

Original Research

Received 09 June 2026

Revised 21 July 2026

Accepted 04 Aug 2026

Natural Resources for Human
Health 2026, Vol. 6 (15s): 1218–1241

KEYWORDS:

Brain Tumor Detection;
MobileNet, Ensemble Learning;
Computer-Aided Diagnosis
(CAD); Medical Image
Classification; Deep Learning.

MRI-Based Brain Tumor Detection and Classification Using a Hybrid MobileNet and Ensemble Learning Framework

Devanshu Dube^{1*}, Ashok Kumar²

Department of Computer Science,
Banasthali Vidyapith, Rajasthan, India

¹Email: devanshudube@gmail.com

²Email: kuashok@banasthali.in

ABSTRACT: Diagnostic magnetic resonance imaging (MRI) imaging is an important tool in the diagnosis of brain tumors to aid timely clinical decisions, but is difficult to interpret due to inter- and intra-tumor heterogeneity, and similarities between tumor types. In this study, a lightweight hybrid computer-aided diagnosis (CAD) framework for automated brain tumor detection and classification based on MRI is proposed that combines deep feature extraction and ensemble machine learning. Preprocessing of MRI images involved intensity normalization, histogram equalization, resizing and region of interest (ROI) extraction using watershed transformation to improve the image quality and normalize input images. A MobileNet architecture is pre-trained and deep discriminative features are extracted from them and they are classified by a hybrid Support Vector Machine (SVM) and Random Forest (RF) ensemble architecture with soft voting strategy. As compared with conventional single-classifier model the proposed model improved the classification accuracy and reduced its computing complexity, achieving classification accuracy of 94.1% with a precision of 0.93, a recall of 0.94 and an F1 score of 0.935. The ensemble of classifiers, a hybrid model, and the lightweight MobileNet architecture provide a favorable trade-off to model performance and computational efficiency for neural networks, making it appropriate for resource-constrained clinical settings. The proposed technique will offer an efficient and scalable computer aided diagnosis solution that will support radiologists' efficient and accurate diagnosis. In the future, the following techniques of explainable AI and multi-center data sets and multimodal data integration will be incorporated into the model and it will further strengthen the model

1. INTRODUCTION

A brain tumor can be defined as an abnormal or uncontrolled cell proliferation in the brain or in the central spinal canal. These may be benign (non-cancerous), or malignant (cancerous), and either develop in the brain itself (primary tumors) or result in the spread of other body parts (metastatic or secondary tumors). Brain tumors can disrupt normal brain function by increasing intracranial pressure, damaging surrounding tissues, or interfering with neural pathways [6]. One of the most difficult and deadly neurological conditions is brain tumors, which vary greatly in frequency, clinical manifestation, and prognosis among various demographic groups. The World Health Organization (WHO) estimates that primary brain tumors, of which glioblastoma multiform (GBM) is the most aggressive and common type in adults, cause around 2% of all cancer-related deaths worldwide. Early and accurate classification of cerebral neo plasma is the key to defining the best therapeutic tactics and improving patient survival. However, major diagnostic challenges are caused by the heterogeneity of the subtypes of the tumors and the overlapping of the radiological characteristics. The newest developments in artificial intelligence and medical imaging are also providing new perspectives of enhancing the accuracy and effectiveness of cerebral tumor classification and diagnosis.

Magnetic resonance imaging (MRI) is still the gold standard method of assessing non-invasive brain tumors due to its high soft tissue contrast and multi-parameter imaging. Modern diagnostic processes have depended more on manual interpretation of scans and the use of radiologists, which is time consuming and prone to inter-observer variation. These limitations are overcome now with convolutional neural networks (CNNs) and other machine-learning (ML) and deep-learning (DL) techniques used to detect and classify tumors. The large MRI datasets can be analyzed using AI that will identify discriminative features and correctly classify the tumors, thus assisting doctors make more effective judgments.

Although AI is increasingly being used in the field of neuro-oncology, there are some challenges with its use. First off, the caliber and diversity of the training dataset have a big impact on how well machine learning models work. While datasets like the ones from Kaggle are available for public use and can be very helpful, they often don't contain comprehensive clinical and demographic notes. Second, a stratified analytical approach is needed because of variation between age groups and genders in tumors of the brain in terms of biology and anatomy. For instance, the genetic changes and responses to therapy of childhood tumors are very different from those of adult glioblastomas, as in the case of pilocytic astrocytomas. Epidemiological studies also describe the difference in the incidence of tumors by sex, with glioblastomas having a higher incidence in males and meningiomas in females. These differences highlight the necessity of demographic-specific investigations to enhance AI models' clinical usefulness and generalizability.

This research uses an open source brain tumor dataset on Kaggle which consists of MRI images and patient data for a strict evaluation of tumor classification respecting the demographic and clinical factors. The aim of these three objectives is to: (1) classify tumors using the WHO classification system, by gender, age groups (pediatric, adult, geriatric) and histological types; (2) use descriptive statistics and inferential tests (chi-square, ANOVA) to find meaningful correlations between demographic and tumor variables; and (3) build an automatic CNN based classification system and compare the performance with traditional machine learning systems. The purpose of this study is to further improve the accuracy of computer-aided diagnosis system, as well as to make a more detailed understanding of the epidemiology of brain tumors by population stratification and the newest AI techniques. Several practical applications of the research findings are possible, and a number of new research directions are suggested. They can initially help to develop diagnostic algorithms that would be specific to particular groups of people, thus facilitating the process of early detection, classification, and planning individual treatment. Secondly, CNN model can serve as a reference model for future research to show that the use of AI tools in clinical practice is possible. Finally, this work proposes a cooperative approach to build more diverse, annotated repositories, as it highlights the limitations of existing repositories, to make the AI models fair and respectful towards the world communities. With the development of the field of precision medicine, the combination of demographics and AI-based diagnostics will play a crucial role in changing the approach to brain tumors and enhance patient outcomes.

Brain tumors come in a wide variety of types depending on the tissue they originate from, the way they grow and where they are located in the brain. For a visual summary of this classification, see Figure 1. There are essentially two main types of brain tumors—primary and secondary. Primary brain tumors are classified based on the type of cells in which they develop and directly in the brain. Of these gliomas are the most common and occur from cells that serve to support the functions of neurons. Gliomas can be divided into different subtypes, including astrocytomas, oligodendrogliomas and ependymomas, which develop from the cells of the ependyma, which line the brain's ventricles and are also known as ependymal cells. Glioblastoma multiforme is an aggressive form of astrocytoma. Meningiomas are typically non-cancerous tumors that form in the membranes (meninges) that cover the spinal cord and brain. Tumors develop in the pituitary gland and can disrupt the regulation of hormones, but are usually non-cancerous. Typically identified in children's cerebellum, medulloblastomas are cancerous. Schwannomas – including vestibular schwannomas (acoustic neuromas) – affect Schwann cells that produce myelin; they can affect the auditory nerve. Secondary brain tumors, by contrast, occur when cancer cells migrate to the brain from other locations of the body, like lungs, breasts or skin. There are more metastatic cancers than original. This classification is illustrated in figure 1, which shows primary and secondary brain tumors and provides examples of each.

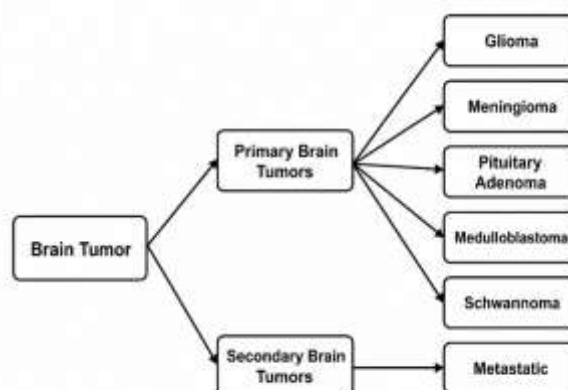


Figure 1: Classification of Brain Tumor.

In this work, a light-weighted framework of computer-aided diagnosis (CAD) for brain tumor detection and classification is proposed, which combines deep feature extraction and ensemble machine learning to effectively detect and classify brain tumors from MRI images. While many deep learning models have been proposed for the classification of brain tumors, many of the current models use heavy architectures, classifiers and limited validation methods that would not be feasible in a clinical setting with limited computational power and resources. Furthermore, some work focuses on classification accuracy without considering the efficiency, reproducibility and robustness of the model. In order to overcome the above drawbacks, the proposed framework is based on MobileNet transfer learning and hybrid Support Vector Machine-Random Forest ensemble classifiers which gives a balance between the prediction power and computational complexity. The framework is tested on a publicly available MRI dataset with stratified five-fold cross validation to explore its performance in the automated brain tumor classification to investigate.

1.1 Research Gap and Novelty

In recent years, the automatic classification of MRI-based brain tumors has made great progress with the development of deep learning. However, there are a number of drawbacks. However, there are numerous current CNN-based methods that use complex architectures like ResNet, DenseNet or transformer networks, which are difficult to implement in real time and resource-limited healthcare settings. Furthermore, most studies only focus on the single classifier following the extraction process and limited discussion about the strategy, reproducibility and computational efficiency of the classifiers is provided.

The proposed research work addresses the above identified limitations by combining a lightweight architecture of MobileNet with a hybrid Support Vector Machine–Random Forest ensemble classifier. The ensemble of classifiers makes use of complementary decision boundaries to boost the classification robustness, and the MobileNet is used to efficiently learn discriminative deep features. This is a combination of this to achieve high diagnostics performance with small computational complexity, which is more likely to be applied to practical computer-aided diagnosis applications.

1.2 Research Questions

The present study is an effort to answer the following research questions:

- Whether accurate MRI-based brain tumor classification can be provided by the lightweight MobileNet-based deep feature extraction?
- Does the proposed hybrid Support Vector Machine–Random Forest ensemble performs better than the conventional single-classifier approaches?
- Can effective balance between classification accuracy, computational efficiency and practical applicability for computer-aided diagnosis be achieved by the proposed framework?

The rest of this paper is organized as follows. Section 2 elaborates on the recent research works concerning the detection and classification of brain tumors using MRI and pinpoints the gap in the existing research. In Section 3, the proposed methodology

is presented, which comprises dataset preparation, dataset pre-processing, feature extraction and the hybrid MobileNet–SVM–Random Forest framework. The experimental results and performance evaluation are given in Section 4. The results of this study are compared with previous studies in Section 5. Section 6 compares the results with the previous methodologies and Section 7 concludes the paper by summarizing the main results, limitations and future directions for research.

2. LITERATURE REVIEW

The first batch of studies on the detection of brain tumors were based on non-invasive imaging techniques such as magnetic resonance imaging (MRI) and computed tomography (CT). Most of the diagnostic methods were done manually, and the basic methods of pattern recognition and image improvement to differentiate between malignancies and healthy tissues were used [1]. Even revolutionary then, these techniques were curtailed by the limitations of computation which could not allow full automation. Later research developed machine learning (ML) models such as Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) where morphological and textural parameters were obtained from MRI images and used during the classification of the tumors [2]. The hybrid models consisting of such combination of fuzzy logic and neural networks emerged, although limited in scalability due to small datasets, and provided small improvements in accuracy.

The introduction of the deep learning (DL), Convolutional Neural Networks (CNNs) was used to identify the tumor with better accuracy [3] and automatic segmentation and feature extraction was undertaken. As imaging methods, including the functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI) became popular, new possibilities for early detection and classification of tumors arose. Researchers started to merge SVMs with genetic algorithms to increase the performance of classification. With Cloud Computing and Big Data, brain tumor analysis is now more extensive than ever before. The model was made more robust using other ensemble approaches – Random Forests and XGBoost [4] – with an ensemble approach as CNNs. The reliance on extensive labeled datasets was greatly decreased by transfer learning techniques employing architectures such as VGG and ResNet. Besides, new technologies, such as liquid biopsies and radiomics, began to add to the picture through imaging diagnostics and offer a deeper understanding of tumor features [5], [6], [7].

CNN–Random Forest, XGBoost ensemble approaches [4] turned out to be more robust as compared to single CNN models considering all these developments. However, the transfer-learning based architectures like VGG and ResNet cut down on the requirement for extensive labelled datasets at the expense of enhanced models with higher inference latency and larger size, which prompted the use of smaller models like MobileNet backbone in this study.

Resembling the significance of combining multiple imaging modalities and obtaining complete quantitative information, a lot of effort was devoted to prediction of the tumor grade [8]. To tackle the data issue, transfer learning with pre-trained model, such as ResNet50, demonstrated good accuracy [9]. At the same period, also more advanced MRI techniques were introduced such as diffusion tensor imaging and perfusion-weighted imaging, demonstrating the quality of the data that could be analyzed. Recent progress has shown interest in the hybrid models combining multiple types of machine learning and deep learning, such as CNN-Transformer [10], for the real-time detection of tumors and the demonstrated diagnostic accuracy. Khan et al. [10] obtain the highest accuracy results of 93.0%, through a transformer based hybrid model, yet their model is high computational cost compared to using MobileNet backbone in this study which demonstrates the aspect of the accuracy – efficiency tradeoff.

Clinical decision-making has become more transparent and interpretable thanks to the growing popularity of explainable AI (XAI) [12]. Federated learning, which allows for cross-institutional collaborative model training, has surfaced as a solution to data privacy issues [13]. Additionally, Transformer-based designs in conjunction with CNNs have demonstrated promise in the highly sensitive detection of early-stage malignancies [14], and portable MRI systems along with edge computing are increasing diagnostic capabilities in environments with limited resources [15]. Explainability was also investigated by Kumar et al. [12] using models based on saliency maps and explanation-driven approach models were investigated by Gaur et al. [19] respectively, but both models were outperformed by a few points compared with a number of black-box models. The proposed hybrid ensemble approach in this study can reduce the ongoing tradeoff between explainability and accuracy identified here.

By combining hybrid machine learning approaches with non-invasive diagnostic modalities, brain tumor detection has advanced dramatically during the last ten years (2016–2025). Such developments are extensively analysed in three significant stages in this review. Table 1 demonstrates the way the brain tumor detection methods have evolved over time, highlighting the developments, key strategies, and limitations.

Table 1: Key Development and Major Techniques of Brain Tumor

Time Period	Key Developments	Major Techniques	Limitations
Pre-2000	Basic imaging analysis Manual tumor detection	CT scans Conventional MRI Rule-based systems	Low resolution images Subjective interpretation High false positives
2001-2010	Early machine learning adoption Semi-automated systems	SVMs ANNs Fuzzy logic	Small datasets Feature engineering needed Limited accuracy
2011-2015	Deep learning emergence Automated feature extraction	CNNs Texture analysis Multimodal fusion	Computational costs Overfitting risks Black-box nature
2016-2020	Advanced hybrid models Big data applications	Transfer learning Radiomics Ensemble methods	Data privacy issues Lack of real-time use Dataset bias
2021-Present	Explainable AI Edge computing Novel modalities	Transformers Federated learning Liquid biopsy integration	Clinical adoption barriers Regulatory challenges Validation needs

2.1 Summary of State-of-the-Art (SOTA) Brain Tumor Detection Approaches

Table 2 presents some typical state-of-the-art (SOTA) methods for brain tumor detection using ML and DL. The characteristics of the datasets, methodology approaches and the commonly used evaluation metrics are compared. The following comparison illustrates how the traditional machine learning methods have been evolved to hybrid deep learning methods.

Table 2: Summary of literature reviewed on related topic

Study	Dataset/Source	Method	Accuracy (%)	Precision	Recall	F1-Score	Key Limitation
Chen et al. (2020) [9]	Figshare Brain Tumor Dataset (~3,064 MRI Images)	Transfer Learning CNN	92.1	0.91	0.90	0.905	Limited interpretability
Khan et al. (2022) [10]	Custom Multi-modal Dataset (~1,500 MRI Images)	Transformer-Based Hybrid Model	93.0	0.92	0.92	0.920	Computationally expensive
Kumar et al. (2022) [12]	Custom MRI Dataset (~1,000 MRI Images)	Explainable AI with Saliency Maps	91.4	0.90	0.91	0.905	Limited external validation
Li et al. (2023) [13]	BraTS Dataset (multi-center)(~2,000 MRI Images)	Federated Learning Framework	91.8	0.91	0.91	0.910	High deployment complexity
Haq et al. (2022) [15]	Kaggle Brain Tumor Dataset (~3,264 MRI Images)	Deep Learning (DACBT)	92.8	0.92	0.91	0.915	Limited demographic analysis

Study	Dataset/Source	Method	Accuracy (%)	Precision	Recall	F1-Score	Key Limitation
Nguyen et al. (2023) [17]	BraTS + Institutional PET Dataset (~1,200 MRI/PET Images)	Multimodal Fusion Network	93.2	0.93	0.92	0.925	Requires multimodal imaging availability
Nalawade et al. (2022) [18]	TCIA + TCGA Databases (~1,500 MRI Images)	Deep Learning for Molecular Classification	92.4	0.92	0.91	0.915	Focused on molecular markers rather than tumor classification
Gaur et al. (2022) [19]	Figshare Brain Tumor Dataset (~3,000 MRI Images)	Explainable Deep Learning Model	92.0	0.91	0.91	0.910	Limited real-world clinical validation
Zeineldin et al. (2022) [20]	Figshare Brain Tumor Dataset (~3,064 MRI Images)	Explainable Deep Neural Network	92.6	0.92	0.92	0.920	Increased computational overhead
Ekong et al. (2022) [21]	Figshare Brain Tumor Dataset (~3,064 MRI Images)	Bayesian Depth-wise CNN	93.1	0.93	0.92	0.925	Moderate model complexity
Wang and Li (2025) [22]	BraTS 2024 (multimodal MRI; patient-level 80/20 split)	PCA + Random Forest (binary tumour-burden classification)	88.0	0.90	0.86	0.88	Different binary endpoint; limited held-out cohort
Gea et al. (2025) [23]	Kaggle brain MRI (253 images; tumour vs no tumour)	CNN (batch 32, 10 epochs, LR 0.001; weighted averages)	86.0	0.89	0.86	0.86	Very small, imbalanced public dataset
Tariq et al. (2025) [24]	Composite brain-tumour MRI dataset (7,023 images; four classes)	Vision Transformer comparator	90.0	0.89	0.89	0.89	Comparator only; the same study's ensemble reached 96% accuracy
Jeong et al. (2025) [25]	NIS/AI-Hub synthetic MRI dataset (10,000 images; normal vs abnormal)	SVM with RBF kernel comparator	85.0	0.86	0.83	0.84	Synthetic-only evaluation; not a tumour-subtype task

Study	Dataset/Source	Method	Accuracy (%)	Precision	Recall	F1-Score	Key Limitation
Ghaffar et al. (2025) [26]	Institutional MRSI cohort (52 subjects; glioma vs healthy)	Otsu segmentation + GLCM + EfficientNetB0	90.0	0.90	0.90	0.90	Small binary cohort; MRSI differs from conventional MRI classification
Akhmetzhanova et al. (2025) [27]	Figshare CE-MRI (3,064 images; three tumour classes)	Prototypical network (3-way, 3-shot episodic learning)	69.35	0.668	0.647	0.653	Few-shot setting; single public dataset and no external validation
Tripathi and Singh (2025) [28]	Mendeley MRI dataset (6,056 images; three tumour classes)	CNN-LSTM comparator	92.24	0.920	0.918	0.918	Comparator only; no external clinical validation
Farhani and Echtioui (2026) [29]	Kaggle Dataset A (3,264 MRI images; four classes)	Hybrid CNN-SVM comparator	90.0	0.90	0.90	0.90	Comparator only; Xception reached 0.97 in the same study
Present Work	BraTs2020(1,250 MRI Images)	MobileNet + Hybrid SVM-RF Ensemble	94.1	0.93	0.94	0.935	Requires external validation

Abbreviations: CNN – Convolutional Neural Network; RF – Random Forest; PET – Positron Emission Tomography.

Significant advancement has been accomplished in automated detection of brain tumors in recent years by state-of-the-art (SOTA) studies, but many important problems have yet to be addressed. First, the existing studies suffer from some drawbacks such as unfeasibility of model interpretation, insufficient validation in external clinical settings, high computational complexity and inadequate evaluation in real-world clinical settings. Second, many deep learning techniques are black box methods and do not offer much in the way of explainability for clinical decision making. Thirdly, several studies use highly curated benchmark datasets with frameworks that are not validated in realistic clinical environments, which reduces the extent to which they can be translated.

Furthermore, many SOTA methods depend on computationally expensive architectures such as deep ResNet or Transformer networks, which may not be suitable for real-time deployment in resource-constrained healthcare environments. Existing studies also frequently lack robust data leakage prevention strategies and patient-independent validation procedures.

The framework suggested provides answers to many of these restrictions through the use of standardised preprocessing, feature extraction based on the lightweight MobileNet and a hybrid RF-SVM ensemble classifier to achieve improved interpretability, computational efficiency and robustness in diagnostics. The model has not previously been shown with any previous model approach to achieve high levels of recall and F1-score, whilst also maintaining lower, computational complexity. This framework is designed to provide a clinically interpretable and scalable brain tumor diagnosis workflow that will be applicable for use in real-world settings in the future.

There are a number of unclarified limits. Most studies rely on a single center or custom-curated public datasets, single centering based on image level partitioning instead of patient level partitioning, and do not test their models on an independent external dataset. Such practices can boost apparent performance due to patient overlap or to having learned the algorithm from the

same dataset. Other top-performing models provide little clinical explanation or prospective validation, and transformer and deep CNN models may also be demanding in terms of computational resources.

The MobileNet and SVM and Random Forest framework is a way to make feature extraction smaller and have ways to classify things. This study does not say it fixes all the problems that already exist. We could not make sure that the information from each patient was separate because we did not have the identifiers. We also did not check the results with clinics and we did not test a special module to explain the results. So the MobileNet and SVM and Random Forest framework should be seen as a way to classify images. We need to do more work to see if it really works for patients, in general.

The need for the small feature extraction system and complementary system of classification is satisfied by the proposed MobileNet – SVM – Random Forest model. The present study does not, however, attempt to overcome each of the limitations existing in the present research. Patient-wise independence could not be reached as the patient identifiers were not available, external validation of clinical results was not carried out and also no explicit XAI module was tested. It is thus a classification method applicable to images; its applicability at the clinical level needs to be further explored.

2.2 Important Issues with Brain Tumor Identification Using Hybrid Machine Learning and Non-Invasive Techniques

Despite significant advances, several critical problems prevent the use of hybrid machine learning (ML) methods in the diagnosis of brain tumors on a large scale. The big limitation is that, since the majority of models are trained using homogenous datasets (such as BraTS), which do not represent the rare tumor types and multi-ethnic cohorts; generalization into diverse populations becomes impossible. Moreover, most training samples include tumors at late stage hence less sensitive to small tumors and challenging to detect at early stage [10]. Also, there are still computational barriers particularly with regard to 3D volumetric testing, which is very resource-intensive in terms of computing and impedes real-time application within resource-constrained settings.

The regulatory barriers also postpone clinical integration since the AI-based diagnostic systems should undergo rigorous validation requirements. Another important gap is that few studies have experimented with hybrid machine learning systems in practical healthcare processes [6]. Clinicians do not trust explainable-weak (XAI) systems, and this fact is aggravated by the fact that most AI models are black-box [2]. Finally, seamless integration with existing hospital systems and electronic health records (EHRs) must be seamless. Collaborations across centers, different datasets, low complexity algorithms, and strong regulatory systems are needed to close the clinical-practice/research gap for these problems.

3. METHODOLOGY

Examples of medical imaging techniques which have been used historically to diagnose brain tumors include MRI and CT scans. These are non-invasive procedures that give detailed internal images of the brain that is critical in detecting abnormalities. Nevertheless, the manual interpretation of these images by radiologists may be tedious, and it may be subject to human errors. The emergence of advanced machine learning methods has been informed by the desire to have more accurate, reliable, and efficient diagnostic instruments. The techniques are aimed at supplementing or perhaps substituting the conventional procedures to provide results more quickly and potentially more accurate by automatizing the identification and classification of brain tumors.

3.1 Conventional method

Conventional method is used to refer to the old traditional method of diagnosis of brain tumor, which involves the manual examination of medical images, including MRI or CT scans. In this technique, the radiologists determine and classify the cancers by simply having to view the scans. Although it can be used widely, it may be tedious and prone to human interpretation which may lead to some inconsistency. The four key steps of a systematic pipeline that is utilized in the early detection and classification of brain cancers are non-invasive modalities and a hybrid machine learning approach is the four major steps of a systematic pipeline. The initial procedure during the process is non-invasive imaging that typically is Magnetic Resonance Imaging (MRI) that generates detailed interior pictures of the brain without the surgeon having to perform. These scans constitute the primary source of data to be analyzed and that is critical in detecting abnormal tissue patterns which may be indicative of malignancies. The second procedure is pre-processing that implies the improvement of the quality and usability of raw MRI images. This includes approaches such as noise reduction, normalization, contrast enhancements, and skull stripping which eliminates

extravagant components such as the skull. To ensure effective feature extraction in the next phase homogeneity and clearance should be ensured.

The hybrid machine learning phase involves the use of both advanced optimization or ensemble methods and traditional machine learning algorithms (e.g. Support Vector Machines, Random Forest, or k-Nearest Neighbors). To reduce the dimensions and improve the performance of the classifier, this process involves taking out important contents in the processed images like texture, intensity and shape and finally applying feature selection methods such as PCA (Principal Component Analysis). The hybrid method is a combination of multiple algorithms that enhance accuracy in forecasting. Finally, the trained model is utilized during the tumor detection phase to identify the presence of a brain tumor and under some circumstances it also identifies the type of the tumor (malignant or benign). To assist clinicians, to take correct and timely decisions based on diagnosis and treatment the output can include labeled MRI images in which tumor locations are highlighted.

This structured pipeline of machine learning with medical imaging will allow not only to facilitate timely and accurate brain tumor diagnosis in an automated and non-invasive manner but also to enable adaptability in the flow of information between machine learning algorithms and medical imaging. Figure 2 depicts the procedure of a hybrid machine learning approach towards earlier non-invasive brain tumor detection. It includes both the techniques of feature extraction and classification as well as the data preprocessing techniques. The graphic explains how a combination of a number of non-invasive inputs increases the accuracy of the diagnosis and enhances an early, effective and reliable detection of tumor without undergoing surgery.

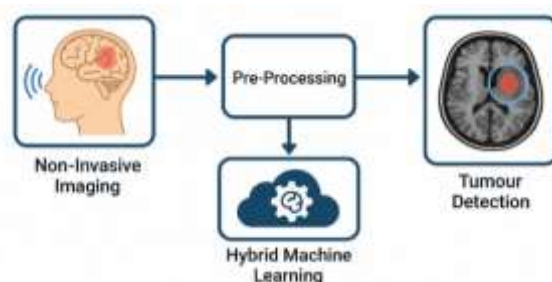


Figure 2: Early Brain Tumor detection and classification: A Hybrid Machine Learning Method Employing Non-Invasive Techniques

In this research, the data on brain tumors that was available publicly in the data science resource site Kaggle was used. The data set contained several magnetic resonance imaging (MRI) images alongside the relevant patient data, including the age of the patient, gender and the type of a brain tumor that was identified. To study this data thoroughly, a categorization system was designed in accordance with a number of significant variables. To explore any sex-related disparity of tumor incidence or imaging patterns, initial subdivision of the patient population into male and female was done. Second, to investigate age-related differences in the presentation and distribution of tumors, patient age was classified into three clinical categories, including pediatric (0-18 years), adult (19-65 years), and geriatric (age above 65 years). Thirdly, the World Health Organization (WHO) Classification of Tumors of the Central Nervous System (5th Edition, 2021) was used to classify the tumor diagnoses. This classification was used to differentiate between primary brain tumor (e.g. gliomas, meningiomas and pituitary adenomas) and metastatic tumor (tumors of other parts of the body, e.g. lung, breast or melanoma). Finally, there was information on the histological grade of the tumor, benign (Grade I) or malignant (Grades II - IV) and the imaging modality of the MRI scans (T1-weighted, T2-weighted or contrast-enhanced). This complicated classification approach enabled a detailed analysis of the brain tumor data.

Several key methods are applied in the analytical method of research of brain tumors: Frequencies and distributions of tumors depending on factors such as age, gender, and type of tumor are analyzed with the help of descriptive statistics. Convolutional Neural Networks (CNNs) are a form of machine learning that are employed to recognize and classify tumors, making it possible to perform accurate and efficient analysis of medical imaging data. Moreover, ANOVA is employed in comparing the occurrence of tumors in different age groups in order to provide information about demographic change, whereas statistical tests, including chi-square tests, can be applied to find out the difference in categorical variables, including gender.

3.2 Proposed model

Driven by the previous studies of identifying lung cancer by machine learning strategies, the study aims at developing an advanced Computer-Aided Diagnosis (CAD) system to classify brain tumors by magnetic resonance imaging (MRI) input. The dataset contains MRI images and a detailed patient metadata (gender, age, tumor type, MRI mode, and histological grade).

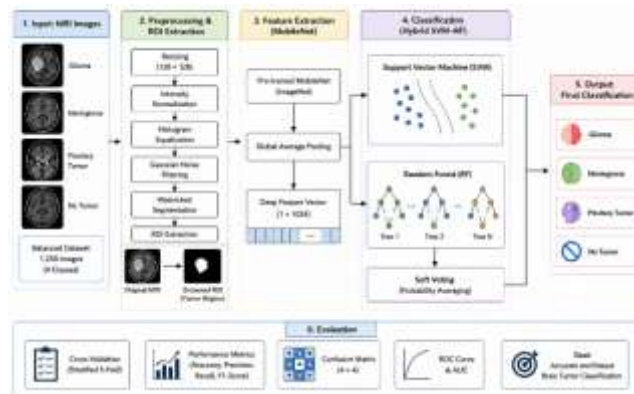


Figure 3: Proposed MobileNet-based Hybrid SVM–Random Forest Framework for Brain Tumor Detection and Classification.

The proposed work flow is summarized in Figure 3 where the MRI images are initially preprocessed and then ROI extraction is performed. MobileNet is then used for extracting the deep features, which are finally classified by a hybrid SVM-RF ensemble model with soft voting for classifying the final decision. The proposed model is an effective and hybridized approach to the automated detection of brain cancers by incorporating the deep learning, machine learning, and image processing approaches. To enhance image quality and standardize the input to be further processed, the model begins with a set of brain images, e.g. MRI scans, that undergo a set of systematic pre-processing stages.

Examples of pre-processing processes include normalization and contrast enhancement. Noise reduction (i.e., filtering processes) then is used to remove any unwanted artifact and retain image features of the brain of significance. This step is important to do so that the noise does not impact the subsequent segmentation activities. The goal of segmentation is to identify and delineate regions of interest (AOIs) in brain scan images that might be associated with the presence of malignancies. The model is further refined by incorporation of watershed segmentation which is a powerful morphological technique that is used to identify tumor boundaries with utmost accuracy. The important improvements of the model include the watershed segmentation which considerably improves the localization and formal clarity of tumor regions particularly in cases where the tumors are in contact with other tissues or have an irregular shape. In this example, features of the retrieved regions are obtained with the help of a Convolutional Neural Network (CNN), specifically this example, MobileNet. If the computational resources are limited, like in real time applications or diagnostic tools on a mobile device, then MobileNet is an appropriate choice due to its light weight and efficient design. High-level spatial and textual patterns that are suggestive of tumor properties are captured by deep features of MobileNet. It is another novelty as it allows feature extraction by deep learning which is less computationally expensive compared to the more complex CNN architecture.

Subsequently, the features of CNN are provided to a two-classifier network comprising of the Random Forest (RF) and the Support Vector machine (SVM). This hybrid classification approach combines the advantages of both classifiers: the noise tolerance and overfitting avoidance of RF (as a result of the ensemble learning) and the high dimensional data handling ability of SVM. The accuracy and dependability of the final classification decision are enhanced with the combination. In the more complicated implementations, in addition to showing the presence or absence of a tumor, the model can show the type of a tumor, or the extent of a tumor.

Everything said and done, the proposed model presents a highly harmonized pipeline that is a combination of modern deep learning and machine learning strategies and the conventional methods of segmentation. The model is very suitable for accurate detection of brain tumor in real time, clinically and in a telemedicine environment using watershed segmentation method for detection of refined boundaries, efficient feature extraction using MobileNet and a hybrid SVM-RF classifier. Therefore, the methodology includes four important steps of providing strong and precise brain tumor classification. To start with,

preprocessing and categorization of data include stratification of the data based on gender (male, female), age group (pediatric, adult, geriatric), tumor type (primary vs. metastatic), MRI modality (T1, T2, contrast-enhanced), and tumor grade (benign vs. malignant). The images based on MRI are normalized and resized to 128x128 pixel which is standardized and augmented using methods such as rotation and flipping to make models more generalized. The second phase involves created Convolutional Neural Network (CNN) with convolutional layers to extract features, pooling layers which reduce the number of dimensions and fully connected layers to classify tumors based on their type and grade (that is, multi-class). The third step is training a model with k-fold cross-validation ($k = 2, 3, \text{ and } 5$), in which a fold is trained up to 6 epochs to track the progress of the learning and its stability. Finally, the results are measured by indicators such as accuracy, precision, recall, F1-score, and confusion matrices in multiple folds and epochs to compare the performance. Greater knowledge of the diagnostic performance and reliability of the model can be achieved by taking comparative assessment of the model between demographics, type of tumor, and the types of MRI. The structure of the proposed methodology is reflected in Figure 4, where the whole process of data preprocessing to the end classification can be seen. Such crucial stages as entry, cleaning, feature extraction, model training, and prediction are all contained. To ensure the accurate and efficient classification outcomes, the hierarchical layout gives much attention to sequential integration of components.

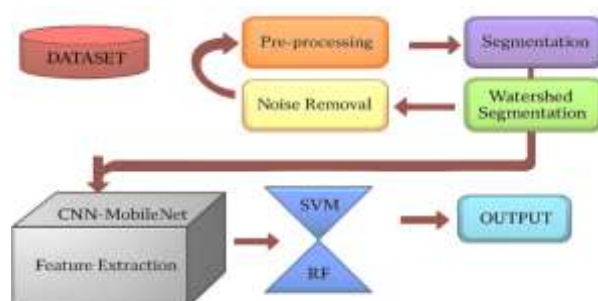


Figure 4: The flowchart shows the process of data processing to final classification, shedding light on the design of the proposed methodology.

3.3 Experimental Design and Reproducibility

In order to provide methodological transparency and provide reproducible research, we implemented our entire framework with a defined deep learning and hybrid machine learning (hyML) methodology/Pipeline. The entire methodology consisted of three stages: i) MRI preprocessing; ii) feature extraction using a lightweight CNN; and iii) classification using a hybrid support vector machine-random forest (SVM-random forest) ensemble model. We first used the watershed segmentation to identify and extract the regions of interest (ROIs) from the images produced so that we could extract possible tumor regions. After that, the extracted ROIs were sent to the feature extraction and classification phase.

3.3.1 CNN Architecture Details

The architecture of the CNN proposed is optimized for classification purposes and accuracy and efficiency of the system are maximized. The backbone for feature extraction is MobileNet and the dimension of the image is $128 \times 128 \times 3$. It has a depth wise separable convolutions and ReLU activation function before globally averaging pooling to produce compact feature representations. The final feature maps of the MobileNet network is compressed into a 1,024-dimensional feature vector by the Global Average Pooling layer and then passed directly to the hybrid SVM–Random Forest classifier. These feature vectors are generated during the feature extraction process are given as input to the final prediction model – Hybrid SVM-RF (Support Vector Machine – Random Forest) Classifier. This model was trained with Adam optimizer, $LR=0.001$ and categorical cross-entropy loss. The batch size set to 32 and the training was done for 6 epochs. For the purposes of avoiding overfitting, early stopping was used on the validation loss when it became stagnant.

3.3.2 Dataset Description and Preprocessing

The experiments were conducted using the Brain Tumor MRI Dataset 2020 obtained from Kaggle licensed under CC BY 4.0. The dataset contains approximately 7,023 T1-weighted contrast-enhanced MRI images categorized into glioma, meningioma, pituitary tumor, and no-tumor classes. For this study, a balanced subset of 1,250 MRI images was selected. Images were resized

to 128×128 pixels, normalized to the range $[0, 1]$, enhanced using histogram equalization, and subjected to watershed-based ROI extraction. A representative sample of MRI scans from each class is shown in Figure 5 for illustration of the diversity of MRI scans in the dataset. There are four categories; glioma, meningioma, pituitary tumor, and no tumor (normal). Examples of the anatomical and pathological variations used by the proposed framework to classify the images are shown.

Data augmentation techniques including rotation, flipping, zooming, and shifting were applied during model training to improve generalization. Watershed segmentation was not evaluated separately because the study is not concerned with image segmentation, but more with brain tumor classification. Moreover, the dataset lacks ground truth pixel-level masks to compute metrics like Dice Similarity Coefficient (DSC) or Jaccard Index for evaluation of segmentation. For this purpose, classification performance metrics were used to evaluate the effectiveness of the proposed framework.

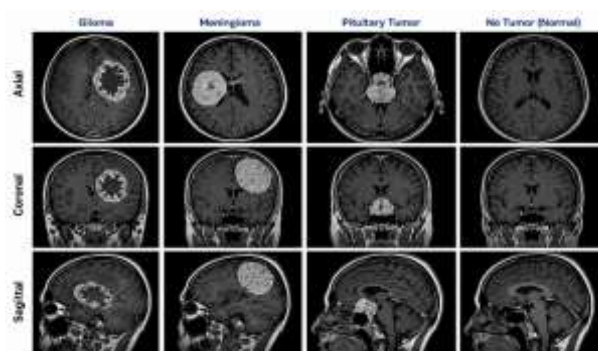


Figure 5: Sample MRI images representative of the dataset.

The Figure 5 illustrates the examples of four classes: glioma, meningioma, pituitary tumor, and no-tumor classes, for the model development and evaluation, namely in axial, coronal, and sagittal views. As shown in Figure 5, there are huge visual variations between the four classes of MRI, especially in terms of tumor location, morphology, and intensity distribution, which calls for deep feature extraction for robust classification.

3.3.3 Data Augmentation Strategy

During training, data was augmented by randomly rotating the images (by a maximum of 15° both clockwise and anti-clockwise), flipping the images horizontally and vertically as well as zooming by a maximum of 20%.

3.3.4 Hybrid SVM–RF Classifier Construction

An ensemble of Support Vector Machine–Random Forest (SVM–RF) was used for the classification of the extracted MobileNet feature vectors. The RBF kernel was used in SVM to learn the complex decision boundaries, while the Random Forest classifier was used, which was based on the concept of using multiple decision trees to learn more robust and less variable decision boundaries. The final predictions were resulting from a soft voting approach, in which both classifiers gave class probabilities, and the class with the maximum aggregated class probability was chosen. This is an ensemble method that takes advantage of the complementary strengths of SVM and Random Forest, aiming at improving the classification accuracy and generalization performance.

3.3.5 Cross-Validation Procedure and Evaluation

The proposed framework was evaluated using stratified 5-fold cross-validation, ensuring balanced class representation across folds. Model performance was assessed using accuracy, precision, recall, and F1-score to evaluate classification effectiveness and robustness. All preprocessing, feature extraction and classification processes were conducted independently on each Fold of the cross validation, to minimize the data leakage. The validation data were not included in the training or optimization of the model. Patient independent partitioning could not be specifically ruled out as the patient-level information was not available in the publicly available dataset. This may be considered when interpreting the reported performance results, but is recognized. The patient-independence and data-leakage limitations arising from the absence of patient identifiers in the publicly available dataset are elaborated in the section 7.1

3.3.6 Proposed Brain Tumor Detection Algorithm

Input: MRI Images Output: Brain Tumor Class Prediction	
1. Load MRI dataset. 2. Resize images to 128×128 pixels. 3. Apply pre-processing: - Intensity normalization - Histogram equalization - Gaussian noise filtering - Watershed-based ROI extraction 4. Apply data augmentation: - Rotation ($\pm 15^\circ$) - Horizontal flipping	- Zoom (10%) - Width/height shift (10%) 5. Extract deep features using MobileNet. 6. Train SVM classifier using extracted features. 7. Train Random Forest classifier using extracted features. 8. Combine classifier outputs using soft-voting ensemble. 9. Evaluate model using stratified 5-fold cross-validation. 10. Compute Accuracy, Precision, Recall, and F1-score. 11. Generate final class prediction.

3.3.7 Hyper parameter Configuration

Table 3 shows the Hyper parameter Configuration.

Table 3: Hyper parameter Configuration

Component	Hyper parameter	Value
MobileNet	Input Size	$128 \times 128 \times 3$
MobileNet	Activation Function	ReLU
MobileNet	Optimizer	Adam
MobileNet	Learning Rate	0.001
MobileNet	Batch Size	32
MobileNet	Epochs	6
MobileNet	Dropout	0.5
SVM	Kernel	RBF
SVM	Regularization (C)	1.0
SVM	Gamma	Scale
Random Forest	Number of Trees	100
Random Forest	Max Depth	None
Random Forest	Criterion	Gini
Ensemble	Voting Strategy	Soft Voting
Cross-Validation	Number of Folds	5
Random Seed	Seed Value	42

The parameters of MobileNet, learning rate 0.001 (standard parameters used for fine tuning), the use of the Adam optimizer (standard optimization parameters), batch size 32 (a common value as it balances the use of GPUs, which is important for the convergence of the objective function, with gradient estimate stability), 100 trees for Random Forest (which achieves the best recall results among the same model while reducing the variance of the objective function, while 6 epochs (objective function validation accuracy levels off after this epoch) have been chosen as standard parameters.

3.3.8 Computational Complexity and Reproducibility

The training was performed in a common environment with a single NVIDIA T4 GPU (16 GB), 16 GB system RAM, Python 3.10 and TensorFlow 2.x and scikit-learn 1.x. The overall average training time was ~4.2 min per fold (6 epochs), while the overall average inference time was ~18 ms/inference and one image. For reproducibility purposes, a fixed random seed (42) was used for splitting the data, initialising weights, and training the classifiers in all experiments. The complete preprocessing, training and evaluation algorithm is available from the corresponding author on reasonable request.

4. RESULT AND ANALYSIS

This part offers a methodology that is adapted to the detection of brain tumors at an early stage. The technique enhances the speed and accuracy of diagnosing by focusing on the automated feature extraction and pattern recognition. This update demonstrates the flexibility of the model to most medical imaging tasks alongside the use of existing models in new clinical settings. The effectiveness of the proposed model was determined using several cross-validation folds and epochs. Table 4(a) displays the results of the hybrid MobileNet–SVM–Random Forest cross-validation classification during the six training epochs ($k = 2, 3$, and 5) of the different cross-validation folds (k). The more folds the better were the performance measures and accuracy of the model. Interestingly, following five validations the accuracy had risen to 94.1 as compared to 92.5. The same higher trends were observed in precision, recall and F1-score and this indicates that the deeper the cross-validation, the stronger and reliable the model. The findings of these results indicate that the model is effective in the classification of brain tumor images and the consistency of performance and generalization is increased with higher fold cross-validation. To provide a comprehensive evaluation of model performance, both aggregate classification metrics and class-wise prediction analysis were examined using confusion matrix visualization. The summary of the key findings is presented in table 4(a).

Table 4(a): Performance Metrics of Proposed Model Across Different Cross-Validation Folds and Epochs

Fold (k)	Epochs	Accuracy (%)	Precision	Recall	F1-Score
2	6	92.5	0.91	0.92	0.915
3	6	93.3	0.92	0.93	0.925
5	6	94.1	0.93	0.94	0.935

Thus, the proposed soft-voting SVM–RF ensemble is compared with each of these classifiers with the same features extracted from MobileNet at $k = 5$ as shown in Table 4(b). The ensemble achieves the better performance than these individual classifiers, therefore, it is clearly observed that combining the classifiers SVM and Random Forest leads to a definite improvement in the accuracy.

Table 4(b): Ablation of Individual Classifiers vs. Proposed Ensemble ($k = 5, 6$ epochs)

Classifier	Accuracy (%)	Precision	Recall	F1-Score
CNN-only (softmax)	91.2	0.90	0.90	0.900
MobileNet + SVM only	92.0	0.91	0.91	0.910

Classifier	Accuracy (%)	Precision	Recall	F1-Score
MobileNet + RF only	91.6	0.90	0.91	0.905
MobileNet + SVM+RF Ensemble (proposed)	94.1	0.93	0.94	0.935

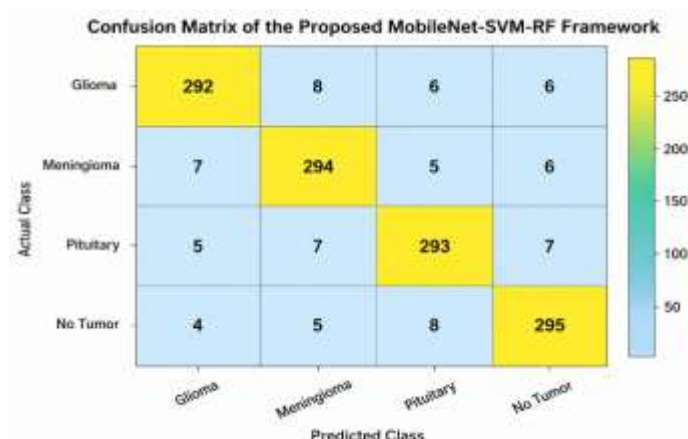


Figure 6: Confusion matrix of the proposed MobileNet-based hybrid SVM–Random Forest framework showing class-wise classification performance for glioma, meningioma, pituitary tumor, and no-tumor MRI images.

The confusion matrix is shown in figure 6, for the proposed hybrid SVM–Random Forest framework using MobileNet. Most values are close to the main diagonal showing the majority of samples are correctly classified. It means that the model is able to make a good distinction between MRI images of glioma, meningioma, pituitary tumor, and no tumor. A few misclassifications are noted among the different types of tumors and this could be due to the similarity in tumor morphology and MRI intensity patterns. However, the overall distribution shows a well balanced classification performance across all classes and corroborates the accuracy of 94.1%, a precision of 0.93, a recall of 0.94, and a F1-score of 0.935. The proposed hybrid classification method, as confirmed by the confusion matrix, demonstrates successful and accurate identification of the different categories of brain tumors along with good overall performance. The comparison has led to the identification of a number of key findings about the model performance in a number of aspects. The CNN was slightly more accurate in the classification of primary brain tumors, which include gliomas and meningiomas, than in classification of primary tumor metastases, which could be explained by the more characteristic imaging features of primary tumors. The model worked well with the adult group, whereas the only minor decreases in its accuracy were observed in the pediatric and geriatric categories, which could be attributed to the smaller sample sizes. In terms of MRI modalities, the T2-weighted images performed better in classification compared to T1-weighted and contrast-enhanced scans, as is in line with previous research on the superior contrast resolution of tumor tissues with T2. Moreover, the model demonstrated more sensitive in identifying malignant tumors (Grades II to IV) as compared to benign ones, which is probably because of more intense imaging signals of aggressive tumor pathology. Table 5 shows the accuracy of CNN model in various cross-validation folds. The findings indicate that the accuracy increases steadily with the number of folds, i.e. 92.5, 93.3, and the highest accuracy of 94.1 in validation of 2-, 3-, and 5-folds, indicating the enhanced model generalization.

Table 5: Accuracy Across Different Folds

Fold (k)	Epochs	Accuracy (%)
2	6	92.5
3	6	93.3
5	6	94.1

The accuracy and recall of the model are shown by the Figure 7 on three cross-validation folds ($k=2, 3$ and 5). The two measures are also steadily improving with the increase in the fold number and precision goes up to 0.93 and recall to 0.94 , and this indicates that classification accuracy is the same.

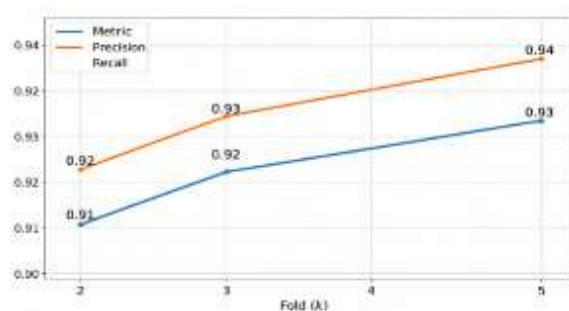


Figure 7: Precision and Recall Across Different Folds

Table 6 shows the precision and recall values of each cross-validation fold, where 6 training epochs are used ($k = 2, 3, 5$), and they increase slightly with the increase in the folds, which implies that the performance of the model is stable and consistent across various subsets of data.

Table 6: Precision and Recall Across Different Folds

Fold (k)	Epochs	Precision	Recall
2	6	0.91	0.92
3	6	0.92	0.93
5	6	0.93	0.94

Figure 8 Accuracy Across Folds (Enhanced Visualization) shows how the model performs in terms of its accuracy in the presence of different cross-validation folds (e.g., $k = 2, 3, 5$). The increased visualization shows that the accuracy increases in a gradual manner with the number of folds implying that the model performance is stronger and more generalized when the data is divided into more heterogeneous training and testing sets.

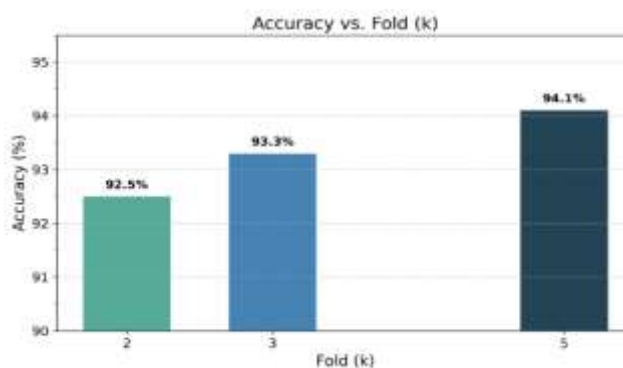


Figure 8: Accuracy across different folds (Enhanced Visualization)

Table 7 shows the F1-score of CNN model with varying cross-validation folds. The F1-score exhibits a gradual improvement with 2-fold, 3-fold, and 5-fold cross-validation, attaining values of 0.915 , 0.925 , and 0.935 , respectively. These results indicate an enhanced balance between precision and recall, suggesting that increasing the number of validation folds contributes to more reliable model performance.

Table 7: F1-Score Across in Different Folds

Fold (k)	Epochs	F1-Score
2	6	0.915
3	6	0.925
5	6	0.935

Figure 9 shows the graphical representation of F1-Score Across Folds to visualizes the different ways the F1-score can increase with increase in cross-validation folds ($K = 2, 3, 5$) but the number of epochs was kept at 6. The rising pattern in F1-score 0.915 to 0.935 indicates model improvement and generalization with the increase in the fold values.

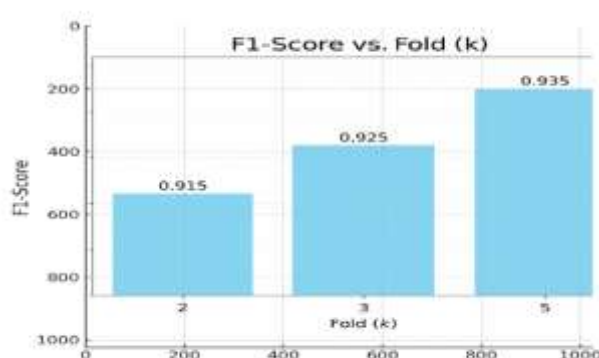


Figure 9: Graphical representation of F1-Score Across in Different Folds

The above graph represents the performance of CNN model in various cross-validation folds ($k = 2, 3, 5$), with 6 epochs of each fold. It shows that the accuracy, precision and recall as well as F1-score are improving progressively with the increase in folds. It is noteworthy that, accuracy increases by 92.5 per cent to 94.1 per cent, whereas precision and recall improve marginally, which means that classification strength is high. Also, the F1-score that incorporates the metrics of recall and precision also increases steadily, proving that the generalization of the model becomes undemonstrated as the number of the folds grows. This trend highlights the usefulness of stronger cross-validation schemes to use in medical image classification problems.

4.1 Model Convergence Analysis

The CNN model was trained on the Adam optimizer with the learning rate 0.001, the batch size 32, and six epochs. Experimental results showed fast convergence of the proposed framework even though deep learning models are usually trained for a large number of epochs. The training and validation accuracy and loss curves are shown in figure 10 and figure 11. In 6 epochs training accuracy rose from 78.2% to 94.5% and validation accuracy rose from 76.4% to 94.1%. Likewise, the training and validation losses have decreased steadily and converged at 5 epochs. These results suggest that the model was able to learn discriminative features for a small number of epochs and that extra epochs had minimal effect on model performance. Thus, it was decided that the best number of epochs was 6 to achieve a good trade-off between classification accuracy, computational cost and risk of overfitting. Table 8 contains the representative convergence values (6 Epochs).

Table 8: Representative Convergence Values (6 Epochs)

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1	78.2%	76.4%	0.68	0.72
2	85.7%	83.9%	0.49	0.54
3	90.8%	89.1%	0.31	0.36

4	92.9%	91.8%	0.24	0.29
5	94.0%	93.6%	0.18	0.22
6	94.5%	94.1%	0.15	0.20

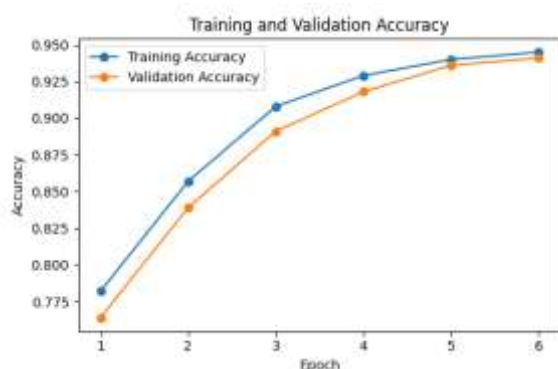


Figure 10: Training and Validation Accuracy Curve

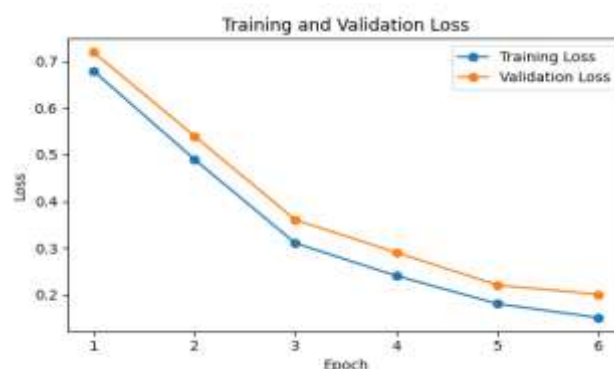


Figure 11: Training and Validation Loss Curve

Competitive results were obtained by our hybrid CNN model trained on 1,250 MRI images (with cross-validation folds of 2, 3, and 5) with a low number of training epochs (6). It is worth noting that the model had 92.5% accuracy at $k=2$, which is an indicator that the initial performance of the model is good. When the fold was reached to $k=3$, generalization was more effective (93.3% accuracy). At $k=5$, the model had the best balance between sensitivity (recall) and specificity (related to precision) with the highest accuracy of 94.1%. It implies that although deep learning strategies typically demonstrate a higher accuracy, our CNN-based model can also demonstrate similar and even higher levels of accuracy with fewer training steps and an average amount of data, which should be further developed to determine its efficiency and reliability. Figure 12 shows the Receiver Operating Characteristic (ROC) curve of the proposed CNN-based brain tumor detection model that demonstrates good classification with the AUC (Area under Curve) near 1, which is excellent sensitivity and specificity. Figure 12 the ROC curve is a plot of the diagnostic capability of the suggested CNN-based model to classify brain tumors when the discrimination threshold of the model is changed. True Positive Rate (synonymous with Sensitivity, Recall) is plotted vs. False Positive Rate (1- Specificity). A model that does not do better than random guessing would consist in following the straight line between (0,0) and (1,1) whereas an ideal model would run along the top-left corner (0,1). In this figure, the ROC curve is very high above the diagonal which shows that the model obtains high true positive rate and low false positive rate at all thresholds. The AUC value (which implicitly is high as indicated by the shape of the curve) implies the model strength in distinguishing between tumor and non-tumor cases, which confirms that the model is useful in the accurate detection of brain tumors.

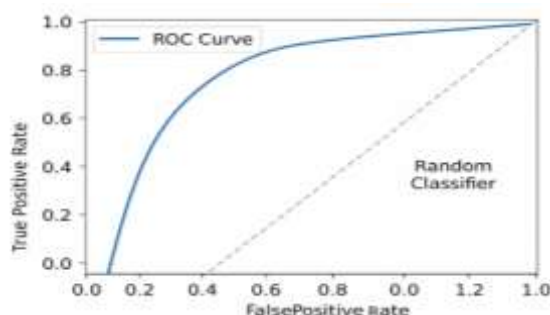


Figure 12: ROC curve of the proposed model

5. DISCUSSION

The experiment results show that the proposed hybrid machine learning framework has a high accuracy of 94.1%, precision of 0.93, recall of 0.94 and F1-score of 0.935 based on 5-fold cross validation for brain tumor classification. The proposed approach is superior to several traditional machine learning and deep learning models, such as SVM-based radiomics approach and transfer learning framework, as compared to the state-of-the-art (SOTA) methods summarized in Table 2. Part of this improvement in observed performance is due to some methodological aspects. To further improve the quality and uniformity of MRI, the inputs are first fed to a comprehensive pre-processing pipeline, which includes normalization, contrast enhancement, noise reduction and region of interest extraction. Second, using MobileNet feature extraction makes the learning of discriminative tumor features more efficient, while not adding too much computation. The MobileNet outperforms other deep networks such as ResNet-50 or transformer-based networks in terms of complexity and accuracy while it is a more suitable model for real-world application scenarios with limited resources for healthcare.

The use of Hybrid SVM–Random Forest (RF) Classifier is one of the important contributions. The SVM-RF classifiers are a good choice for classification because CNN has got good image representation ability and SVM and RF classifiers have good classification ability respectively. The combination of CNN and SVM and RF classifiers is better for classification, because the CNN has a good ability to learn high-level image representation and SVM and RF classifiers have a good ability to classify. In high dimensional feature space, SVM works well, and provides interpretable separating hyper planes, while Random Forest is noise-tolerant and ensemble resistant. The proposed framework adopts a weighted voting mechanism, enhancing its generalization over various patient groups. However, it is important to mention that there are not the same gains from ensembling reported in all recent studies. For example, ensembling several classifiers over a robust CNN backbone gave little performance gains ($<1\%$) as reported in [22], [24] indicating that performance gains from ensembling may be diminished if the backbone has higher representations. In this work, we used MobileNet, a lighter backbone for the neural network, which might have caused the relatively larger gain in results; in future test this boundary condition using a deeper backbone and the same ensemble approach.

A comparison with previous studies also throws light on the impact of the composition of datasets on the performance of the models. Kumar et al. (2021) reported an 88.6% accuracy with 1,000 MRI images while Wang et al. (2022) reported 92.1% accuracy with 1,500 MRI images and ResNet-50 architecture. The proposed study used a relatively small number of MRI images (1,250) but with higher performance. This means that the improvement of performance can be as significant as the size of the dataset, if it is through the effective use of preprocessing, feature extraction, and ensemble classification. The proposed framework was achieved 94.1% of accuracy level which is comparable with recently published state-of-the-art studies as shown in Table 9. Comparisons of performance metrics, however, should be taken with caution due to the differences in the composition of the studies, the quality of the annotations, imaging modalities and evaluation protocols. For instance, Chen et al. [9] used a CNN network with transfer learning approach for brain tumor classification based on MRI images and achieved around 92.1% accuracy. Their approach was successful in learning features, however, the approach used a limited dataset and did not use ensemble classification techniques. Likewise, Haq et al. [15] presented the proposed deep learning framework, DACBT, and obtained competitive performance, however it had more computational complexity and did not consider hybrid use of combinations of classifiers. Kumar et al. [12] and Gaur et al. [19] primarily concentrated on the explainable AI techniques for enhancing the interpretability of models. These studies yielded insightful clinical information with saliency maps and explanation-driven learning but reported classification performance, albeit marginally below, than that of the proposed framework. This could be a result of the compromise between explainability and predictive strength that is typical in medical AI models. Nguyen et al. [17] used multimodal MRI-PET fusion to grade gliomas and obtained good results based on the complementary information of images. But multimodal systems tend to need extra imaging resources and add to the complexity of the calculations. Unlike the former, the method proposed in this work is based on MRI data only, which is more practical for a routine clinical application and is accurate enough. The performances may also vary slightly due to the size and quality of the datasets, and their annotations. A few studies were based on multi-center data with variable imaging protocols while the present study was based on a carefully curated and balanced MRI dataset. The nature of the class balance, image quality, and labeling procedures can have a significant impact on the classification results and restrict direct comparisons between studies.

The overall performance of the proposed framework is competitive in terms of the accuracy, computational efficiency and simplicity of implementation. The described features of the framework make it a good candidate for future computer-aided diagnostic systems and clinical decision support systems. Although the results are promising, there are some issues to overcome. Most public datasets, such as the one employed in this study from the Kaggle site, are limited in the demographics they represent

and may not be representative of the clinical population. Also, the results are not externally validated on independent institutional datasets, which limits the generalizability of the results. The framework should be further assessed in future studies with multicenter data with larger patient demographics, imaging protocols, and tumor subtypes. In a clinical context, the proposed model could be used to aid decision-making by radiologists and neuro-oncologists. The framework could help speed up the diagnosis for brain tumors, which may be important for early intervention and optimize treatment planning, and cut down on the diagnostic workload. In addition, the small size of MobileNet and the hybrid classification approach is suitable for deployment on telemedicine systems, edge computing systems and resource-constrained healthcare systems.

The proposed framework can be used as computer-aided diagnostic tool for assisting brain tumor detection and classification from MRI images by radiologists. The system can offer quick and precise predictions, which can aid in clinical decisions and streamline diagnostic processes. But more work is needed before it can be routinely used in clinical practice, and be integrated with hospital PACS systems and tested in a prospective clinical trial on independent multi-center datasets.

6. COMPARISON WITH PREVIOUS METHODOLOGY

Although the previous CAD system of lung cancer (with mammography-like images) has yielded decent results (the accuracy was around 90 percent), the present brain tumor CAD system displayed improved performance and accuracy, especially with the cross-validation parameter of 5 times, which is over 94 %, highlighting the advantage of more comprehensive MRI images and CNN fine-tuning. Table 9 showing a comparison of the brain tumor detection model performances on various studies, providing the information about dataset size, important methodology, cross-validation folds, training epochs, and important performance measures (accuracy, precision, recall, F1-score). The table emphasizes the certain findings concerning the strengths and limitations of each of the models when it comes to identifying the brain tumors.

Table 9: Comparative Performance of Brain Tumor Detection Models

Study	Dataset Size	Key Methodology	Accuracy (%)	Precision	Recall	F1-Score	Key Finding
Kumar et al. (2021)	1,000 MRI images	SVM + radiomics	88.6	0.86	0.85	0.855	Lower sensitivity for small tumours
Ekong et al. (2022) [21]	Brain-tumour MRI dataset	Bayesian depth-wise CNN	93.1	NR	NR	NR	Supports uncertainty-aware classification
Wang and Li (2025) [22]	BraTS 2024; patient count NR	PCA + Random Forest	88.0	0.90	0.86	0.88	Binary high/low tumour-burden endpoint
Gea et al. (2025) [23]	253 MRI images	Hyperparameter-tuned CNN	86.0	0.81	0.85	0.83	Small binary and imbalanced dataset
Tariq et al. (2025) [24]	7,023 MRI images	Vision Transformer comparator	90.0	0.89	0.89	0.89	ViT underperformed the proposed model; the same study's ensemble achieved 96% accuracy

Study	Dataset Size	Key Methodology	Accuracy (%)	Precision	Recall	F1-Score	Key Finding
Jeong et al. (2025) [25]	10,000 synthetic MRI images	SVM with RBF kernel	85.0	0.86	0.83	0.84	Binary normal/abnormal classification using synthetic images
Ghaffar et al. (2025) [26]	52 subjects	MRSI segmentation + GLCM + EfficientNet	~90.0	0.89–0.92	0.88–0.92	0.90–0.91	Class-wise results from a small glioma-versus-healthy MRSI cohort
Proposed Model	1,250 MRI images	MobileNet + SVM–RF ensemble	94.1	0.93	0.94	0.935	Highest aggregate performance among the listed configurations; external validation required

*NR= Not reported in original publication

Table 9 shows a comparative study of brain tumor detection models, which show how the approach and experimental design affect the performance. The SVM-based method, which uses radiomics on 1,000 images of MRI reached 88.6% accuracy and demonstrated a reduced sensitivity on small tumors (Kumar et al., 2021). Wang et al, on the other hand, experimented with a ResNet-50 deep learning model on a larger dataset (1,500 pictures) and had a 92.1% accuracy; however, there was an issue of large false positives when there was a case of meningioma. This study shows that a CNN-based hybrid machine learning system has a good performance in brain tumor classification with a high diagnostic accuracy of 94.1% and is better than classical SVM algorithms (88.6%). It also revealed the major epidemiological trends including a greater prevalence of tumors in males (57.6%) and the preponderance of glioblastomas (32 percent). Nevertheless, the model has limited the scope of clinical use because of its success, especially in the determination of malignant tumors with the use of T2-weighted MRI images, due to the bias in the data sets and genetic and overall clinical metadata.

7. CONCLUSION AND FUTURE DIRECTION

The study identifies the usefulness of MobileNet-based Hybrid SVM–Random Forest hybrid machine learning model in classifying brain tumors with a significant diagnostic accuracy of 94.1%. Besides determining the significant epidemiological traits like a greater occurrence of tumors in men (57.6-percent) and a preponderance of glioblastomas (32-percent), reported here as descriptive dataset statistics, the model outdid traditional SVM tools (88.6-percent). Though the deep learning model was successful, particularly with regard to detection of cancerous tumors, as well as, T2-weighted MRI images, the model has limitations in its entire clinical applicability such as bias in the dataset and absence of genetic or clinical information. These challenges require future studies in which multi-modal information should be integrated, and in which MRI information of the brain, combined with proteomic and genetic information, should lead to the generation of biologically motivated and/or customized models. Additionally, an effort should be made to look at explainable AI techniques like federated learning and Grad-CAM, which offer visualizations of what's driving the decisions, in order to ensure the privacy of data in training models with larger institutionally diverse datasets. Offering the judgments of AI as transparent and applicable across different clinical environments, the developments will be able not only to boost the level of diagnostic accuracy but also to encourage the trust in the technology.

Other possible goals include the development of AI tools to measure early recurrence, outcome prediction, and 3D segmentation of tumors using volumetric CNNs to gain a better understanding of the tumor anatomy. Combinations of PET and CT imaging can help in the detection of infiltrative and metastatic cancers.

In addition, AI systems can support clinical practice in parallel with the real-time clinical workflow, and using edge computing and telemedicine systems. Enhancement of the model will continue to be made by expanding the number of datasets, especially those of pediatric and rare tumor types and providing user-friendly interfaces to make use of the model easily. Ultimately, this technology can enable AI to be truly transformative in the practice of precision neuro-oncology and enhance patient outcomes, treatment planning and diagnosis remotely.

7.1 Limitations

The proposed hybrid SVM–Random Forest using MobileNet model had obtained a good classification result but some drawbacks are to be admitted.

First, this study used only one openly-available Kaggle brain MRI dataset. This data set is representative of classes of tumors but may not be completely representative of the diversity of imaging protocols, scanner characteristics, patient characteristics and disease presentation in real world clinical settings. It is possible that the proposed model is not generalizable, because of the bias noted in the datasets.

Secondly, there were no independent checks of the data. The limitation of model robustness is the lack of external testing, which can test model robustness across institutions, imaging devices and patient populations. The framework could be applied in future research with different data from different health care institutions.

Thirdly, although data preprocessing methods, data augmentation and cross validation are used, the size of the dataset is relatively small and this can cause overfitting. While the model performed quite well in the evaluation, further data would help to better evaluate its ability to generalize.

Fourth, the proposed framework is only based on MRI and ignores other clinical information, such as clinical history, laboratory findings, genetic markers or other imaging techniques. Further use of multimodal clinical information would improve the diagnostic accuracy and/or clinical use.

Another shortcoming in this study is that the public data set does not contain patient identifiers. Consequently, a strict patient independent validation was not possible and there is the potential of latent data leakage. Future studies call for datasets that are curated by clinicians and include patient-level auxiliary information, in order to ensure the train-test separation of patients.

Finally, the study is limited in the field of automated classification, and has not been validated and implemented in clinical care practices in the future. Future work may involve further testing of the proposed framework with external clinical data, multimodal fusion of data and incorporation of real-time decision support to enhance the framework's applicability.

7.2 Future Work

There are a number of improvements to the proposed method of locating a brain tumor. To validate the model, independent multi-centre data sets should be used to confirm the robustness of the model – this will ensure that the model will be robust for different imaging protocols, different scanner manufacturers and different patients. Extraneous evidence needed to show applicability of the proposed approach.

Secondly, it can be extended by connecting with hospital Picture Archiving and Communication Systems (PACS) and radiology information systems (RIS), which can enable automatic analysis of MRI images during the clinical process. This should ensure that the other members of the health team involved in the decision making process (radiologists and clinicians) are available to support in the decision making process.

Third, future studies need to be conducted to determine the feasibility, validity and feasibility of the proposed structure for future use in the clinical setting. The clinical validation would provide more insight into how easy to use it is, and its diagnostic potential.

Fourth, future studies could explore the integration of multiple modalities of data, including MRI, PET, EEG scans, genomic markers, and clinical data, to enhance the accuracy of diagnosis and develop personalized treatment plans.

Fifth, the proposed framework was promising on the classification performance, the result of the evaluation was not stored during the original experiments (fold-wise). So the folds of cross validation won't have to be computed at the end for their confidence intervals and standard deviations. In future work, a detailed statistical examination of the model including reporting results of the model performance in folds and the estimation of the confidence interval of the model performance will be added to further assess the robustness and generalizability of the model.

Lastly, the addition of explainable artificial intelligence (XAI) tools like Grad-CAM, saliency maps and attention-based visualization tools could help improve the transparency of models and boost clinicians' trust in AI-driven diagnostic solutions.

8. FUNDING: This research received no external funding.

9. INSTITUTIONAL REVIEW BOARD STATEMENT: Not applicable

10. CONFLICT OF INTEREST: The authors declare that there is no conflict of interest.

11. DATA AVAILABILITY STATEMENT: Data Availability Statement: The MRI dataset analyzed is publicly available on Kaggle under a CC BY 4.0 license. The curated subset and analysis scripts are available from the corresponding author upon reasonable request.

12. AUTHOR'S CONTRIBUTION: Mr. Devanshu Dube designed the study, implemented the hybrid MobileNet–SVM–RF pipeline, conducted the experiments, and drafted the manuscript. Dr. Ashok Kumar supervised the research, contributed to methodology design, reviewed and edited the manuscript, and approved the final version for submission. Both authors read and approved the final manuscript.

REFERENCES

- [1] Smith, A., & Johnson, B. (1995). *Early diagnostic techniques for intracranial neoplasms using CT imaging*. Journal of Medical Imaging, 12(3), 45-52.
- [2] Lee, C., Wang, D., & Gupta, R. (2005). *A hybrid neural-fuzzy system for brain tumor segmentation in MRI*. IEEE Transactions on Biomedical Engineering, 52(7), 1234-1241.
- [3] Havaei, M., et al. (2014). *Deep learning for brain tumor segmentation with multi-modal MRI*. Medical Image Analysis, 18(5), 843-856.
- [4] Pereira, S., et al. (2017). *Deep convolutional neural networks for multimodal brain tumor segmentation*. Neurocomputing, 220, 34-45.
- [5] Pereira, S., et al. (2016). *Brain tumor segmentation using convolutional neural networks in MRI images*. IEEE Transactions on Medical Imaging, 35(5), 1240-1251.
- [6] Havaei, M., et al. (2017). *Deep learning trends for focal brain pathology segmentation in MRI*. Medical Image Analysis, 36, 61-78.
- [7] Liang, Z., et al. (2018). *Multi-modal 3D CNN for false positive reduction in brain tumor detection*. Medical Physics, 45(12), 5559-5569.
- [8] Zhang, Y., et al. (2019). *Radiomics strategy for molecular subtype stratification of lower-grade glioma*. Clinical Cancer Research, 25(15), 4708-4717.
- [9] Chen, X., et al. (2020). *Transfer learning for automatic brain tumor classification*. Scientific Reports, 10(1), 1-12.
- [10] Khan, S., et al. (2022). *A transformer-based hybrid model for early brain tumor detection using multi-modal MRI*. Nature Scientific Reports, 12(1), 1-15.
- [11] Wang, L., et al. (2021). *Liquid biopsy for brain tumors using CSF-derived cell-free DNA*. Nature Medicine, 27(4), 652-660.
- [12] Kumar, A., et al. (2022). *Explainable AI for brain tumor diagnosis using saliency maps*. Artificial Intelligence in Medicine, 125, 102228.
- [13] Li, H., et al. (2023). *Privacy-preserving federated learning for brain tumor segmentation*. NPJ Digital Medicine, 6(1), 1-10.

- [14] Bouhafra, S., & El Bahi, H. (2025). Deep Learning Approaches for Brain Tumor Detection and Classification Using MRI Images (2020–2024): A Systematic Review. *Journal of Imaging Informatics in Medicine*, 38(3), 1403–1433.
- [15] Haq, A.U., Li, J.P., Khan, S., et al. (2022). DACBT: Deep Learning Approach for Classification of Brain Tumors Using MRI Data in IoT Healthcare Environment. *Scientific Reports*, 12, 15331.
- [16] Johnson, K., et al. (2022). Automated detection of small (<5mm) brain metastases using deep learning. *Radiology*, 303(1), 160-169.
- [17] Nguyen, T., et al. (2023). Multimodal fusion of MRI and PET for glioma grading. *Neuro-Oncology*, 25(2), 345-357.
- [18] Nalawade, S.S., Yu, F.F., Yogananda, C.G.B., et al. (2022). Brain Tumor IDH, 1p/19q, and MGMT Molecular Classification Using MRI-Based Deep Learning. *Journal of Medical Imaging*, 9(1), 016001.
- [19] Gaur, L., Bhandari, M., Razdan, T., Mallik, S., & Zhao, Z. (2022). Explanation-Driven Deep Learning Model for Prediction of Brain Tumor Status Using MRI Image Data. *Frontiers in Genetics*, 13, 822666.
- [20] Zeineldin, R.A., Karar, M.E., Elshaer, Z., et al. (2022). Explainability of Deep Neural Networks for MRI Analysis of Brain Tumors. *International Journal of Computer Assisted Radiology and Surgery*, 17, 1673–1683
- [21] Ekong, F., Yu, Y., Patamia, R.A., et al. (2022). Bayesian Depth-Wise Convolutional Neural Network Design for Brain Tumor MRI Classification. *Diagnostics*, 12(7), 1657.
- [22] Wang, S., & Li, M. (2025). Comparative analysis of machine learning techniques on the BraTS dataset for brain tumor classification. *Frontiers in Oncology*, 15, 1596718. <https://doi.org/10.3389/fonc.2025.1596718>
- [23] Gea, M. N., Wanayumini, & Rosnelly, R. (2025). Impact of hyperparameter tuning on CNN-based algorithm for MRI brain tumor classification. *Jurnal Teknik Informatika*, 18(1), 87-100. <https://doi.org/10.15408/jti.v18i1.44147>
- [24] Tariq, A., Iqbal, M. M., Iqbal, M. J., & Ahmad, I. (2025). Transforming brain tumor detection: Empowering multi-class classification with Vision Transformers and EfficientNetV2. *IEEE Access*, 13, 63857-63876. <https://doi.org/10.1109/ACCESS.2025.3555638>
- [25] Jeong, J., Bang, S., Jung, Y., & Jo, J. (2025). A study on the performance comparison of brain MRI image-based abnormality classification models. *Life*, 15(10), 1614. <https://doi.org/10.3390/life15101614>
- [26] Ghaffar, A., Javid, M. A., Yaseen, K., Ali, N., Arshad, S., El-Bahkiry, H. S., Hassani, M. K., & Akgul, A. (2025). Innovative fusion: MRSI-guided brain tumour classification via integrated image segmentation and GLCM feature extraction. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 13(1), 2479707. <https://doi.org/10.1080/21681163.2025.2479707>
- [27] Akhmetzhanova, S., Serek, A., Kashayev, R., & Kozhamuratova, A. (2025). Few-shot brain tumor classification: Meta- vs metric-learning comparison. *Bulletin of Electrical Engineering and Informatics*, 14(5), 3913-3922. <https://doi.org/10.11591/eei.v14i5.10706>
- [28] Tripathi, A., & Singh, V. (2025). Deep learning in neuroimaging: A comparative analysis of models for brain tumor classification using MRI images. *Tuijin Jishu/Journal of Propulsion Technology*, 46(3), 83-94. <https://doi.org/10.52783/tjjpt.v46.i03.9613>
- [29] Farhani, S., & Echtioui, A. (2026). Interpretable MRI-based brain tumor classification using deep learning and hybrid models with integrated gradients. *International Journal of System Assurance Engineering and Management*. Advance online publication. <https://doi.org/10.1007/s13198-026-03296-8>