Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-024-09284-z

RESEARCH ARTICLE-ELECTRICAL ENGINEERING

Customized CNN for Multi-Class Classification of Brain Tumor Based on MRI Images

Bentahar Heythem[1] · Mohamad Djerioui[2] · Tawfiq Beghriche[2] · Azzedine Zerguine[3] · Azeddine Beghdadi[4]

Received: 14 February 2024 / Accepted: 9 June 2024

Abstract

In this paper, we propose a new strategy to exploit the advantages of Deep Neural Network-based architectures for brain tumor classification using MRI images for a better diagnosis. This was achieved by analyzing and evaluating pre-trained models on three different datasets**.** To better design the optimal architecture for solving the classification of brain tumor using MRIs, we have conducted extensive experiment-based analysis on how different layers of Convolutional Neural Network (CNN) process the inputs. Four distinct architectures are then built, each with its specific hyperparameters and layers. The images are fed into the convolutional layers for feature extraction followed by a softmax function before applying the classification process. An extensive experimental study carried out clearly demonstrates that our novel CNN-based classification approach achieves state-of-the-art accuracy, precision, recall and an F1-score of 99.76% 99.64% 99.62% and 99.64%, respectively. Also, a higher performance in terms of Micro-Avg Matthew correlation coefficient (MCC) of 0.929 is achieved. This exceptional performance is achieved thanks to the new proposed model's architecture. Indeed, unlike conventional methods, that often rely on complex transfer learning models or hybrid architectures, our approach utilizes a custom and non-hybrid scheme. Consequently, this streamlined architecture offers a significant advantage of being remarkably lightweight, enabling efficient operation on resource-constrained computing systems.

Keywords Brain tumor classification · Custom CNN · Transfer learning · MRI

Part of this work was presented at the *International Conference of Advanced Technology in Electronic and Electrical Engineering, ICATEEE 2022*, 2022 [1].

✉ Azzedine Zerguine
azzedine@kfupm.edu.sa

Bentahar Heythem
heythembentahar@gmail.com

Mohamad Djerioui
mohamed.djerioui@univ-msila.dz

Tawfiq Beghriche
tawfiq.beghriche@univ-msila.dz

Azeddine Beghdadi
beghdadi@sorbonne-paris-nord.fr

1 Department of Electronics, Faculty of Technology, University of Mostaganem, Mostaganem, Algeria

2 LASS Laboratory, Department of Electronics, Faculty of Technology, University of M'Sila, M'Sila, Algeria

3 Electrical Engineering Department, and Center for Communication Systems and Sensing, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia

4 L2TI, Institut Galilée, Université Sorbonne Paris Nord, Paris, France

1 Introduction

One of the major causes of increased mortality rates in the world is due to brain tumors [2]. According to the National Brain Tumor Society (NBTS), about 700,000, people developed a brain tumor in the U.S. only and 250,000 people died from it in 2020 in the rest of the world [3, 4]. The early diagnosis of a brain tumor plays an important role in scheduling an effective treatment plan and patient care. Moreover, early detection of brain tumors aids radiologists in achieving an accurate prognosis and hence may increase the chances of long-term survival and the possibility of full recovery. However, the task of classifying MRI images of a brain tumor is a delicate and challenging one for radiologists. Indeed, this manual or even semi-automatic brain tumor classification is impractical, unreproducible, time-consuming, costly, and prone to errors, when processing a large amount of MRI images [5]. This is due to several factors, including the fact

Published online: 12 July 2024

that tumors may be present in different locations, sometimes difficult to access, and in different shapes, contrasts, sizes, and visual appearances.

Therefore, to improve the quality of early detection and classification that would lead to a successful and timely prognosis, it is of the utmost importance to bring the recent powerful tools of machine learning (ML) methods especially DL [6]. This would help to bear on this important field by automating the whole diagnosis process with the use of computer-based tumor classification systems. This can be used to classify MRI images of brain tumors automatically, rapidly, and accurately by minimizing radiologist intervention in repetitive tasks that can lead to diagnostic errors due to fatigue and visual discomfort during MRI image analysis. This imaging technique is frequently utilized by scientists to automatically detect brain tumors and monitor their progression over time. Indeed, MRI is an essential imaging modality, used to evaluate a range of morphological and functional targets, particularly in soft tissue and organs. Its effectiveness in terms of clinical outcomes is indisputable and has been widely proven in numerous studies, both in the early diagnostic phase and as an effective means of image-guided therapeutic treatment [7]. MRI is considered as the most appropriate non-invasive imaging modality for exploring the anatomical structures of the human brain structure, thereby offering comprehensive information about them which facilitates the reliable detection of brain tumors [8]. Scientists have developed various methods to identify and classify brain tumor using MRI scans. These methods encompass a wide range of techniques, ranging from traditional medical image processing to state-of-the-art ML methods [9–13]. In recent years, a variety of DL methods have been employed in healthcare systems [9–11, 14–19]. In this context, many researchers have used methods such as CNN, to classify the different brain tumors. In [20], the authors developed a brain tumor classification model using CNN techniques. They created seven different CNN variants. On the FigShare brain MRI dataset and without any prior region-based segmentation, the second variant of their model achieves the best training and testing accuracies, with 98.51 and 84.19 percent, respectively. Sultan et al. [21] presented a DL-based model to classify various types of brain tumors using two publicly available datasets. Their proposed system architecture achieves training and testing accuracy levels of 96.13% and 98.7%, respectively.

On the other hand, Deep transfer learning is used in some research for categorization [22, 23] classification [24–26], and segmentation [27, 28] purposes. A proposal for an ensemble approach that combines deep features extracted from pre-trained deep CNN with ML classifiers appeared in [29] based on a small dataset. The DenseNet-169 and ResNeXt- 101 models had an average classification accuracy of 92.37% and 96.13%, respectively. A real-time medical classification system would need knowledge-distillation techniques in order to be successfully applied to these models. Noreen et al. [30] developed a DL-based concatenation mechanism for the detection of brain tumors. In order to classify the detected tumors, the features from the two pre-trained models DenseNet201 and InceptionV3 were first extracted independently, then concatenated, and finally transmitted to different classification layers (Softmax). The DenseNet201-based features concatenation approach demonstrated a remarkable 99.51% accuracy in the case of the Figshare dataset comprising 3064 T1-weighted contrast MR images collected from 233 patients [42].

In this paper, we proposed a brain tumor prediction method from MRI images for a better diagnosis than has previously been achieved. By introducing three different datasets and making a comparison with pre-trained models, we succeeded in developing a better CNN architecture (in terms of accuracy) for this specific task, namely brain tumor classification.

In brief, contrast to the research outlined in [1], this study distinguishes itself through the following contributions:

- Proposing a new method for efficient classification of brain tumors based on MRI images outperforming existing solutions in terms of accuracy and complexity.
- Introducing four distinct architectures and conducting an extensive compsarative study to determine the most effective one. This streamlined architecture offers the significant advantage of being remarkably lightweight, enabling efficient operation on resource-constrained computing systems.
- Creating a novel custom CNN (hand-crafted) model. A fast, simple and non-hybrid architecture that can operate effectively even on resource-constrained systems. Unlike conventional methods that often rely on complex transfer learning models or hybrid architectures.
- Conducting a performance-based comparative analysis against various Transfer Learning variants for the classification of brain malignancies, encompassing glioma, meningioma, and pituitary tumors.
- Emphasizing the experimental results achieved with a notably high accuracy using brain tumor datasets and showcasing the superior performance of the proposed model compared to state-of-the-art modern methods. A promising solution for learning and classifying brain tumors is both replicable, adaptable, and scalable to other brain diseases that enjoy non-invasiveness, cost and time efficiency.

The rest of the paper is organized as follows: Sect. 2 gives a literature review of some previous and relevant research in this area. Section 3 deals with the brain tumor system architecture, dataset description, pre-processing techniques, and the proposed CNN-, and DL-based approaches. Section 4 discusses some key evaluation metrics and their application to the proposed methods, and presents the various detailed results obtained for comparison purposes. Section 5 presents a comparative study between the proposed methods and the state-of-the-art ones in terms of brain tumor detection and classification and clearly demonstrates the superior classification accuracy achieved by our proposed methods. Finally, Sect. 6 concludes the paper.

2 Literature Review

Numerous research studies have been carried out on the important topic of brain tumor detection and classification, and some of the key ones that are mainly based on CNN and DL models are reviewed here. For the purpose of brain tumor classification, a deep CNN with multiple layers was developed by Ayadi et al. [31]. Their model demonstrated impressive performance, requiring significantly less pre-processing compared to previous methods. However, a notable drawback is the inconsistency and data imbalance present within the datasets; this can lead to bias issues in prediction later. Deepak et al. [32] applied a pre-trained model to evaluate three classes of brain tumors, based on MRI images obtained from FigShare. Their model achieved a classification accuracy of 98% using a small number of training cases; with fewer samples, the model may memorize the training data rather than learning patterns and they also offered a misclassification analysis. A binary classification problem was discussed about brain tumor detection by the authors in [33]. Their approach consisted of merging the extracted features from two pre-trained models, AlexNet and VGG16, then fusing them into a recurrent feature elimination (RFE) scheme and finally using an SVM classifier that gives a 96% accuracy. Although the model's concept is innovative, its complexity could pose significant challenges regarding computational resources and interpretability.

Another multi-class approach for classifying brain tumors using a deep neural network is provided in [34]. With the use of augmentation techniques on data obtained from FigShare, the method collects information from images and learns their structure in order to pre-train a CNN as a discriminator in a generative adversarial network (GAN). This network is not overfitting, thanks to the augmentation techniques. The model has been trained to work as a classifier, and the fully connected (FC) layers of the network have been replaced with a Softmax layer, and then, the whole resulting deep neural network is trained as a classifier to discriminate between the tumor types. Using five-fold cross-validation, the model is judged to have an accuracy of 93.1% and an accuracy of 95.6% on a random split. While this model boasts high accuracy, it is crucial to prioritize addressing data quality concerns, with a particular focus on ethical considerations, when dealing with GAN models. Kumar et al. [35] have proposed a DCNN model for brain tumor categorization. Using the BRATS dataset as an input, the images were first pre-processed to remove noise. The method given here was utilized to identify cancers using non-tumor cells, the approach begins by pre-processing using an NLM filter, followed by segmentation employing a fusion model termed Dolphin-SCA, and subsequent feature extraction via various methods. Finally, a Deep CNN is applied. Executing these steps on a dataset containing just 65 samples may substantially lead to overfitting and data memorization. Nonetheless, they managed to attain an accuracy of 96.3%. The work in [36] considers transfer learning as a crucial component for dealing with small dataset challenges. The authors of this study created an advanced algorithm utilizing VGGnet as the foundation architecture. Two publicly available datasets (BRATS and CE-MRI) were used in this study. The findings from these 2 datasets produced high accuracies of 97.28% and 98.69%, respectively.

Khairandish et al. [37] proposed a brain tumor detection and classification model based on a deep-learning hybrid model. They implemented a CNN network used as a feature extractor and an SVM as a classifier with threshold-based segmentation for the detection. Despite its hybrid architecture, achieving a 98.49% accuracy falls short of expectations given the critical nature of the problem.

Kubra et al. [38] built what they called "A novel DL model radiologist emulator" for the diagnosis of brain normality/abnormality from brain CT images. This model labels the region of the identified object borders using a faster ResNet50-modified R-CNN, which achieved a very high accuracy of 99.75%. This result was achieved by utilizing a novel brain dataset containing both normal and abnormal images. With the aid of segmentation techniques Gopal et al. [39], managed to successfully achieve excellent results amounting to an accuracy of 99.06%. The segmented images based on techniques like (RSM, SSM) were passed through some pre-trained models such as ResNet18 and GoogleNet, before reaching the final step of classification. The study carried out by Sindhiya Devi et al. [40] offers a hybrid DL-based method for classifying and diagnosing brain tumors. They first introduce an image segmentation phase and then the resulting images are fed through the entire features extraction procedure by applying a stationary wavelet packet transform (SWPT). Thanks to its use of the wavelet transform, this hybrid technique is capable of extracting fine features. Finally, they used a hybrid CNN-LSTM network for accuracy optimization, which led to a high accuracy of 97.85%.

Using hybrid techniques in AI can also have drawbacks. Integrating different techniques can introduce complexities and inefficiencies that can impact system performance. It can be challenging to balance and optimize the various components. Additionally, hybrid techniques may require more computational resources than individual techniques, which can lead to higher memory requirements, longer processing times, or the need for more powerful hardware infrastructure. Despite this, researchers may prefer hybrid over non-hybrid techniques because of the capacity to detect fine features by combining the strengths of different techniques to overcome the weaknesses of each individual. For example, a hybrid technique that combines edge detection with texture analysis can be used to detect fine features that would be difficult to detect with either technique on its own.

Abd-Ellah et al. [41] present a two-phase DL system for brain tumor detection and localization in MRIs. The first phase uses CNN and ECOC-SVM for feature extraction and classification, while the second phase utilizes a five-layer R-CNN for tumor localization. The study aims to classify MRIs into normal and abnormal images and accurately identify the tumor's location in abnormal MRIs. The proposed method achieves a very significant accuracy of 99.55% based on evaluation with 349 MRIs from the RIDER Neuro MRI database. The system's performance evaluation relies on specific datasets, such as the BraTS 2013 database, which is a recognized benchmark. However, it is crucial to ensure that the selected dataset represents real-world clinical scenarios to demonstrate the system's effectiveness across a wider range of cases. Additionally, the system's ability to generalize and handle different types of brain tumors and imaging conditions is not explicitly addressed. It is important to assess the system's performance on various tumor types, sizes, locations, and imaging protocols to ensure its reliability and practicality in real clinical settings. Also, the consideration of data imbalance is crucial and should be taken into account, as it can introduce bias into the model's predictions.

It is important to note that despite these remarkable results for some of the solutions discussed here, they do have some limitations. Some studies are limited to well-targeted databases and present increased levels of complexity. On the other hand, it is always important to further increase accuracy in the context of medical diagnosis. This is the main objective of the solution proposed and presented in this article. A summary of the literature review is depicted in Table 1.

3 Proposed System

The objective of this study is to establish an efficient CNN-based architecture that can accurately classify various types of brain tumors. Firstly, the input MRI dataset is pre-processed and split into training and testing sets. Then, the training part is fed into both the pre-trained and custom CNNs (CNNs built from scratch). By the end of the training process, the trained model will be capable to predict the testing data images of the brain tumor. Implementing such a system greatly facilitates the diagnostic process making it faster, easier and most importantly significantly enhancing accuracy and reliability. Figure 1 depicts the overall flowchart of our proposed system.

3.1 Datasets

To evaluate the performance of this work, we used three different datasets. The datasets used in this project were limited; however, it underwent a thorough cleaning and were appropriately represented prior to being used in the classification process. Given that our prediction and classification study has been based on limited-scope datasets, and for higher reliability, prior to being employed in real-life applications, it may be essential for professionals to subject the data to a secondary validation process. Batches of 2D (sagittal, coronal and horizontal) slices have been fed to our models (Fig. 2). Each of the datasets used, namely (3264, 7022, 11,119) T1 [42], T2 [43] and FLAIR MRI [44], respectively, contains four classes (Glioma, Meningioma, Pituitary, No tumor). More Information on this can be found in Table 2.

3.2 Pre-processing

This first stage, known as data augmentation, is used to enrich the database by artificially increasing its size. Image data augmentation, widely recognized as a prominent form of data augmentation, focuses on generating modified versions of images from the training dataset while retaining their original class. This can be achieved by using some elementary image transformation. This data pre-processing is a necessary and crucial stage that transforms the data into a usable and efficient format so that it can fit as an input to the DL algorithm. Pre-processing encompasses all the transformations performed on raw data prior to feeding it into a ML or DL algorithm. Image resizing and augmentation are the only used pre-processing approaches in this work. Data pre-processing starts with resizing the images to a 224 × 224 format before feeding them into our models as input. While image augmentation involves generating transformed versions of the original images within the training dataset while maintaining the same class label. Four types of augmentation techniques, including horizontal flipping, zoom, shear, and scaling, are applied to different datasets [45]. Figure 3 illustrates some samples from the datasets belonging to various image pre-processing techniques. Hereafter, we will elaborate on four augmentation techniques employed in our system.

- The rescale factor is applied to the data as a multiplier before proceeding with subsequent processing steps. In our

Table 1 Different brain tumor classification techniques

Reference	Used method	Dataset	Accuracy	Advantages	Disadvantages
[31]	DCNN	FigShare/REMBRANDT	94.74%	–	Classification accuracy needs to be increased
[32]	GoogLeNet	FigShare	98%	–	Small dataset
[33]	AlexNet and VGG16 fusion	Kaggle	96.77%	The fused deep feature set is considered to exploit the generalization abilities	Model complexity/Accuracy needs to be increased
[34]	GAN model (CNN discriminator)	FigShare	95.6%	This model tackles low dataset size	Input size of 64^2 due to GAN limitation
[35]	Dolphin-SCA-based CNN	BRATS/simBRATS	96.3%	A segmentation is also done after classification	Very small dataset utilized with a complex model
[36]	Fine-Tuned VGG19 (BT-VGG-Net)	BRATS 2018/CE-MRI	98.69%	Using Transfer learning knowledge is highly effective	–
[37]	Hybrid CNN-SVM	BRATS 2015	98.49%	Consideration of both CNN and SVM model advantages	Classification accuracy needs to be increased
[38]	Faster R-CNN	Real-World dataset	99.75%	Very high classification accuracy	Binary classification
[39]	ResNet18/GoogleNet	TCIA Repository	99.06%	A majority voting-based ensemble algorithm is proposed	Requires additional human intervention
[40]	Deep CNN-LSTM	–	97%	Adaptive Black widow optimization with Moth Flame optimization is introduced	Classification accuracy needs to be increased
[41]	CNN/ECOC-SVM R-CNN	BRATS 2013	99.55%	Very high classification accuracy	Only Binary Classification

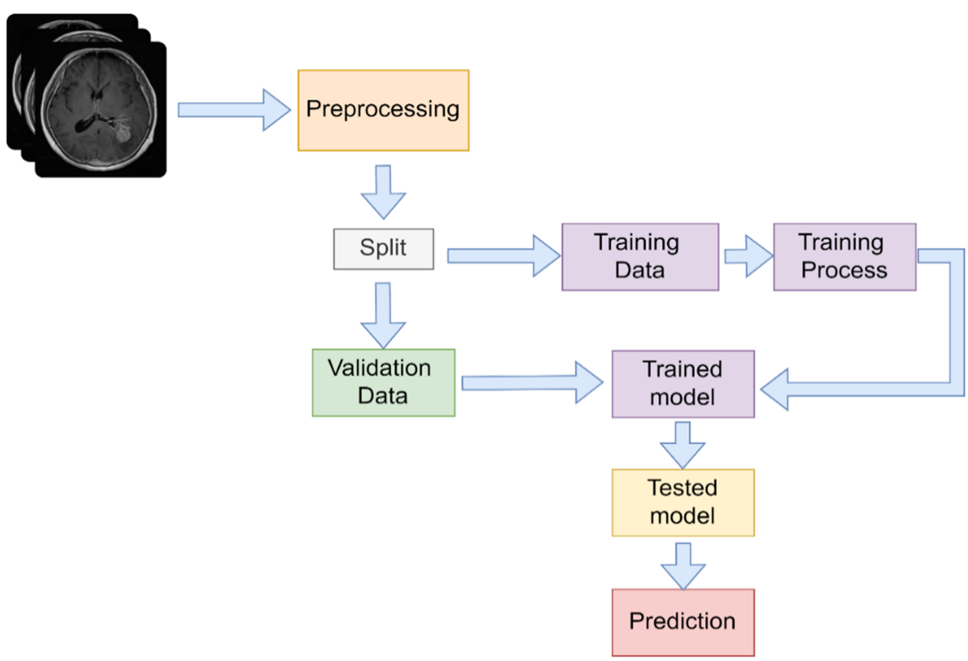


Fig. 1 Schematic diagram of the Workflow

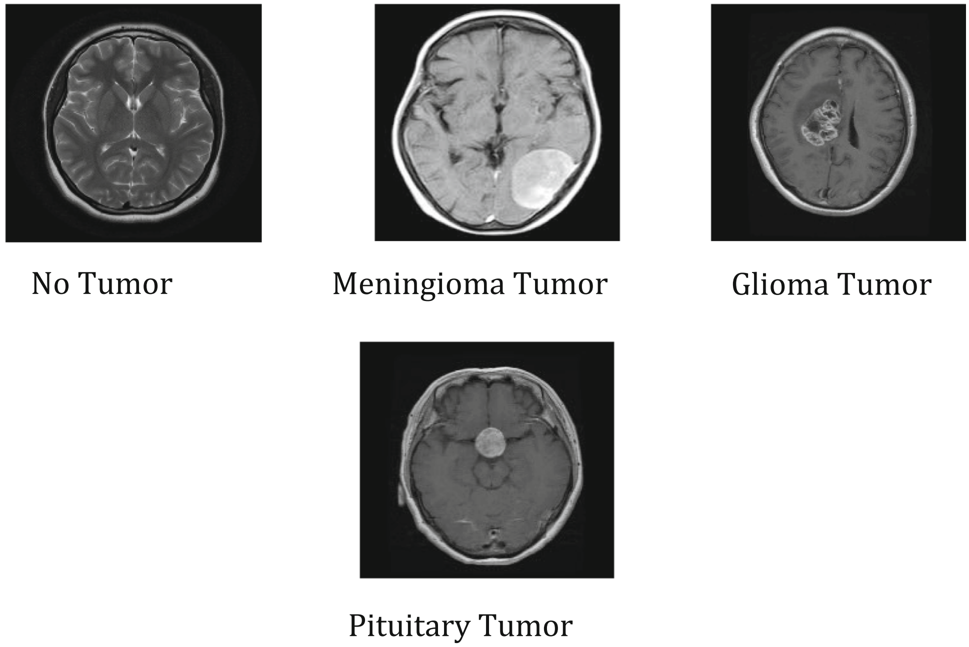


Fig. 2 Brain tumor image sampless

Table 2 Dataset description

Dataset	Total sampsles	Samples after Augmenta-tion	Classes	No. of samples in each class			
				Glioma	Meningioma	Pituitary	No tumor
Dataset 1 T1	3264	13,056	4	962	937	901	500
Dataset 2 T2	7022	28,088	4	1621	1645	1757	2000
Dataset 3 Flair MRI	11,119	44,476	4	2772	2774	2873	2700

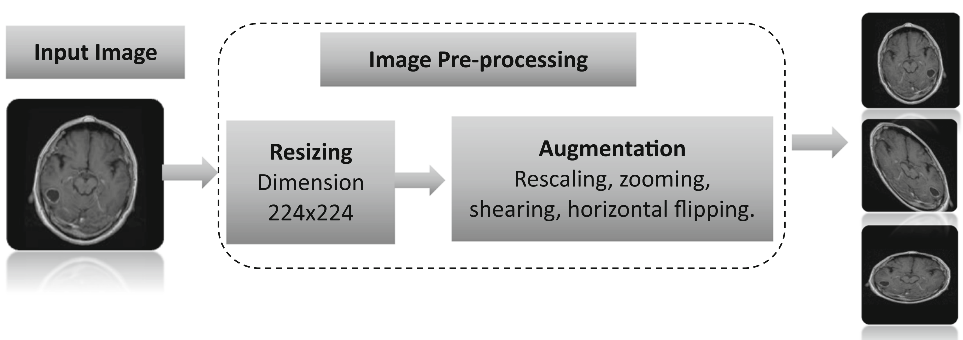


Fig. 3 Data pre-processing stages

case, the original images contain gray levels ranging from 0 to 255. However, these values are not suitable for our models to effectively process with a typical learning rate. To address this, the input image gray levels are linearly scaled on the range [0,1].

- Shear range refers to the extent of distortion or slanting applied to an image along a specific axis.
- Zoom range represents the allowable range of magnification or reduction applied to an image during the data augmentation process.
- The horizontal flip operation randomly mirrors half of the images horizontally, which is particularly useful in situations where there are no inherent assumptions of horizontal asymmetry [46].

3.3 Transfer Learning

Transfer learning is a technique that leverages the knowledge acquired from a pre-trained model to learn and adapt

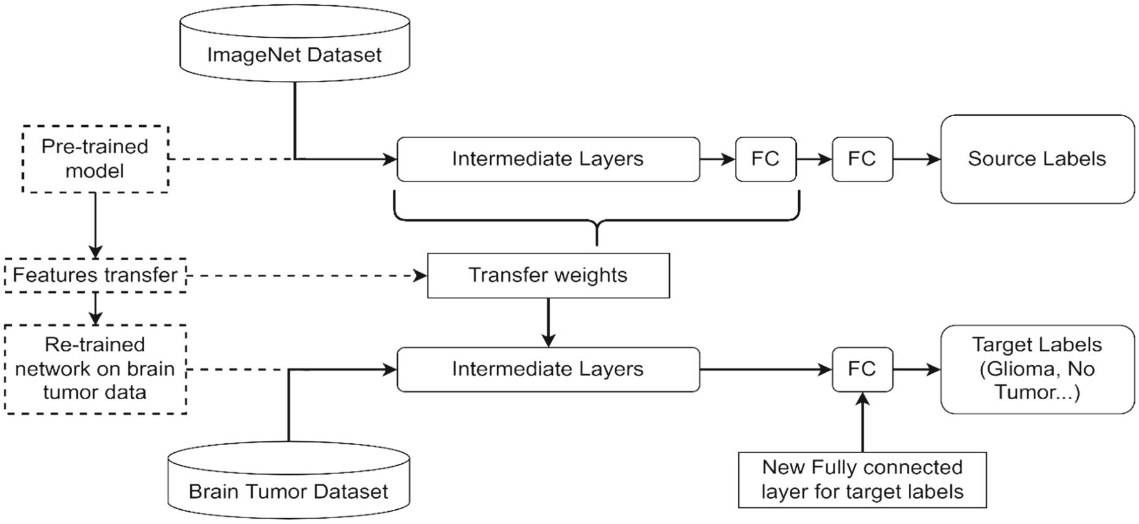


Fig. 4 An illustration of a transfer learning process

to a new dataset [47]. The acquired knowledge from previous training is utilized to aid in the execution of a new task. When labeled data are accessible in both the source and target domains for a classification task, the transfer learning approach is referred to as inductive transfer learning [48]. For the original trained model to effectively adapt to new, unseen data, it typically needs to possess a high level of generalization ability [49]. Transfer learning eliminates the need to start training models from scratch for each new task, which can be resource-intensive, time-consuming, and costly. By leveraging pre-trained models, transfer learning significantly saves resources, cost, and time (Fig. 4).

In this study, we used four pre-trained CNN models. These models were selected based on their parameters and popularity in order to make a comparison with previous similar research. These models can be found on the Keras application site.

3.3.1 VGG16 Model

VGG16 (Visual Geometry Group) is one of the widely used CNN models, and yet, it has a simple architecture that only consists of 16 layers [50]. It has been trained on ImageNet, which is a large visual database project used in visual object recognition software research. The VGG16 architecture was first developed and introduced by Karen Simonyan and Andrew Zisserman in 2014 [50].

3.3.2 VGG19 Model

The concept of the VGG19 model (i.e., VGGNet-19) is similar to the VGG16 (Fig. 5) except that it contains 19 convolutional layers rather than sixteen. We will further discuss the characteristics of VGG16 and VGG19 networks in a later part of this study [51]. Figure 6 illustrates the model architecture of VGGNet-19 (Fig. 7).

3.3.3 ResNet50 and ResNet152V2 Models

The ResNet (Residual Network) is a ground-breaking neural network architecture introduced in 2015 by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun in [52]. A deep Residual Network (ResNet) shares similarities with networks composed of convolutional, pooling, activation, and fully connected layers stacked sequentially. The distinctive aspect of a ResNet is the inclusion of identity connections between the layers as shown in Fig. 8, which enables it to be classified as a residual network [53]. The ResNet model achieved remarkable success, demonstrated by its ensemble securing first place in the ILSVRC 2015 classification competition with a mere 3.57% error rate. Moreover, it also emerged as the winner in ImageNet detection, ImageNet localization, COCO detection, and COCO segmentation categories in the ILSVRC and COCO competitions of 2015. ResNet has several variations based on the same concept but with varying numbers of layers. ResNet50, denoted by Fig. 9, refers to a variant that operates with 50 neural network layers [53].

3.4 Customized Models

To better design the optimal architecture for solving the problem of classifying brain tumor using MRI images, we have conducted extensive experiment-based analysis on how different layers of CNNs process the input images. For the sake of clarity and to make this paper self-contained, we briefly recall some basic concepts and notions. One of the key elements in a standard CNN architecture is undoubtedly the convolution block present at different levels of the network

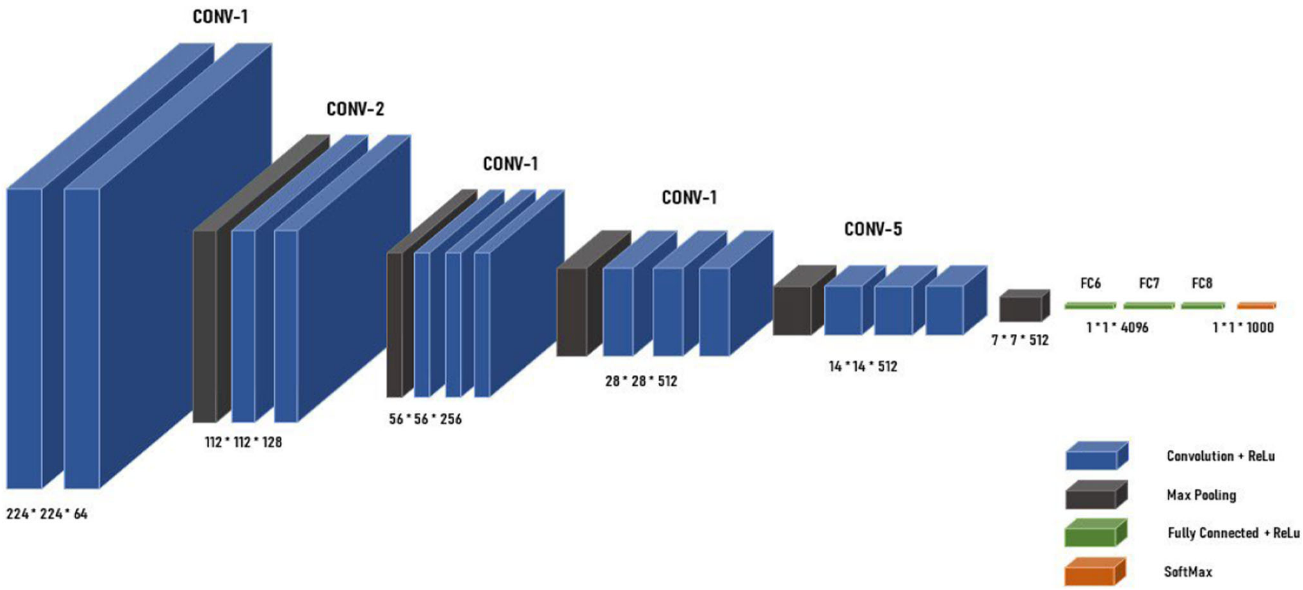


Fig. 5 VGG16 model architecture

Fig. 6 VGG19 model architecture

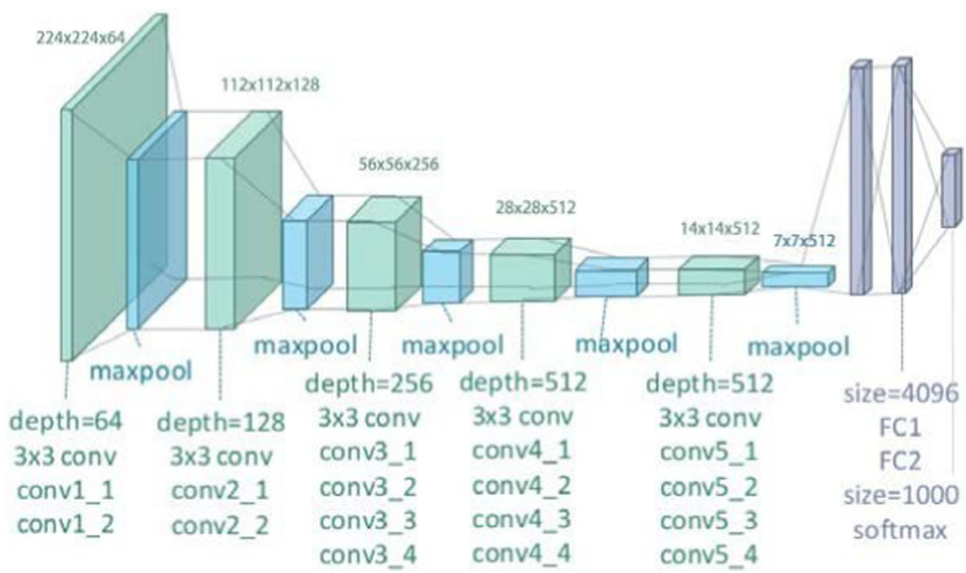
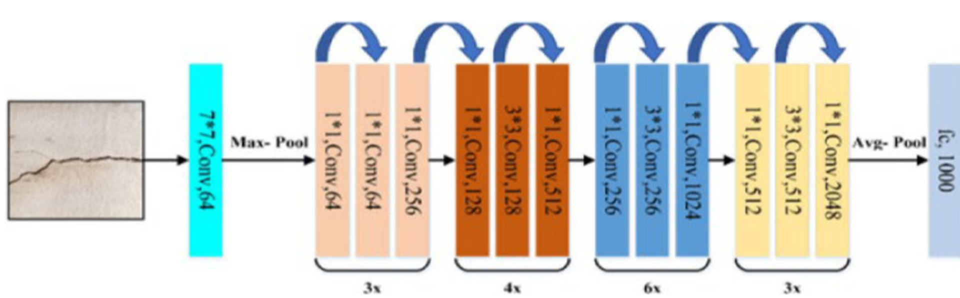


Fig. 7 ResNet50 model architecture



layers. A convolutional layer in a CNN applies filters to input data through convolutional operations. These filters highlight spatial patterns, and their weights are learned during training. The convolution involves sliding filters' masks across an image, producing feature maps that highlight different features and then stride and padding control the filter movement and spatial dimensions. Multiple filters in a layer capture

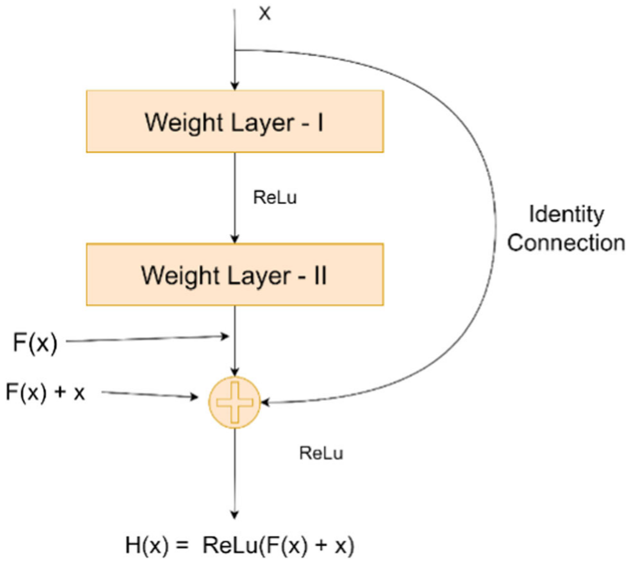


Fig. 8 A Residual Block of Deep Residual Network

diverse features after that optional pooling layers are introduced to reduce spatial dimensions. Convolutional layers are crucial for learning hierarchical features in spatial data. As a result, we have adopted four distinct architectures, each with unique parameters and layers, tailored to suit our specific objectives. Each one of these four is an improvement over its previous one in terms of accuracy and precision. Ultimately leading to improved model generalization for this particular task. The advantage of building a CNN model specifically for this task makes it more performant and more generalized than a pre-trained model with tweaks its last layers. Table 3 presents the proposed architectures, and Fig. 10 depicts an illustration of these architectures (Fig. 11).

4 Results & Discussion

To assess the effectiveness of our system, we employed a range of widely recognized performance metrics, which included precision, recall, f1-score, support, and accuracy. These metrics were utilized to evaluate and measure different aspects of the system's performance and provide comprehensive insights into its classification capabilities. While accuracy is a commonly used metric for comparing models, it can sometimes be misleading, especially in imbalanced datasets. In our evaluation, we recognized the importance of fairness and utilized accuracy as our primary metric. However, we also took into account other performance metrics to ensure a comprehensive assessment of the classification model's performance. By considering multiple metrics, we aimed to gain a deeper understanding of the model's effectiveness and make informed decisions regarding its performance and suitability for our application. The fourth model paired with the third dataset gives us the following results. The experimental setup for our study is divided into Software and Hardware; in the hardware part, we used a Ryzen 3–3100 CPU paired with a GTX 1050 GPU. On the software side, we relied primarily on Tensor-flow and Keras API.

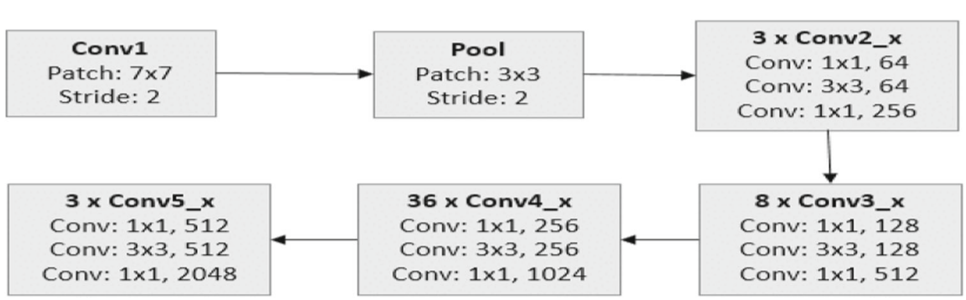
Table 4 Classification performance evaluation metrics

s	Mathematical formsula
Accuracy	$Acc = \frac{Tp+Tn}{Tp+Fp+Tn+Fn}$
Precision	$Pre = \frac{Tp}{Tp+Fp}$
Recall	$Rec = \frac{Tp}{Tp+Fn}$
F1-Score	$F1.Score = 2.\frac{Precision x Recall}{Precision+Recall}$

4.1 Performance Evaluation of the Proposed Method

The classification report serves as an essential performance evaluation metric for classification-based ML models. It presents crucial parameters of the model, including accuracy, precision, recall, F1-score, and support. This report offers a comprehensive overview and deeper insights into the overall performance of the trained model, aiding in a better understanding and assessment of its effectiveness [54] and hence provides the user with a useful and reliable model selection criterion. For the sake of completeness of the paper, we recall the mathematical expressions of these metrics in Table 4, where TP, TN, FP, and FN mentioned in these expressions

Fig. 9 Architecture of ResNet152V2 model



refer to True Positive, True Negative, False Positive, and False Negative, respectively.

4.2 Prediction Methods

We consider two approaches in our work: First, we start with the use of the Transfer learning approach using already-trained models and then, making suitable changes to the last few layers of these pre-trained models so that they can predict exactly the data we fed into them, which, in our case, is our own data set made of Brain tumor MRI images. The modifications made to the last few layers of the model included adding a fully connected layer and a hidden layer with 4 nodes, each representing one of the tumor classes. In the second approach, we create our own Custom CNN model, specifically designed for our own dataset so that we can achieve a high-performance model characterized by higher precision and accuracy and a low prediction time. We shall now discuss in some detail each of these two approaches. In summary, the combination of transfer learning and custom CNN models allows us to capitalize on the strengths of both approaches: leveraging pre-trained models for their general feature learning capabilities and fine-tuning them for the specific task, while also having the flexibility to design custom architectures optimized for the dataset's characteristics. Also, we wanted to include a generalized model to have baseline results to make a comparison with.

4.2.1 Transfer Learning Approach

In this approach, we used four of the most popular models available in the literature and retrained them with our datasets to test their respective performance. Also, in this approach, we used three different dataset sizes (from small to large), so we can get a clear idea of how the dataset size affects our DL process. Starting with the first dataset, we note from Table 5 that VGG16, VGG19, and ResNet152V2 give close results when testing for their respective accuracies, i.e., 80.63%, 75.42%, 82.56%, and ResNet50 ranked last with an accuracy of 56.67%.

The second dataset has more elements than the first. Technically, although it took more time to be processed, it produced much better accuracy results, with the highest accuracy of 94.03% still being retained by ResNet152V2, followed by VGG16, then VGG19 and finally ResNet50 ranking last, as was the case with the first data set. Such an improvement in accuracy results is to be expected as the number of samples in the second dataset is far greater than that in the first data set and amounted to around 7022 samples that were divided into four different classes of brain tumors. With the third data set, which accounts for around 11,119 samples, ResNet152V2 retained its position as the top performer with 98.17% accuracy. Except for ResNet50 which seems to struggle with dealing with larger datasets, the other 3 learning models improved appreciably their accuracy performances as the number of images fed into them grew. The expected conclusion to infer from training these 4 models with the differently sized datasets is that the larger the training data, the better the classification performance. This conclusion assumes the availability of large datasets whenever needed. However, if for some reasons related to either the difficulty of data collection or cost, etc., the number of training data available is insufficient, then, better classification results may still be obtained through using some other techniques such as random sampling, bootstrapping of k-fold validation techniques [55, 56]. An alternative way of handling insufficient datasets relies on doing multiple experiments and choosing the best results from them [57].

4.2.2 Customized CNN-Based Approach

Table 6 shows the performance of the four different CNN-based models evaluated through different metrics using the third dataset as an input. We used the last dataset, which is the largest to achieve the best possible all-round performance. These architectures differ in the number of hidden layers used in the model and in the kernel size of each layer. Each architecture has more layers in it but not necessarily more parameters. We deliberately used different numbers of parameters in the layers to illustrate that the number of parameters does not affect the performance of the model directly as shown through the results summarized below in Table 6. As mentioned in Table 6, the first two models achieved 89% and 92.55% accuracy, respectively.

Then, the third and fourth models achieve 99.52% and 99.76% accuracy, respectively. The closeness of the performances of the third and fourth models is reflected by the close similarity of their respective architectures. Though the difference in their performances is quite small, it may have a significant impact on the final decision-making process which will determine which model would produce the best and most reliable classification results.

4.2.2.1 Models' Complexity In this sub-section, we compare the effectiveness of our suggested CNNs based on the complexity of each model (i.e., number of trainable parameters). Figure 12 shows that the testing accuracy of these models increases by increasing their complexity. The size of each ball corresponds to the model complexity.

4.2.3 Performance Evaluation of the Classification

For an evaluation of the performance of previously tested models, let us recall some performance indicators such as accuracy, to provide a comparison of the third dataset for brain tumor classification. We used accuracy as our primary

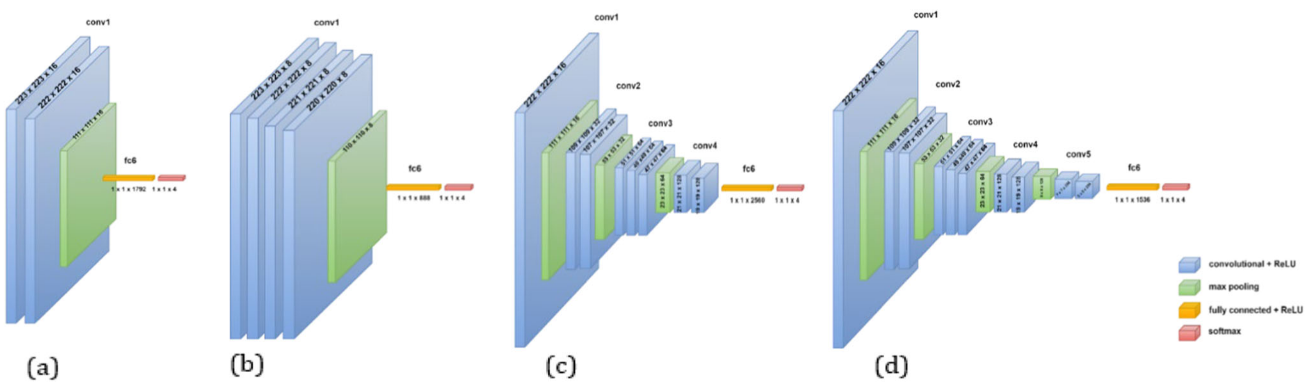


Fig. 10 Customized CNN-based architectures. **a** First architecture, **b** second architecture, **c** third architecture, **d** fourth architecture

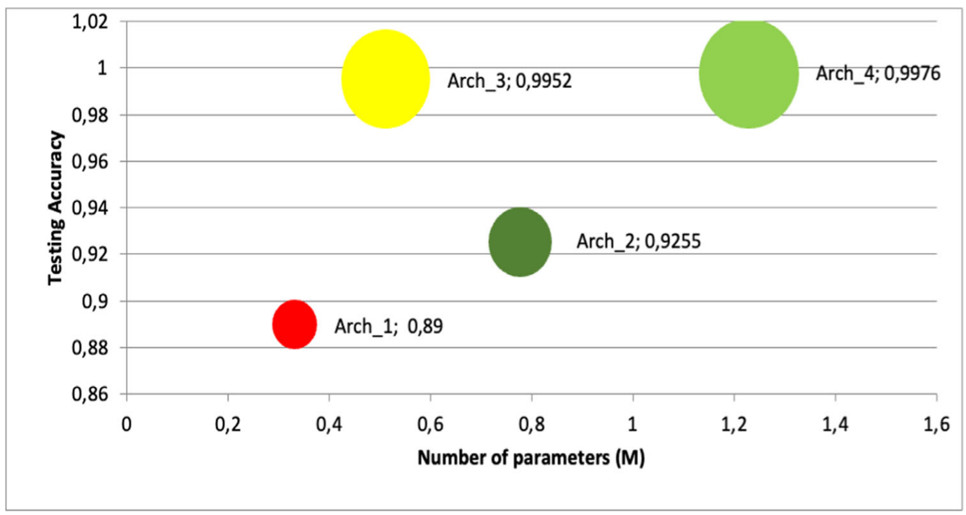


Fig. 11 Ball chart reporting the testing accuracy vs models complexity

performance metric because it is widely used in various fields of applied research to measure the efficiency of image or object classification or detection and recognition processes. This metric is appropriate for evaluating the performance of the proposed solution to the classification problem presented in this work. However, we acknowledge that there are some drawbacks to using accuracy as the sole metric, and our primary goal was to make a fair comparison between the different models. By highlighting the performance of VGG16, VGG19, ResNet50, ResNet152V2, and the best proposed CNN-based model, which is our fourth architecture (model 4), a study of the accuracy obtained with the third dataset of the brain tumor classification is carried out. Table 7 summarizes this and shows that the best accuracy is obtained with the proposed CNN-based model. In addition, to measure the quality of the prediction of this model, a classification report and a confusion matrix are shown in Table 8 and Fig. 12, summarizing the performance of our model through the main metrics used to quantify the efficiency of the classification process is provided.

5 Comparison with the State of the Art

To compare the performance of our proposed CNN-based method of classifying reliably and accurately brain tumor images, against those of the published SOTA methods, it is important to point out here that the input images fed into our proposed system are different to those fed into SOTA methods used in our comparison study. It is worth noticing that though the input images used are intrinsically different, they present the same classification challenges to the system as they share similar features. One may view these images as simply being different members of similar classes of brain tumor images. Therefore, this does not reduce the usefulness of the comparative study carried out in this work. Table 9 presents a comprehensive comparison of the results obtained from various SOTA methods. The classification techniques employed in these methods are primarily binary or multi-class processes, with CNNs being the predominant choice, including both transfer learning and customized CNN-based architectures. The results, focusing on classification accuracy as the metric, clearly highlight the superior performance of the proposed method over the previous works reported in the

Table 3 Details of the key elements of the proposed CNN-based architecture

Models		Conv layer		Max pooling layer		Activation
		N. kernel	Kernel size	Pool size	strid	
Model 1	Conv1	16	3 × 3	2 × 2	2	softmax
		16	3 × 3			softmax
Model 2	Conv1	8	2 × 2	2 × 2	2	softmax
		8	2 × 2			softmax
		8	2 × 2			softmax
		8	2 × 2			softmax
Model 3	Conv 1	16	3 × 3	2 × 2	2	softmax
	Conv 2	32	3 × 3	2 × 2	2	softmax
		32	3 × 3			softmax
	Conv 3	64	3 × 3	2 × 2	2	softmax
		64	3 × 3			softmax
		64	3 × 3			softmax
	Conv4	128	3 × 3	2 × 2	2	softmax
		128	3 × 3			softmax
Model 4	Conv 1	16	3 × 3	2 × 2	2	softmax
	Conv 2	32	3 × 3	2 × 2	2	softmax
		32	3 × 3			softmax
	Conv 3	64	3 × 3	2 × 2	2	softmax
		64	3 × 3			softmax
		64	3 × 3			softmax
	Conv4	128	3 × 3	2 × 2	2	softmax
		128	3 × 3			softmax
	Conv5	256	3 × 3	2 × 2	2	softmax
		256	3 × 3			softmax

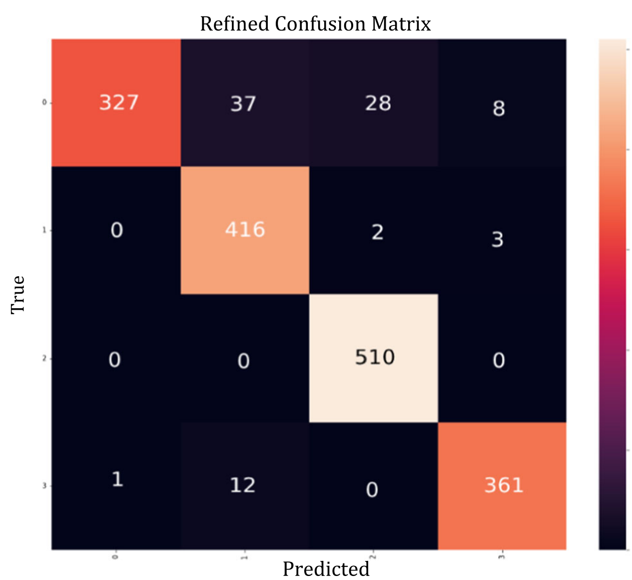


Fig. 12 Confusion Matrix of the 4th CNN-based model

literature. Table 9 serves as a valuable reference for showcasing the advancements achieved by the proposed method in comparison with existing approaches.

6 Conclusion

In this paper, we proposed a new automated non-invasive CNN-based classification system for early brain tumor diagnosis using MRI images. The system can be used to enhance patient care by providing quick early diagnosis, assisting radiologists in their evaluations, prioritizing cases, aiding in medical research, and serving as an educational resource for healthcare professionals. The study carried out and presented in this paper is organized into two main parts. In the first part, we used pre-trained models like VGG16, VGG19, ResNet50, and ResNet152V2 which were all trained with three different datasets and produced different results, with notable differences. The best results were achieved by the model ResNet152V2 when fed with the third dataset. In the second part, we proposed four different CNN architectures to

Table 5 Performance of transfer learning models for the 3 datasets used

Parameters			Evaluation metrics			
		Epoch	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Dataset 1	VGG16	15	76.40	81.13	75.97	71.86
		30	80.63	82.11	76.40	71.37
	VGG19	15	68.27	77.11	67.67	65.01
		30	**75.42**	75.35	68.27	65.65
	ResNet50	15	47.72	52.17	45.43	39.92
		30	56.67	53.32	45.63	40.87
	ResNet152V2	15	79.70	86.41	79.68	75.73
		30	82.56	86.56	79.70	76.53
Dataset 2	VGG16	15	92.89	94.67	94.66	94.52
		30	**93.04**	95.10	94.74	94.78
	VGG19	15	91.85	92.85	91.25	91.34
		30	93.75	92.96	91.85	91.61
	ResNet50	15	63.84	73.86	60.77	56.19
		30	70.57	74.38	63.84	58.85
	ResNet152V2	15	93.55	94.07	93.08	93.21
		30	94.03	94.20	93.55	93.40
Dataset 3	VGG16	15	96.26	96.32	95.93	95.94
		30	96.82	96.54	96.26	96.23
	VGG19	15	94.81	94.58	94.40	94.41
		30	95.49	94.83	94.81	94.75
	ResNet50	15	73.54	78.54	73.98	72.89
		30	75.95	78.79	75.25	73.55
	ResNet152V2	15	97.25	98.04	98.05	98.02
		30	98.17	98.20	98.17	98.16

Table 6 Customized CNN models performances (Dataset 3)

	Epoch(s)	Evaluation msetrics			
		Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Model 1	15	83.28	82.45	80.74	81.89
	30	**89**	**90.20**	**89.32**	**88.96**
Model 2	15	91.35	91.25	91.43	90.82
	30	**92.55**	**91.58**	**92.11**	**92.63**
Model 3	15	99.04	98.54	98.80	98.71
	30	**99.52**	**99.25**	**99.48**	**99.31**
Model 4	15	99.40	99.28	99.22	99.35
	30	99.76	99.64	99.62	99.64

carry out the main brain tumor classification task. Our extensive performance study showed that our proposed automated CNN-based classification system outperformed, in terms of accuracy, all of the state-of-the-art methods used in our study. It is also worth noting that one of the lessons to be learned from this study, through the analysis of the complexity of the different architectures considered is that achieving a high level of accuracy comes at a price in terms of increased complexity. Recall here that, even though the datasets used in our comparison study were different from those used in the state-of-the-art methods, our datasets were just as rich in discriminative features as they were challenging to our proposed classification systems. Because of the statistical and time-evolution nature of the images used and the crucial importance of the decision-making process which would dictate the ensuing therapy to be followed by the patients, we recommend here that expert radiologists further validate our results to ensure the safety and reliability of the medical advice to be derived from the results of our proposed automated brain tumor classification systems. As part of our future endeavors, we aim to further enhance our model by

Table 7 Performances evaluation on the third dataset

	Evaluation metrics				
	Epoch	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
VGG16	15	96.26	96.32	95.93	95.94
	30	96.82	96.54	96.26	96.23
VGG19	15	94.81	94.58	94.40	94.41
	30	95.49	94.83	94.81	94.75
ResNet50	15	73.54	78.54	73.98	72.89
	30	75.95	7s8.79	75.25	73.55
ResNet152V2	15	97.25	98.04	98.05	98.02
	30	**98.17**	98.20	98.17	98.16
Our model	15	99,40	99.28	99.22	99.35
	30	**99,76**	**99.64**	**99.62**	**99.64**

Table 8 Classification report of the 4th CNN-based model

	Precision (%)	Recall (%)	F1-Score (%)	Support
Glioma	99.79	100	99.88	400
Meningioma	99.82	99.48	99 s.74	421
No Tumor	99.50	99	99.25	510
pituitary	99.08	100	99.54	374
Accuracy (%)			**99.76**	**1705**
Macro Average (%)	99.77	99.62	99.60	1705
Weighted Average (%)	99.71	99.57	99.58	1705

Table 9 Comparison with the stat of the art

References	Technique	Type of classification	Dataset	Accuracy (%)
P. Afshar et al. 2019 [58]	CNN	Multi	T1-weighted CE-MRI, 708 meningioma's, 1426 gliomas, and 930 pituitary tumors	90.89
Amin Kabir A et al. 2019[59]	CNN	Binary	TCIA (REMBRANDT)	94.2
Hossam H. Sultan et al. 2019 [21]	CNN	Multi	T1-weighted CE-MRI, 708 meningioma's, 1426 gliomas, and 930 pituitary tumors	96.13
S. Deepak et al. 2019 [32]	CNN (Google Net), KNN, SVM	Multi	Figshar	98
Y. Guan et al. 2021[60]	CNN	Multi	T1-weighted CE-MRI, 708 meningioma's, 1426 gliomas, and 930 pituitary tumors	98.04
M. A. Ansari et al. 2020 [61]	DWT + PCA + GLCM + SVM	Binary	T1-weighted CE-MRI, 140 tumor affected, 60 normal	98.91
Muhammad. A et al. 2022 [62]	CNN	Multi	T1-weighted CE-MRI, 708 meningioma's, 1426 gliomas, and 930 pituitary tumors	98.95
Proposed model	CNN	Multi	T1-weighed CE-MRI, 2774 meningioma's, 2772 gliomas, and 2874 pituitary tumors, 2700 no tumor	**99.76**

launching an online platform dedicated to medical professionals and hospitals. This initiative is poised to significantly benefit our model's generalization by evaluating its performance across a wider range of MRI-related pathologies. Through this platform, medical doctors and healthcare institutions will have access to our model, enabling them to utilize its capabilities for accurate and efficient diagnosis of various medical conditions. Finally, oversight remains crucial to address potential errors and biases and ensure responsible integration into healthcare practices. Collaboration between AI developers and medical experts is essential to align these models with clinical standards and ethical considerations.

Funding SB190112.

References

1. H. Bentahar, M. Djerioui, and N. Bouaziz: An Efficient Classification System for Brain Tumor Based on Convolutional Neural Network. In: *Proceedings of the 2022 International Conference of Advanced Technology in Electronic and Electrical Engineering, ICATEEE 2022*, 2022.
2. Khazaei, Z., et al.: The association between incidence and mortality of brain cancer and human development index (HDI): an ecological study. BMC Pub. Health (2020). https://doi.org/10.1186/S12889-020-09838-4
3. "Brain Tumor Facts." Accessed: Dec. 13, 2022. [Online]. Available: https://braintumor.org/brain-tumors/about-brain-tumors/brain-tumor-facts/
4. "Brain Tumor: Statistics | Cancer.Net." Accessed: Dec. 13, 2022. [Online]. Available: https://www.cancer.net/cancer-types/brain-tumor/statistics
5. Escal, L., et al.: MRI-based score to predict surgical difficulty in patients with rectal cancer. Br. J. Surg. **105**(1), 140–146 (2018). https://doi.org/10.1002/BJS.10642
6. Zhu, G.; Jiang, B.; Tong, L.; Xie, Y.; Zaharchuk, G.; Wintermark, M.: Applications of deep learning to neuro-imaging techniques. Front. Neurol. **10**, 869 (2019). https://doi.org/10.3389/FNEUR.2019.00869/BIBTEX
7. van Beek, E.J.; Kuhl, C.; Anzai, Y.; Desmond, P.; Ehman, R.L.; Gong, Q.; Gold, G.; Gulani, V.; Hall-Craggs, M.; Leiner, T.; Lim, C.T.: Value of MRI in medicine: more than just another test? J. Magn. Reson. Imagin. **49**(7), e14-25 (2019)
8. Tandel, G.S., et al.: A review on a deep learning perspective in brain cancer classification. Cancers (Basel) **11**(1), 111 (2019)
9. Angermueller, C.; Pärnamaa, T.; Parts, L.; Stegle, O.: Deep learning for computational biology. Mol. Syst. Biol. **12**(7), 878 (2016)
10. Qayyum, A.; Anwar, S.M.; Awais, M.; Majid, M.: Medical image retrieval using deep convolutional neural network. Neurocomputing **266**, 8–20 (2017)
11. Xiao, C.; Choi, E.; Sun, J.: Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. J. Am. Med. Inform. Assoc. **25**(10), 1419–1428 (2018)
12. N. F. Aurna, M. A. Yousuf, and K. A. Taher, "Multi-classification of brain tumors via feature level ensemble of convolutional neural networks," *2021 3rd International Conference on Sustainable Technologies for Industry 4.0, STI 2021*, 2021,
13. Aurna, N.F.; Yousuf, M.A.; Taher, K.A.; Azad, A.K.M.; Moni, M.A.: A classification of MRI brain tumor based on two stage feature level ensemble of deep CNN models. Comput. Biol. Med. **146**, 105539 (2022)
14. Sharif, M.I.; Khan, M.A.; Alhussein, M.; Aurangzeb, K.; Raza, M.: A decision support system for multimodal brain tumor classification using deep learning. Complex Intell. Syst. **8**(4), 3007–3020 (2022)
15. Mohsen, H.; El-Dahshan, E.-S.A.; El-Horbaty, E.-S.M.; Salem, A.-B.M.: Classification using deep learning neural networks for brain tumors. Future Comput. Inform. J. **3**(1), 68–71 (2018)
16. Trebeschi, S.; van Griethuysen, J.: Deep learning for fully-automated localization and segmentation of rectal cancer on multiparametric MR. Sci. Rep. (2022). https://doi.org/10.1038/s41598-017-05728-9
17. Betancur, J., et al.: Deep learning for prediction of obstructive disease from fast myocardial perfusion SPECT: a multicenter study. JACC Cardiovasc. Imagin. **11**(11), 1654–1663 (2018)
18. J. Lee, S. Jun, Y. Cho, H. L.-K. Journal of, and undefined 2017, "Deep learning in medical imaging: general overview," *synapse.koreamed.org*, Accessed: Dec. 13, 2022. [Online].
19. Litjens, G., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)
20. Abiwinanda, N.; Hanif, M.; Hesaputra, S.T.; Handayani, A.; Mengko, T.R.: Brain tumor classification using convolutional neural network. IFMBE Proc **68**(1), 183–189 (2019)
21. Sultan, H.H.; Salem, N.M.; Al-Atabany, W.: Multi-classification of brain tumor images using deep neural network. IEEE Access **7**, 69215–69225 (2019)
22. Aderghal, K.; Afdel, K.; Benois-Pineau, J.; Catheline, G.: Improving Alzheimer's stage categorization with convolutional neural network using transfer learning and different magnetic resonance imaging modalities. Heliyon **6**(12), e05652 (2020)
23. Niu, S.; Liu, Y.; Wang, J.; Song, H.: A decade survey of transfer learning (2010–2020). IEEE Trans. Artif Intell. **1**(2), 151–166 (2020)
24. Lu, Y.; Wang, W.; Yuan, C.; Li, X.; Lai, Z.: Manifold transfer learning via discriminant regression analysis. IEEE Trans. Multimed. **23**, 2056–2070 (2021)
25. Shin, H.-C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M.: Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imagin. **35**(5), 1285–1298 (2016)
26. Niu, S.; Liu, M.; Liu, Y.; Wang, J.; Song, H.: Distant domain transfer learning for medical imaging. IEEE J. Biomed. Health Inform. **25**(10), 3784–3793 (2021)
27. Fan, X.; Cao, P.; Shi, P.; Chen, X.; Zhou, X.; Gong, Q.: An underwater dam crack image segmentation method based on multi-level adversarial transfer learning. Neurocomputing **505**, 19–29 (2022)
28. Wu, R.; Xin, Y.; Qian, J.; Dong, Y.: A multi-scale interactive U-Net for pulmonary vessel segmentation method based on transfer learning. Biomed. Signal Process. Control **80**, 104407 (2023)
29. Kang, J.; Ullah, Z.; Gwak, J.: MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers. Sensors (Basel) **21**(6), 1–21 (2021)
30. Noreen, N.; Palaniappan, S.; Qayyum, A.; Ahmad, I.; Imran, M.; Shoaib, M.: A deep learning model based on concatenation approach for the diagnosis of brain tumor. IEEE Access **8**, 55135–55144 (2020)
31. Ayadi, W.; Elhamzi, W.; Charfi, I.; Atri, M.: Deep CNN for Brain tumor classification. Neural. Process. Lett. **53**(1), 671–700 (2021)
32. Deepak, S.; Ameer, P.M.: Brain tumor classification using deep CNN features via transfer learning. Comput. Biol. Med. **111**, 103345 (2019)
33. Toğaçar, M.; Cömert, Z.; Ergen, B.: Classification of brain MRI using hyper column technique with convolutional neural network

and feature selection method. Expert Syst. Appl. **149**, 113274 (2020)

34. Ghassemi, N.; Shoeibi, A.; Rouhani, M.: Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images. Biomed. Signal Process. Control **57**, 101678 (2020)
35. S. Kumar, D. M.-B. and B. Engineering, and undefined 2020, "Optimization driven deep convolution neural network for brain tumor classification," *Elsevier*, Accessed: Dec. 02, 2022. [Online].
36. Anilkumar, B.; Rajesh Kumar, P.: Tumor classification using block wise fine tuning and transfer learning of deep neural network and KNN classifier on MR brain images. Int. J. Emerg. Trends Eng. Res. **8**(2), 574–583 (2020)
37. Khairandish, M.O.; Sharma, M.; Jain, V.; Chatterjee, J.M.; Jhanjhi, N.Z.: A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images. IRBM **43**(4), 290–299 (2022)
38. Uyar, K.; Taşdemir, Ş; Ülker, E.; Öztürk, M.; Kasap, H.: Multi-Class brain normality and abnormality diagnosis using modified Faster R-CNN. Int. J. Med. Inform. **155**, 104576 (2021)
39. Tandel, G.S.; Tiwari, A.; Kakde, O.G.: Performance enhancement of MRI-based brain tumor classification using suitable segmentation method and deep learning-based ensemble algorithm. Biomed. Signal Process. Control **78**, 104s018 (2022)
40. Sindhiya Devi, R.; Perumal, B.; Pallikonda Rajasekaran, M.: A hybrid deep learning-based brain tumor classification and segmentation by stationary wavelet packet transform and adaptive kernel fuzzy c means clustering. Adv. Eng. Software **170**, 103146 (2022)
41. Abd-Ellah, M.K.; Awad, A.I.; Khalaf, A.A.M.; Hamed, H.F.A.: Two-phase multi-model automatic brain tumor diagnosis system from magnetic resonance images using convolutional neural networks. EURASIP J. Image Video Process **2018**(1), 1–10 (2018)
42. J. Cheng, Brain tumor dataset. figshare. dataset. https://doi.org/10.6084/m9.figshare.1512427.v5. Accessed 30 May 2018
43. "Brain Tumor MRI Image Classification Dataset | Kaggle." Accessed: May 19, 2022. [Online].
44. "Brain Tumor Classification (MRI) | Kaggle." Accessed: Oct. 03, 2022. [Online].
45. Le Cun, Y.A.; Bottou, L.; Orr, G.B.; Müller, K.R.: Efficient backprop. Lect. Notes Comput. Sci. **7700**, 9 (2012)
46. Shorten, C.; Khoshgoftaar, T.M.: A survey on Image Data Augmentation for Deep Learning. J Big Data **6**(1), 1–48 (2019)
47. C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11141 LNCS, pp. 270–279, 2018.
48. Pan, S.J.; Yang, Q.: A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**(10), 1345–1359 (2010)
49. P. Goldsborough, "A Tour of TensorFlow," Oct. 2016, https://doi.org/10.48550/arxiv.1610.01178.
50. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014 Sep 4..
51. IWARI and R, "Transfer Learning — Part — 4.0!! VGG-16 and VGG-19 | by RAVI SHEKHAR TIWARI | Becoming Human: Artificial Intelligence Magazine." Accessed: May 15, 2022. [Online].
52. Boesch and G, "Deep Residual Networks (ResNet, ResNet50) - Guide in 2022 - viso.ai." Accessed: May 15, 2022. [Online].
53. He K, Zhang X, Ren S, Sun J. "Deep residual learning for image recognition", in Proceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 770–778).
54. Hossin, M.; Sulaiman, M.N.: A review on evaluation metrics for data classification evaluations. Int. J. Data Min. Knowl. Manage. Process **5**(2), 01–11 (2015)
55. Ferdinandy, B.; Gerencsér, L.; Corrieri, L.; Perez, P.; Újváry, D.; Csizmadia, G.; Miklósi, Á.: Challenges of machine learning model validation using correlated behaviour data: Evaluation of cross-validation strategies and accuracy measures. PLoS ONE **15**(7), e0236092 (2020)
56. Wong, T.T.; Yeh, P.Y.: Reliable accuracy estimates from k-fold cross validation. IEEE Trans. Knowl. Data Eng. **32**(8), 1586–1594 (2019)
57. C. Luo, X. Li, L. Wang, J. He, D. Li, and J. Zhou, How does the data set affect CNN-based image classification performance? *2018 5th International Conference on Systems and Informatics, ICSAI 2018*, pp. 361–366, Jan. 2019.
58. P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries," ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, vol. 2019-May, pp. 1368–1372, May 2019, https://doi.org/10.1109/ICASSP.2019.8683759.
59. Kabir Anaraki, A.; Ayati, M.; Kazemi, F.: Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. Biocybern. Biomed. Eng. **39**(1), 63–74 (2019). https://doi.org/10.1016/J.BBE.2018.10.004
60. Guan, Y., et al.: A framework for efficient brain tumor classification using MRI images. Math. Biosci. Eng. **18**(5), 5790–5815 (2021). https://doi.org/10.3934/MBE.2021292
61. Ansari, M.A.; Mehrotra, R.; Agrawal, R.: Detection and classification of brain tumor in MRI images using wavelet transform and support vector machine. J. Interdiscip. Math. **23**(5), 955–966 (2020)
62. Aamir, M., et al.: A deep learning approach for brain tumor classification using MRI images. Comput. Electr. Eng. **101**, 108105 (2022). https://doi.org/10.1016/J.COMPELECENG.2022.108105