Correlation Coefficient (MCC). These findings underline the value of selecting the right AI model and suggest that KNN may be particularly useful in pinpointing tumors in MRI scans. This research highlights how AI-driven methods, particularly those that can incorporate richer imaging data, may offer a promising path to more reliable and early detection of brain tumors.

**KEYWORDS:** Magnetic resonance imaging; classification; kaggle dataset; machine learning models; K-Nearest Neighbor

# **1** Introduction

The global mortality rate is alarming, and it is rising due to various factors. Brain tumors are one of the major contributors to this alarming rate. Cancers of the nervous system, including brain tumors, rank as the 10th leading cause of death worldwide. According to estimates from the World Health Organization (WHO), about 400,000 people are currently affected by brain tumors globally. Moreover, over the past few years, brain tumors have been responsible for about 120,000 deaths [1].

Brain tumors occur from the abnormal growth of cells in the brain. There is a variation in gender and age, and in terms of prevalence, there is a significant disparity [2,3]. Generally, brain tumors are classified into benign and malignant [4]. Benign tumors develop slowly and are not aggressive in their progression while malignant tumors develop rapidly, invading the adjacent tissues [5].

Understanding the etiology of brain tumors requires detailed molecular and genetic factors contributing to brain tumor development. This opens up the critical question of whether specific gene dysfunctions form brain tumors. The reply to this question forms part of the justification for pursuing further research in this topic. Future studies intend to look further into genetic mechanisms of growth in the tumor, and particularly how to identify the genes responsible for cancer and how they may function in the proliferation of abnormal



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

cells within the brain [5]. The study of brain tumors by genetic knowledge is supposed to lead towards a more complete understanding of their classification and diagnosis, especially in regard to imaging methods.

Detecting brain tumors is still a challenging task due to the complexity of tumor morphology, size, type, and location [6]. Early detection is essential and often involves methods such as biopsies and Magnetic Resonance Imaging (MRI). MRI has become a preferred tool for brain tumor detection, relying on variations in image intensity to distinguish tumors. Tumors usually appear brightly on MRI scans, though the intensity of tumor cells may be similar to normal areas of the brain, whereupon misclassification can often occur. Such errors, which can have serious results, influence decisions regarding potential surgical interventions or biopsies, thus underlining that differentiation between tumor and normal brain tissue, while accurate, is greatly needed [5].

Many risk factors have been associated with the formation of brain tumors, although their presence does not mean that a tumor will occur. The age factor plays a major role in many types of brain tumors and is more prevalent in some age groups. Some of the genetic predispositions, such as neurofibromatosis or Li-Fraumeni syndrome, and family history of brain tumors may be at increased risk. Other risk factors include exposure to ionizing radiation, disorders of the immune system, and occupational hazards. It has been shown that there are gender differences in tumor prevalence. Most individuals with risk factors, however do not have brain tumors, showing how complex the factors must be interplaying into tumor genesis. These risk factors need continuous research studies to elucidate their interaction, hence guiding efforts on mitigating risks involved in brain tumors. Consulting healthcare professionals for personal advice on risk factors remains advisable.

One of the major focus areas of the recent studies is the subtypes of brain tumors including glioma, meningioma, pituitary, and non-tumor. For furthering the work in this field, the current study carries out an elaborate investigation of diverse machine learning (ML) models for classifying brain tumors. These are Naïve Bayes (NB), K-Nearest Neighbor (KNN), J-48 Decision Tree (J48), Random Forest (RF), Ripple-Down Rule Learner (RIDOR), Decision Table (DT), Credal Decision Trees (CDT), Random Tree (RT), and Hoeffding Tree (HT). F1-Score, recall, precision, accuracy, and MCC are used as performance metrics to evaluate and rank the models, bringing out the strengths and weaknesses of each.

This study contributed to the field of brain tumor classification by conducting a comprehensive empirical investigation of multiple ML models using MRI data. Among the models examined, the KNN algorithm emerged as a standout performer, demonstrating reliable diagnostic accuracy and minimizing misclassifications. The research also highlights the challenges some models face, particularly in capturing complex decision boundaries, and provides valuable insights into their limitations. A key contribution of this study lies in its novelty, achieved by tailoring existing ML models to address domain-specific challenges such as imbalanced datasets and noisy samples. The integration of unique preprocessing techniques and domain-specific features further enhances the performance and applicability of these models. Additionally, the study conducts a systematic and reproducible comparative analysis of state-of-the-art methods under consistent experimental setups, emphasizing superior performance metrics. These contributions provide practical insights for researchers and practitioners in medical image analysis, supporting informed model selection and paving the way for improved brain tumor diagnosis accuracy and better clinical outcomes.

In the subject of brain tumor analysis, the study makes several cutting-edge advances. To begin with, it uses a sizable dataset of MRI scans to do a thorough empirical examination of many ML models, offering insightful information about the dependability and performance of the models. To further improve the efficiency of the classification process, the research presents a unique method that combines feature extraction with model training and testing using Orange and Weka. In addition, the study's incorporation of genetic insights provides a more comprehensive understanding of the subject by acknowledging the intricate interactions between genetic and imaging elements in brain tumor identification. Lastly, although the paper's

main focus is model analysis, it also highlights the significance of potential future research avenues, such as adding extra factors like specificity and sensitivity for a more thorough assessment. All things considered, these features add to the study paper's uniqueness by expanding our knowledge of how to classify brain tumors and suggesting creative approaches for further studies.

Combining new imaging techniques with AI has exciting potential for brain tumor detection and treatment. One approach, direct myelin imaging, could be especially valuable in identifying subtle changes in brain tissue that might not show up on traditional MRI scans [7]. This method focuses on the brain's white matter integrity, which could reveal early-stage or spreading tumors before they become visible through standard imaging. By integrating direct myelin imaging with AI models, we could improve the accuracy of tumor detection, helping doctors catch these issues sooner and offering a fuller picture of how tumors affect the brain. This could lead to earlier, more targeted interventions for patients.

The outline of this paper includes: Section 2 outlines a review of related literature. Section 3 sets out the research design and methodology, while Section 4 discusses the results of the study. Finally, Section 5 concludes this study.

#### 2 Literature Study

In many different methods, numerous studies have made observations on the classification and detection of brain tumors, often using various datasets and methodologies. The research projects also use a variety of sources, from data found within healthcare institutions to open datasets available online. Following are some of the more relevant contributions in this domain:

A research in [1] utilized the BRATS 2015 and BRATS 2017 datasets along with six ML techniques, including J48, Linear Regression (LR), K-Nearest Neighbor (KNN), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and ensemble methods, that combine features created by the human brain with deep learning. The performance was evaluated with accuracy, sensitivity, specificity, Dice Similarity Coefficient (DSC), and the Jaccard coefficient.

Amin et al. [2] used statistical and ML-based techniques, SVM, J48, KNN, RF, and classifiers as an ensemble approach for their evaluation. Various performances like Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), Structural Similarity Index Measure (SSIM), specificity, sensitivity, Area Under Curve AUC, DS and accuracy were used. All of the above approaches revealed Convolutional Neural Networks (CNN) as the most viable technique with an accuracy percentage of 91%. Hemanth et al. used a combination of CNN with other techniques, namely Genetic Algorithm (GA), SVM and Conditional Random Field (CRF), based on University of California Irvine (UCI) data. CNN gained an accuracy of 91%. A paper [6] applied CNN and observed an improved accuracy of 96.08%. Additionally, Irsheidat et al. [8] proposed a concept called Artificial Convolutional Neural Network and made an achievement with the peak accuracy of 96.7%.

Raju et al. [5] proposed the application of CNN in brain tumor detection from the BRATS dataset, with two classes, namely one with the cases containing tumors and the other with the cases without tumors. They concluded that CNN outperformed the models, which included SVM, KNN, Multilayer Perceptron (MLP), LR, NB, and RF, at a 97.87% accuracy rate. Similarly, in [6], MRI images were applied to 153 volunteers by taking 73 tumor-affected images and 80 normal images. Then deep neural networks were utilized for reporting an accuracy of 91%.

In [9], a K-means clustering approach was proposed to transform grayscale images to color-space representations for detecting brain tumors. Another recent work in [8] emphasized the use of MRI for tumor detection and addressed the difficulties arising from various morphologies of tumors. Abbas et al. [3]

developed an automated method for detecting tumors using a 3D brain tumor dataset [10] based on K-means clustering. The evaluation metrics included AUC, Dice Score (DS), precision, and Jaccard Score (JS), with the study reporting an AUC of 89% and a DS of 0.95.

In other study [2], the author developed a methodology with a combined statistical and machine learning-based approach along with SVM, J48, KNN, RF, and ensemble classifiers. All the approaches were checked through various PSNR, MSE, SSIM, specificity, sensitivity, AUC, DS, and accuracy metrics. The model accuracy was established by CNN to be as high as 91% better than others. Likewise, Hemanth et al. have proposed several machine learning approaches involving CNN, GA, SVM, and CRF together, and also have CNN reported an accuracy of 91%. Another implementation of CNN achieved an improved accuracy of 96.08% [6]. Recently, Irsheidat et al. proposed an Artificial Convolution Neural Network (ACNN) [8], that achieves 96.7%.

Using Deep Learning (DL) and ML techniques, Saeedi et al. [11] offer MRI-based brain tumor identification methods. They introduce a unique 2D CNN and compare it with conventional ML approaches using a dataset of 3264 MRI brain images. Their CNN outperforms a convolutional auto-encoder network and shows notable distinctions from conventional techniques, achieving high accuracy (96.47%). Khan et al. [12] address the manual MRI image processing by computer-assisted diagnosis, and they provide two models for binary and multiclass tumor classification. They also describe an accurate approach for detecting brain tumors using deep CNNs. Khan et al. [13] provide an intelligent model that uses deep learning to identify brain tumors. They suggest a hierarchical approach to tumor classification and achieve a 92.13% accuracy rate and a 7.87% miss rate. Their approach, which makes use of CNNs, provides substantial clinical support, facilitating prompt and precise tumor detection.

These works collectively demonstrate the diverse methods, datasets, and performance metrics used by the research in the field of brain tumor detection. The previous studies have varied approaches and performance measures that can be summarised nicely in Figs. 1 and 2. Table 1 offers a comparison between the suggested method and previous research, stressing the unique qualities, benefits, drawbacks, and quantitative evaluation of each strategy. The technical weight of the article is raised by explicitly stating the strengths and differences of the suggested technique compared to earlier research, successfully addressing the reviewer's remark.



Figure 1: Count of each technique discussed in literature



Figure 2: Count of each assessment measure

Technique	Distinguishing features	Pros	Cons	Quantitative analysis			
Proposed	Utilizes a	- Incorporates a	- Reliance on	Achieved an			
tech-	combination of	novel approach of	external tools may	accuracy of 94.09%			
nique	feature extraction	combining feature	introduce	in classifying brain tumors.			
	using Orange and	extraction and ML	complexity.				
	ML model	model					
	training/testing with Weka.	training/testing.					
[1]	Combination of	- Incorporates deep	- Limited datasets	Achieves accuracy,			
	handmade and deep	learning for feature	may restrict	sensitivity, and			
	learning features	extraction.	generalizability.	specificity metrics,			
	using datasets			with DSC and			
	BRATS 2015 and			Jacquard coefficient			
	BRATS 2017.			included.			
[2]	Utilizes CNN with	- Achieves high	- Performance is	Metrics include			
	statistical and ML	accuracy (91% and	highly dependent on	PSNR, MSE, SSIM,			
	methods on various	above) with CNN.	dataset selection.	specificity,			
	datasets including			sensitivity, AUC,			
	the UCI dataset.			DS, and accuracy.			
[5]	Utilizes CNN for	- Achieves high	- Limited to binary	Outperforms SVM,			
	brain tumor	accuracy (97.87%) in	classification of	KNN, MLP, LR, NB,			
	identification with a	brain tumor	brain tumors.	and RF classifiers.			
	binary BRATS	identification.					
	dataset.						

# Table 1: Comparison of proposed study with literature

(Continued)

		<b>D</b>	0			
Technique Distinguishing		Pros	Cons	Quantitative		
	features			analysis		
[7]	Proposes ACNN for early brain tumor detection using CNN and convolutional auto-encoder	- Achieves high training accuracies (96.47% and 95.63%) in early tumor detection.	- Limited discussion on generalization to diverse datasets.	Outperforms traditional ML approaches like KNN.		
	network.					
[8]	Introduces WBM-DLNets for MRI-based brain tumor detection using wrapper-based metaheuristic DL petworks	- Utilizes metaheuristic optimization techniques for improved classification.	- Complexity introduced by multiple optimization techniques.	Outperforms existing approaches in classification through feature selection and optimization.		
[3]	Utilizes K-means for automated brain tumor identification with a 3D brain tumor dataset.	- Achieves high AUC and DS in automated tumor identification.	- Limited discussion on scalability to larger datasets.	Achieved an AUC of 89% and a DS of 0.95.		

#### Table 1 (continued)

#### 3 Research Design and Procedure

This is an empirical study on machine learning based classifiers regarding the classification of brain tumor using MRI data. A process shown in Fig. 3 is imported MRI image into Orange tool. Further, features are extracted for each image and then to generate a structured Comma-Separated Values CSV file. This provides efficient analysis for ML models in the following way: These extracted features are saved into the CSV file to train and test models. Standard evaluation metrics are taken, with each of the steps being elaborated further in the study. The methodology followed multiple tools. The features were extracted using Orange 3-3.27.1-Miniconda-x86\_64. The ML models were trained and tested on Weka 3.9.5. The system has a Core i5 Central Processing Unit (CPU), 12 Gigabytes (GB) of Random Access Memory (RAM), and windows 10 running in a 64-bit mode.

# 3.1 Data Training and Testing

This research uses a dataset of 5707 MRI images, which are obtained from the Kaggle repository<sup>1</sup>. Table 2 summarizes the dataset, including the percentage of glioma, pituitary, meningioma, and non-tumor images. Every MRI image has been translated into 1006 features that are extracted. The partitioning of the dataset for training and evaluation purposes is done with the widely used 10-fold cross-validation technique.

<sup>&</sup>lt;sup>1</sup>https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset (accessed on 21 January 2025).



Figure 3: Methodology workflow

Table 2: Statics of brain tumour classes in the dataset

S. No.	MRI of	No. of MRIs
1	Glioma	1321 (23%)
2	Pituitary	1457 (26%)
3	Meningioma	1339 (23%)
4	Non-Tumor	1590 (28%)
5	Total MRIs	5707 (100%)

## 3.2 Feature Extraction

Since it establishes the foundation for efficient ML model training and testing, the feature extraction stage of the research paper is a crucial component of the technique. To extract features from MRI scans, the work uses the Orange tool. This tool allows features to be taken from each image and is structured into a CSV file. In all these after importing the images, the next important phase is image embeddings.

**Image embedding** is a way used to represent images as vectors of numbers in a high-dimensional area. These vectors are designed to capture significant features of the image, inclusive of shapes, colors, textures, and styles, in a manner that is beneficial for ML tasks. This procedure carries several steps which are:

- 1. CNNs: Image embeddings are typically extracted with the use of pre-skilled CNNs. CNNs are a category of deep neural networks specifically designed for reading visual records which include pictures.
- 2. Feature Extraction Layers: CNNs consist of a couple of layers, together with convolutional layers, pooling layers, and linked layers. When an image is passed through a CNN, the early layers carry out operations like convolution and pooling to extract low-level features which include edges, corners, and textures.
- 3. Hierarchical Representation: As the image progresses through the network, the layers grow to be deeper and more abstract, capturing higher-degree features and styles. For example, early layers would possibly discover easy shapes like traces and curves, whilst deeper layers might apprehend complicated objects.
- 4. Image Embeddings: Image embeddings are representations of images in a lower-dimensional area that capture their semantic means or visible characteristics. In the context of Orange, the image embeddings are usually extracted from one of the intermediate layers of the CNN, earlier than the related layers.

- 5. Dimensionality Reduction: The output of the selected intermediate layer is an excessive-dimensional vector representing the image within the embedding area. However, those vectors frequently contain redundant or inappropriate data. So, the dimensionality reduction of the use of Principal Component Analysis (PCA) is carried out to reduce the dimensionality of the embeddings at the same time as retaining their essential structure.
- 6. Similarity Measurement: Once the picture embeddings are received, they may be used to measure the similarity between images. Images with similar embeddings are possibly to be visually comparable or belong to the same magnificence.

By using pre-trained CNNs to extract image embeddings, Orange leverages the power of deep learning to know to routinely research meaningful representations of pix, which could then be used for diverse responsibilities that include type, clustering, or visualization.

**Data Table:** Once the image embeddings are extracted for all of the pix for the dataset, they're prepared right into a data table in Orange. Each row inside the data table represents an image, and the columns represent the features extracted from that image. The image embeddings are normally represented as numerical values inside the data table. Each column corresponds to a selected feature or dimension of the embedding space.

After the image features are extracted and organized right into a data table, it is then smooth to proceed with in additional analysis, the use of ML models. This methodology guarantees the effective portrayal of pertinent data from the MRI scan, hence streamlining the ensuing model analysis procedure. The research improves the accuracy and efficacy of feature extraction by utilizing Orange, a potent tool renowned for its adaptability and capacity to handle complicated data structures. Additionally, the structured CSV file streamlines the whole workflow and guarantees consistency throughout research stages by providing a common input format for the ML models that are trained and evaluated using the Weka platform. Overall, the study paper's feature extraction phase shows a methodical and reliable technique for getting the MRI data ready for further analysis, providing a strong basis for assessing ML models for classifying brain tumors.

#### 3.3 Performance Assessment

Performance evaluation is a very significant phase in any academic discipline. In this research work, multiple metrics are used for evaluating and comparing the performances of the models used. The performance metrics include accuracy [14], recall [15], precision [16], F1-Score [17], and Matthews Correlation Coefficient (MCC) [18]. The formulas for all of these metrics are described as follows:

$$Precision = \frac{TP}{TP + FP}$$
(1)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$
(2)

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall}$$
(3)

$$MCC = \frac{(TN * TP) - (FN * FP)}{\sqrt{(EP + TP)(FN + TP)(TN + EP)(TN + EN)}}$$
(4)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
(5)

where the rate of true negative values (TN), true positive values (TP), false positive values (FP), and false negative values (FN) are expressed by the equations.

#### 3.4 Summarization of Employed Models

The study uses various ML models to analyze and compare the classification techniques. These are NB [19], a probabilistic model based on Bayes' theorem, and KNN [20], which classifies instances by considering the nearest examples in the feature space. The J-48 [21] implements the C4.5 algorithms, and RF [22] uses a combination of decision trees for improved accuracy and robustness. RIDOR [23] is used for incremental knowledge acquisition, while the DT [24] model gives a simple tabular classification approach. CDTs [25] are also given to handle uncertainty by imprecise probabilities. This involves the RT [26] model variant of a single decision tree and the HT [27], designed for incremental learning on large datasets [28,29]. This collection of models provides an all-inclusive framework through which different ML methods can be tested.

#### **4** Results Analysis and Discussion

This section discusses the simulation results of the models implemented within this study. The models use the brain tumor MRI dataset from the Kaggle repository. The key evaluation metrics used to test the models are accuracy, recall, precision, F1-Score, and MCC. These can be obtained directly from the confusion matrix, which is a standard ML tool for analyzing the performance of classification. The confusion matrix summarizes predicted vs. actual labels, showing counts for True Positives, False Positives, True Negatives, and False Negatives. Fig. 4 presents the confusion matrices for all the models analyzed in this paper.

NB	a	b	с	d	KNN	a	b	с	d	HT	a	b	С	d	< classified as
COMPACT:	1007	239	18	57		1188	118	1	14	57655	1044	206	17	54	a = glioma
	316	732	108	183		58	1224	18	39		380	677	107	175	b = meningioma
	5	132	1396	57		18	22	1526	24		4	132	1397	57	c = notumor
	26	61	37	1333		5	18	2	1432		31	61	36	1329	d = pituitary
RF	a	b	с	d	RIDOR	a	b	с	d	CDT	a	b	с	d	< classified as
	1111	175	7	28		1024	233	7	57		1000	268	4	49	a = glioma
	72	1072	70	125		213	853	101	172		181	916	87	155	b = meningioma
	14	21	1529	26	1	3	24	1550	13		1	42	1531	16	c = notumor
	3	34	12	1408		27	70	29	1331		29	138	21	1269	d = pituitary
J48	a	b	с	d	DT	a	b	с	d	RT	a	b	с	d	< classified as
	1027	255	2	37		988	253	13	67		825	282	139	75	a = glioma
	245	886	77	131		339	681	107	212		268	719	172	180	b = meningioma
	0	46	1533	11		208	16	1359	7		66	61	1394	69	c = notumor
	37	132	16	1272		89	174	37	1157		79	173	100	1105	d = pituitary

Figure 4: Confusion matrix of each employed model

Fig. 5 presents, based on true positive rate (TPR) and false positive rate (FPR) values, how well the models work for the classification of brain tumors. Of all the models tested, KNN stands out with the best performance. It has a very high TPR at 0.941 and a very low FPR at 0.019. That is, it classifies the positive case-that is, the case of brain tumors and does not classify negative cases as positive or non-tumor instances. The success of KNN is attributed to its instance-based learning approach, where classification is based on the classes of neighboring data points. KNN classifies brain tumors, and it can find local patterns within the data very well because nearby pixels or areas contain important information for diagnosis purposes. Besides, the generalization capability of the model would depend on the nature of the dataset since relationships in different parts of the data are important. It is important to remember that how well KNN performs can

depend upon issues like the type of distance measure used, the number of neighbors looked at, and also the spread of the data. KNN performed very well with this dataset, but it would be great to look into other machine learning methods for the best choice for such tasks. RF and J48 models also performed very well, thus the comparison of different techniques is a must.



Figure 5: TPR and FPR score achieved via each employed model

Fig. 6 plots the performance evaluation of the models based on their accuracy, recall, and F1-Score. The maximum values of accuracy, recall, and F1-Score, respectively, produced by KNN are clearly evident (0.942, 0.941, and 0.941). This shows a potential capability of KNN in distinguishing both the positive and negative cases without substantial loss in accuracy. Nonetheless, there are potential problems such as overfitting along with the resultant computational cost of using KNN with the entire data set. In contrast, the models DT and RT have considerably low performance metrics, viz., precision: 0.737/0.703, recall: 0.733/0.708, F1-Score: 0.733/0.705, which point to a serious issue of finding a balance between false positives and false negatives. The accuracy and recall of the model NB are highly balanced at 0.778/0.783, while the F1-Score measures 0.779. RF and J48 models show consistency but with modest accuracy, recall, and F1-Score values (0.897/0.897/0.896 for RF, 0.825/0.827/0.826 for J48) and indicate the strength of model generalization capability to the unseen data. All such considerations related to dataset characteristics, feature selection process, and hyperparameter optimization that were involved in improving model performance are considered important aspects of this research. Third, the model that eventually emerges for selection should serve needs specific to an application environment, taking into consideration, for example, considerations regarding interpretability, computationally efficiency, and consequences for costs of misclassification.

Fig. 7 illustrates the MCC results for the models applied in brain tumor classification. The MCC is a single, all-encompassing measure of model performance, encompassing true positives, true negatives, false positives, and false negatives. Of the models, KNN has the highest MCC value at 0.922, which indicates it has a strong ability to classify brain tumors accurately. However, generalization ability of KNN and effects of dataset imbalance on MCC score need to be considered. Comparing MCC scores of DT and RT, both are found to be low at 0.649 and 0.611, respectively, which makes it difficult to describe complex decision boundaries for precise classification of brain tumors. MCC scores for the NB and J48 models were 0.711 and 0.771, respectively, meaning that both are able to handle true and false predictions reasonably well though extra fine-tuning would further improve their performance considerably. MCC analysis gives a well-balanced view of the strength and weakness of each model in terms of various aspects of classification performance. The results obtained should be interpreted in a cautious manner, as this depends on the quality of the dataset, complexity of the model, and implications of the misclassifications, especially in medical applications.



Figure 6: Models comparison through Precision, Recall, and F1-Score



Figure 7: Model comparison using MCC

Fig. 8 depicts the accuracy analysis of the ML models used in this work for brain tumor classification. There was a significant result for KNN, reaching an accuracy of 94.09%, meaning that this model could efficiently detect a brain tumor. This outcome is possibly because KNN is simple and has the concept of instance-based learning, which allows this model to recognize specific patterns in data. However, it must be said that KNN is very sensitive to the choice of distance measure, and even noise or outliers can play a role in its performance. The other model, RT, that has achieved 70.85% accuracy, is the least accurate among all and relates to the inherent unpredictability of such models and to overfitting the data. The DT model, even with an accuracy of 73.34%, also suffers from overfitting because it tends to generate complex trees that are not easy to generalize. However, the NB and J48 models have relatively high accuracy rates of 78.29% and 82.67%, respectively, but may not really capture the complex relationships between the data. Both RF and CDT models performed moderately, with an accuracy score of 89.72% for RF and 82.64% for CDT; this showed that they were effective in reducing overfitting in single DT models. From the analysis, accuracy clearly shows both the strength and weakness of each model; hence, there is the need to identify an ideal balance between complexity, robustness, and adaptability in tasks concerning the classification of brain tumors.



Figure 8: Accuracy analysis of each employed model for brain tumour classification

This study significantly expands the area of brain tumor classification since it addresses increasing global mortality rates due to brain tumors, which currently rank as the 10th leading cause of death. The research is distinct in its use of a large dataset of 5707 MRI scans that integrates genetic insights into recognizing complex relationships between imaging and genetic factors in brain tumor diagnosis. Empirical comparison of several ML models, namely NB, KNN, and RF, underlines the KNN method as one of the top performers and thus reliable for the accurate diagnosis of brain tumor. This work greatly contributes to improving the detection of brain tumors by providing precious guidance on the selection of the model, emphasizing the role of genetic factors, and encouraging future research with an aim toward improving the accuracy of diagnostic results based on machine learning.

The study offers an in-depth empirical research on brain tumor classification through ML algorithms. Apart from achieving improved accuracy of the detection of tumors, the research aims at increasing global mortality rates associated with brain tumors and highlights how much these contribute to the problem of public health. What distinguishes this work is the analysis of the various ML models, KNN, RT, etc. It assesses crucial performance metrics such as accuracy, recall, precision, FI-Score, and MCC, hence providing an overall assessment of the abilities of each model. The KNN algorithm has been discovered to be the most efficient for the correct classification of brain tumors, whereas the RT model cannot identify very subtle decision boundaries. In addition, the study considers genetic factors because of the complex interrelation between imaging data and genetics in the diagnosis of brain tumors. In conclusion, this research makes a significant contribution to the field by offering insights that will guide future advancements, stressing the importance of selecting appropriate models for brain tumor classification, and advocating for an integrated approach that considers both imaging and genetic data to enhance diagnostic precision. KNN stood out in this study because of its straightforward yet powerful approach to handling complex, high-dimensional data like MRI scans. By comparing new images directly to labeled examples, KNN can pick up on subtle differences that are essential for accurate medical diagnoses. The model's success was also driven by the high-quality, well-prepared dataset used in the research, which allowed KNN to effectively assess similarities between tumor and non-tumor cases. This combination of a robust dataset and KNN's ability to leverage all available information without getting too complicated made it particularly effective for classifying the intricate patterns found in medical imaging. As a result, KNN achieved impressive accuracy, precision, and F1-Scores in detecting brain tumors, showcasing its effectiveness in this critical area of healthcare. Besides all these advantages, some limitations of this study are: it primarily focuses on a fixed configuration for the KNN algorithm and does not explore the impact of varying the number of nearest neighbors, which could provide further insights into model optimization. Additionally, while the study demonstrates promising results for brain tumor classification using MRI data, it does not investigate the model's performance on

larger, more diverse datasets or across different types of brain tumors, which could affect the generalizability of the findings.

#### Threats to Validity

While the study is useful in presenting findings, concern does remain for generalizability. A dataset would be a very specific set based on what was placed in the Kaggle repository and may not represent all the variations of brain tumor patients experienced in clinical practice. The limited scope of the dataset may affect the performance of the model on unseen data, thus reducing the external validity of the results. Furthermore, though the chosen evaluation metrics are comprehensive, they may not capture all the nuances of medical diagnostic tasks, which may impact construct validity. The effect of parameter tuning and dataset preparation may cause researcher bias that would lower the internal validity of this study. This study has only looked into a particular subset of the machine learning models; more can be learned if new or hybrid models are to be probed. Hence, although the results bring significant contributions, these weaknesses emphasize the need for more validation on various datasets, other metrics for evaluation, and a larger variety of models to improve the strength and relevance of the derived conclusions.

#### 5 Conclusion

This research provides a detailed analysis of ML models for the classification of brain tumors. It addresses a critical need for accurate detection in a world where global mortality rates are increasing. The paper explored various models and their performance metrics to shed light on their strengths and weaknesses. Among the models considered, KNN was the best-performing model with good accuracy, recall, precision, F1-Score, and MCC, thus making it efficient for the classification of brain tumors due to its instance-based learning method. In contrast, the RT model performed poorly on all of the above metrics, indicating that complex decision boundaries were difficult to capture. This study stresses the need for careful selection of models and the complexities in the process of classifying brain tumors. These findings of the study provide insights for further improvement in detection processes regarding brain tumors, thus urging further research for better accuracy in diagnostic practices by enhancing machine learning techniques.

#### 6 Future Work

Future research could be extended based on the work done in this study to explore different avenues for furthering brain tumor classification. One of the ways is through the use of more diverse and comprehensive datasets, which would help improve model generalization by incorporating the complexities of different brain tumor types and symptoms. Furthermore, the use of multiple imaging modalities, for example, a combination of MRI with PET or functional MRI (fMRI), may offer a more complete view of tumor features. In addition, inclusion of explainable AI techniques could further improve the interpretability and transparency of the model's decision-making, an important aspect for its clinical adoption. Last but not least, collaboration with clinical staff can foster the inter-reinforcement of domain-specific expertise and result in creating hybrid models that bring out both computational power and clinic-usable knowledge. Improving these directions could hence offer a pathway to improvement for future research on more accurate, reliable, and practical brain tumor classification models to be applied in medical practices.

## 7 The Retrieval Framework

This research study provides in-depth empirical analysis of diverse ML models for classifying brain tumors, particularly focusing on enhancing the precision of diagnosis. It established that the KNN algorithm is the most effective model for the reliable detection of brain tumors with minimal errors in misclassifications.

The study uses established metrics for evaluating performance and delivers an overall view of performance, pointing out both strong and weak points of each model. It includes an in-depth examination of a large MRI dataset that is important for the understanding of the distribution of brain tumor types in clinical practice. It stresses a balance of model complexity, robustness, and interpretability while offering valuable insights into generalization models. The study also introduces MCC, an evaluation metric to evaluate the model in more detail, and openly discusses possible limitations of the study, such as dataset constraints and the effects of parameter tuning. Lastly, it lists future research directions, such as increased dataset size, incorporation of multi-modal imaging, and cooperation with clinicians for even more effective brain tumor classification methods.

Acknowledgement: We would like to acknowledge the support of our team members and the Kaggle repository for providing the data. Our sincere gratitude goes to the team leader for his invaluable guidance and feedback throughout this research.

Funding Statement: The authors received no specific funding for this study.

**Author Contributions:** The authors confirm their contribution to the paper as follows: study conception and design: Khalid Jamil, Wahab Khan; data collection: Bilal Khan, Sarwar Shah Khan; analysis and interpretation of results: Khalid Jamil, Wahab Khan, Bilal Khan; draft manuscript preparation: Khalid Jamil, Bilal Khan, Sarwar Shah Khan. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data supporting this study's findings are openly available in the Kaggle repository at https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset (accessed on 21 January 2025).

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

# References

- 1. Saba T, Sameh Mohamed A, El-Affendi M, Amin J, Sharif M. Brain tumor detection using fusion of hand crafted and deep learning features. Cogn Syst Res. 2020;59:221–30. doi:10.1016/j.cogsys.2019.09.007.
- 2. Amin J, Sharif M, Raza M, Saba T, Anjum MA. Brain tumor detection using statistical and machine learning method. Comput Meth Programs Biomed. 2019;177:69–79. doi:10.1016/j.cmpb.2019.05.015.
- Abbas K, Khan PW, Ahmed KT, Song WC. Automatic brain tumor detection in medical imaging using machine learning. In: 2019 International Conference on Information and Communication Technology Convergence (ICTC); 2019 Oct 16–18; Jeju Island, Republic of Korea: IEEE. p. 531–6. doi:10.1109/ICTC46691.2019.
- 4. Toğaçar M, Ergen B, Cömert Z. BrainMRNet: brain tumor detection using magnetic resonance images with a novel convolutional neural network model. Med Hypotheses. 2020;134:109531. doi:10.1016/j.mehy.2019.109531.
- 5. Raju P, Gayatri P, Nagaraju G. Brain tumour detection using convolutional neural network. Int J Recent Technol Eng. 2019;8(3):77–80.
- Siar M, Teshnehlab M. Brain tumor detection using deep neural network and machine learning algorithm. In: 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE); 2019 Oct 24–25; Mashhad, Iran: IEEE. p. 363–8. doi:10.1109/iccke48569.2019.8964846.
- Sedaghat S, Jang H, Athertya JS, Groezinger M, Corey-Bloom J, Du J. The signal intensity variation of multiple sclerosis (MS) lesions on magnetic resonance imaging (MRI) as a potential biomarker for patients' disability: a feasibility study. Front Neurosci. 2023;17:1145251. doi:10.3389/fnins.2023.1145251.
- Irsheidat S, Duwairi R. Brain tumor detection using artificial convolutional neural networks. In: 2020 11th International Conference on Information and Communication Systems (ICICS); 2020 Apr 7–9; Irbid, Jordan: IEEE. p. 197–203. doi:10.1109/icics49469.2020.239522.

- Lalitha B, Ramashri T. Brain tumor detection using adaptive K-Means clustering segmentation. Int J Adv Eng Res Dev. 2017;4(7):597–601.
- Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, et al. The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging. 2015;34(10):1993–2024. doi:10.1109/TMI. 2014.2377694.
- Saeedi S, Rezayi S, Keshavarz H, Niakan Kalhori RS. MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. BMC Med Inform Decis Mak. 2023;23(1):16. doi:10. 1186/s12911-023-02114-6.
- Khan MSI, Rahman A, Debnath T, Karim MR, Nasir MK, Band SS, et al. Accurate brain tumor detection using deep convolutional neural network. Comput Struct Biotechnol J. 2022;20:4733–45. doi:10.1016/j.csbj.2022.08.039.
- 13. Khan AH, Abbas S, Khan MA, Farooq U, Khan WA, Siddiqui SY, et al. Intelligent model for brain tumor identification using deep learning. Appl Comput Intell Soft Comput. 2022;2022:8104054. doi:10.1155/2022/8104054.
- Madhavi KR, Soujanya K, Inaganti R, Swamy KA, Yalavarthi S, Munagala M. Brain tumor classification and segmentation using transfer learning from MRI images. Int J Comput Inform Syst Indus Manage Appl. 2025;17:14. doi:10.70917/2025004.
- Baali M, Bourbia N, Messaoudi K, Bourennane EB. Efficient deep learning approach for brain tumor detection and segmentation based on advanced CNN and U-Net. Indones J Electr Eng Comput Sci. 2025;37(2):1365. doi:10. 11591/ijeecs.v37.i2.pp1365-1375.
- Belard A, Buchman T, Forsberg J, Potter BK, Dente CJ, Kirk A, et al. Precision diagnosis: a view of the clinical decision support systems (CDSS) landscape through the lens of critical care. J Clin Monit Comput. 2017;31(2):261–71. doi:10.1007/s10877-016-9849-1.
- 17. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over Fl score and accuracy in binary classification evaluation. BMC Genomics. 2020;21(1):6. doi:10.1186/s12864-019-6413-7.
- Chicco D, Tötsch N, Jurman G. The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. BioData Min. 2021;14(1):13. doi:10.1186/s13040-021-00244-z.
- Zaw HT, Maneerat N, Win KY. Brain tumor detection based on Naïve Bayes classification. In: 2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST); 2019 Jul 2–5; Luang Prabang, Laos: IEEE. p. 1–4. doi:10.1109/iceast.2019.8802562.
- Havaei M, Jodoin PM, Larochelle H. Efficient interactive brain tumor segmentation as within-brain kNN classification. In: 2014 22nd International Conference on Pattern Recognition; 2014 Aug 24–28; Stockholm, Sweden: IEEE. p. 556–61. doi:10.1109/ICPR.2014.106.
- Naik J, Patel S. Tumor detection and classification using decision tree in brain MRI. Int J Comput Sci Netw Secur. 2014;14(6):87–91.
- 22. Lefkovits L, Lefkovits S, Vaida MF. Brain tumor segmentation based on random forest. Mem Sci Sect Rom Acad. 2016;39(1):83–93.
- Papageorgiou EI, Spyridonos PP, Glotsos DT, Stylios CD, Ravazoula P, Nikiforidis GN, et al. Brain tumor characterization using the soft computing technique of fuzzy cognitive maps. Appl Soft Comput. 2008;8(1):820–8. doi:10.1016/j.asoc.2007.06.006.
- 24. Hall M, Frank E. Combining naive Bayes and decision. In: Proceedings of the Twenty-First International FLAIRS Conference; 2008. p. 318–9.
- Asiri AA, Khan B, Muhammad F, ur Rahman S, Alshamrani HA, Alshamrani KA, et al. Machine learningbased models for magnetic resonance imaging (MRI)-based brain tumor classification. Intell Autom Soft Comput. 2023;36(1):299–312. doi:10.32604/iasc.2023.032426.
- Yadav LN, Marve SR, Sadrani VH, Shende SR, Gedam NJ. Automated diagnosis of brain tumors from MRI scans using U-Net segmentation. In: Artificial intelligence revolutionizing cancer care. Boca Raton: CRC Press; 2024. p. 239–52. doi:10.1201/9781003571339-15.
- Kiranmayee BV, Rajinikanth TV, Nagini S. Effective analysis of brain tumor using hybrid data mining techniques. Int J Adv Res Comput Sci. 2017;8(7):286–93. doi:10.26483/ijarcs.v8i7.4237.

- 28. Loeber S. Brain MRI protocol and systematic approach to interpretation of brain tumors on MRI. Vet Clin North Am Small Anim Pract. 2025;55(1):11–21. doi:10.1016/j.cvsm.2024.07.003.
- 29. Muhammad Danyal M, Shah Khan S, Shah Khan R, Jan S, Rahman NU. Enhancing multi-modality medical imaging: a novel approach with Laplacian filter + discrete Fourier transform pre-processing and stationary wavelet transform fusion. J Intell Med Healthcare. 2024;2:35–53. doi:10.32604/jimh.2024.051340.