



A transfer learning-based brain tumor classification using magnetic resonance images

Ishwari Singh Rajput¹ · Aditya Gupta² · Vibha Jain²  · Sonam Tyagi¹

Received: 2 September 2022 / Revised: 31 May 2023 / Accepted: 3 July 2023 /
Published online: 2 August 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

The brain is one of the most important and complex organs responsible for controlling the functions of the human body. Brain tumors are among the most lethal malignancies in the world. Analyzing MRI images of the patient's brain is one of the usual approaches for diagnosing brain tumors. However, misdiagnosis of brain tumor forms hinders effective responses to medical interventions and reduces patients' chances of survival. With recent technological advancements, deep learning-based medical image analysis has become increasingly popular in the research community. Transfer learning reuses a pre-trained model, trained on huge volume datasets to solve a new problem in distinguished application domains. This research article attempts to diagnose the three most prevalent forms of brain tumors using pre-trained CNN models such as VGG19, Inception-v3, and ResNet50 using transfer learning. The features extracted using pre-trained models are supplied into fully connected layers that fine-tune the model for classifying multi-class tumors. The performance of the presented approach is evaluated on a benchmark MRI results reveal that the proposed approach leads to superior performance in comparison to conventional techniques with an average accuracy of 90%.

Keywords Transfer learning · Brain tumor · Deep learning · Feature extraction · Classification

1 Introduction

Tumors are large aggregates of cells that grow irregularly in any organ of the body [22]. Amongst varying forms of tumors, brain tumors are regarded as one of the world's worst concerns, capable of afflicting anybody regardless of age, gender, or ethnicity. These tumors damage the nervous system and brain tissues, resulting in a shortened life expectancy for patients. Besides, it degrades the patients' quality of life and severely curtails their regular activities. Tumors reported in the literature are divided into two stages, benign

✉ Vibha Jain
vibha.jain@jaipur.manipal.edu

¹ Graphic Era Hill University, Haldwani, Uttarakhand, India

² Manipal University Jaipur, Jaipur, Rajasthan, India

and malignant [1]. Benign tumors, the primary stage of the tumor is usually non-cancerous, and their recurrence rate is relatively low. Conversely, malignant tumors grow faster and are extremely dangerous compared to benign tumors. The mortality rate of patients with brain tumor is on the rise [15]. According to a recent report from the National Brain Tumor Society, approximately 700,000 people in the United States have brain tumors [26]. In 2020, more than 87 thousand people were diagnosed with brain tumors. Moreover, the survival rate of patients with brain tumors is as low as 36%. Furthermore, work carried out by [12] states that the mortality rate has increased up to 3 times in the last 3 decades.

Tumors are often diagnosed manually by radiologists using a combination of imaging techniques, such as computed tomography (CT) and magnetic resonance imaging (MRI) scans. However, such tumor diagnostic procedure is exceedingly time intensive, error-prone, and relatively dependent on the radiologist's abilities and expertise [19]. Moreover, since the tumor progresses without any signs and is realized only at an advanced stage, hence, early diagnosis of this disease is cumbersome. Therefore, it is necessary to develop automated healthcare support systems which can help radiologists identify brain tumors more accurately [2, 3].

In recent years, the advancements in artificial intelligence in association with machine learning have revolutionized the healthcare domain by providing real-time healthcare solutions [11]. Nevertheless, the complexity of operations such as pre-processing, segmentation, and feature extraction in traditional machine learning algorithms reduces the model's efficiency and accuracy [10]. To address the limitations of existing machine learning methods, the concept of deep learning is introduced to extract useful features from input images and use them effectively for diagnosis and classification. Deep learning can assist doctors in identifying and classifying tumors more accurately. Convolutional neural networks (CNNs) are the most commonly used deep learning techniques and have a wide range of applications in distinct domains [27]. Accurate CNN-based classification systems typically require substantial amounts of visual data for training. To boost the performance of individual CNN architectures by pooling knowledge, the concept of transfer learning can be used to achieve higher classification accuracy. Transfer learning attempts to re-use a CNN model trained on a generic image dataset such as ImageNet [20], and apply it to domain-specific and smaller datasets, and finally, network parameters are fine-tuned for better performance. The benefit of using transfer learning is that it improves classification accuracy, while also speeding up the training process.

In this work, we attempt to use pre-trained CNN models such as VGG19, Inception-V3, and ResNet-50 to transfer their learned parameters for multi-class classification of brain tumor.

The primary objectives of the proposed framework are listed below.

1. To design a deep learning-based framework for the detection and classification of brain tumor.
2. To demonstrate the concept of transfer learning using four pre-trained CNN architectures.
3. To compare the performance of the proposed framework with the existing classification models.

The rest of the paper is organized as follows. Section 2 presents the work done in the current domain. In Section 3, a brief description of the preliminaries is given. The proposed methodology utilized to solve the current problem is provided in Section 4. Section 5

describes the simulation setup and performance assessment of the proposed strategy. Finally, Section 6 concludes the paper with some future avenues.

2 Literature review

Numerous strategies for identifying brain tumors using MRI scans have been proposed by a variety of researchers throughout the world. These methods range from traditional machine learning algorithms to deep learning models. Consequently, this section discusses the various findings of the researchers in the diagnosis of brain tumor based on MRI images.

Noreen et al. [23] presented a tumor diagnostic framework based on pre-trained Inception-V3 and DensNet201 deep learning techniques. To diagnose brain tumors, the authors used 11 inception modules in Inception-V3 and 4 dense blocks in DensNet201, with varying numbers of convolutional layers for feature extraction, followed by a softmax classifier. This approach utilizes multi-level feature extraction and concatenation and the model was tested using 3064 T1-weighted contrast MR images. The proposed strategy achieved a greater level of accuracy in comparison to past works.

Musallam et al. [21] (2022) presented a framework that included three phases of pre-processing to eliminate confusing objects, denoising the MRI images, and histogram equalization to enhance MRI image quality. A novel Deep Convolutional Neural Network model was utilized to classify the brain tumor. The approach was tested on both the Sartaj brain MRI image dataset, and the Navoneel brain tumor dataset, which comprised both T1- and T2-weighted MRI images.

Assam et al. [4] (2021) developed a method that used a median filter to pre-process MRI images, followed by a Discrete wavelet transform to extract features and color moments to minimize features. These features were fed into feed-forward artificial neural networks, random subspace with random forests, and random subspace classifiers with Bayesian networks for MRI brain image classification. The method was validated using a real-world dataset of 70 T2-weighted pictures from Harvard Medical School.

Sekhar et al. [25] (2021) proposed a deep learning approach that integrates pre-trained GoogleNet for feature extraction from brain MRI images, as well as softmax, SVM, and K-NN for classification. The authors relied on data from the CE-MRI Figshare repository and the Harvard Medical Archives. The proposed model exhibited the highest classification accuracy for 3-class tumor classification and improved classification results for 4-class tumor classification. Compared with previous works, the proposed model outperforms all other models.

Irmak [13] (2021) introduced a framework for multi-class classification of brain tumor MRI images using three fully automated convolutional neural network models followed by a Softmax classifier. The first CNN model (Classification-1) identifies brain tumors, the second CNN model (Classification-2) classifies brain tumors into categories, and the third CNN model (Classification-3) classifies glioma brain tumors into Class II, Class III, and Class IV. To optimize hyperparameters and build the most successful CNN models, a grid search optimizer was used.

By incorporating three neurons into a fully connected layer, Ismael et al. [14] (2020) developed a method for classifying brain tumors using a ResNet 50-based augmented deep learning algorithm. The authors employed data augmentation techniques such as horizontal and vertical flipping, rotation, shifting, scaling, ZCA whitening, clipping, and brightness

modification to expand the dataset. The model was validated using 3064 T1-weighted contrast-enhanced MRI images from a publicly accessible brain tumor dataset.

Ghassemi et al. [9] (2020) devised a framework in which a deep convolutional neural network was pre-trained as a discriminator in a GAN to differentiate between fraudulent and genuine MR images generated by the generative model. Consequently, the discriminator identified the structure of MR images and extracted trustworthy MRI scan features. The pre-trained CNN was therefore fine-tuned as a classifier for brain tumor classification by training it on the original dataset. The last fully connected layer of the GAN discriminator was replaced with a SoftMax layer of three neurons for classification purposes. The author employed data augmentation techniques such as rotation and mirroring to extend the dataset. The approach was utilized on 3064 T1-CE MR images and whole brain volume MR images. Another transfer learning based image classification approach is proposed in [18].

Khairandish et al. [16] (2021) presented a technique for identifying and classifying MRI brain images that combine CNN and support vector machine. The proposed model classifies brain images as benign or cancerous using a supervised hybrid CNN and SVM techniques. The input images were normalized during primary preprocessing, and significant features were extracted from the preprocessed image using the Maximally stable extremal regions technique, followed by the threshold-based segmentation methodology. The labeled segmented features were fed into hybrid CNN and SVM algorithms to categorize brain MRI images. The researcher utilized the BRATS 2015 dataset for data analysis.

Khan et al. [17] (2022) developed a brain tumor classifying model based on hierarchical deep learning. The suggested method was divided into three stages: data collection, preparation, and application. The MR images were collected initially with IoMT devices, then processed by the data acquisition layer, preprocessing, and subsequently CNN. Table 1 presents the comparative analysis of the previous works.

3 Preliminaries

3.1 CNN

Convolutional Neural Network (CNN) is a deep learning architecture that stands out from other neural networks due to its superior performance on image and speech input. Due to the reduced number of parameters involved and the reusability of weights, CNN architectures can better fit image datasets. They are commonly used in computer vision/image recognition applications such as image analysis in healthcare, object recognition in self-driving cars, and a range of other fields. CNN architectures have three types of layers: convolutional layers, pooling layers, and fully connected layers.

- *Convolutional layer:* The initial layer extracts basic features from the input image. The convolution operation is applied between the image input matrix and the f-dimensional filter. The filter slides over the input matrix and computes the dot product of the filter elements and the input matrix area. The resulting computation is referred to as a feature map, and it includes information about the vertical and horizontal edges. These feature maps are fed into the subsequent layers, which extract more specific features from an input image.
- *Pooling layer:* The pooling layer down samples the feature maps acquired by the convolutional layer in order to minimize the number of parameters in the input matrix.

Table 1 Comparative analysis of recent related work

Reference	Objective	Dataset	Classes of tumors	Total Images	Model
[23]	To diagnose brain tumor using a deep learning model based on concatenation approach	3064 T1-weighted contrast MR images	Glioma, Meningioma, and Pituitary	3064	Inception-v3 and DensNet201
[21]	To detect brain tumor using deep Convolutional Neural Network	Sartaj brain MRI images dataset and the Navoneel brain tumor dataset which contains two types of MRI images	Glioma, Meningioma, and Pituitary	3394	DCNN
[4]	To classify MRI brain images into normal and abnormal using individual and hybrid classifiers	Real-world datasets (taken from Harvard Medical School) (T2 weighted images)	-	Normal=25, Abnormal=45	Classifiers: i.LFF-ANN ii.RSwthRF iii.RSwthBN
[25]	To detect brain tumor using IoMT enabled CAD system	Figshare dataset (CE-MRI dataset)	Glioma, Meningioma, and Pituitary	3064	i. GoogLeNet with SVM ii. GoogLeNet with KNN
[13]	Multi-classification of brain tumors using convolutional neural network	RIDER, REMBRANDT, TCGA-LGG	Classification I: Brain tumor and non-tumor; Classification II: lioma, meningioma, pituitary, normal and metastatic; Classification III: glioma grade	RIDER = 70,220, REMBRANDT = 110,020, TCGA-LGG = 241,183	CNN
[14]	Classification of brain cancer using an Enhanced Deep Learning Approach using Residual Networks	Publicly available brain tumor dataset of 3064 T1-weighted contrast-enhanced MRI images	Meningioma, Glioma, and Pituitary	3064	ResNet-50
[9]	To classify brain tumor using Deep neural network with generative adversarial networks	i. 3064 T1-CE MR Images, ii. Whole brain volume MR images (contains 373 sets of longitudinal images from 150 subjects).	Meningioma, Glioma, and Pituitary	3064, 165	Deep CNN

Table 1 (continued)

Reference	Objective	Dataset	Classes of tumors	Total Images	Model
[16]	To detect and classify MRI Brain Images using a hybrid CNN-SVM Threshold Segmentation Approach	BRATS 2015	-	220 HGG and 54 LGG MRIs	CNN-SVM
[17]	To identify Brain Tumor using Intelligent Deep Learning	Publicly available dataset	Normal, Glioma, Meningioma, and Pituitary	3264	Hierarchical Deep Learning-Based Brain Tumor Classifier HDL2B-TUMOR-CLASSIFIER

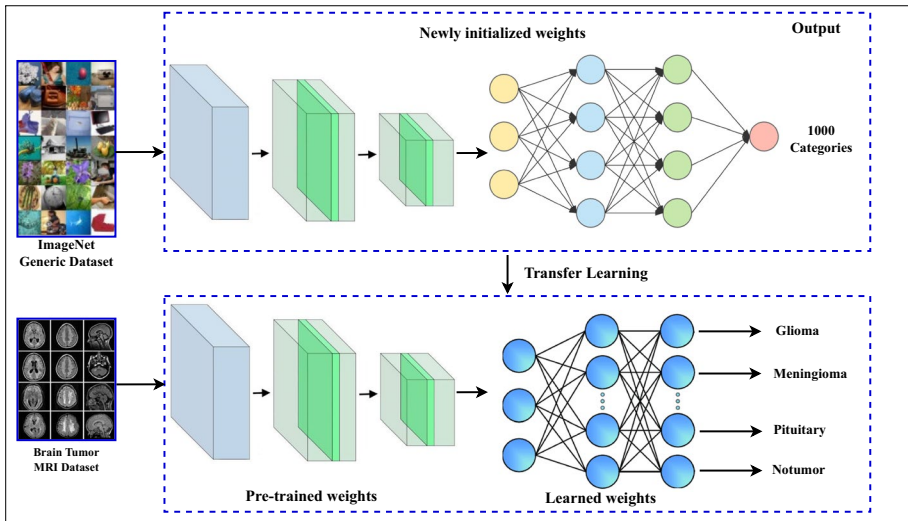


Fig. 1 Transfer learning

Table 2 Summary of pre-trained CNN architectures

CNN	Image input size	Filter size in various layers	No. of parameters	Depth
VGG19	224 × 224	(3 × 3)	138M	19
Inception-v3	299 × 299	(1 × 1), (3 × 3), (5 × 5), (1 × 7), (7 × 1)	23.9M	48
ResNet50	224 × 224	(1 × 1), (3 × 3), (7 × 7)	25.6M	50

Pooling reduces complexity and computational cost while increasing efficiency and reducing the risk of overfitting. Depending on the method used, max pooling and average pooling are two widely used categories.

- *Fully connected layer:* This layer consists of weights and biases and precedes the output layer. The preceding layer’s feature maps are flattened and fed into this layer. It then performs classification using the extracted features from the initial layers.

3.2 Transfer learning

In essence, a substantial amount of data is required to train a CNN model, but in certain scenarios, gathering a large dataset of relevant domains is quite challenging. Hence, the paradigm of transfer learning is introduced. Transfer learning is a well-known machine learning technique that attempts to learn the basic features needed to solve a problem and then apply them to problems in other domains as depicted in Fig. 1.

Transfer learning has the advantages of faster training processes, prevention of overfitting, training with fewer data, and increased performance. GoogleNet, VGG19, Inception-V3, and ResNet-50 were the pre-trained CNN models employed in our experimental process. Table 2 provides the description of pre-trained CNN models. The aforementioned CNN architectures are trained on the ImageNet dataset comprising 1000 images classes

and transfer learning is adopted. The subsequent subsections provide description of different pre-trained CNN models.

3.2.1 VGG19

VGG19 is a basic and efficient network for a variety of object recognition models. It consists of 16 convolutional layers followed by pooling layers and 3 fully connected layers with a total of 138 M parameters. Because of its simple architecture, it outperforms other models.

3.2.2 Inception-V3

Inception-V3 is a frequently used CNN architecture used for classification purposes. It was designed by modifying the Inception module and is composed of several blocks of convolutional, pooling, and fully connected layers. It also uses dropout layer to mitigate the problem of overfitting. It comprises 42 layers with a total of 23.9M parameters.

3.2.3 ResNet50

A residual network often abbreviated as ResNet is a deep architecture that provides good accuracy in image classification. It also took first place in the ILSVRC challenge in 2015. ResNet 50 contains 49 convolutional layers and a fully connected layer at the end with a total of 25.6 learnable parameters. By leveraging skip connections, it alleviates the problem of vanishing gradient.

4 Proposed methodology

The proposed methodology used for brain tumor classification using MRI images is presented in Fig. 2. It comprises mainly three stages, namely data preprocessing, feature extraction, and finally classification. The description of each stage is explained in detail in subsequent sub-sections.

4.1 Data preprocessing

Preprocessing of data is an important stage before performing data analytics operations. Preprocessing is performed to enhance the quality of images. It includes the following steps:

- *Cropping and Resizing*: The aim of cropping are to extract the brain contour for the MRI images. Extra space out of the image is removed by cropping. Normalization is done to resize the image according to the permissible input size of the particular CNN model. For VGG19 and ResNet-50 the image size is normalized to 224×224 and for Inception-V3 images are resized to 299×299 .
- *Data Augmentation*: For higher performance in terms of accuracy CNN models requires a large number of the labeled image dataset. But in a few circumstances, it is not feasible, or only small sized dataset is available. The data augmentation method is used to increase the number of images by applying some geometric transformations

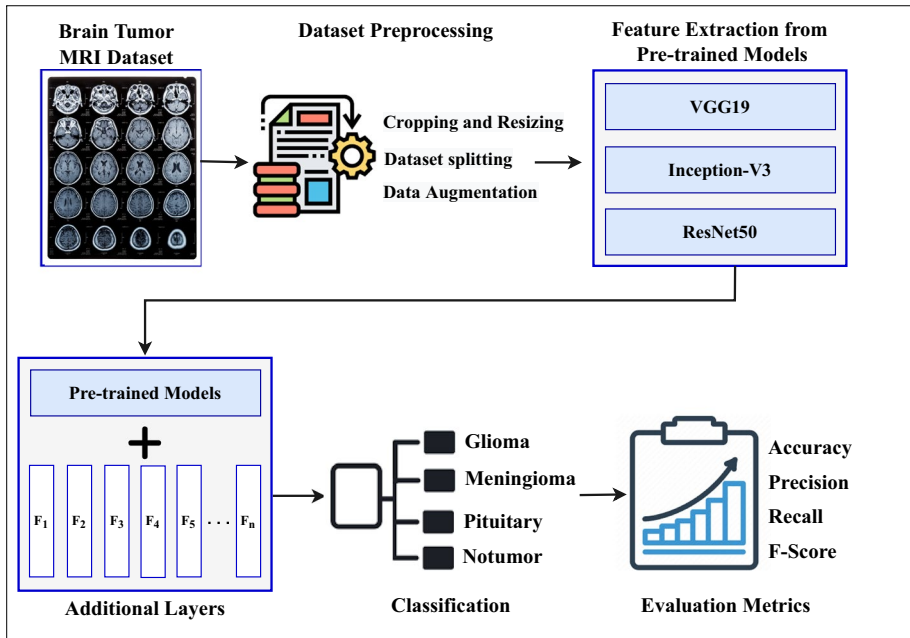


Fig. 2 Proposed model

such as image flipping, rotation by some angle, color transformation, change in brightness and scaling, etc.

- *Dataset splitting*: For evaluating the performance of the proposed model, the dataset is split into 70%-30%, 80%-20%, and 90%-10% train-test split. To validate the proposed model 20% of the test data is used as a validation set.

4.2 Feature extraction

The initial layers of CNN models i.e. group of convolution and pooling layers are responsible for extracting image features. The extracted features are then fed into the series of fully connected layers for the task of classification. The proposed framework uses three widely used CNN models: VGG19, ResNet-50, and Inception-V3 for feature extraction from the generic ImageNet dataset which is then used for the classification of brain tumor MRI image dataset. The concept of transfer learning is used to learn the features from MRI images without any training and classify the MRI images in Glioma, Meningioma, Pituitary, and No tumor classes.

4.3 Classification

CNN has shown substantial potential in the field of medical image classification in recent decades, where it has been quite effective in tackling image classification challenges. The CNN model performs better when learning global and local features from medical visuals. The number of nodes in the classification output layer in our proposed framework is equal to the number of classes in the dataset. CNN model automatically learns features during



Fig. 3 Brain Tumor images

Table 3 Dataset summary

Tumor type	Number of images	Training set (80%)	Testing set (20%)
Glioma	1621	1297	324
Meningioma	1645	1316	329
Pituitary	1757	1405	352
Notumor	2000	1600	400

the training stage from each image. Each output has a different probability for the input image; the model then selects the output with the highest probability as its prediction of the class. Finally, the output layer determines which class of brain tumor is present in the patient using the pre-trained model.

5 Experimental results

This section discusses the simulation setup and performance evaluation of the proposed work. The proposed model is implemented on a system equipped with Intel Core-i7 processor with 32 GB RAM. Simulation results are generated with 95% confidence interval. The Keras library available in latest version of Python is used for training and validation.

5.1 Dataset description

To assess the performance of the proposed framework, a brain tumor MRI dataset is required. The required dataset is obtained from the Kaggle [7] which consists of 7023 images with four classifications, namely glioma, meningioma, no tumor, and pituitary. The explored dataset is combination of three widely used datasets: figshare [8], SARTAJ [24], and Br35H [6]. Figure 3 presents the sample images from each of the labels. Table 3 provides the dataset description which contains 1621 brain MRI images with glioma, 1645 brain MRI images belonging to class meningioma, 1757 MRI images containing pituitary and the rest 2000 images have no tumor. Initially, size of each image is 512×512 which is resized into 224×224 for VGG19 and ResNet-50 as well as to 299×299 for Inception-V3. However, image quality can be improved using the super-resolution scheme proposed in [3, 5]. In order to evaluate the performance of the proposed framework, the dataset is split into 70%-30% , 80%-20%, and

90%-10% train-test splits. Moreover, to control over-fitting difficulties, an initial stopping condition based on performance validation is devised, i.e., to terminate the training process when the system exhibits no or little progress after a few iterations.

5.2 Evaluation measures

Several evaluation measures are used to validate the performance of classifier. Out of all the defined measures classifier accuracy is the prominently used quality index which measures the number of correctly classified samples to the total number of data samples. Table 4 illustrate the features of the confusion matrix with four partitions each depicting a unique features True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The following performance measures are calculated to evaluate the proposed framework.

5.2.1 Accuracy

It measures the ability of the classifier to predict sample data accurately. Accuracy is calculated as a ratio of corrected predictions to the total number of classification predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

5.2.2 Precision

Precision is calculated as a ratio of the correct prediction (TP) to the total number of infected patients.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

5.2.3 Recall or Sensitivity

Recall calculates classifier quality to make accurate positive predictions. It is calculated as the ratio between total correct positive predictions to the total positively infected sample.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

5.2.4 Specificity

Opposite to sensitivity, specificity measures the ability to negative predictions.

Table 4 Confusion matrix

Actual	Predicted	
	Yes	No
Yes	True positive	False negative
No	False positive	True negative

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4)$$

5.2.5 F1-score

F1-score quantifies the accuracy of the classifier in terms of recall and precision.

$$F1 - score = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (5)$$

5.3 Results and discussions

The proposed framework is trained on three different CNN architectures, namely VGG19, Inception-V3, and ResNet-50. The learned data is then transferred into ensemble feature extraction using transfer learning. The generated results from a single CNN are compared to a combined feature set upon several classifier quantifiers. Hyperparameters of CNN model are tuned iteratively to achieve the convergence point in training process and minimize the loss function. We have selected the Adam optimizer for learning rate adoption due to its adaptive nature. A mini-batch of size 32 is selected to save the results and loss function calculation. To avoid the condition of overfitting, number of epochs is set to 10. Table 5 enlist the parameters used for experimentation.

The confusion matrix of different pre-trained models and proposed transfer learning model is given from Figs. 4, 5, 6 and 7. To evaluate the classifier performance more efficiently, three different training and test splits are considered which divide sample data into 70%-30%, 80%-20%, and 90%-10% sets.

Figure 4 shows three different confusion matrices under different data splits. Each matrix contains four labels corresponding to each class in the dataset. Figure 4(a) shows that ResNet-50 correctly predict 2047 sample data to their corresponding class when the total number of test samples is 2110. Therefore, the system gives 97% accuracy with 70%-30% data split. Figure 4(b) is for data split 80%-20% which depicts that ResNet-50 accurately predicts 1380 sample data to correct class with the accuracy of 98%. At last, Fig. 4 gives the confusion matrix for data split 90%-10%. Out of the total sample, ResNet-50 correctly predict 704 MRI images 679 to their correct classes. From the experimental results it is evident that ResNet-50 gives the best results when 5618 images MRI images are used for training and 1405 brain MRI images are used for testing of classifier model.

Table 5 Hyper-parameter for fine-tuning

Parameter	Value
Type of pooling layer	Average
Learning rate	0.001
Mini-batch size	32
Optimizer algorithm	Adam
Loss function	Cross-Entropy
Maximum epoch	10

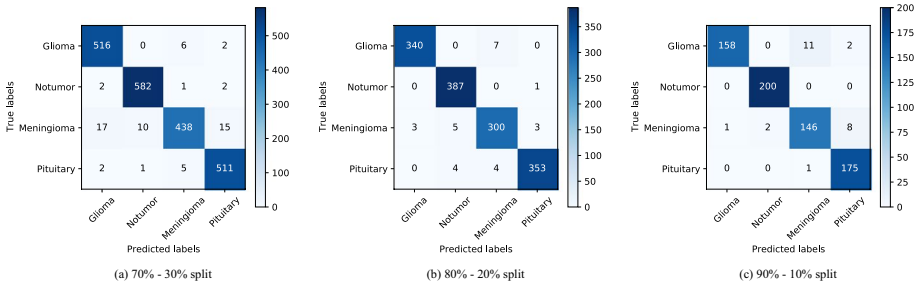


Fig. 4 Confusion Matrix for ResNet-50

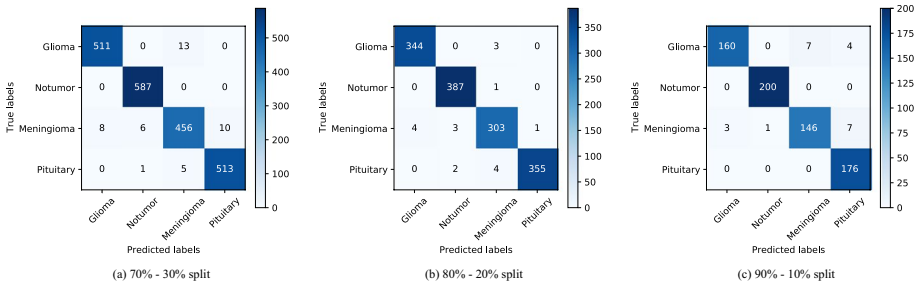


Fig. 5 Confusion Matrix for Inception V3

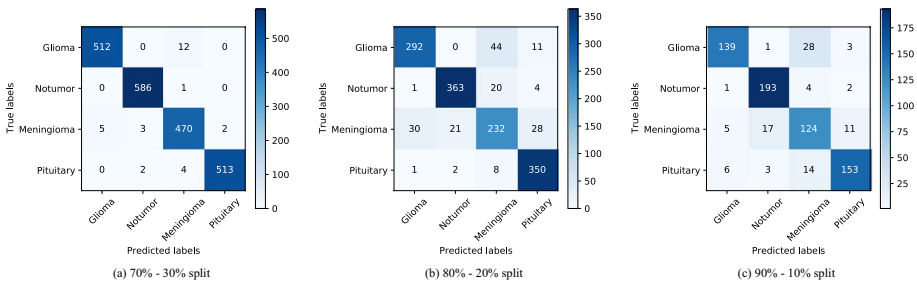


Fig. 6 Confusion Matrix for VGG19

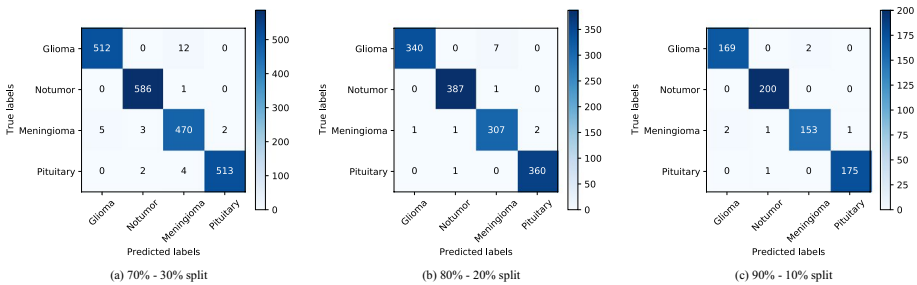


Fig. 7 Confusion Matrix for proposed model

Figure 5 shows the confusion matrix on a similar parameter for Inception-V3 network. Figure 5(a) shows that 2067 images are classified correctly from the total of 2110 images giving an accuracy of 97.9%. The second Fig. 5(b) gives the confusion matrix for the data split of 80%-20%. Results show that Inception V3 gives 99% by classifying 1389 MRI images accurately to the given class. Lastly, with the data split of 90%-10%, Inception V3 gives 97% accuracy.

Figure 6 shows the confusion matrix for VGG19 model. Contrary to the earlier pre-trained model, VGG19 has comparatively less accurate. VGG19 gives 81% with 70%-30%, 88% with 80%-20% and 87% with 90%-10% .

Figure 7 gives the confusion matrix of the proposed classifier with transfer learning. Figure 7(a) gives the classification results when training and test data are split in 70 and 30 ratio. Total of 2110 brain MRI images is used for testing purpose whereas the proposed classifier accurately classify 2081 images in their corresponding class giving an accuracy of 98.6%. Best results are obtained with a data split of 80% - 20%. The optimum accuracy obtained in the proposed classifier method is 99% with 1394 brain MRI images classified accurately.

Figure 8 show the accuracy and loss during training and validation process under different pre-trained models and the proposed model. The experiment is carried out by setting number of epochs equal to 10. In each epoch, accuracy, training, and validation error is calculated. The proposed model achieves highest classification accuracy and lowest training and validation loss compared with other architectures.

Table 6 gives a summary of experimentation results with pre-trained model and designed transfer learning classifier model. From the results, it is evident that the proposed architecture outperforms existing CNN architectures and achieves highest classification accuracy of 99% when trained using 90-10 train-test split. On the contrary, ResNet-50, Inception-V3, and VGG 19 attain average accuracy of 97%, 98%, and 85% respectively.

5.3.1 Comparative analysis with recent works

The overall performance of the proposed method is also compared with existing works in the current domain. Table 7 reports the experimental results of the proposed framework and other existing studies in terms of accuracy and other measures. From the Table 7, it is evident that the performance approach attains higher values of accuracy in comparison to previous studies.

6 Conclusion

Brain tumors are the prominent cause of death in developing and underdeveloped countries. Early identification and classification of a brain tumor can potentially save a patient's life by recommending an appropriate and timely cure. The proposed work presents an efficient decision support system by leveraging the concept of transfer learning. The framework involves extracting features using a generic ImageNet dataset from the three most widely available CNN models, namely VGG19, RestNet-50, and Inception-V3. The proposed model is fine-tuned using pre-trained weights on the brain tumor MRI dataset. Finally, the performance of our proposed model is compared with state-of-the-art CNN models, namely ResNet-50, Inception-V3, and VGG19 using various classification metrics.

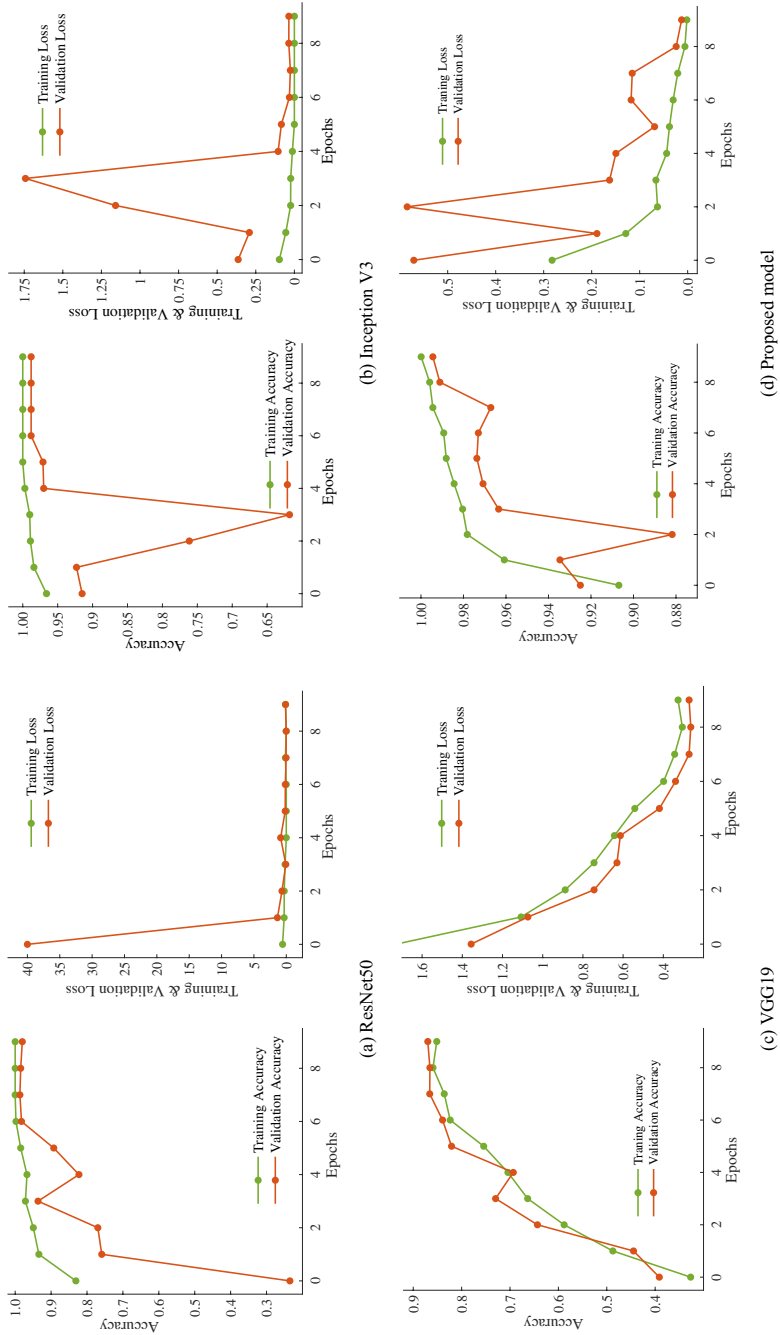


Fig. 8 Epochs vs. training and validation Accuracy/Loss

Table 6 Comparative analysis of different decision making algorithms

Classifier	Train-Test data split	Class Type	Precision	Recall	F1-Score	Accuracy	Average Accuracy
ResNet50	70%-30%	G	0.96	0.98	0.97	0.97	0.97
		N	0.98	0.99	0.99		
		M	0.97	0.91	0.94		
		P	0.96	0.98	0.97		
	80%-20%	G	0.99	0.98	0.99	0.98	
		N	0.98	1.0	0.99		
		M	0.96	0.96	0.96		
		P	0.99	0.98	0.98		
	90%-10%	G	0.99	0.92	0.96	0.96	
		N	0.99	1.0	1.0		
		M	0.92	0.93	0.93		
		P	0.95	0.99	0.97		
Inception V3	70%-30%	G	0.98	0.98	0.98	0.98	0.98
		N	0.99	1.0	0.99		
		M	0.96	0.95	0.96		
		P	0.98	0.99	0.98		
	80%-20%	G	0.99	0.99	0.99	0.99	
		N	0.99	1.0	0.99		
		M	0.97	0.97	0.97		
		P	1.0	0.98	0.99		
	90%-10%	G	0.98	0.94	0.96	0.97	
		N	1.0	1.0	1.0		
		M	0.95	0.93	0.94		
		P	0.94	1.0	0.97		
VGG19	70%-30%	G	0.96	0.64	0.77	0.81	0.85
		N	0.88	0.98	0.92		
		M	0.61	0.74	0.67		
		P	0.86	0.85	0.86		
	80%-20%	G	0.90	0.84	0.87	0.88	
		N	0.94	0.94	0.94		
		M	0.76	0.75	0.75		
		P	0.89	0.97	0.93		
	90%-10%	G	0.92	0.81	0.86	0.87	
		N	0.90	0.96	0.93		
		M	0.73	0.79	0.76		
		P	0.91	0.87	0.89		
Proposed Model	70%-30%	G	0.99	0.98	0.98	0.99	0.99
		N	0.99	1.0	0.99		
		M	0.97	0.98	0.97		
		P	1.0	0.99	0.99		
	80%-20%	G	1.0	0.98	0.99	0.99	
		N	0.99	1.0	1.0		
		M	0.97	0.99	0.98		
		M	0.97	0.99	0.98		

Table 6 (continued)

Classifier	Train-Test data split	Class Type	Precision	Recall	F1-Score	Accuracy	Average Accuracy
		P	0.99	1.0	1.0		
	90%-10%	G	0.99	0.99	0.99	0.99	
		N	0.99	1.0	1.0		
		M	0.99	0.97	0.98		
		P	0.99	0.99	0.99		

The proposed framework outperforms all the other deep learning architectures by attaining highest classification accuracy of 99%.

The proposed approach utilized a different set of advantages inherited from pre-trained transfer learning models such as reduced training time, improved performance, and high accuracy. Pre-trained models have acquired beneficial properties from vast datasets and can generalize effectively to new, similar applications. We have obtained better results than training from scratch with minimal data by fine-tuning these models with limited data. The proposed model also solves the problem of overfitting found in previous literature. Although in addition to the advantages listed above, the proposed model also has some presented disadvantages. Deep neural networks usually have millions of parameters, and fine-tuning such models requires substantial computational resources and expertise. In the future, patients' health attributes can be integrated with image features to further improve classification accuracy.

Table 7 Comparative analysis with recent works

Reference	Evaluation Metrics			
	Accuracy	Sensitivity	Specificity	
			F1-score	
[23]	Overall accuracy = 98.22%, G=99%, M=99.13%, P=97.3%	G=96.12%, M=98.28%, P=94.74%	G=99.65%, M=99.64%, P=99.37%	G=97.85%, M=98.95%, P=97%
[4]	FF-ANN = 95.83%, RSwthRF = 97.14%, RSwthBN = 95.71%	-	-	-
[25]	GoogLeNet +SVM = 97.6%, GoogLeNet +KNN = 98.3%	GoogLeNet +SVM: G=97.24%, M=95.80%, P=99.20%; GoogLeNet +KNN: G=98.02%, M=94.57%, P=99.10%	GoogLeNet +SVM: G=98.93%, M=98.40%, P=99.30%; GoogLeNet +KNN: G=98.63%, M=98.65%, P=99.01%	GoogLeNet +SVM = 97.35% GoogLeNet +KNN = 97.24%
[13]	Classification I = 99.33%, Classification II = 92.66%, Classification III = 98.14%	-	-	-
[14]	99%	99%	-	99%
[9]	With GAN = 93.01% & 95.60%; Without GAN = 91.70%	With GAN = 92.01 & 94.91%; Without GAN = 90.16%	With GAN = 96.39% & 97.69%; Without GAN = 95.58%	With GAN = 92.10% & 95.10%; Without GAN = 90.54%
[16]	98.495%	-	-	-
[17]	92.13%	-	-	-

Abbreviation:: G: Glioma, M: Meningioma, P: Pituitary

Data Availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

References

1. Aamir M, Rahman Z, Dayo ZA, Abro WA, Uddin MI, Khan I, Imran AS, Ali Z, Ishfaq M, Guan Y et al (2022) A deep learning approach for brain tumor classification using mri images. *Comput Electr Eng* 101:108105
2. Amirkhani D, Bastanfard A (2021) An objective method to evaluate exemplar-based inpainted images quality using jaccard index. *Multimed Tool Appl* 80(17):26199–26212
3. Amirkhani D, Bastanfard A (2019) Inpainted image quality evaluation based on saliency map features. In: 2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), IEEE, pp 1–6
4. Assam M, Kanwal H, Farooq U, Shah SK, Mehmood A, Choi GS (2021) An efficient classification of mri brain images. *IEEE Access* 9:33313–33322
5. Bastanfard A, Amirkhani D, Mohammadi M (2022) Toward image super-resolution based on local regression and nonlocal means. *Multimed Tool Appl* 81(16):23473–23492
6. Br35h : Brain tumor detection 2020 | kaggle. <https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection>
7. Brain tumor classification | kaggle. <https://www.kaggle.com/code/duonggiakhanhb/brain-tumor-classification/notebook#CNN-3>
8. Figshare brain tumor dataset. https://figshare.com/articles/dataset/brain_tumor_dataset/1512427/5
9. Ghassemi N, Shoeibi A, Rouhani M (2020) Deep neural network with generative adversarial networks pre-training for brain tumor classification based on mr images. *Biomed Signal Process Control* 57:101678
10. Gupta A, Bhagat M, Jain V (2023) Blockchain-enabled healthcare monitoring system for early monkeypox detection, *J Supercomputing*:1–25
11. Gupta A, Singh A (2023) Edl-nsga-ii: Ensemble deep learning framework with nsga-ii feature selection for heart disease prediction, *Expert Syst:e13254*
12. Iqbal S, Khan M, Saba T, Rehman A (2018) Computer-assisted brain tumor type discrimination using magnetic resonance imaging features. *Biomed Eng Lett* 8(1):5–28
13. Irmak E (2021) Multi-classification of brain tumor mri images using deep convolutional neural network with fully optimized framework. *Iranian J Sci Technol Trans Elect Eng* 45(3):1015–1036
14. Ismael SAA, Mohammed A, Hefny H (2020) An enhanced deep learning approach for brain cancer mri images classification using residual networks. *Artificial Intell Med* 102:101779
15. Jena B, Nayak GK, Saxena S (2022) An empirical study of different machine learning techniques for brain tumor classification and subsequent segmentation using hybrid texture feature. *Mach Vis Appl* 33(1):1–16
16. Khairandish M, Sharma M, Jain V, Chatterjee J, Jhanjhi N, A hybrid cnn-svm threshold segmentation approach for tumor detection and classification of mri brain images, *IRBM*
17. Khan AH, Abbas S, Khan MA, Farooq U, Khan WA, Siddiqui SY, Ahmad A (2022) Intelligent model for brain tumor identification using deep learning, *Appl Comput Intell Soft Comput*
18. Minoofam SAH, Bastanfard A, Keyvanpour MR, Trcla: a transfer learning approach to reduce negative transfer for cellular learning automata, *IEEE Transactions on Neural Networks and Learning Systems*
19. MirMashhour A, Bastanfard A, Amirkhani D (2022) Collecting a database for emotional responses to simple and patterned two-color images. *Multimed Tool Appl* 81(13):18935–18953
20. Morid MA, Borjali A, Del Fiol G (2021) A scoping review of transfer learning research on medical image analysis using imagenet. *Comput Biology Med* 128:104115
21. Musallam AS, Sherif AS, Hussein MK, A new convolutional neural network architecture for automatic detection of brain tumors in magnetic resonance imaging images, *IEEE Access*
22. Nawaz SA, Khan DM, Qadri S (2022) Brain tumor classification based on hybrid optimized multi-features analysis using magnetic resonance imaging dataset, *Appl Artif Intell*:1–27
23. 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

24. Sartaj brain tumor classification (mri) | kaggle. <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>
25. Sekhar A, Biswas S, Hazra R, Sunaniya AK, Mukherjee A, Yang L, Brain tumor classification using fine-tuned googlenet features and machine learning algorithms: Iomt enabled cad system, IEEE Journal of Biomedical and Health Informatics
26. Sharif MI, Li JP, Khan MA, Saleem MA (2020) Active deep neural network features selection for segmentation and recognition of brain tumors using mri images. *Pattern Recog Lett* 129:181–189
27. Tulbure A-A, Tulbure A-A, Dulf E-H (2022) A review on modern defect detection models using dcnn-deep convolutional neural networks. *J Adv Res* 35:33–48

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.