Biomedical Engineering: Applications, Basis and Communications, (2024) 2430006 (16 pages)
DOI: 10.4015/S1016237224300062

World Scientific
www.worldscientific.com

EARLY DETECTION OF BRAIN TUMOR USING MRI AND TRANSFER LEARNING

Wael Korani*, Shyam Sundar Domakonda and Priyan Malarvizhi Kumar
Department of Information Science, University of North Texas
Texas, Denton, USA

Accepted 1 October 2024
Published

ABSTRACT

Brain tumors pose significant risks to cognitive functions, making early detection crucial for improving patient survival rates. Accurate tumor detection aids neuro-oncologists in diagnosing tumor types and recommending appropriate treatments. However, manual detection is challenging, time-consuming, and prone to human error. Recently, Deep Learning (DL) models have demonstrated substantial potential in efficiently classifying large image datasets. This paper presents three novel and efficient multi-class DL architectures leveraging transfer learning to classify different brain tumors. Our primary contribution is to enhance the predictive accuracy while minimizing the usage of computational resource compared to other state-of-art models in the literature by forgoing preprocessing, segmentation, hybrid models, and augmentation techniques. Additionally, we reduce the number of FC layers to streamline computation. We conduct extensive experiments to evaluate the performance of our models using the brain tumor Figshare T1-weighted contrast-enhanced MRI dataset, comprising 3064 images from three distinct tumor types. The InceptionV3 model records a 98.36% accuracy level with five-fold cross-validation. By incorporating batch normalization and optimizing the learning rate for the Adam optimizer, the Xception model reached 99.19% accuracy. Finally, we utilize Particle Swarm Optimization (PSO) to fine-tune the learning rate of the Stochastic Gradient Descent (SGD) optimizer. The Xception model attained 98.85% accuracy. These results highlight the novelty of our approach, offering a practical solution for neuro-oncologists, particularly through the fine-tuned Xception model, which delivers early and accurate tumor detection with minimal computational resources.

Keywords: Brain tumor; Transfer learning; Particle Swarm Optimization (PSO); Computer aided diagnosis; Convolutional neural network; Brain surgery.

INTRODUCTION

Brain is the main organ in human that controls all functions of the body, and it helps human to share thoughts and feelings. It is located in the human head and protected by the skull. The abnormal growth of cells in human brain might cause degenerative brain disease. Brain tumor is an abnormal mass of cells that can be developed within the brain or central nervous system. These tumors can be either benign (non-cancerous) or malignant (cancerous).[1] The exact causes of brain tumors are often unclear, but a combination of genetic factors, environmental influences, and exposure to radiation or certain chemicals may contribute to their development. Brain tumors manifest a range of symptoms including frequent and severe headaches, seizures, changes in cognitive function (i.e. memory problems and confusion), sensory disturbances, motor function changes (i.e. weakness or difficulty walking), nausea and vomiting. These symptoms are caused due to increased

*Corresponding author: Wael Korani, Department of Information Science, University of North Texas, Texas, Denton, USA. E-mail: wael.korani@unt.edu

intracranial pressure (ICP) inside human skull, which reduces the blood flow to the healthy tissue.[2] In Ref. 3, the American Cancer society reported in 2023 about 25,400 (14,420 males and 10,980 females) only malignant tumors of the brain and spinal cord. They also reported the death possibility of 18,760 patients from brain and spinal cord tumors.

Reducing the risk of brain tumors primarily involves healthy lifestyle and avoiding harmful environments. Human brain is a sensitive organ in human body that should not be exposed to harmful radiation and chemicals. In addition, adopting a healthy lifestyle that includes a balanced diet, regular exercise, and overall well-being has significant effect on human brain; however, brain tumor cannot be avoided in all cases. Brain tumors are classified into three main categories: glioma, meningioma, and pituitary tumors. Glioma originates in the glial cells, and can vary in malignancy. Meningioma is a benign tumor that forms in the meninges, which is the protective layer surrounding the brain. Pituitary is a benign tumor in most of the cases, and it is developed in the pituitary gland that affects hormone regulation.[4]

Identifying brain tumors typically involves physical exams, neurological exams, and images tests. The accurate ways to detect brain tumor and its stages are the medical imaging techniques such as magnetic resonance imaging (MRI), computed tomography (CT) scans, positron emission tomography (PET), Magnetic Resonance Spectroscopy (MRS), Single-Photon-Emission Computed Tomography (SPECT), or X-rays, where these techniques provide a clear image of human brain.[5] MRI is considered the best imaging technique because it provides clear image of human brains with more details about the structure of the cells and vascular supply. There are four MRI images: T1-weighted MRI (for healthy tissues), T2-weighted MRI (for edema boarder regions), T1-weighted contrast-enhanced MRI (for tumor borders), and Fluid-Attenuated Inversion Recovery (for edema regions in cerebrospinal fluid). These brain images are then interpreted by well-trained neuro-oncologists to identify the presence, type, and location of a tumor. The experienced neuro-oncologists are not available everywhere, and the delay in diagnosing brain tumor might cause serious problems. In addition, the diagnosis is subjective and it depends on the neuro-oncologists, which is difficult to quantify. In some cases, doctors cannot safely obtain a tissue sample (biopsy) to decide the best treatment plan, so that results are prone to human errors and variability. In addition, the biopsy procedure is painful and expensive. Thus, developing a computer-aided diagnosis (CAD) system to diagnose and locate brain tumor is essential.

CAD systems help doctors in diagnosing different diseases.[6] Machine learning (ML), including DL models, can process large volume of medical imaging data swiftly and consistently in a short time and high accuracy. These models excel in identifying intricate patterns and features within images, enabling highly accurate classification of objects based on their type, shape, and location. CAD systems with the aid of ML offer significant advantages in the field of brain tumor detection compared to traditional methods. They are successful in determining the stages of brain tumors as well, which allows doctors to decide the best course of treatment. Thus, CAD systems can be used to help neuro-oncologists in cases where manual tumor detection is time-consuming or even unavailable. The brain tumor CAD systems' results are consistent, efficient, and accurate; therefore, these systems enhance the diagnosis and treatment of brain tumors process, ultimately leading to better patient outcomes.[7] The motivation for this work stems from the need to bridge the gap between highly accurate DL models and their practical deployment in brain tumor diagnosis. Our goal is to propose a set of novel, simple, and efficient multi-class DL architectures that utilize transfer learning (TL) to classify brain tumors with minimal computational overhead. We aim to develop models that maintain high accuracy while reducing the complexity involved in model training, such as avoiding preprocessing techniques, segmentation, or data augmentation, which are common in many existing approaches. By minimizing the number of FC layers and optimizing the learning rate with techniques like Particle Swarm Optimization (PSO), this research seeks to present a cost-effective solution for accurate tumor classification. This not only enhances diagnostic capabilities but also makes early and precise detection more accessible, especially in clinical environments with limited computational resources, ultimately improving patient outcomes in brain tumor diagnosis and treatment.

The contributions of the paper are as follows:

- We propose novel highly accurate deep learning (DL) architectures based on TL to classify three common brain tumors: glioma, meningioma, and pituitary tumors.
- A novel architecture incorporating PSO is introduced to optimize the learning rate of the Stochastic Gradient Descent (SGD) optimizer.

- The models are designed to minimize computational resource usage compared to state-of-the-art proposed models in the literature by avoiding preprocessing, segmentation, extra hidden layers, and augmentation techniques.
- We fine-tuned several architectures by optimizing the number of FC layers and learning rate, while incorporating batch normalization and global average pooling to prevent overfitting.
- The architectures are evaluated using cross-validation (CV) strategy to void any bias and to overcome the shortcoming in previous studies.

RELATED WORK

In the past years, a plethora of research efforts related to the figshare brain tumor dataset (T1-weighted contrast enhanced MRI)[8] were undertaken. These existing works can be categorized into distinct approaches, including traditional ML models, custom Convolutional Neural Network (CNNs), Transfer Learning (TL) of pre-trained architectures, and hybrid models. This categorization provides a comprehensive overview of the diverse methodologies that were employed in previous research using the figshare dataset.

In Ref. 9, Saraswathi *et al.* introduced a multiclass brain tumor model (RF-PCA) using a random forest (RF) classifier and principal component analysis (PCA). The authors extracted the region of interest (ROI) using segmentation technique to isolate the tumor regions, and the output ROI was used for feature extraction using: gray level co-occurrence matrix (GLCM), shape and local binary pattern (LBP). Then, they applied PCA for dimensionality reduction. The proposed model was evaluated using figshare dataset, which contains 3064 T1-weighted contrast-enhanced brain tumor images. In fact, RF-PCA outperformed other approaches achieving a testing accuracy level of 85.56%. The experiments were conducted without CV. In Ref. 10, Jun Cheng *et al.* proposed a novel method to augment the tumor region. The authors used SVM after extracting the features using different techniques: intense histogram, GLCM, and bag of words. However, the best accuracy result was 91.28% with CV. In Ref. 11, a novel MGMT Promoter Methylation Prediction (MGMT-PMP) system is introduced for glioblastoma classification. The study employs a Deep Learning Radiomic Feature Extraction (DLRFE) module for feature extraction and uses a rejection algorithm to refine training data. Classifiers k-NN and SVM are applied to the BraTS-2021 dataset, achieving an accuracy of 96.84% with 10-fold CV, demonstrating the effectiveness of combining radiomic and DL features for predicting MGMT methylation status. In Ref. 12, the paper reviews ML techniques for automated pipeline condition assessment. It covers methods for detecting, classifying, locating, and quantifying pipeline anomalies using routine operation data, non-destructive testing, and computer vision. The review discusses performance metrics, SWOT analysis of different approaches, and provides guidance for implementing automated assessment systems. The study highlights the potential of ML to enhance efficiency and reliability in pipeline monitoring. In Ref. 13, a DL approach is proposed for real-time automatic interpretation of strain distributions from distributed fiber optic sensors (DFOSs) for crack monitoring. The method includes DL-based detection and localization of cracks, TL to handle limited data, and a split-and-merge technique to improve multi-crack detection accuracy. The approach achieved a mean average precision (mAP) of 0.968 for crack detection and processed DFOS data with 10,000 measurement points in under 0.05 s. The study demonstrates the effectiveness of DL for efficient and accurate crack monitoring. In Ref. 14, a novel deep neural network model for audio sentiment analysis is introduced. The model combines CNN and LSTM networks to generate Audio Sentiment Vectors (ASVs) from both spectrum graphs and traditional acoustic features like spectral centroid and MFCC. BiLSTM with an attention mechanism is used for effective feature fusion. The approach significantly outperforms previous methods, achieving a 9.33% improvement on the MOSI dataset compared to state-of-the-art techniques. The study demonstrates the model's capability to accurately and swiftly recognize audio sentiment.

In Ref. 15, the authors proposed Ultra-light Brain Tumor Detection (UL-BTD) for deep features extraction along with Gray Level Co-occurrence Matrix. The extracted features were utilized using SVM to propose a predictive tool, and the performance of the model was evaluated using figshare dataset. The average accuracy level was 99.23% using 10-fold CV. In Ref. 16, Ismael *et al.* developed a brain tumor classification framework using figshare images. The framework combined statistical features along with neural network algorithms. The authors implemented 2D Discrete Wavelet Transform (DWT) and 2D Gabor filter techniques on extracted ROI for features extraction. The proposed model was evaluated on figshare dataset and achieved an accuracy level of 91.90% without CV strategy.

In Ref. 17, Sunanda Das *et al.* proposed a custom CNN model and evaluated the model using figshare dataset. The authors performed pre-processing equalization to enhance the contrast of the images. The results show a testing accuracy of 94.39% without using CV strategy. In Ref. 18, Nyoman *et al.* proposed five custom CNN models and evaluated the models using figshare dataset. The authors did not perform any pre-processing operation. The results show training accuracy of 98.51% and a validation accuracy of 84.19% without using CV strategy. In Ref. 19, Francisco Javier *et al.* introduced a custom CNN using a sliding window of size 65×65 and three pathways using different filters' size, which is time consuming. The results show an accuracy level of 97.3% using five-fold CV strategy. In Ref. 20, Sultan *et al.* proposed a custom CNN model using 16 layers and evaluated the model on the figshare dataset. The results show that the proposed model achieved accuracy levels of 96.13%. However, the authors used different augmentation method to increase the number of images to 15,320 images. In addition, they did not use CV strategy to avoid bias in their results. In Ref. 21, Ali Mohammad Alqudah modified his old custom CNN architecture that consists of 18 layers. The authors conducted different experiments using the modified architecture and evaluated the model on the figshare dataset. The best accuracy was 99% for the uncropped lesions without performing CV. In Ref. 22, Farhad Hossain *et al.* proposed a custom CNN architecture and evaluated the performance of the model using figshare dataset. The authors used normalization and segmentation to pre-process the images. The accuracy was 96.90% without employing CV in his work. In Ref. 23, Badža *et al.* introduced a custom CNN architecture and used two different augmentation transformations. The authors conducted different reliable experiments without using CV and with 10-fold CV using figshare dataset. The best accuracy was 96.56% using 10-fold CV. In Ref. 24, Paul *et al.* proposed a custom CNN and used four augmentation transformations. The performance was evaluated using figshare dataset and using five-fold CV. The results show that accuracy level was 91.43%. In Ref. 25, Zhiguan Huang *et al.* proposed a custom CNN based complex network and optimized different activation functions. The model was evaluated using figshare dataset and achieved an accuracy of 95.49% without CV. The model outperformed several pretrained architectures. In Ref. 26, Mesut Toğaçar introduced a custom CNN architecture. The authors used otsu for segmentation and used six different transformations for augmentation. The model was evaluated using figshare, and the best accuracy was 9.57% without CV and using 20% testing. In Ref. 27–29, Afshar *et al.* proposed few custom CNNs along with the capsules structure (CapsNets). The authors evaluated the proposed model using figshare dataset, and the best accuracy was 86.56%, 90.89%, and 78%, respectively, without using CV. However, the result was not promising in using capsules concept in brain tumor detection. In Ref. 30, Xiaohao Du *et al.* proposed a custom CNN using randomly wired neural network (RWNN) and figshare dataset. The results was 95.33% level of accuracy without CV.

In Ref. 31, Mohamed Ait Amou *et al.* used Bayesian Optimization to optimize several hyperparameters: activation function, batch size, dropout rate, number neurons in flatten later, and the optimizer. The authors evaluated different pretrained architectures: VGG16, VGG19, ResNet50, InceptionV3, and DenseNet201. The optimized pretrained architecture using 90% training and 10% testing and without using CV achieved 98.7%. In Ref. 32, Benjamin Maas *et al.* implemented a modified version of QuickNAT architecture for segmentation and classification. The performance was evaluated using figshare dataset, and the accuracy of the model was 99.35% without CV. In Ref. 33, Alamin Talukder *et al.* developed a novel DL models based on TL. The authors used extensive preprocessing, TL architectures such as Xception, ResNet50V2, InceptionResNetV2, and DenseNet201. The performance of the models was evaluated using figshare dataset to find the best model. The best accuracy level was 99.68% for ResNet50V2 without CV. In Ref. 34, Prince Priya Malla *et al.* introduced a TL-based model utilizing VGG16 architecture. The proposed model assessed on the figshare dataset and achieved accuracy level of 98.93% without using CV. In Ref. 35, Sudhakar Tummala *et al.* adopted four vision transformer (ViT) models that were pretrained on ImageNet dataset. The models were evaluated on figshare dataset using five-fold CV. The best reported accuracy was 98.7%. In Ref. 36, Zar Nawab *et al.* adopted TL from VGG19 architecture. The authors evaluated the performance using figshare dataset and achieved an accuracy level of 94.81% using five-fold CV. In Ref. 37, Amin ul Haq *et al.* used TL from pretrained architecture such as VGG16, Inception V3, and Xception, and they used two augmentation transformations. The proposed models were evaluated using figshare dataset, and the best accuracy was 99.9 using ResNet50 without CV. In Ref. 38, Arumaiththurai *et al.* used a preprocessing segmentation to isolate IOR. The authors extracted 16 statistical features and implemented SVM, Decision and

Tree (DT) classifiers. The results were compared to VGG19 and ResNet152. The models were evaluated using figshare dataset and CV strategy. The results show the best accuracy was 94.67% using Resnet151. In Ref. 39, Momina Masood *et al.* introduced a custom Mask Region-based Convolution neural network (Mask RCNN) with a densenet-41 backbone for brain tumor classification and segmentation. The model was evaluated using figshare dataset and achieved an accuracy 98.34% without CV. In Ref. 40, the authors proposed a model using VGG-16 and figshare dataset. The authors used GLCM and the original images as an input. The results showed that the accuracy was 96.5% without CV. In Ref. 41, Arshia Rehman *et al.* used different pretrained models: AlexNet, GoogLeNet, and VGG16. The authors used different augmentation transformation, and evaluated the models using figshare dataset. The Best result was 98.69% using VGG16 without CV.

In Ref. 42, Deepak *et al.* proposed an approach to overcome the imbalanced data problem. The authors used the features fusion from different models that were created using different loss functions. Then, the output features were used as inputs for Support Vector Machine (SVM) and *K*-Nearest Neighbor classifiers. They assessed the models using figshare dataset, and the best results was 95.4% without CV. In Ref. 43, Xiaoqing Gu *et al.* introduced an approach called Convolutional Dictionary Learning with Local Constraint (CDLLC). The authors used the extracted features as an input to *k*NN classifier, and the performance was evaluated using figshare dataset, and achieved level of 96.93% using five-fold CV. In Ref. 44, Deepak *et al.* introduced three models using pretrained GoogLeNet, GoogLeNet + SVM, and GoogLeNet + *k*NN. The models were evaluated on figshare dataset, and the best accuracy was reported using a hybrid model of GoogLeNet and *k*NN. The accuracy level was 98.0% using five-fold CV. In Ref. 45, Pashaei *et al.* proposed a hybrid model from a custom CNN and several classifiers. The proposed models were evaluated using figshare dataset. The results show that best model (custom CNN and kernel Extreme Learning Machines) achieved accuracy level of 93.68% using 10-fold CV. In Ref. 46, Almetwally Mohamad Mostafa *et al.* used segmentation and Gaussian filter for preprocessing on a portion of figshare dataset. Then, the authors used the fusion of features principle from Histogram of Oriented Gradient (HOG) and DarkNet 19. The fused features were utilized as input to SVM, Naive Bayes, and Long–Short-Term Memory (LSTM). The best result was 99.3% using LSTM. In Ref. 47, Ahmad Kazemi *et al.* combined AlexNet and VGGNet for feature extraction, and they used a feature selection technique to select the best group of features. The extracted features were classified using different classifier: SVM, DT, and *K*NN. The proposed framework was evaluated using different dataset dataset. The results show an accuracy level of 98.49% for figshare dataset using 10-fold CV. In Ref. 48, Asaf Raza *et al.* introduced a hybrid DL classification model, and evaluated the model on figshare dataset. The results show an accuracy level of 99.67% without using CV strategy.

List of limitations in previous studies is mentioned as follows:

- Prepossessing for images using equalization requires time and computational resources as in Ref. 17.
- Segmentation methods and their parameter selections are required more computational resources. In Refs. 9, 16, 22, 26, 38, 46 and 39, different segmentation approaches were implemented.
- Sliding window approach takes more time and resources.[19] In addition, we use ViT require large dataset and more computational resources because it divides images into many batches.[35]
- Increasing the number of images using augmentation transformation consumes more computational resources. In Refs. 20, 23, 24, 26, 33, 37 and 41, authors increased the number of images between three to five times to improve the performance of their proposed model.
- CV strategy has significant effect in reducing the bias in the results. In Refs. 9, 17, 18, 20–22, 25, 27–29, 31–34, 37, 41, 42, 46, 48 and 30, authors did not use CV. In Refs. 23 and 39, the authors showed the significant effect of using CV in their exterminates. The results showed that in all cases that did not implement CV, the reported accuracy was higher than if it was implemented in CV.
- A small portion of the dataset was used for testing to improve the accuracy, and when a lager portion was used, the accuracy was dropped.[31]
- Using hybrid models requires more computational resources, as in Refs. 42–48.

To overcome the aforementioned limitations, we conduct extensive experiments using CV to avoid any bias in our results and to propose robust and reliable models. To optimize the computational resources, we do not use any technique that might consume more computational resources. In addition, our models maintain a leaner structure with less number of layers, contributing in reducing the computational expenses and faster training. Furthermore, we run many experiments

to show the significant effect of the learning rate to improve the performance of our proposed models.

The rest of this paper is organized as follows.

In the first section, we discuss different models and their limitations: traditional ML, DL, and hybrid models from previous studies. In the second section, we present the dataset and preprocessing, TL architectures, our proposed architectures, our experiments, classifier settings and performance metrics and evaluation. In the third section, we outline our experiments results, an extensive comparison with previous studies and regarding misclassifications. In the fourth section, we provide a discussion for our results and the best recommended models. In the fifth section, we draw the conclusion of the paper.

MATERIALS AND METHODS

In this section, we discuss the proposed methods that includes data preprocessing and pre-trained CNNs architectures. The concept of TL is implemented using different successful architectures. The block diagram of our proposed research is shown in Fig. 1(A).

Materials

In 2017, Cheng published a brain tumor dastaset at figshare.[8] The dataset includes T1-weighted contrast-enhanced MRI images that are presented as grey scale images with three channels. It consists of 3064 images that were taken from 233 patients who were diagnosed with one of three brain tumors: glioma, meningioma, or pituitary tumor. The number of images in each brain tumor category is: 708 images (Meningioma), 1426 images (Glioma), 930 images (Pituitary) as shown in Fig. 1(B). Each image uniformly maintains a dimensions of (512×512), adding a crucial consistency to the dataset's structure. The dataset is formatted using MATLAB in file format (.mat), which efficiently stores multidimensional arrays and complex data structures. Each matlab file contains a description of an image: tumor type ($0 \rightarrow$ meningioma, $1 \rightarrow$ glioma, $2 \rightarrow$ pituitary), patient ID (PID), image data, tumor boarder, tumor mask.

Transfer Learning

TL, a specialized form of DL, enhances brain tumor classification by leveraging pre-trained models. This approach allows models to transfer knowledge from tasks that are previously learned. There are many models that are proposed using ImageNet Large Scale

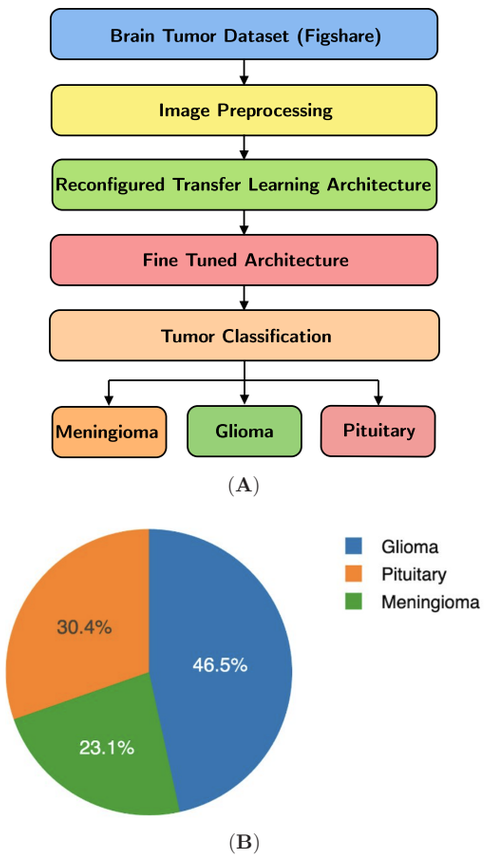

Fig. 1. (**A**) The process of our proposed framework and (**B**) the distribution of tumor types in the figshare dataset.

Visual Recognition Challenge (ILSVRC) database such as AlexNet, VGG16, VGG19, GooglNet and Inception. These models can quickly adapt to new problem (dataset) and achieve high levels of accuracy, even when working with limited datasets. This is particularly valuable in medical imaging, where obtaining a substantial amount of labeled data for training is often a challenge.

In this paper, we utilize various TL models, including Xception, Inceptionv3, DenseNet201, and EfficientNetV2L. The FC layers in the pre-trained architectures are dropped. Instead, we introduce an optimized group of FC layers with an output size of three classes. These architectures play a crucial role in our study that facilitate the adaptation of pre-trained models to address our specific research objectives. By utilizing TL models, we introduce three novel architectures that perform exceptionally well across all metrics. These proposed models are implemented to efficiently and effectively classify brain tumors and presenting promising potential for various other medical applications. We conduct three experiments to introduce and validate our three novel developed models.

Proposed Classification Frameworks

In architecture-1, we utilize a pre-trained models as the basic part of our architecture and introduce several modifications to create a novel framework, as shown in Fig. 2(A). One of the key features is the integration of a global average pooling layer, which effectively condenses the extracted features to reduce spatial dimensions while retaining essential information. In addition, a dense output layer utilizing softmax activation function is incorporated to enable precise multi-class classification, specifically distinguishing among meningioma, glioma, and pituitary tumors. These architectural adjustments are pivotal in transforming rich feature representations into meaningful outcomes, significantly improving the accuracy of our model for classification purposes.

In architecture-2, we utilize the first model and some other adaptations. In the proposed model, we add to the base model a global average pooling layer, Batch Normalization to enhance feature extraction, and a custom dense output layer tailored for the precise classification of brain tumor images, as shown in Fig. 2(B). The proposed model, rooted in TL techniques, introduces meticulous adjustments and refinements, optimizing its capacity for effectively classify image brain tumors. These modifications are added and evaluated in extensive experimentation to elevate the architecture performance for the intricate task of brain tumor classification. In this architecture, we adapt and optimize different mechanisms to choose the best mechanism for brain tumors. For instances, Early Stopping mechanism is employed to monitor the validation loss, mitigating the risk of overfitting problem. Simultaneously, the ReduceLROnPlateau strategy dynamically optimizes the learning rate during the training process, promoting enhanced convergence and performance. Our approach highlights the significance of customizing TL methodologies to specifically cater to the domain of brain tumor classification. The adaptability of the proposed architecture introduces a robust and precise classification outcomes.

In architecture-3, a novel architectural framework is introduced using TL and utilized PSO to fine-tune the learning rate. We use PSO to optimize the learning rate of the SGD optimizer. The PSO algorithm is used to iteratively search and determine the best learning rate for the SGD optimizer within a specified range [0.0001, 0.1]. The PSO algorithm aimed to enhance the model convergence and performance by fine-tuning the learning rate for optimal training. In this architecture, we add the learning rate optimization to the previous model. Thus, we include what we

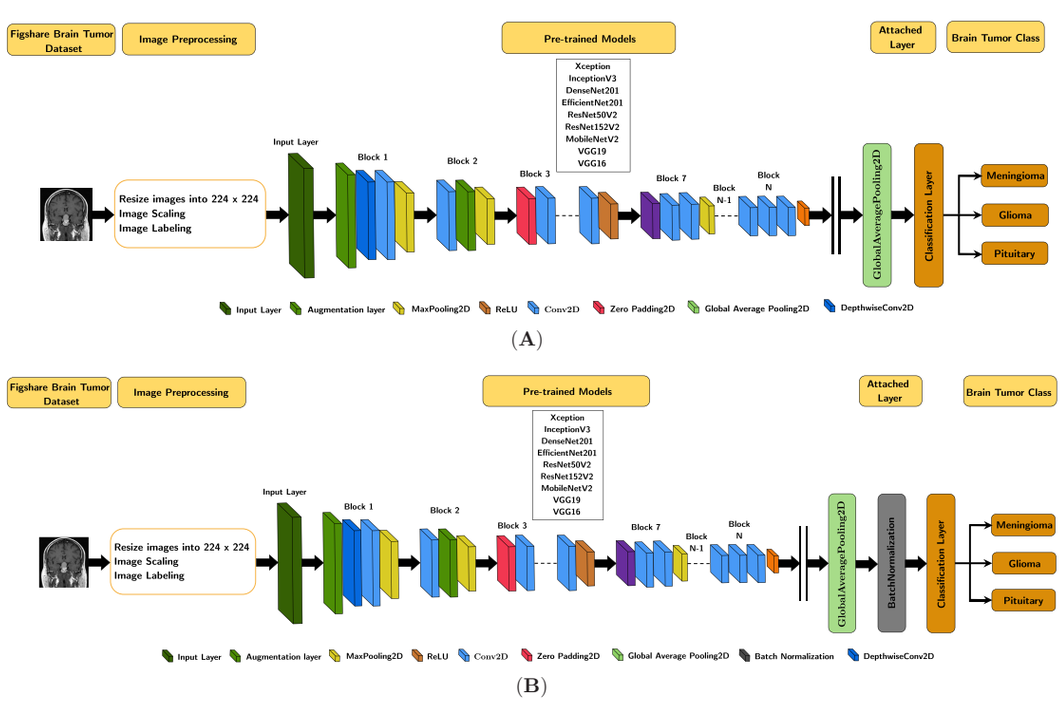

Fig. 2. (**A**) model-1 The pretrained architectures and global average pooling layer and (**B**) model-2 the pretrained architecture and global average pooling along with batch normalization.

optimized in the previous experiments and a global average pooling layer. The modifications and the integration of PSO for learning rate optimization are implemented to enhance the adaptability and accuracy of the our proposed model.

Experiments

Our experiments are conducted on a MacBook M1 that has eight-core CPU, comprising four performance cores and four efficiency cores, an eight-core GPU, and a 16-core Neural Engine. The system's computational power, coupled with the utilization of TL techniques, enabled the efficient training and evaluation of our models.

In the data preprocessing, we extract the images data and their labels from (.mat) files and convert them into JPEG format. Then, we reduce the size of the images to (224×224) in order to optimize the usage of the computational resources, which helps the CNN to perform well in short time. The images are then shuffled before splitting to avoid any bias. The shuffled dataset is divided into 80% (training and validation) and 20% testing. This partitioning aims to ensure a balance between training and testing. In addition, the five-fold CV strategy is implemented to avoid bias and to ensure a robust assessment of our proposed models.

Our framework evaluation utilizes with the testing set, which accounted for 20% of the original dataset. Collectively, these subsets: training, validation, and testing form a meticulously designed hierarchy for training, refining, and evaluating our classification models. The size of these subsets emphasizes the structure and rigorous nature of our approach to robustly validate our model's efficacy and real-world applicability. Then, we standardize the images for model training and evaluation. The pixel values are standardized in range [0, 1]. This foundational step ensured consistent data representation and enhanced model convergence during subsequent training and testing phases.

Classifier Settings

The effectiveness of an image classification system is contingent upon the interplay between image features and the chosen classifier model. In the context of our brain tumor classification system, the concluding phase involves the utilization of classifier models. We conduct three distinct experiments, highlighting our systematic approach to evaluate the proposed brain tumor classification models.

Architecture-1: In our pursuit of an effective brain tumor classifier, we optimize a pre-trained architecture for image-related tasks. We adapt several TL models by adding a Global Average Pooling layer to consolidate and refine feature extraction. The final design comprises a dense output layer with three neurons, aligned with the classes in our brain tumor classification task, ensuring a specialized model for accurate categorization.

The parameters employed in this architecture are selected for optimal convergence and including softmax activation function ensured that the model generated probability distributions across these classes, facilitating informed classification decisions. In this architecture, we utilize Adam optimizer, which is well-regarded for its learning rate adaptability and efficiency. The loss function employed is sparse categorical cross-entropy that is suitable for multi-class classification tasks. We tune and optimize other hyperparameters including batch size and number of epochs, and we conclude the best batch size is 32 and number of epochs is 10. It is important to note that the default initial learning rate for the Adam optimizer is 0.001. This configuration allows our model to learn and adapt from the dataset effectively, ultimately leading to improve our classification performance. Table 1 summarizes the hyperparameters settings employed in our experiments.

Architecture-2: To fine-tune the model for our specific brain tumor classification task, we propose significant architectural adjustments. In this architecture, we integrate a Global Average Pooling layer that is known for its feature extraction and dimensionality reduction capabilities. This layer plays a pivotal role in preserving essential information while mitigating overfitting risks. In addition, we introduce Batch-Normalization, a crucial technique that stabilizes training by normalizing activation within the network. These aforementioned modifications enhance our model convergence and generalization. The dense output layer of our model has three neurons, each corresponding to one of the brain tumor classes. Employing the softmax activation function ensured that the model generated probability distributions across these classes, facilitating confident and informed classification. Our models are optimized using Adam optimizer with a lowered initial learning rate, specifically 1e-6. For optimal model performance during training, we adjust the learning rate to 1e-6. This fine-tuning effectively reduces the risk of learning divergent patterns from the pre-trained weights and is built on careful experimentation and observation of convergence dynamics. Our training process includes 10 epochs, with data batches comprising 32 samples

Table 1. Experimental Parameters.

Models	Pre-trained Models	Parameters	Classifier Setting
Architecture-1	InceptionV3	optimizer	adam
	DenseNet201	initial-learning rate	0.001
	ResNet50V2	batch size	32
	ResNet152V2	loss function	sparse categorical entropy
	Xception	epochs	10
	MobileNetV2		
	VGG19		
	VGG16		
	EfficientNetV2L		
Architecture-2	Xception	optimizer	adam
	DenseNet201	initial-learning rate	1e-6
	EfficientNetV2L	batch size	32
	ResNet50V2	loss function	sparse categorical entropy
	ResNet152V2	ReduceLROnPlateau	min_lr = 1e-8
	InceptionV3	epochs	10
	MobileNetV2		
	VGG19		
	VGG16		
Architecture-3	Xception	optimizer	SGD
	InceptionV3	learning_rate	max(0.0001, min(0.1, learning_rate))
	MobileNetV2	momentum	0.9
	DenseNet201	batch size	32
	ResNet152V2	loss function	sparse_categorical_crossentropy
	ResNet50V2	epochs	10
		ps.single.GlobalBestPSO	n_particles = 10, dimensions = 1
		iterations	10

each. To enhance training efficiency and effectiveness, we incorporate two essential callbacks: Early-Stopping and ReduceLROnPlateau. Early-Stopping diligently monitors validation loss, allowing us to halt training if it plateaued for an extended period to prevent overfitting. ReduceLROnPlateau dynamically adjusts the learning rate during training, ensuring smoother convergence by reducing the learning rate when validation loss stagnated. Importantly, we set the minimum learning rate lr_{min} to 1e-8 to avoid excessively diminishing the learning rate, which could hamper convergence. We provide a concise summary of the parameters employed in conjunction with the experiment-2 in Table 1.

Architecture-3: In this architecture, we aim to optimize the training process of TL architectures, with a specific focus on enhancing the learning rate of the SGD optimizer. To achieve this, we employ a novel approach by incorporating PSO to dynamically adjust the learning rate during model training. Additionally, we incorporate momentum into the SGD optimizer, with a value set at 0.9 to improve the convergence process. This parameter further influences the optimization process, contributing to the model's performance.

Our basic architecture is built on TL. To adapt these architectures to our classification task, we introduce a global average pooling layer followed by a dense output layer. To effectively optimize the learning rate, we leverage PSO, which is a population-based optimization algorithm. The primary objective is to identify the optimal learning rate that would accelerate convergence without causing any overfitting. We initialize the PSO with a population size of 10 particles, and run it for a total of 10 iterations. In addition, we set the model batch size to 32 and training spanned 10 epochs. Table 1 shows all setting parameters for this experiments. This study highlights the potential of PSO as an effective tool for fine-tuning SGD learning rates, ultimately leading to more efficient and proficient training of deep neural networks for diverse applications.

Performance metrics and evaluation

Various performance metrics can be employed to assess classifier performance; however, classification accuracy remains a fundamental and frequently utilized metric. In our research, we place a primary focus on accuracy as the key evaluation metric for our proposed models. Classification accuracy is defined as the proportion of correctly classified data samples to the total number of data samples evaluated. The accuracy metric

underscore the efficacy and reliability of our proposed models within the context of our research.

Our Brain tumor classification centers on an imbalanced dataset, demanding a broader evaluation. To comprehensively assess the proposed models performance, we employ confusion matrices as a vital tool to delve deeper into the intricacies of our tumor classification system. Confusion matrix provides a concise representation, tabulating both actual and predicted classifications. we introduce an in-depth analysis of precision, recall, and specificity, all pivotal performance metrics pertinent to our classifier's evaluation. These equations are derived from the confusion matrix, which provides valuable insights into our classifier's effectiveness.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}, \tag{1}$$

where True Positives means TP, False Positives means FP, False Negatives means FN and True Negatives means TN. Precision quantifies the accuracy of positive predictions by calculating the ratio of TP to the total predicted positives.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{2}$$

Recall recognizes sensitivity, measures the classifier's capacity to identify all relevant positive samples. It is calculated as the ratio of TP to the total actual positives:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{3}$$

Specificity gauges the model's aptitude in recognizing true negatives among the total actual negatives. It is determined as the ratio of TN to the sum of true negatives and false positives:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \tag{4}$$

The harmonic mean of precision and recall, which yields the critical statistical measure, is known as the *F*-score for each class:

$$F_{\text{score}} = \frac{2}{3} \sum_{c=1}^{3} \frac{\text{precision}_c \cdot \text{recall}_c}{\text{precision}_c + \text{recall}_c}. \tag{5}$$

Given the inherent class imbalance among the three categories, we compute the average *F*-score (F_{avg}) using the following formula:

$$F_{\text{avg}} = \frac{F_0 + F_1 + F_2}{3}. \tag{6}$$

This metric offers a comprehensive assessment of our models performance, considering the nuances of the imbalanced classes.

RESULTS

In our "Experiment 1", we run our proposed architecture with various TL architectures, including InceptionV3, DenseNet201, ResNet50V2, ResNet152V2, Xception, MobileNetV2, VGG19, VGG16, and EfficientNetV2L. Our results are listed in Table 2. The results show that InceptionV3 architecture yielded the highest classification accuracy, achieving an impressive 98.36%. This model consists of 21.81 million parameters, showcasing its capability to handle complex feature extraction and classification tasks effectively. Following closely, the ResNet50V2 model achieved an accuracy of 97.71%. These results signify InceptionV3 as the top-performing model among the executed architectures with our proposed architecture, demonstrating its effectiveness in solving our classification problem. Furthermore, InceptionV3 shows superior performance while maintaining computational efficiency, making it a compelling choice for our application. Our proposed model shows high performance in comparison to other architectures in the literature. In addition, our model is more efficient than other previous studies from computational resources optimization, which makes it suitable for real-world deployment in medical image classification task.

In "Experiment 2", we enhance the performance of various TL architectures with our proposed architectures. We run our proposed architecture with various TL architectures, including Xception, DenseNet201, EfficientNetV2L, ResNet50V2, ResNet152V2, InceptionV3, MobileNetV2, VGG19, and VGG16. The outcomes of our experiments are presented in Table 2. The results show remarkable improvements in classification performance. Particularly, Xception architecture achieves an outstanding accuracy level of 99.19% and consists of 20.88 million parameters, demonstrating its efficiency in handling the classification task with a reduced parameter count compared to other models. DenseNet201 and EfficientNetV2L are closely followed with an accuracy level of 98.69%, establishing them as the second and third top-performing models among the tested architectures. Our models contribute to the field of brain tumor classification, providing highly accurate and efficient models that are well-suited for real-world applications in medical imaging. These models not only

Table 2. Classification Results for Brain Tumor Classification.

	Precision			Recall			Specificity			*F*1_Score			
Model	**M**	**G**	**P**	**M**	**G**	**P**	**M**	**G**	**P**	**M**	**G**	**P**	**Acc**
Experiment-1													
InceptionV3	98.09	97.79	99.46	95.65	98.88	100	99.34	98.26	99.79	96.86	98.34	99.73	**98.36**
DenseNet201	98.70	97.10	99.45	94.41	99.63	99.49	99.56	97.67	99.77	96.51	98.75	99.45	**98.20**
ResNet50V2	98.04	96.38	99.46	93.17	98.88	100	99.34	97.09	99.77	95.54	97.61	99.73	97.71
ResNet152V2	98.64	93.40	99.44	90.06	100	96.72	95.56	94.48	99.77	94.16	96.59	98.06	96.41
Xception	95.06	99.21	81.33	80.12	93.31	100	98.67	99.42	90.23	87.16	96.17	89.77	91.84
MobileNetV2	44.64	60.59	100	31.06	100	31.15	86.28	49.13	100	36.63	75.46	47.59	61.33
VGG19	NaN	43.88	NaN	0	100	0	100	0	100	NaN	61.00	NaN	43.88
VGG16	NaN	43.88	NaN	0	100	0	100	0	100	NaN	61.00	NaN	43.88
EfficientNetV2L	23.61	45.10	NaN	10.56	90.71	0	87.83	13.66	100	14.59	13.66	NaN	42.57
Experiment-2													
Xception	99.37	98.53	100	97.52	99.63	100	99.78	98.84	100	99.43	99.08	100	**99.19**
DenseNet201	99.35	97.81	99.46	95.65	99.63	100	99.78	98.26	99.77	97.47	98.71	99.75	**98.69**
EfficientNetV2L	99.35	98.17	98.12	95.65	99.63	100	99.78	98.55	99.53	97.47	98.89	99.46	**98.69**
ResNet50V2	97.47	98.15	98.92	95.65	95.51	100	99.12	98.55	99.53	96.55	98.33	99.46	98.20
ResNet152V2	97.45	97.44	99.45	95.03	99.88	99.45	99.12	97.97	99.77	96.23	98.15	99.45	98.04
InceptionV3	98.05	96.73	99.46	93.79	98.88	100	99.34	97.38	99.77	95.87	97.79	99.73	97.87
MobileNetV2	96.23	98.50	97.86	95.03	97.77	100	98.67	98.84	99.07	95.63	98.13	98.92	97.71
VGG19	82.31	86.91	94.38	75.16	92.94	91.80	94.25	88.95	97.67	78.57	89.77	93.07	87.92
VGG16	78.29	87.00	94.57	73.91	89.59	95.08	92.70	89.53	97.67	76.04	88.28	94.82	87.11
Experiment-3													
Xception	98.12	98.52	100	97.52	98.88	100	99.34	98.84	100	97.82	98.70	100	**98.85**
*Inception*V3	97.50	99.25	98.39	96.89	98.51	100	99.12	99.42	99.30	97.20	98.88	99.19	98.53
MobileNetV2	94.51	98.87	98.37	96.27	97.40	98.91	98.01	99.13	99.30	95.38	98.13	98.64	97.55
DenseNet201	94.41	96.30	100	94.41	96.65	99.45	98.01	97.09	100	94.41	96.47	99.73	96.90
ResNet152V2	94.63	93.21	97.83	87.58	97.03	98.36	98.23	94.48	99.07	90.97	95.08	98.09	94.94
ResNet50V2	86.36	91.14	96.81	82.61	91.82	99.45	95.35	93.02	98.60	84.44	91.48	98.11	91.68

excel in performance, but also minimize computational costs, making them practical models for brain tumor diagnostics.

In "Experiment 3", we delve into an exploration of various TL architectures, encompassing well-established models: Xception, InceptionV3, MobileNetV2, DenseNet201, ResNet152V2, ResNet-50V2. In these experiments, our primary objective is to enhance the efficiency and performance of these models. The results are recorded in Table 2. The results show that the Xception model demonstrated exceptional performance, achieving a classification accuracy level of 98.85% and consisting of 20.87 million parameters. This highlights the model's ability to deliver high accuracy with a relatively compact architecture. Following closely, the InceptionV3 model attained an accuracy level of 98.53%, and the MobileNetV2 model secured an accuracy level of 97.55%. It is important to emphasize that these models exhibit outstanding performance while remaining computationally efficient. This combination of high accuracy, efficiency, and adaptability positions them as strong candidates for practical deployment in real-world medical image classification tasks.

Performance of Fine-Tuned Xception Model Using ROC Analysis

The best model in all experiments is Xception model-2. In the context of evaluating the performance of ML models, the Receiver Operating Characteristic (ROC) curve plays a significant role in evaluating the performance of the proposed model. An ideal ROC curve is the one where the Area Under the Curve (AUC) equals unity, which indicates perfect classification. This metric is critical in assessing how effectively a model can distinguish between different classes or categories.

The ROC curves and their corresponding AUC values for a finely-tuned Xception model of our model-2 are particularly noteworthy, with the AUC value for all three classes being equal to 1. The results signify that the model achieves the highest level of accuracy and precision in classifying brain tumors. This performance shows the model's extraordinary performance. The recorded performance underscores its potential to revolutionize the medical field by providing fast and accurate brain tumor classification. The ability to swiftly and accurately categorize brain tumors has the potential to greatly impact patient care by enabling

early diagnosis and precise treatment planning. This underlines the importance and utility of our model in the medical field, where it can significantly improve the efficiency and accuracy of brain tumor classification, ultimately benefiting patient outcomes and advancing the field of medical imaging.

Comparison with Related Works

We conduct an extensive performance comparison of our proposed models with previous proposed models using the same dataset T1-weighted contrast-enhanced MRI of brain tumors including most recent publications. Table 3 shows a comparative analysis primarily hinges on classification accuracy as the central metric. It also records some other significant parameters: the name of the TL architecture, number of epochs, CV strategy implementation, and size of the training dataset that was implemented by other researchers.

Our best proposed model outperforms all previous studies to date. It is important to highlight that very few studies achieved slightly higher accuracy; however, those studies either used: hybrid architectures, many number of epochs, extensive preprocessing and augmentation, and/or without cross validation strategy. In Ref. 37, the authors implemented hybrid architectures (ResNet50 and CNN) and 100 epochs, which are computationally expensive and time-consuming, as additional complexities. In Ref. 15, the authors utilized two different features techniques: custom CNN of 15 layers and GLCM. In addition, the authors have used two different preprocessing techniques and augmentations to increase the number of images, which require more computational resources. In Refs. 33, 34, 48 and 27, the authors did not implement the cross validation strategy, which causes bias results.

We consider more evaluation of our proposed models in terms of precision, recall, sensitivity, specificity, and $F1$ score. Table 4 shows a comparative analysis

Table 3. Related Works and Comparative Analysis on the Figshare Dataset.

Work	Year	Method	Training Data	Cross-Validation	Epochs	Accuracy
Talukder *et al.*[33]	2023	ResNet-50	80%	Without	40	99.68%
Malla *et al.*[34]	2023	VGG-16 CNN	70%	Without	100	98.93%
Raza *et al.*[48]	2022	Hybrid DeepTumorNet	70%	Without	10	99.67%
Amou *et al.*[31]	2022	Optimized CNN	90%	Without	1200	98.70%
Deepak and Ameer[42]	2022	CNN-KNN	60%	Without	—	95.6%
Benjamin *et al.*[32]	2021	QuickTumorNet	70%	Without	200	99.27%
Masood *et al.*[39]	2021	Mask-RCNN	70%	Without	45	98.34%
Hossain *et al.*[22]	2021	ConvNet	70%	Without	—	96.90%
Arumaiththurai *et al.*[38]	2021	ResNet-152	70%	Without	30	94.67%
Rehman *et al.*[41]	2020	Fine-tune VGG16	—	Without	30	98.69%
Mesut *et al.*[26]	2020	BrainMRNet	80%	Without	100	97.69%
Belaid and Loudini[40]	2020	VGG16	—	Without	400	96.5%
Alqudah *et al.*[21]	2019	CNN	70%	Without	—	99.19%
Sultan *et al.*[20]	2019	CNN	—	Without	—	96.13%
Das *et al.*[17]	2019	CNN	—	Without	100	94.39%
Afshar *et al.*[28]	2019	CapsNet	—	Without	50	90.89%
Ismael *et al.*[16]	2018	DWT-Gabor-NN	70%	Without	—	91.90%
Afshar *et al.*[27]	2018	CapsNet	—	Without	10	86.56%
Abiwinanda *et al.*[18]	2018	CNN	—	Without	10	84.19%
Afshar *et al.*[29]	2020	Bayescap	80%	Without	500	78%
Huang *et al.*[25]	2020	modified CNNBCN+ER	—	—	—	95.49%
Du *et al.*[30]	2020	RWNN+WS	—	—	—	95.33%
Haq *et al.*[37]	2022	ResNet50-CNN	70%	With	100	99.90%
Kazemi *et al.*[47]	2022	Parallel CNN	75%	With ($k=10$)	—	98.49%
Tummala *et al.*[35]	2022	Ensemble ViT	70%	With ($k=5$)	25	98.7%
Ayan and Shrivastava Vimal[49]	2022	BMRI-Net + PFpM	70%	With($k=5$)	50	98.45%
Daz-Pernas *et al.*[19]	2021	Multiscale CNN	80%	With ($k=5$)	80	97.3%
Gu *et al.*[43]	2021	CDLLC	70%	With ($k=5$)	—	96.93%
Badža *et al.*[23]	2020	CNN	—	With ($k=10$)	—	96.56%
Deepak and Ameer[44]	2019	Deep CNN-SVM	56%	With ($k=5$)	10	97.1%
Pashaei *et al.*[45]	2018	CNN-ELM	70%	With ($k=10$)	10	93.68%
Paul *et al.*[24]	2017	CNN	—	With ($k=5$)	100	91.43%
Cheng *et al.*[10]	2015	BOW-SVM	80%	With ($k=5$)	—	91.28%
Proposed	**2023**	**Xception**		**With ($k=5$)**	**10**	**99.19%**

Table 4. Comparing the Performance using Additional Metrics for Meningioma (M), Glioma (G), and Pituitary tumor (P) Classification.

	Precision			Recall			Specificity			
Work	**M**	**G**	**P**	**M**	**G**	**P**	**M**	**G**	**P**	F_{avg}
Mostafa *et al.*[46]	—	—	—	—	—	—	—	—	—	0.99
Talukder *et al.*[33]	—	—	—	—	—	—	—	—	—	0.99
Haq *et al.*[37]	—	—	—	—	—	—	—	—	—	0.99
Raza *et al.*[48]	—	—	—	—	—	—	—	—	—	0.99
Amou *et al.*[31]	97.0	99.0	99.0	98.0	99.0	99.0	—	—	—	0.98
Ayan *et al.*[49]	—	—	—	—	—	—	—	—	—	0.98
Deppak and Ameer[42]	—	—	—	—	—	—	—	—	—	0.95
Hossain *et al.*[22]	94.0	98.0	97.0	94.0	96.0	100	—	—	—	0.97
Gu *et al.*[43]	—	—	—	—	—	—	—	—	—	0.94
Daz-Pernas *et al.*[19]	—	—	—	96.1	90.7	95.4	—	—	—	—
Badža *et al.*[23]	—	—	—	—	—	—	—	—	—	0.96
Mesut *et al.*[26]	92.4	97.3	98.9	94.8	98.4	95.3	97.3	97.9	99.5	0.96
Belaid and Loudini[40]	97.4	93.3	98.9	96.0	97.5	96	98.7	96.5	99.5	—
Deepak and Ameer[44]	94.7	99.92	98.0	96.0	97.9	98.9	98.4	99.4	99.1	0.97
Sultan *et al.*[20]	95.8	97.2	95.2	95.5	94.4	93.4	98.7	95.1	0.97	—
Das *et al.*[17]	94.0	88.0	98.0	95.0	85.0	99.0	—	—	—	0.93
Pashaei *et al.*[45]	94.5	91.0	98.3	76.8	97.5	100	—	—	—	0.93
Ismael *et al.*[16]	—	—	—	86.9	95.1	91.2	96.0	96.3	95.7	—
Cheng *et al.*[10]	—	—	—	86.0	96.4	87.3	95.5	96.3	95.3	—
Proposed Xception	**99.37**	**98.53**	**100**	**97.52**	**99.63**	**100**	**99.78**	**98.84**	**100**	**0.99**
Proposed DenseNet201	**99.35**	**97.81**	**99.46**	**95.65**	**99.63**	**100**	**99.78**	**98.26**	**99.77**	**0.98**
Proposed EfficientNetV2L	**99.35**	**98.17**	**98.12**	**95.65**	**99.63**	**100**	**99.78**	**98.55**	**99.53**	**0.98**

with respect to all evaluated metrics compared to state of the art methods in previous studies. In these aspects as well, our proposed models demonstrate their superiority over other previous studies on the same dataset.

Regarding Misclassifications

Upon conducting comprehensive performance evaluations and detailed analysis, it becomes an evident that our proposed models exhibit varying degrees of accuracy. Specifically, architecture-2 demonstrates a notably higher accuracy compared to both architecture-1 and architecture-3. This observation suggests that instances where misclassifications occurred in architecture-1 and architecture-3 were successfully rectified by the improved capabilities of the architecture-2. The results show the efficacy of Model-2 in correctly classifying cases that may have been challenging for Model-1 and Model-3. This result implies the superiority of Model-2 in handling specific classification challenges within the scope of our research.

DISCUSSION

In this paper, we conduct three extensive experiments to optimize the classification of brain tumors in T1-CE Figshare images using different TL architectures and techniques. Each experiment employs distinct strategies and architectures to achieve the highest possible accuracy and minimize the usage of the computational resources. We ensure the robustness of our results by using a five-fold CV approach. Here, we discuss the key findings and comparisons among these models.

In model-1, we investigate different TL architectures for brain tumor classification. In our proposed architecture, we run various TL architectures, and the InceptionV3 architecture shows the best accuracy level. The model achieves an accuracy level of 98.36%, demonstrating the potential of TL in medical image analysis. Adam optimizer and sparse categorical entropy as loss functions contribute to this success. This experiment establishes a strong baseline for our research.

In model-2, we adapt several pre-trained architecture through fine tuning the TL architectures and hyperparameters optimization. The proposed models show that the Xception model demonstrates its efficiency in extracting the most discriminative features. The Xception model achieves the highest accuracy among other models in the same experimental group. The results show that the accuracy level is 99.19%. Several factors contribute to this achievement, including architectural adjustments, BatchNormalization, and strategic changes to the learning rate. EarlyStopping and ReduceLROnPlateau techniques show significant role in the convergence process.

In model-3, we optimize the learning rate of SGD using PSO to enhance the convergence process. In model-3, we introduce an innovative approach by integrating PSO to dynamically adjust the learning rate of the SGD optimizer. We run the PSO for only 10 iterations because it is time consuming process. However, the recorded accuracy of this experiment is 98.53% using the Xception architecture, showcasing the potential of optimization algorithm in refining the learning rate. In addition, we include the use of momentum along with adjusting the learning rate using PSO optimization.

Comparative Analysis of our Experiments

The results show that the fine-tuned Xception model in Experiment-2 is the best proposed model. The model achieves the highest level of accuracy 99.19%. This outcome underscores the importance of architectural optimization and fine-tuning for achieving high level of accuracy while optimizing the usage of computational resources. In Experiment-3, the implementation of PSO to optimize the learning rate of SGD also achieves comparable accuracy, showcasing the potential of using optimization technique to adapt learning rate. Although it achieves slightly lower accuracy than the results in Experiment-2, it highlights the significance of exploring novel methods for adapting learning rate.

Although few studies achieved a slightly higher accuracy, it is crucial to note that these models often consist of more layers or involve hybrid architectures. These models tend to incur significantly higher computational costs, making them less practical for real-world applications, especially in the medical field. In contrast, our proposed architectures in Experiment-2, the fine-tuned Xception model achieves remarkable accuracy without excessive model complexity. The accuracy of 99.19% surpasses the performance of other experiments while remaining computationally efficient. This results show the practicality and effectiveness of our methodology for real-world medical image analysis applications. Our research presents a thorough exploration of TL techniques for brain tumor classification. The results of Experiment-2 emphasize the importance of fine-tuning and architectural adjustments in achieving outstanding accuracy while maintaining computational efficiency. The proposed model shows an optimal balance between accuracy and computational cost, a critical factor for deploying models in real-world medical settings.

In our study, we evaluate various critical factors, not only accuracy but also the pivotal aspects of computational efficiency and model complexity. In our proposed models, we avoid using prepossessing techniques and segmentation that consume more computational resources. In addition, we did not increase the number of images using different augmentation techniques to keep the computations to minimum level. We conduct a thorough analysis using essential metrics such as precision, recall, specificity, and the $F1$ score. We conduct an extensive comparative analysis to show the efficacy of our model in the brain tumor classification. In addition, our model consistently asserts its dominance by delivering superior performance.

CONCLUSION

This paper presents three fully automatic DL architectures designed for brain tumor detection, emphasizing minimal computational resource usage. By employing TL, these architectures effectively capture and utilize the most relevant features for classifying brain tumors. Our extensive evaluation using the Figshare dataset demonstrates the reliability and robustness of the proposed models, with the Xception model achieving an exceptional accuracy of 99.18%, surpassing all other models reported in the literature. The study highlights the significant impact of architectural modifications, fine-tuning, and learning rate optimization techniques on enhancing model performance. In addition, the five-fold CV strategy confirms the models' robustness and generalization capability.

Although the results obtained using PSO and SGD are competitive with those of the Adam optimizer, further improvements could be achieved by increasing the number of PSO iterations. Overall, the proposed computer-aided diagnosis system offers substantial benefits to brain surgeons, facilitating more accurate and timely treatment decisions. This work underscores the effectiveness of leveraging advanced DL techniques to address critical challenges in brain tumor detection and diagnosis, contributing valuable insights and practical solutions to the field.

ORCID

Wael Korani
https://orcid.org/0009-0008-7663-1200
Shyam Sundar Domakonda
https://orcid.org/0009-0008-6807-0159

REFERENCES

1. Srinivas C, Nandini Prasad KS, Zakariah M, Alothaibi YA, Shaukat K, Partibane B, Awal H, Deep transfer learning approaches in performance analysis of brain tumor classification using MRI images, *J Healthc Eng* **2022**, 2022.
2. Sinha S, Saraswat A, Bansal S, Sharan S, Brain tumour segmentation techniques from MR images using machine learning: An analysis technology (AIST), *2022 4th Int Conf Artificial Intelligence and Speech Technology (AIST)*, pp. 1–6, 2022.
3. American Cancer Society, Key statistics for brain and spinal cord tumors, 2021.
4. Bhanothu Y, Kamalakannan A, Rajamanickam G, Detection and classification of brain tumor in MRI images using deep convolutional network, *2020 6th Int Conf Advanced Computing and Communication Systems (ICACCS)*, pp. 248–252, 2020.
5. Langfitt TW, Obrist WD, Alavi A, Grossman RI, Zimmerman R, Jaggi J, Uzzell B, Reivich M, Patton DR, Computerized tomography, magnetic resonance imaging, and positron emission tomography in the study of brain trauma: Preliminary observations, *J Neurosurg* **64**(5):760, 1986.
6. Yanase J, Triantaphyllou E, A systematic survey of computer-aided diagnosis in medicine: Past and present developments, *Expert Syst Appl* **138**:112821, 2019.
7. Arabahmadi M, Farahbakhsh R, Rezazadeh J, Deep learning for smart Healthcare — A survey on brain tumor detection from medical imaging, *Sensors* **22**(5):1960, 2022.
8. Cheng J, Brain magnetic resonance imaging tumor dataset, Figshare MRI dataset version 5, 2017.
9. Vishlavath S, Deep G, Classification of brain tumor using PCA-RF in MR neurological images, *2019 11th Int Conf Communication Systems & Networks (COMSNETS)*, pp. 440–443, 2019.
10. Cheng J, Huang W, Cao S, Yang R, Yang W, Yun Z, Wang Z, Feng Q, Enhanced performance of brain tumor classification via tumor region augmentation and partition, *PLoS One* **10**(10):e0140381, 2015.
11. Qureshi SA, Hussain L, Ibrar U, Alabdulkreem E, Nour MK, Alqahtani MS, Nafie FM, Mohamed A, Mohammed GP, Duong TQ, Radiogenomic classification for MGMT promoter methylation status using multi-omics fused feature space for least invasive diagnosis through mpMRI scans, *Sci Rep* **13**(1):3291, 2023.
12. Liu Y, Bao Y, Review on automated condition assessment of pipelines with machine learning, *Adv Eng Inform* **53**:101687, 2022.
13. Liu Y, Bao Y, Intelligent monitoring of spatially-distributed cracks using distributed fiber optic sensors assisted by deep learning, *Measurement* **220**:113418, 2023.
14. Luo Z, Xu H, Chen F, Audio sentiment analysis by heterogeneous signal features learned from utterance-based parallel neural network, *Proc 2nd Workshop on Affective Content Analysis (AffCon 2019) co-located with Thirty-Third AAAI Conf Artificial Intelligence (AAAI 2019)*, pp. 80–87, 2019.
15. Qureshi SA, Raza SE, Hussain L, Malibari AA, Nour MK, Rehman A, Al-Wesabi FN, Hilal AM, Intelligent ultra-light deep learning model for multi-class brain tumor detection, *Appl Sci* **12**(8):3715, 2022.
16. Ismael MR, Abdel-Qader I, Brain tumor classification via statistical features and back-propagation neural network, *2018 IEEE Int Conf Electro/Information Technol (EIT)*, pp. 0252–0257, 2018.
17. Das S, Riaz Rahman Aranya OFM, Labiba NN, Brain tumor classification using convolutional neural network, *2019 1st Int Conf Advances in Science, Engineering and Robotics Technology (ICASERT)*, pp. 1–5, 2019.
18. Abiwinanda N, Hanif M, Hesaputra ST, Handayani A, Mengko TR, Brain tumor classification using convolutional neural network, in *World Congress on Medical Physics and Biomedical Engineering 2018*, Vol. 1 (2022), p. 3715.
19. Daz-Pernas FJ, Martnez-Zarzuela M, Antón-Rodríguez M, González-Ortega D, A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network, *Healthcare* **9**(1):153, 2021.
20. Sultan HH, Salem NM, Al-Atabany W, Multi-classification of brain tumor images using deep neural network, *IEEE Access* **7**:69215, 2019.
21. Alqudah AM, Alquraan H, Qasmieh IA, Alqudah A, Al-Sharu W, Brain tumor classification using deep learning technique — A comparison between cropped, uncropped, and segmented lesion images with different sizes, 2020, arXiv:2001.08844.
22. Hossain MF, Islam MA, Hussain SN, Das D, Amin R, Alam MS, Brain tumor classification from MRI images using convolutional neural network, *2021 IEEE Int Conf Artificial Intelligence in Engineering and Technology (IICAIET)*, pp. 1–6, 2021.
23. Badža MM, Barjaktarović C, Classification of brain tumors from MRI images using a convolutional neural network, *Appl Sci* **10**(6):1999, 2020.
24. Paul JS, Plassard AJ, Landman BA, Fabbri D, Deep learning for brain tumor classification, *Med Imaging Biomed Appl Mol Struct Funct Imaging* **10137**:253, 2017.
25. Huang Z, Du X, Chen L, Li Y, Liu M, Chou Y, Jin L, Convolutional neural network based on complex networks for brain tumor image classification with a modified activation function, *IEEE Access* **8**:89281, 2020.
26. Mesut T, Burhan E, Zafer C, Tumor type detection in brain MR images of the deep model developed using hypercolumn technique, attention modules, and residual blocks, *Med Biol Eng Comput* **59**(1):57, 2021.
27. Afshar P, Mohammadi A, Plataniotis KN, Brain tumor type classification via capsule networks, *2018 25th IEEE Int Conf Image Processing (ICIP)*, pp. 3129–3133, 2018.
28. Afshar P, Plataniotis KN, Mohammadi A, Capsule networks for brain tumor classification based on MRI

images and coarse tumor boundaries, *ICASSP 2019–2019 IEEE Int Conf Acoustics Speech and Signal Processing (ICASSP)*, pp. 1368–1372, 2019.

29. Afshar P, Mohammadi A, Plataniotis KN, BayesCap: A bayesian approach to brain tumor classification using capsule networks, *IEEE Signal Process Lett* **27**:2024, 2020.
30. Du X, Chen L, Liu Z, Li S, Liu M, Yang J, Jin L, Brain tumor image classification by randomly wired neural networks with a modified method, *2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS)*, pp. 1141–1146, 2020.
31. Amou MA, Xia K, Kamhi S, Mouhafid M, A novel MRI diagnosis method for brain tumor classification based on CNN and bayesian optimization, *Healthcare* **10**(3):494, 2022.
32. Benjamin M, Erfan Z, Soroush A, QuickTumorNet: Fast automatic multi-class segmentation of brain tumors, *2021 10th Int IEEE/EMBS Conf Neural Engineering (NER)*, pp. 81–85, 2021.
33. Talukder MA, Islam MM, Uddin MA, Akhter A, Pramanik MAJ, Aryal S, Almoyad MAA, Hasan KF, Moni MA, An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning, *Expert Syst Appl* **230**:120534, 2023.
34. Malla PP, Sahu S, Alutaibi AI, Classification of tumor in brain MR images using deep convolutional neural network and global average pooling, *Processes* **11**(3):679, 2023.
35. Tummala S, Kadry S, Bukhari SAC, Rauf HT, Classification of brain tumor from magnetic resonance imaging using vision transformers ensembling, *Curr Oncol* **29**(10):7498, 2022.
36. Swati ZN, Zhao Q, Kabir M, Ali F, Ali Z, Ahmed S, Lu J, Brain tumor classification for MR images using transfer learning and fine-tuning, *Comput Med Imaging Graph* **75**:34, 2019.
37. Haq AU, Li JP, Khan S, Alshara MA, Alotaibi RM, Mawuli C, DACBT: Deep learning approach for classification of brain tumors using MRI data in IoT healthcare environment, *Sci Rep* **12**(1):15331, 2022.
38. Arumaiththurai T, Mayurathan B, The effect of deep learning and machine learning approaches for brain tumor recognition, *2021 10th Int Conf Information and Automation for Sustainability (ICIAfS)*, pp. 185–190, 2021.
39. Masood M, Nazir T, Nawaz M, Mehmood A, Rashid J, Kwon HY, Mahmood T, Hussain A, A novel deep learning method for recognition and classification of brain tumors from MRI images, *Diagnostics* **11**(5):744, 2021.
40. Belaid ON, Loudini M, Classification of brain tumor by combination of pre-trained VGG16 CNN, *J Inf Technol Manag* **12**(2):13, 2020.
41. Rehman A, Naz S, Razzak MI, Akram F, Imran M, A deep learning-based framework for automatic brain tumors classification using transfer learning, *Circuits Syst Signal Process* **39**:757, 2020.
42. Deepak S, Ameer PM, Brain tumor categorization from imbalanced MRI dataset using weighted loss and deep feature fusion, *Neurocomputing* **520**:94, 2023.
43. Gu X, Shen Z, Xue J, Fan Y, Brain tumor MR image classification using convolutional dictionary learning with local constraint, *Front Neurosci* **15**:679847, 2021.
44. Deepak S, Ameer PM, Brain tumor classification using deep CNN features via transfer learning, *Comput Biol Med* **111**:103345, 2019.
45. Pashaei A, Sajedi H, Jazayeri N, Brain tumor classification via convolutional neural network and extreme learning machines, *2018 8th Int Conf Computer and Knowledge Engineering (ICCKE)*, pp. 314–319, 2018.
46. Mostafa AM, El-Meligy MA, Alkhayyal MA, Alnuaim A, Sharaf M, A framework for brain tumor detection based on segmentation and features fusion using MRI images, *Brain Res* **1806**:148300, 2023.
47. Kazemi A, Shiri ME, Sheikhahmadi A, Classifying tumor brain images using parallel deep learning algorithms, *Comput Biol Med* **148**:105775, 2022.
48. Raza A, Ayub H, Khan JA, Ahmad I, Salama AS, Daradkeh YI, Javeed D, Ur Rehman A, Hamam H, A hybrid deep learning-based approach for brain tumor classification, *Electronics* **11**(7):1146, 2022.
49. Ayan M, Shrivastava Vimal K, A novel parametric flatten-p mish activation function based deep CNN model for brain tumor classification, *Comput Biol Med* **150**:106183, 2022.