IJEMS
The International Journal for Engineering and Modern Science
Year: 2024 Volume: 2 Article ID: **2023006**
MODERN ACADEMY FOR ENGINEERING & TECHNOLOGY

Brain Tumors MRI Classification Through CNN Transfer Learning Models - An Overview

Hend Fathy Khalil [1,*], Eman Mohammed Mahmoud [1], Ashraf M. Mahrous [2], Hesham F. A. Hamed [3,4]

1Communication and Electronics Department, Modern Academy for Engineering and Technology, Cairo, Egypt
2 Biomedical Department, Faculty of Engineering, Minia University, Minia, Egypt
3 Faculty of Engineering, Minia University, Minia, Egypt
4 Faculty of Engineering, Egyptian-Russian University, Cairo, Egypt

ARTICLE INFO.

Article history:
Received 3 September 2023
Revised 31 December 2023
Accepted 1 January 2024
Available online 2024

Keywords:

Brain Tumors
Deep Neural Network
CNN transfer learning models

ABSTRACT

One of the most prevalent methods employed in medical research involves identifying brain tumors and monitoring their growth through brain MRI scans. By examining the internal structure of the human brain, valuable insights regarding tumor development can be obtained. However, manually detecting brain tumors from MRI scans poses a significant challenge within the medical research field, as tumors can lead to substantial alterations in both the internal and external brain structure. To address this issue, it is suggested to explore recent classifier approaches for the detection of brain tumors in MRI images. By utilizing these advanced techniques, the performance and analysis of brain tumor growth can be described, enabling the identification of general symptoms and facilitating a targeted diagnosis for an effective treatment plan. This discussion encompasses various classification approaches derived from existing research papers, ultimately leading to conclusive findings on brain tumor detection from MRI scans.

1. Introduction

Abnormal cell growths in the brain or surrounding tissues are known as brain tumors. Brain tumors can be classified as malignant (cancerous) or benign (non-cancerous), with the latter being more aggressive and perhaps fatal[1][2]. The classification system established by the World Health Organization (WHO) categorizes brain tumors into several primary groups. Gliomas are one of the most common forms; they are derived from glial cells. Gliomas can be further subdivided into ependymomas, oligodendrogliomas, and astrocytomas based on the particular glial cell that is implicated and the tumor's histological features. Another category of brain tumors is meningiomas, which develop from the meninges, the protective membranes enveloping the brain and spinal cord. Meningiomas are typically benign tumors characterized by a slow growth rate. Furthermore, pituitary tumors are identified by the WHO classification scheme. These tumors originate in the pituitary gland, which is located near the base of the brain. The pituitary gland

---

*** Corresponding author**

E-mail address: eng.hend2025@gmail.com

is responsible for regulating hormone production. Pituitary tumors can disrupt hormonal balance and manifest in various symptoms[3].

Brain imaging techniques refer to the various methods used to visualize the brain's structure, function, and connectivity. These techniques allow researchers and clinicians to study the brain and diagnose neurological conditions. Magnetic resonance imaging (MRI), computed tomography (CT), electroencephalography (EEG), positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) are a few frequently utilized brain imaging modalities[4]. Among these techniques, MRI is considered one of the most versatile and widely used methods due to its exceptional anatomical detail, non-invasiveness, ability to capture functional and physiological data, and superior soft tissue contrast[5].

In order to interpret radiological and pathological pictures, sophisticated software systems called computer-aided detection and diagnosis (CAD) combine computer vision and artificial intelligence approaches. To help radiologists correctly diagnose a range of disorders in distinct anatomical locations, these cutting-edge technologies have been developed[6].

Machine learning has significantly accelerated the advancement of CAD systems. In recent times, machine learning has been applied to classify objects of interest, such as lesions, by leveraging input features. Machine learning enables the discovery and learning of informative features that effectively capture patterns and regularities in data. Unlike traditional approaches, where human experts design features based on domain knowledge, machine learning offers the ability to automate feature extraction. However, it should be noted that the complexity of living organisms far exceeds the superficial linear relationships detectable by traditional machine learning methods[7]. This highlights the need for more sophisticated approaches to uncover the intricate biology underlying disease detection and diagnosis.

Deep learning, inspired by the neural networks in the human brain, has revolutionized various challenges by integrating feature extraction and selection into the training process. Deep learning models consist of multiple layers, where each layer combines elements from the previous layer using weighted sums. This layered structure enables deep learning models to effectively capture complex mapping functions, surpassing traditional machine learning methods. With minimal human intervention, deep learning has consistently outperformed alternative approaches and has gained widespread adoption in medical image analysis for a range of tasks[8]. It offers unprecedented potential for advanced analysis and diagnosis in the medical field.

Convolutional neural networks, or CNNs, are a kind of deep learning architecture that are frequently applied to tasks involving images[9]. CNNs use several layers of connected nodes to automatically learn and extract pertinent characteristics from images. They utilize convolutional layers to apply filters and extract patterns, followed by pooling layers to down sample the extracted features. Finally, fully connected layers are used for classification or regression tasks. Conversely, transfer learning is a method that uses large-scale datasets to make use of pre-trained CNN models. Instead of training a CNN from scratch, transfer learning allows us to take the learned feature representations from a pre-trained model and apply them to a new task or domain

In this research , An overview of the use of Convolutional Neural Network (CNN) transfer learning models for the classification of MRI images of brain tumors is provided. CNNs have demonstrated amazing effectiveness in a variety of computer vision tasks, including medical image analysis, thanks to the quick developments in deep learning. We investigate the application of transfer learning, a method that makes use of CNN models that have already been trained, to the categorization of brain tumor MRIs. By fine-tuning pre-trained CNNs on a dataset of brain tumor MRI images, we demonstrate the effectiveness of transfer learning in extracting relevant features and improving classification accuracy. Additionally, we discuss the benefits of transfer learning, such as reduced training time and the ability to handle limited datasets. In this overview, we demonstrate CNN transfer learning's potential as a useful tool to help radiologists and other medical professionals correctly diagnose brain cancers from MRI scans.

2. Material and method

The fundamental difference between a traditional classifier and a deep Convolutional Neural Network (CNN) lies in their approach to feature extraction and pattern recognition. Traditional classifiers rely on manually engineered features, which are typically designed by domain experts[10]. These handcrafted features are based on prior knowledge and understanding of the problem domain. However, deep CNNs are made to automatically identify and extract pertinent features from unprocessed input, especially from material that resembles a grid, like pictures. Convolutional, pooling, and fully connected layers are just a few of the layers that CNNs use to learn hierarchical feature representations. Deep CNNs are able to automatically extract non-trivial features from the data, including spatial correlations and minute details that would be difficult for conventional classifiers to identify. As a result, deep CNNs are particularly effective in image-related tasks and have demonstrated superior performance in various domains specially in medical domain.

This paper provides an overview of research articles published between 2018 and 2022 that delve into the realm of brain tumor classification This overview focuses on using transfer learning techniques in conjunction with Convolutional Neural Networks (CNNs). Through the use of a transfer learning methodology, the research reviewed in this paper improves CNN models' classification performance in brain tumor analysis by utilizing prior information gleaned from extensive datasets.The paper aims to present a scientific overview of the advancements made in this domain during the specified time frame. This section is formulated as follows.

Figure 1 illustrate the basic structure of CNN.

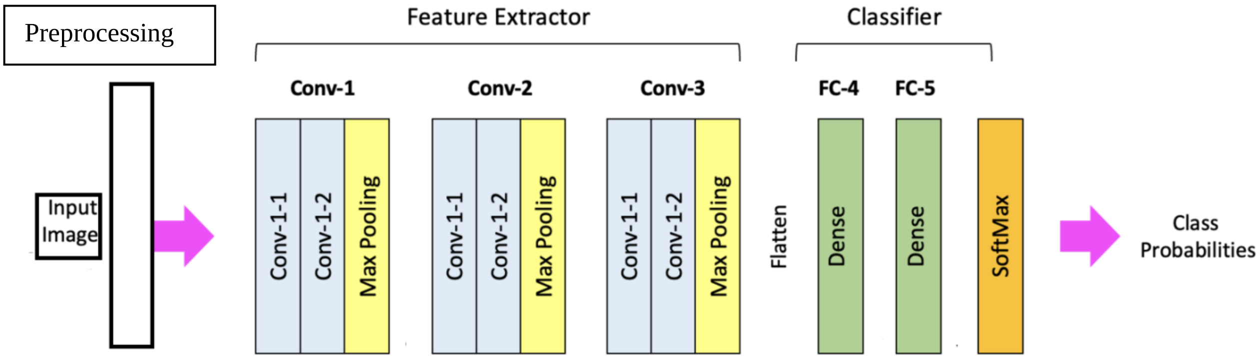

Fig. 1 Basic structure of CNN

The utilization of Convolutional Neural Networks (CNNs) involves a systematic four-step. Firstly, input images are acquired from the dataset, serving as the initial data representation. Secondly, the acquired images undergo preprocessing techniques to enhance their quality and suitability for analysis, utilizing various methods such as normalization, resizing, and color space conversion. Thirdly, a pre-trained CNN model that was trained on a sizable dataset is used to apply CNN transfer learning. This method makes it possible to apply learnt feature representations from the pre-trained model to the current task, which makes training and testing the CNN on the particular dataset more effective. Ultimately, the CNN model's performance is assessed, usually through the use of metrics such as F1 score, accuracy, precision, and recall. The assessment procedure guarantees the robustness and dependability of the outcomes by offering insights into how well the model classifies events that have not yet been observed.

3. Dataset

Datasets can be found on particular websites. The datasets and their URLs are shown in Table 1.

Table 1 URLs for Datasets

Dataset Name	**Available website**
Figshare brain tumor dataset	https ://www.figshare.com/articles/brain tumor dataset
Kaggle brain tumor dataset	https ://www.kaggle.com/brain tumor dataset
BraTS_brain tumor dataset	http://www.smir.ch/BRATS
The Cancer Imaging Archive (TCIA)	https://www.cancerimagingarchive.net

4. Preprosessing

4.1. Normalization

Normalization of brain tumor MRI images involves adjusting the intensity values of the images to a standardized range[11]. This is done to account for variations in intensity caused by different factors such as scanner settings and imaging protocols. Two common methods used for normalization are linear normalization and non-linear normalization. In linear normalization, the intensity values are rescaled to a predefined range, such as 0 to 255, improving the visibility of tumor features. Non-linear normalization methods, such as histogram equalization and histogram matching, redistribute intensity values to enhance contrast and ensure consistency. Normalization improves visual interpretation, facilitates quantitative analysis, and serves as a foundation for further processing steps like segmentation and classification, aiding in tumor characterization and treatment planning.

4.2. Resizing

All of the images must be resized before being entered into CNN classification models in order to comply with the fixed input size requirement of deep neural networks. Resizing ensures that the images conform to the specific dimensions expected by the network. This process involves adjusting the size of the images while maintaining their aspect ratio or by applying cropping or padding techniques to match the desired dimensions. By resizing the images, they become compatible with the network architecture, allowing for efficient processing and ensuring consistent input across different images.

4.3. Bias Field Correction

Unwanted artifacts in medical imaging, such as scan position, instrument used, and unidentified problems, might lead to the bias field. This artifact can seriously impair the efficacy of medical image analysis methods. It appears as fluctuations in brightness across the image. Hence, before employing distorted MRI images for CNN classification models, a preprocessing step is required to rectify the bias field signal. The Statistical Parametric Mapping (SPM) module and the N4 bias field correction algorithm are two frequently used techniques for addressing the intensity inhomogeneity and bias field correction in MR images[12]. A popular technique for addressing the low-frequency intensity non-uniformity in MR image data is the N4 bias field correction algorithm. However, SPM is a complete tool for brain segmentation tasks; it consists of software packages that contain segmentation routines, intensity non-uniformity (bias) correction, and skull stripping. With the SPM module or the N4 bias field correction algorithm, the inhomogeneity in the intensity of MR images can be effectively adjusted. This preprocessing step enhances the accuracy and reliability of subsequent analysis tasks, including classification using CNN models. Correcting the bias field artifact ensures that the image data used for classification is more standardized and reliable, ultimately improving the performance of the CNN classification model.

4.4. Skull stripping

In brain tumor classification problems, distinguishing between the tumor and the skull can be challenging due to the similarity in intensity between them. To overcome this issue and improve the performance of classification models, skull stripping is commonly employed as a preprocessing step[13]. Skull stripping involves removing the surrounding skull region from the brain image, isolating the tumor and enhancing its visibility. By eliminating the interference from the skull, the classification model can focus more accurately on the tumor region, leading to improved performance and more reliable classification results. Skull stripping is an essential technique that helps to address the intensity overlap between the tumor and the skull, enabling better differentiation and enhancing the overall effectiveness of brain tumor classification systems.

4.5. Data Augmentation

Large datasets are important for CNN-based classification tasks. To achieve good problem generalization, a common recommendation recommends having about ten times as many samples as the network's parameters. Overfitting is a possibility in cases where the database is small. Data augmentation approaches are commonly used, especially in brain tumor classification research, to solve challenges of unbalanced distribution and data shortage. Applying geometric adjustments to the current dataset is known as data augmentation. Rotation, scaling, translation, mirroring, and cropping are some of these transformations. By performing these operations, the dataset is expanded, creating variations of the original samples. This augmentation helps to diversify the dataset, making it more robust and reducing the risk of over fitting. Data augmentation enhances the model's capacity to learn various features and generalize effectively to previously unseen data by purposefully growing the sample size. [14].

In the context of brain tumor classification, data augmentation techniques have been extensively utilized to address the limited availability of labeled data. By applying geometric transformations, the augmented dataset provides additional training samples, enabling the CNN model to learn more effectively and improve its classification performance.

5. Convolutional Neural Network (CNN)

Specifically created for processing organized grid-like input, like images or sequences, Convolutional Neural Networks (CNNs) are a form of deep learning model. picture categorization, object identification, and picture segmentation are just a few of the computer vision tasks that frequently use it.

The convolutional layer is a crucial component of a CNN. Convolutional layers are made up of several filters or kernels that slide over the input data and extract local characteristics by summing and multiplying the elements at a time. These filters aid in identifying various aspects and patterns in the supplied data. Multiple convolutional layers are stacked to enable the network to learn ever-more-abstract and complex features. Figure 2 illustrate the convolutional layer of CNN.

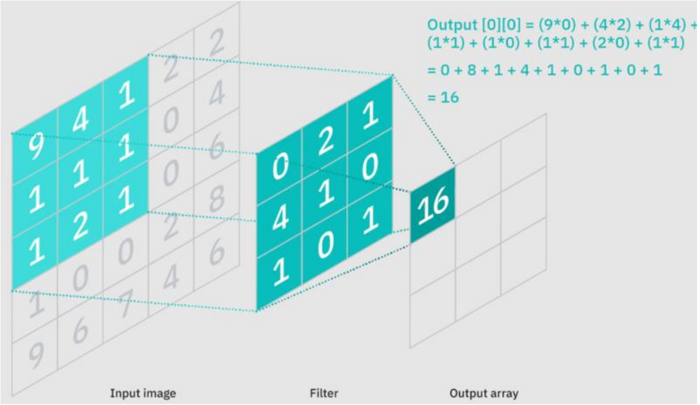

Fig.2 The convolutional layer of CNN

CNNs frequently have pooling layers after the convolutional layers. By downsampling, pooling layers (usually by max pooling or average pooling) minimize the spatial dimensions of the features. By lowering computational complexity and strengthening the network's resistance to input fluctuations, pooling aids in the extraction of the most pertinent and instructive features. Fig.3 illustrate the pooling layer of CNN.

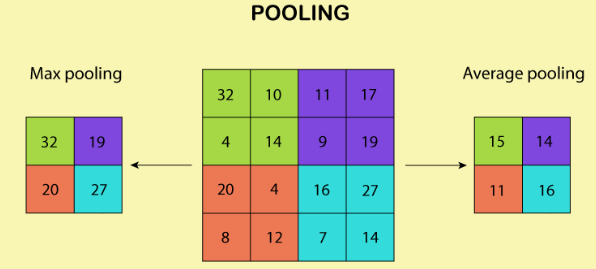

Fig. 3 The pooling layer of CNN

CNNs have fully connected layers in addition to convolutional and pooling layers. Like a standard neural network, these layers link every neuron in the preceding layer to every neuron in the following layer. The network can learn high-level representations and generate predictions using the retrieved information thanks to fully linked layers.

CNNs usually end with an output layer, which can have different configurations depending on the task at hand. The output layer for image classification frequently comprises of softmax activation, which provides class probabilities for every potential category [15]. In other tasks like object detection or segmentation, the output layer may involve different activation functions and structures to produce the desired output format.

In machine learning, transfer learning is the process of applying knowledge from addressing one problem to another that is related yet unrelated. Transfer learning in the context of Convolutional Neural Networks (CNNs) is the use of pre-trained models that were trained on a large-scale dataset such as ImageNet. ImageNet is an image classification dataset with millions of labeled images spanning a thousand different classes. These models have been trained on ImageNet to learn general visual representations, capturing features like edges, shapes, and textures. As a result, these pre-trained models have developed a strong ability to extract meaningful features from images. When using ImageNet models for transfer learning, the previously trained layers are often kept and utilized as a feature extractor. The final fully connected layers of the original model, which are responsible for classifying images into ImageNet classes, are often replaced with new layers suited to the target task. These new layers are trained using a smaller, task-specific dataset.

Transfer learning models of Convolutional Neural Networks (CNNs) offer several advantages, including leveraging learned visual representations, addressing limited labeled data challenges, improving convergence and generalization, and facilitating the transfer of domain-specific knowledge. By starting with pre-trained models, we avoid the need to train the entire network from scratch, saving computational time and reducing the overall training burden. The computational time in transfer learning is mainly spent on fine-tuning the task-specific layers, while the pre-trained layers, acting as feature extractors, require minimal computation. This approach speeds up the training process and contributes to the efficiency of using transfer learning models in CNNs.

VGG16, VGG19, ResNet, Inception models, MobileNet, and DenseNet are a transfer learning models that are based on convolutional neural networks (CNNs).

5.1. VGG Transfer learning model

Deep convolutional neural network designs VGG16 and VGG19 were created by the University of Oxford's Visual Geometry Group (VGG). They are well-known for being straightforward and efficient when it comes to classifying images. Convolutional, pooling, and fully linked layers are among the many layers that make up both models.

There are 16 layers in all in VGG16: 3 fully linked layers and 13 convolutional layers [16]. Mostly, 3x3 filters with a stride of 1 and padding of 1 are used in the convolutional layers of VGG16. After these layers come max pooling layers with a stride of two and 2x2 filters. At the end of the network, the completely connected layers carry out the categorization operation. VGG16 normally requires an input image with 224 x 224 pixel size.

VGG19 extends VGG16 by adding four additional convolutional layers, resulting in a total of 19 layers. The additional layers enhance the model's depth and representation capacity. VGG19 employs 2x2 filters with a stride of 2 in the max pooling layers and 3x3 filters with a stride of 1 in the convolutional layers, similar to VGG16. There are no changes to the completely connected layers. Typically, the dimensions of the input image for VGG19 are 224 by 224 pixels.

Both VGG16 and VGG19 are known for their uniform architecture, where the layers are stacked one after another, making them easy to understand and implement. The deeper layers capture more complex and abstract features, while the pooling layers downsample the spatial dimensions, reducing computational complexity.

5.2. ResNet transfer learning model

A popular transfer learning model in the deep learning space is called ResNet, short for Residual Network. By addressing the vanishing gradient issue, it developed the idea of residual connections, which permits the training of even deeper networks.

ResNet architectures come in various depths, such as ResNet50, ResNet101, and ResNet152. These numbers represent the number of layers in the network. For example, ResNet50 has 50 layers, ResNet101 has 101 layers, and so on. The deeper versions of ResNet typically offer improved performance but come at the cost of increased computational complexity[16].

Multiple "residual blocks" serve as the building blocks of ResNet models. Every residual block uses skip connections, sometimes referred to as shortcut connections, and has two or more convolutional layers. These skip connections enable the gradient flow to bypass one or more layers, allowing the network to learn residual mappings. By incorporating these residual connections, ResNet models can effectively train deeper architectures. The input image dimensions for ResNet models are typically 224x224 pixels.

5.3. Mobile Net Transfer learning

MobileNet was created especially for embedded and mobile devices with constrained processing power. It provides a fair balance between accuracy and model size, which makes it ideal for real-time applications on devices with limited computing power.

MobileNet consists of several building blocks called depth-wise separable convolution blocks. Depth-wise convolution, batch normalization, and ReLU activation are the first three steps in each block. Point-wise convolution, additional batch normalization, and ReLU activation come next. These blocks are stacked to form the overall MobileNet architecture. The input image dimensions for MobileNet are typically 224x224 pixels, similar to other popular CNN architectures[17].

Because of their small size, low computing cost, and high accuracy, mobile net models are a good fit for situations with constrained computer resources.

5.4. Inception transfer learning

Inception is another transfer learning model in convolutional neural networks (CNNs) that was introduced by Christian Szegedy et al. in 2014. The Inception architecture is known for its use of multiple filter sizes in parallel, allowing it to capture features at various scales and improve the network's ability to recognize complex patterns.

In Inception, the network consists of multiple "Inception modules" stacked on top of each other. Each module performs different convolutions with different filter sizes, including 1x1, 3x3, and 5x5 filters. These filters are applied to the input in parallel, and their outputs are concatenated to form the module's output. This parallel approach allows the network to capture both local and global features effectively. The input dimension of images for Inception is typically 224x224 pixels. However, variations of the Inception architecture, such as Inception-v3, can also handle larger input sizes like 299x299 pixels. Table 2 illustrates the comparison between transfer learning models.

Table 2 The comparison between transfer learning models.

Model	**Year**	**Number of Layers**	**Parameters (according to ImageNet)**	**Notable Feature**
VGG16	2014	16	138 million	Deep architecture, uniform structure
VGG19	2014	19	144 million	Deeper version of VGG16
MobileNet	2017	28	4.2 million	Depth wise separable convolutions
Residual connections	ResNet	2015	Varies (up to 152)	Varies
Inception	2014	Varies (up to 4)	Varies	Multiple parallel

6. Performance Measurement

In brain tumor classification research, a CNN algorithm's performance evaluation is essential. Evaluation measures like accuracy, precision, sensitivity, F1 score, and area under the curve (AUC) are frequently seen. These measurements offer insightful information about the categorization performance[18]. Accurately classifying an image as positive is known as true positive (TP), while correctly classifying an image as negative is known as true negative (TN). When an image is mistakenly categorized as positive, it is called false positive (FP), and when an image that should be classified as negative is mistakenly labeled as false negative (FN). These measures aid in evaluating how well the CNN algorithm classifies brain tumor images.

6.1. Accuracy

A performance indicator called accuracy is used to assess how well a transfer learning model in CNNs performs in categorization. It calculates the percentage of cases that are correctly classified relative to the total number of instances.

$$\mathrm{ACC}=\frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

6.2. Specificity

It calculates the percentage of accurately identified negative cases relative to the total number of negative cases that actually occur.

$$\mathrm{SPE}=\frac{TN}{TN+FP} \tag{2}$$

6.3. Precision

It calculates the percentage of genuine positives—that is, positively classified instances—among all instances that the model has determined to be positive.

$$\text{PRE}=\frac{TP}{TP+FP} \tag{3}$$

6.4. Sensitivity

Also referred to as true positive rate (TPR) or recall. It calculates the percentage of real positive cases—that is, true positive instances—that the model properly identifies out of all actual positive instances.

$$\text{SEN}=\frac{TP}{TP+FN} \tag{4}$$

6.5. F1 Score

The model's capacity to deliver both high precision (correct positive predictions) and high recall (precise identification of positive examples) is balancedly evaluated by the F1 score. It is particularly helpful when memory and precision are equally critical for the given task or when there is an imbalance between positive and negative examples.

$$\text{F1 Score}=2\frac{PRE*SEN}{PRE+SEN} \tag{5}$$

6.6. Area under the Curve

It gauges how well the model predicts things overall across various classification thresholds. As the classification threshold changes, the AUC measures the trade-off between the true positive rate (sensitivity) and the false positive rate (specificity). A higher AUC value shows that the transfer learning model can reliably distinguish between positive and negative occurrences across a range of thresholds, indicating improved discriminative ability.

7. Overview of using Transfer learning models of CNN

Research publications have been produced in large quantities due to the popularity for CNN-based deep learning techniques. These publications summarize the efficacy of CNN algorithms in brain tumor classification and provide suggestions for future research directions.

Kaur and Gandhi conducted a study focusing on the development of an automated system for brain image classification using the VGG-16 model and transfer learning techniques[19]. The main objective was to assess how well a transfer learning-trained VGG-16 model performed in classifying diseased brain images. In order to achieve this, the final few layers of the original VGG-16 model were altered to fit the particular image categories that were pertinent to the investigation. The assessment stage made use of the AANLIB database, which included a set of brain scans from Harvard Medical School. A selection of 160 axial T2-weighted MR brain pictures was selected from this database. Twenty of these photos were normal cases, and the other 140 were aberrant cases. A 256 × 256 pixel in-plane resolution was used to capture the images.

The suggested approach consisted of a number of consecutive phases. First, in Step 1, the Harvard repository provided the dataset, which was arranged in the "imgData/train" directory and included both normal and aberrant T2-weighted MR images. Then, in Step 2, all subfolders inside the directory were included, suitable labels were assigned based on the folder names, and an image DataStore was created to read the images.

Step 3 involved resizing the input images to make sure they fit the input layer size of the VGG-16 model. In Step 4, the data was divided into training and test sets. Ninety percent of the photos in each category were designated as training images, and the remaining ten percent were designated as test datasets to assess the network's performance.

Step 5 then involved a review procedure for the network design, specifically for the "vgg16" model. In this stage, the final three layers of the pre-trained network were changed or replaced with a predetermined set of layers: a "classification output" layer, a "softmax" layer, and a "fc" layer (fully connected). Based on earlier research findings, this update made it easier to classify images into normal and pathological groups.

Step 6 involved training the modified network using the training dataset. The network learned from the input images and adjusted its parameters to enhance its ability to accurately classify between normal and abnormal brain images. Subsequently, in Step 7, the newly trained classifier was tested using a separate testing dataset to evaluate its performance in a real-world scenario.

Finally, in Step 8, performance metrics were computed and reported to assess the efficacy of the proposed methodology. These metrics covered a range of assessment parameters, including recall, accuracy, precision, F1 score, and others, and they offered insightful information about the network's classification performance on the testing dataset. Notably, the proposed methodology aimed to achieve exceptional performance in terms of sensitivity, specificity, and accuracy, with all metrics set at the ideal value of 100%. Deepak, S., & Ameer, P. M. addresses a classification problem involving three prominent types of brain tumors, namely glioma, meningioma, and pituitary tumors[20]. Using a pre-trained GoogLeNet model, the suggested approach applies the idea of deep transfer learning to extract characteristics from brain MRI images. The collected features are subsequently classified using tried-and-true classifier models. Using an MRI dataset from figshare, the study applies a five-fold cross-validation procedure at the patient level. The publicly available figshare dataset includes 3064 brain MRI scans from 233 patients who were diagnosed with one of three brain malignancies: meningioma, glioma, or pituitary tumors. It is frequently used to assess classification and retrieval methods. Coronal, sagittal, and axial views are included in the T1-CE MRI modality images. The dataset includes 930 images of pituitary tumors, 708 images of meningiomas, and 1426 images of gliomas. The images are sent as .mat files, complete with a size of 512x512 pixels. The MRI images were preprocessed in order to use the GoogLeNet model, which was initially created for RGB color images with an input layer of size 224x224x3. This included scaling the values between 0 and 1 using a min-max normalization technique, normalizing the intensity values, resizing to 224x224 pixels, and creating three channels by replicating the grayscale values three times. Using the figshare dataset, the suggested approach was evaluated using a patient-level five-fold cross-validation method Five roughly equal-sized subsets were created from the 233 patient dataset; one subset was designated as the test set, and the other four as the training set. To make sure that a patient's data did not show up in both the test and training sets, this procedure was repeated. Using the preprocessed training set, the updated GoogLeNet model was trained, and its hyperparameters were heuristically tuned to promote the loss function's convergence during training. With an initial learning rate of 0.0003 set to balance convergence and training time, the Adam optimizer was chosen because of its adaptable nature and good learning rate. The training speed and computing needs were balanced by using a mini-batch size of 30, and the loss function of cross-entropy was selected to quantify the degree of similarity between the anticipated and real distributions. The updated FC layer was trained with a learning factor of 10 to identify features unique to MRI images. Ten epochs were the maximum number allowed in order to avoid overfitting. Following the modified GoogLeNet's final inception module, features were taken out of the pooling layer and classified using Support Vector Machines (SVM). To support multi-class classification, an ECOC model-based multi-class SVM was used. Additionally, the KNN classifier was employed, with the distance metric being Euclidean distance and the square root of the number of samples in the training set being 49. The five trials were averaged and given in a mean±standard deviation manner. The experiments were carried out five times, with each trial undergoing a five-fold cross-validation procedure. An array of performance metrics is commonly employed to assess a classifier's efficacy. Classification accuracy, defined as the ratio of correctly classified samples to total data samples, is the most widely used metric. The deep transfer learning (standalone) model achieved 92.3±0.7% classification accuracy in our trials, SVM achieved 97.8±0.2% on deep CNN features, and KNN achieved 98.0±0.4% on deep CNN features. The outcomes show that deep CNN feature classification with SVM or KNN performs better. However, if the test dataset is uneven, as the dataset under consideration in this study is, then classification accuracy might not be the right metric to use.

Rehman, Arshia, et al. proposed framework involves three studies that employ three convolutional neural network architectures, namely AlexNet, GoogLeNet, and VGGNet, to classify three types of brain tumors: meningioma, glioma, and pituitary[21]. Each study investigates the application of transfer learning methods on MRI slices of a brain tumor dataset downloaded from Figshare, such as freezing and fine-tuning. The categorization of brain cancers from 233 patients with three different categories of brain tumors using T1-weighted, contrast-enhanced images. MRI scans were processed using a variety of machine learning and image processing algorithms to improve contrast. To create high-resolution contrast images, contrast stretching—a preprocessing method that expands the dynamic range of gray levels for low-contrast MRI images—was used. With 8 bits per pixel, the MRI pictures in the brain tumor dataset downloaded from Figshare have 256 levels of grayscale, ranging from 0 to 255.

Different image variations were obtained by using typical data augmentation techniques, which increased the size of the dataset and decreased overfitting during CNN training. A number of data augmentation methods, including as flipping and rotations, were used to increase the size of the training dataset and give CNNs a lot more input space. One fundamental method of data augmentation was rotation, which involved rotating input images at angles of 90, 180, and 270 degrees. Flipping, which mirrors images both vertically and horizontally, was another technique used. In order to extract distinctive visual characteristics for the categorization of brain tumors, three pretrained CNN architectures were implemented: AlexNet, GoogLeNet, and VGGNet. For every architecture, two distinct transfer learning scenarios were used: freezing and fine-tuning. In order to improve the CNN network's productivity and efficiency, fine-tuning entailed just replacing the final layers of the pretrained network. The pretrained network's weights were transferred from the source dataset (ImageNet) to our target dataset (Figshare) by initializing the ConvNet weights from the top of the pretrained network. A new softmax layer pertinent to the classification problem of brain tumors was added after the pretrained network's softmax layer was trimmed. The last fully connected layer of each CNN architecture was replaced with neurons corresponding to the three classes of the brain tumor dataset (Figshare), replacing the 1000 classes of ImageNet.

In the first case, brain tumor classification was accomplished by employing the softmax layer and initializing the number of neurons to three classes. The learnt visual features from each CNN architecture were then adjusted to the target dataset. In order to enhance performance, the fine-tuning parameters were established and refined based on the outcomes of training the MRI images. For every architecture, 0.9 momentum was used for the training of stochastic gradient descent momentum (SGDM), with a batch size of 10, an initial learn rate of 1e−4, and a maximum of 30 epochs. The validation frequency was set at 300 iterations, and the number of optimal epochs was adjusted according to the validation criteria. Using VGG16 on epoch 7, the best network's greatest accuracy was reached up to 98.69%.

In the second case, an SVM linear classifier was given the pretrained networks' frozen layers, also known as ConvNets or fully linked layers. For the purpose of classifying brain tumors, distinct feature vectors from each architecture were sent separately to the SVM. Conv5 characteristics were given to the SVM by vectorization, and they were more expressive, abstract, and discriminative for AlexNet. In this instance, the maximum accuracy of 96.73% was attained. To unify dimension, the GoogLeNet inception layers were investigated and vectorized throughout the activation process. In this instance, inception-4e-output characteristics yielded the maximum accuracy of 97%. Comparably, Fc7 characteristics for VGG16 were sent to the SVM for classification since they were more expressive. The greatest accuracy that was attained was 89.76%.

Polat, Özlem, and Cahfer Güngen proposed solution involves utilizing transfer learning networks to classify brain tumors in MR images[22]. To identify the most prevalent kinds of brain cancers, they specifically used transfer learning to deploy the VGG16, VGG19, ResNet50, and DenseNet21 networks. Adadelta, ADAM, RMSprop, and SGD are the four optimization algorithms that were used to train and test deep transfer learning networks on the Figshare dataset. This dataset consists of 3064 T1-weighted MR images from 233 patients who were diagnosed with glioma (1426 images), meningioma (708 images), and pituitary (930 images). The images were first acquired with a resolution of 512 × 512 pixels and then reduced to 224 × 224 pixels, which is the input size that the transfer deep learning network can handleFour different transfer learning networks (VGG16, VGG19, ResNet50, DenseNet121) were used to classify three different types of brain cancers using MR images. To extract the most informative features, these networks were tested using four alternative learning algorithms: Adadelta, ADAM, RMSprop, and SGD, with learning rates of 1.0, 0.0001, 0.0001, and 0.01 respectively. Each transfer learning network had a three-layer fully connected neural network appended to the end after feature extraction. ReLU activation functions were present in 1024 and 512 neurons, respectively, in the first and second fully connected layers. The third completely linked layer, which has a Softmax activation function and as many neurons as classes, was used for classification.

For every optimizer, all four approaches were tried ten times using a batch size of sixteen and 50 epochs. An complete training dataset is processed through the neural network once, forward and backward, to define an epoch. In this investigation, an epoch was completed in 134 iterations using a batch size of 16 and 2145 training samples. First, the VGG16 transfer learning network was evaluated; the SGD optimizer produced the best classification score of 97.49%. Second, employing four optimizers, the VGG19 architecture was evaluated ten times for the classification of brain tumors; the SGD optimizer produced the best results, with a performance of 97.93%. As opposed to VGG16 and VGG19, the ResNet50 network produced different results, with the Adadelta optimizer achieving the highest performance of 99.02%. Ultimately, the DenseNet121 network was tested ten times and four different optimizers were applied in order to assess the network's ability to classify brain tumors. The Adadelta optimizer yielded the best categorization performance. It may be inferred that the classification performances attained using alternative optimizers were likewise fairly impressive.

Srinivas, Chetana, et al. presented a comparative performance analysis of transfer learning-based CNN models, namely VGG-16, ResNet-50, and Inception-v3, for automatic prediction of tumor cells in the brain using MRI images[23]. There are 233 images of brain tumors in the dataset used in this investigation. This study compares the VGG-16 pretrained CNN model's performance with that of the ResNet-50 and Inception-v3 models in order to determine the most accurate location of brain tumors. Brain MRI images in a JPEG format that were gathered via Kaggle dataset used in this study. 256 unprocessed MRI images with different sizes expressed in terms of pixel values are included in the collection. The presence of the tumor in an MRI brain imaging determines whether portion of the images gets labeled as Yes or No158 of the 256 images show benign tumors, and the remaining 98 show malignant tumors. The dataset is divided into three parts in order to train, test, and validate the models. False positives and incorrect analyses may result from anomalies in the MRI images, such as poor image quality, distortion, inhomogeneity, misinterpretation, and motion heterogeneity. Options for patient care may be significantly impacted by this. The brain MRI images are scaled to a specified dimension of 224 × 224 × 3 in order to accommodate the pre-trained CNN models. The validation loss in the VGG-16 model demonstrates a stable increase of 0.95% at each epoch interval and tends to decrease until the final epoch, whereas both training and validation accuracy consistently grow up to 0.96% at each epoch interval. Of contrast, the validation accuracy of the Inception-v3 model is unstable while the training accuracy exhibits an increasing tendency. After the eighth epoch, validation accuracy increases the highest, from 0.78% in the fourth epoch to 0.86% in the eighth. From the second to the last epoch, the training accuracy in the ResNet50 model is consistently greater and stable, attaining accuracy above 0.95%.

Ahamed, Md Atik, and Rabeya Tus Sadia explored the performance of state-of-the-art CNN models on two different types of datasets[24]. Images of both normal and tumor categories are included in the first dataset, which is binary. Images of tumors that have been diagnosed as glioma, meningioma, or pituitary are included in the second dataset, which is multi-class. In the trials, pre-trained weights from ImageNet were used for transfer learning, and the weights were initialized randomly. All of the models employed in this study have identical experimental environments to provide a fair comparison. 40% of the data was utilized for validation and 60% of the data was used for training in both datasets, which was constant across all models. This research utilizes two distinct datasets. The first is a binary classification dataset consisting of 3000 samples, while the second is a multi-class comprising 3064 samples. The data processing pipeline comprises two steps. The first step involves data augmentation techniques such as rotation up to 20 degrees, as well as horizontal and vertical flips. An image flip, whether vertical or horizontal, is flipping the rows or columns of pixels. The images are subjected to a random flip operation. Normalization is the second phase, in which every image is downsized to 224 x 224 pixels, the standard 2-dimensional size. With an accuracy of 99.75% in the binary classification dataset, the EfficientNet-B5 architecture outperformed all other state-of-the-art models thanks to the suggested strategies used in this study. Similar to this, the EfficientNet-B5 architecture outperformed other models with an accuracy of 98.61% for the multi-class dataset.

Swati, Zar Nawab Khan, et al. employed pre-trained deep CNN models and propose a block-wise fine-tuning strategy based on transfer learning[25]. A benchmark dataset of T1-weighted contrast-enhanced magnetic resonance images (CE-MRI) is used to assess the suggested approach. The CE-MRI dataset used in this investigation was sourced from figshare, a website that is open to the public. As opposed to 3D volumes, the brain tumor classification system uses two-dimensional (2D) images or slices since, in clinical practice, acquired and available CE-MRI images are usually 2D slices with a significant slice gap. Thier 2D MR image-based classification system is therefore useful for clinical settings. The dataset was gathered from General Hospital at Tianjin Medical University in China and Nanfang Hospital in Guangzhou, China, between 2005 and 2010. In all, 3064 images from 233 patients are included, representing three different tumor types (glioma, meningioma, and pituitary tumor) in axial, coronal, and sagittal views. The images in the dataset are given in matrix format, with 512 x 512 pixel sizes and 49 mm x 49 mm pixel sizes for each image. The CE-MR images are scaled to a range of [0, 1] using min-max normalization in order to make them ready for analysis. After normalization, the image is resized to 224 x 224 pixels and duplicated three times, resulting in three channels that match the pre-trained VGG19 model's input size. The CNN layers' weights are updated following each iteration during training. There are 19 layers and 144 million trainable parameters (weights) in the VGG19 architecture. A sizable dataset is needed for such a deep network's training and optimization. Overfitting could happen, though, and it becomes difficult to identify the right local minima for the cost function with smaller datasets. Consequently, we use the CE-MRI dataset to fine-tune the pre-trained VGG19 model after initializing its weights. Three fully connected layers and sixteen convolutional layers make up the VGG19 model.

It is necessary to fine-tune 19 layers for each of the five-fold cross-validations when using the layer-wise fine-tuning approach, which adds layers one at a time. This yields a total of 95 VGG19 architectures. Layer-by-layer fine-tuning of the VGG19 architecture would take more than a week, assuming that each architecture takes about 30 minutes to train. Even more time-consuming would be figuring out the ideal parameters for the layer-wise fine-tuning technique. Using the layer-wise fine-tuning approach only resulted in a marginal improvement in the results. As a result, pooling layers are used to divide the VGG19 architecture into six blocks, and block-wise fine-tuning is used in its place.

There were initially 1000 neurons in the last fully connected layer of VGG19, which corresponded to the classes in the ImageNet dataset. The final fully connected layer for the CE-MRI dataset is altered to include three neurons, which represent the three classes in the dataset. Block by block, the deep CNN is trained. The last block is fine-tuned first, and all subsequent blocks (layers) are kept fixed by freezing their learning. In pre-trained CNNs, domain-specific features of natural images are found in later layers, while generic features are found in earlier layers. Because these layers contain low-level features, the learning of the earlier layers may become frozen. Block-wise fine-tuning is started from the top block in order to discover the domain-specific characteristics of MRI brain tumors.

The suggested model's application of transfer learning and fine-tuning techniques leads to a decrease in overfitting and a quicker rate of convergence. This is demonstrated by our deep block-wise fine-tuned CNN model's accuracy and loss history, which shows consistency in training-validation accuracy and loss and rapidly reaches its maximum performance. An important aspect of the performance gap between shallow and deep fine-tuning is observed during the experiments. Low-level features are determined by the earlier layers of CNN, whereas domain-specific features are determined by the later layers. In computer vision applications, where the low-level features of the fine-tuned CNN and pre-trained CNN datasets are similar, shallow fine-tuning—fine-tuning only the final few fully connected layers—has produced state-of-the-art results. However, since the dataset contains medical brain MR images that differ from natural images, the CNN may find it challenging to extract pertinent features from medical brain MR images using only natural images if the last few layers are fine-tuned. Particularly in shallow models like FT: B6 and FT: B5-B6, this effect is noticeable. Deep fine-tuning is necessary to improve performance; this involves progressively adding more blocks to the fine-tuning process, which raises performance over time. Experimental results highlight the efficacy of our proposed method by showing a gradual increase in classification performance with incremental block-wise fine-tuning.

Arbane, Mohamed, et al. proposed for accurate classification of brain tumors from MRI images using a convolutional neural network (CNN) based on transfer learning[26]. The system that has been developed looks into a number of cutting-edge CNN architectures, such as ResNet, Xception, and MobileNet-V2. The dataset that was used consists of 253 magnetic resonance imaging (MRI) images, of which 155 samples show the presence of a tumor and 98 samples do not. The original dataset is subjected to a variety of randomized transformations, including rotations, height and width shifts, brightness adjustments, and other techniques made possible by Keras TensorFlow's ImageDataGenerator tool, in order to increase the quantity of MRI images that are available for training the suggested system. The augmentation parameters are carefully selected to ensure that the proposed classifier is never presented with duplicate images during training, thereby improving generalization and mitigating the risk of overfitting. As a result of this augmentation process, the original dataset of 253 images is expanded to 1516 images, providing a more diverse and extensive dataset for training the classifier. Then, the images are resized to dimensions of 224x224x3. This resizing operation serves to normalize the images and ensure that they are uniformly represented across the dataset. all convolutional neural networks are trained for 20 epochs with a batch size of 20. To update the neural weights based on the training data, the Adam optimization algorithm is utilized instead of the conventional stochastic gradient descent optimizer. The Adam optimizer employs a learning rate of 0.001 to optimize the model during the training process. the MobileNet-v2 architecture demonstrated superior performance, achieving F1-score and accuracy metrics of 98.42% and 98.24%, respectively. These results indicate that the MobileNet-v2 classifier is highly effective in accurately classifying brain tumors from MRI images, surpassing the performance of other investigated classifiers.

Saxena, Priyansh, Akshat Maheshwari, and Saumil Maheshwari introduced a transfer learning approach based on convolutional neural networks (CNN) for the classification of brain MRI scans into two distinct classes, utilizing three pre-trained models[27]. The effectiveness of these models is assessed and contrasted with each other in order to ascertain how well each performs in the task of classifying brain tumors. A set of 253 brain MRI images is used as the brain tumor detection dataset. The dataset includes 98 samples of benign tumors and 155 samples of MRI images showing malignant tumors. This study's preprocessing pipeline consists of a few essential steps. To begin with, the dataset is split into three separate parts: the train, test, and validation sets. The validation set is used for hyper-parameter tuning and objective assessment of the trained model, whereas the train set is used for model fitting. The test set is used to conduct the model's last, objective assessment. There are 183 training images, 50 validation images, and 20 testing images from the original dataset, which had 253 images. Crop normalization is a component of the preprocessing pipeline in addition to data splitting. To find the farthest north, south, east, and west (x, y) coordinates along a given contour, one must determine the contour's extreme points. Both rotated bounding boxes and raw contours can be created using this method.

In addition, the dataset's images are resized to a predetermined format to guarantee uniformity for the pre-trained models used in the research. As required by the pre-trained models, the input images are resized to 224 × 224 × 3. The purpose of this resizing operation is to help the deep convolutional neural network learn by standardizing the images. Data augmentation techniques were used to expand the training dataset's size and diversity. During data augmentation, a specific rotation angle of 15° was applied to the images. By adding more variations of the input images to the dataset, this rotation operation helps the deep convolutional neural network learn and classify the input images more efficiently.

Several performance metrics, such as Cohen's kappa (κ), F1-score, area under the receiver operating characteristic (ROC) curve (AUC-ROC), and test accuracy, are used to evaluate the effectiveness of the suggested methodology. ResNet-50 outperformed the other two models under investigation in terms of F1-score because of its high recall and precision values, which are added together to determine the F1-score's harmonic mean. Furthermore, ResNet-50 showed no false-negative rate on the test data, which qualifies it as a viable option for real-world uses where precise tumor identification is essentialResNet-50 had the largest area under the ROC curve, while Inception-V3 had the lowest. Its AUC value of 0.55 suggested that the latter model performed similarly to a random classifier. ResNet-50 outperformed Inception-V3 in terms of test accuracy, with the latter suffering from overfitting. Cohen also had the highest kappa score for ResNet-50 and the lowest for Inception-V3. Ultimately, a thorough analysis and comparison of the performance metrics reveals that ResNet-50 performs better than the other two models in correctly classifying brain tumors.

Table 3 illustrates the summarized of studies exploring CNN-based deep learning approaches in brain tumor classification.

Table 3 The summarized of studies exploring CNN-based deep learning approaches in brain tumor classification.

Author	Datasets	Size of datasets	Preprocessing	Classification task	CNN Model	Performance
Kaur and Gandhi	AANLIB database, Harvard Medical School	160 (140 abnormal ,20 normal)	resizing	classification to normal and abnormal brain images	VGG16	Sensitivity, specificity, and accuracy, with all metrics set at the ideal value of 100%.
Deepak, S., & Ameer, P.	figshare	3064 (930 pituitary tumors images , 708 meningiomas images , and 1426 gliomas images)	Normalization , resizing	classification of brain tumors types, namely glioma, meningioma, and pituitary tumors	GoogLeNet	Accuracy 98.0±0.4%
Rehman, Arshia, et al.	Figshare	3064 (930 pituitary tumors images , 708 meningiomas images , and 1426 gliomas images)	data augmentation	classification of brain tumors types, namely glioma, meningioma, and pituitary tumors	AlexNet, GoogLeNet, and VGGNet	AlexNet acc accuracy of 96.73% GoogleNet accuracy of 97% VGG16 acc89.76%.
Polat, Özlem, and Cahfer Güngen	Figshare dataset	3064,glioma (1426 images), meningioma (708 images), and pituitary (930 images).	———	classification of brain tumors types, namely glioma, meningioma, and pituitary	VGG16, VGG19, ResNet50,	VGG16 97.49%. VGG1997.93% . ResNet 99.02%
Srinivas, Chetana, et al.	Kaggle dataset	256 unprocessed MRI images	———	classification to benign and malignant	VGG-16, ResNet-50, and Inception-v3	ResNet50 accuracy above 0.95%.
Ahamed, Md Atik, and Rabeya Tus Sadia	Figshare dataset	3064,glioma (1426 images), meningioma (708 images), and pituitary (930 images).	data augmentation	first classification to timorous and non timorous then classification tumorours to glioma and meningioma and pituitary	EfficientNet-B5 architecture	accuracy of 99.75%

Swati, Zar Nawab Khan, et al.	Figshare dataset	3064, glioma (1426 images), meningioma (708 images), and pituitary (930 images).	———	classification tumorours to glioma and meningioma and pituitary	the VGG19 model	
Arbane, Mohamed, et al.	Kaggle	253 images(155t and 98 not)	data augmentation	classification tumorous and non tumorous	ResNet, Xception, and MobileNet-V2.	MobileNet high acc 98.24%,
Saxena, Priyansh, Akshat Maheshwari, and Saumil Maheshwari	Kaggle	253 images(155t and 98 not)	resizing, data augmentation	classification to benign and malignant	ResNet-50, Inception-V3	ResNet high performance

8. Conclusion:

Transfer learning models are a promising approach for brain tumor detection in medical imaging. Compared to traditional machine learning algorithms and CNNs, transfer learning models offer several advantages, such as improved accuracy, reduced training time, and the ability to generalize well to new datasets.

Recent studies have demonstrated the effectiveness of transfer learning models in detecting brain tumors in MRI scans, achieving high accuracy levels even with small datasets. By leveraging pre-trained models that have already learned features from large datasets, transfer learning models can effectively extract meaningful features from MRI scans and enable accurate tumor detection.

While CNNs and traditional machine learning algorithms have also shown promise in detecting brain tumors, transfer learning models offer a more efficient and effective approach for this task. Moreover, transfer learning models can be easily adapted to different types of imaging modalities and can be fine-tuned on new datasets to achieve even higher accuracy levels.

References

[1] D. N. Louis, A. Perry, G. Reifenberger, A. Deimling, D. Figarella-Branger, W. K. Cavenee, H. Ohgaki, O. D. Wiestler, P. Kleihues, D. W. Ellison, The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary, *Acta Neuropathol*, 131(6), (2016) 803–820. https://doi.org/10.1007/s00401-016-1545-1

[2] A. Darlix, S. Zouaoui, V. Rigau, F. Bessaoud, D. Figarella-Branger, H. Mathieu-Daudé, B. Trétarre, F. Bauchet, H. Duffau, L. Taillandier, L. Bauchet, Epidemiology for primary brain tumors: A nationwide population-based study, *Journal of Neurooncol*, 131(3), (2017) 525–546. https://doi.org/10.1007/s11060-016-2318-3

[3] Q.T. Ostrom, H. Gittleman, G. Truitt, A. Boscia, C. Kruchko, J. S. Barnholtz-Sloan, Neuro-Oncology CBTRUS Statistical Report : Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011–2015, 20(4), (2018) iv1–iv86. https://doi.org/10.1093/neuonc/noy131

[4] J. Gonzalez-Castillo and P. A. Bandettini, Task-based dynamic functional connectivity: Recent findings and open questions, *Neuroimage*, 180, (2018) 526–533. https://doi.org/10.1016/j.neuroimage.2017.08.006

[5] B. H. Menze *et al.*, The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS), *IEEE Trans. Med. Imaging*, 34(10), (2015) 1993–2024. https://doi.org/10.1109/TMI.2014.2377694

[6] D. Rammurthy and P. K. Mahesh, Whale Harris hawks optimization based deep learning classifier for brain tumor detection using MRI images, *Journal of King Saud University - Computer and Information Sciences*, 34(6), (2022) 3259–3272. https://doi.org/10.1016/j.jksuci.2020.08.006

[7] J. Ker, Y. Bai, H. Y. Lee, J. Rao, and L. Wang, Automated brain histology classification using machine learning, *Journal of Clinical Neuroscience*, 66, (2019) 239–245. https://doi.org/10.1016/j.jocn.2019.05.019

[8] M. A. Mazurowski, M. Buda, A. Saha, and M. R. Bashir, Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI, *Journal of Magnetic Resonance Imaging (JMRI)*, 49(4), (2019) 939–954. https://doi.org/10.1002/jmri.26534

[9] G. S. Tandel, A. Tiwari, and O. G. Kakde, Performance optimisation of deep learning models using majority voting algorithm for brain tumour classification, *Computers in Biology and Medicine*, 135, (2021) 104564. https://doi.org/10.1016/j.compbiomed.2021.104564

[10] S. Ahuja, B. K. Panigrahi, and T. Gandhi, Transfer Learning Based Brain Tumor Detection and Segmentation using Superpixel Technique, *2020 International Conference on Contemporary Computing and Applications (IC3A)*, (2020) 244–249.

https://doi.org/10.1109/IC3A48958.2020.233306

[11] S. Pirzada *et al.*, Spatial normalization of multiple sclerosis brain MRI data depends on analysis method and software package, *Magnetic Resonance Imaging*, 68, (2020) 83–94. https://doi.org/10.1016/j.mri.2020.01.016

[12] E. Kondrateva, P. Druzhinina, A. Kurmukov, and K. Net, Do we really need all these preprocessing steps in brain MRI segmentation?, *Medical Imaging with Deep Learning*, 2022. https://openreview.net/pdf?id=7ub0rd8h_Ie

[13] M. B. V. Reddy, D. P. B. Reddy, D. P. S. Kumar, and D. S. S. Reddy, Developing an Approach to Brain MRI Image Preprocessing for Tumor Detection, *International Journal of Research*, 1(6), (2014) 725–731.

[14] P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, and A. Haworth, A review of medical image data augmentation techniques for deep learning applications, *Journal of Medical Imaging and Radiation Oncolgy*, 65(5), (2021) 545–563. https://doi.org/10.1111/1754-9485.13261

[15] S. Deepak and P. M. Ameer, Automated Categorization of Brain Tumor from MRI Using CNN features and SVM, *Journal of Ambient Intelligence and Humanized Computing*, 12(8), (2021) 8357–8369. https://doi.org/10.1007/s12652-020-02568-w

[16] S. Kuraparthi *et al.*, Brain tumor classification of MRI images using deep convolutional neural network, *Traitement du Signal*, 38(4), (2021) 1171–1179. https://doi.org/10.18280/ts.380428

[17] R. A. Zeineldin, M. E. Karar, J. Coburger, C. R. Wirtz, and O. Burgert, DeepSeg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images, *International Journal of Computer Assisted Radiology and Surgery*, 15(6), (2020) 909–920. https://doi.org/10.1007/s11548-020-02186-z

[18] R. Pugalenthi, M. P. Rajakumar, J. Ramya, and V. Rajinikanth, Evaluation and classification of the brain tumor MRI using machine learning technique, *Control Engineering and Applied Informatics*, 21(4), (2019) 12–21.

http://www.ceai.srait.ro/index.php?journal=ceai&page=article&op=view&path%5B%5D=6505&path%5B%5D=1549

[19] T. Kaur and T. K. Gandhi, Automated brain image classification based on VGG-16 and transfer learning, *2019 International Conference on Information Technology (ICIT)*, (2019) 94–98. https://doi.org/10.1109/ICIT48102.2019.00023

[20] S. Deepak and P. M. Ameer, Brain tumor classification using deep CNN features via transfer learning, *Computers in Biology and Medicine*, 111, (2019) 103345. https://doi.org/10.1016/j.compbiomed.2019.103345

[21] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, A Deep Learning-Based Framework for Automatic Brain Tumors Classification Using Transfer Learning, *Circuits, Systems, and Signal Processing*, 39 (2), (2020) 757–775. https://doi.org/10.1007/s00034-019-01246-3

[22] Ö. Polat and C. Güngen, Classification of brain tumors from MR images using deep transfer learning, *The Journal of Supercomputing*, 77(7), (2021) 7236–7252. https://doi.org/10.1007/s11227-020-03572-9

[23] C. Srinivas *et al.*, Deep Transfer Learning Approaches in Performance Analysis of Brain Tumor Classification Using MRI Images, *Journal of Healthcare Engnierring*, 2022, (2022) 3264367. https://doi.org/10.1155/2022/3264367

[24] M. A. Ahamed and R. T. Sadia, Examining the behaviour of state-of-the-art convolutional neural networks for brain tumor detection with and without transfer learning, (2022) 1–15 [Online]. http://arxiv.org/abs/2206.01735

[25] Z. N. K. Swati *et al.*, Brain tumor classification for MR images using transfer learning and fine-tuning, *Computerized Medical Imaging and Graphics*, 75, (2019) 34–46. https://doi.org/10.1016/j.compmedimag.2019.05.001

[26] M. Arbane, R. Benlamri, Y. Brik, and M. Djerioui, Transfer Learning for Automatic Brain Tumor Classification Using MRI Images, *2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH)*, (2021) 210–214. https://doi.org/10.1109/IHSH51661.2021.9378739

[27] P. Saxena, A. Maheshwari, and S. Maheshwari, Predictive Modeling of Brain Tumor: A Deep Learning Approach, *Advances in Intelligent Systems and Computing*, 1189, (2021) 275–285. https://doi.org/10.1007/978-981-15-6067-5_30