

An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning

Md. Alamin Talukder^a, Md. Manowarul Islam^a, Md Ashraf Uddin^b, Arnisha Akhter^a, Md. Alamgir Jalil Pramanik^c, Sunil Aryal^b, Muhammad Ali Abdullah Almoyad^d, Khondokar Fida Hasan^e, Mohammad Ali Moni^f

^aDepartment of Computer Science and Engineering, Jagannath University, Dhaka, Bangladesh

^bSchool of Information Technology, Deakin University, Geelong Warrn Ponds Campus, Australia

^cDept of Surgery, Rangpur Medical College Hospital, Rangpur

^dDepartment of Basic Medical Sciences College of Applied Medical Sciences in Khamis Mushyt King Khalid University, 47 Abha, Mushait, PO Box. 4536, ZIP. 61412, (Saudi Arabia)

^eSchool of Computer Science, Queensland University of Technology (QUT), 2 George Street, Brisbane 4000, Australia

^fArtificial Intelligence & Data Science, School of Health and Rehabilitation Sciences, Faculty of Health and Behavioural Sciences, The University of Queensland St Lucia, QLD 4072, Australia

Abstract

Brain tumors are among the most fatal and devastating diseases, often resulting in significantly reduced life expectancy. An accurate diagnosis of brain tumors is crucial to devise treatment plans that can extend the lives of affected individuals. Manually identifying and analyzing large volumes of MRI data is both challenging and time-consuming. Consequently, there is a pressing need for a reliable deep learning (DL) model to accurately diagnose brain tumors. In this study, we propose a novel DL approach based on transfer learning to effectively classify brain tumors. Our novel method incorporates extensive pre-processing, transfer learning architecture reconstruction, and fine-tuning. We employ several transfer learning algorithms, including Xception, ResNet50V2, InceptionResNetV2, and DenseNet201. Our experiments used the Figshare MRI brain tumor dataset, comprising 3,064 images, and achieved accuracy scores of 99.40%, 99.68%, 99.36%, and 98.72% for Xception, ResNet50V2, InceptionResNetV2, and DenseNet201, respectively. Our findings reveal that ResNet50V2 achieves the highest accuracy rate of 99.68% on the Figshare MRI brain tumor dataset, outperforming existing models. Therefore, our proposed model's ability to accurately classify brain tumors in a short timeframe can aid neurologists and clinicians in making prompt and precise diagnostic decisions for brain tumor patients.

Keywords: Neuroscience; Deep Learning; Transfer Learning; Brain MRI Image; Brain Tumor; Classification

1. Introduction

The brain is an essential and highly complex component of the human body, responsible for governing both intentional and unintentional activities (Quader et al., 2022). As one of the most intricate and delicate organs, it controls various critical functions, including cognition, emotion, vision, hearing, and response (Asif et al., 2022). Brain tumors, which result from abnormal tissue growth within the skull, are among the most lethal brain disorders.

*Mohammad Ali Moni

**Md Manowarul Islam

Email addresses: alamintalukder.cse.jnu@gmail.com (Md. Alamin Talukder), manowar@cse.jnu.ac.bd (Md. Manowarul Islam), ashraf.uddin@deakin.edu.au (Md Ashraf Uddin), arnisha@cse.jnu.ac.bd (Arnisha Akhter), alamgirjalilpramanik@gmail.com (Md. Alamgir Jalil Pramanik), sunil.aryal@deakin.edu.au (Sunil Aryal), maabdullah@kku.edu.sa (Muhammad Ali Abdullah Almoyad), fida.hasan@qut.edu.au (Khondokar Fida Hasan), m.moni@uq.edu.au (Mohammad Ali Moni)

These tumors can be categorized into primary and secondary types. Primary brain tumors, accounting for 70% of cases, only spread within the brain, whereas secondary tumors originate in other organs like the breast, kidney, or lung before metastasizing to the brain (Kibriya et al., 2022).

Common types of brain tumors include gliomas, meningiomas, and pituitary tumors. Gliomas result from abnormal proliferation in glial cells, which constitute approximately 80% of the brain. Meningiomas develop in the meninges spinal cord, the protective layer of the brain, while pituitary tumors arise within the pituitary gland, responsible for producing essential hormones. Although pituitary tumors are typically benign, they can cause hormonal imbalances and irreversible vision impairment (Komninos et al., 2004).

Numerous disorders can damage the brain, notably brain tumors, which are primarily accompanied by abnormal proliferation within the nervous system (Naki and Aderibigbe, 2022). These abnormalities were extremely difficult to cure; therefore, prompt detection is critical to the human's wellbeing (Almadhoun and Abu-Naser, 2022). World Health Organization (WHO) suggests that the brain tumors will expand by 5% each year globally (Çinar and Yildirim, 2020). Brain tumors are more deadly and difficult to diagnose than a tumor in any other section of the body. Since the brain is surrounded by the blood-brain barrier, typical radioactivity detectors are unable to detect tumor cell impulsivity (Graber et al., 2019; Naseer et al., 2021). As a result, magnetic resonance imaging (MRI) as well as computed tomography (CT) images are considered the most effective clinical tracers for detecting brain disturbance (Naseer et al., 2021). MRI is a widely used technology for diagnosing and prognosing brain tumors in a variety of neurological disorders and situations (Gurbină et al., 2019). Many clinical disorders now require MRI as the main diagnostic evaluation (Abd-Ellah et al., 2019). The premise underlying MRI is to achieve higher cross-sectional images of organs using non-ionizing radio-frequency electromagnetic waves in the context of regulated magnetic fields (Katti et al., 2011). It is also thought to be superior to CT since it exposes patients to less radiation, has less dimensional inaccuracy, and has no adverse effects (Niraj et al., 2016).

In neuroscience, brain tumors are a hot topic, as the detection of brain tumors in the early stages is very important to protect against loss of human life (Islam et al., 2021). Although there are several approaches for diagnosing abnormalities in brain magnetic resonance scans, there is still scope for enhancing performance and making the classification within a reasonable amount of time (Mandle et al., 2022). Due to the growing volume of medical data, attempt in analyzing and extrapolating them using conventional techniques are becoming increasingly difficult (Abiwinanda et al., 2019; Badža and Barjaktarović, 2020). Scientists now have a new incentive to optimize present methodologies for more complete clinical research (Bruton et al., 2020). Deep learning is a popular technique for evaluating biomedical data in the healthcare field (Naser and Deen, 2020; Talukder et al., 2022).

Deep learning (DL) is an advanced categorization and prediction innovation that has demonstrated outstanding performance in domains that require multilevel data processing, such as classification, detection, and voice recognition. (Pyrkov et al., 2018). It can obtain valuable underlying patterns from images that have been shown to achieve provincial efficiency (Ciregan et al., 2012). It is the most notable ML achievement, capable of managing complex pattern recognition and object detection from image dataset (Avci et al., 2021). Traditional ML-based techniques are not applicable for image classification because they rely heavily on hand-made features (Le et al., 2019). The essential factor to make them attractive to complex biomedical applications is their ability to extract optimized features directly from raw data to the nature of the problem to enhance classification performance (Kiranyaz et al.,

2021).

Transfer learning (TL) refers to the process that uses the knowledge of previously trained models to discover a new set of data to deal with a precise scenario (Tan et al., 2018). Moreover, we do not require a lot of processing power. The model employs the convent weights of the pre-trained model and trains only the final dense layer (Talo et al., 2019). There are three ways in which it can be utilized, namely as a baseline model for object classification that can be used to train TL models on imagenet data (Morid et al., 2021); as a feature extractor that extracts features from image data and then uses deep learning or machine learning for labels. (Abbasi et al., 2020); as a fine-tuning, which requires changing the last layer to suit the classes of the preferred data source and retraining the network's layers (Montalbo, 2020).

Numerous efforts have been identified in the related works, each based on a unique strategy for classifying brain tumors (Belaid and Loudini, 2020; Rehman et al., 2020; Afshar et al., 2018; Sadad et al., 2021; Ayadi et al., 2021). Various medical images have already been identified and represented using DL approaches. Its procedures have enabled machines to evaluate multidisciplinary pathology scans, high-dimensional image data, and video recordings. (Talukder et al., 2022; Andresen et al., 2022). As they can handle biomedical image data, many deep-learning techniques have been applied to disease prediction. (Khan et al., 2022; Islam et al., 2022; Savaş, 2022).

Manually assessing and analyzing a vast array of brain MRI data is not only time-consuming and costly, but it can also be prone to errors, as the processing and classification of MRI images require expertise. Accurate diagnosis and classification of brain tumors are vital, as they inform prognostic predictions and enable medical professionals to select the most suitable treatment options.

To help medical experts in selecting the best line of action and stopping the early death of life due to brain tumors, we need to build a robust DL model to accurately predict brain tumor with less amount of time. Hence, our research focuses on creating a productive and well-organized framework to classify brain tumors in which we use numerous preprocessing steps to prepare our dataset, reconfigure the TL architecture, and add some extra layers. The proposed DL approach is tested on the publicly available Figshare MRI brain tumor dataset. In order to build a robust model, we used our novel DL approach for effective improvement in brain tumor classification. In this research, we assess the effectiveness of our proposed deep learning model by utilizing various performance metrics, such as Accuracy, Precision, Recall, F1-score, Confusion Matrix, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Squared Error (MSE). The findings demonstrate that our deep learning model is capable of classifying brain tumors with an exceptional accuracy rate surpassing 99%.

The main contributions of this research are as follows:

- It proposes a novel deep learning model for brain tumor classification, incorporating comprehensive preprocessing, transfer learning architecture modification, and fine-tuning to enhance efficiency.
- Reconfiguration architecture is modified by involving image augmentation to solve overfitting problems and utilize the GPU speed. Furthermore, to get instantaneously standardize images based on the configuration, which helps the initiative of reimplementing the augmentation process.
- Fine-tuning is the process of adding layers with the modified architecture that lets us build a new DL architecture to classify brain tumor efficiently.

- Finally, assess the effectiveness of our proposed models using multiple criteria, including accuracy, precision, recall, f1-score and confusion matrix and finding the best model to categorize brain tumor.

The subsequent sections of the paper are structured as follows: In Section 2, an overview of related work on brain tumor diagnosis using deep learning is presented. The methodology and dataset used in our research are elaborated in Section 3. Section 4 outlines the experimental setup and performance evaluation. Lastly, the paper concludes with Section 5.

2. Related Works

Classification of brain tumours is essential for evaluating tumours and deciding on medication options based on their categories. Brain cancers are detected using a variety of neuroimaging methods. Conversely, MRI is frequently employed because of its greater image quality and lack of ionising radiation. DL is a machine learning discipline that has consistently demonstrated outstanding results, particularly in categorization and detection issues. Table 1 shows the tree diagram of related works for brain tumor classification.

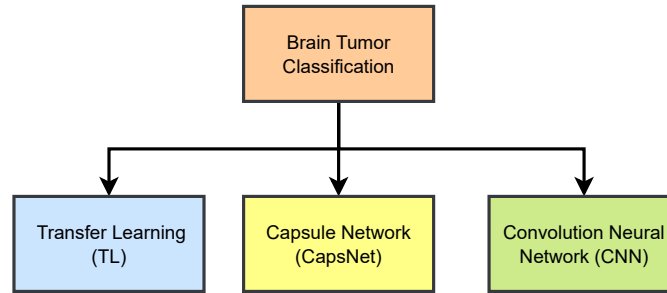


Figure 1: The tree diagram of related works for brain tumor classification

2.1. Brain tumor classification using Transfer Learning (TL)

Belaïd and Loudini (2020) explored the use of transfer learning networks to categorize brain tumors in MRI images. The TL networks were trained and evaluated using several optimization strategies on the Figshare brain tumor dataset to identify the most frequent brain lesions. With ResNet50 and Adadelta optimisation, the presented transfer learning model got the greatest classification accuracy of 99.02 percent. The classification results showed that the most frequent brain tumor may be classified with excellent accuracy. As a result, the transfer learning paradigm in medicine holds promise and can help physicians make rapid and precise diagnoses.

Rehman et al. (2020) used three different designs of convolutional neural networks to diagnose brain lesions in three experiments. They employed MRI slices from the Figshare brain tumor dataset; each study investigates TL strategies, including fine-tuning and freezing. The MRI slices are subjected to data augmentation procedures to improve the generalization of results and expand data sampling to lower the overfitting chance. The optimized VGG16 model produced the best results in categorization and had a prediction rate of 98.69% in the tests.

Sadad et al. (2021) utilized NASNet and ResNet50-UNet TL architects to achieve brain tumor classification. To improve the recognition results, the pre-processing and data augmentation idea was established. According to the

findings, the research proposal paradigm outperformed the current state of the art. Among the many TL models used for tumor categorization, NASNet was the greatest accuracy rate of 99.6 percent.

Tummala et al. (2022) used ImageNet-based ViT models that had been pre-trained and fine-tuned for categorization assignment. The Figshare brain tumor data set was utilized to evaluate the performance of the ensemble ViT model in cross-validation (CV) and testing for a three-class classification task. The combination of all four ViT variants, L/16,L/32, B/16 and B/32, yielded a 98.7% total testing accuracy. As a result, a collection of ViT models could be used to support the computer-assisted identification of brain cancers based on MRI scans, easing the burden on radiologists.

Swati et al. (2019) proposed a blockwise fine-tuning approach utilizing TL with a pre-trained CNN model. On the Figshare brain tumor dataset, the offered approach was tested. Their strategy was more versatile because it did not employ any feature extractor, required little preprocessing, and had an average accuracy of 94.82 percent when tested five times. They contrasted their findings with classic CNN-based ML and DL methods. The developed VGG19 technique exceeded state-of-the-art classification, according to experimental results.

2.2. Brain tumor classification using Capsule Network (CapsNet)

Afshar et al. (2018) exploited CapsNets to increase the performance of the categorization issue on a real series of brain MRI images to identify brain tumors. Due to the handling of a limited training set and their units being increasingly adaptable, they exceed CNNs in the tumor diagnosis challenge by 86.56 percent. Their findings demonstrated that the suggested method could efficiently defeat CNNs in the categorization of brain tumors.

Afshar et al. (2019) developed a modified CapsNet design that included both raw MRI brain images and tumor borders in order to categorize cancers. The tumor rough borders are added as new inputs to the CapsNet's workflow to boost the CapsNet's emphasis. They utilized the figshare brain MRI data with 3064 pictures to test their recommended CapsNet framework and achieved a 90.89 percent accuracy rate, which exceeds its competitors significantly. They showed that, in contrast to previous CapsNets and CNNs, the new technique improved the classification accuracy. It was also notable to observe that CapsNets were equipped with features that improved their interpretability.

Afshar et al. (2020) introduced a Bayesian CapsNet framework, termed the BayesCap, capable of providing not only mean estimates, but also entropy as an indicator of forecasting uncertainties, by taking advantage of the ability of capsule networks to regulate small datasets and control uncertainty. To test the BayesCap model, they used Figshare brain tumor dataset and obtained a 78 percent accuracy rate in detecting brain tumors. They demonstrated that screening out the uncertain forecasts improves accuracy, indicating that releasing the uncertain forecasts was a good method for increasing network comprehensibility.

2.3. Brain tumor classification using Convolution Neural Network (CNN)

Badža and Barjaktarović (2020) demonstrated a new CNN model for brain tumor categorization that is easier to use than pre-trained networks and was evaluated using Figshare MRI data. The network's capacity to generalize was evaluated through various methods, including the use of the 10-fold CV technique. The record-wise CV for the augmented data yielded the highest results among the different approaches. They had a 96.56 percent accuracy

rate. The novel CNN architecture could be employed as an excellent verdict utility for radiologists in diagnostic purposes, together with its high generalization potential and processing speed.

Leveraging two publically accessible datasets, a DL model premised on a convolutional neural network is presented by (Sultan et al., 2019) to diagnose various brain tumor kinds. For the 2 experiments, the presented network topology produces remarkable results, with the greatest overall accuracy of 96.13 percent and 98.7 percent, respectively. The findings proved that the proposal could be used to detect multiple types of brain tumors.

Ait Amou et al. (2022) presented an improved hyperparameter optimization strategy for CNN that relies on Bayesian optimization. This strategy was tested by categorizing MRI scans of the brain into three classes of cancers. Five well-known deep pre-trained models are examined to optimize CNN's efficiency using TL. Despite the use of data augmentation or cropping lesion procedures, their CNN was capable of achieving an accuracy rate of 98.70 percent at most after utilizing Bayesian optimization. Furthermore, using the MRI sample, the suggested model exceeds the existing work, proving the viability of hyperparameter optimization automation.

Abiwinanda et al. (2019) attempted to train CNN to recognize three different types of brain tumors: gliomas, meningiomas, and pituitary tumors. The authors used a simple CNN architecture that included only one convolution layer, a max-pooling layer, and a flattening layer over each hidden layer, followed by a full connection from one hidden layer. Their model achieved 98.51 percent training accuracy and 84.19 percent validation accuracy despite its simplicity and lack of prior region-based segmentation. This method has the potential to be a straightforward tool for doctors in accurately diagnosing brain tumours.

Das et al. (2019) worked on constructing a CNN model for the classification of brain tumors, and the designed scheme was made up of two major steps: preprocessing images using various image processing strategies and then categorizing the processed images using CNN. Using their CNN model on a brain tumor dataset, they were able to attain a high accuracy of 94.39 percent. In the data set, the designed scheme demonstrated satisfactory accuracy and exceeded the number of well-known current approaches.

Paul et al. (2017) employed 989 axial photos to develop a convolutional neural network over the axial data, which proved to be efficient in categorization with a 5-fold CV of 91.43 percent on the finest CNN model. This finding showed that CNN outperformed specialized approaches in tumors that need image dilation and ring-forming subareas.

Ayadi et al. (2021) demonstrated a novel CNN-based model with multiple layers for classifying MRI brain tumors. It was an intelligent model that needed very little preprocessing and was evaluated on three different brain tumor datasets. To test the accuracy of the model and determine the resilience of the system, a variety of performance metrics were examined. With an accuracy rate of 94.74% for Figshare, 93.71% for Radiopaedia, and 97.22 percent for Rembrandt datasets, the proposed scheme achieved the best classification and recognition accuracies relative to previous relevant studies along the same data.

3. Methodology

This section presents our proposed methodology, which includes a description of the various transfer learning (TL) algorithms that are utilized in our approach. First, we explain the working principle of our proposal and then we briefly describe the transfer learning algorithms.

To ensure a brain tumor prognosis, we constructed our proposed approach using image data collection, preprocessing, reconstruction transfer learning architecture and fine-tune by attaching some extra layers such as global avg. pooling, batch normalization and dense layers to classify brain tumors on a brain tumor dataset. The block diagram and architecture of our proposed paradigm are depicted in Fig 2 and Fig 3. The following are the steps of our proposed approach:

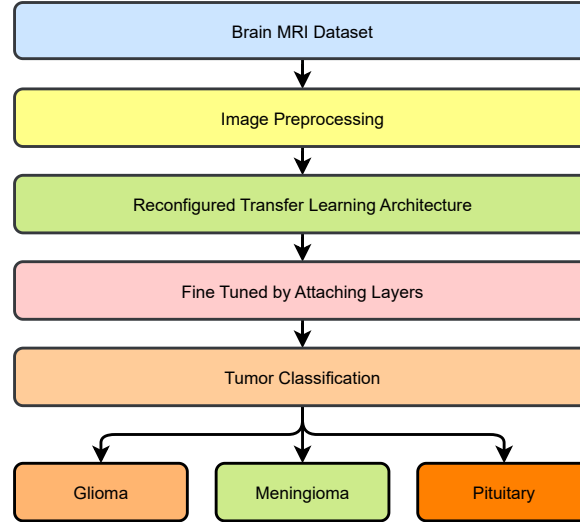


Figure 2: The block diagram of our proposed research

- Step-1: Initially, we take the brain tumor dataset to conduct our experiment. There are three types of brain tumors available in our dataset such as glioma, meningioma and pituitary.
- Step-2: During the pre-processing stage, the image is resized to achieve the desired size of 256x256, applied a filter to sharpen the image, then complements the image and performed image scaling to normalize the image data.
- Step-3: In the reconstruction transfer learning architecture step, we add image augmentation after the input layer and then truncate the last few layers after the activation layer.
- Step-4: In this step, we fine-tune by attaching some layers including global average pooling, batch normalization and dense layer to make it more suitable to classify brain tumors.
- Step-5: In this step, some well-known transfer learning algorithms such as Xception, ResNet50V2, InceptionResNetV2 and DenseNet201 are utilized in our approach.
- Step-6: Finally, the performance is evaluated for each transfer learning model and selecting the best one based on various performance metrics, including Accuracy, Precision, Recall, F1-score, Confusion Matrix, RMSE, MAE and MSE. Furthermore, a comparison analysis is performed with other existing works.

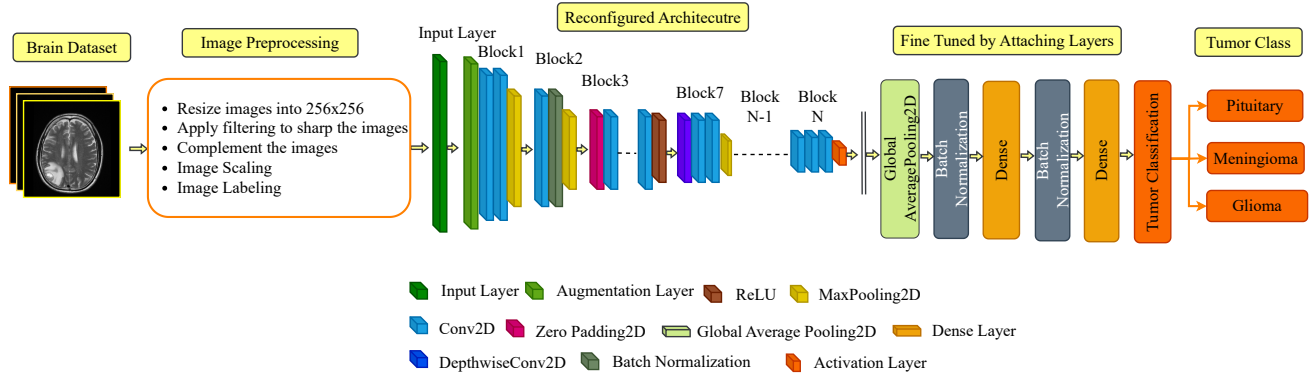


Figure 3: The proposed architecture for brain tumor classification

3.1. Data collection

The dataset (Cheng, 2017) contains 3064 T1-weighted contrast-enhanced images of the brains of 233 patients who had been diagnosed with one of three types of brain tumours: meningioma (708 slices), glioma (1426 slices), or pituitary tumour (930 slices). The information can be downloaded in the form of Matlab files (.mat files). Each image file includes a struct that contains pertinent information about the image, such as the label (1 for meningioma, 2 for glioma, and 3 for pituitary tumour), patient ID (PID), image data, and tumour Border. The tumor border is a vector that contains the coordinates of distinct points on the tumor's edge, and it is obtained by manually tracing the tumor border. Due to the availability of this information, the generation of a binary image of the tumor mask is made simple. In addition, the dataset contains a tumor Mask, which is a binary image with the tumor region represented by a string of ones.

The distribution of the dataset is depicted in Fig. 4

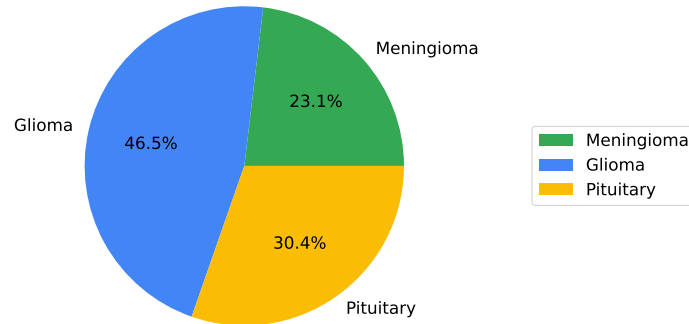


Figure 4: The distribution of brain tumor dataset

3.2. Data preprocessing

In the data preprocessing, we took the dataset and prepare it for processing by taking the image and label information from the dataset as the dataset was in Matlab (.mat) file format. Then we start image preprocessing by utilizing the resize the images into 256x256, applying a sharp filter to sharpen the images and complement the images to make them more visible. After that, we scale the image by dividing the images by 255. Moreover, to

fit into the CNN model, we split our dataset into train, test and validation parts in 80%, 10% and 10% and also shuffle 1000 times to minimize loss, reduce the variance and generalize the model.

The preprocessed images are sharper, brighter, and have more detectable details than the original images, making them appropriate for driving into the model and achieving greater performance than existing works. Figure 5 illustrates the image preprocessing steps for brain tumor types, including glioma, meningioma, and pituitary. The top section of the figure (a,b,c) shows the images before preprocessing, while the bottom section (d,e,f) displays the images after the preprocessing steps have been applied.

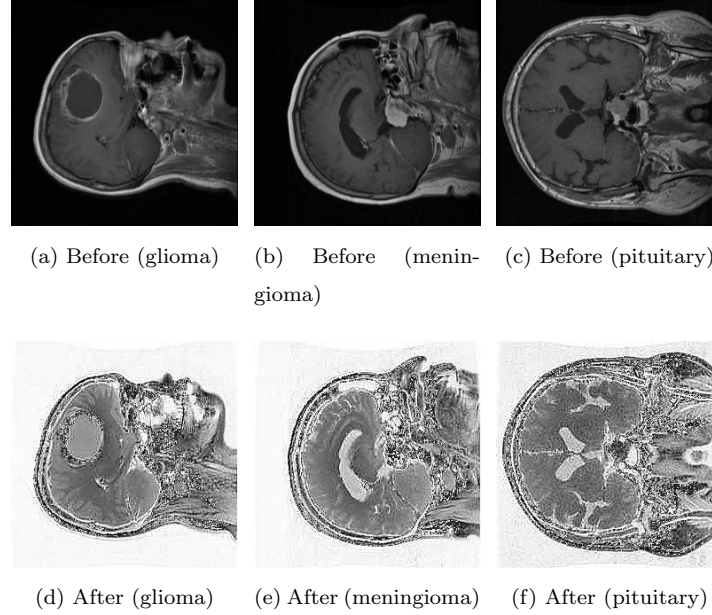


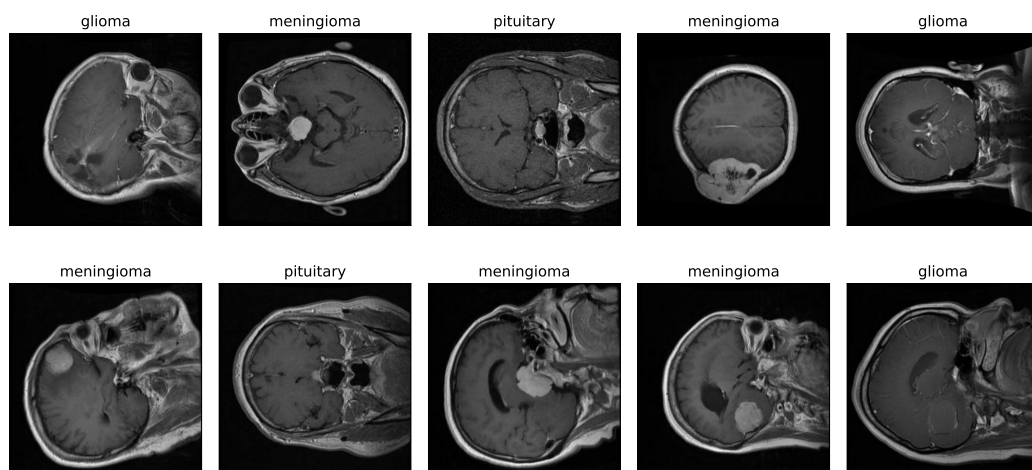
Figure 5: Brain glioma, meningioma and pituitary image before and after image preprocessing

We have further enhanced the experimental setup by incorporating additional photographs to measure the impact of the image preprocessing techniques proposed in this paper. In Figure 6a and 6b, we present a selection of before-and-after images randomly selected from our dataset, illustrating the demonstrable effects of the image processing methods employed. These visual examples serve to provide compelling evidence of the efficacy of our image preprocessing approach, enhancing the attractiveness and visual appeal of our research

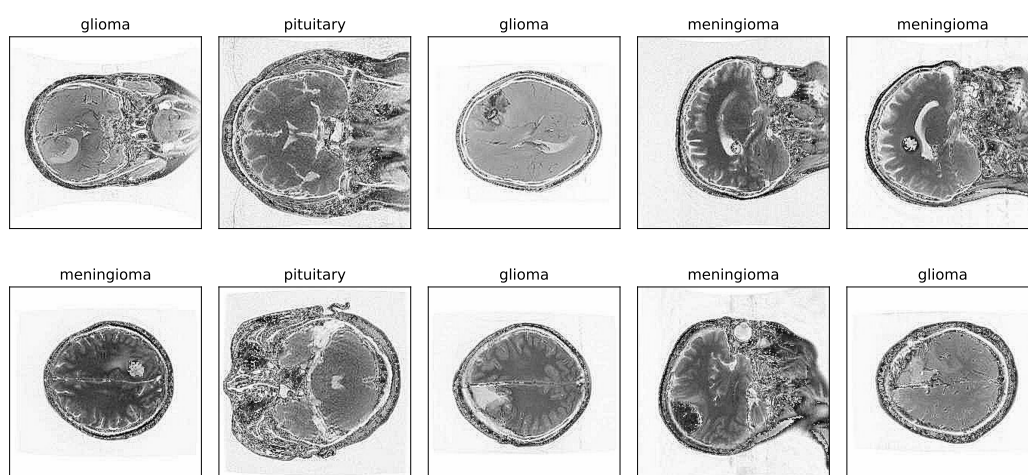
3.3. Reconstruction transfer learning architecture

In reconstruction transfer learning architecture, we reconstruct the architecture of transfer learning as we have already known that the transfer learning algorithms are pre-trained with ImageNet data (You et al., 2020) so to use it in our dataset we need to reconstruct the architecture for better predictions. We reconstruct the architecture so that we can utilize all transfer learning algorithms in our modified architecture. This procedure follows two steps:

- **Image Augmentation:** Initially, we take the input layer then we add the image augmentation layer means we make the augmentation layer part of our architecture so that the modified architecture can use preprocessed images to perform augmentation on-device, simultaneously with the remainder of the layers, and advantage from GPU speed. In addition, when we extract our model, the preprocessing layers are stored alongside the



(a) Before Image Preprocessing



(b) After Image Preprocessing

Figure 6: Sample of before and after image preprocessing

rest of the model. When we afterward deploy this model, it will instantaneously standardize images (based on the configuration of our layers), saving us the initiative of reimplementing that logic server-side.

- **Truncate Layers:** After that, we keep all the layers of transfer learning algorithms except after the last activation layers in our architecture as we want to add more layers to make it a more efficient architecture to predict brain tumors.

Fig. 7 shows the original and reconfigured transfer learning architectures where Fig. 7(a) represents the original architecture of transfer learning and Fig. 7(b) represents the reconfigured architecture of transfer learning.

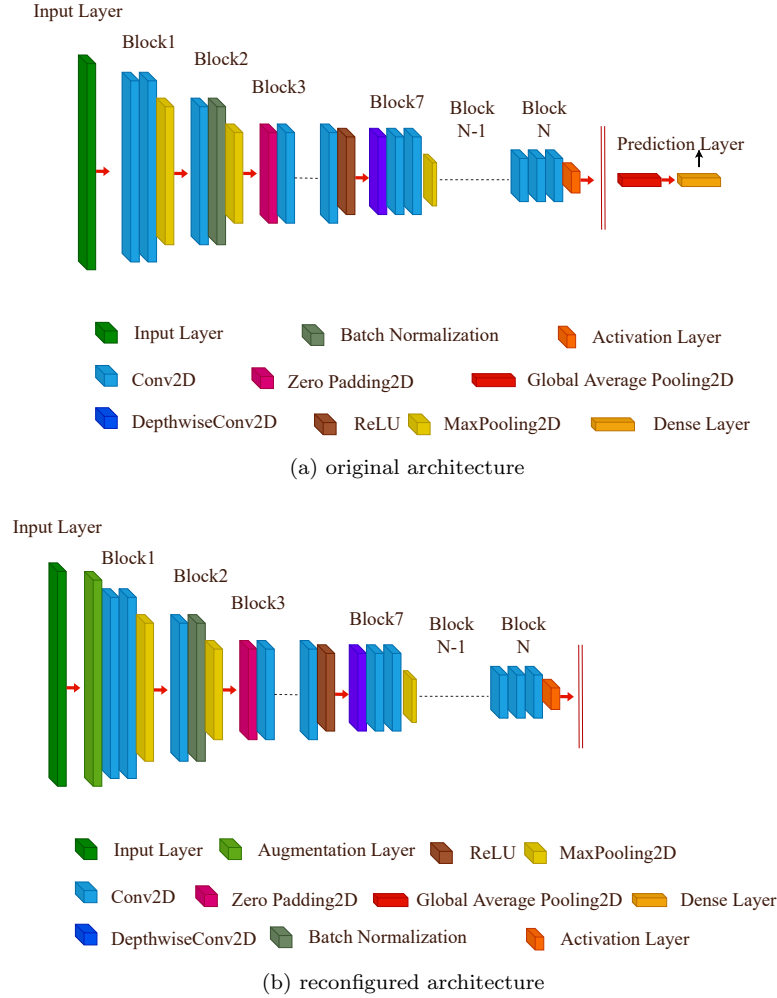


Figure 7: The original and reconfigured architecture of transfer learning model

3.4. Image Augmentation

Figure 8 illustrates the image augmentation process used in our proposed architecture. Image augmentation is a frequently employed approach to enhance the scale and variety of a dataset, which in turn can enhance the efficacy of deep learning models. In our approach, we applied a series of image processing techniques to the input images to generate augmented images. The augmentation process includes horizontal flipping, rotation, zooming, and contrast adjustment. The input images are also rescaled to 0 to 1 to facilitate model training. Additionally, we

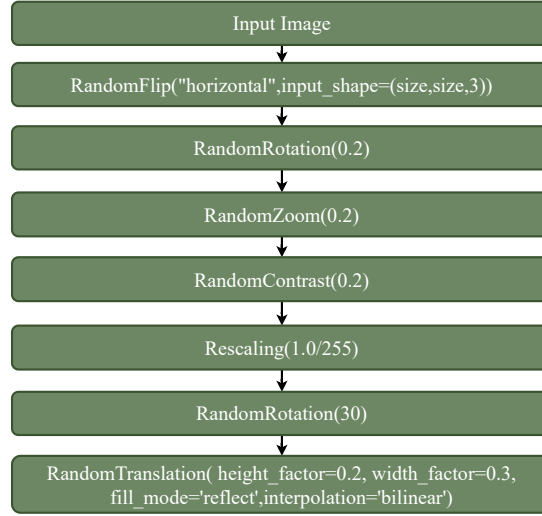


Figure 8: The augmentation process of our proposed architecture

randomly rotated and translated the images to increase the diversity of the augmented dataset further. Firstly, we take the input image from the input layer of TL architecture, later we perform the following image augmentation processing techniques as follows:

- `RandomFlip("horizontal",input_shape=(size,size,3))`: This technique randomly flips the input image horizontally with image shape of 256x256x3, where 256 is height, 256 is width and 3 is the channel for rgb image.
- `RandomRotation(0.2)`: This technique randomly rotates the input image by a maximum of 0.2 radians.
- `RandomZoom(0.2)`: This technique randomly zooms in or out of the input image by a maximum of 20%.
- `RandomContrast(0.2)`: This technique randomly adjusts the contrast of the input image by a maximum of 20%.
- `Rescaling(1.0/255)`: This technique rescales the input image pixel values to a range of 0 to 1.
- `RandomRotation(30)`: This technique randomly rotates the input image by a maximum of 30 degrees.
- `RandomTranslation(height_factor = 0.2, width_factor = 0.3, fill_mode = 'reflect', interpolation = 'bilinear')`: This technique randomly translates the input image horizontally and vertically by a maximum of 20% and 30% of the image height and width, respectively. The "fill_mode" parameter determines how the pixels outside the image boundary are filled, and the "interpolation" parameter determines how the image is interpolated after the translation.

The augmentation process was carefully designed and adjusted to ensure the augmented dataset was diverse and representative of the original dataset. The resulting augmented dataset was used to train our proposed model, which achieved state-of-the-art performance on the task of interest. Thus, these image-processing techniques can help augment the dataset's size and diversity, which can upgrade the performance of deep-learning models.

3.5. Fine-tuning

In Fine-tuning, we add some layers to suit the classes of the preferred brain MRI data making a better architecture. We add Global Average Pooling2D and twice Batch Normalization and Dense layer to complete the proposed architecture. In the Global Average Pooling2D layer, we add the output of the base model after that, we add Batch Normalization and Dense layer of neuron 1280, activation of 'relu', kernel initializer of 'glorot uniform' with seed 1377 and bias initializer is 'zeros'. Then we again add another Batch Normalization and later we add a prediction Dense layer with 3 neurons which is the class number. 'softmax' activation is utilized as it is a multilabel classification (Kini et al., 2022; Thilagaraj et al., 2022), 'random uniform' kernel initializer and 'zeros' bias initializer. we use the pre-trained trainable weighted to utilize the knowledge in our architecture. Finally, the architecture is compiled with 'Adamax' optimizer with a learning rate of 0.0001 and 'sparse categorical cross' entropy with 'accuracy' metrics. The 'Adamax' is utilized as it is an Adam variant relying on the infinity norm that outperforms Adam, particularly in models with embeddings. The learning rate is 0.0001, since at this rate, the output models can ensure durability with less loss than others (Mustapha et al., 2020). We utilized 'sparse categorical cross' entropy as the output labels are in integer form (performs label encoding) and It helps in saving memory and computation time by using a single integer for a class rather than an entire vector (Kakarla et al., 2021; Andrei-Alexandru and Henrietta, 2020; Chai et al., 2022).

3.6. Transfer leaning algorithms

- Xception: The Xception architecture, also known as "Extreme Inception" (Chollet, 2017), is a convolutional neural network design consisting of a sequence of depthwise separable convolution layers with residual connections. The architecture comprises 36 convolutional layers grouped into 14 blocks, where all but the first and last blocks feature linear residual connections between them. This appears to consider the architecture very simple to interpret and customize; utilizing a top-level library such as Keras (Joseph et al., 2021) or TensorFlow-Slim (Silberman, 2017), it requires only 30 to 40 lines of code, similar to VGG-16 (Simonyan and Zisserman, 2014), but unlike architectures such as InceptionV2 or V3, which are far more difficult to delineate.
- ResNet50V2: ResNet is a novel neural network that was first invented by (He et al., 2016a). The success of this model is undeniable, as demonstrated by the fact that its ensemble was able to secure the first position in the ILSVRC 2015 classification contest, with an impressively low defect rate of only 3.57 percent. It has many varieties using the same principle but has various numbers of layers. Resnet50 is a variation that can work with 50 neural network layers. Deep residual nets employ residual blocks to enhance model accuracy. The central idea behind residual blocks, known as "skip connections," is the robustness of this neural network architecture. ResNet50V2 (He et al., 2016b) is an adapted variant of ResNet50 that demonstrates superior performance over ResNet50 and ResNet101 in the ImageNet dataset. Specifically, ResNet50V2 (Rahimzadeh and Attar, 2020) introduces a modified inter-block connection structure to improve information flow between blocks.
- InceptionResNetV2: According to (Szegedy et al., 2017), the InceptionResNetV2 architecture, based on the Inception block, extracts features using transformation and merging functions. It outperforms inceptionResNetv1 with less computation. Residual learning and inception blocks underpin it. Residual connections link

different-sized convolution filters. Residual connections avoid degradation and shorten training. This network uses Stem, InceptionResNet, and Reduction blocks to improve detection accuracy. According to (Asif et al., 2022), the deep network connects one main block, five Inception ResNetA blocks, ten ResNet-B blocks, five ResNet-C blocks, one Reduction-A block, and one Reduction-B block.

- DenseNet201: The architecture of DenseNet, as introduced in (Huang et al., 2017), utilizes feedforward connections to establish interactions between each layer and all subsequent layers. This is in contrast to traditional L-layer CNNs that only have L connections. With DenseNet, the number of direct connections between layers increases significantly to $(L(L+1))/2$, leading to improved feature propagation and gradient flow throughout the network. A feature map is included in each layer of the model. Each layer's feature map is utilized as the next layer's input. It maximizes information transfer within the network by directly connecting all layers. It significantly reduces the dimensionality, lessens gradient runaway, improves feature diffusion, and encourages feature reusability. When contrasted with traditional CNN. DenseNet needs fewer parameters since the feature map is not discovered twice. Furthermore, by using regularisation, DenseNet minimizes the possibility of overfitting. DenseNet121 is made up of four dense blocks, each with six, twelve, twenty-four, and sixteen convolution blocks (Asif et al., 2022).

4. Results and Discussion

We have reconfigured transfer learning architecture and fine-tuning by attaching some extra layers, used four transfer learning algorithms, and evaluated the performance of our proposed scheme to identify brain tumors. The performance is evaluated using a variety of performance indicators. The experiment setup, performance evaluation metrics, results analysis, and discussion are provided in the following section.

4.1. Experiment Setup

The research is carried out using a computer that is powered by an Intel Xeon CPU with 2 Cores, 13GB of RAM, a 16GB GPU, and a 73GB hard drive. With the help of a Jupyter notebook, the experiment has been conducted. Python is used to implement the proposed approach, together with a number of widely used libraries such as Scikit-learn, Keras, TensorFlow, Seaborn, Matplotlib, Numpy, and Pandas.

4.2. Performance Evaluation Metrics

Several performance indicators, such as accuracy, precision, recall, f1-score, confusion matrix, MSE, MAE, and RMSE, evaluate the performance of our proposed approach. The following are the metrics developed for evaluating performance:

- A method for evaluating the effectiveness of machine learning categorization is the confusion matrix. This table-like structure displays four different combinations of predicted and actual values, namely TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative). The confusion matrix, depicted in Table 1, uses these labels to represent correct and incorrect predictions for positive and negative values. A confusion matrix is a valuable tool for assessing accuracy, precision, recall, and f1-score in evaluating dependability, as cited by Talukder Talukder et al. (2023).

	Actual positive	Actual negative
Predicted positive	TP	FP
Predicted negative	FN	TN

Table 1: Confusion Matrix

- The most considered highly statistical is accuracy, which depends on the number of proper outcome expectations over the total number of observations.

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

- Precision is defined as the proportion of accurately predicted positive values among the total number of predicted positive values. It is visually represented as:

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

- Recall is the proportion of positively predicted values that are accurate to all other actual values. As seen, it is:

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

- The F1-score, which is a measure of performance for classification tasks, is calculated as the harmonic mean of precision and recall scores. It is typically represented in the following formula:

$$F1_{score} = 2 * \frac{(Recall * Precision)}{(Recall + Precision)} \quad (4)$$

- The mean absolute error (MAE) is a metric used to compare errors in observations that are connected and represent the same phenomenon. The arithmetic mean of the expected and actual numbers, as stated in Willmott and Matsuura (2005), is utterly inaccurate.

$$MAE = \frac{\sum_{i=1}^n Predict(i) - Actual(i)}{n} \quad (5)$$

- The mean squared error (MSE) calculates the average of the squared residuals or the average squared difference between the values that were actually observed and those that were predicted. The most crucial aspect of MSE is it's frequently completely positive (rather than zero) due to unpredictability or if the classifier does not permit data that could produce a reasonable forecast (Lehmann and Casella, 2006; Das et al., 2004).

$$MSE = \frac{\sum_{i=1}^n (Predict(i) - Actual(i))^2}{n} \quad (6)$$

- The root mean square error is the assessment measure most frequently employed in regression issues (RMSE). Its foundation is that mistake is objective and appears to have a normal distribution. The impact of outlier traits on RMSE is fairly significant.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (predict(i) - actual(i))^2}{n}} \quad (7)$$

where n is the total number of values.

- An classifier evaluation can be made with the help of the Matthews correlation coefficient (MCC). The value can be anything from -1 to 1, with -1 indicating that the expected and actual results are completely at odds with one another, 0 indicating that the predictions are completely random, and 1 indicating that the predictions are spot on.

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \quad (8)$$

- Kappa is a measurement for evaluating the degree of agreement between a classification model's anticipated and observed results, controlling for the possibility that the observed agreement is due to chance alone. It goes from -1 to 1, with -1 denoting total disagreement, 0 denoting chance agreement, and 1 denoting perfect agreement.

$$Kappa = \frac{(P_o - P_e)}{(1 - P_e)} \quad (9)$$

where P_o = observed agreement, and P_e = expected agreement.

$$P_o = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (10)$$

$$P_e = \frac{((TP + FP) * (TP + FN) + (TN + FP) * (TN + FN))}{(TP + TN + FP + FN)^2} \quad (11)$$

- The Classification Success Index (CSI) is a measurement tool that assesses the effectiveness of a classification model by counting the percentage of samples that were properly classified out of all the samples. CSI has a scale from 0 to 1, with 0 denoting incorrect classification and 1 denoting flawless classification.

$$CSI = \frac{TP}{TP + FP + FN} \quad (12)$$

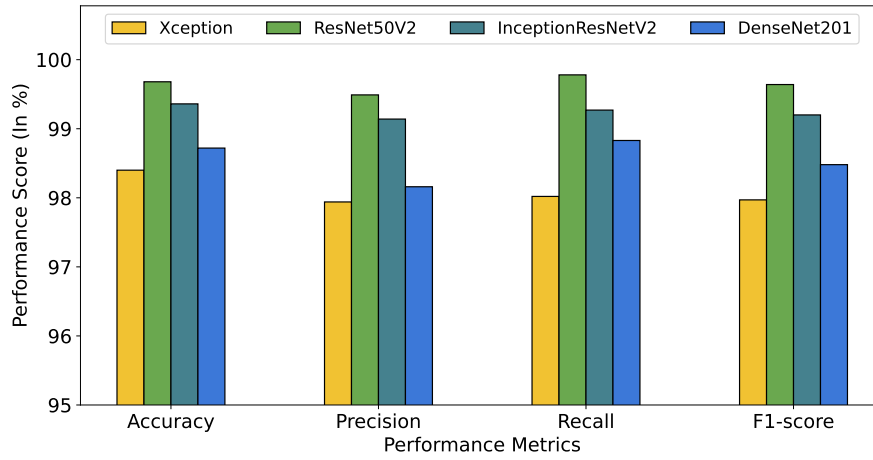
4.3. Results and Performance Analysis

Our proposed approach uses four transfer learning algorithms on the brain tumor dataset to classify brain tumors efficiently. Table 2 and Fig. 9 shows the performance and error analysis of our deployed transfer learning models. In 9(a), we can see that the accuracy rates are 98.40%, 99.68%, 99.36%, 98.72%; the precision rates are 97.94%, 99.49%, 99.14%, 98.16%; the recall rates are 98.02%, 99.78%, 99.27%, 98.83%; the f1-score rates are 97.97%, 99.64%, 99.20%, 98.48%; for Xception, ResNet50V2, InceptionResNetV2 and DenseNet201 respectively. Among all the models, ResNet50V2 achieves the highest performance rate with 99.68% accuracy, 99.49% precision, 99.27% recall and 99.20% f1-score rate.

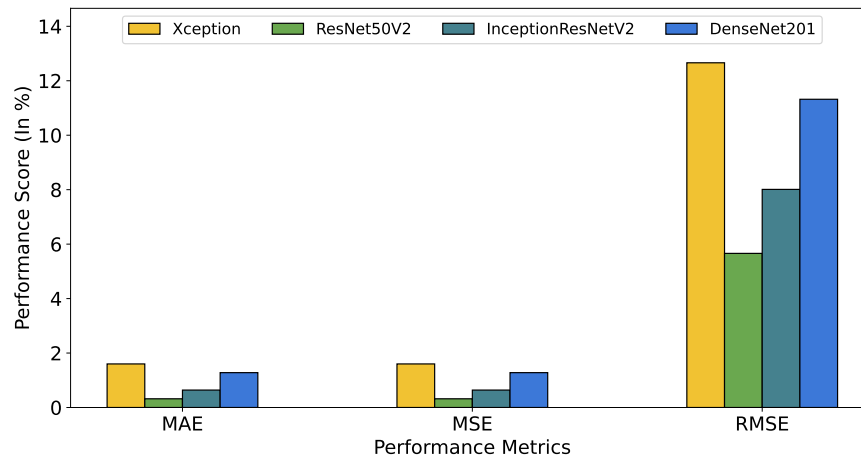
Similarly, in 9(b), we can see that the MAE rates are 1.6%, 0.32%, 0.64%, 1.28%; the MSE rates are 1.6%, 0.32%, 0.64%, 1.28%; the RMSE rates are 12.66%, 5.66%, 8.01%, 11.32%; for Xception, ResNet50V2, InceptionResNetV2

Table 2: Performance analysis of Transfer learning models

Proposed Model	Performance Metrics									
	Accuracy	Precision	Recall	F1-score	MAE	MSE	RMSE	MCC	Kappa	CSI
Xception	98.40	97.94	98.02	97.97	1.6	1.6	12.66	98.39	98.36	98.40
ResNet50V2	99.68	99.49	99.78	99.64	0.32	0.32	5.66	99.69	99.67	99.68
InceptionResNetV2	99.36	99.14	99.27	99.2	0.64	0.64	8.01	99.35	99.34	99.36
DenseNet201	98.72	98.16	98.83	98.48	1.28	1.28	11.32	98.70	98.68	98.72



(a) Performance



(b) Error

Figure 9: Analysis of Transfer learning models

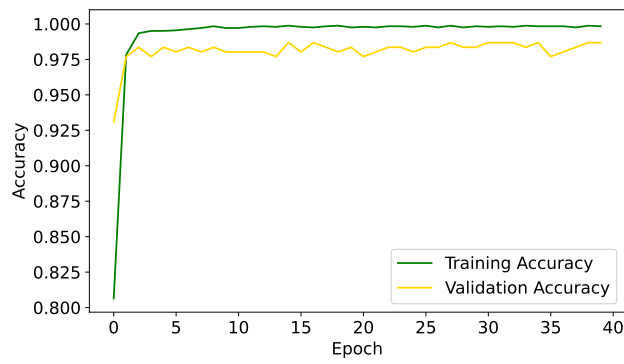
and DenseNet201 respectively. Among all the models ResNet50V2 attains the lowest error rate with 0.32% MAE, 0.32% MSE and 5.66% RMSE rate.

Fig. 10 shows all transfer learning models' accuracy and loss graphs. With the number of epochs, the accuracy rate rises while the loss rate falls. The learning curves also reveal that the models are not overfitting because the models learn the given input very well at each epoch. The augmentation procedure addresses the overfitting issue. The following is a brief analysis of all accuracy and loss graphs:

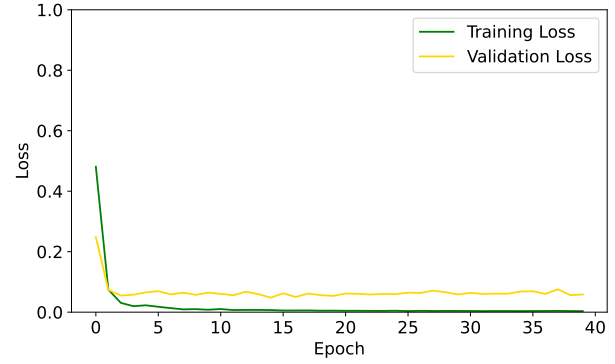
- Xception: In Fig. 10(a), we can visualize that the training and validation accuracy rate are slightly far away where the training accuracy is close to 99.5% and the testing accuracy is close to 98.5%. In Fig. 10(b), we can see that the training and validation error rate are also slightly far away, where the training accuracy is close to 0.3% and the testing accuracy is close to 1.5%.
- ResNet50V2: In Fig. 10(c), the training and validation accuracy rates are smooth and close to each other where the training accuracy is close to 99.99% and the testing accuracy is close to 99.7%. In Fig. 10(d), we can see that the training and validation error rate are also close to each other where the training accuracy is close to 0% and the testing accuracy is close to .3%.
- InceptionResNetV2: In Fig. 10(e), the training and validation accuracy rate isn't smooth where the training accuracy is close to 99.8% and the testing accuracy is close to 99.4%. In Fig. 10(f), we can see that the training and validation error rate are slightly far away where the training accuracy is close to 0.1% and the testing accuracy is close to .6%.
- DenseNet201: In Fig. 10(g), the training and validation accuracy rates are slightly close to each other and the training accuracy is close to 99.8% and the testing accuracy is close to 98.7%. In Fig. 10(h), we can see that the training and validation error rate are also slightly close where the training accuracy is close to 0.2% and the testing accuracy is close to 1.2%.

Fig. 11 represents the confusion matrix for all transfer learning models. The following is a brief description of all confusion matrices: Fig. 11(a) shows the confusion matrix of Xception model where considering the glioma TP, TN, FP, FN rates are 48.08%, 51.28%, 0%, 0.64%; considering meningioma TP, TN, FP, FN rates are 19.87%, 78.53%, 0.64%, 0.96%; and considering pituitary rates are 30.45%, 68.59%, 0.96%, 0% respectively. In Fig. 11(b), the confusion matrix of the ResNet50V2 model where considering glioma TP, TN, FP, FN rates are 48.4%, 51.28%, 0%, 0.32%; considering meningioma TP, TN, FP, FN rates are 20.83%, 78.85%, 0.32%, 0%; and considering pituitary TP, TN, FP, FN rates are 30.45%, 69.55%, 0%, 0%. Fig. 11(c) shows the confusion matrix of the InceptionResNetV2 model where considering glioma TP, TN, FP, FN rates are 48.4%, 51.28%, 0%, 0.32%; considering meningioma TP, TN, FP, FN rates are 20.51%, 78.85%, 0.32%, 0.32%; and considering pituitary TP, TN, FP, FN rates are 30.45%, 69.23%, 0.32%, 0%. Fig. 11(c) shows the confusion matrix of DenseNet201 model where considering glioma TP, TN, FP, FN rates are 47.76%, 51.28%, 0%, 0.96%; considering meningioma TP, TN, FP, FN rates are 20.51%, 78.21%, 0.96%, 0.32%; and considering pituitary TP, TN, FP, FN rates are 30.45%, 69.23%, 0.32%, 0%.

After analyzing all the performance metrics, we can conclude that among all the transfer learning models, ResNet50V2 provides the best performance and lowest error rate as well as high TP and TN rates and less number



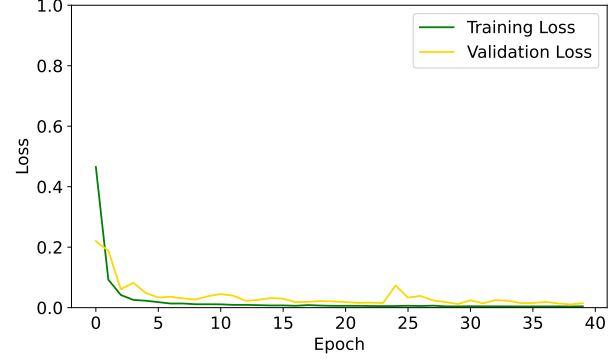
(a) Xception Accuracy



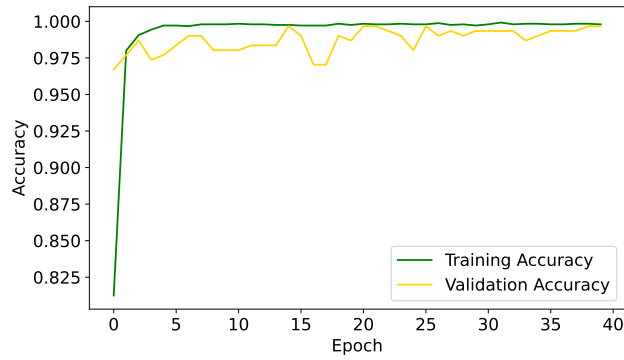
(b) Xception Loss



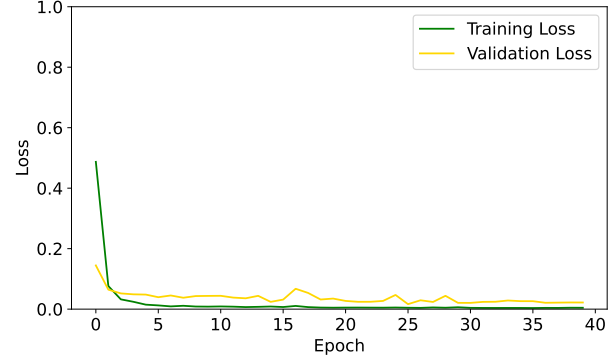
(c) ResNet50V2 Accuracy



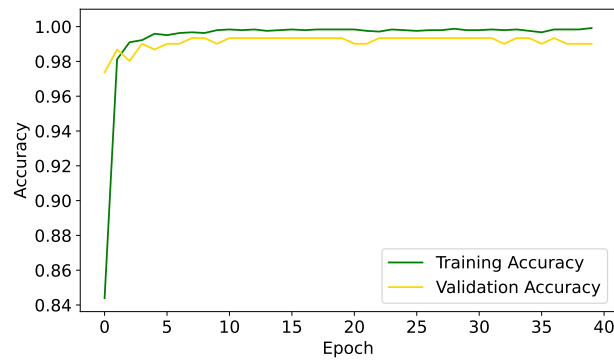
(d) ResNet50V2 Loss



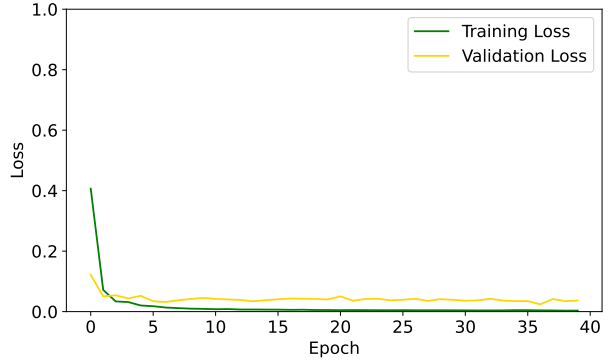
(e) InceptionResNetV2 Accuracy



(f) InceptionResNetV2 Loss



(g) DenseNet201 Accuracy



(h) DenseNet201 Loss

Figure 10: Accuracy and Loss for Transfer learning models

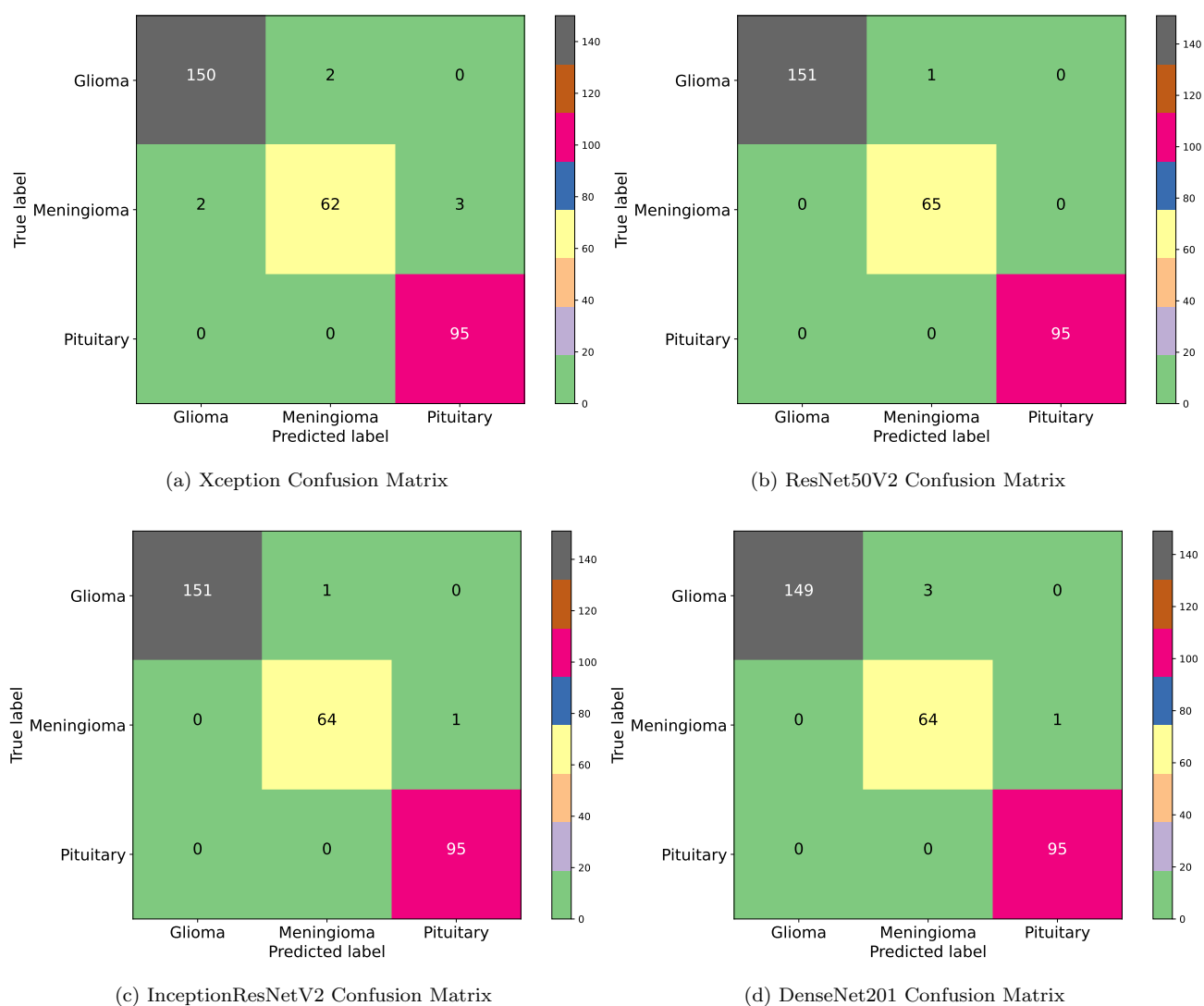


Figure 11: Confusion matrix for Transfer learning models

of FP and FN rates. ResNet50V2 (He et al., 2016b) is a revised version of ResNet50 which uses residual nets that utilize residual blocks to improve model performance. 'skip connections' is the spirit of residual blocks which render it much simpler for the layers to gain knowledge identification operations in skip connections. ResNet50v2, similar to ResNet50, increases the effectiveness of deep neural networks with more neural layers while reducing the portion of errors (Shafiq and Gu, 2022).

The analysis results reveal that ResNet50V2 achieved the highest overall performance with an accuracy of 99.68%, followed by InceptionResNetV2 with an accuracy of 99.36%, DenseNet201 with an accuracy of 98.72%, and Xception with an accuracy of 98.40%. ResNet50V2 also achieved the highest precision, recall, F1-score, MCC, Kappa, and CSI among the proposed models. Overall, the analysis indicates that ResNet50V2 is the most effective transfer learning model for the given task, followed by InceptionResNetV2, DenseNet201, and Xception.

Further, a test performance measurement analysis as depicted in Fig 12 we have also included a measure of our proposal where we provide a performance of our efficient brain tumor classification model by illustrating a visualization of performance metrics for some random images to prove the effectiveness of our model in random input images. The results of the random test images (a total of 8 images) were astonishing, with a perfect accuracy rate of 100% for each image!

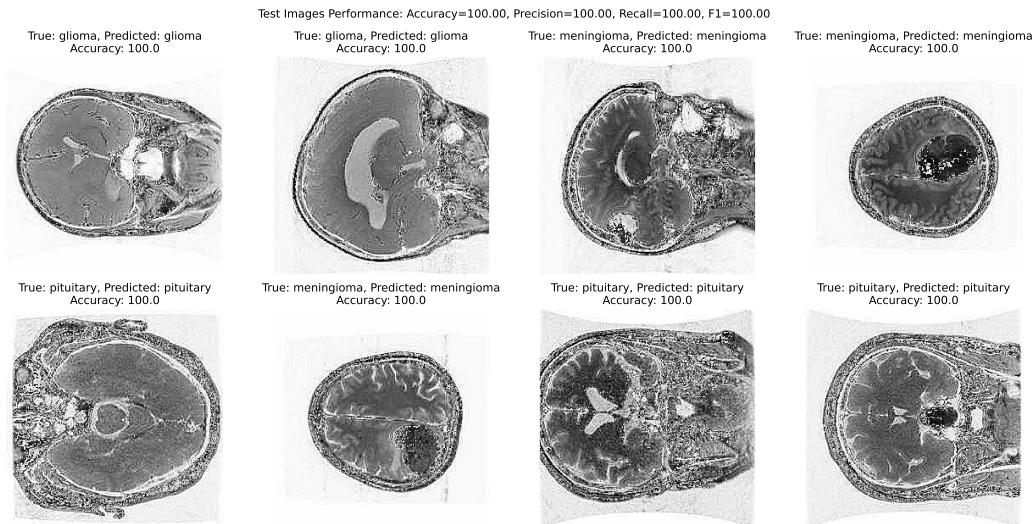


Figure 12: Performance measurements of our proposed model

4.4. Complexity analysis

We ascertain the complexity of our experiments by calculating the prediction time in seconds using an Intel Xeon CPU with 2 Cores, 13GB RAM, and a 16GB GPU. The experiments have been conducted utilizing 312 test images to classify brain tumors. The predicted times are 19s, 16s, 27s, 23s for Xception, ResNet50V2, InceptionResNetV2 and DenseNet201 respectively. Among all the models, our proposal ResNet50V2 takes only 16s which underperforms others. Table 3 and Fig 13 show the analysis of prediction time in tabular and graphical format.

Proposed Model	Prediction Speed (In sec)
Xception	19
ResNet50V2	16
InceptionResNetV2	27
DenseNet201	23

Table 3: Prediction speed analysis

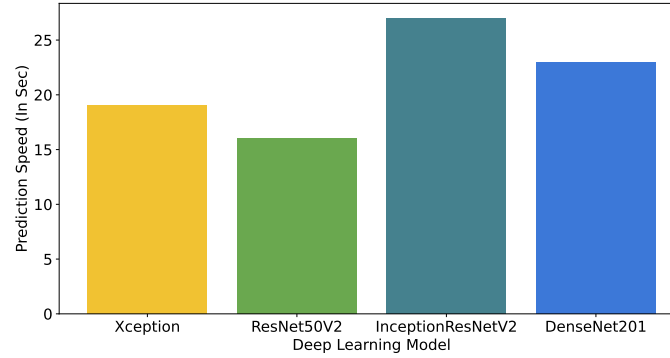


Figure 13: Prediction speed analysis

4.5. Discussion

In our proposal, we have conducted preprocessing, architecture reconfiguration as well as fine-tuning by attaching some extra layers to build a novel model to detect brain tumors efficiently. Our proposed efficient brain tumor classification model has been evaluated against other existing works and significantly achieved a notably superior level of accuracy. Table 4 shows the comparison study of brain tumor classification. Among all the accuracy performances our proposal provides the greatest accuracy with the same (brain tumor (Cheng, 2017)) dataset as well as the same number of images (3064). The use of the same dataset and number of images allows for a fair comparison to predict efficiency. The study demonstrates that by using RestNet50V2, we can reach the highest 99.68 percent accuracy rate for the brain tumor dataset that significantly outperforms others.

There are several methods available for categorizing brain tumors, each with its advantages and disadvantages. Here, we compare the following approaches: Transfer Learning (TL) Belaid and Loudini (2020); Rehman et al. (2020); Tummala et al. (2022), Capsule Network (CapsNet) Afshar et al. (2020, 2019) and Convolution Neural Network (CNN) Ait Amou et al. (2022); Ayadi et al. (2021); Badža and Barjaktarović (2020) with our proposed approach. Table 5 illustrates the advantages and disadvantages of other approaches.

Besides, the proposed research can be applied to various classification tasks related to neural networks, CNN, and deep learning beyond brain tumor classification. The techniques and methods proposed in the paper, such as preprocessing, reconstructing transfer learning architecture, and fine-tuning, can be applied to various classification tasks in medical imaging and beyond. For example, the proposed model can be adapted for the detection of COVID-19 in medical imaging data, as demonstrated in the papers Irfan et al. (2021); Almalki et al. (2021). The application of deep learning models in the diagnosis of COVID-19 has become an important area of research during the pandemic and the techniques proposed in the paper can be utilized to improve the accuracy of diagnosis and facilitate timely

SI. NO	Author	Architecture	Dataset	No. of Images	Accuracy (In %)
1	Belaïd and Loudini (2020)	VGG16	Brain tumor (Cheng, 2017)	3064	96.5
2	Rehman et al. (2020)	Fine-tune VGG16	Brain tumor (Cheng, 2017)	3064	98.69
3	Badža and Barjaktarović (2020)	CNN	Brain tumor (Cheng, 2017)	3064	96.56
4	Sultan et al. (2019)	DL	Brain tumor (Cheng, 2017)	3064	98.7
5	Ait Amou et al. (2022)	Optimized CNN	Brain tumor (Cheng, 2017)	3064	98.7
6	Tummala et al. (2022)	Ensemble ViT	Brain tumor (Cheng, 2017)	3064	98.7
7	Abiwinanda et al. (2019)	CNN	Brain tumor (Cheng, 2017)	3064	84.19
8	Paul et al. (2017)	CNN	Brain tumor (Cheng, 2017)	3064	91.43
9	Das et al. (2019)	CNN	Brain tumor (Cheng, 2017)	3064	94.39
10	Afshar et al. (2018)	CapsNet	Brain tumor (Cheng, 2017)	3064	86.56
11	Afshar et al. (2019)	CapsNet	Brain tumor (Cheng, 2017)	3064	90.89
12	Afshar et al. (2020)	BayesCap	Brain tumor (Cheng, 2017)	3064	78
13	Swati et al. (2019)	VGG19	Brain tumor (Cheng, 2017)	3064	94.82
14	Sadad et al. (2021)	NASNet	Brain tumor (Cheng, 2017)	3064	99.6
15	Ayadi et al. (2021)	CNN	Brain tumor (Cheng, 2017)	3064	94.74
16	Our proposal	DL (ResNet50V2)	Brain tumor (Cheng, 2017)	3064	99.68

Table 4: The comparison analysis of brain tumor classification

treatment. Additionally, the proposed model can be applied to fault diagnosis in industrial settings, as shown in the paper Glowacz (2022). By utilizing deep learning models in conjunction with thermal imaging techniques, the proposed model can facilitate the early detection and diagnosis of faults in machinery, leading to reduced downtime and increased productivity. Overall, the proposed research can serve as a valuable basis for various classification tasks involving medical imaging, fault diagnosis, and other applications that require deep learning models.

4.6. Application and Profitable Implications on Society

The proposed research aims to develop a fine-tuned deep-learning model for brain tumor classification, leveraging the power of deep learning to identify different types of brain tumors accurately. The potential applications of this research are numerous and could have a significant clinical impact in neuro-oncology. Figure 14 illustrates the potential application of our proposed brain tumor classification research.

- **Improved Brain Tumor Diagnosis:** The proposed research can assist radiologists and clinicians in accurately diagnosing brain tumors from medical imaging data, such as MRI scans. It provides accurate and consistent tumor classification results, reducing the risk of misdiagnosis and enabling early detection of brain tumors.
- **Personalized Treatment Planning:** Accurate classification of brain tumors can help tailor personalized treatment plans for patients. The model aids in identifying the specific tumor type, enabling clinicians to design targeted treatment plans, leading to more effective and precise treatment outcomes.
- **Clinical Decision Support System:** The developed model can serve as a clinical decision support system, assisting healthcare providers in making informed decisions about patient management. It provides accurate

Model	Description	Advantages	Disadvantages
Transfer Learning (TL)	Uses a pre-trained model as a starting point for training a new model on a related task	Reduces the amount of data needed for training, speed up the training process, can improve accuracy by using pre-trained models trained on large datasets	Difficult to find a pre-trained model suitable for the specific task of brain tumor classification, the model may not be optimized for the specific features of brain tumor images
Capsule Network (CapsNet)	Uses capsules to represent the features of an image, can handle variations in pose and deformation, can handle multiple viewpoints of the same object	Can improve accuracy for analyzing brain tumor images	Can be computationally expensive to train, may require a large amount of data for training
Convolutional Neural Network (CNN)	Commonly used for image classification tasks, can handle spatial relationships between pixels in an image, can identify features such as edges and textures, can identify complex patterns in an image	Can improve the accuracy of classification for brain tumor images	Can require a large amount of data for training, can be computationally expensive
Proposed Model	Utilizes extensive preprocessing, reconfiguration of transfer learning architecture, and fine-tuning for classification of brain tumors	Efficient utilizes image augmentation to solve overfitting problems and utilize GPU speed, can standardize images based on configuration, allows building of new DL architecture	Require advanced techniques in image processing, and the suggested framework would enhance the appropriateness of the task.

Table 5: Advantages and disadvantages of other approaches

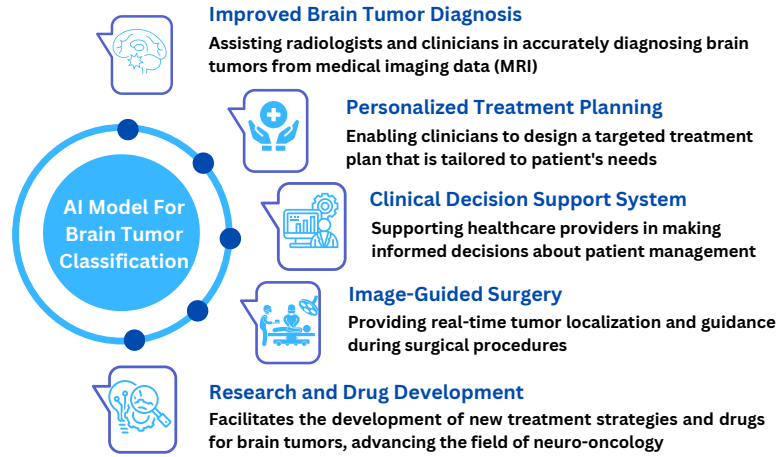


Figure 14: Application of proposed research

and reliable tumor classification results, aiding clinical decision-making for more precise and personalized patient care.

- **Image-Guided Surgery:** The proposed model can be integrated into image-guided surgery systems, providing real-time tumor localization and guidance during surgical procedures. This helps neurosurgeons accurately identify tumor margins, optimize tumor resection, and minimize damage to healthy brain tissue, improving surgical outcomes.
- **Research and Drug Development:** The model can aid brain tumor research and development by accurately classifying tumors and identifying specific genetic mutations or molecular characteristics associated with different tumor types. This provides insights into tumor biology and potential therapeutic targets, facilitating the development of new treatment strategies and drugs for brain tumors.

Hence, the proposed research has significant potential applications in improving brain tumor diagnosis, guiding personalized treatment planning, facilitating research and drug development, image-guided surgery, and serving as a clinical decision support system. These applications can positively impact patient care, outcomes, and neuro-oncology advancement.

Furthermore, our innovative research has the potential to reshape the medical imaging field, providing significant benefits to society. Our proposed deep learning model accurately classifies brain tumors and employs creative reconstruction and fine-tuning methods. The model's direct clinical relevance and easy integration into existing workflows allow for smooth implementation in clinical settings, enhancing brain tumor diagnosis, treatment planning, and patient outcomes. Our approach also reduces the need for extensive manual annotation, lowering costs and development time and making it more suitable for clinical use. The model's capacity for personalized medicine is groundbreaking, as tailored treatment plans based on tumor characteristics lead to more effective therapies, improved results, and better patient well-being. In addition to clinical impact, our research holds significant societal implications by improving patient care, reducing healthcare costs, and addressing global healthcare system challenges. Overall, our model's innovative reconstruction and fine-tuning techniques contribute to deep learning

research and inspire further progress in medical imaging across various applications.

5. Conclusion

This article presents a novel deep learning (DL) method for classifying brain tumours that combine preprocessing, transfer learning (TL) architecture reconstruction, and fine-tuning. Four TL algorithms were utilized in our methodology: Xception, ResNet50V4, InceptionResNetV4, and DenseNet201. We evaluated the performance of the model using a variety of metrics, including accuracy, recall, precision, f1 score, MAE, MSE, and RMSE, to demonstrate the substantial progress made. Using the Figshare Brain Tumor Image dataset, we demonstrated that our proposed model is highly effective in accurately diagnosing brain tumors. The classification accuracy for brain tumors was 98.40% for Xception, 99.68% for ResNet50V4, 99.36% for InceptionResNetV2, and 98.72% for DenseNet201. Further investigation revealed that ResNet50V2 outperforms other models and existing methods in terms of precision. We hope that our developed model can be implemented in clinical settings to facilitate faster and more accurate diagnosis of brain tumours. Despite the architecture's enhanced precision, further advances in image processing and the proposed architecture could improve its suitability for this task. The absence of clearer images and an improved DL architecture is the primary limitation of this study, which hinders the achievement of even higher performance results.

The following statements about what has been accomplished in the paper can be made in light of the conclusion:

- Developed a novel DL approach for brain tumor classification that integrates preprocessing, reconstruction of TL architectures, and fine-tuning. Employed four TL algorithms, namely Xception, ResNet50V2, InceptionResNetV2, and DenseNet201 in the approach.
- Conducted an extensive experiment using the Figshare MRI brain tumor image dataset. Multiple performance metrics, such as precision, precision, recall, f1 score, confusion matrix, root mean square error, mean absolute error, and mean squared error, were used to evaluate the effectiveness of the suggested method.
- Found that ResNet50V2 provides better accuracy than other models and existing works for brain tumor classification.
- Identified the insufficiency of the work, which is the lack of more clear images with improved DL architecture that resist getting higher performance outcomes.

In the future, we intend to improve our proposed DL model by adopting more advanced hybrid ensemble techniques with newly available brain tumor datasets. Moreover, we desire to employ explainable AI techniques to provide greater insight into the decision-making process of our DL model, thus improving the confidence and trust of clinicians and patients in diagnosis.

Acknowledgments

This work is partially supported by Ministry of Science and Technology, NST Fellowship Program.

References

- Abbasi, A. A., Hussain, L., Awan, I. A., Abbasi, I., Majid, A., Nadeem, M. S. A., and Chaudhary, Q.-A. (2020). Detecting prostate cancer using deep learning convolution neural network with transfer learning approach. *Cognitive Neurodynamics*, 14(4):523–533.
- Abd-Ellah, M. K., Awad, A. I., Khalaf, A. A., and Hamed, H. F. (2019). A review on brain tumor diagnosis from mri images: Practical implications, key achievements, and lessons learned. *Magnetic resonance imaging*, 61:300–318.
- Abiwinanda, N., Hanif, M., Hesaputra, S. T., Handayani, A., and Mengko, T. R. (2019). Brain tumor classification using convolutional neural network. In *World congress on medical physics and biomedical engineering 2018*, pages 183–189. Springer.
- Afshar, P., Mohammadi, A., and Plataniotis, K. N. (2018). Brain tumor type classification via capsule networks. In *2018 25th IEEE international conference on image processing (ICIP)*, pages 3129–3133. IEEE.
- Afshar, P., Mohammadi, A., and Plataniotis, K. N. (2020). Bayescap: a bayesian approach to brain tumor classification using capsule networks. *IEEE Signal Processing Letters*, 27:2024–2028.
- Afshar, P., Plataniotis, K. N., and Mohammadi, A. (2019). Capsule networks for brain tumor classification based on mri images and coarse tumor boundaries. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1368–1372. IEEE.
- Ait Amou, M., Xia, K., Kamhi, S., and Mouhafid, M. (2022). A novel mri diagnosis method for brain tumor classification based on cnn and bayesian optimization. In *Healthcare*, volume 10, page 494. MDPI.
- Almadhoun, H. R. and Abu-Naser, S. S. (2022). Detection of brain tumor using deep learning. *International Journal of Academic Engineering Research (IJAER)*, 6(3).
- Almalki, Y. E., Qayyum, A., Irfan, M., Haider, N., Glowacz, A., Alshehri, F. M., Alduraibi, S. K., Alshamrani, K., Alkhalik Basha, M. A., Alduraibi, A., et al. (2021). A novel method for covid-19 diagnosis using artificial intelligence in chest x-ray images. In *Healthcare*, volume 9, page 522. MDPI.
- Andrei-Alexandru, T. and Henrietta, D. E. (2020). Low cost defect detection using a deep convolutional neural network. In *2020 IEEE International conference on automation, quality and testing, robotics (AQTR)*, pages 1–5. IEEE.
- Andresen, J., Kepp, T., Ehrhardt, J., Burchard, C. v. d., Roeder, J., and Handels, H. (2022). Deep learning-based simultaneous registration and unsupervised non-correspondence segmentation of medical images with pathologies. *International Journal of Computer Assisted Radiology and Surgery*, 17(4):699–710.
- Asif, S., Yi, W., Ain, Q. U., Hou, J., Yi, T., and Si, J. (2022). Improving effectiveness of different deep transfer learning-based models for detecting brain tumors from mr images. *IEEE Access*, 10:34716–34730.

- Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., and Inman, D. J. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. *Mechanical systems and signal processing*, 147:107077.
- Ayadi, W., Elhamzi, W., Charfi, I., and Atri, M. (2021). Deep cnn for brain tumor classification. *Neural Processing Letters*, 53(1):671–700.
- Badža, M. M. and Barjaktarović, M. Č. (2020). Classification of brain tumors from mri images using a convolutional neural network. *Applied Sciences*, 10(6):1999.
- Belaid, O. N. and Loudini, M. (2020). Classification of brain tumor by combination of pre-trained vgg16 cnn. *Journal of Information Technology Management*, 12(2):13–25.
- Bruton, S. V., Medlin, M., Brown, M., and Sacco, D. F. (2020). Personal motivations and systemic incentives: Scientists on questionable research practices. *Science and Engineering Ethics*, 26(3):1531–1547.
- Chai, X., Nie, W., Lin, K., Tang, G., Yang, T., Yu, J., and Cao, W. (2022). An open-source package for deep-learning-based seismic facies classification: Benchmarking experiments on the seg 2020 open data. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–19.
- Cheng, J. (2017). Brain magnetic resonance imaging tumor dataset. *Figshare MRI Dataset Version*, 5.
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258.
- Çinar, A. and Yildirim, M. (2020). Detection of tumors on brain mri images using the hybrid convolutional neural network architecture. *Medical hypotheses*, 139:109684.
- Ciregan, D., Meier, U., and Schmidhuber, J. (2012). Multi-column deep neural networks for image classification. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3642–3649. IEEE.
- Das, K., Jiang, J., and Rao, J. (2004). Mean squared error of empirical predictor. *The Annals of Statistics*, 32(2):818–840.
- Das, S., Aranya, O. R. R., and Labiba, N. N. (2019). Brain tumor classification using convolutional neural network. In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, pages 1–5. IEEE.
- Glowacz, A. (2022). Thermographic fault diagnosis of shaft of bldc motor. *Sensors*, 22(21):8537.
- Graber, J. J., Cobbs, C. S., and Olson, J. J. (2019). Congress of neurological surgeons systematic review and evidence-based guidelines on the use of stereotactic radiosurgery in the treatment of adults with metastatic brain tumors. *Neurosurgery*, 84(3):E168–E170.
- Gurbină, M., Lascu, M., and Lascu, D. (2019). Tumor detection and classification of mri brain image using different wavelet transforms and support vector machines. In *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*, pages 505–508. IEEE.

- He, K., Zhang, X., Ren, S., and Sun, J. (2016a). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016b). Identity mappings in deep residual networks. In *European conference on computer vision*, pages 630–645. Springer.
- Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.
- Irfan, M., Iftikhar, M. A., Yasin, S., Draz, U., Ali, T., Hussain, S., Bukhari, S., Alwadie, A. S., Rahman, S., Glowacz, A., et al. (2021). Role of hybrid deep neural networks (hdnns), computed tomography, and chest x-rays for the detection of covid-19. *International Journal of Environmental Research and Public Health*, 18(6):3056.
- Islam, K. T., Wijewickrema, S., and O’Leary, S. (2022). A deep learning framework for segmenting brain tumors using mri and synthetically generated ct images. *Sensors*, 22(2):523.
- Islam, M. K., Ali, M. S., Miah, M. S., Rahman, M. M., Alam, M. S., and Hossain, M. A. (2021). Brain tumor detection in mr image using superpixels, principal component analysis and template based k-means clustering algorithm. *Machine Learning with Applications*, 5:100044.
- Joseph, F. J. J., Nonsiri, S., and Monsakul, A. (2021). Keras and tensorflow: A hands-on experience. *Advanced Deep Learning for Engineers and Scientists: A Practical Approach*, pages 85–111.
- Kakarla, J., Isunuri, B. V., Doppalapudi, K. S., and Bylapudi, K. S. R. (2021). Three-class classification of brain magnetic resonance images using average-pooling convolutional neural network. *International Journal of Imaging Systems and Technology*, 31(3):1731–1740.
- Katti, G., Ara, S. A., and Shireen, A. (2011). Magnetic resonance imaging (mri)–a review. *International journal of dental clinics*, 3(1):65–70.
- Khan, A. H., Abbas, S., Khan, M. A., Farooq, U., Khan, W. A., Siddiqui, S. Y., and Ahmad, A. (2022). Intelligent model for brain tumor identification using deep learning. *Applied Computational Intelligence and Soft Computing*, 2022.
- Kibriya, H., Amin, R., Alshehri, A. H., Masood, M., Alshamrani, S. S., and Alshehri, A. (2022). A novel and effective brain tumor classification model using deep feature fusion and famous machine learning classifiers. *Computational Intelligence and Neuroscience*, 2022.
- Kini, A. S., Gopal Reddy, A. N., Kaur, M., Satheesh, S., Singh, J., Martinetz, T., and Alshazly, H. (2022). Ensemble deep learning and internet of things-based automated covid-19 diagnosis framework. *Contrast Media & Molecular Imaging*, 2022.
- Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., and Inman, D. J. (2021). 1d convolutional neural networks and applications: A survey. *Mechanical systems and signal processing*, 151:107398.

- Komninos, J., Vlassopoulou, V., Protopapa, D., Korfias, S., Kontogeorgos, G., Sakas, D. E., and Thalassinou, N. C. (2004). Tumors metastatic to the pituitary gland: case report and literature review. *The Journal of Clinical Endocrinology & Metabolism*, 89(2):574–580.
- Le, E., Wang, Y., Huang, Y., Hickman, S., and Gilbert, F. (2019). Artificial intelligence in breast imaging. *Clinical radiology*, 74(5):357–366.
- Lehmann, E. L. and Casella, G. (2006). *Theory of point estimation*. Springer Science & Business Media.
- Mandle, A. K., Sahu, S. P., and Gupta, G. (2022). Intelligent brain tumor detection system using deep learning technique. In *2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T)*, pages 1–6. IEEE.
- Montalbo, F. J. P. (2020). A computer-aided diagnosis of brain tumors using a fine-tuned yolo-based model with transfer learning. *KSII Transactions on Internet and Information Systems (TIIS)*, 14(12):4816–4834.
- Morid, M. A., Borjali, A., and Del Fiol, G. (2021). A scoping review of transfer learning research on medical image analysis using imagenet. *Computers in biology and medicine*, 128:104115.
- Mustapha, A., Mohamed, L., and Ali, K. (2020). An overview of gradient descent algorithm optimization in machine learning: Application in the ophthalmology field. In *International Conference on Smart Applications and Data Analysis*, pages 349–359. Springer.
- Naki, T. and Aderibigbe, B. A. (2022). Efficacy of polymer-based nanomedicine for the treatment of brain cancer. *Pharmaceutics*, 14(5):1048.
- Naseer, A., Yasir, T., Azhar, A., Shakeel, T., and Zafar, K. (2021). Computer-aided brain tumor diagnosis: performance evaluation of deep learner cnn using augmented brain mri. *International Journal of Biomedical Imaging*, 2021.
- Naser, M. A. and Deen, M. J. (2020). Brain tumor segmentation and grading of lower-grade glioma using deep learning in mri images. *Computers in biology and medicine*, 121:103758.
- Niraj, L. K., Patthi, B., Singla, A., Gupta, R., Ali, I., Dhama, K., Kumar, J. K., and Prasad, M. (2016). Mri in dentistry-a future towards radiation free imaging–systematic review. *Journal of clinical and diagnostic research: JCDR*, 10(10):ZE14.
- Paul, J. S., Plassard, A. J., Landman, B. A., and Fabbri, D. (2017). Deep learning for brain tumor classification. In *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 10137, pages 253–268. SPIE.
- Pyrkov, T. V., Slipensky, K., Barg, M., Kondrashin, A., Zhurov, B., Zenin, A., Pyatnitskiy, M., Menshikov, L., Markov, S., and Fedichev, P. O. (2018). Extracting biological age from biomedical data via deep learning: too much of a good thing? *Scientific reports*, 8(1):1–11.

- Quader, S., Kataoka, K., and Cabral, H. (2022). Nanomedicine for brain cancer. *Advanced Drug Delivery Reviews*, page 114115.
- Rahimzadeh, M. and Attar, A. (2020). A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2. *Informatics in medicine unlocked*, 19:100360.
- Rehman, A., Naz, S., Razzak, M. I., Akram, F., and Imran, M. (2020). A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circuits, Systems, and Signal Processing*, 39(2):757–775.
- Sadad, T., Rehman, A., Munir, A., Saba, T., Tariq, U., Ayesha, N., and Abbasi, R. (2021). Brain tumor detection and multi-classification using advanced deep learning techniques. *Microscopy Research and Technique*, 84(6):1296–1308.
- Savaş, S. (2022). Detecting the stages of alzheimer’s disease with pre-trained deep learning architectures. *Arabian Journal for Science and Engineering*, 47(2):2201–2218.
- Shafiq, M. and Gu, Z. (2022). Deep residual learning for image recognition: a survey. *Applied Sciences*, 12(18):8972.
- Silberman, N. (2017). Tf-slim: A lightweight library for defining, training and evaluating complex models in tensorflow.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sultan, H. H., Salem, N. M., and Al-Atabany, W. (2019). Multi-classification of brain tumor images using deep neural network. *IEEE Access*, 7:69215–69225.
- Swati, Z. N. K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., and Lu, J. (2019). Brain tumor classification for mr images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, 75:34–46.
- Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*.
- Talo, M., Baloglu, U. B., Yıldırım, Ö., and Acharya, U. R. (2019). Application of deep transfer learning for automated brain abnormality classification using mr images. *Cognitive Systems Research*, 54:176–188.
- Talukder, M. A., Hasan, K. F., Islam, M. M., Uddin, M. A., Akhter, A., Yousuf, M. A., Alharbi, F., and Moni, M. A. (2023). A dependable hybrid machine learning model for network intrusion detection. *Journal of Information Security and Applications*, 72:103405.
- Talukder, M. A., Islam, M. M., Uddin, M. A., Akhter, A., Hasan, K. F., and Moni, M. A. (2022). Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning. *Expert Systems with Applications*, page 117695.

- Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., and Liu, C. (2018). A survey on deep transfer learning. In *International conference on artificial neural networks*, pages 270–279. Springer.
- Thilagaraj, M., Arunkumar, N., and Govindan, P. (2022). Classification of breast cancer images by implementing improved dcnn with artificial fish school model. *Computational Intelligence and Neuroscience*, 2022.
- Tummala, S., Kadry, S., Bukhari, S. A. C., and Rauf, H. T. (2022). Classification of brain tumor from magnetic resonance imaging using vision transformers ensembling. *Current Oncology*, 29:7498 – 7511.
- Willmott, C. J. and Matsuura, K. (2005). Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, 30(1):79–82.
- You, K., Kou, Z., Long, M., and Wang, J. (2020). Co-tuning for transfer learning. *Advances in Neural Information Processing Systems*, 33:17236–17246.