

Nile Journal of Communication & Computer Science

Volume 9, June 2025 Journal Webpage: https://njccs.journals.ekb.eg



# Improved Classification of Brain Tumors Via Fine-Tuned Transfer Learning Approaches

Shaimaa E. Nassar<sup>1\*</sup>, Hanan M. Amer<sup>2</sup>, Ibrahim Yasser<sup>2</sup>, Mohamed A. Mohamed<sup>2</sup>

<sup>1</sup> Communication and Electronics Engineering Department, Nile Higher Institute of Engineering and Technology, Mansoura 35511, Egypt. <sup>2</sup> Electronics and Communication Engineering Department, Mansoura University, Mansoura 35516, Egypt.

Corresponding author: Shaimaa E. Nassar, shaimaaelsabahy@nilehi.edu.eg.

## Abstract

Brain cancer, a perilous disease, underscores the critical need for brain tumor classification to enhance treatment outcomes and increase patient survival rates. Nevertheless, the challenging task of classifying brain tumors in their initial stages is compounded by variations in size, shape, and appearance. Deep learning (DL) gained prominence as a promising solution, particularly in the healthcare sector, utilizing brain magnetic resonance (MR) images for effective detection and classification. The prevalent use of transfer learning via fine-tuning addresses this challenge, where specific layers of a pre-trained architecture are adapted for a related target task. Despite its efficacy, selecting the optimal fine-tuning layers remains a key issue. This study presents a novel system employing a fine-tuning approach with manually chosen layers across five diverse architectures (EfficientNetV2s, Inception ResNetV2, MobileNetV2, RegNetY-320, and ConvNeXt-Large). A Global Average Pooling (GAP) layer was implemented at the output to address overfitting and vanishing gradient challenges, while a dropout layer was added to improve generalization. A comprehensive evaluation of multiple models on the BT-Large-4C dataset, which consists of 3,264 brain MRI images, shows that the fine-tuned EfficientNetV2s architecture outperforms other models. It achieved an impressive test accuracy of 97.86% while using only image resizing as the preprocessing step. Additionally, EfficientNetV2s outperforms state-of-the-art methods, making it a highly efficient and effective choice for classification of brain tumors. This study underscores the effectiveness of tailored fine-tuning in improving brain tumor classification.

*Keywords:* Deep learning (DL), Brain cancer, Transfer learning (TL), Fine-tuning, Magnetic resonance imaging (MRI).

# **1.INTRODUCTION**

The human brain is a critical organ that controls a variety of essential functions, including cognition, hearing, emotion, vision, and reflexes [1,2]. Brain tumors, which develop from abnormal growth of tissues within the skull, are extremely fatal diseases. They are classified into two categories: primary tumors which remain confined inside the brain and secondary cancers that originate in other organs and metastasize to the brain [3]. Common forms include meningiomas, gliomas, and pituitary tumors. These tumors may lead to hormone imbalances and visual problems. Addressing these disorders poses considerable obstacles in treatment, stressing the necessity of early detection for general health [4,5]. In 2023, the American Cancer Society predicts that 24,810 people will be diagnosed with malignant tumors,

resulting in 18,990 deaths [6]. Techniques of brain tumor treatment differ based on the size and kind of the tumor. It may be challenging to identify brain tumors due to their varying sizes. Precisely evaluating a brain tumor's size and resolution in its early stages can be challenging, limiting detection efforts. However, early detection greatly improves patient survival rates. Brain tumors are more difficult to diagnose when compared to cancers in other regions of the body. The existence of the blood-brain barrier complicates the detection of tumor cell activity, as traditional radioactivity markers are useless in this condition [7].

Recent imaging techniques have achieved significant progress in diagnosing serious human illnesses such as skin cancer, brain tumors, and stomach cancer [7-9]. Among these, MRI and computed tomography (CT) scans are widely regarded as the most effective tools for detecting brain malignancies [10]. MRI, in particular, is preferred over CT due to its ability to capture detailed information about tumor texture and shape, offering high-resolution, non-invasive images of brain tissue. Despite this, the complex nature of brain tumor tissue and the inherent variability in patient images often present significant challenges for manual interpretation. These challenges highlight the need for more advanced methods to improve diagnostic accuracy and reduce human error. In response, artificial intelligence (AI) and deep learning (DL) techniques have emerged as transformative tools in medical image analysis. These AIdriven methods not only expedite the diagnostic process but also enhance accuracy by automatically detecting intricate patterns that may elude human expertise, offering a more reliable and efficient approach to brain tumor diagnosis. Among the AI techniques, transfer learning (TL) has gained considerable attention for its ability to leverage pre-trained models, reducing the need for large, labeled datasets and computational resources [11,12]. TL allows models that have already been trained on large datasets to be adapted for specific tasks, such as brain tumor classification, enhancing both the speed and accuracy of diagnoses. Building on recent advancements in artificial intelligence (AI), deep learning (DL), and finetuning pre-trained models specifically for brain tumor classification allows for further optimization by adapting these models to the unique characteristics of brain MRI scans. By adjusting specific layers and incorporating custom features such as Global Average Pooling (GAP) and dropout, classification accuracy can be significantly improved, ensuring more reliable diagnostic outcomes [13-15]. Transfer learning (TL) further enhances this process by leveraging knowledge from pre-trained models to new, task-specific data, making these systems particularly valuable in medical diagnostics. Early and accurate detection is essential for effective treatment, and this study aims to improve brain tumor classification by utilizing TL to optimize model performance for this critical application [16]. Incorporating a Global Average Pooling (GAP) layer improves training efficiency and model accuracy by mitigating overfitting and addressing vanishing gradient issues [17],[18]. Additionally, the inclusion of a dropout layer with a rate of 0.2 is essential for enhancing the model's ability to generalize to unseen data, which is particularly important in clinical applications where the model must handle diverse patient cases [19]. Dense layers are also added to further boost performance. This combination of GAP, dropout, and dense layers not only helps prevent overfitting but also significantly enhances accuracy compared to the model's baseline performance without these enhancements.

This study focuses on improving brain tumor classification by first applying transfer learning to five pre-trained architectures: EfficientNetV2 [20], Inception ResNetV2 [21], MobileNetV2 [22], RegNetY-320 [23], and ConvNeXt-Large [24]. Fine-tuning these models with minimal preprocessing, freezing particular layers, and incorporating new ones, including GAP and dropout layers, helped prevent overfitting and improved model generalization. The effectiveness of these fine-tuned networks was evaluated using the BT-large-4c MRI brain tumor dataset, which includes meningioma, glioma, pituitary tumors, and normal tissue. Standard evaluation metrics such as precision, accuracy, F1-score, recall, and the confusion matrix, were employed to evaluate performance. Results indicate that fine-tuned networks offer significant improvements in classifying different types of brain tumors, demonstrating the potential of TL and fine-tuning approaches in medical diagnostics [25].

In summary, the significant contributions of the paper are outlined as follows:

• An approach for the automatic classification of brain tumors and healthy brain tissue was developed using transfer learning followed by fine-tuning on five pre-trained architectures. Using the fine-tuning method, important features were effectively extracted from the BT-large-4c dataset with minimal preprocessing, achieving high accuracy;

• EfficientNetV2s, when fine-tuned, achieved the highest accuracy among the architectures, demonstrating its efficiency and capability to deliver excellent performance with fewer computational resources. This makes it suitable for applications requiring high efficiency;

• Fine-tuning with layers such as a Global Average Pooling (GAP) layer and a 0.2 dropout layer was used to prevent overfitting and address vanishing gradient issues, enhancing the architecture's ability to generalize to unseen data;

• Comprehensive evaluation of the models on the BT-large-4c dataset was conducted using various metrics, including accuracy, accuracy curve, loss curve, F1-score, confusion matrix, sensitivity, and specificity.

The subsequent sections of the research article are organized as follows: Section 2 conducts a comparative analysis of prior related work. Section 3 outlines the pre-trained architectures and transfer learning with fine-tuning. Section 4 provides a comprehensive description of the dataset utilized in the experiment, as well as the evaluation metrics, training procedure, results, and discussion. Finally, Section 5 presents the conclusion.

## 2. RELATED WORKS

The problem of brain tumor detection and classification has been a prominent focus in research for the past two decades due to its significant medical implications. Timely identification, diagnosis, and classification of brain tumors are crucial for effective treatment planning, which contributes to patient recovery and life extension. Traditional neural networks are enhanced by deep learning models through the incorporation of additional hidden layers between the input and output layers, enabling the establishment of more complex and nonlinear relationships. Medical imaging applications extensively utilize various deep learning models, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep neural networks (DNNs). This section provides a summary of various existing studies in the field of brain tumor classification that leverage deep learning approaches. Kujur et al. [26] introduced a tumor classification methodology that combines CNN and the genetic algorithm (GA). The proposed method achieved an accuracy of 90.9% in the classification of three Glioma grades. Munira et al. [27] utilized preprocessing techniques, including thresholding, resizing, cropping, and rescaling. They created a novel 23-layer CNN architecture designed for extracting deep features from brain MRI images. The features extracted from the flattened layer of the CNN model underwent evaluation using random forest (RF) and support vector machine (SVM) classifiers. They employed CNN, CNN-SVM, CNN-RF, and fine-tuned Inception V3 deep learning models for multi-class brain MRI datasets. The hybrid approach proposed in this research is tested on two publicly accessible datasets. Out of the four models examined, the CNN-RF model attains an accuracy of 96.52% on the Figshare 3C dataset, as the CNN-SVM model records a 95.41% accuracy on the BT-large-4c dataset, which includes four categories (glioma, normal, meningioma, and pituitary).

Hossain et al. [28] incorporated various pre-trained models in their research, including VGG16, VGG19, InceptionV3, ResNet50, Xception, InceptionResNetV2, and a model denoted as IVX16. IVX16 was formulated by combining the top-performing transfer learning models: VGG16, InceptionV3, and Xception. Their preprocessing steps involved applying data augmentation techniques, including horizontal flipping, rescaling, zooming, and shearing. These techniques effectively increased the dataset size from 3264 to 13056 images. The experiments resulted in peak accuracy scores of 95.11% for VGG16, 94.19% for VGG19, 93.88% for InceptionV3, 93.88% for ResNet50, 94.5% for Xception, 93.58% for

InceptionResNetV2, and an impressive 96.94% for IVX16. Evaluations were conducted on a dataset divided into: 80% for training, 10% for validation, and 10% for testing. The primary goal was to classify four specific categories: glioma, pituitary, no tumor, and, meningioma sourced from the BT-large-4c dataset.

Vankdothu et al. [29] introduced a model termed CNN-LSTM, which combined a convolutional neural network (CNN) via a long short-term memory (LSTM). They achieved accuracy scores of 80% and 92% through two distinct dataset splits: one with 80% for training and 20% for testing, and the other with 90%, 10% for training and testing, respectively. The primary objective was to classify four specific categories—glioma, no tumor, meningioma, and pituitary—derived from the BT-large-4c dataset. Xiao et al. [30] introduced the Dual Suppression and Enhancement (DSE) block, which employed two distinct methods to identify precise features that captured the distinctions among similar classes. One approach involved refining global, orderless features to enhance valuable clues by mitigating the influence of negative ones. They focused on redirecting attention to salient regions derived from local spatial features. Subsequently, a fusion layer named FBE (Feature Bilinear Enhancement) was employed to generate compact and discriminative representations. They achieved an accuracy of 97.43% on the BT-large-4c dataset.

Chitnis et al. [31] introduced a methodology for accurate brain tumor classification employing a neural architecture search technique known as Learning by Self-Explanation (LeaSE). Within LeaSE, the main objective was to enable an effective neural architecture to be explored and identified by an explainer model. This was accomplished by requiring coherent explanations from the explainer. The underlying concept of LeaSE was rooted in the notion that a thorough comprehension of a subject matter by a model was necessary before it could provide lucid explanations. To formalize LeaSE, a four-level optimization framework was established, outlining four consecutive stages: a specific topic was learned by the explainer; insights on this topic were articulated by the explainer; the topic was comprehended by the audience based on the explainer's explanations; and the explainer refined its understanding of the topic based on the learning outcomes of the audience. A test accuracy of 90.6% was attained in the classification of four specific categories—glioma, no tumor, meningioma and pituitary—utilizing data from the BTlarge-4c dataset. Balaji et al. [32] applied transfer learning methodology to multiple architectures, including VGG16, MobileNet, Xception, ResNet, and EfficientNet-B0, in the study. They applied preprocessing steps that encompassed cropping and resizing, bias correction, and various denoising techniques such as Gaussian filtering, BM3D denoising, and total variation denoising. Following these steps, skull stripping was performed, and the data augmentation process was implemented for further refinement. Remarkably, the EfficientNet-B0 architecture outperformed others, achieving an accuracy of 97.61% based on the BT-large-4c dataset.

Kang et al. [33] utilized an ensemble of deep features from 13 pre-trained deep CNNs with 9 various machine learning classifiers across three datasets: BT-large-2c, BT-small-2c, and BT-large-4c. They found that using the DenseNet-169 deep feature with an SVM with an RBF kernel on the BT-small-2c dataset achieved an accuracy of 94.12%. For the BT-large-2c dataset, the ensemble of DenseNet-169, Inception V3, and ResNeXt-50 deep features with a fully connected classifier achieved an accuracy of 98.83%. Using the BT-large-4c dataset, an ensemble of DenseNet-169, ShuffleNet V2, and MnasNet deep features with an SVM with an RBF kernel achieved an accuracy of 93.72%. Nassar et al. [34] employed a Vision Transformer (ViT) and achieved an accuracy of 95.4% using the BT-large-4c dataset. Bin and Uddin [35] compared a newly developed CNN with pre-trained models such as VGG-16, AlexNet, ResNet-50, and Inception-v3, using transfer learning. The test accuracies of the models were 95.52%, 92.59%, 93.31%, and 89.40%, respectively.

Recent research indicates a predominant reliance on conventional machine learning (ML), sequential deep learning (DL) architectures, and transfer learning (TL) for brain tumor classification in MRI studies. Traditional ML methods often prove ineffective due to their reliance on extensive domain knowledge and manual feature extraction from MRIs. DL models, on the other hand, are limited by the uniform kernel

sizes in convolutional layers, which hinders their ability to extract relevant features from brain MRIs. TL addresses these challenges by leveraging pre-trained models, thereby reducing the need for large annotated datasets and extensive training times. The study utilized five fine-tuned pre-trained architectures with custom layes—EfficientNetV2s [20], Inception ResNetV2 [21], MobileNetV2 [22], RegNetY-320 [23], and ConvNeXt-Large [24]—both without and with fine-tuning. Notably, all models achieved higher accuracy when fine-tuned. EfficientNetV2s demonstrated superior results due to its computational efficiency, achieving superior performance with fewer parameters and reduced computational costs compared to many other models. This makes it ideal for deployment on devices with limited computational resources. Moreover, the system requires minimal preprocessing, involving only image resizing, further enhancing its efficiency and practicality for real-world applications. This research significantly contributes to the field, particularly on the BT-Large-4C dataset, by improving the efficiency and accuracy of classification methods. The findings highlight the potential for enhanced patient outcomes and a reduced workload for medical professionals.

## **3. METHODOLOGY**

The proposed methodology is outlined in Figure 1, providing an overview of the approach for classifying brain tumors and healthy brain MR images. First, we utilized a freely available brain tumor classification dataset from Kaggle, which consists of four categories: pituitary, glioma, meningioma, tumors, and healthy brain images. Secondly, we applied resizing as the only preprocessing step, adjusting all images to the specific input dimensions required for each model: EfficientNetV2s at 384×384 pixels, Inception ResNetV2 at 299×299 pixels, and MobileNetV2, RegNetY-320, and ConvNeXt-Large at 224×224 pixels. The proposed method is divided into two phases. The first phase relies on transfer learning (TL) using five pre-trained architectures. This approach enables us to leverage prior knowledge from large datasets to classify brain MRI images into four distinct categories: pituitary, meningioma, glioma tumors, and normal tissue, as shown in Figure 1 (a). This method operates within the framework of source domain  $\phi_s = (d_1^S, e_1^S), (d_2^S, e_2^S), ..., (d_n^S, e_n^S), d_n^S, e_n^S \in R$  and learning task LS, adjusting to the target domain  $\psi^T = (d_1T, e_1^T), (d_2^T, e_2^T), ..., (d_n^T, e_n^T), ..., (d_n^T, e_n^T), where <math>d_n^T, e_n^T \in R$  and learning task (L<sub>T</sub>) [36]. Finally, we evaluate the model's performance through a detailed set of metrics, including specificity, accuracy, sensitivity, and confusion matrix analysis. This thorough evaluation ensures that our proposed method delivers reliable and precise results in brain tumor classification.

In the second phase, fine-tuning was performed to further optimize the performance of these models. This involved adding custom layers, such as replacing the fully connected (FC) output layer with a Global Average Pooling (GAP) layer, which converted the multi-dimensional feature maps ( $M \times N \times N$ ) into onedimensional vectors ( $1 \times N$ ) for improved feature interpretation. The GAP layer enhanced training efficiency and model accuracy by mitigating overfitting and addressing vanishing gradients [17,18]. Additionally, a dropout layer set at a 0.2 rate played a key role in improving the models' generalization to unseen data, which is crucial for clinical applications [19]. Dense layers were also added to further classify the four categories classification accuracy. This combination of transfer learning, fine-tuning, and the inclusion of GAP and dropout layers worked synergistically to achieve superior accuracy, making the architecture more adept at handling the complexities of brain tumor classification. The two phases of this approach are illustrated in Figure 1 (b).



Fig. 1. Proposed model for brain tumor classification: a) Utilizing Transfer Learning, b) Enhancing performance through finetuning by adding layers based on pre-trained models.

#### 3.1 EfficientNetv2s model

The EfficientNetv2s [20] model efficiently converges with a stable training process and dynamic regularization adjustment. It demonstrates progressive learning and addresses memory consumption by resizing input images to 384×384 pixels. The evaluation of the tumor diagnosis system before and after fine-tuning demonstrates enhanced accuracy, improved feature learning, and reduced overfitting. It has proven adept at tackling challenges in medical image analysis, with prevalent applications including automated detection of tuberculosis in chest X-ray images [37], breast cancer classification [38], and COVID-19 detection through X-ray and CT imaging [39].

The EfficientNetv2 architecture, developed in alignment with EfficientNetv1, demonstrates significant superiority over its predecessor in terms of parameter efficiency and floating-point operations per second (FLOP) performance. FLOPs, a metric for model complexity, reveal that EfficientNetv2 surpasses EfficientNetv1 due to enhanced FLOP efficiency. EfficientNetv1, trained with a large image size, presents a challenge due to substantial memory consumption, requiring the use of smaller batch sizes in CNN models and impeding training speed. A distinctive feature of EfficientNet is its utilization of depth-wise convolutions, which sets it apart from other models [30]. The model was fine-tuned by replacing the conventional flattening layer with Global Average Pooling (GAP) in the neural network, aiming to minimize parameters and accelerate computations. A dropout layer with a 0.2 rate was introduced to enhance model flexibility and robustness. To categorize the data into four categories— pituitary, no tumor, meningioma, glioma and—a dense layer was added. Overall, these adjustments contribute to improving efficiency and classification accuracy in the fine-tuned model.

## 3.2 Inception ResNetV2 model

The Inception-ResNet model, introduced by Szegedy, combines elements from the Inception and ResNet network backbones [21]. The Inception module demonstrates an advantageous local topology by performing several convolution or pooling operations are performed simultaneously on the input image in

parallel. It employs various convolution kernels simultaneously, merging the output results to create a more profound feature map, enhancing image representation [40]. ResNet, proposed by Kaiming He, is a residual neural network with 152 layers, addressing challenges of deep networks by introducing shortcut connections, mitigating issues like gradient dispersion [41]. The Inception-ResNetV2 model combines the strengths of the Inception module with the residual network architecture, increasing both the depth and width of the network while mitigating the issue of gradient vanishing [42].

## 3.3 MobileNetV2 Model

Deep learning methods are constantly evolving, with MobileNetV2 emerging as one of the latest and most widely adopted architectures due to its lightweight design [22]. It is worth noting that MobileNetV2 includes an innovative layer module, the inverted residual with linear bottleneck, which effectively minimizes the required memory for processing [43]. This research presents a brain tumor image classification approach based on the MobileNetV2 network to achieve accurate and efficient results. The approach utilizes MobileNetV2 as the base model for the transfer learning procedure, incorporating a GAP layer. Subsequently, A dropout layer set at a rate of 0.2 was incorporated, followed by a dense layer designed for the classification of four distinct categories: glioma, no tumor, meningioma, and pituitary. The utilization of this lightweight architecture allows for the implementation of the model on mobile devices, facilitating the detection of tumors. The objective is to enable early detection of this disease, thereby enhancing life expectancy [44].

## 3.4 RegNetY-320 Model

RegNetY stands out as a robust convolutional network, deriving its foundation from the original RegNetX model with notable distinctions. RegNetY features linear block width parameters, adaptive depth and initial width configurations, and the incorporation of squeeze-and-excitation blocks, making it a versatile model well-suited for a variety of tasks. Its popularity in the field of machine learning is rapidly growing, cementing its status as one of the most favored models. ResNet marked a pivotal advancement by enabling the efficient training of neural networks with an unprecedented depth exceeding 150 layers. The bottleneck RegNet module, built upon the foundation of the bottleneck ResNet building block, was introduced to overcome challenges associated with large-scale datasets [23].

## 3.5 ConvNeXt-Large Model

ConvNeXt, introduced by Facebook AI Research in 2018, constitutes a deep neural network aimed at enhancing the efficiency of Convolutional Neural Networks (CNN). This improvement is achieved through the implementation of group convolutions and concatenation. The extended version, ConvNeXt-Large, builds upon the original architecture by augmenting both depth and width. ConvNeXt, which attains cutting-edge performance on various image classification benchmarks, including ImageNet, utilizes group of convolutions and concatenation to enhance parameter efficiency and reduce computational overhead during network training. This strategic approach enhances efficiency and scalability [24, 45].

## 4. **RESULTS and DISCUSSION**

The effectiveness of the transfer learning-based framework is assessed in this study on the widely adopted BT-large-4c dataset. This section provides a comprehensive overview of the datasets, evaluation metrics, network training procedures, and subsequent assessment of performance.

## 4.1. Datasets

The BT-large-4c, a freely accessible resource, is commonly used to evaluate classification algorithms. It consists of 3,264 JPEG images featuring MRI scans that depict three categories of brain tumors: pituitary, pituitary, meningioma tumors, as well as images of brain scans exhibiting no tumors. Figure 2 visually

represents a sample of the MRI images presented in sagittal, coronal, and axial views. Specifically, the dataset encompasses 500 MRI scans depicting a tumor-free brain, 901 scans of pituitary tumors, 937 images of meningioma tumors, and 926 images of glioma tumors [46]. The dataset was divided into three subsets: 80%, 10%, 10% for training, testing, and validation, respectively. All images were resized according to the specific input dimensions required by each model: EfficientNetV2s (384×384 pixels), Inception ResNetV2 (299×299 pixels), and MobileNetV2, RegNetY-320, and ConvNeXt-Large (224×224 pixels). Excessive preprocessing may change the inherent properties of the images, which may hinder the model's ability to extract features.



Fig. 2. Visual representations of glioma, meningioma, pituitary tumors, and no tumors in the BT-large-4c dataset.

## **4.2 Evaluation Metrics**

To evaluate the efficacy of the proposed method for the classification of brain tumors, we employed six essential performance metrics: precision, accuracy, sensitivity, F1-score, specificity, and confusion matrices.

Accuracy, a fundamental performance measure, quantifies the percentage of correctly classified image samples relative to the total number of samples, irrespective of specific class labels. Mathematically, accuracy is determined using the following: Equation 1: Equation 5 as follows:

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

Sensitivity, a crucial performance metric, gauges the model's ability to correctly identify brain tumor instances. It is quantified using the following Eq. 2:

Sensitivity (Recall) = 
$$\frac{TP}{TP + FN}$$
 (2)

Specificity, an essential performance metric, assesses the model's capacity to accurately classify negative samples. This is computed using the following Eq. 3:

Specificity = 
$$\frac{TN}{TP + FP}$$
 (3)

Precision, a pivotal performance metric, characterizes the correctness of the positive predictions made by the model. It is computed using the following Eq. 4:

$$Precision = \frac{TP}{TP + FP}$$
(4)

The F1-score, a combined metric, provides a balanced assessment of both precision and recall in a model. It is determined using the following Eq. 5:

$$F1 - Score = \frac{2 * (Precision * Sensitivity)}{Precision + Sensitivity}$$
(5)

Here, TN, FP, FN, and TP represent true negatives, false positives, false negatives, and true positives, respectively, within the context of our study.

## 4.3 Hyper-parameters

Hyperparameter optimization aims to increase a specific deep learning model's performance by choosing the best hyperparameter. Five models were fine-tuned: EfficientNetV2s, Inception ResNetV2, MobileNetV2, RegNetY-320, and ConvNeXt-Large. In this study, the five utilized models were initially applied without fine-tuning and then with fine-tuning to enhance accuracy. During the training process, the learning rate was dynamically adjusted using the ReduceLROnPlateau callback, which monitored validation accuracy to optimize convergence. This adaptive approach, combined with a training duration of 30 epochs and the utilization of the Adam optimizer, significantly enhanced training performance and accelerated convergence towards an optimal solution. This research project was conducted using Python as the primary programming language in the Kaggle Notebook environment, which is an open-source platform that originated from the Python Notebook project. Python, chosen for its object-oriented structure, high-level functionality, and interpretative design, played a central role in the implementation of various tasks.

## 4.4 Qualitative Results

In the first experimental phase, we evaluated transfer learning-based models—EfficientNetV2s, Inception ResNetV2, MobileNetV2, RegNetY-320, and ConvNeXt-Large—using the BT-Large-4C dataset. The utilized models of EfficientNetV2s, Inception ResNetV2, MobileNetV2, RegNetY-320, and ConvNeXt-Large can achieve accuracies of 96.64%, 95.4%, 95.10%, 96.33%, and 95.41%, respectively. EfficientNetV2s achieved the highest an accuracy of 96.64%, while MobileNetV2 showed the lowest an accuracy of 95.10%, as illustrated in Table 1. In the second phase, fine-tuning the models resulted in notable improvements: EfficientNetV2s recorded the highest accuracy at 97.86%, followed by Inception ResNetV2 at 97.25%, RegNetY-320 at 96.94%, MobileNetV2 at 96.64%, and ConvNeXt-Large at 96.01%. The fine-tuning process included the integration of a Global Average Pooling (GAP) layer to alleviate overfitting and reduce issues related to vanishing gradients, with an optimized dropout rate of 20%. EfficientNetV2s surpassed all other models in various metrics, achieving an accuracy of 97.86%, sensitivity of 97.87%, specificity of 98.00%, precision of 98.25%, and an F1-score of 98.00%, as shown in Table 2. Figure 3 shows the accuracy curves (a) and loss curves (b) for both training and validation for the EfficientNetV2s architecture after fine-tuning. The accuracy curve shows significant improvements during training, while the loss curve demonstrates a steady decrease, indicating effective model optimization.

Figure 4 presents the confusion matrices for the EfficientNetV2s architecture: (a) displays the results before fine-tuning, and (b) highlights the enhanced outcomes following fine-tuning.



**Fig.3**. Illustrates the training and validation (a) accuracy curves for the Efficient-netV2s architecture with fine-tuning., (b) loss curves for the Efficient-netV2s architecture with fine-tuning.

Model	ACC.	Sen.	Spec.	Prec.	F1-score
EfficientNetV2s	96.64	96.79	97.00	97.03	97.12
RegNetY_320	96.33	96.26	96.42	96.51	96.43
MobileNetV2	95.1	95.31	95.38	95.45	95.38
ConvNeXt_Large	95.41	95.39	95.50	95.73	95.77
Inception_ResNet	95.40	95.24	95.61	95.59	95.49

Table 1. Results of five pre-trained models utilizing transfer learning, before the fine-tuning process.

Table 2. Evaluation results of five pretrained models post fine-tuning for enhanced performance.

Model	ACC.	Sen.	Spec.	Prec.	F1-score
EfficientNetV2s	97.86	97.87	98.00	98.25	98.00
RegNetY_320	96.94	96.81	97.01	97.25	97.20
MobileNetV2	96.64	96.70	96.81	97.10	96.95
ConvNeXt_Large	96.01	95.94	96.13	96.51	96.49
Inception_ResNet	97.25	97.23	97.33	97.42	97.39

glioma meningioma pituitary no tumor no tumor pituitary glioma meningioma no tumor pituitary meningioma glioma no tumor pituitary meningioma glioma (a) (b)

Fig. 4. Showcases the confusion matrices for the EfficientnetV2s architecture, (a) the outcomes without fine-tuning; (b) the improved results with fine-tuning.

## 4.5 Discussion

The EfficientNetV2s model demonstrated outstanding performance, achieving a 97.86% accuracy, surpassing several leading methods. For comparison, LeaSE by Chitnis et al. [31] achieved an accuracy of 90.60%, a customized CNN by Munira et al. [27] reached 95.41%, Hossain et al. [28] utilized IVX16 and attained 96.94%, and Kang et al. [33] employed an ensemble of pre-trained CNN models with an SVM (RBF), reporting an accuracy of 93.72%, as illustrated in Table 3. Additionally, Nassar et al. [34] using a ViT model achieved an accuracy of 95.4%, and Bin and Uddin [35] tested several transfer learning-based models, with VGG-16 achieving 95.52%, AlexNet achieving 92.59%, ResNet-50 achieving 93.31%, and Inception-v3 achieving 89.40%. These results highlight the effectiveness of fine-tuning in improving brain tumor classification accuracy. For instance, Hossain et al. [28] applied the IVX16 model, integrating three models with additional layers, achieving 96.94% accuracy. In contrast, the individually fine-tuned EfficientNetV2s and Inception ResNetV2 models demonstrate that simpler approaches can often outperform complex systems dependent on multiple models. Furthermore, the fine-tuned RegNetY-320 delivered performance comparable to Hossain et al. [28]. Crucially, the EfficientNetV2s model, even without fine-tuning and relying solely on transfer learning, achieved an accuracy of 96.64%, further underscoring its effectiveness. The proposed system's primary advantage lies in its impressive performance with minimal preprocessing, only requiring resizing, which significantly reduces the computational cost compared to models that necessitate complex preprocessing or ensemble methods.

Ref.s	Technique	ACC. (%)	
Chitnis et al. [31]	LeaSE	90.60%	
Munira et al. [27]	Customized CNN	95.41	
Hossain et al. [28]	IVX16	96.94	
Kang et al. [33]	Ensemble of pre-trained CNN models with an SVM (RBF)	93.72	
Nassar et al. [34]	ViT	95.4	
Bin and Uddin [35]	VGG-16 based TL	95.52	
	AlexNet based TL	92.59	
	ResNet-50 based TL	93.31	
	Inception-v3 based TL	89.40	
Proposed systems	Fine-tuned ConvNeXt_Large	96.01	
	Fine-tuned RegNetY_320	96.94	
	Fine-tuned MobileNetV2	96.64	
	Fine-tuned Inception_ResNetV2	97.25	
	Fine-tuned EfficientnetV2s	97.86	

 Table 3. Comparative results with state-of-the-art approaches on the BT-large-4c dataset.

# 5. CONCLUSION

Precise classification of brain tumors plays a vital role in developing effective treatment methods and improving patient outcomes. This study effectively addresses the inefficiencies and inaccuracies inherent in traditional manual diagnoses by leveraging advanced deep learning techniques to classify MR images of brain tumors. A customized transfer learning strategy with fine-tuning was employed to optimize specific layers of five different models: EfficientNetV2s, Inception ResNetV2, MobileNetV2, RegNetY-320, and ConvNeXt-Large. Of all the models, EfficientNetV2s emerged as the leading model, attaining the highest accuracy. The system underwent rigorous evaluation using the BT-large-4C dataset, and both EfficientNetV2s and Inception ResNetV2 surpassed state-of-the-art techniques across all performance metrics. These results highlight the potential of this approach to significantly enhance brain tumor

classification, ultimately contributing to improved diagnostic accuracy and better treatment outcomes for patients.

EfficientNetV2s demonstrated superior results, as it is designed to be computationally efficient, achieving better performance with fewer parameters and a lower computational cost compared to many other models. This makes it particularly suitable for deployment on devices with limited computational resources. Additionally, this system benefits from minimal preprocessing, requiring only image resizing, which further enhances its practicality for real-world applications without compromising the inherent features of the data.

## 6. FUTURE WORK and LIMITATIONS

Future research should concentrate on various key areas to enhance the effectiveness of brain tumor classification models. Training with a larger and more varied dataset is crucial for enhancing the model's robustness and ability to generalize. Addressing feature dimensionality issues during the transfer learning process will be crucial for simplifying training and enhancing performance. Additionally, optimizing the computational efficiency of the model will be necessary to facilitate practical deployment. Despite the promising results, several limitations exist within the current approach. The dataset utilized was limited in both size and diversity, which may not adequately capture the variability present in real-world scenarios, potentially affecting the model's generalizability.

#### REFERENCES

- 1. Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., Shinohara, R.T., Berger, C., Ha, S.M., Rozycki, M., et al.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. arXiv preprint, arXiv:1811.02629 (2018).
- 2. Asif, S., Yi, W., Ain, Q.U., Hou, J., Yi, T., Si, J.: Improving effectiveness of different deep transfer learning-based models for detecting brain tumors from MR images. In: IEEE Access, vol. 10, pp. 34716-34730. IEEE (2022).
- Kibriya, H., Amin, R., Alshehri, A.H., Masood, M., Alshamrani, S.S., Alshehri, A., et al. A novel and effective brain tumor classification model using deep feature fusion and famous machine learning classifiers. In: Computational Intelligence and Neuroscience, vol. 2022. Hindawi (2022). <u>https://doi.org/10.1155/2022/7897669</u>.
- 4. Komninos, J., Vlassopoulou, V., Protopapa, D., Korfias, S., Kontogeorgos, G., Sakas, D.E., Thalassinos, N.C.: Tumors metastatic to the pituitary gland: case report and literature review. In: The Journal of Clinical Endocrinology & Metabolism, vol. 89, no. 2, pp. 574-580. Oxford University Press (2004).
- Nassar, S.E., Yasser, I., Amer, H.M., Mohamed, M.A. A robust MRI-based brain tumor classification via a hybrid deep learning technique. In: The Journal of Supercomputing, pp. 1-25. Springer (2023). <u>https://doi.org/10.1007/s11227-023-05549-w</u>
- 6. Cancer.Net: Brain Tumor Statistics. http://www.cancer.net/cancer-types/brain-tumor/statistics (2023). Accessed: 2025-1-25.
- Rehman, A., Khan, M. A., Saba, T., Mehmood, Z., Tariq, U., Ayesha, N. Microscopic brain tumor detection and classification using 3D CNN and feature selection architecture. Microscopy Research and Technique 84(1), 133-149 (2021). DOI: 10.1002/jemt.23597.
- Khan, M. A., Sharif, M., Akram, T., Bukhari, S. A. C., & Nayak, R. S.: Developed Newton-Raphson based deep features selection framework for skin lesion recognition. Pattern Recognition Letters 129, 293–303 (2020). DOI: 10.1016/j.patrec.2019.11.034.
- 9. Khan, M. A., Sarfraz, M. S., Alhaisoni, M., Albesher, A. A., Wang, S., & Ashraf, I.: StomachNet: Optimal deep learning features fusion for stomach abnormalities classification. IEEE Access 8, 197969–197981 (2020).
- Kumar, S., Dabas, C., & Godara, S.: Classification of brain MRI tumor images: A hybrid approach. Procedia Computer Science 122, pp. 510–517 (2017).
- 11. Nassar, S. E., Mohamed, M. A. E.-A., & Elnakib, A.: MRI brain tumor segmentation using deep learning. MEJ. Mansoura Engineering Journal 45(4), pp. 45–54 (2021).
- 12. Wang, J., Lu, S.-Y., Wang, S.-H., & Zhang, Y.-D.: Ranmerformer: Randomized vision transformer with token merging for brain tumor classification. Neurocomputing 573, pp. 127216 (2024). <u>https://doi.org/10.1016/j.neucom.2023.127216</u>.
- Talo, M., Baloglu, U. B., Yıldırım, "O., & Acharya, U. R.: Application of deep transfer learning for automated brain abnormality classification using MR images. Cognitive Systems Research 54, pp. 176–188 (2019). <u>https://doi.org/10.1016/j.cogsys.2018.12.007</u>.
- 14. Morid, M. A., Borjali, A., & Del Fiol, G.: A scoping review of transfer learning research on medical image analysis using ImageNet. Computers in Biology and Medicine 128, pp. 104115 (2021).
- 15. Abbasi, A. A., Hussain, L., Awan, I. A., Abbasi, I., Majid, A., Nadeem, M. S. A., & Chaudhary, Q.-A.: Detecting prostate cancer using deep learning convolution neural network with transfer learning approach. Cognitive

Neurodynamics 14, pp. 523-533 (2020). https://doi.org/10.1007/s11571-020-09587-5.

- Talukder, M. A., Islam, M. M., Uddin, M. A., Akhter, A., Pramanik, M. A. J., Aryal, S., Almoyad, M. A. A., Hasan, K. F., & Moni, M. A.: An efficient deep learning model to categorize brain tumor using reconstruction and fine-tuning. Expert Systems with Applications 230, pp. 120534 (2023). <u>https://doi.org/10.1016/j.eswa.2023.120534</u>.
- 17. Bazi, Y., Al Rahhal, M. M., Alhichri, H., & Alajlan, N.: Simple yet effective fine-tuning of deep CNNs using an auxiliary classification loss for remote sensing scene classification. Remote Sensing 11(24), pp. 2908 (2019).
- Malla, P. P., Sahu, S., & Alutaibi, A. I.: Classification of tumor in brain MR images using deep convolutional neural network and global average pooling. Processes 11(3), pp. 679 (2023).
- Ali, I., Muzammil, M., Haq, I. U., Amir, M., & Abdullah, S.: Efficient lung nodule classification using transferable texture convolutional neural network. IEEE Access 8, pp. 175859–175870 (2020). DOI: 10.1109/ACCESS.2020.3026080.
- Anagun, Y.: Smart brain tumor diagnosis system utilizing deep convolutional neural networks. Multimedia Tools and Applications, pp. 1–27 (2023).
- Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.: Inception-v4, inception-resnet and the impact of residual connections on learning. Proceedings of the AAAI Conference on Artificial Intelligence 31 (2017).
- 22. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 (2017).
- Alam, T.M., Shaukat, K., Khan, W.A., Hameed, I.A., Almuqren, L.A., Raza, M.A., Aslam, M., Luo, S.: An efficient deep learning-based skin cancer classifier for an imbalanced dataset. Diagnostics 12(9), 2115 (2022). doi: 10.3390/diagnostics12092115
- Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., Xie, S.: A convnet for the 2020s. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11976–11986 (2022).
- Noreen, N., Palaniappan, S., Qayyum, A., Ahmad, I., Imran, M., & Shoaib, M.: A deep learning model based on concatenation approach for the diagnosis of brain tumor. IEEE Access 8, pp. 55135–55144 (2020). DOI: 10.1109/ACCESS.2020.2978629.
- Kujur, A., Raza, Z., Khan, A. A., & Wechtaisong, C.: Data complexity based evaluation of the model dependence of brain MRI images for classification of brain tumor and Alzheimer's disease. IEEE Access 10, pp. 112117–112133 (2022). DOI: 10.1109/ACCESS.2022.3216393.
- 27. Munira, H. A., & Islam, M. S.: Hybrid deep learning models for multi-classification of tumour from brain MRI. Journal of Information System and Emerging Technologies 8(2), pp. 162–174 (2021). DOI: 10.20473/jisebi.8.2.162-174.
- Hossain, S., Chakrabarty, A., Gadekallu, T. R., Alazab, M., & Piran, M. J.: Vision transformers, ensemble model, and transfer learning leveraging explainable AI for brain tumor detection and classification. IEEE Journal of Biomedical and Health Informatics (2023). DOI: 10.1109/JBHI.2023.3266614.
- Vankdothu, R., Hameed, M. A., & Fatima, H.: A brain tumor identification and classification using deep learning based on CNN-LSTM method. Computers and Electrical Engineering, vol. 101, pp. 107960 (2022). DOI: 10.1016/j.compeleceng.2022.107960.
- Xiao, G., Wang, H., Shen, J., Chen, Z., Zhang, Z., & Ge, X.: Synergy factorized bilinear network with a dual suppression strategy for brain tumor classification in MRI. Micromachines, vol. 13, no. 1, pp. 15 (2021). DOI: 10.3390/mi13010015.
- Chitnis, S., Hosseini, R., & Xie, P.: Brain tumor classification based on neural architecture search. Scientific Reports, vol. 12, no. 1, pp. 19206 (2022). DOI: 10.1038/s41598-022-22172-6.
- Balaji, G., Sen, R., & Kirty, H.: Detection and classification of brain tumors using deep convolutional neural networks. arXiv preprint arXiv:2208.13264 (2022).
- Kang, J., Ullah, Z., & Gwak, J.: MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers. Sensors, 21(6), 2222 (2021). <u>https://doi.org/10.3390/s21062222</u>.
- Nassar, S. E., Yasser, I., Amer, H. M., & Mohamed, M. A.: Comparative analysis of vision transformers and fine-tuned transfer learning models for brain tumor classification. Imaging and Radiation Research 7(1), 8514 (2024). <u>https://doi.org/10.24294/irr8514</u>.
- 35. Bin, S. M., & Uddin, M.: NeuroSight: A deep-learning integrated efficient approach to brain tumor detection. Engineering Reports, vol. 7, no. 1, e13100 (2025). https://doi.org/10.1002/eng2.13100.
- Khan, M.A., Mostafa, R.R., Zhang, Y.-D., Baili, J., Alhaisoni, M., Tariq, U., Khan, J.A., Kim, Y.J., Cha, J.: Deep-net: Fine-tuned deep neural network multi-features fusion for brain tumor recognition. Comput. Mater. Contin. 76(3), 3029– 3047 (2023). doi: 10.32604/cmc.2023.038838
- 37. Ahmad, I., Shin, S.: A perceptual encryption-based image communication system for deep learning-based tuberculosis diagnosis using healthcare cloud services. Electronics 11(16), 2514 (2022). doi: 10.3390/electronics11162514
- Sarker, M. M. K., Akram, F., Alsharid, M., Singh, V. K., Yasrab, R., Elyan, E.: Efficient breast cancer classification network with dual squeeze and excitation in histopathological images. Diagnostics 13(1), 103 (2022). doi: 10.3390/diagnostics13010103
- Huang, M.-L., Liao, Y.-C.: A lightweight CNN-based network on COVID-19 detection using X-ray and CT images. Computers in Biology and Medicine 146, 105604 (2022). doi: 10.1016/j.compbiomed.2022.105604
- Wu, C., Wen, W., Afzal, T., Zhang, Y., Chen, Y., et al.: A compact DNN: approaching GoogLeNet-level accuracy of classification and domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5668–5677 (2017).
- 41. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016).
- 42. Li, Z., Tian, X., Liu, X., Liu, Y., Shi, X.: A two-stage industrial defect detection framework based on improved-YOLOv5 and optimized-Inception-ResNetV2 models. Applied Sciences 12(2), 834 (2022).

https://doi.org/10.3390/app12020834

- 43. Eyiokur, F. I., Ekenel, H. K., Waibel, A.: Unconstrained face-mask & face-hand datasets: Building a computer vision system to help prevent the transmission of COVID-19. Signal, Image and Video Processing 17, 1027–1034 (2023). https://doi.org/10.1007/s11760-022-02308-x.
- 44. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp. 4510–4520.
- 45. Rani, K. A. & Gowrishankar, S. Pathogen-based classification of plant diseases: A deep transfer learning approach for intelligent support systems. IEEE Access (2023).