DOI: 10.6977/IJoSI.202506_9(3).0007
N. F. Othman, S. W. Kareem//Int. J. Systematic Innovation, 9(3), 109-123 (2025)

Ensemble learning for enhanced brain tumor diagnosis: A new approach for early detection

Nayla Faiq Othman[1]*, Shahab Wahhab Kareem[1,2]

[1]Department of Technical Information Systems Engineering, Erbil Technical Engineering College, Erbil Polytechnic University, Erbil, Iraq

[2]Department of Computer Science, Bayan University, Erbil, Iraq

*Corresponding author E-mail: nayla.othman@epu.edu.iq

(Received 06 January 2025; Final version received 11 March 2025; Accepted 23 April 2025)

Abstract

Brain tumors represent one of the most extreme and complex types of cancer, requiring unique analysis for powerful remedy and management. Accurate and early identification of brain tumors can greatly enhance patient outcomes and decrease mortality. Nowadays, deep learning aids the medical field a lot by diagnosing magnetic resonance imaging images in brain tumors. The potential of deep learning architectures to improve brain tumor diagnosis accuracy was explored in this work. This study evaluated three different convolutional neural network architectures: AlexNet, VGG16, and ResNet18 as an ensemble model. By leveraging the complementary strengths of these models and applying them to a dataset sourced from local hospitals and public repositories, this research aims to address the challenges in accurate and early brain tumor detection. Our ensemble technique achieved excessive accuracy, demonstrating its potential for reliable computer-aided diagnosis (CAD) in medical imaging. However, while the results indicate an improvement in class overall performance, the novelty of this approach is restrained because it builds upon existing methodologies as opposed to offering a completely new framework. The gathered dataset was used to train and test the models. To enhance the dataset's balance and the models' performance, data were collected from Rizgary Hospital (Erbil) and Hiwa Hospital (Slemani), addressing the underrepresentation of cases from the Kurdistan Region of Iraq (KRI). These image enhancement techniques were applied to two categories: normal and abnormal brain tumors. Several brain tumor datasets are available online for the development of CADs, but not KRI cases, which pose challenges in their classification through deep learning models. This study was implemented with Python programming language. Out of the three models, ResNet had the highest accuracy of 98.66%, VGG16 had an accuracy of 97.8%, and AlexNet had an accuracy rate of 97.666%. The ensemble, using both majority voting and weighting voting strategies, achieved an accuracy of 98.33%.

Keywords: Brain Tumor, Computer-aided Diagnosis, Magnetic Resonance Imaging, Transfer Learning

1. Introduction

This document is a template for Microsoft Word. Absolute confidence, human needs, technological paradigms, hospitals, and organizations have all significantly improved in the twenty-first century, with computing and records technology displacing other fields like an octopus (Alfonse & Salem, 2016; Nawaz et al., 2022). The human brain, situated within the skull, is an essential organ responsible for various functions, governed by a network of billions of neurons that coordinate electrical and chemical impulses, influencing our experiences and lives (Mathivanan et al., 2024). The brain is an extraordinary organ, serving as a cornerstone in the domains of cognition, emotion, and humanity. The mind consists of remarkable components, each with distinct functions, exemplifying complexity. The cerebral cortex, a convoluted outer layer, governs awareness, whereas the cerebellum is responsible for balance and coordination. The formation of an abnormal cellular proliferation, manifested as a mass or lump, is referred to as a tumor or neoplasm (Lauko et al., 2022). It is an aberrant proliferation of cells in the essential spinal canal or the brain. Two types of brain tumors exist, classified as benign and malignant. Meningiomas and low-grade gliomas are alternative designations for benign tumors, whereas malignant tumors are referred to as glioblastoma multiforme and high-grade gliomas. The most common type of brain neoplasm is referred to as a malignant tumor (Leena & Jayanthi, 2020). Brain tumors characterized by a homogeneous structure are classified as benign, as they do not include malignant cells. This can be entirely resolved with surgical removal or radiological monitoring, ensuring they do not recur. The life-threatening tumor is a malignant neoplasm characterized by a heterogeneous structure containing the majority of cancerous cells. The treatment for malignant tumors involves chemotherapy, radiation, or a combination of both. Consequently, timely analysis of brain tumors is a crucial element in the advancement of therapy (Harnod et al., 2014). Consequently, it is essential to ascertain the dimensions of the brain tumor and establish the associated grade. Recently, advancements in magnetic resonance imaging (MRI) have significantly improved the detection rate of brain malignancies. It has established a crucial foundation in the study of tumor diagnosis and image registration (El Kader et al., 2021). Tasked with fostering awareness of self and surroundings, including influencing muscular actions (Khairandish et al., 2022). Every responding notion, emotion, and plan is facilitated by the brain. MRI and computed tomography scans are diagnostic modalities utilized to reveal the internal structure of the brain. MRI is advantageous for delineating soft tissues and revealing the internal architecture of the framework. The MRI delineates the contrast between normal and abnormal tissues. This work utilizes MRI data to identify the damaged areas in the brain. MRI utilizes a powerful magnet and radio waves to generate a distinctive image of the patient's internal organs (Praveen & Agrawal, 2016). MRI poses no danger to the human body as it does not involve radiation. It provides statistics regarding anomalous tissues for diagnostic purposes. MRI is non-invasive, making it highly popular among individuals and commonly utilized for assessing tumor size, shape, and type (Ramtekkar et al., 2023).The MRI scans can accurately reflect the brain's structure and function, allowing for multi-angle and multi-modal imaging with minimal harm to the human body. Consequently, it is extensively utilized in the identification of neurological disorders (Fang & Wang, 2022). The amalgamation of deep learning and artificial intelligence has markedly enhanced medical image processing, resulting in substantial progress in the detection, diagnosis, and characterization of diverse medical disorders. This has allowed healthcare providers to make better-informed decisions, especially in the precise classification of cancer kinds, including lung and breast cancer (Siva Raja & Rani, 2020). This integration has led to earlier diagnoses, enhanced treatment decisions, and better patient outcomes. Artificial intelligence is essential in surgical planning, facilitating accurate segmentation of lesion margins and cerebral structures while balancing intervention with the preservation of quality of life (Arumugam et al., 2024). It forecasts problems, recurrence rates, and therapeutic responses, directing appropriate follow-up tactics and facilitating personalized patient management through customized screening protocols. Transfer learning is a machine learning methodology that has garnered considerable interest in the medical domain, emphasizing the utilization of pre-trained models on extensive datasets for particular tasks. Transfer learning is an essential instrument in medical image analysis, facilitating the development of high-performance models while minimizing training duration and computing expenses (Pacal, 2024). With the evolution of the field, transfer learning is anticipated to assume a more pivotal role in enhancing patient care. Numerous transfer learning models, such as VGG16, ResNet, and AlexNet, have demonstrated significant effectiveness in this domain. Transfer learning models, employing the depth and intricacy of neural networks, are utilized to discern complicated patterns in medical images (Mathivanan et al., 2024). This versatile approach extends beyond these well-known architectures, with numerous other models contributing to the growing range of tools for medical imaging analysis (Kumar & Ma, 2024). Transfer learning in medical imaging has markedly accelerated the development process and enhanced the performance and accuracy of pre-trained models, facilitating swifter and more precise diagnoses of malignant tumors, especially in their detection and categorization (Remzan et al., 2024). The efficiency improvements realized by transfer learning models have substantial implications for patient care, as early diagnosis and accurate categorization of cancer types are crucial for commencing prompt and focused treatment methods (Mandle et al., 2024). The interplay among deep learning, artificial intelligence, and transfer learning is set to revolutionize medical picture analysis. The integration of these technologies enhances the capacities of healthcare workers and has the potential to improve patient outcomes and transform medical diagnostics (Gül & Kaya, 2024). This study makes its key contribution through pre-trained convolutional neural network (CNN) ensemble methods that enhance brain tumor diagnostic accuracy. This study combines the AlexNet, ResNet18, and VGG16 networks through an ensemble system, which shows the advantages of multiple architecture implementations during prediction. Through the combination of majority and weighted voting systems, the classification precision increases. The first section of this paper is the introduction, and the second section is the related work, where we show the work of other papers, their methods, and results. Section 3 of the paper discusses the methods of our paper and how they are utilized. Section 4 is about our preparations and the experiments we have done. The fifth section is the result of our experiments and comparison with other datasets. Finally, Section 6 concludes the paper.

2. Related Work

The main goal of this study is to review and comprehend brain tumor classification and recognition strategies established globally between 2015 and 2024. The current study reviews the most widespread procedures for detecting brain cancer that have been made accessible universally, in addition to observing how effective computer-aided diagnosis systems are in this process. Current relevant review papers, along with their respective specifics and highlights, are discussed in this section.

ZainEldin et al. (2023) propose a method for diagnosing and classifying brain tumors using CNNs, and an adaptive dynamic sine-cosine fitness grey wolf optimizer is presented in this paper. The proposed model, Brain Connectivity Matrix-CNN, outperforms other models when evaluated on the Brain Tumor Segmentation Challenge (BRATS) 2021 dataset with an accuracy of 99.99%. It features both hyperparameter tuning of a CNN and segmentation capabilities utilizing a 3D volumetric data segmentation (3D UNet) architecture. In a study by Abdusalomov et al. (2023), a new method of brain tumor detection using deep learning was developed, the central idea of which is based on the enhancement of the You Only Look Once (YOLOv7) model integrated with additional components such as the Convolutional Block Attention Module, Spatial Pyramid Pooling Fast Plus(), and Bidirectional Feature Pyramid Network. This model is implemented and trained using a dataset of MRI images, which allowed for the achievement of satisfactory accuracy equal to 99.5%. It focuses on the correct detection of glioma, meningioma, and pituitary tumors and performs better than previously existing methods in this regard. Problems of small tumor size and instability of localization were resolved due to effective feature extraction. Ranjbarzadeh et al. (2021) suggest an innovative method for brain tumor segmentation based on CNNs using a Distance-Wise Attention Mechanism. The model achieves computational efficiency and lower overfitting by concentrating only on the more localized areas of interest and using pre-processing for unnecessary information. Experiments on the BRATS 2018 dataset provide competitive results with good performance in tumor localization and segmentation. Mahmud et al. (2023) aim to improve the efficiency of deep learning in detecting brain tumors from MRI scans. A CNN framework is presented and evaluated against other architectures, including ResNet-50, VGG16, and Inception V3. It used a sample of 3264 MRI scans and emerged with the CNN model, achieving the highest performance at 93.3% accuracy and an area under the curve (AUC) at 98.43%. Based on the study, it focuses on the early stage of tumor detection to further reduce the mortality rate of the patients using the pre-processing and data augmentation methods. Noreen et al. (2020) recommend using deep learning methodology, attempting an automated diagnosis of brain tumors focused on differentiating among the three types: gliomas, meningiomas, and pituitary tumors, utilizing MRI. It uses two pre-trained models, Inception-v3 and DenseNet201, to extract features from different layers and combine them in order to improve the classification. The feature-level fusion technique is known to provide multiscale and dense embeddings, which are better than other existing techniques. The Inception-v3 model achieved an accuracy of 99.34%, and the DenseNet 201 model achieved 99.51% accuracy on the testing dataset. The results support the second hypothesis and show that brain cancer classification through feature fusion is effective and can be considered a viable approach for medical imaging tasks. Mostafa et al. (2023) tackled the problem of brain tumor segmentation through deep learning techniques with MRI images. Khan et al. (2022) present a solution that offers complete automation based on a CNN model, which is a modern deep-learning technique making use of datasets containing multimodal MRI images, including the BRATS. Major results include the segmentation of brain tumors into classes: Necrotic and edema enhancing, with the achievement of validation accuracy of 98% even in different settings. The methodology includes: preprocessing, data augmentation, model training including U-Net sampling techniques, and model optimization with cross-entropy loss and categorical method. The research focuses on progress made in the area of automated brain tumor diagnosis and efforts made to enhance efficiency and the level of precision to be used in assisting medical personnel. The paper describes a novel hierarchical deep-learning model for brain tumor identification. The developed hierarchical deep learning with 2D and 3D features for brain tumor segmentation (HDL2B-TUMOR-CLASSIFIER) system consists of CNNs that divide the studied brain tumors into four types: Glioma, meningioma, pituitary, and no tumor. The model achieved a 92.13% accuracy, which is significantly higher than many of the methods known: the model's miss rate was 7.87%. A dataset of 3,264 images was used in three stages: Preparation, training, and validation. It focuses on assisting clinical diagnoses of brain tumors by increasing the speed of detection and the accuracy of classification. Ghaffari et al. (2020) show that the classification and segmentation of brain tumors demonstrate substantial development when deep learning and machine learning models are applied to the process. The researchers utilized multimodal MRI together with machine learning and CNN on the BRATS dataset to tackle the problem of costly and time-consuming physical brain tumor segmentation. The research analyzed multimodal MRI to develop a benchmark that showed more than 90% success through previous work comparisons. In a study by Gao et al. (2022), a multi-scale CNN was used on the BRATS dataset for dealing with heavy computational needs that affect 3D convolution networks and the independent nature of single-view 2D slices. This model implemented multi-scale approaches to effectively show how neighborhood size affected segmentations in three-dimensional CNNs while achieving superior tumor segmentation performance at enhancing tumor (75%), whole tumor (90%), and tumor core (84%). According to a study by Albalawi et al. (2024), the application of CNN technology to process the Kaggle database successfully solved problems with manually detecting brain tumors from MRI images because of diverse tumor dimensional characteristics and variations in shape and brightness. The study created a high-performing classification system that differentiated glioma from meningioma and pituitary tumors with a 98.04% success rate. The hybrid deep learning framework by Rasheed et al. (2023) combined AlexNet with ResNet-18 using a support vector machine (SVM) to analyze MRI data collected at Nanfang Hospital and Tianjin Medical University General Hospital. Researchers worked on tumor brain classification through the combination of CNN algorithms with SVM methods because they wanted improved precision in detecting tumor and non-tumor patterns. The combined model reached a performance level of 95.1% accuracy coupled with 95.25% sensitivity and 98.50% specificity. Jader et al. (2024) tackled painstaking issues associated with the segmentation and classification of brain tumors that are prone to human error. Its tasks include the classification of MRI images into four categories: pituitary, glioma, meningioma, and nontumorous. Subsequently, the research carried out employed VGG-16, ResNet-50, and AlexNet models, which were based on transfer learning, and consolidated them into an ensemble model in order to increase classification accuracy. Achieving greater classification accuracy than the older methods, the ensemble model performed better than other methods such as Naïve Bayes, decision trees, random forests, and deep neural networks. The results returned were 99.16% accuracy, 98.47% sensitivity, 98.57% specificity, 98.74% precision, 98.49% recall, and 98.18% F1 score. The primary innovation of this paper is brain tumor classification ensemble transfer learning, which is a new diagnosis approach. The results prove this method is effective. The deep learning method for classifying various stages of Alzheimer's disease (AD) presented in Ramzan et al. (2020) was conducted on a sample that consists of resting state functional MRI images collected by the AD Neuroimaging Initiative, which is comprised of 138 subjects. The model classifies six stages of AD, which include: Cognitively normal, significant memory concern, early mild cognitive impairment, mild cognitive impairment, late mild cognitive impairment, and AD. The goal of this study's research is to enhance the diagnostic process for the classification of AD at its early stages, which poses a great challenge due to similarities in symptom presentation and a lack of multiclass classification approaches. Doing so allows for the contribution of this research to be more clinically relevant. Residual learning, transfer learning, and deep learning methods of this work have been shown to significantly improve classification performance. Results show that the average accuracy of the ResNet-18 model that was fine-tuned is 97.88%. This achieves the best results to date and exceeds accuracy in classifier development for all stages of AD. In Roopa et al. (2022), a CNN algorithm is suggested for detecting brain tumors in MRI images. The study employed a dataset of 3,264 MRI scans, classified into four categories: Glioma, meningioma, pituitary tumor, and no tumor. The challenge in focus is the timely and precise identification of brain tumors for effective treatment management. The study analyzes the performance of the designed CNN against the ResNet-50, VGG16, and Inception V3 benchmark models, measuring the accuracy, recall, AUC, and loss. The main contribution of the paper is uncovering the effectiveness and benefit of the proposed CNN model, which outshone other models by achieving an accuracy of 93.3%, an AUC of 98.43%, a recall of 91.19%, and a loss of 0.25. The authors' conclusions indicate that the CNN algorithm put forward is a dependable answer for detecting brain tumors from MRI images and is more accurate and robust than popular transfer learning algorithms. Putzu et al. (2020) present a classifier model that is based on CNNs and is built on top of AlexNet. Their research attempts to solve the content-based image retrieval problem using CNNs for feature extraction and relevance feedback (RF). The dataset encompasses Caltech-101, Caltech-256, Flowers-102, and SUN-397, which have different classes of images for the retrieval tests. The most important part of this paper is the introduction of two CNN architectures for RF, one of which has an original AlexNet depth but a last layer adapted to feedback, and the other one has an additional layer for better user feedback incorporation. The work also investigates some query refinement approaches such as relevance scoring and mean feature computation to improve retrieval accuracy. The experimental results confirm that the proposed CNN based RF methods increase the retrieval accuracy, where RF using classification outperformed feature extraction in the later iterations, which proves the effectiveness of tuned CNNs on interactive image search systems. The study by Al-Hadidi et al. (2020) features an advanced technique that employs a multi-model CNN using Xception, DenseNet-201, and EfficientNet-B3 as classifiers to identify brain tumors. The study seeks to address the issue associated with low classification accuracy in brain MRI images, which stems from variations in tumor size, shape, and position. The data set contains THOMAS (Dataset 1) and NICKPARVAR (Dataset 2), which each contain MRI photographs of glioma, meningioma, pituitary tumor, and other no tumor categories. The algorithm developed in this work is innovative in the sense that it bases model selection on the validation accuracy and false positive rate to combine multiple CNN models and thus results in better classification efficiency. The results show that multi-model CNN outperformed single CNN models, achieving 97.74% accuracy on Dataset 1 and 99.69% on Dataset 2, which is an improvement between 1.29% and 4.19% for Dataset 1 and 0.22–0.61% for Dataset 2. Based on these results, it can be concluded that traditional single-model approaches to brain tumor detection are less effective than the multi-model approach. Abdullah et al. (2024) suggest incorporating deep CNN (DCNN) techniques, which are based on the VGG-16 model for tumor identification in MRI scans. This dataset has 253 MRI images of the brain—155 with identified tumors and 98 without. These images were obtained from Kaggle. The research is aimed at providing a solution for automated and accurate brain tumor detection, which circumvents manual methods that are labor-intensive and highly subjective. The primary alteration the authors made to the architecture of VGG-16 was a substitution of the last max-pooling with Global Average Pooling, which mitigated the effects of overfitting and improved generalization. The accuracy achieved by the proposed DCNN model is 96%, which is higher than the accuracy achieved using conventional methods. The model performed remarkably well, as it achieved a precision of 0.93, a sensitivity of 1.00, an F1-score of 0.97, Cohen's kappa of 0.91, and an AUC of 0.95, an indicator of the effectiveness of the tool for clinical experts to improve brain tumor detection and expedite treatment measures. Asif et al. (2022) focus on the classification of brain tumors using pre-trained DCNNs like VGG-19, VGG-16, ResNet50, and Inception V3 on MRI images. The dataset consists of 305 brain MRI images, including tumorous and non-tumorous cases, collected from publicly available sources. The study addresses the problem of fully automated and accurate classification of brain tumors, which is significant for prompt diagnosis and therapeutic intervention. The main effort stems from determining how effective transfer learning is with the pre-trained DNNs with the small sample set, and the poster shows that high accuracy can be achieved. The results indicate that VGG-19 achieved the highest accuracy of 99.48%, followed by VGG-16 (99%), ResNet50 (97.92%), and Inception V3 (81.25%). These results support the claim that the automated classification of brain tumors with models assisted by a transfer learning framework is practical and does not require elaborate feature extraction procedures.

3. Methodology

3.1. Dataset

Three deep learning models, namely AlexNet, ResNet18, and VGG16, were utilized for the implementation of this proposed approach. The trained deep-learning models operated on data obtained from both local hospitals and public repositories. The main research objective of this study does not revolve around transfer learning innovations but rather deploying ensemble voting methods to enhance classification accuracy. The ensemble framework carried out predictive analysis by using majority voting and weighted voting methods to boost classification outcomes. The proposed approach uses existing ensemble learning techniques for its effectiveness while mainly depending on established principles. The main advancement occurs from the practical application of these models in medical imaging contexts instead of creating new computational structures. Future explorations need to use deep features or advanced fusion techniques to improve methodological innovation.

For developing the model, a dataset was collected, which had 200 cases of patient MRI images using Digital Imaging and Communication in Medical. Of those 200 cases, 100 of them were normal, and 100 of them were abnormal; the dataset normal cases had 38 male patients and 62 female patients, and in the abnormal 67 male and 33 female, which means males are infected with brain cancer, the age of the patients was in the range of (19–98) years old. It required 3 months of data collection from Rizgary Hospital and Hiwa Hospital from the cases of 2024; 2 months were spent for data cleaning, which gave the outcome of 3000 images 1500 of them were normal, and 1500 were abnormal with a resolution of 512 × 512, and 15 images from each case, and Table 1 shows the detail of dataset split, which was divided into 80% for training and 20% for testing. The dataset was organized into two main folders: Normal and abnormal. Within each folder, images were further divided into test and train subfolders. Specifically, 300 images were allocated for testing and 1,200 images for training in each category, as shown in Table 1.

Overall, training and testing the data for each model took between 2 and 5 h, different for each model, using a Legion 13th Gen Intel (R) Core (TM) i7-13650HX 2.60 GHz having an Nvidia RTX 4060 GPU with 16 GB of RAM DDR5 4,800 MHz and 1.5 TB SSD NVME hard disk 3,500 mb/s.

3.2. Methods

The proposed model, illustrated in Fig. 1, employs three well-known transfer learning approaches—ResNet, VGG16, and AlexNet, an ensemble model—to create three classes for analyzing and estimating: the recommended frame. The data undergoes three transfer learning techniques, and following analysis, it's divided into an 80% training set and a 20% testing set. Using Pytorch, this code implements and assesses three deep learning models, AlexNet, ResNet, and VGG16, based on an image dataset. To begin with, the code imports and processes the image dataset by creating a training and testing split. The models are adjusted for their required number of output classes, and weights have been initialized from pre-trained models. At the phase of training, loss and accuracy scores are computed for each batch and overall raw and weighted accuracy statistics are recorded across a number of epochs. After training has been completed, it tests the models on the test data and scores them in terms of accuracy, confusion matrix, as well as a classification report. Also, the code performs an ensemble of the models by using majority and weighted voting and computes the ensemble accuracy as well. Multi-class receiver operating characteristic (ROC) and AUC metrics to analyze the model's multi-class classification performance are also calculated and presented. Confusion matrices and comprehensive classification reports for training and test sets are presented to evaluate the performance of the models according to the various classes. Fig. 2 shows the normal and abnormal images of brain MRI.

Table 1. Data split detail

Phase	Abnormal (50%)	Normal (50%)	Total
Train	1,200	1,200	2,400
Test	300	300	600
Total	1,500	1,500	3,000

3.3. Pre-processing

Pre-processing refers to the input images that will go for further analysis to improve the effective analysis. It comprises eliminating artifacts for better focus, concentrating on the area of the brain by excluding non-brain tissues, and cutting up the image with segmentation to define its meaningful parts (Khairandish et al., 2022). In the context of image processing, pre-processing regularly includes resizing pixels to a hard and fast size, normalizing pixel values to a trendy range, applying data augmentation techniques like flipping or rotating to growth variety, and converting photos into tensor formats appropriate for computation. These steps collectively ensure the statistics are optimized for effective model learning and evaluation (Rasheed et al., 2023).

3.3.1. ResNet

ResNet, which stands for residual network, is a deep-learning architecture proposed by Ramzan et al. (2020) as a way of solving problems posed during classification. The primary advancement of this system is the deployment of the so-called residual or skip connections, which make it possible to directly add portions of a layer back into a layer (Jader et al., 2024; (Mahmud et al., 2023). The network is able to learn residual mappings, which makes the whole optimization process less complicated and also reduces the problems of vanishing gradients that are experienced in deep networks. ResNet architecture is built using modular residual blocks each of which contains convolutional layers that perform batch normalization and rectified linear unit (ReLU) activation in conjunction with skip connections (Ramzan et al., 2020). ResNet reliability and versatility have thus made it a more commonly used model to transfer learning between several tasks. With the ability to train very deep networks, ResNet has served as a point of reference in deep learning, proving that if done well within trained deeper models, they yield better results (Ramtekkar et al., 2023). The model consists of several blocks, as shown in Fig. 3.

3.3.2. VGG16

The structure is distinguished by a uniform and deep architecture, which contains 16 weight layers of three fully connected and thirteen convolutional layers. The unique feature of VGG16 is the use of small 3 × 3 convolutional kernels in each layer of the network. A smaller kernel is not a drawback, as many convolutional layers may be accustomed to compute

DOI: 10.6977/IJoSI.202506_9(3).0007
N. F. Othman, S. W. Kareem//Int. J. Systematic Innovation, 9(3), 109-123 (2025)

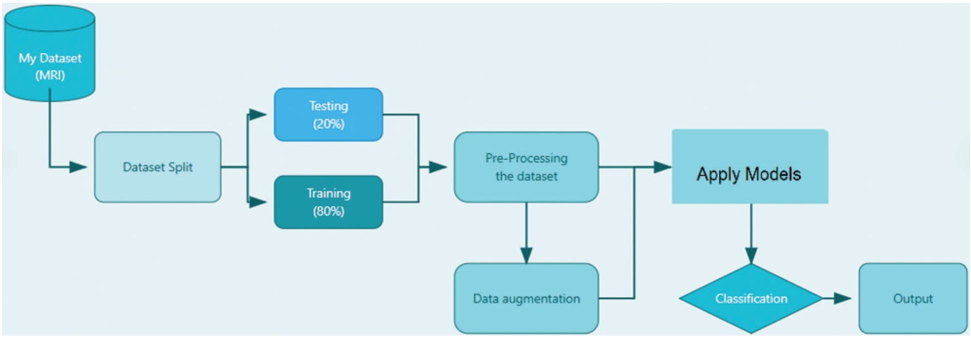

Fig. 1. Proposed model architecture

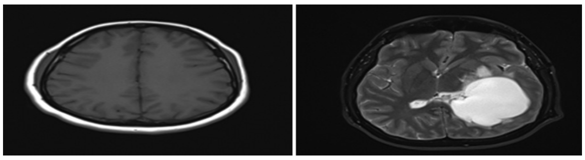

Fig. 2. Normal (left) and abnormal (right) images of brain magnetic resonance imaging

complex and hierarchical features (Roopa et al., 2022). Definitely a typical approach for the mentioned architectures supporting wider kernels as opposed to the deeper network (Mahmud et al., 2023); (Sulistyowati et al., 2023). The structural design of VGG16 is quite simple, adopted sequential architecture without skip or residual connections, and this makes it very easy to build and comprehend. The architecture's depth and its design enable it to perform effectively in retrieving the details of visual features; hence, it is widely used for image classification, transfer learning, and feature extraction (Bansal et al., 2023). Fig. 4 depicts the architecture of VGG16.

3.3.3. AlexNet

AlexNet features eight layers in total, namely, five convolutional and three fully connected layers. The model proposed multiple measures to train deep neural networks more efficiently (Mathivanan et al., 2024). As for the architecture, the authors employed ReLU activations, which considerably sped up the training process, dealing with the vanishing gradient problem that is extremely common for the use of sigmoid or tanh activations. Moreover, AlexNet included dropout into the network to prevent the issue of overfitting and also used data augmentation to extend the limited training set and its variability. The biggest advancement, however, was the training on GPUs, which allowed for harnessing their parallelism in order to tackle the numerous computationally heavy operations needed in the model (Jader et al., 2024). It was also the case that AlexNet implemented local response normalization, which sought to increase the feature learning of the model through neuronal competition (Putzu et al., 2020). The success of the architecture proved the ability of deep learning to tackle complicated vision problems, which in turn led to the construction. AlexNet allowed for the development of the modern architecture of CNNs, establishing it as a landmark in the evolution of deep-learning algorithms (Nawaz et al., 2022). The model consists of several layers, as shown in Fig. 5.

4. Preparation and Assessment Experiments

In this experiment, a large dataset of 3,000 images was collected from 200 cases of patients, 1,500 of them were benign, and 1,500 were Malignant. All the data was collected within 3 months in Erbil Rzgray Hospital and Sulaymaniyah Hewa Hospital to ensure the effectiveness of the training and testing phases. We also collected two more datasets to access a robust computing environment. The datasets that we collected from Kaggle were Brain Tumor Image Dataset with Grayscale Normalization and Zoom and Brain MRI Images for Brain Tumor Detection, which were used to compare with the data we collected. Importantly, the same datasets were utilized for all advanced models, encompassing both the training set and the testing set. The success of our models can be attributed to the collaborative contributions of Sklearn, TensorFlow, and PyTorch. For optimal performance in all high-end models, a block size of 32 was determined to be the most effective. Table 2 illustrates the hyperparameter details of transfer learning models.

The evaluated model, ResNet18, exhibited superior performance, achieving the lowest testing loss of 0.0235 at epoch 30 and testing accuracy of 99.33.00%. Among the evaluated models, AlexNet exhibited the lowest testing loss of 0.1026 at epoch 32 but experienced the most fluctuation in testing accuracy. It ultimately achieved testing accuracy of 98.17%. VGG16 demonstrated promising results, achieving training accuracy of 98.83% and loss of 0.0426 at epoch 39, respectively, Figs. 6-8 describe the result of the three models from each epoch.

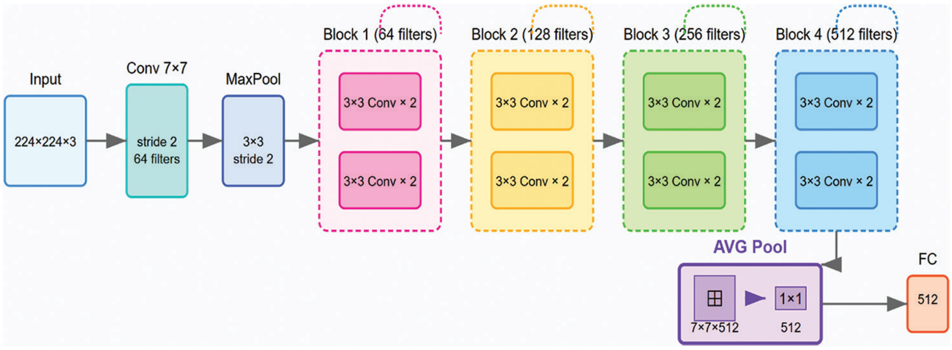

Fig. 3. Architecture of ResNet 18

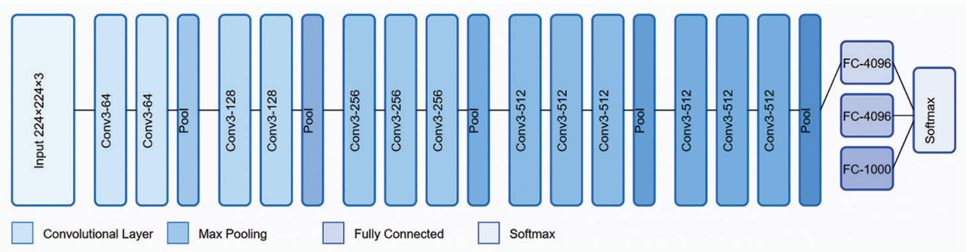

Fig. 4. Architecture of VGG16

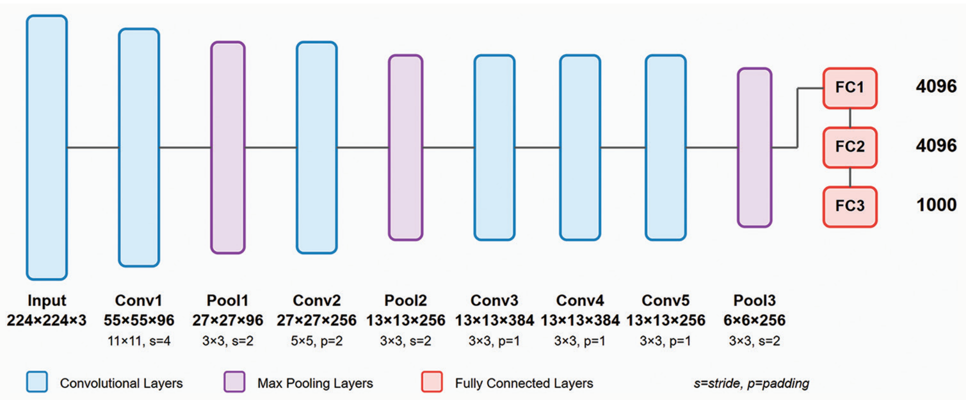

Fig. 5. Architecture of AlexNet

Table 2. Hyperparameters of transfer learning models for image classification

Quantifying performance and evaluation	**Assessing measurement outcomes**
Size of the batch	32
Optimizer	Adam
No. of epochs	50
Rate of learning	0.0001
Evaluation criterion	Cross-entropy loss function
Training	Confusion matrices

An ensemble of three models combines their predictions to enhance accuracy and robustness. This can be executed through strategies such as voting (majority or weighted), where the final output is primarily based on the consensus or confidence of each model. For brain tumor classification, an ensemble of ResNet18, VGG16, and AlexNet should use majority or weighted voting to supply extra reliable predictions. The pseudocode below shows the process of combining all three models and running them together. It also shows the result of the training and testing data accuracy for each of the models. Furthermore, we got the ensemble predictions of all the models together; the majority voting was (98.33%), and the weighting voting was (98.33%). Lastly, we also found the ROC curves, an AUC of 1, and plot accuracies for the models over the epochs as shown in Figs. 9-11, which show the architecture of the ensemble model.

DOI: 10.6977/IJoSI.202506_9(3).0007
N. F. Othman, S. W. Kareem//Int. J. Systematic Innovation, 9(3), 109-123 (2025)

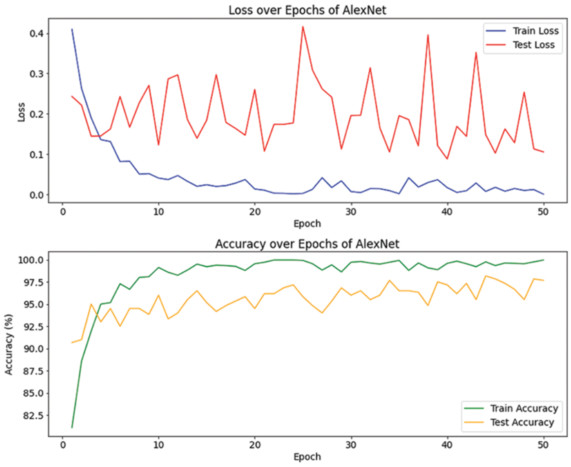

Fig. 6. Training and testing loss and accuracy of AlexNet

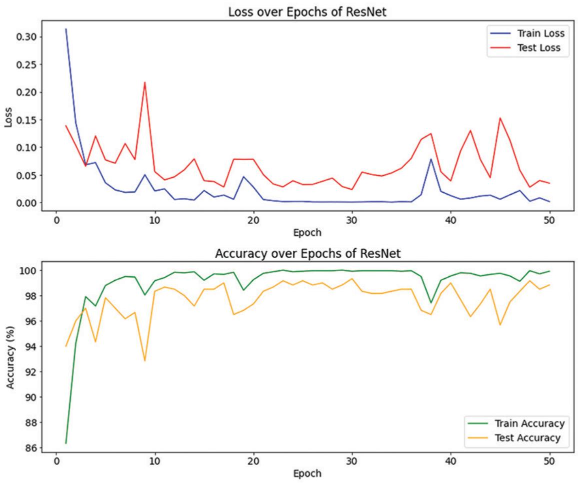

Fig.7. Training and testing loss and accuracy of ResNet

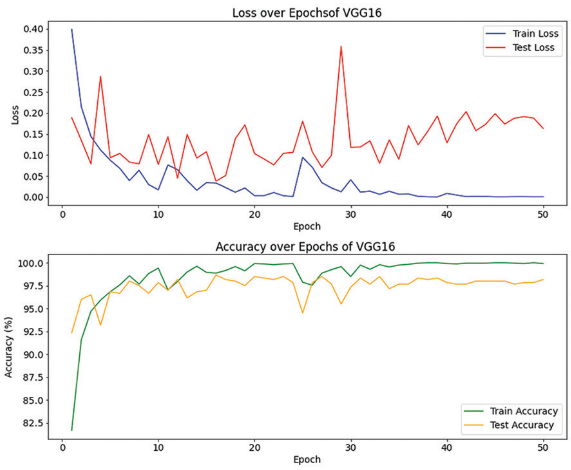

Fig.8. Training and testing loss and accuracy of VGG16

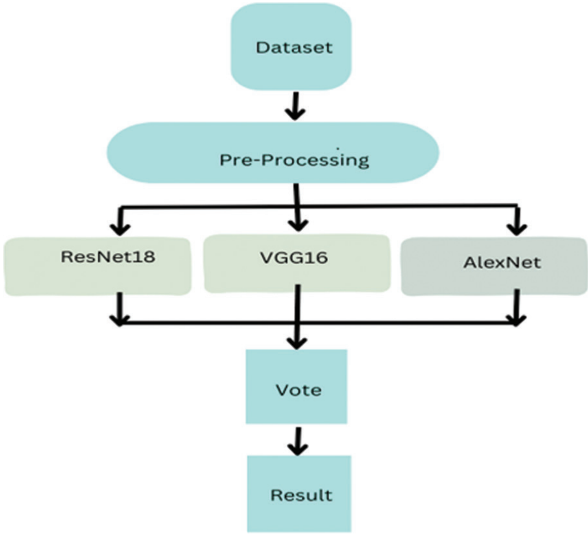

Fig. 9. Ensemble of AlexNet, ResNetv18, and VGG16

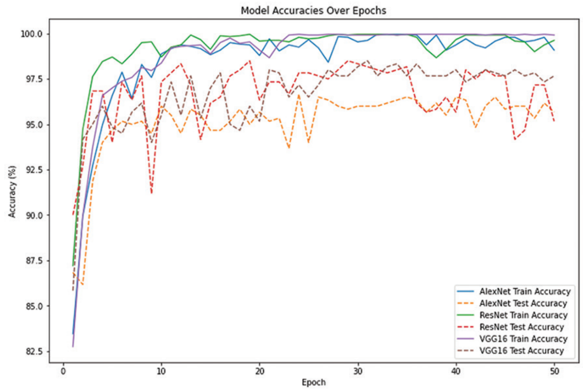

Fig. 10. Ensemble accuracy of AlexNet, ResNetv18, and VGG16

5. Experimental Results and Comparison

5.1. Results and Discussion

The primary aim of the learning technique presented is to develop several models in this study. The initial ResNet represents a specific learning methodology. The second model is VGG16, and the third is AlexNet. The accuracy, specificity, and sensitivity computed during the testing in this work serve as the evaluation metrics for the three models in this system.

Table 3 proposes an evaluation metric for the ResNet18, VGG16, and AlexNet models' performance in training and testing out datasets, differentiating between "normal" and "abnormal" instances of the dataset. These metrics contain precision, F1-rating, and support for both datasets.

- Accuracy: Accuracy refers to the ratio of the true patterns to the summation of entire patterns. It can be expressed as

1 Setup and Preprocessing
- Define dataset path and image transformations.
- Load dataset, split into training (80%) and testing (20%).
- Create Data Loaders for training and testing.

2 Define Models
- Create AlexNetModel, ResNetModel, VGG16Model classes.
- Modify last layer of each model to match the number of classes.

3 Initialize
- Load models and move to CPU/GPU.
- Set training parameters (epochs, learning rate, loss function).

4 Train Model
- For each batch in training data:
- Forward pass, calculate loss, backpropagate, update weights.
- Track training loss and accuracy.

5 Evaluate Model
- For each batch in testing data:
- Forward pass, calculate loss, track predictions.
- Compute test loss and accuracy.

6 Train and Evaluate All Models
- For each model (AlexNet, ResNet, VGG16):
- Train and evaluate for each epoch.
- Store training/testing metrics.

7 Ensemble Predictions
- Predict on test set using all models.
- Combine predictions using majority voting and weighted voting.
- Calculate ensemble accuracy.

8 Performance Visualization
- Plot accuracies for all models over epochs.
- Compute and plot ROC curves (binary or multi-class).

9 Metrics
- Calculate and display confusion matrix, accuracy, and classification report for training and testing datasets.

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positives} + \text{False Positives} + \text{True Negative} + \text{False Negative}} \times 100$$

- Precision: The percentage of accurately projected positive observations to the total projected positives

Table 3. Training and testing all models

Model	Type	Precision	Recall	F1-score
ResNet18				
Testing Abnormal	Normal	0.99	0.98	0.99
	0.98	0.99	0.99	
Training Abnormal	Normal	1	1	1
	1	1	1	
VGG16				
Testing Abnormal	Normal	0.97	0.99	0.98
	0.99	0.97	0.98	
Training Abnormal	Normal	1	1	1
	1	1	1	
AlexNet				
Testing Abnormal	Normal	0.97	0.99	0.98
	0.99	0.97	0.98	
Training	Normal	1	1	1
	Abnormal	1	1	1

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positives} + \text{False Positives}}$$

Greater precision shows fewer false positives.

- Recall (or sensitivity): The percentage of accurately projected positive observations to all actual positives

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positives} + \text{False Negative}}$$

Greater recall shows fewer false negatives.

- F1-Score: The harmonic mean of precision and recall

$$F1_{score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

A high F1-score shows that the model balances precision and recall in a good way.

Fig. 12 illustrates the ROC curves for all three models. ResNet is exceptionally proficient in differentiating between the two classes, exhibiting minimal false positives and a high rate of true positives across various thresholds. An AUC of 1 demonstrates high model performance. VGG16 also had a proficient ROC curve with a result of 1, and lastly, AlexNet had an ROC curve result of 1.

5.2. Comparison

Table 4 provides a detailed performance evaluation for the testing of three transfer learning

DOI: 10.6977/IJoSI.202506_9(3).0007
N. F. Othman, S. W. Kareem//Int. J. Systematic Innovation, 9(3), 109-123 (2025)

Table 4. Performance comparison of the author's dataset with two other datasets

a-our dataset (3000imges)				b-Brain Tumor Image Dataset with Grayscale Normalization and Zoom (3096 images)				c-Brain MRI images for brain tumor recognition (252 images)			
	Precision	Recall	F1-score		Precision	Recall	F1-score		Precision	Recall	F1-score
Normal	0.98	0.99	0.98	Normal	0.97	0.99	0.98	Normal	0.86	0.90	0.88
abnormal	0.99	0.97	0.98	Abnormal	0.99	0.98	0.99	Abnormal	0.93	0.90	0.92
AlexNetAccuracy: 97.666%				AlexNet Accuracy: 98.13%				AlexNet Accuracy: 94.11%			
ResNet Accuracy: 98.66%				ResNet Accuracy: 99.25%				ResNet Accuracy: 96.07%			
VGG16 Accuracy: 97.8%				VGG16 Accuracy: 98.88%				VGG16 Accuracy: 94.117%			
Ensemble (majority vote) accuracy: 98.33%				Ensemble (Majority Vote) Accuracy: 98.5%				Ensemble (Majority Vote) Accuracy: 92.15%			
Ensemble (weighted vote) Accuracy: 98.33%				Ensemble (Weighted Vote) Accuracy: 98.5%				Ensemble (Weighted Vote) Accuracy: 92.15%			
AlexNet: Epoch 32 is the best, with a Test Loss of 0.10261 and a Test Accuracy of 98.17%.				AlexNet: Epoch 30 is the best, with a Test Loss of 0.1883 and a Test Accuracy of 98.88%.				AlexNet: Epoch 4 is the best, with a Test Loss of 0.2836 and a Test Accuracy of 96.08%			
ResNet: Epoch 30 is the best, with a Test Loss of 0.0235 and a Test Accuracy of 99.33%.				ResNet: Epoch 35 is the best, with a Test Loss of 0.0345 and a Test Accuracy of 99.63%.				ResNet: Epoch 16 is the best, with a Test Loss of 0.1488 and a Test Accuracy of 96.08%			
VGG16: Epoch 39 is the best, with a Test Loss of 0.0426 and a Test Accuracy of 98.83%.				VGG16: Epoch 27 is the best, with a Test Loss of 0.0493 and a Test Accuracy of 99.63%.				VGG16: Epoch 10 is the best, with a Test Loss of 0.1074 and a Test Accuracy of 96.8%.			

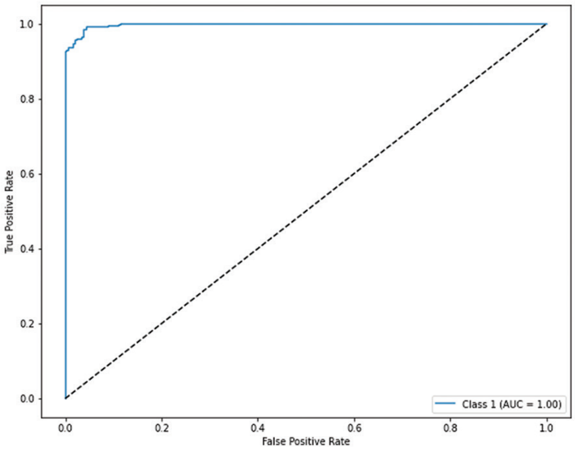

Fig. 11. Receiver operating characteristic curve Ensemble of AlexNet, ResNetv18, and VGG16

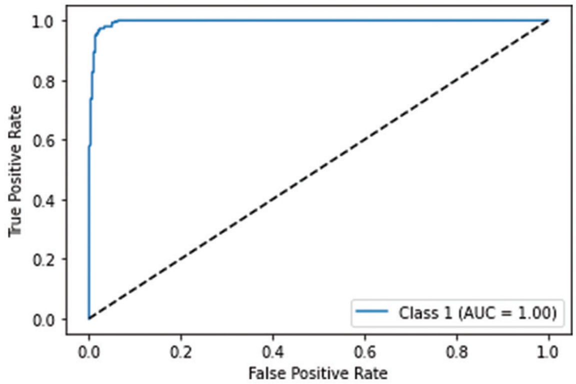

Fig. 12. Receiver operating characteristic curve

models, AlexNet, ResNet, and VGG16, on three datasets of different sizes and complexity. Below is a performance comparison focusing on the key targets. I used two other datasets to compare with my dataset. The results show that my dataset has a better accuracy rate than the other two datasets.

- Dataset Size Impact:
    - My dataset (3000 images) enhances the overall.
    - Performance due to higher model generalization, as visible in the 98.33% accuracy on your dataset the usage of ensemble strategies.
    - Smaller datasets display barely decreased accuracy due to constrained schooling variety.
- Model Performance:
    - ResNet continuously achieves the highest accuracy across datasets, highlighting its capability to handle complex datasets successfully
    - VGG16 and AlexNet carry out comparably, with slight differences in smaller datasets.

- Ensemble Models:
    - Ensemble (Majority Vote) and Ensemble (Weighted Vote) methods continually outperform character models, emphasizing their robustness in combining a couple of predictions.
- Optimal Epochs:
    - Models obtain their fine overall performance at varying epochs depending on the dataset complexity and size.

5.2.1. Brain tumor image dataset with grayscale normalization and zoom (3096 images)

The dataset was brought from Kaggle; it consists of 3096 brain images. That contains normal and abnormal. As shown in Table 5, the method used in my model had the highest accuracy and the method used in (Al-Hadidi et al., 2020). Had the lowest accuracy. Furthermore, (Wang et al., 2024) Had the highest precision.

5.2.2. Brain MRI images for brain tumor detection (252 images)

The dataset was also brought from Kaggle; it contains 252 images of normal and abnormal brain images. As shown in Table 6, the methods used in (Asif et al., 2022) had the highest accuracy among all the methods, with InceptionResNetV2 having the highest accuracy, while (Bakr Siddiaue et al., 2020) had the lowest accuracy among all. Moreover, InceptionResNetV2 had 100% in precision.

Table 5. Performance comparison of brain tumor image dataset with grayscale normalization

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Multi-model of Xception, DenseNet-201, and EfficientNet-B3. (Santoso *et al*., 2024)	97.74	97.7	98.06	97.92
U-Net, CNN, VGG19 (Dhiman & Satpute, 2019)	93.7	-	93.1	-
CNN (Al-Hadidi *et al*., 2020)	75	-	-	-
AlexNet, VGG, ResNet (Wang *et al*., 2024)	96.94	99.32	-	-
ResNet50 and VGG16 (Abdullah *et al*., 2024)	92.6	-	-	-
Ensemble model (ResNet18, VGG16, and AlexNet)	98.5	98	98.5	98.5

Table 6. Performance comparison with the dataset of brain MRI images for brain tumor detection

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Deep convolutional neural network (DCNN) (Bakr Siddiaue *et al*., 2020)	96	93	100	97
VGG-19, VGG-16, ResNet50, Inception V3) (Krishnapriya & Karuna, 2023)	99.48	100	98.76	99.17
	99	100	98.18	99.08
	97.92	77.77	87.27	82.24
	81.25	53.84	63.25	58.16
Xception, NasNet Large, DenseNet121, InceptionResNetV2 (Asif *et al*., 2022)	99.67	99.68	99.68	99.68
	99.34	99.36	99.36	99.36
	99.00	98.72	99.36	99.04
	99.67	100.00	99.36	99.68
DCNN (Ramtekkar *et al*., 2023)	98.9	-	-	-
Detection (Phase 01: DBFS-EC Framework), Classification (Phase 02: HFF-BTC Framework) (Khan *et al*., 2022)	99.56	99.91	98.99	99.45
	99.20	99.13	99.06	99.09
Ensemble model (ResNet18, VGG16, and AlexNet)	92.15	89.5	0.90	90
Ensemble model (VGG-16, ResNet-50, and AlexNet) (Jader *et al*., 2024)	99.16	98.74	98.49	98.18

6. Conclusion

This research has proven the efficiency of employing transfer learning architectures, ResNet18, VGG16, and AlexNet, for the identification of brain tumors through MRI images. Of the three models, ResNet18 outperformed the others thanks to its outstanding accuracy of 98.66%. This enabled it to proficiently navigate complex data patterns as well as solve common issues experienced by convolutional networks, one of them being the vanishing gradient problem. The performance of VGG16 was equally promising, as the model recorded an accuracy of 97.8%, aided by its deep convolutional configuration that enabled the model to learn advanced features. Yet, AlexNet, a less complicated architecture, also performed reasonably well with an accuracy of 97.66%, demonstrating its applicability in less demanding environments. The technique of ensemble learning further brought to the surface the fact that improving the accuracy of individual models enhances the overall model accuracy to 98.33%, thereby supporting the case for combined models in improving diagnosis. This work not only supports the credibility of deep learning approaches to medical imaging but also the transformative nature of transfer learning in medicine, where the challenges posed by low datasets have been streamlined. In the future, studies might investigate the addition of further transfer learning frameworks, hyperparameter tuning, and multi-modal imaging data to enhance their diagnostic capabilities. This study's findings are valuable in improving the diagnosis and management of a brain tumor, which provides a basis for better medical solutions that are more accurate and efficient. Future work could discover innovative alternatives, fusion strategies, architectural modifications, or hybrid models to further enhance diagnostic accuracy and model robustness.

References

Abdullah, F., Jamil, A., Alazawi, E.M., & Hameed, A.A. (2024). Exploring deep learning-based approaches for brain tumor diagnosis from MRI images. In: *2024 IEEE 3rd International Conference on Computing and Machine Intelligence, ICMI 2024 - Proceedings*. https://doi.org/10.1109/ICMI60790.2024.10585851

Abdusalomov, A.B., Mukhiddinov, M., & Whangbo, T.K. (2023). Brain tumor detection based on deep learning approaches and magnetic resonance imaging. *Cancers (Basel)*, 15(16), 4172. https://doi.org/10.3390/cancers15164172

Al-Hadidi, M.R., AlSaaidah, B., & Al-Gawagzeh, M.Y. (2020). Glioblastomas brain tumour segmentation based on convolutional neural networks. *International Journal of Electrical and Computer Engineering*, 10(5), 4738–4744. https://doi.org/10.11591/ijece.v10i5.pp4738-4744

Albalawi, E., Thakur, A., Dorai, D.R., Khan, S.B., Mahesh, T.R., Almusharraf, A., et al. (2024). Enhancing brain tumor classification in MRI scans with a multi-layer customized convolutional neural network approach. *Frontiers in Computational Neuroscience*, 18, 1-19. https://doi.org/10.3389/fncom.2024.1418546

Alfonse, A., & Salem, A.B. (2016). An automatic classification of brain tumors through MRI using support vector machine. *Egyptian Computer Science Journal*, 40(3), 1110–2586

Arumugam, M., Thiyagarajan, A., Adhi, L., & Alagar, S. (2024). Crossover smell agent optimized multilayer perceptron for precise brain tumor classification on MRI images. *Expert Systems with Applications,* 238, 121453. https://doi.org/10.1016/j.eswa.2023.121453

Asif, S., Yi, W., Ain, Q.U., Hou, J., Yi, T., & Si, J. (2022). Improving effectiveness of different deep transfer learning-based models for detecting brain tumors from MR images. *IEEE Access*, 10, 34716–34730. https://doi.org/10.1109/ACCESS.2022.3153306

Bakr Siddiaue, M.A., Sakib, S., Rahman Khan,M.M., Tanzeem, A.K., Chowdhury, M., & Yasmin, N. (2020). Deep Convolutional Neural Networks Model-Based Brain Tumor Detection in Brain MRI Images. In: *Proceedings of the 4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020*, p909–914. https://doi.org/10.1109/I-SMAC49090.2020.9243461

Bansal, M., Kumar, M., Sachdeva, M., & Mittal, A. (2023). Transfer learning for image classification using VGG19: Caltech-101 image data set. *Journal of Ambient Intelligence and Humanized Computing*, 14(4), 3609-3620. https://doi.org/10.1007/s12652-021-03488-z

Dhiman, A., & Satpute, P.B.S. (2019). Brain tumor segmentation in MRI images. *International Journal of Research in Advent Technology*, 7(9), 10-14. https://doi.org/10.32622/ijrat.78201916

El Kader, I. A., Xu, G., Shuai, Z., & Saminu, S. (2021) Brain tumor detection and classification by hybrid CNN-DWA model using MR images. *Current Medical Imaging,* 17(10), 1248-1255. https://doi.org/10.2174/157340561766621022411 3315

Fang, L., & Wang, X. (2022). Brain tumor segmentation based on the dual-path network of multi-modal MRI images. *Pattern Recognition,* 124, 108434. https://doi.org/10.1016/j.patcog.2021.108434

Gao, P., Shan, W., Guo, Y., Wang, Y., Sun, R., Cai, J., *et al.* (2022). Development and validation of a deep learning model for brain tumor diagnosis and classification using magnetic resonance imaging. *JAMA Network Open*, 5(8), e2225608. https://doi.org/10.1001/jamanetworkopen.2022.25608

Ghaffari, M., Sowmya, A., & Oliver, R. (2020). Automated brain tumor segmentation using multimodal brain scans: A survey based on models submitted to the BraTS 2012-2018 Challenges. *IEEE Reviews in Biomedical Engineering*, 13, 156-168. https://doi.org/10.1109/RBME.2019.2946868

Gül. M., & Kaya, Y. (2024). Comparing of brain tumor diagnosis with developed local binary patterns methods. *Neural Computing and Applications*, 36(13), 7545-7558. https://doi.org/10.1007/s00521-024-09476-6

Harnod, T., Lin, C.L., Sung, F.C., & Kao, C.H. (2014). An association between benzodiazepine use and occurrence of benign brain tumors. *Journal of the Neurological Sciences*, 336(1-2), 8-12. https://doi.org/10.1016/j.jns.2013.11.009

Jader, R.F., Kareem, S.W., & Awla, H.Q. (2024). Ensemble deep learning technique for detecting MRI brain tumor. *Applied Computational Intelligence and Soft Computing*, 2024, 6615468. https://doi.org/10.1155/2024/6615468

Khairandish, M.O., Sharma, M., Jain, V., Chatterjee, J.M., & Jhanjhi, N.Z. (2022). A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images. *Irbm*, 43(4), 290–299. https://doi.org/10.1016/j.irbm.2021.06.003

Khan, A., Ghafoor, U., & Bhutta, M.R. (2022). A new deep hybrid boosted and ensemble learning-based brain tumor analysis using MRI. *Sensors*, 22(7), 2726.

Krishnapriya, S., & Karuna, Y. (2023). Pre-trained deep learning models for brain MRI image classification. *Frontiers in Human Neuroscience*, 17, 1150120. https://doi.org/10.3389/fnhum.2023.1150120

Kumar, A., & Ma, Y. (2024). *Localization and Classification of Brain Tumor Using Multi-Layer Perceptron*. Vol. 1132. Singapore: Springer Nature. https://doi.org/10.1007/978-981-99-8853-2_6

Mahmud, M.I., Mamun, M., & Abdelgawad, A. (2023). A deep analysis of brain tumor detection from MR images using deep learning networks. *Algorithms*, 16(4), 1–19. https://doi.org/10.3390/a16040176

Mathivanan, S.K., Sonaimuthu, S., Murugesan, S., Rajadurai, H., Shivahare, B.D., & Shah, M.A. (2024). Employing deep learning and transfer learning for accurate brain tumor detection. *Scientific Reports*, 14(1), 1–15. https://doi.org/10.1038/s41598-024-57970-7

Mandle, A. K., Sahu, S. P., & Gupta, G. P. (2024). WSSOA: Whale social spider optimization algorithm for brain tumor classification using deep learning technique. *International Journal of Information Technology,* 16(1), 1–17. https://doi.org/10.1007/s41870-024-01782-5

Mostafa, A.M., Zakariah, M., & Aldakheel, E.A. (2023). Brain tumor segmentation using deep learning on MRI images. *Diagnostics,* 13(9), 1-22. https://doi.org/10.3390/diagnostics13091562

Nawaz, S.A., Khan, D.M., & Qadri, S. (2022). Brain tumor classification based on hybrid optimized multi-features analysis using magnetic resonance imaging dataset. *Applied Artificial Intelligence*, 36(1), 2031824. https://doi.org/10.1080/08839514.2022.2031824

Noreen, N., Palaniappan, S., Qayyum, A., Ahmad, I., Imran, M., & Shoaib, M. (2020). A deep learning model based on concatenation approach for the diagnosis of brain tumor. *IEEE Access,* 8, 55135-55144. https://doi.org/10.1109/access.2020.2978629

Lauko, A., Lo, A., Ahluwalia, M.S., & Lathia, J.D. (2022). Cancer cell heterogeneity and plasticity in glioblastoma and brain tumors. *Seminars in Cancer Biology*, 82, 162–175. https://doi.org/10.1016/j.semcancer.2021.02.014

Leena, B., & Jayanthi, A. (2020). Brain tumor segmentation and classification via adaptive CLFAHE with hybrid classification. *International Journal of Imaging Systems and Technology*, 30(4), 874–898. https://doi.org/10.1002/ima.22420

Pacal, I., (2024). A novel Swin transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images. *International Journal of Machine Learning and Cybernetics*., 15(9), 3579-3597. https://doi.org/10.1007/s13042-024-02110-w

Praveen, G. B., & Agrawal, A. (2016). *Hybrid approach for brain tumor detection and classification in magnetic resonance images*. In: 2015 Communication, Control and Intelligent Systems, CCIS. IEEE, United States, p162-166. https:///doi.org/10.1109/ccintels.2015.7437900

Putzu, L., Piras, L., & Giacinto, G. (2020). Convolutional neural networks for relevance feedback in content based image retrieval:

A Content based image retrieval system that exploits convolutional neural networks both for feature extraction and for relevance feedback. *Multimedia Tools and Applications*, 79(37–38), 26995–27021. https://doi.org/10.1007/s11042-020-09292-9

Ramtekkar, P.K., Pandey, A., & Pawar, M.K. (2023). Accurate detection of brain tumor using optimized feature selection based on deep learning techniques. *Multimedia Tools and Applications*, 82(29), 1-31. https://doi.org/10.1007/s11042-023-15239-7

Ramzan, F., Khan, M.U.G., Rehmat, A., Iqbal, S., Saba, T., Rehman, A., *et al*. (2020). A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks. *Journal of Medical Systems*, 44(2), 37. https://doi.org/10.1007/s10916-019-1475-2

Ranjbarzadeh, R., Kasgari, A.B., Ghoushchi, S.J., Anari, S., Naseri, M., & Bendechache, M. (2021). Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. *Scientific Reports*, 11(1), 1-17. https://doi.org/10.1038/s41598-021-90428-8

Rasheed, Z., Ma, Y.K., Ullah, I., Al Shloul, T., Tufail, A.B., Ghadi, Y.Y., *et al*. (2023). Automated classification of brain tumors from magnetic resonance imaging using deep learning. *Brain Science*, 13(4), 602. https://doi.org/10.3390/brainsci13040602

Remzan, N., Tahiry, K., & Farchi, A. (2024). 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 and Applications*, 83, 0123456789. https://doi.org/10.1007/s11042-024-18780-1

Roopa, Y.M., Kumar, G.N., Harsha, Y.V.S., & Aditya, P.S. (2022). Detection of Brain Tumor Types Using Deep Learning. In: *Proceedings of the 2nd International Conference on Artificial Intelligence and Smart Energy, ICAIS*. Vol. 6, p459–465. https://doi.org/10.1109/ICAIS53314.2022.9742731

Santoso, I.B., Supriyono, & Utama, S.N. (2024). Multi-model of convolutional neural networks for brain tumor classification in magnetic resonance imaging images. *International Journal of Intelligent Engineering and Systems*, 17(5), 741–758. https://doi.org/10.22266/ijies2024.1031.56

Siva Raja, P. M., & Rani, A.V. (2020). Brain tumor classification using a hybrid deep autoencoder with Bayesian fuzzy clustering-based segmentation approach. *Biocybernetics and Biomedical Engineering*, 40(1), 440–453. https://doi.org/10.1016/j.bbe.2020.01.006

Sulistyowati, T., Purwanto, P., Alzami, F., & Pramunendar, R.A. (2023). VGG16 deep learning architecture using imbalance data methods for the detection of apple leaf diseases. *Moneter: Jurnal Keuangan Dan Perbankan*, 11(1), 41–53. https://doi.org/10.32832/moneter.v11i1.57

Wang, J., He, L., & Zhou, X. (2024). Optimizing Inception-V3 for Brain Tumor Classification Using Hybrid Precision Training and Cosine Annealing Learning Rate. In: *2024 7th International Conference on Advanced Algorithms and Control Engineering, ICAACE*, p528–532. https://doi.org/10.1109/ICAACE61206.2024.10548577

ZainEldin, H., Gamel, S.A., El-Kenawy, E.M., Alharbi, A.H., Khafaga, D.S., Ibrahim, A., Talaat, F.M. (2023). Brain tumor detection and classification using deep learning and sine-cosine fitness grey wolf optimization. *Bioengineering,* 10(1), 1-19. https://doi.org/10.3390/bioengineering10010018

AUTHOR BIOGRAPHY

Shahab Wahhab Kareem. He received his BSc in Control and Computer Engineering from the University of Technology, Baghdad, in 2001, MSc in Software Engineering from Salahadeen University in 2009, and Ph.D. in Yasar University Izmir, Turkey in 2020. My research interests include Machine learning and BIG DATA. He is a lecturer at the Information System Eng. (ISE) Department.

Nayla Faiq Othman. She received her BSc in Information Systems Engineering from Erbil Polytechnic University in 2018 and an MSc student in Information Systems Engineering from Erbil Polytechnic University in 2024. Her research interests include deep learning and machine learning.