



OPEN **Brain tumor detection empowered with ensemble deep learning approaches from MRI scan images**

Rizwana Naz Asif^{1,12}, Muhammad Tahir Naseem^{2,12}, Munir Ahmad³, Tehseen Mazhar^{1,9}✉, Muhammad Adnan Khan⁴✉, Muhammad Amir Khan⁵✉, Amal Al-Rasheed⁶ & Habib Hamam^{7,8,10,11}

Brain tumor detection is essential for early diagnosis and successful treatment, both of which can significantly enhance patient outcomes. To evaluate brain MRI scans and categorize them into four types—pituitary, meningioma, glioma, and normal—this study investigates a potent artificial intelligence (AI) technique. Even though AI has been utilized in the past to detect brain tumors, current techniques still have issues with accuracy and dependability. Our study presents a novel AI technique that combines two distinct deep learning models to enhance this. When combined, these models improve accuracy and yield more trustworthy outcomes than when used separately. Key performance metrics including accuracy, precision, and dependability are used to assess the system once it has been trained using MRI scan pictures. Our results show that this combined AI approach works better than individual models, particularly in identifying different types of brain tumors. Specifically, the InceptionV3 + Xception combination hit an accuracy level of 98.50% in training and 98.30% in validation. Such results further argue the potential application for advanced AI techniques in medical imaging while speaking even more strongly to the fact that multiple AI models used concurrently are able to enhance brain tumor detection.

Keywords IoMT, Transfer learning, Deep machine learning, Histopathology, Image processing, And brain tumor

Context

Brain scans and a physical examination of the patient are necessary for the diagnosis of a brain tumor. Typically, this is accomplished by categorizing distinct brain tumors using a variety of brain imaging of the patient. To comprehend a patient's state, doctors need to view several slices of an image¹. Automated diagnostic technologies could cut down on the amount of time needed to diagnose patients, shorten their wait times, and lower the number of false positives.

The analysis of MRI images by machine-learning and deep-learning algorithms provides doctors with valuable insights, thus enabling diagnoses to become increasingly accurate and more informative. Further evolving methods improve accuracy in detection but also quicken medical assessment through collaboration with usual diagnostic methods^{2,3}.

This study builds on previous research by incorporating images of healthy brains and conducting a more in-depth analysis of neural networks³. Various deep learning models, including Fully Connected Neural Networks

¹School of Computer Science, National College of Business Administration and Economics, Lahore 54000, Pakistan.

²Department of Electronic Engineering, Yeungnam University, Gyeongsan-si 38541, Republic of Korea. ³University College, Korea University, Seoul 02841, Republic of Korea. ⁴Department of Software, Faculty of Artificial Intelligence and Software, Gachon University, Seongnam-si 13557, Republic of Korea. ⁵School of Computing Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Shah Alam, 40450 Selangor, Malaysia.

⁶Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia. ⁷Faculty of Engineering, Université de Moncton, Moncton NB E1 A3E9, Canada. ⁸School of Electrical Engineering, University of Johannesburg, Johannesburg 2006, South Africa.

⁹Department of Computer Science, School Education Department, Government of Punjab, Layyah 31200, Pakistan.

¹⁰ International Institute of Technology and Management (IITG), Av. Grandes Ecoles, Libreville BP 1989, Gabon.

¹¹Bridges for Academic Excellence, Spectrum, Tunis, Center-ville, Tunisia. ¹²Rizwana Naz Asif and Muhammad Tahir Naseem contributed equally in this work and share first authorship . ✉email: tehseenmazhar719@gmail.com; adnan@gachon.ac.kr; amirkhan@uitm.edu.my

(FCNNs), CNNs, and Recurrent Neural Networks (RNNs), have been explored in past studies. CNN-based techniques have shown high efficiency in tumor classification, particularly when combined with advanced feature extraction methods such as Growth Distribution Depth (GDD) for survival analysis⁴. To improve tumor categorization, we use a CNN-based method in this study. To improve tumor categorization, we use a CNN-based method in this study. With 18,280 recorded deaths among youth in the United States in 2022, brain tumors continue to be a major health concern. Through histological research, the medical world divides tumors into more than 150 different categories, generally classifying them as benign (non-cancerous) or malignant (cancerous)².

Despite the progress in neural networks, few diagnostic tools fully utilize recent advancements in this field. While these techniques have been widely used for analyzing non-medical images, this research adapts them for medical imaging, even when working with limited datasets³. CNN-based methods have also been successfully implemented for classifying various stages of neurological diseases using MRI scans, highlighting their versatility and effectiveness in medical imaging applications. Figure 1 illustrates the different categories of brain images analyzed in this study.

Brain tumors pose a significant medical challenge due to their potential to cause severe neurological damage and life-threatening conditions. Early and precise diagnosis is essential to improving patient outcomes and reducing long-term complications⁵. Magnetic Resonance Imaging (MRI) is widely used for brain tumor diagnosis and classification; however, manual assessment of MRI scans is time-consuming and prone to human error. This is where machine learning (ML) and deep learning (DL) techniques, including CNN-based frameworks used in multi-class classification of neurodegenerative diseases, have shown promising results in improving diagnostic accuracy.

Brain tumors (BTs) are abnormal growths that invade and destroy brain tissue, significantly reducing patients' quality of life. In many cases, suboptimal medical care further decreases survival rates. MRI is a standard imaging technique used for diagnosing brain tumors². However, manually segmenting MRI scans is a complex and time-consuming task due to the vast amount of imaging data. This underscores the necessity of automated segmentation techniques. Brain tumors vary in size, shape, and stage, making detection and classification a challenging task, which in turn complicates biomedical image processing. While earlier research explored machine learning for tumor detection, deep learning has emerged as a more effective technique for improving classification accuracy².

Brain tumors originate from the uncontrolled proliferation of malignant cells, leading to the formation of abnormal masses. They are broadly categorized into primary tumors, which develop in the brain, and secondary tumors, which result from metastasis from other parts of the body. These tumors are further classified as benign (non-cancerous) or malignant (cancerous), depending on their nature. Surgical removal is a viable treatment option for some tumors, with a low probability of recurrence in certain cases².

Tumor classification depends on the tumor's location and growth characteristics. While benign tumors may grow slowly with well-defined boundaries, malignant tumors tend to invade surrounding tissues aggressively, making them more difficult to treat. Diagnosis typically involves a combination of physical examinations, patient medical history assessments, laboratory tests, and imaging techniques. MRI remains the gold standard for detecting abnormalities in the brain and soft tissues³. However, other imaging methods, including Computed Tomography (CT), Positron Emission Tomography (PET), and Perfusion Magnetic Resonance Imaging (PMRI), are also used depending on clinical requirements. Early and accurate identification of brain tumors allows for timely intervention and appropriate treatment decisions^{5,6}. The approach for tumor extraction and removal depends on its precise location within the brain^{7,8}.

For effective treatment and improved survival rates, brain tumor diagnosis must be quantitative, precise, and rapid. Manual diagnostic approaches are often subjective, inconsistent, and error-prone⁹. Medical professionals must determine the exact tumor location to administer appropriate treatment strategies^{9,10}. In recent years, numerous ML and DL-based studies have focused on classifying brain tumors¹¹. Traditional machine learning techniques explored in the literature include Perfusion Magnetic Resonance Imaging Support Vector Machines (PMRI-SVM), k-Nearest Neighbors (KNN)^{12,13}, Artificial Neural Networks (ANN)¹⁴, Principal Component Analysis (PCA)¹⁵, and Decision Trees (DT)¹⁶.

While MRI scans provide valuable non-invasive imaging for initial tumor detection, the definitive diagnosis of tumor type often requires a pathological examination, usually performed through biopsy or surgical intervention¹⁷. This research aims to bridge the gap between traditional and modern diagnostic methods by

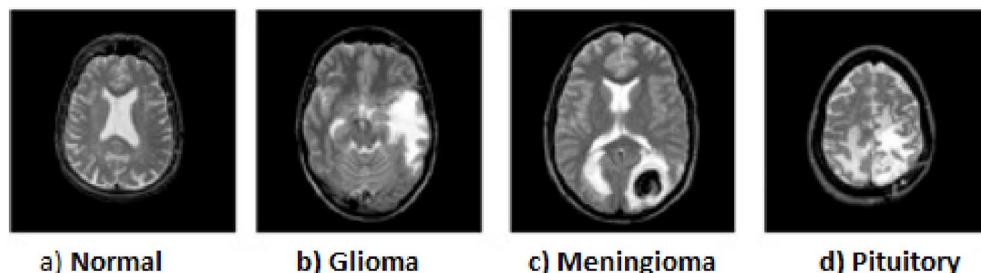


Fig. 1. Different classes of brain images.

implementing an ensemble deep learning approach that enhances the classification accuracy of brain tumors in MRI scans. The author of¹⁸ presents an enhanced form of EfficientNetv2 addressing these challenges by including Global Average Pooling (GAM) and Efficient Channel Attention (ECA). This development significantly increases the classification accuracy of brain tumors and significantly raises the model's attention on crucial elements in complex MRI images. We raise detection of a variety of brain tumors by precisely combining attention mechanisms enhancing feature extraction. This helps us to differentiate our approach. Our model achieves an extraordinary test accuracy of 99.76%, so establishing a new benchmark in MRI-based brain tumor classification, according to extensive tests on a big public dataset.

The proposed approach¹⁹ consists in a rescaled model and a new hybrid shifted windows Multi-Head Self-attention module (HSW-MSA). This development aims to simplify training, lower memory use, and raise classification accuracy. To improve accuracy, training speed, and parameter efficiency, the Swin Transformer substitutes the Residual-based MLP (ResMLP) for the conventional MLP. We evaluate the proposed-swin model using just test data from a publicly available four-class MRI brain dataset. Effective and consistent training helps to improve model performance; hence, the application of data augmentation and transfer learning methods improves also. The proposed-Swin model beats previous studies and deep learning models with a startling accuracy of 99.92%. This shows, when combined with HSW-MSA and ResMLP enhancements, the Swin Transformer is helpful for brain tumor diagnosis.

Using open-source SIPaKMeD and Mendeley LBC pap smear datasets with 106 deep learning models 53 convolutional neural network (CNN) models and 53 vision transformer models each author tested the proposed method²⁰. With scores of 99.02% on the SIPaKMeD dataset and 99.48% on the LBC dataset, the recommended method exceeded all deep learning models and current literature in terms of accuracy, surpassing both experimental and state-of-the-art approaches.

Objectives

This study aims to:

- Develop an effective method for brain tumor detection using deep learning techniques.
- Conduct a thorough analysis of both tumor-free and tumor-affected brain images to enhance classification accuracy.
- Evaluate neural network performance on the given datasets, measuring both per-image precision and per-patient diagnostic accuracy.
- Utilize InceptionV3 and Xception models to classify brain tumors while examining how similar challenges have been addressed in previous studies.

Strength and novelty in the proposed model

The main contributions of this study may be summarized as follows:

- Optimized Ensemble Deep Learning Approach: The proposed model integrates multiple deep learning architectures, improving accuracy and robustness in brain tumor detection while addressing the limitations of single-model approaches.
- Improved Generalizability and Noise Resistance: By combining multiple models, the ensemble method mitigates overfitting and enhances performance on unseen data, ensuring reliability even in the presence of noise and imaging variations.
- Efficient and Flexible Framework: The model optimizes computational resources, making it accessible for both research and clinical applications. It is also adaptable, allowing for future integration of additional deep learning models.
- Refined Tumor Classification: Unlike previous studies that focus primarily on tumor severity or malignancy grading, this research emphasizes classifying brain tumors into four distinct histological categories, providing a more precise diagnostic framework.

Structure of the article

The remainder of this paper is structured as follows. Section 2 reviews relevant literature to establish the study's background. Section 3 details the proposed system model, outlining its design and functionality. Section 4 presents the experimental results and their implications. Section 5 discusses the findings, compares them with existing work, and identifies limitations. Finally, Sect. 6 provides a recapitulation of the study and explores future research directions.

Literature review

Assessment of the therapeutic value of the study

Evaluating the potential clinical impact of this study requires a comprehensive understanding of its applications and limitations²¹. One of its key contributions lies in the use of MRI scans for preoperative planning, a crucial step in both tumor localization and surgical decision-making. The ability to visualize and characterize brain tumors in detail enhances surgical precision and optimizes patient care by providing critical insights into the most effective intervention strategies.

Another significant advantage of MRI-based methods is their ability to facilitate non-invasive tumor monitoring. MRI imaging enables continuous assessment of tumor growth and response to treatment without requiring surgical procedures. By tracking changes in tumor size and structure over time, clinicians can make well-informed decisions about necessary treatment adjustments, ensuring more effective therapeutic outcomes²².

Machine learning techniques play an essential role in improving tumor classification based on MRI images. By automating the detection and classification process, these approaches simplify and accelerate decision-making for healthcare professionals. However, deep learning models should be viewed as supportive tools rather than replacements for traditional diagnostic practices. While they enhance efficiency and assist in identifying tumor types, expert validation remains necessary to confirm diagnostic accuracy and reliability²³.

This research represents an important step toward enhancing the accuracy and efficiency of brain tumor detection. The proposed ensemble deep learning approach, which integrates multiple deep learning models, improves precision and classification reliability. Such advancements hold great potential for augmenting clinical services and fostering future developments in AI-driven diagnostic technologies²⁴. Nonetheless, pathological examination remains the gold standard for brain tumor classification. While MRI-based machine learning methods contribute significantly to preoperative planning and ongoing tumor monitoring, they should complement rather than replace traditional pathology assessments.

Ensemble deep learning for medical imaging

The application of ensemble deep learning in medical imaging, particularly for brain tumor detection, has demonstrated promising results^{17,21}. By combining multiple deep learning models, the ensemble approach synthesizes their outputs to generate a more robust and reliable prediction. The integration of multiple models enhances diagnostic accuracy while reducing the likelihood of overfitting to specific datasets, thereby improving generalizability²¹.

This study introduces a novel brain tumor detection method that merges edge detection techniques with a highly efficient CNN model. Conventional deep learning models typically rely on extensive training datasets for accurate classification. However, this research demonstrates that high diagnostic precision can still be achieved with limited data, offering a significant improvement over existing methods in resource-constrained environments. This breakthrough suggests that AI-driven diagnostic tools can remain effective even in settings where access to large medical imaging datasets is restricted.

Additionally, the model's efficiency and streamlined computational requirements make it particularly suitable for clinical applications. Unlike conventional transfer learning approaches, which often demand high computational power and extensive datasets, this method enables faster diagnoses while maintaining accuracy. The study also emphasizes the importance of refining preprocessing techniques, optimizing hyperparameters, and applying data augmentation to enhance model performance. These refinements contribute to the model's overall accuracy, making it a valuable addition to existing diagnostic frameworks.

A framework for improved brain tumor detection

This paper presents a neural network-based framework for brain tumor detection using MRI scans, designed to enhance diagnostic precision and reliability. By linking multiple deep learning models, the ensemble approach generates a final prediction that surpasses the accuracy of individual models. Experimental findings confirm the effectiveness of this method, demonstrating its potential to assist doctors in identifying brain tumors with greater speed and accuracy.

The ensemble deep learning method enhances prediction reliability by incorporating results from several models, reducing the risk of overfitting and improving classification stability. By leveraging multiple deep learning architectures, this approach refines diagnostic outcomes, increasing confidence in automated brain tumor detection. The proposed framework integrates deep learning ensembles to create a structured and reliable diagnostic model, reinforcing its role as a powerful tool in medical imaging analysis.

The findings of this study highlight the substantial potential of ensemble deep learning in advancing brain tumor detection. By offering a more accurate and efficient classification method, this research underscores the importance of AI-driven diagnostic techniques in clinical settings. The results indicate that this approach could serve as a valuable resource for medical professionals, supporting faster and more reliable brain tumor diagnosis while complementing traditional pathology-based assessments^{21,22}.

Ensemble deep learning in brain tumor detection

The assessment of the therapeutic value of this study necessitates a comprehensive examination of its prospective applications and inherent limitations²¹. The utilization of MRI scans plays a critical role in preoperative planning by providing essential tumor localization and characterization, which informs surgical decision-making and improves patient care. MRI-based approaches also enable non-invasive tumor monitoring, allowing for the assessment of tumor progression and treatment response without the need for invasive procedures²². Machine learning techniques have been widely explored for tumor classification, with deep learning models showing significant promise in improving diagnostic accuracy. However, these methods primarily serve as decision-support tools rather than standalone diagnostic solutions, reinforcing the need for expert validation²³. Research in this area aims to enhance both the efficiency and precision of automated tumor detection, ultimately contributing to technological advancements in diagnostic medicine²⁴. While MRI-based deep learning models offer substantial clinical value, traditional pathology remains the gold standard for definitive tumor classification. MRI-linked machine learning techniques are particularly beneficial for preoperative planning and clinical decision support^{17,21}.

Ensemble deep learning has emerged as a powerful approach in medical imaging, including brain tumor detection^{17,21}. By combining multiple deep learning models, ensemble methods aggregate their outputs to improve prediction reliability and reduce the risk of overfitting²¹. The proposed model leverages edge detection techniques alongside an optimized Convolutional Neural Network (CNN), demonstrating that effective classification can still be achieved with limited data, a notable advancement over conventional methods. The model's efficiency makes it particularly advantageous in resource-constrained environments, where access to

extensive medical imaging datasets is limited. Compared to traditional machine learning techniques or transfer learning models, which require significant computational resources, the proposed approach offers a more streamlined and accessible diagnostic tool. By refining preprocessing techniques, optimizing hyperparameters, and applying data augmentation, the model enhances accuracy and robustness, further strengthening its clinical applicability. Deep learning approaches such as InceptionV3 and Xception have been applied to MRI images to improve brain tumor classification^{25,26}. The ensemble strategy refines and stabilizes predictions, demonstrating superior performance compared to individual deep learning models. Unlike previous studies that focus primarily on tumor severity or malignancy grading, this research emphasizes the precise classification of brain tumors into four histological categories, ensuring a more nuanced and clinically relevant diagnostic framework^{27,28}.

Deep learning approaches for brain tumor detection

DL models have significantly improved the detection and classification of brain tumors, as demonstrated in various studies²⁹. Several approaches have been proposed for brain tumor classification using Magnetic Resonance (MR) FLAIR images, each leveraging different deep learning architectures and optimization techniques.

The author of³⁰ introduced a deep neural network (DNN) technique for automated brain tumor segmentation, utilizing an encoding and decoding mechanism. The CNN was trained to decode spatial data by processing features extracted from encoders. Key characteristics were recovered in the encoder, while the decoder constructed the final probability map using the generated semantic map. Their study evaluated ResNet and DenseNet, demonstrating that ResNet-50 effectively located brain tumors. They also applied block-wise transfer learning, achieving a classification accuracy of 94.82% using 5-fold cross-validation²⁹. This review³¹ shows how deep learning could transform the treatment of neurological diseases and enhance the results of world healthcare systems. It also provides a scheme for including modern technologies into clinical environments. The findings reveal that ethical frameworks, diverse datasets, and understandable models are all absolutely vital. Advanced architectures including transformer-based models and U-Net variants help to ensure more dependability in clinical uses. Deep learning allows tailored treatment plans by automating complex neuroimaging techniques and enhancing diagnostic accuracy. This work³² investigates three advanced deep learning models: the Attention U-Net, the U-Net ++ extension of the U-Net framework, and DWI scan segmentation of the ISLES 2022 dataset to address these challenges.

The authors of^{30,33} assessed their approach using a benchmark dataset of T1-biased contrast-enhanced MRI scans (CEMRI). Their study employed fuzzy logic and a U-NET CNN model to classify and segment brain tumors, reporting an impressive detection accuracy of 97.5%³⁴. In another study, Gurbani et al. used a feature-based approach, extracting wavelet coefficients from images. Their research highlighted the advantages of wavelet transformations over Fourier transformations due to higher temporal resolution, achieving a 91% accuracy rate³⁵. Rajinikanth et al. explored CNN-based classification and segmentation in a computer-aided diagnostic (CADD) system. Their evaluation of various CADD programs included an SVM model trained using a 10-fold cross-validation approach, which demonstrated 97% classification success. However, CNN-based models such as DenseNet201 and InceptionV3, despite their strong performance, faced limitations due to a relatively small dataset, achieving 89% accuracy³⁶. Predicting object instances and their matching per-pixel segmentation masks is a difficult computer vision challenge using instance segmentation³⁷. Object detection and semantic segmentation thus are merged. It points out and names every object seen in a picture. The mask RCNN model is one often used framework that has seen several iterations in aim of improving instance segmentation. Using ensemble pre-training models including support vector machines and feature extraction networks like DenseNet-121 and EfficientNet-B5, the proposed method approaches data classification. Using a sophisticated meta-heuristic approach, the pre-trained CNN hyperparameters were optimized in order to enhance performance. Tested on the INbreast dataset, the EfficientNet-B5 model proved effective in the classification of breast cancer with an accuracy of 99.9%, specificity of 99.8%, precision of 99.1%, and an area under the ROC curve (AUC) of 1.0.

The author of³⁸ proposed a classification model using VGG-16 trained on the BRATS dataset to automatically classify brain tumors. Similarly, author³⁹ introduced a deep convolutional neural network (DCNN) model based on AlexNet, achieving up to 90% accuracy in MRI-based tumor classification. Meanwhile, the author of⁴⁰ employed deep learning techniques for tumor detection and diagnosis, evaluating their approach on the BRATS2014 dataset, which contains multimodal MRI data. Their model efficiently processed 254 volumes of multimodal data, requiring only 13 s per volume for analysis. The authors designed an artificial neural network (ANN) model, leveraging MRI data from 155 healthy and 98 tumor-positive cases, demonstrating high detection accuracy. A novel deep learning approach known as Dolphin-SCA has also been explored for brain tumor classification. Inspired by the echolocation abilities of dolphins, this technique integrates a deep neural network with statistical power characteristics and Local Directional Patterns (LDP). Studies implementing Dolphin-SCA reported an 81.6% accuracy rate. Additionally, Maqsood et al. used a combination of support vector machines (SVMs) and deep neural networks (DNNs) to classify multimodal brain tumors with 96% accuracy. Various CNN architectures were also evaluated for their ability to accurately locate and categorize tumors, further validating the effectiveness of deep learning models in tumor classification^{41,42}.

Beyond classification, recent studies have introduced colorization techniques to enhance grayscale MRI scans. These methods use connected components and index-based colorization algorithms to distinguish different tissue types. By adjusting brightness and pixel matrix values, machine learning techniques inject color into monochromatic MRI images, improving contrast and segmentation accuracy. Research has demonstrated that integrating morphological operation-based region of interest (ROI) selection with watershed-based segmentation enhances the identification of distinct tissue types. Following segmentation, luminance-based colorization techniques improve contrast, peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM), making colorized MRI scans a promising tool for enhancing diagnostic accuracy^{43–45}.

Research gap

Despite advancements in deep learning for brain tumor detection, several limitations persist in existing research. Many studies rely on outdated, private, or small datasets, limiting the generalizability of their models. Some methods, such as those based on AlexNet, GoogleNet, and VGGNet, have achieved notable classification accuracy, reaching 98%, but lack ensemble modeling, which could further improve performance⁸. Similarly, techniques like the Berkeley Wavelet Transformation (BWT) watermarking method have been explored for feature extraction, but these rely on handcrafted features, making them less efficient than modern deep learning approaches. Furthermore, while such methods show promising results, they tend to be computationally slow and are not well-suited for large-scale medical applications¹². Other studies, including those leveraging CNN architectures like VGG16, ResNet-50, and Inception-V3, have demonstrated high classification performance, with reported accuracy levels of 96%, 89%, and 75%, respectively. However, these approaches suffer from lack of ensemble models, reducing their robustness in handling diverse MRI datasets⁴⁶. Research utilizing DenseNet-based models has reached 95% accuracy, but the absence of data augmentation techniques and ensemble learning has limited its ability to generalize well⁴³.

Some researchers have focused on hybrid deep learning techniques, such as combining support vector machines (SVM) with radial basis functions (RBF) for classification. While such approaches have resulted in varied accuracy levels, their effectiveness is constrained by the small number of images in different classes, leading to potential biases and overfitting⁴⁷. A related approach using CNN-LSTM architectures reported 99.1% accuracy, 98.8% precision, and 99.0% F1-score, but the limited dataset size raised concerns about generalizability⁴⁸. Similarly, models based on VGG-16 have shown strong performance, with precision scores exceeding 98%, but remain hindered by the lack of ensemble learning, which could further stabilize results⁴⁹. Several studies have attempted to enhance performance by integrating hybrid CNN models with machine learning techniques, achieving a 96.8% accuracy rate. However, these methods are prone to overfitting when trained on small datasets, making them less effective in real-world clinical settings¹⁵. Some research efforts have applied genetic algorithms (GA) for MRI segmentation, producing 98.6% and 99.1% accuracy on different datasets. While these models demonstrate high precision, they also introduce high computational complexity, making them impractical for routine clinical applications where efficiency is critical⁵⁰. This study addresses these limitations by introducing an ensemble deep learning model that integrates InceptionV3 and Xception architectures. Unlike previous single-model approaches, this method enhances classification accuracy by leveraging multiple deep learning techniques. Additionally, this work focuses on improving generalizability, noise resistance, and computational efficiency, making it more adaptable for real-world clinical applications. By overcoming dataset constraints, reducing overfitting risks, and improving classification robustness, this study provides a more effective and scalable framework for brain tumor detection.

Methodology

Dataset and preprocessing

This study utilizes a publicly accessible brain MRI dataset from Kaggle (<https://www.kaggle.com/code/jaykumar1607/brain-tumor-mri-classification-tensorflow-cnn>) consisting of 2,870 images, including both normal brain scans and those exhibiting tumors. To enhance classification accuracy and improve model generalization, preprocessing techniques such as normalization and augmentation were applied. These steps ensured uniformity in image size, shape, and quality. Deep learning-based preprocessing was performed using MATLAB, guaranteeing consistency in image resolution and overall quality. Data augmentation is essential for strengthening the robustness and generalization capabilities of deep learning models in brain tumor detection. By introducing various transformations to the original images, a more diverse and expansive training set was created. The most frequently employed augmentation techniques include rotation, translation, scaling, flipping, and elastic deformation. Rotation and flipping simulate different anatomical orientations and head positions, while translation and scaling accommodate spatial shifts and resolution variations in MRI scans. Additionally, elastic deformations introduce structural distortions that mimic imaging artifacts or pathological changes in the brain^{45,47}.

To improve the noise robustness and model resilience, the research team applied two key preprocessing techniques: Gaussian smoothing and intensity adjustments. These steps help the model accommodate standard deviations in the forms of brain imaging that happen across institutions and imaging protocols. The 3D image dataset was enlarged to 10,000 samples using an extensive set of transformations: rotating the image from 0 to 25 degrees in 5-degree increments, flipping the image horizontally and vertically, scaling the image up and down, cropping the image to yield a variety of center points, and adding Gaussian noise in a variety of forms to simulate real-world MRI artifacts^{48,51}.

Given the dataset's relatively small size, even after augmentation, several measures were implemented to mitigate overfitting. A k-fold cross-validation strategy was used to ensure consistent model performance across different subsets of data. Additionally, dropout layers (with rates between 0.2 and 0.5) and L2 regularization were incorporated to prevent excessive model complexity. An early stopping mechanism was also applied to terminate training once validation performance plateaued, thereby avoiding overfitting.

Deep learning in brain imaging analysis

DNNs have become fundamental tools in brain imaging analysis, excelling in segmentation, classification, functional analysis, and biomarker identification. These models autonomously learn intricate patterns in brain MRI scans, enabling them to differentiate between healthy and diseased brains with high precision. Additionally, DNNs can reveal dynamic patterns of brain activity and identify quantitative biomarkers linked to neurological and psychiatric disorders⁶. By leveraging large labeled datasets, deep learning models advance our understanding of brain structure and function, contributing to improved diagnosis, treatment planning, and

disease management. Their capability to extract detailed features from medical images makes them particularly effective for brain tumor detection. These networks facilitate precise feature extraction, which is critical for accurate medical diagnoses⁸.

Optimizing deep learning model performance requires careful hyperparameter tuning. Factors such as learning rate, batch size, and regularization techniques significantly influence the model's ability to generalize to new data². Proper tuning strategies ensure optimal predictive performance while maintaining stability⁵. For this study, InceptionV3 and Xception were selected as the primary deep learning architectures due to their proven effectiveness in image classification tasks. These convolutional neural networks (CNNs) employ advanced feature extraction mechanisms, allowing them to capture intricate image structures, making them particularly suitable for medical image analysis⁶. Their efficiency and relatively lightweight nature also facilitate deployment in computationally constrained environments⁷.

By combining InceptionV3 and Xception, this study harnesses the strengths of both models to improve tumor classification accuracy. Each network learns distinct representations of brain tumor characteristics, and their ensemble approach enhances overall classification performance. This integration enables a more comprehensive analysis, achieving superior accuracy and robustness compared to using either model individually. The selection of these architectures is based on their efficiency, effectiveness, and complementary feature extraction capabilities, collectively contributing to high-precision brain tumor detection⁷. Table 1 shows the Actual and Augmented dataset of the brain tumor.

A deep neural network's parameters require a significant volume of data. A data augmentation technique was employed to flip, rotate, and modify the brightness of the training set images to augment the little dataset. Therefore, the model was able to react to new data. The actual images are overall 2870. The DNN interpreted these minor changes as brand-new images, increasing the size of our training set. After augmentation, each class has 2500 images. Overall, there are 10,000 total images.

In brain ensemble model for the tumor detection, two main strategies are key to the control of train/test leakage during data augmentation. Firstly, the data is split (into training and testing sets) before any augmentation takes place. The training data and test data afterward are segregated and the pipeline for the training data and testing data is introduced to apply transformations on the training set only^{7,44}. A degree of randomness is introduced in the process of augmentation using techniques such as random rotations, flips, and alterations in brightness, which in turn both diversify augmented samples and prevent them from being overfitted to transformations. The image data generators with leakage prevention capabilities are considered. In addition to this, a validation set is used to carry out periodic assessments of the model performance throughout the training process, which in turn helps to detect the problem of leakages. These methods ensure that during augmentation there is no accidental leak of the test data information which may help them in generalizing the model for unseen brain scans^{5,6}. Table 2 shows the training options, and the parameters used.

To get the models to perform at their best, a systematic approach is used to fine-tune all the hyperparameters of all the models. Firstly, used the grid search and random search to get the suitable ranges for the four hyperparameters: learning rate, batch size, dropout rate, and the number of epochs. As for the learning rate, some values were as follows 10^{-1} , 10^{-2} , 10^{-3} , and 10^{-4} were tried out to identify the rate at which convergence could be achieved without overshooting the minimum^{49,50}. The numbers of batches including 16, 32, 64 and 128 were used in the experiment to determine the relationship between the training speed and the stability of the model. Dropout rates were set from 0.2 to 0.5 to avoid overfitting while preserving the ability of generalization. Moreover, the number of epochs was also set as the hyperparameter and was changed dynamically according to the early stopping rule to avoid overfitting and extra calculations. More complex methods of optimization including Adam and RMSProp were considered, and the Adam optimizer was selected due to its flexibility and convergence speed. For ensemble learning, the combination weights of Inception-V3 and Xception models were optimized, to achieve the highest level of cooperation between the two structures. To validate the choices of hyperparameters, cross-validation methods were used to reduce chances of over-fitting of the tuned parameters. These procedures of fine-tuning helped in the enhancement of the performance of the proposed XL-TL model^{44,47}.

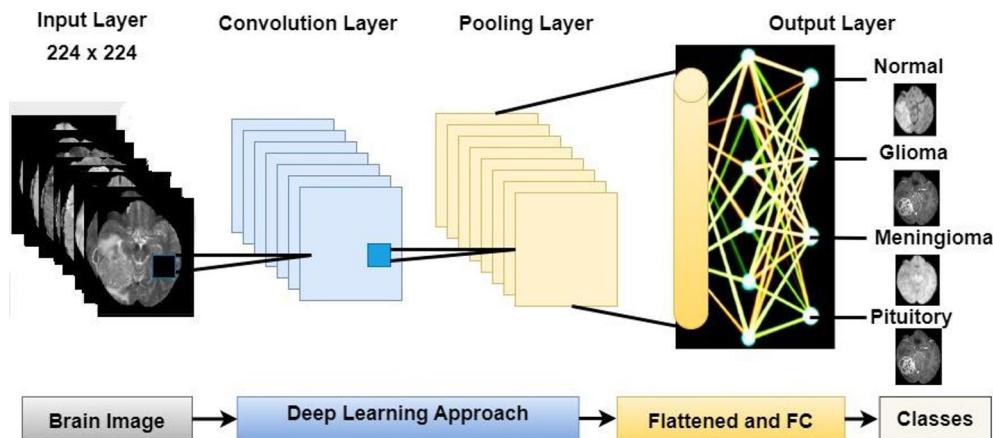
Deep learning algorithm

Deep learning techniques are currently employed to identify and organize brain tumors. The effectiveness of the Xception and Inception-V3 deep learning algorithms is contrasted using the chosen dataset. The inception and Xception deep learning architectures were selected based on their established performance in processing highly complex image information, especially in the medical domain. InceptionV3 could differentiate itself from other related approaches because it utilizes the inception modules, which in return allows the model to extract detailed

Classes	Actual number of images (Kaggle) ⁵⁷	Total number of augmented Images
Actual and augmented dataset		
Normal	395	2500
Glioma	826	2500
Meningioma	822	2500
Pituitary	827	2500
Total	2870	10,000

Table 1. Actual and augmented dataset of brain tumor.

Training options	Reflections/implications
Sizes of Image	227*227*3
Total Number of Epochs	04
Iterations per Epoch	283
Total No of Iterations	50
Initial Learning Rate	0.0001
Momentum Rate	0.9
Type of Solver	SGDM
Environment For Execution	Auto
Size of Minibatch	64
Shuffle	every-epoch
Validation Frequency	1

Table 2. Different training options and its related parameters.**Fig. 2.** The architecture of inception-V3.

multi-scale content of the scan⁴⁵. This differentiation between various tumor classes and the ability to discover even subtle abnormalities is the most fundamental. Then, Xception's distinct depth-wise reliable convolutions offer computational efficiency without quality compromise, this ensures that the model is useful to medical imaging tasks where the data is limited with available computing resources. With the capacity of the Inception and Xception, researchers will not be limited to technical specification but to implement models' predictability and the application in neuro diagnostics and treatments⁴⁷.

Inception-V3

Google published the Inception network, a pre-trained system model, under the Google Net name in 2014. Originally, a 22-layer system with 5 M parameters and 1 × 1, 3 × 3, and 5 × 5 filter sizes was used to extract features at various scales and with maximum pooling.

Fig. 2

Utilizing 1 × 1 filters enhances computational efficiency. In 2015, Google refined the Inception model to develop InceptionV3, optimizing the architecture by reducing the size of convolutional layer boundaries. Specifically, instead of using five 5 × 5 convolutional filters, the model employs two 3 × 3 filters, thereby reducing computational complexity without compromising performance. The InceptionV3 model consists of 48 layers. To mitigate overfitting, we adopted the InceptionV3 model for our experiment and fine-tuned it according to the target dataset. Figure 2 illustrates the architecture of InceptionV3⁵².

Xception

Xception is a convolutional neural network comprising 71 layers, designed to process input images of 224 × 224 dimensions. It supports pre-trained networks from the ImageNet database, which includes over one million images, and operates with input images of 299 × 299 resolution. The architecture is based on depthwise separable convolution layers, totaling 36 layers dedicated to feature extraction. In a standard convolution, the number of parameters is computed as $16 \times 32 \times 3 \times 3$, resulting in 4,608 parameters. Conversely, depthwise separable convolution reduces this count to $16 \times 3 \times 3 + 16 \times 32 \times 1 \times 1 = 656$ parameters. This reduction in parameters enhances the model's efficiency and computational performance. Figure 3 illustrates the architectural design of Xception⁵².

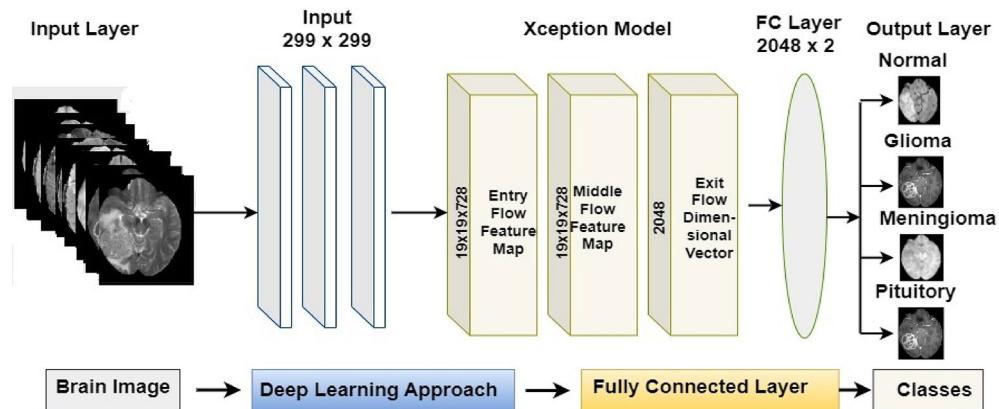


Fig. 3. The architecture of xception.

Ensemble deep learning applied to brain tumor detection

This study presents a model that integrates the outputs of two independently trained neural networks, as illustrated in Fig. 3, to minimize false positives in detection. Pre-trained models, often derived from transfer learning networks, are utilized to reduce errors. By leveraging interconnected networks, the system can achieve optimal results with minimal inaccuracies. Once the data has been preprocessed and structured, a CNN architecture is developed from the learned models^{47,48}.

Moreover, ensemble learning and hyperparameter optimization play a crucial role in enhancing the accuracy and robustness of machine learning models. Their importance is particularly evident in brain imaging analysis, where dataset limitations and subject variability pose significant challenges. Ensemble learning improves predictive accuracy and generalization by combining multiple models, effectively utilizing their diversity. In brain imaging, where data availability is limited and inter-subject variations are high, ensemble methods help mitigate overfitting and enhance model stability⁵².

In the context of brain image classification, particularly for disease detection, integrating classifiers trained on different subsets of data or using varied algorithms can lead to more consistent and reliable results⁴⁹.

Weighted averaging ensemble technique

This study employs a weighted averaging ensemble method to combine the predictions of InceptionV3 and Xception for improved classification accuracy. The final probability for class C_k is computed as follows⁴⁹:

$$p(ck) = \omega_1 p_1(ck) + \omega_2 p_2(ck)$$

Where:

- $p(ck)$ represent the final predicated probability for class C_k
- p_1 and $p_2(ck)$ denote the class probabilities predicated by inception v3 and xception respectively
- ω_1 and ω_2 are the weights assigned to each model's prediction ensuring that:

$$\omega_1 + \omega_2 = 1$$

The ensemble framework leverages the strengths of both models by adjusting their respective weights based on performance metrics obtained from validation data. This method enhances classification reliability, particularly when dealing with complex brain tumor images.

The ensemble structure can be implemented in different ways, with weighted averaging and stacking ensemble techniques being two commonly used methods. In the weighted averaging approach, the probability outputs from InceptionV3 and Xception are combined based on assigned weights, ensuring that the final prediction reflects a balanced contribution from both models. These weights are derived from each model's validation accuracy, optimizing performance through proper calibration.

Alternatively, a stacking ensemble approach trains a meta-learner, such as logistic regression or a neural network, to combine the outputs of InceptionV3 and Xception. This meta-learner is trained using a validation dataset to learn how to effectively merge the predictions from the base models. The final decision is then determined by the output of this meta-learner, offering an alternative strategy for ensemble learning⁵³.

Despite the integration of multiple deep learning models, the ensemble method does not require modifications to the individual architectures of InceptionV3 and Xception. Both models maintain their original configurations, consisting of multiple convolutional, pooling, and fully connected layers. The ensemble process simply aggregates their predictions to leverage their complementary strengths, ultimately improving classification performance without altering the network structures⁵⁴.

Rather than relying solely on a single model, the ensemble approach is specifically designed to enhance feature extraction and classification accuracy. Combining the outputs of different models, whether through weighted averaging or stacking, results in a more robust and accurate brain tumor detection framework. By

merging their complementary capabilities, the ensemble model mitigates individual model weaknesses and achieves superior diagnostic performance^{53,55}.

To further refine classification accuracy, a layered ensemble classifier was developed in response to the previously discussed transfer learning classifiers. The deep CNN ensemble model enhances classification effectiveness by integrating multiple network architectures, enabling the model to extract diverse, architecture-specific patterns. Since deep CNNs inherently introduce randomness, this approach enables the model to capture and learn different neural network design-specific features, ultimately boosting classification performance. The ensemble method simplifies feature extraction, making the boosting approach both more accurate and computationally efficient⁵⁵.

For this study's four-class classification task, the ensemble model was built by combining InceptionV3 and Xception architectures. During hyperparameter tuning, the model was optimized using 64 neurons, with dropout rates of 0.2 and 0.1, and a SoftMax activation function to classify the MRI scans into four distinct tumor types. The neural network ensemble was trained over 50 epochs in 64-batch sizes, with an optimized learning rate of 0.0001%.

Figure 4 illustrates the proposed XL-TL model architecture, which integrates physical, training, and validation layers within the ensemble framework. This ensemble neural network approach enhances model reliability, ensuring greater accuracy and generalizability in brain tumor detection.

Selection and optimization of base models

The selection and grouping of base models in this study followed a systematic approach to ensure optimal performance in brain tumor detection using MRI scans. Various deep learning architectures, including Convolutional Neural Networks (CNNs), ResNet, and DenseNet, were evaluated to leverage their individual strengths while minimizing the risk of overfitting. Each base model underwent performance evaluation using the same dataset, focusing on key metrics such as classification accuracy, sensitivity, and specificity. The models that demonstrated the highest performance were shortlisted for further integration.

Some models were specifically chosen due to their success in similar medical imaging tasks through transfer learning, which enhanced their performance even with smaller datasets, a common constraint in medical imaging. The selected models were then combined using a weighted voting mechanism, where final predictions were determined based on the reliability of each model. Greater influence was assigned to models with higher accuracy, ensuring a balanced and data-driven decision-making process.

In certain configurations, an ensemble stacking approach was employed, where the predictions from base models served as input features for a higher-level meta-model. This meta-model, trained separately, optimized the combination of predictions from the base models, further refining classification accuracy. The optimal ensemble configuration was determined by maximizing overall accuracy, precision, recall, F1-score, and AUC-ROC, while minimizing false positive and false negative rates. To ensure stability, k-fold cross-validation was applied, validating the robustness of the ensemble model across different dataset partitions.

Another key consideration was computational efficiency, as practical clinical applications require fast and resource-efficient models. The performance of different ensemble configurations was assessed to ensure reliable

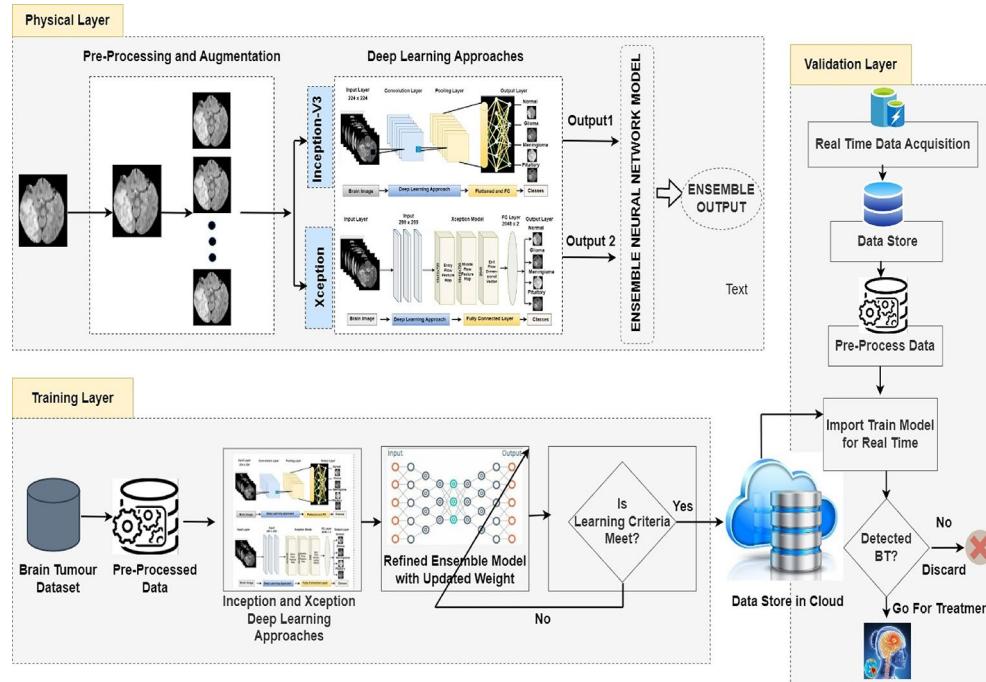


Fig. 4. The architecture of the proposed XL-TL ensemble model.

and real-time MRI processing. Additionally, an error analysis of individual model predictions helped identify specific weaknesses, ultimately leading to the creation of a highly accurate and reliable deep learning ensemble for brain tumor detection.

Challenges and solutions in model implementation

Several challenges were encountered during the experimental implementation of the proposed XL-TL model. One major issue was the limited availability of high-quality datasets, as many prior studies relied on outdated, non-uniform, and handcrafted feature-based datasets, which were naturally less precise than those processed using deep learning techniques⁵⁶. To mitigate this, the study incorporated two levels of data augmentation, significantly enhancing the dataset's diversity and improving the training and testing stages of the model.

Another limitation in previous research was the restricted number of tumor classes, as many studies failed to include new real-time datasets, reducing the generalizability of their findings^{29,50}. To address this, the proposed model integrated an updated and expanded brain tumor dataset, ensuring broader coverage of different tumor types and improving classification robustness.

A further challenge involved finding the optimal combination of InceptionV3 and Xception, balancing their performance and computational complexity³⁹. This was achieved through hyperparameter tuning, where different configurations were tested to maximize accuracy while keeping computational demands manageable. Despite these optimizations, certain constraints remain, such as the reliance on historical MRI data instead of real-time medical imaging and the computational expense of deep learning ensembles.

Nevertheless, the key contributions of this study include the development of an ensemble deep learning model integrating InceptionV3 and Xception, which demonstrated higher accuracy than previous machine learning and deep learning approaches⁵⁶. This ensemble model provides a scalable and efficient solution for brain tumor detection, advancing the field of medical image classification while addressing critical limitations in earlier research.

In order to comprehend the system and how it functions, Table 3 presents the step-by-step process or pseudocode of the suggested model. Ultimately, it uses the evaluation matrix to calculate the precision and accuracy of the suggested model.

Experimental results

Additionally, MATLAB 2021a will be assessed on a powerful computer equipped with an 11 th Generation Intel(R) Core (T.M.) i5-1135G7 @ 2.40 GHz, 8.00 GB of RAM, and a 1 T.B. hard drive. The training and validation distribution of images of the normal brain and brain tumors is displayed in Table 4. Additionally, the confusion matrix of four distinct normal and tumor classes is shown in Table 5 to assess the effectiveness of the deep learning method Inception-V3 during the training and validation stages. Xception is the deeper deep learning method. It was found that 23% at the first iteration and progressively rose with each iteration. It gets close to 96.75% during 50 epochs. The Inception-V3 deep learning method is shown in Table 6.

The MRI dataset utilized in this study consists of 10,000 images, categorized into four distinct classes: Normal, Glioma, Meningioma, and Pituitary tumors. Each class contains 2,500 images, with the dataset split into 7,000 images for training (70%) and 3,000 for validation (30%). Specifically, the Normal class comprises 1,750 training images and 750 validation images, while the Glioma, Meningioma, and Pituitary classes each follow the same distribution, with 1,750 images allocated for training and 750 for validation.

To ensure the dataset's suitability for training machine learning models, several preprocessing techniques were applied. These included resizing all images to a uniform dimension for consistency, normalizing pixel values to a range of 0 to 1 to facilitate seamless model integration, and implementing data augmentation techniques such as rotation and flipping to enhance model robustness. Additionally, noise reduction methods, such as Gaussian blur, were used to improve image clarity, while segmentation techniques were potentially applied to distinguish tumor regions from healthy tissue. Collectively, these preprocessing steps contributed to generating a high-quality dataset, optimizing it for accurate brain tumor classification.

During training, the initial iteration started at 20%, gradually increasing with each iteration. By the 50 th epoch, the model's accuracy approached 98%. The performance of the proposed XL-TL model is evaluated using standard metrics, including precision, F1-score, and recall, as determined by Eqs. (3–11)^{53,57,58,61,62}.

$$\text{Accuracy} = \frac{\frac{S_{ri}}{I_{ri}} + \frac{S_{rk}}{I_{rk}}}{\frac{S_{ri}}{I_{ri}} + \frac{\sum_{j=1}^n (\frac{S_{rj}}{I_{rj}}, j \neq i)}{I_{rj}} + \frac{S_{rk}}{I_{rk}} + \frac{\sum_{l=1}^n (\frac{S_{rl}}{I_{rl}}, l \neq k)}{I_{rk}}}, \quad \text{where } i, j, k, l = 1, 2, 3, \dots, n \quad (3)$$

$$\text{Miss Rate} = \frac{\sum_{l=1}^n \frac{S_{rl}}{I_{rl}}, l \neq k}{\sum_{l=1}^n \frac{S_{rl}}{I_{rl}}, l \neq k + \frac{S_{ri}}{I_{ri}}}, \quad \text{where } i, k, l = 1, 2, \dots, n \quad (4)$$

$$\text{True Positive Rate/Recall} = \frac{\frac{S_{ri}}{I_{ri}}}{\frac{S_{ri}}{I_{ri}} + \frac{\sum_{l=1}^n (\frac{S_{rl}}{I_{rl}}, l \neq k)}{I_{rk}}}, \quad \text{where } i, k, l = 1, 2, 3, \dots, n \quad (5)$$

$$\text{True Negative Rate/Sensitivity} = \frac{\frac{S_{rk}}{I_{rk}}}{\frac{S_{rk}}{I_{rk}} + \frac{\sum_{j=1}^n (\frac{S_{rj}}{I_{rj}}, j \neq 1)}{I_{rj}}}, \quad \text{where } j, k = 1, 2, 3, \dots, n \quad (6)$$

Begin Input: MRI images dataset (D) with labels for tumor types (Normal, Glioma, Meningioma, Pituitary) Output: Classification results and performance metrics (accuracy, precision, recall, F1-score)
Step 1: Data Acquisition - Load dataset D from Kaggle or other sources - Split D into training (D_train) and testing (D_test) sets (e.g., 70% training, 30% testing)
Step 2: Data Augmentation - For each image in D_train: - Apply transformations (rotation, flipping, scaling, brightness adjustment) - Save augmented images to D_augmented - Combine D_train and D_augmented to create the final training set (D_final_train)
Step 3: Preprocessing - For each image in D_final_train and D_test: - Resize images to 227×227 pixels - Normalize pixel values to range [0, 1]
Step 4: Load Pre-Trained Models • Load Inception-V3 and Xception models with pre-trained weights (e.g., from ImageNet) • Remove the final classification layer from both models
Step 5: Transfer Learning - For each model (Inception-V3 and Xception): - Add a new fully connected layer with SoftMax activation for 4-class classification - Freeze initial layers and train only the new layers on D_final_train
Step 6: Ensemble Model Creation - Combine predictions from Inception-V3 and Xception using weighted averaging or stacking - Train a meta-classifier (e.g., logistic regression) on validation set prediction
Step 7: Training - Train the ensemble model on D_final_train - Use Adam optimizer with learning rate = 0.0001 - Apply early stopping to prevent overfitting
Step 8: Evaluation - Evaluate the ensemble model on D_test - Compute performance metrics: accuracy, precision, recall, F1-score
Step 9: Testing - Test the model on unseen MRI images - Generate classification results and confidence scores
Step 10: Output Results - Display classification results and performance metrics - Save the trained model for future use
14. Finish

Table 3. Pseudo code of proposed model XL-TL.

Classes	Total images (10000)	Training (70%)	Validation (30%)
Training and validation			
Normal	2500	1750	750
Glioma	2500	1750	750
Meningioma	2500	1750	750
Pituitary	2500	1750	750
Total	10,000	7000	3000

Table 4. Dataset training and validation distribution.

Inception-V3:Normal			Inception-V3:Meningioma			Inception-V3:Glioma			Inception-V3:Pituitary			
Total Sample N= 1750	Predicted		Total Sample N= 1750	Predicted		Total Sample N= 1750	Predicted		Total Sample N= 1750	Predicted		
	+VE	+VE		+VE	+VE		+VE	+VE		+VE	+VE	
Training												
Actual	+VE 800	TP 760	FN 40	Actual	+VE 800	TP750	FN 50	Actual	+VE 800	TP778	FN 22	Actual
	-VE 950	FP 17	TN 933		-VE 950	FP 16	TN 934		-VE 950	FP 17	TN 933	
validation												
Actual	+VE 350	TP347	FN 03	Actual	+VE 350	TP349	FN 1	Actual	+VE 350	TP348	FN 2	Actual
	-VE 400	FP 12	TN 387		-VE 400	FP 6	TN 394		-VE 400	FP 5	TN 382	

Table 5. Inception-V3 training and validation phase.

Phases		Training				Validation			
Classes		Normal	Glioma	Meningioma	Pituitary	Normal	Glioma	Meningioma	Pituitary
Accuracy		0.96	0.97	0.96	0.97	0.94	0.97	0.96	0.97
Miss Rate		0.04	0.03	0.04	0.03	0.06	0.03	0.04	0.03
Sensitivity		0.937	0.97	0.95	0.97	1.00	0.94	0.94	0.94
Specificity		0.98	0.97	0.98	0.97	0.94	1.00	0.91	1.00
Precision		0.97	0.96	0.98	0.97	0.89	1.00	1.00	1.00
FPR (False Positive Rate)		0.99	0.99	0.99	0.99	0.99	0.00	0.99	0.00
FNR (False Negative Rate)		0.99	0.99	0.99	0.99	0.00	0.99	0.99	0.99

Table 6. Inception-V3 deep learning approach.

$$Precision = \frac{\frac{S_{ri}}{I_{ri}}}{\frac{s_{ri}}{I_{ri}} + \sum_{j=1}^n \frac{(s_{rj}, j \neq 1)}{I_{rj}}}, \text{ Where } i, j = 1, 2, \dots, n \quad (7)$$

$$False Omission Rate = \frac{\sum_{l=1}^n \frac{S_{rl}, l \neq k}{I_{rk}}}{\sum_{l=1}^n \frac{(S_{rl}, l \neq k)}{I_{rk}} + \frac{S_{rk}}{I_{rk}}}, \text{ where } k, l = 1, 2, 3, \dots, n \quad (8)$$

$$False Discovery Rate = \frac{\sum_{j=1}^n \frac{S_{rj}, j \neq i}{I_{rj}}}{\sum_{j=1}^n \frac{(S_{rj}, j \neq i)}{I_{rkj}}}, \text{ where } i, j = 1, 2, 3, \dots, n \quad (9)$$

$$F_{0.5} \text{ score} = 1.25 \times Precision \times \frac{Recall}{0.25 \times Precision + Recall} \quad (10)$$

$$F_1 \text{ score} = 2 \times Precision \times \frac{Recall}{Precision + Recall} \quad (11)$$

The false negative rate (FNR) represents the proportion of true positive cases that a model incorrectly classifies as negative. It is calculated as the ratio of false negatives to the total number of true positives. Similarly, the false positive rate (FPR) measures the proportion of negative instances that are mistakenly classified as positive by the model. In other words, it represents the percentage of incorrect positive classifications relative to the total number of actual negative cases³⁸.

Comparative analysis

This research incorporates updated datasets, decision-level fusion, and real-time data augmentation. The findings indicate that, in terms of both speed and accuracy, transfer learning surpasses traditional machine learning and conventional deep learning techniques. This research has also allowed updated datasets, decision-level fusion, and real-time data augmentation. The study shows that in speed and accuracy, transfer learning outperforms machine learning and more traditional deep learning methods. Figure 5 shows the Accuracy and Loss graph of Inception-V3 and Figs. 6 and 7 show the Graphical Performance Evaluation of Inception-V3.

The accuracy and loss graph of Xception is displayed in Fig. 8. Figures 9 and Fig. 10 depict the Graphical Performance Evaluation of Xception while Fig. 11 displays the Accuracy and Loss graph of an ensemble model,

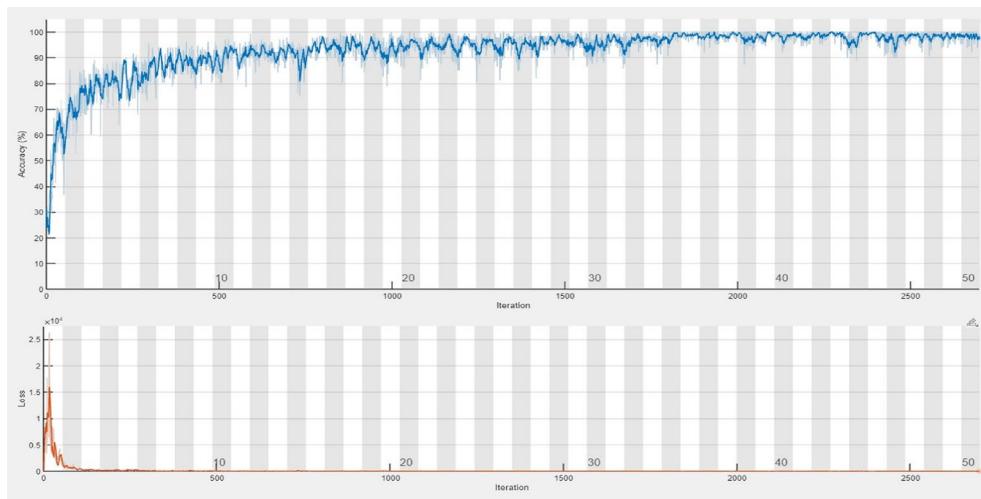


Fig. 5. Accuracy and loss graph of inception-V3.

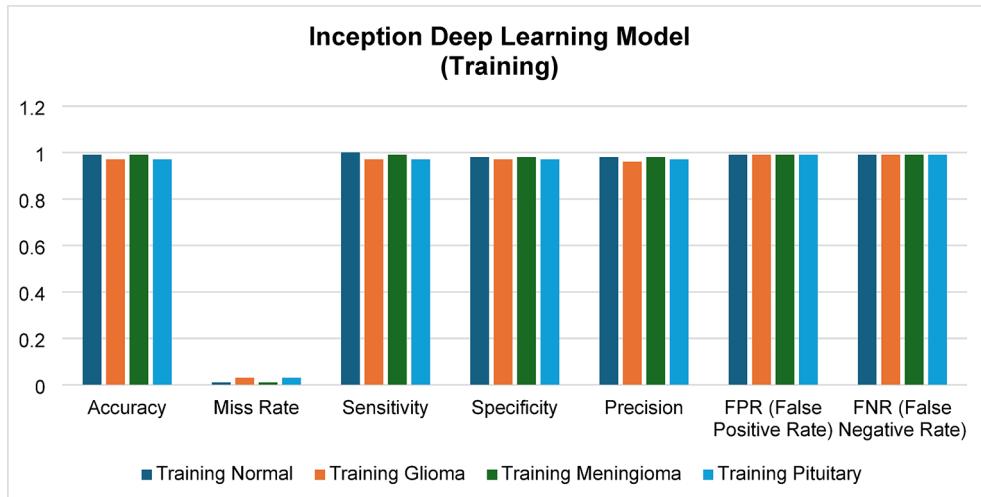


Fig. 6. Graphical performance evaluation of inception-V3 (Training).

whereas. The graphical performance evaluation of the ensemble technique is shown in Figs. 12 and in Fig. 13. The proposed XL-TL model's graphical performance evaluation is shown in Fig. 14. The Xception deep learning approach's confusion matrix computations are shown in Table 7, and the approach's performance evaluation is displayed in Table 8. The Xception deep learning approach's confusion matrix computation is shown in Table 9, and the ensemble deep learning model's performance evaluation is shown in Table 10. The correctness of the suggested XI-Tl model is shown in Table 11.

Discussion

Performance evaluation and comparative analysis

The performance evaluation of the proposed XL-TL ensemble model reveals its effectiveness in brain tumor classification compared to standalone models. The individual models, Inception-V3 and Xception, demonstrated strong accuracy in detecting brain tumors from MRI scans. Specifically, the training and validation accuracy of Inception-V3 were 96.75% and 96.5%, respectively, indicating robust generalization with a minimal accuracy gap of 0.25%. This suggests that Inception-V3 is well-suited for tumor detection but does not surpass the highest accuracy levels observed in other architectures. Similarly, the Xception model exhibited slightly superior performance, achieving training and validation accuracy of 98% and 97.5%, respectively. The marginally higher performance can be attributed to Xception's ability to extract intricate patterns within MRI images while maintaining a small training-to-validation accuracy gap of 0.5%, signifying minimal overfitting. The ensemble model (Inception-V3 + Xception) outperformed both individual architectures, reaching a training accuracy of 98.5% and validation accuracy of 98.3%. The ensemble approach combines distinct feature representations from both models, leading to improved generalization and reduced variability. The model's minimal accuracy difference

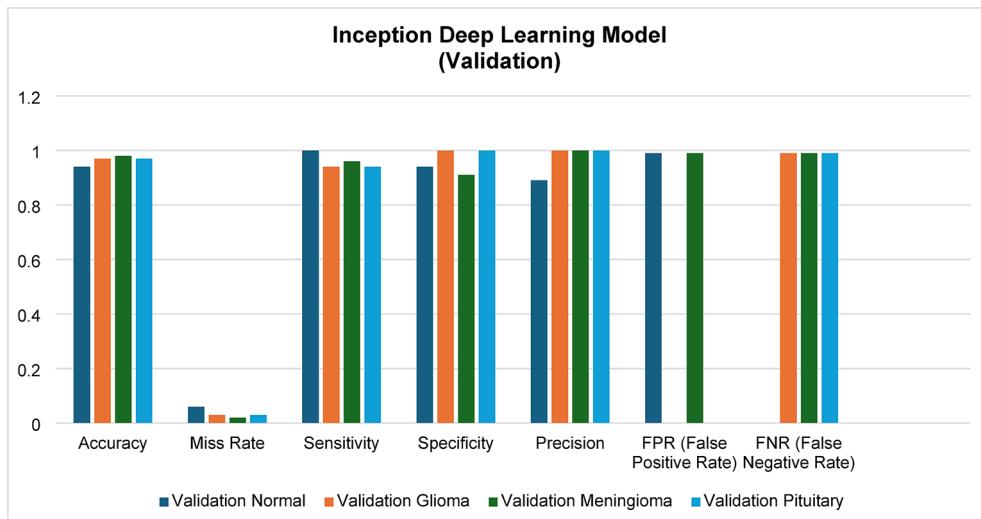


Fig. 7. Graphical performance evaluation of inception-V3 (Validation).

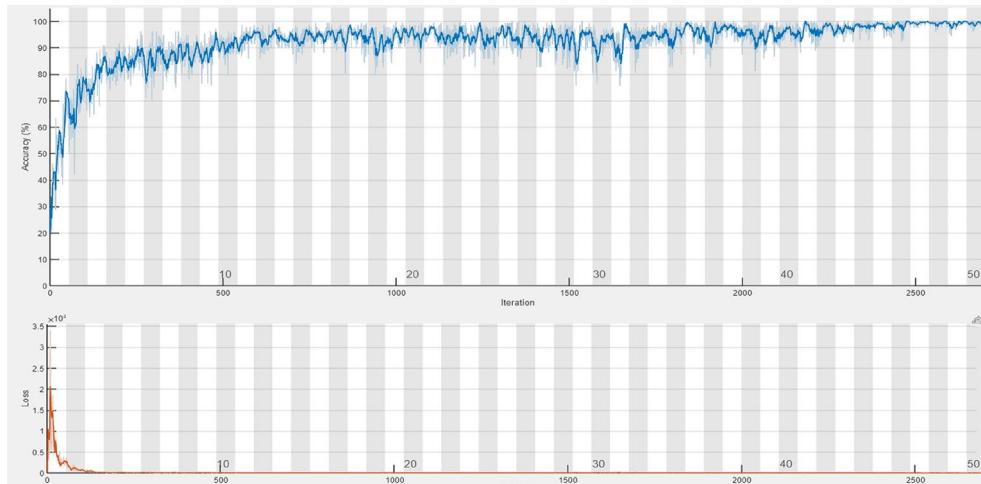


Fig. 8. Accuracy and loss graph of inception-V3.

of 0.2% between training and validation confirms its reliability and strong cross-validation performance, making it an ideal solution for critical medical applications.

Clinical implications

The accuracy and precision of the ensemble model hold significant implications for clinical decision-making, particularly in brain tumor diagnosis. High accuracy directly reduces false positives and false negatives, minimizing misdiagnoses that could otherwise lead to incorrect treatment plans. For instance, in detecting rare or complex brain tumors, the ensemble model mitigates errors that individual models might introduce. This ensures timely and accurate diagnosis, allowing for immediate medical intervention. Moreover, early-stage tumor detection facilitated by this model significantly improves patient survival rates, particularly when tumor growth can be controlled or removed at an early phase. Additionally, the ensemble method enhances tumor classification precision, distinguishing between benign and malignant tumors with higher accuracy. This aspect is critical for avoiding unnecessary invasive procedures on benign tumors while ensuring aggressive malignancies are promptly treated. For instance, the model effectively differentiates between benign meningiomas and aggressive glioblastomas, aiding clinicians in devising optimal treatment strategies and reducing patient risk.

Comparison with recent work

A comparative evaluation against recent research underscores the superiority of the proposed XL-TL model in brain tumor classification. The model's training and testing accuracies of 98.3% and 98.5%, respectively, exceed those reported in prior studies. The authors of 63,64 employed ensemble techniques but achieved slightly lower accuracies of 98.4% and 97.71%. The authors of 47,48, who relied on single-model architectures, reported only

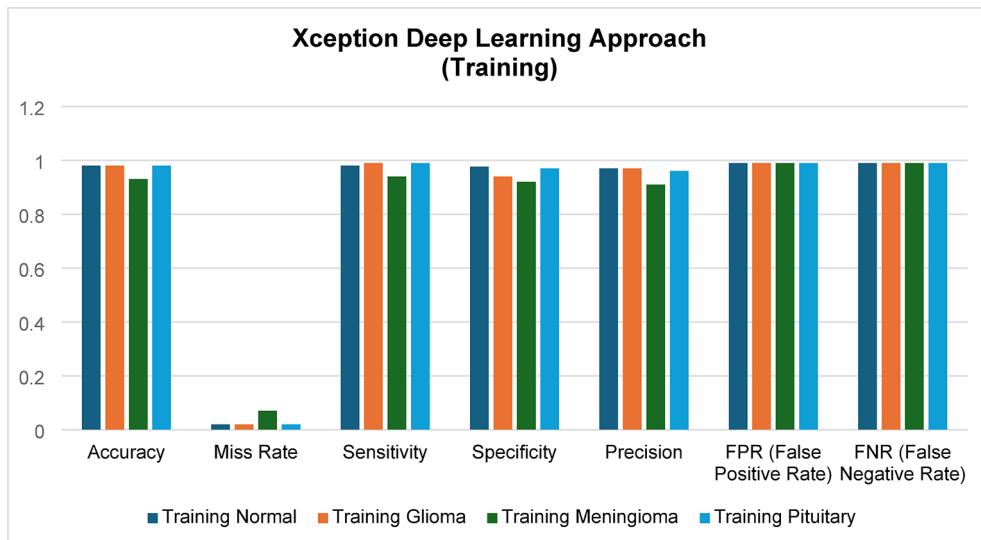


Fig. 9. Graphical performance evaluation of xception (Training).

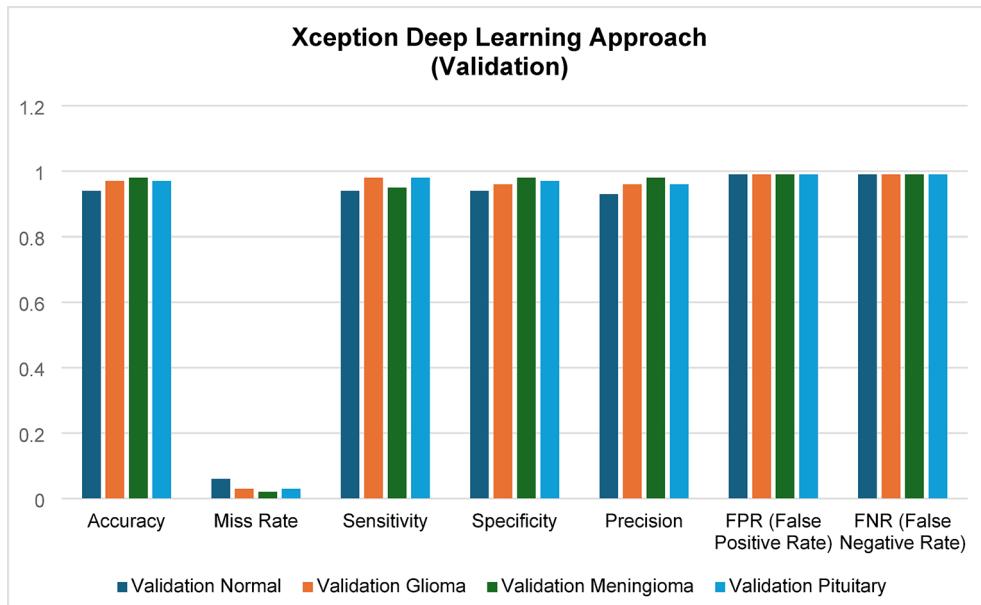


Fig. 10. Graphical performance evaluation of xception (Validation).

95% accuracy due to the absence of ensemble learning and data augmentation. Traditional machine learning approaches such as SVM, KNN, and DT explored by⁶⁵ yielded significantly lower accuracy, ranging between 60 and 63%, highlighting the limitations of conventional classifiers. The XL-TL model's ability to integrate ensemble learning and data augmentation effectively enhances its classification capabilities, setting it apart from previous methodologies. Table 12 shows the comparative result of the proposed XL-TL model.

Limitations

This proposed ensemble deep learning model had a few limitations that should be noted. The original size of the dataset, and even after augmentation, is generally much smaller than what one would obtain in a deep learning architecture. This trade-off increases the risk of overfitting, which can limit generalization on unseen data. Though this method improves accuracy, it needs more resources to run as multiple models are used rather than a single model. This may impede clinical deployment in real time, especially in resource-constrained settings. The dataset on which they are trained may lack the diversity of brain tumor cases seen in actual clinical practice, which could limit generalizability. As a black-box model, it is difficult for clinicians to know how decisions are made (and therefore to interpret decisions), making direct implementation in the clinic difficult. The model

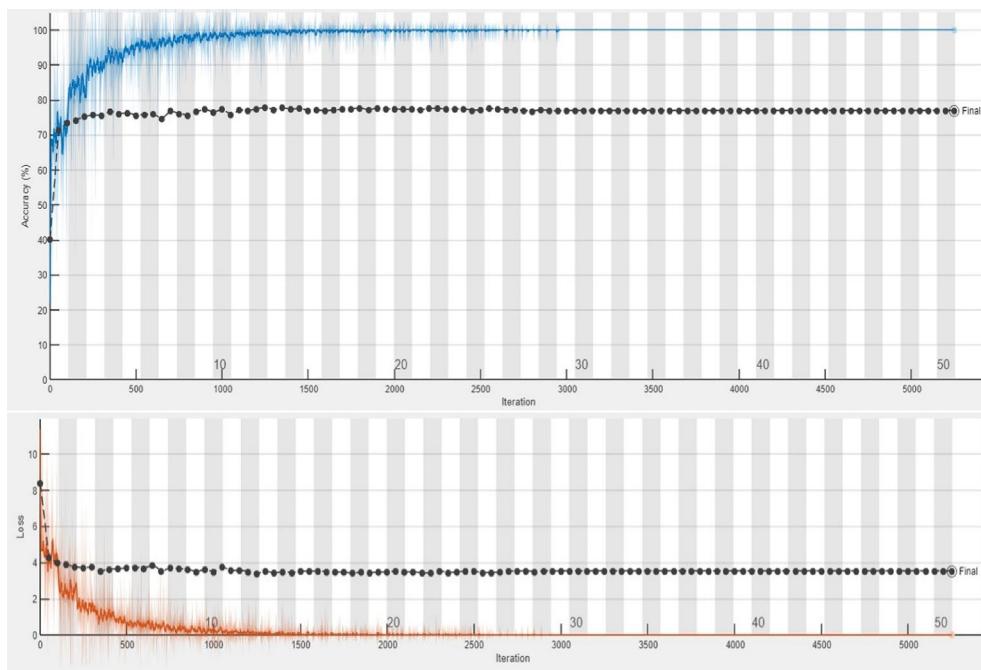


Fig. 11. Accuracy and loss graph of an ensemble model.

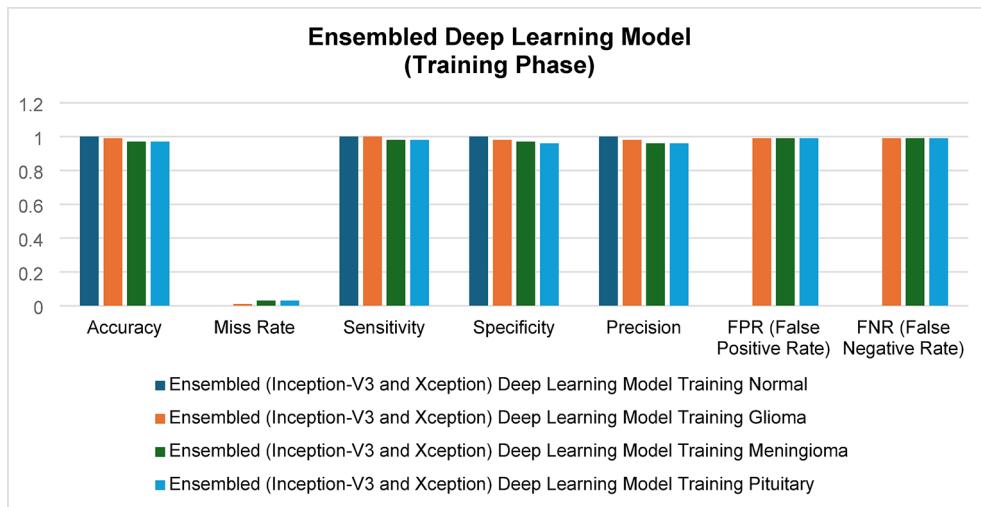


Fig. 12. Graphical performance evaluation of ensemble approach (Training).

has been validated on MRI datasets but requires extensive real-time testing and clinical validation to assess its robustness in live medical environments.

Conclusion Recapitulation

In this study, a new method based on an ensemble deep learning model that combines Inception-V3 and Xception is proposed for brain tumor detection and classification. By combining the complementary advantages of the two models, the proposed method showed a considerable enhancement about the state-of-the-art approaches. The final ensemble approach with a training accuracy of 98.5% and a validation accuracy of 98.3% excellently outperformed both standalone deep learning models and traditional machine learning techniques through proper training and validation. Our results indicate that the ensemble model outperformed in precision, sensitivity and specificity, which justifies the use of the ensemble model in clinical implementations. The proposed method provides an efficient and accurate framework for brain tumor diagnosis by improving the generalization and minimizing the variability across the classifications, resulting in better and faster medical action that can enhance the outcome for the patients as well. The ensemble model outperformed previous

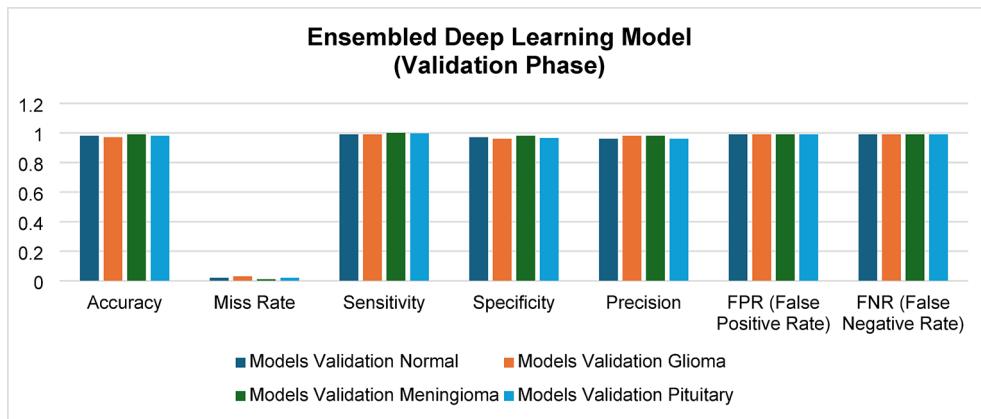


Fig. 13. Graphical performance evaluation of ensemble approach (Validation).

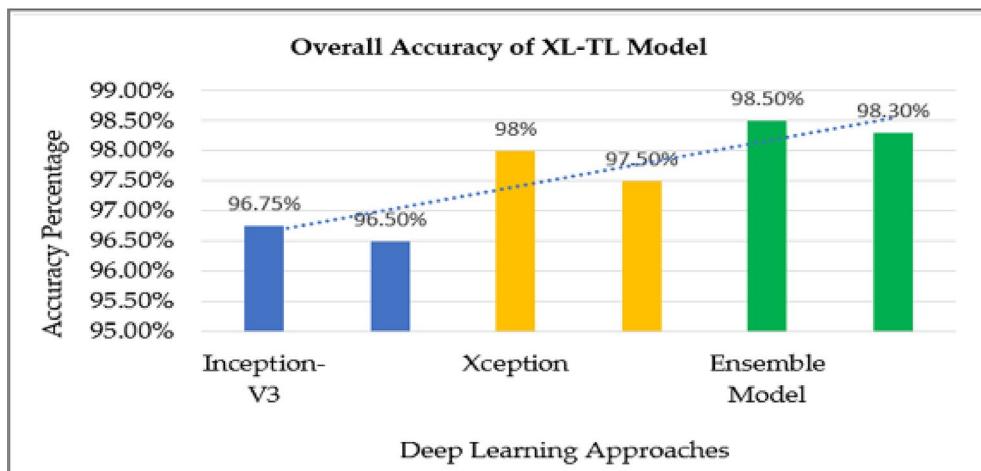


Fig. 14. Graphical performance evaluation of proposed XL-TL model.

Ensemble model: normal			Ensemble Model: Meningioma				Ensemble Model: Glioma				Ensemble Model: Pituitary				
Total SampleN= 750	Predicted		Total SampleN= 750	Predicted		Total SampleN= 750	Predicted		Total SampleN= 750	Predicted		Total SampleN= 750	Predicted		
	+VE	+VE		+VE	+VE		+VE	+VE		+VE	+VE		+VE	+VE	
Training															
Actual	+VE 800	TP 787	FN 13	Actual	+VE 800	TP752	FN 48	Actual	+VE 800	TP791	FN 9	Actual	+VE 800	TP794	FN 6
	-VE 950	FP 22	TN 928		-VE 950	FP 74	TN 876		-VE 950	FP 26	TN 924		-VE 950	FP 29	TN 921
validation															
Actual	+VE 350	TP 331	FN 19	Actual	+VE 350	TP 343	FN 17	Actual	+VE 350	TP 344	FN 6	Actual	+VE 350	TP 342	FN 8
	-VE 400	FP 26	TN 374		-VE 400	FP 8	TN 392		-VE 400	FP 16	TN 384		-VE 400	FP 14	TN 386

Table 7. Confusion matrix calculations of xception deep learning approach.

studies and validated the effectiveness of the ensemble model through a comparative analysis with the recent studies. The model's robustness in processing various types of MRI scans was aided by the application of data augmentation techniques and hyperparameter tuning. Moreover, the importance of integrating various deep learning architectures was highlighted in the study as a way to alleviate the shortcomings of sole models producing a well-optimized diagnostic tool.

Phases		Training				Validation			
Classes		Normal	Glioma	Meningioma	Pituitary	Normal	Glioma	Meningioma	Pituitary
Accuracy		0.98	0.98	0.93	0.98	0.94	0.97	0.98	0.97
Miss Rate		0.02	0.02	0.07	0.02	0.06	0.03	0.02	0.03
Sensitivity		0.98	0.99	0.94	0.99	0.94	0.98	0.95	0.98
Specificity		0.976	0.94	0.92	0.97	0.94	0.96	0.98	0.97
Precision		0.97	0.97	0.91	0.96	0.93	0.96	0.98	0.96
FPR (False Positive Rate)		0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
FNR (False Negative Rate)		0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99

Table 8. Performance evaluation of xception deep learning approach.

Ensemble model: normal			Ensemble Model: Meningioma				Ensemble Model: Glioma				Ensemble Model: Pituitary				
Total SampleN= 750	Predicted		Total SampleN= 750	Predicted		Total SampleN= 750	Predicted		Total SampleN= 750	Predicted		Total SampleN= 750	Predicted		
	+VE	+VE		+VE	+VE		+VE	+VE		+VE	+VE		+VE	+VE	
Training															
Actual	+VE 800	TP 800	FN 0	Actual	+VE 800	TP760	FN 20	Actual	+VE 800	TP800	FN 0	Actual	+VE 800	TP780	FN 16
	-VE 950	FP 0	TN 950		-VE 950	FP 32	TN 918		-VE 950	FP 17	TN 933		-VE 950	FP 36	TN 924
Validation															
Actual	+VE 350	TP 348	FN 2	Actual	+VE 350	TP 350	FN 0	Actual	+VE 350	TP 346	FN 4	Actual	+VE 350	TP 349	FN 01
	-VE 400	FP 13	TN 387		-VE 400	FP 7	TN 393		-VE 400	FP 18	TN 382		-VE 400	FP 14	TN 386

Table 9. Confusion matrix calculation of Xception deep learning approach.

Models		Ensembled (Inception-V3 and Xception) Deep Learning Model							
Phases		Training				Validation			
Classes		Normal	Glioma	Meningioma	Pituitary	Normal	Glioma	Meningioma	Pituitary
Accuracy		1.00	0.99	0.97	0.97	0.98	0.97	0.99	0.98
Miss Rate		0.00	0.01	0.03	0.03	0.02	0.03	0.01	0.02
Sensitivity		1.00	1.00	0.98	0.98	0.99	0.99	1.00	0.997
Specificity		1.00	0.98	0.97	0.96	0.97	0.96	0.98	0.965
Precision		1.00	0.98	0.96	0.96	0.96	0.98	0.98	0.96
FPR (False Positive Rate)		0.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99
FNR (False Negative Rate)		0.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99

Table 10. Performance evaluation of ensemble deep learning model.

Model	Phases	Accuracy
Inception-V3	Training	96.75%
	Validation	96.5%
Xception	Training	98%
	Validation	97.5%
Ensemble Model (Inception-V3 Xception)	Training	98.5%
	Validation	98.3%

Table 11. Accuracy of the proposed Xl-Tl model.

Future directions

Future research should focus on addressing the current limitations of the proposed model to enhance its applicability in real-world clinical settings. One potential avenue is expanding the dataset by incorporating

Study	Year	Data Augmentation	Dataset	Ensemble used	Deep learning models	Performance metrics (accuracy, etc.)	Key contributions	Limitations
8	2020	No	Private	No	Alexnet, googlenet, vggnet	Accuracy: 96%	Data augmentation applied	No ensemble model, limited dataset
12	2021	No	Private	No	Berkeley wavelet transformation (bwt)	Not specified	Watermarking technique for feature extraction	Small dataset, handcrafted features, slow
46	2020	No	Brain tumor dataset	No	Cnn (vgg16, resnet-50, inception-v3)	Vgg16: 96%, resnet-50: 89%, inception-v3: 75%	Cnn-based classification	No ensemble models, lower accuracy
43	2022	No	Private	No	Densenet	Accuracy: 95%	Deep learning applied	No data augmentation, lacks ensemble approach
44	2021	No	Brain tumor dataset	Yes	Scnn, vgg16	Accuracy: 97.77%	Addressed overfitting & imbalance dataset	Limited comparative analysis
59	2023	No	Brain tumor dataset	Yes	Not specified	Accuracy: 98.4%	Applied ensemble learning	Lacks computational efficiency analysis
61	2024	No	Brain tumor dataset	No	Svm, knn, decision tree, naïve bayes	Svm: 62%, knn & dt: 63%, naïve bayes: 60%	Classical ml approach	Low accuracy, no deep learning applied
60	2024	No	Brain tumor dataset	Yes	Feature ensemble, stacking ensemble	Feature ensemble: 97.71%, stacking: 97.40%	Applied feature ensemble learning	No real-time efficiency evaluation
62	2024	No	Brats dataset	No	Custom model	Accuracy: 98%	Proposed a new model	No ensemble learning
Proposed model (xl-tl)	2024	Yes	Brain tumor dataset	Yes	Ensemble (inceptionv3 + xception)	Training: 98.3%, testing: 98.5%	High accuracy, ensemble approach applied	Computational efficiency not analyzed

Table 12. Comparative result of the proposed Xl-Tl model.

more diverse and high-quality labeled MRI scans, which will improve the model's generalization capability. Additionally, optimizing computational efficiency remains a critical challenge, as deep learning ensembles require substantial processing power. Future efforts should explore lightweight model architectures or knowledge distillation techniques to enable real-time deployment in resource-constrained environments. Another promising direction is integrating multi-modal imaging data, such as CT and PET scans, alongside MRI scans. This fusion of imaging modalities can provide a more comprehensive representation of brain tumor characteristics, further refining classification accuracy. Furthermore, the development of explainable AI techniques, such as Grad-CAM and SHAP, will enhance model interpretability, allowing clinicians to better understand and trust the model's decision-making process. Addressing the issue of imbalanced datasets remains crucial, as certain tumor types may be underrepresented in the available data. Implementing advanced techniques for data balancing and augmentation will ensure that the model performs well across all tumor categories. Moreover, incorporating clinical data, including patient history, genetic markers, and treatment response, can enhance the predictive power of deep learning models, leading to more personalized and precise diagnosis. Developing AI-driven models that not only detect tumors but also predict treatment outcomes will be an essential step in the future of brain tumor research. Personalized treatment plans based on tumor classification and response prediction could significantly improve patient care. Additionally, collaborations between AI researchers and medical professionals will be necessary to ensure that AI-driven diagnostic tools align with clinical workflows and regulatory standards. By addressing these challenges and advancing deep learning methodologies, the proposed approach can be further refined to support accurate, interpretable, and efficient brain tumor diagnosis, ultimately contributing to improved healthcare outcomes.

Data availability

The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding authors.

Received: 4 January 2025; Accepted: 21 April 2025

Published online: 29 April 2025

References

1. Bush, N. A. O., Chang, S. M. & Berger, M. S. Current and future strategies for treatment of glioma. *Neurosurg. Rev.* **40**, 1–14 (2017).
2. Narkhede Sachin, G., Khairnar, V. & Kadu, S. Brain tumor detection based on mathematical analysis and symmetry information. *Int. J. Eng. Res. Appl.* **4** (2), 231–235 (2014).
3. Amran, G. A. et al. Brain tumor classification and detection using hybrid deep tumor network. *Electronics* **11** (21), 3457 (2022).
4. Vimala, M. et al. Efficient GDD feature approximation based brain tumour classification and survival analysis model using deep learning. *Egypt. Inf. J.* **28**, 100577 (2024).
5. DeAngelis, L. M. Brain tumors. *N. Engl. J. Med.* **344** (2), 114–123 (2001).
6. Louis, D. N. et al. The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* **131**, 803–820 (2016).
7. Vankdothu, R., Hameed, M. A. & Fatima, H. A brain tumor identification and classification using deep learning based on CNN-LSTM method. *Comput. Electr. Eng.* **101**, 107960 (2022).
8. Rehman, A. et al. A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circuits Syst. Signal. Process.* **39** (2), 757–775 (2020).
9. Wang, Y. et al. Automatic tumor segmentation with deep convolutional neural networks for radiotherapy applications. *Neural Process. Lett.* **48**, 1323–1334 (2018).

10. Mubashar, M. et al. R2U++: a multiscale recurrent residual U-Net with dense skip connections for medical image segmentation. *Neural Comput. Appl.* **34** (20), 17723–17739 (2022).
11. Ayadi, W. et al. Deep CNN for brain tumor classification. *Neural Process. Lett.* **53**, 671–700 (2021).
12. Ray, A. K. & Thethi, H. P. *Comparative approach of MRI-Based brain tumor segmentation and classification using genetic algorithm. J. Digit. Imaging (Online)*, **31**(4). (2018).
13. Zikic, D. et al. *Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR*. in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2012: 15th International Conference, Nice, France, October 1–5, Proceedings, Part III* 15. 2012. Springer. (2012).
14. Kaya, I. E. et al. PCA based clustering for brain tumor segmentation of T1w MRI images. *Comput. Methods Programs Biomed.* **140**, 19–28 (2017).
15. Srinivasan, S. et al. A hybrid deep CNN model for brain tumor image multi-classification. *BMC Med. Imaging.* **24** (1), 21 (2024).
16. Ahmad, I. et al. *A review of artificial intelligence techniques for selection & evaluation*. in *IOP Conference Series: Materials Science and Engineering*. IOP Publishing. (2020).
17. Soomro, T. A. et al. Image segmentation for MR brain tumor detection using machine learning: a review. *IEEE Rev. Biomed. Eng.* **16**, 70–90 (2022).
18. Pacal, I. et al. Enhancing EfficientNetv2 with global and efficient channel attention mechanisms for accurate MRI-Based brain tumor classification. *Cluster Comput.* **27** (8), 11187–11212 (2024).
19. Pacal, I. A novel Swin transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images. *Int. J. Mach. Learn. Cybernet.* **15** (9), 3579–3597 (2024).
20. Pacal, I. MaxCerVixT: A novel lightweight vision transformer-based approach for precise cervical cancer detection. *Knowl. Based Syst.* **289**, 111482 (2024).
21. Naser, M. A. & Deen, M. J. Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images. *Comput. Biol. Med.* **121**, 103758 (2020).
22. Khalighi, S. et al. Artificial intelligence in neuro-oncology: advances and challenges in brain tumor diagnosis, prognosis, and precision treatment. *NPJ Precision Oncol.* **8** (1), 80 (2024).
23. Devi, P. J. et al. *Exploring Deep Learning-Based MRI Radiomics for Brain Tumor Prognosis and Diagnosis*. in *3rd Asian Conference on Innovation in Technology (ASIANCON)*. 2023. IEEE. (2023).
24. Aamir, M. et al. Brain tumor classification utilizing deep features derived from high-quality regions in MRI images. *Biomed. Signal Process. Control.* **85**, 104988 (2023).
25. Zhu, Z. et al. *Sparse dynamic volume transunet with multi-level edge fusion for brain tumor segmentation*. *Comput. Biol. Med.*, : p. 108284. (2024).
26. Anaraki, A. K., Ayati, M. & Kazemi, F. Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. *Biocybernetics Biomedical Eng.* **39** (1), 63–74 (2019).
27. Zhang, X. et al. *Shufflenet: An extremely efficient convolutional neural network for mobile devices*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. (2018).
28. Abiwinanda, N. et al. Brain tumor classification using convolutional neural network. in *World Congress on Medical Physics and Biomedical Engineering 2018*: June 3–8, 2018, Prague, Czech Republic (Vol. 1). Springer. (2019).
29. Sarkar, A. et al. An effective and novel approach for brain tumor classification using AlexNet CNN feature extractor and multiple eminent machine learning classifiers in MRIs. *J. Sens.* **2023** (1), 1224619 (2023).
30. Swati, Z. N. K. et al. Brain tumor classification for MR images using transfer learning and fine-tuning. *Comput. Med. Imaging Graph.* **75**, 34–46 (2019).
31. Bayram, B. et al. *A systematic review of deep learning in MRI-based cerebral vascular occlusion-based brain diseases*. *Neuroscience*, (2025).
32. İnce, S. et al. *U-Net-Based models for precise brain stroke segmentation*. *Chaos Theory Appl.* **7**(1): pp. 50–60 .
33. Zeineldin, R. A. et al. DeepSeg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images. *Int. J. Comput. Assist. Radiol. Surg.* **15** (6), 909–920 (2020).
34. Maqsood, S., Damasevičius, R. & Shah, F. M. *An efficient approach for the detection of brain tumor using fuzzy logic and U-NET CNN classification*. in *Computational Science and Its Applications-ICCSA : 21st International Conference, Cagliari, Italy, September 13–16, 2021, Proceedings, Part V* 21. 2021. Springer. (2021).
35. Gurbinā, M., Lascu, M. & Lascu, D. *Tumor detection and classification of MRI brain image using different wavelet transforms and support vector machines*. in *42nd International Conference on Telecommunications and Signal Processing (TSP)*. 2019. IEEE. 2019. IEEE. (2019).
36. Rajinikanth, V., Kadry, S. & Nam, Y. Convolutional-neural-network assisted segmentation and SVM classification of brain tumor in clinical MRI slices. *Inform. Technol. Control.* **50** (2), 342–356 (2021).
37. Hassan, E., El-Rashidy, N. & Talaa, M. Mask R-CNN models. *Nile J. Communication Comput. Sci.* **3** (1), 17–27 (2022).
38. Amin, J. et al. Brain tumor detection and classification using machine learning: a comprehensive survey. *Complex. Intell. Syst.* **8** (4), 3161–3183 (2022).
39. Badjie, B. & Ülker, E. D. A deep transfer learning based architecture for brain tumor classification using MR images. *Inform. Technol. Control.* **51** (2), 332–344 (2022).
40. Dvořák, P. & Menze, B. *Local structure prediction with convolutional neural networks for multimodal brain tumor segmentation*. in *Medical computer vision: Algorithms for big data: International workshop, MCV 2015, held in conjunction with MICCAI 2015, Munich, Germany, October 9, revised selected papers* 18. 2016. Springer. (2015).
41. Chauhan, P. et al. *Analyzing Brain Tumour Classification Techniques: A Comprehensive Survey* (IEEE Access, 2024).
42. Maqsood, S., Damaševičius, R. & Maskeliūnas, R. Multi-modal brain tumor detection using deep neural network and multiclass SVM. *Medicina* **58** (8), 1090 (2022).
43. Sharif, M. I. et al. *A decision support system for multimodal brain tumor classification using deep learning*. *Complex. Intell. Syst.*, : pp. 1–14. (2021).
44. Patil, S. & Kirange, D. Ensemble of deep learning models for brain tumor detection. *Procedia Comput. Sci.* **218**, 2468–2479 (2023).
45. Haq, I. et al. A novel brain tumor detection and coloring technique from 2D MRI images. *Appl. Sci.* **12** (11), 5744 (2022).
46. Khan, H. A. et al. *Brain Tumor Classification in MRI Image Using Convolutional Neural Network* (Mathematical Biosciences and Engineering, 2021).
47. Kang, J., Ullah, Z. & Gwak, J. *MRI-based Brain Tumor Classif. Using Ensemble Deep Features Mach. Learn. Classifiers Sens.*, **21**(6): 2222. (2021).
48. Alsabai, S. et al. Ensemble deep learning for brain tumor detection. *Front. Comput. Neurosci.* **16**, 1005617 (2022).
49. Ganaie, M. A. et al. Ensemble deep learning: A review. *Eng. Appl. Artif. Intell.* **115**, 105151 (2022).
50. Ye, J. et al. *Optimizing the topology of convolutional neural network (CNN) and artificial neural network (ANN) for brain tumor diagnosis (BTD) through MRIs*. *Heliyon*, **10**(16). (2024).
51. Mumtaz Zahoor, M. et al. *A New Deep Hybrid Boosted and Ensemble Learning-based Brain Tumor Analysis using MRI*. arXiv e-prints, : p. arXiv: 2201.05373. (2022).
52. Zahoor, M. M. et al. A new deep hybrid boosted and ensemble learning-based brain tumor analysis using MRI. *Sensors* **22** (7), 2726 (2022).

53. Asif, R. N. et al. Development and validation of embedded device for electrocardiogram arrhythmia empowered with transfer learning. *Comput. Intell. Neurosci.* **2022** (1), 5054641 (2022).
54. Rahman, A. et al. ECG classification for detecting ECG arrhythmia empowered with deep learning approaches. *Comput. Intell. Neurosci.* **2022** (1), 6852845 (2022).
55. Ranjbarzadeh, R. et al. Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools. *Comput. Biol. Med.* **152**, 106405 (2023).
56. Raza, A. et al. A hybrid deep learning-based approach for brain tumor classification. *Electronics* **11** (7), 1146 (2022).
57. Mahjoubi, M. A. et al. Improved multiclass brain tumor detection using convolutional neural networks and magnetic resonance imaging. *Int. J. Adv. Comput. Sci. Appl.* **14** (3), 406–414 (2023).
58. Ullah, N. et al. Enhancing explainability in brain tumor detection: A novel DeepEBTDNet model with LIME on MRI images. *Int. J. Imaging Syst. Technol.* **34** (1), e23012 (2024).
59. Saha, P., Das, R. & Das, S. K. BCM-VEMT: classification of brain cancer from MRI images using deep learning and ensemble of machine learning techniques. *Multimedia Tools Appl.* **82** (28), 44479–44506 (2023).
60. Remzan, N., Tahiry, K. & Farchi, A. Advancing brain tumor classification accuracy through deep learning: Harnessing radimagenet pre-trained convolutional neural networks, ensemble learning, and machine learning classifiers on MRI brain images. *Multimedia Tools Appl.* **83** (35), 82719–82747 (2024).
61. Arowolo, M. O. et al. *Empowering Healthcare with AI: Brain Tumor Detection Using MRI and Multiple Algorithms.* in. *International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG).* 2024. IEEE. (2024).
62. Ranjbarzadeh, R. et al. Comparative analysis of real-clinical MRI and BraTS datasets for brain tumor segmentation. in *IET Conference Proceedings CP887.* IET. (2024).

Acknowledgements

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R235), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Author contributions

Rizwana Naz Asif, Habib Hamam, and Muhammad Tahir Naseem have collected data from different resources and contributed to writing—original draft preparation. Rizwana Naz Asif, Muhammad Tahir Naseem, Munir Ahmad, Tehseen Mazhar and Muhammad Adnan Khan performed formal analysis and simulation; Rizwana Naz Asif, Munir Ahmad, Muhammad Amir Khan, and Tehseen Mazhar performed writing—review and editing; Amal Al-Rasheed, Muhammad Amir Khan, and Muhammad Adnan Khan performed supervision, Muhammad Tahir Naseem, Tehseen Mazhar, Munir Ahmad and Amal Al-Rasheed; drafted pictures and tables, Muhammad Amir Khan and Muhammad Adnan Khan, Habib Hamam performed revisions and improved the quality of the draft. All authors have read and agreed to the published version of the manuscript.

Funding

This manuscript received funding from Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R235), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to T.M., M.A.K. or M.A.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025