



AI System for Detecting Brain Tumor

Prepared By

Name: Alaa Mohammed Hassan

ID No. 252000381

Submitted in fulfillment of the requirements for the degree Master of
Science in Artificial Intelligence

At

Collage of Informatics

Midocean University

UAE

Supervisor Name: **Dr. Ashraf Elsayed**

Collage of Informatics, Midocean University, Comoros

Email: 252000381@st.midocean.edu.km

Acknowledgements

Special thanks to Dr. Ashraf El-Sayed.

First and foremost, we offer our deepest gratitude to God for His continued guidance in enabling us to complete this work on time. We then extend our profound thanks and appreciation to the beacon of knowledge and role model for researchers, Professor Dr. Ashraf El-Sayed, who graciously supervised this research. His knowledge, time, and invaluable effort played a crucial role in refining and perfecting this work. We are deeply grateful for his patience and unwavering support throughout the research process. I learned from him the fundamentals of scientific research, as well as the virtues of good character. I offer him my utmost respect and appreciation. We ask God to reward him abundantly, to grant him success in his pursuit of knowledge, to bless him in his life and knowledge, and to guide him in all that is good. We also extend our heartfelt thanks to Dr. Hussein.

Abstract

Brain tumors are a major cause of psychological and physical problems, making early and accurate diagnosis crucial for successful treatment. Therefore, we used deep learning to detect and classify brain tumors using a large collection of MRI images of various types. This is essential for accelerating treatment.

In this study, medical images were processed using convolutional neural networks (CNNs), which achieved excellent image classification results. MobileNetV2, Inception v3, and EfficientNetB0 were used for rapid tumor detection and classification, along with the Adam optimizer using sparse cross-entropy loss function and a precision metric. Among the tested models, MobileNetV2 achieved the highest Acc with 82%, Precision 81%, Recall 79%, and F1-score 79%. EfficientNetB0 demonstrated a good balance between network depth, width, and Acc, allowing the model to learn more efficiently without overfitting.

This study discusses deep learning methods for detecting brain tumors using artificial intelligence, obtaining results in minutes or less, which saves time and effort and provides accurate results that help the doctor make informed decisions.

Key Word:

Brain tumors, deep learning , CNN , MRI , classification , Kaggle , MobileNetV2 , Inception v3, EfficientNetB0, accuracy .

Table of Contents

ii	
AI System for Detecting Brain Tumor.....	ii
Acknowledgements	iii
Abstract	iv
Table of Contents	v
List of Abbreviations	vii
List of Figures.....	viii
List of Tables	ix
Chapter 1 Introduction	1
1.1 Problem Definition.....	2
1.2 Thesis Objectives	3
1.3 Thesis Contribution.....	4
1.4 Thesis Organization	4
Chapter 2	5
Background and Related Work	5
2.1 Background	6
2.2 Related Work:	7
Chapter 3	13
Proposed Approach (Methodology).....	13
3.1 Methodology:	14
3.2 Proposed approach	20
Chapter 4:	21
Experimental Results and Discussion	21
4.1 Experimental Results:	22
Chapter 5	31

Conclusion and Future Work	31
5.1 Conclusion:	32
5.2 Future Work:	32
References	33

List of Abbreviations

- **Acc.:** Accuracy.
- **AI:** Artificial Intelligence.
- **CNN:** Convolutional Neural Network.
- **MRI:** Magnetic Resonance Imaging.
- **CT:** Computed Tomography.
- **DL:** Deep Learning.
- **SVM:** Support Vector Machine.
- **VGG 16:** Visual Geometry Group 16-layer.
- **XAI :** eXplainable Artificial Intelligence
- **SPPE :** Stacked Parallel-Parallel Ensemble/Network
- **BiFPN:** Bidirectional Feature Pyramid Network.
- **BraTs:** Brain Tumor Segmentation Challenge.
- **EHR:** Electronic health records.
- **SGD:** Stochastic Gradient Descent.
- **CPU:** Central Processing Unit.
- **RAM:** Random-Access Memory.

List of Figures

Figure 3.1	The sample image from each class	14
Figure 3.2	Data Augmentation [17].	15
Figure 3.3	Detailed of the MobilenetV2 architecture [18].	17
Figure 3.4	Detailed of the EfficientNetB0 architecture [16].	18
Figure 3.5	Detailed of the Inception v3 architecture [17].....	18
Figure 3.6	Proposed approach.....	20
Figure 4.1	Accuracy and loss in MobileNetV2	23
Figure 4.2	Confusion Matrix over result after apply MobileNetV2	24
Figure 4.3	Accuracy and Loss InceptionV3	25
Figure 4.4	Confusion Matrix over result after apply InceptionV3	26
Figure 4.5	Accuracy and Loss EfficientNetB0	27
Figure 4.6	Confusion Matrix over result after apply EfficientNetB0.....	28

List of Tables

Table 2.1 Comparison of previous studies.....	10
Table 3.1 Dataset Brain Tumor	15
Table 4.1 : Models performance average.....	28
Table 4.2: Training Time for each Model.....	29
Table 4.3: Models performance in each class	29

Chapter 1

Introduction

1.1 Problem Definition

The brain is the central organ responsible for all bodily activities and issuing commands, aided by nerve cells. The human brain is incredibly complex, regulating numerous biological functions, such as decision-making and command issuing. Many diseases can affect the brain, including stroke, epilepsy, migraines and Alzheimer's disease, Parkinson's disease, and brain tumors.

Brain tumors are abnormal growths of brain cells that accumulate to form masses. These tumors can be benign or malignant and can cause a variety of problems and disrupt vital functions.

Brain tumors are among the most serious diseases and require accurate and rapid detection methods.

There are many symptoms associated with brain tumors, whether benign or malignant, including cognitive decline and problems resulting from pressure on the optic nerve. However, the most common symptom is headache. A tumor can cause various problems leading to impaired hearing, smell, or vision, and may cause weakness or paralysis in an arm or leg, or in some cases, coma.

Benign tumors are abnormalities in the cell life cycle and are the least harmful type of tumor, rarely recurring after removal. Malignant tumors, on the other hand, consist of cancerous cells that spread rapidly and cause serious functional problems. Brain tumors are classified into categories such as pituitary tumors, meningiomas, and gliomas. For example, the pituitary gland is located at the base of the brain, and its function is to regulate the secretion of hormones from other glands in the body, such as hormones that regulate metabolism, water and salt balance, growth, blood pressure, skin pigmentation, and more. It also controls the hormones of other glands in the body, such as the adrenal and thyroid glands. When a tumor forms in this important and sensitive area, its effects will not be limited to the pituitary gland alone, but will directly or indirectly affect other

glands in the body, causing significant disruption to their functions and hormone secretion.

1.2 Thesis Objectives

Most diagnostic methods rely on the evaluation of MRI images by a neurologist and a radiologist. This is a complex and time-consuming process that can lead to inaccurate results and errors. This is in addition to traditional biopsy techniques, which are painful, time-consuming, and also prone to inaccuracy.

Effective treatment planning depends on rapid and accurate diagnosis. For example, benign tumors can be treated with complete surgical excision, while malignant tumors are treated with radiation, chemotherapy, or both.

With the rapid development of medical imaging technologies, such as magnetic resonance imaging (MRI) and computed tomography (CT), which play a vital role in disease detection and diagnosis, early detection of tumors has become crucial. This facilitates early treatment, supports clinical decision-making, improves patient outcomes, and God willing and increases the likelihood of recovery.

This project aims to help clinicians achieve highly accurate diagnoses using deep learning algorithms (convolutional neural networks) and MRI image classification. These algorithms can accurately detect tumors of all sizes within the brain.

This enables doctors to analyze thousands of images in just a few minutes, significantly reducing human error during diagnosis and contributing to faster decision-making and the selection of the most appropriate treatment, thus increasing the patient's chances of survival, God willing. This is our ultimate goal.

1.3 Thesis Contribution

Artificial intelligence (AI) has revolutionized brain tumor detection and classification, making it an indispensable tool for neurologists. Healthcare researchers have employed a range of deep learning algorithms, such as convolutional neural networks (CNNs), MobileNetv2, and EfficientNetB0, to aid in tumor detection and classification, along with the Inception v3 model and the Adam optimizer for enhanced training. This has been facilitated by the availability of vast amounts of medical imaging data for analysis, such as that offered by platforms like Kaggle, highlighting the importance of AI in supporting physicians in tumor detection and its speed in identifying tumor type compared to traditional, painful biopsy methods.

1.4 Thesis Organization

This study is divided into several chapters, organized as follows:

- Chapter 1: Provides general background on brain tumors, the importance of early diagnosis, and presents the research problem, objectives, and contributions.
- Chapter 2: Reviews and compares relevant background and research on brain tumor classification using artificial intelligence and deep learning techniques.
- Chapter 3: Explains the proposed methodology, including data preparation, preprocessing, and the models used.
- Chapter 4: Presents and discusses the experimental results, including performance analysis and comparisons between different models.
- Chapter 5: Presents the final conclusions and recommendations for future research.

Chapter 2

Background and Related Work

2.1 Background

Tumors of all types are among the most serious diseases of our time due to their widespread prevalence. They affect all age groups and ethnicities, and brain tumors are considered among the most dangerous because of their direct impact on the nervous system, spinal cord, and certain glands. A tumor grows and develops within or around the brain and is classified into several grades based on its aggressiveness and rate of growth. The most common types of tumors include gliomas, meningiomas, and pituitary tumors.

Diagnostic methods rely primarily on computed tomography (CT) and magnetic resonance imaging (MRI). These tests are performed first when a person experiences headaches or other related symptoms. A tissue sample (brain biopsy) is also taken and examined in the laboratory. Despite advancements in the medical field, the interpretation of the results depends heavily on the experience of the specialists involved (neurologist, radiologist, and laboratory specialist), which can lead to errors.

The field of artificial intelligence (AI) has witnessed remarkable advancements in the medical field in recent years, with a growing interest in developing diagnostic models using deep learning and convolutional neural networks. These technologies have proven highly effective in analyzing and classifying medical images due to their ability to extract distinctive features.

Several advanced models are used for image classification, such as VGG16, MobileNetV2, and EfficientNetB0. Their strength lies in their ability to automatically learn complex patterns within data and extract features, contributing to improved diagnostic accuracy and reduced analysis time.

However, AI applications in the medical field face several challenges, including data imbalances, limited medical data, and patient privacy concerns. Training models also suffer from over-customization and other issues.

Despite these challenges, numerous studies in recent years have demonstrated significant progress in using AI for the accurate and rapid diagnosis of brain tumors. This has prompted researchers to conduct studies analyzing MRI images to detect tumorous areas and classify them into different categories. These studies point to the use of strategies based on the deep convolutional neural network framework, which uses multiple layers and variables to accurately identify brain cancer.

2.2 Related Work:

Recent years have witnessed remarkable progress in the use of artificial intelligence in the accurate and rapid diagnosis of brain tumors. This has prompted researchers to conduct studies analyzing MRI images to detect tumorous areas and classify them into different categories. These studies indicate the use of strategies based on the deep convolutional neural network framework (CNN), which employs multiple layers and variables to accurately identify brain cancer.

The researchers in [1] proposed a model designed to divide MRI images depicting brain tumors into two categories. The model (CNN + Adam) demonstrated excellent performance, achieving 98% Acc, in addition to feature extraction (VGG16). They processed 3,000 MRI images downloaded from (Kaggle), despite the challenges they encountered in ensuring compatibility between the medical record systems (EHR) and (AI). These challenges were resolved by developing interfaces that scan (AI) with (EHR) for information exchange and integration.

This model, as described in the paper [2], is distinguished by its classification, detection, and localization of brain tumors. It utilizes a fully automated model employing the DL&YOLOv7 algorithm to minimize false positives and losses, and integrates SPPF + BiFPN components to address the challenge of small tumors. The total dataset for the various categories comprised 10,288 Kaggle MRI images. The model demonstrated superiority in detecting gliomas,

meningiomas, and pituitary tumors in the brain, achieving a verification Acc of 99.5%.

Reference [3] presents the identification and classification of brain tumors into four types using a deep learning-based hierarchical classification system. MATLAB software was used for intelligent detection using CNNs, and the initial function involved a RELU layer on 3264 Kaggle MRI images. The model achieved an Acc of 92.13%, and the images were processed using intermediate filtering to remove noise with a multi-layered hierarchical classification.

In this study [4], the researchers proposed comparing different DL models, based on various types of neural networks (CNN, ANN, TL), to identify and classify brain tumors. They used ANN, CNN (ReLU), and TL algorithms from the Keras library on 3190 Kaggle MRI images. The results showed that CNN outperformed ANN in classifying medical images, achieving 90% Acc. The best TL model was VGG16, achieving the highest accuracy among all models at 94%.

Brain tumors were classified using deep learning (CNN) and x-ray analysis (XAI) techniques, as described in the paper [5], which significantly outperformed the Grad-Cam model, which focused on and highlighted tumors. The model achieved 99% Acc in the first dataset and 95% in the second. Methods such as XAI Shap and LIME were used to verify local interpretations and model decisions, and to test the model's results. The model demonstrated accuracy and balance, but it was trained and tested using only MRI images, and the results may be unsatisfactory with other imaging techniques such as computed tomography. The dataset used was created by Nikparvar and included 7023 images, while the second dataset, NeuroMRIs, included 3264 images.

In this research [6], GoogleNet and AlexNet were used as learning networks for feature extraction, but the results were unsatisfactory. Therefore, a hybrid approach was employed, combining them with other networks (AlexNet-SVM and AlexNet-KKN) to improve verification and testing accuracy. The study utilized a dataset compiled from various sources (Figshare, SARTAJ, and Br35h) containing 2880 MRI images. The KFold data splitting technique was used, and the results showed that AlexNet-KKN exhibited higher verification and testing

accuracy compared to GoogleNet, with the hybrid networks achieving accuracy rates of 96.9% and 98.6%, respectively.

In this study [9], researchers proposed a model characterized by a two-layer architecture combining CNN for classification and DeepLabV3 for image segmentation. This resulted in high accuracy (99.31%) using the Adam enhancer. The study also relied on Transfer and various enhancers (SGD, InceptionV3), as well as employing different strategies to combat overfitting. They processed 755 MRI images from BraTs MRI 2015.

A limitation of the study was its reliance on a single dataset, which reduces generalizability. Therefore, their future plan is to conduct multi-class classification of tumor subtypes and validate multi-institutional and multi-modal data.

This study [10] relied on a Kaggle dataset that combines several deep learning models, including CNN, Efficient Net V3, Mobile Net V2, Inception- ResNet V2, in addition to the Vision Transformer model, achieving an accuracy of 95%. One of its drawbacks is its reliance on CNN, which requires a large dataset and longer training time. This study used only one dataset, thus reducing its generalizability. One of the advantages is that an innovative neural network was built with the aim of overcoming training limitations, and their future plans include further improving the models and exploring ways to enhance classification performance even more.

The authors [11] proposed an innovative classification system based on CNNs and a deep learning architecture that integrates EfficientNetB7, ResNet-152, and enhanced Xception to improve feature extraction and classification. They also utilized XAI techniques, including Grad-CAM and LIME, to provide visual insights into decision-making and treatment planning, which are essential for clinical applications.

Five sets containing 20,286 MRI images were used in the Kaggle platform, demonstrating high accuracy ranging from 97% to 99%. However, the model faces several limitations, including the need for high computational power and the need for multicenter studies and further clinical evaluations.

In this study [12], researcher's integrated 6420 MRI images downloaded from Kaggle and developed a hybrid deep learning model by combining CE-EEN-B0 and ResGANet to enhance classification performance. The model demonstrated 99.11% Acc, outperforming SOTA models. The use of Grad-CAM also enhanced visual interpretability, increasing the reliability of its predictions for healthcare professionals and providing valuable insights that improve treatment planning and patient care.

Future plans focus on expanding the dataset, exploring its application in other medical imaging fields, and conducting multicenter clinical trials to ensure its adoption in real-world practice.

Table 2.1 Comparison of previous studies

Datasets used	Accuracy	Methodology	advantages	Disadvantages	Reference
	achieving 98% accuracy	VGG16 +Adam	Interfaces have been developed that integrate electronic health record (EHR) systems with artificial intelligence (AI) for information exchange and integration.	- Lack of clarity regarding comparisons with other models - Challenges in integration with EHR - Failure to address data diversity	[1]
10,288 Kaggle MRI images	accuracy of 99.5%	DL&YOLOv7 and integrates SPPF + BiFPN	-Its ability to classify, detect, and locate brain tumors. -Addressing the problem of small tumors. -Very high accuracy. -Large data volume.	-Lack of data source diversity -No comparison with other models	[2]
3264 Kaggle MRI images	accuracy of 92.13%	CNN	Hierarchical classification based on deep learning. Using intermediate filtering to remove noise and	Lack of modern techniques. Low data volume. Lower accuracy compared to previous studies.	[3]

			improve image quality.		
3190 Kaggle MRI images	VGG16= 94%. CNN= 90%	CNN & VGG16	Comparison between different models. Using deep learning techniques. - Adopting the Keras library.	- Relying on data from a single source (Kaggle),	[4]
by Nikparvar and included 7023 images, while the second dataset, NeuroMRIs, included 3264 images	The first data set has 99% accuracy. The second data set has 95% accuracy.	-DL→CNN -XAI→ Grad-CAM - XAI Techniques→ SHAP- LIME	-Its accuracy is high, and the model demonstrated precision and balance. - Using two different datasets. - Integrating XAI techniques.	- Relying solely on MRI images and not using real clinical data	[5]
various sources (Figshare, SARTAJ, and Br35h) containing 2880 MRI images	96.9% and 98.6	DL→ GoogleNet & AlexNet Hybrid Models→ AlexNet-SVM and AlexNet-KKN The KFold data splitting technique	- Using diverse data sources. The hybrid approach achieved higher verification and testing accuracy. Comparison between models.	- The model was not tested on other imaging types such as CT.	[6]
755 MRI images from BraTs MRI 2015		- CNN for classification and DeepLabV3 + Adam. - SGD, InceptionV3	It employs a two-layer structure that integrates classification and segmentation. Its high accuracy is ensured. It utilizes transfer learning techniques and overfitting strategies. It relies on reliable and well-established data in	The data volume is very small, and there is only one data source.	[9]

			the field of brain tumors.		
Kaggle	accuracy of 95%	CNN , Efficient Net V3 , Mobile Net V2 , Inception-ResNet V2 and Vision Transformer	A wide variety of models are used. Their accuracy is high. They use modern models. They have future plans to improve the model.	It used a single dataset. It relied on CNN, which requires a massive dataset and longer training time.	[10]
20,286 MRI images were used in the Kaggle		CNN→ EfficientNetB7, ResNet-152, and enhanced Xception. XAI→ Grad-CAM and LIME.	Utilizing robust algorithms. - Improving the quality of extracted characteristics, feature extraction, and classification. - Employing XAI techniques to provide visual insights in decision-making and treatment planning.	- Failure to specify the size or variety of the data. - Using large models such as ResNet-152 and EfficientNetB7 requires significant computing resources. - Lack of knowledge regarding the model's accuracy.	[11]
6420 MRI images downloaded from Kaggle	99.11% accuracy	DL→ CE-EEN-B0 and ResGANet. XAI→ Grad-CAM hybrid deep learning model→ CE-EEN-B0 and ResGANet	- Very high accuracy, surpassing and outperforming the latest available models. - A hybrid model combining CE-EEN-B0 and ResGANet to improve classification performance. - Visual explanations enhance the model's reliability.	A single dataset. - The model is not tested on multi-mode data (such as CT or PET), which limits its practical application.	[12]

Chapter 3

Proposed Approach (Methodology)

3.1 Methodology:

The methodology begins with collecting images from reliable platforms and converting the visual data into a format that can be understood and processed by deep learning models. This allows for precise tumor classification, type identification, and evaluation. The methodological structure of this study comprises several key steps or phases: setting up the software environment, data collection, data preprocessing, classification models, and evaluation models.

3.1.1 Data collection

This thesis presents a dataset of magnetic resonance imaging (MRI) images for brain tumor classification, downloaded from Kaggle. The first dataset, titled "Brain Tumor Classification (MRI) [7]," contains 2,656 JPG images, divided into two categories: training (80%) and testing (20%). Each category is further divided into four subcategories: pituitary tumor, glioma, no tumor, and meningioma.

The second dataset, "Brain Tumor Dataset," contains [8] (10,560) images, also divided into four subcategories, in JPG and PNG formats.

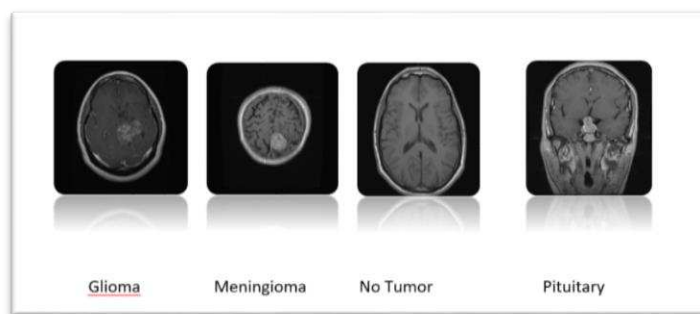


Figure 3.1 The sample image from each class

Table 3.1 Dataset Brain Tumor

	Group2	Group 1		
		Training	Testing	Total
Glioma	3754	826	100	4680
Meningioma	2343	247	127	2717
No- tumor	1757	327	104	2188
Pituitray	2706	827	98	3631
Total	10560	2227	429	13216

3.1.2 Data Preprocessing

In this study, a variety of data preprocessing techniques were applied to improve model stability, training quality, and speed, and to reduce overfitting caused by data imbalances between folders. The techniques used will be briefly described.

3.1.3 Image Normalization:

Normalization was applied to transform pixel values from the 0–255 range to the 0–1 range, which helps accelerate the training process and improve model stability. Furthermore, normalization ensures a uniform distribution of values, enabling the model to learn more efficiently.

Data Augmentation:

To reduce overfitting, several image enhancement techniques were used to increase sample diversity, such as Random Flip, Random Rotation, Random Zoom, and Random Contrast.

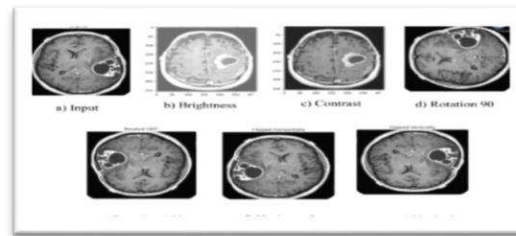


Figure 3.2 Data Augmentation [17].

3.1.4 Data Balance Using Oversampling:

Due to the imbalance in the number of images among the four categories (No Tumor, Glioma, Meningioma, and Pituitary), oversampling was applied to increase the number of images in the underrepresented categories until they were equal to the largest category. New images were generated using ImageDataGenerator.

In addition, simple transformations such as Rotation, Shift, Zoom, and Horizontal Flip were applied. The importance of this step lies in reducing overfitting caused by data imbalances, improving the model's generalization ability, and preventing the model from favoring larger categories.

3.1.5 Classification Models

In this study, we used and compared several different models to determine the best one for accurately classifying brain tumors. The main models used in this study were pre-trained convolutional neural networks, specifically MobileNetV3, EfficientNetB0, and InceptionV2.

The results from these models help us understand how to differentiate tumor types based on important spatial patterns and overall characteristics derived from brain MRI images. The results varied in terms of accuracy, loss, and speed. A detailed description of each model used in this research follows.

3.1.6 Pre-trained CNNs models:

MobileNetV2 is a lightweight and fast machine learning model (CNN) developed by Google, designed for vision applications on mobile and embedded devices while maintaining excellent accuracy. It is available in several versions: MobileNetV1, MobileNetV2, and MobileNetV3, and consists of 53 trainable layers. The MobileNetV2 architecture uses an inverted residual structure where the input and output of the residual blocks are thin bottleneck layers. These begin with an expansion layer (a 1x1 convolutional layer that expands the data),

followed by a convolutional layer (a 3x3 layer) used for data processing, and finally a projection layer (a linear bottleneck) to conserve information (also a 1x1 layer).

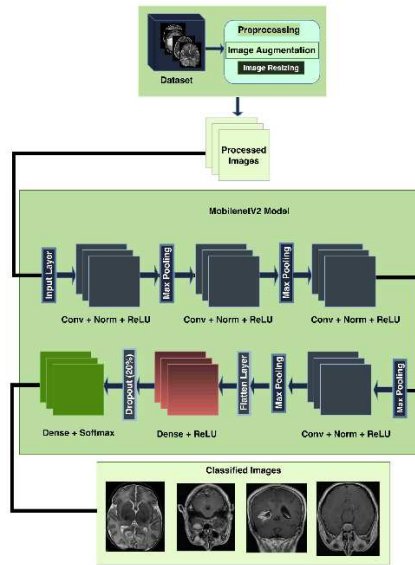


Figure 3.3 Detailed of the MobilenetV2 architecture [18].

EfficientNet is a family of CNN models that operate on the principle of balancing width, depth, and resolution, using minimal computation. Developed by Google, it is characterized by its flexibility, accuracy, and efficiency. It has several versions, starting with B0, with subsequent versions expanding to EfficientNet-B7. We used EfficientNetB0, which features a small number of parameters, moderate computational cost, and simplicity, making it a good choice for image classification tasks. It consists of 237 layers and seven main MBConv modules. The EfficientNetB0 architecture consists of iterative blocks that perform deep decoupling convolutions to reduce computation, batch normalization (improving compression and excitation), and the Swish nonlinear activation function to enhance efficiency and performance. It is used in complex tasks.



1-Figure 3.4 Detailed of the EfficientNetB0 architecture [16].

Inception v3 is an advanced convolutional neural network (CNN) architecture developed by Google, designed to enhance computational efficiency and accuracy in image classification tasks. It has several distinct and improved versions: Inception-v1 (GoogLeNet), Inception-v2, Inception-v3, and Inception-v4. This version is based on the modular concept of Inception and relies heavily on batch normalization. Inception v2 introduces segmented convolutions to reduce computational costs and incorporates batch normalization after each convolution to accelerate training and improve stability. The architecture includes convolutional, aggregate, and fully connected layers, designed to achieve an effective balance between depth and width.

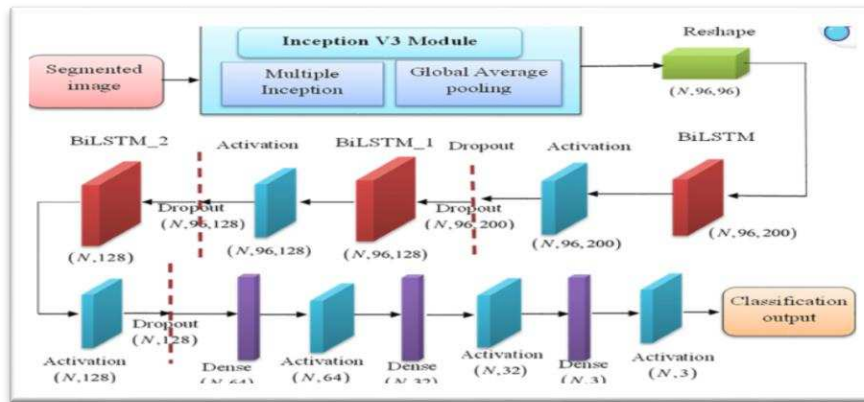


Figure 3.5 Detailed of the Inception v3 architecture [17].

3.1.7 Evaluation Models:

In this study, we have several evaluation criteria. We don't rely solely on the term "accuracy" due to the sensitivity of the topic and the potentially costly consequences of diagnostic errors. Therefore, several criteria must be considered to evaluate model performance and obtain a comprehensive assessment:

- Accuracy: The percentage of images correctly classified.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.1)$$

- Precision (Overall Mean): The ability to minimize false positives across all categories, ensuring that a healthy individual is not mistakenly diagnosed with cancer.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3.2)$$

- Recall (Overall Mean): The model's sensitivity in detecting correct positive results.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3.3)$$

- F1 Scale (Overall Mean): A balanced assessment of accuracy and recall performance, representing their harmonic average.

$$\text{F1 Scale} = \frac{2*(\text{Precision}*\text{Recall})}{(\text{Precision}+\text{Recall})} \quad (3.3)$$

The Confusion Matrix is another metric used to evaluate the efficiency of classification models. The formula:

TP = Correct positives (sick, and the model said sick (correct))

TN = Correct negatives (healthy, and the model said healthy (correct))

FP = False positives (healthy, and the model said sick (incorrect))

FN = False negatives (sick, and the model said healthy (big false))

3.2 Proposed approach

In this section, we will explain the complete pipeline for a brain tumor detection system using MRI images, from data input to the final prediction stage. Using different models – MobileNetV2, EfficientNetB0, and InceptionV3 – we began by extracting the data, followed by preprocessing. This involved preparing the images before feeding them into the model, removing noise, resizing the image, segmenting the data, and normalizing it to improve data quality. Once the data was ready, we used our models. MobileNetV2 was distinguished by its speed, accuracy, and lightweight design, while EfficientNetB0 stood out for its balance between accuracy and speed. InceptionV3, it was distinguished by its deep and robust model. Next, in the Classification Layers, the data was passed to classification layers such as Flatten, Dense layers, Dropout, and Softmax (to produce the final class). The next stage is Compile, where a Loss function and Optimizer (such as Adam) are defined to prepare the model for training. Finally, in the Training Model stage, the model is trained on the training data using a series of epochs. Next comes the Evaluate phase. In this phase, the model is tested on previously unseen data to measure several parameters. The Prediction phase, the final stage of the modeling process, uses a new MRI image to determine the tumor type.

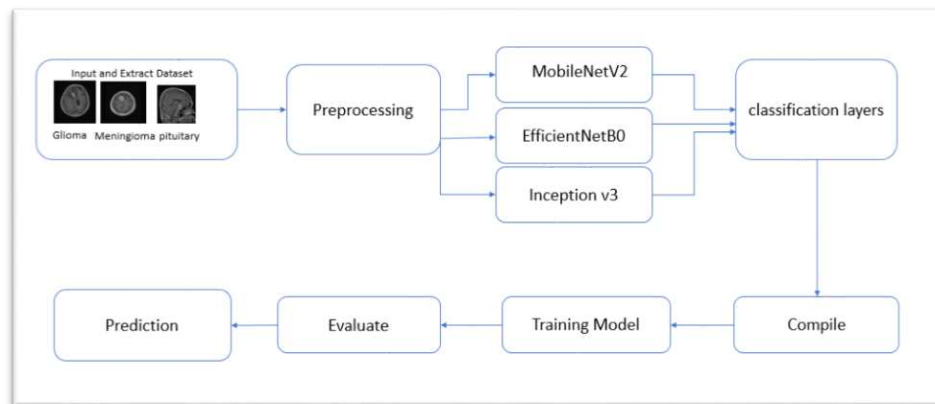


Figure 3.6 Proposed approach

Chapter 4:

Experimental Results and Discussion

4.1 Experimental Results:

4.1.1 Setting up the software environment (Google Colab work environment specifications):

The project was implemented using Google Colab, a cloud computing environment. It provides an Intel(R) Xeon(R) CPU @ 2.20GHz, 56320 KB of cache memory, a physical kernel, and support for a wide range of advanced processing instructions (flags) such as: fpu, vme, de, pse, tsc, msr, pae, mce, cx8, apic, sep, sse, and sse2. These are crucial for accelerating mathematical operations and deep learning. The address sizes are 46 bits physical and 48 bits virtual.

The environment was set up using Python 3, with the necessary libraries installed using pip install. Pre-installed libraries were used and connected to Google Drive to read, load, and process data. The appropriate model (MobileNetV2) was then built, trained, and loss and accuracy curves plotted. The model was evaluated, and finally, prediction testing was performed. A desktop equipped with an Intel Core i7-1255U processor and 16GB RAM was used for the experiment.

The data was divided into two main parts: a training set containing 80% of the data and a test set containing 20%. The first set is used to train the model, enabling it to learn patterns, extract features, characteristics, and relationships from the data. The second set, the test set, is used to evaluate the model after training and how it performs on real-world data. Each training or test data set contains four categories: glioma, meningioma, no tumor, and pituitary gland.

4.1.2 Results of model performance:

In this project, we used MobileNetV2, a high-performance, efficient, and lightweight CNN model.

The experiments were conducted using a deep learning path built on TensorFlow and Keras. The dataset was divided into two sets: training and testing. The input

images, which were $224 \times 224 \times 3$, underwent a series of downsizing stages. All pixel values were transformed from the range $[0, 255]$ to $[0, 1]$ using a resizing layer ($1./255$) to ensure stability and accelerate convergence. A common problem was an imbalance in the size of the data folders, so I increased the data. Several techniques were applied to increase the data during training, including: random horizontal flip, random rotation ($\pm 10\%$), random zoom (10%), and random contrast adjustment (10%). These improvements were implemented as a preprocessing layer. The model was then initialized with pre-trained weights on ImageNet, with the top classification layers deleted (`include_top=False`). In addition, all convolutional layers were frozen (`trainable=False`) to keep the learned features low and mid-level. The model consists of 53 layers, including 17 inverted bottleneck units. Each bottleneck unit contains an expansion layer, a deep detachable wrap layer, and a dropout layer, making the structure computationally efficient while maintaining strong feature extraction capabilities.

A custom classification head consisting of four layers (GlobalAveragePooling2D, Dense-128, Dropout-0.4, Dense-4) was added, bringing the total number of layers in the final model to 57. The softmax output layer corresponds to the four tumor classes.

The model was compiled using the optimizer Adam, with the sparse categorical cross-entropy loss function and the precision metric. The model was then trained for four epochs, with the training time measured using the Python ``time`` module.

Two sets of training and test data were fed via the optimized ``tf.data`` path with randomization and prefetching enabled.

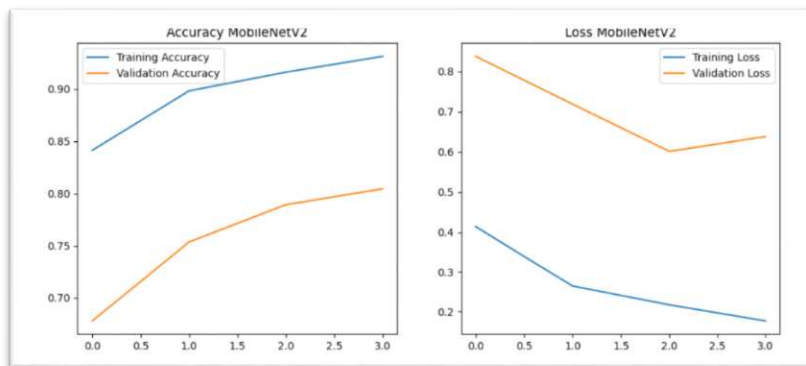


Figure 4.1 Accuracy and loss in MobileNetV2

The model in the graph showed that in Training Accuracy it started to rise until it reached 0.93, while Validation Accuracy reached 0.80, and this indicates that there is a gap between training and validation. As for Training Loss, it continued to decrease until it reached 0.17, and this indicates that the model is improving on new data. In Validation Loss, it decreased to 0.60 and then rose slightly to 0.64, and here it shows the beginning of overfitting.

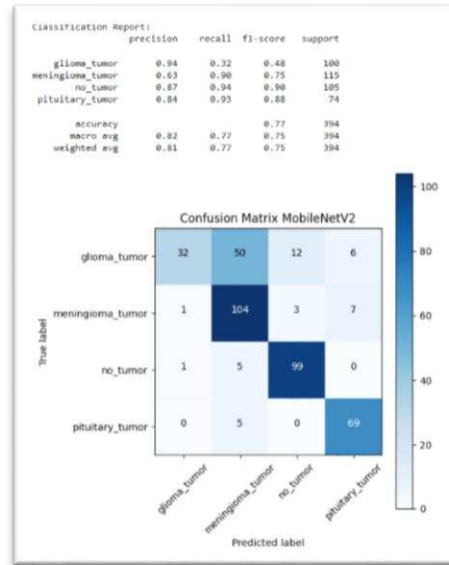


Figure 4.2 Confusion Matrix over result after apply MobileNetV2

The confusion matrix in Figure 4.2 shows the predicted classifications of brain tumor types. The model performed exceptionally well for pituitary tumor, followed by no tumor, and then meningioma. However, glioma tumor showed the highest error rate. The overall accuracy of the MobileNetV2 model is 82%.

The diagonal values represent the correct values for Glioma-32, Meningioma-104, No Tumor-99, and Pituitary-69. Other non-diagonal values represent incorrect classifications. The sum of each row represents the actual number of brain tumor types, while the sum of each column represents the predicted number of brain tumor types.

We also used the Inception V3 model, known for its accuracy and efficiency, which typically has between 48 and 300 layers. In our model, all pre-trained layers were frozen to minimize the risk of over-allocation and improve performance. Custom layers were added to enhance performance, such as the input layer, Inception V3, GlobalAveragePooling2D, Dropout (0.4), Dense (256), Dropout (0.3), and Dense. The learning rate was set to 1e-4, a small value that helps ensure training stability and improve data accuracy. Mixed accuracy was then enabled using the mixed_float16 policy to reduce memory consumption and further improve performance. The output layer was set to dtype="float32" to ensure the accuracy of softmax calculations. The model was trained for 8 cycles, and we observed improved accuracy, reduced loss, and no significant gap between training and test accuracy. This indicates that the model did not suffer from over-allocation, and the training time was 663 minutes and 3 seconds.

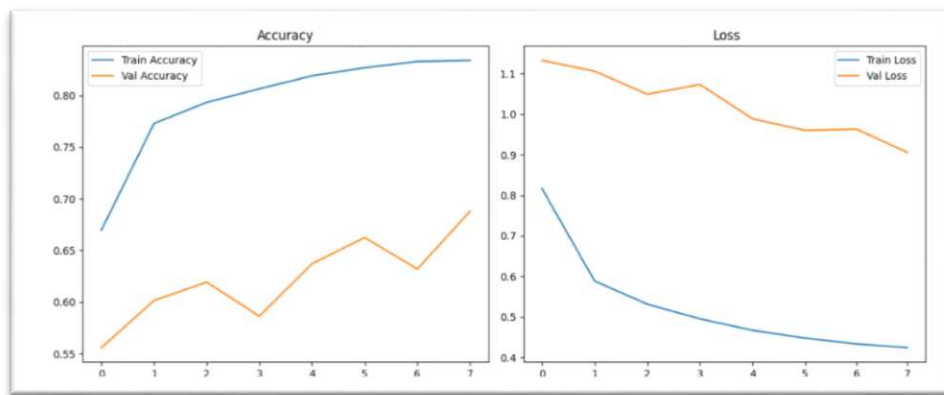


Figure 4.3 Accuracy and Loss InceptionV3

The InceptionV3 model, as shown in the graph, exhibited that Training Accuracy gradually increased to 0.83, while Validation Accuracy showed fluctuating improvement, reaching 0.69. This indicates that the model is learning. Training Loss continued to decrease to 0.42, suggesting that the model is improving on new data. Validation Loss started high at 1.13 and then decreased to 0.90.

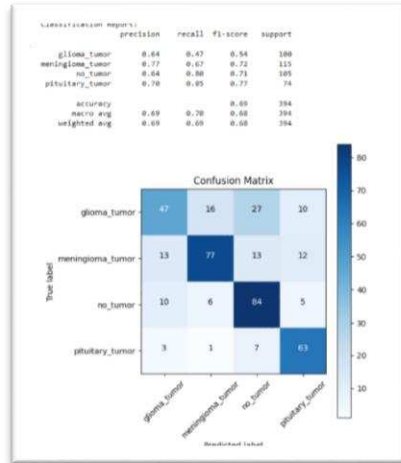


Figure 4.4 Confusion Matrix over result after apply InceptionV3

The confusion matrix in Figure 4.4 shows the predicted classifications of brain tumor types on the x-axis, while the correct classifications are shown on the y-axis. The diagonal represents the exact values for each type. The overall accuracy of the InceptionV3 model is 68%.

The maximum value indicates high accuracy. The diagonal values for Glioma-47, Meningioma-77, No Tumor-84, and Pituitary-63 indicate that the No Tumor classification is more accurate than the others. Other non-diagonal values represent incorrect classifications.

The sum of each row represents the actual number of brain tumor types. The sum of each column represents the predicted number of brain tumor types.

We used the EfficientNetB0 model, which is characterized by its lightness and efficient composite structure. It consists of 237 layers, with the top classification layers (include_top=False) removed. Additionally, all convolutional layers in the base structure were frozen during initial training to prevent weight updates and preserve the extracted general visual features. A custom classification head was added on top of the fixed base structure, along with a GlobalAveragePooling2D layer to reduce spatial dimensions, the Dropout (0.3) technique to reduce over-allocation, a four-neuron density layer, and a softmax activation function for multi-class prediction. The images were standardized to 224×224 pixels, and the batch size was set to 32. The model was trained for 16 epochs, with a training time of 472 minutes and 19 seconds. The Adam optimizer and a sparse

categorical cross-entropy loss function were used. The accuracy and loss curves plotted showed no significant variance, indicating no overfitting and increased training accuracy.

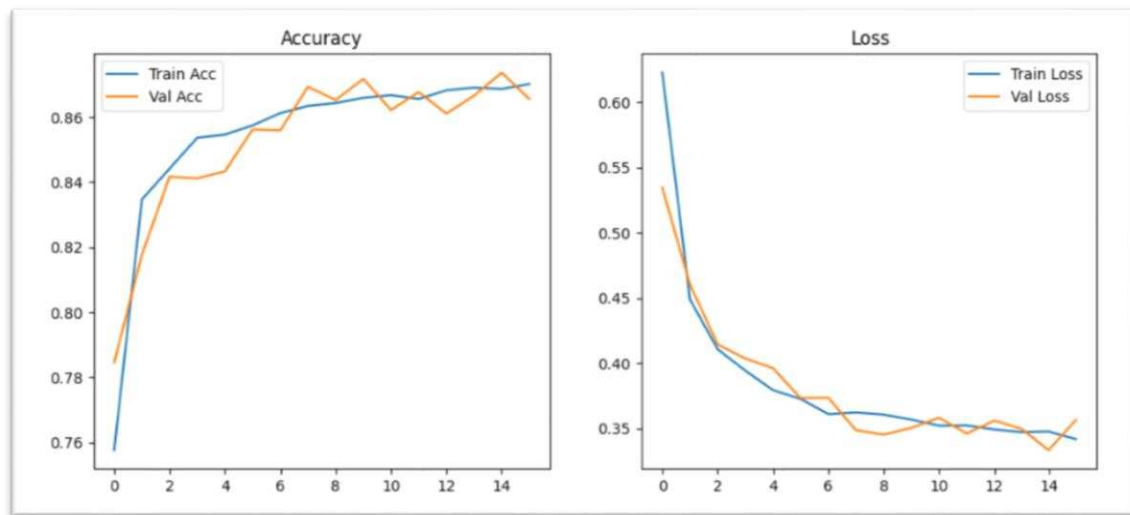


Figure 4.5 Accuracy and Loss EfficientNetB0

The EfficientNetB0 model demonstrated balanced and stable performance during the training process, outperforming all other graphs. Training Accuracy gradually increased to 0.87, while Validation Accuracy also improved to 0.87. The results were remarkably consistent and stable, with no significant gap between them. Training Loss and Validation Loss decreased smoothly to 0.34, exhibiting a similar pattern. The model showed no signs of overfitting, indicating the efficiency of the EfficientNetB0 model's architecture in extracting features while maintaining a balance between learning processes.

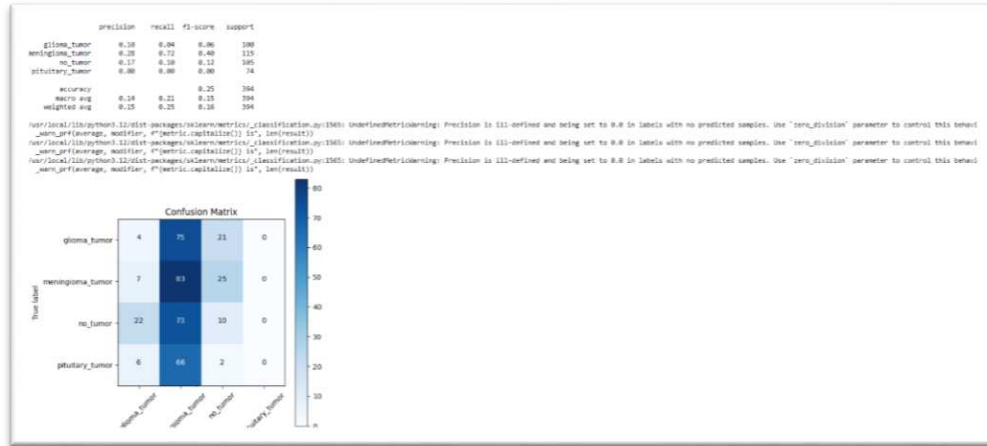


Figure 4.6 Confusion Matrix over result after apply EfficientNetB0

The confusion matrix in Figure 4.6 shows that EfficientNetB0 performed poorly in classifying brain tumor images. For example, it completely failed to recognize and predict pituitary tumors. In the no-tumor category, it could not distinguish between healthy and affected images. The model clearly had a strong bias towards classifying most images as meningioma. For example, 75 glioma images were classified as meningioma, 66 pituitary images as meningioma, and 73 no-tumor images as meningioma. This indicates that the model suffers from several problems and challenges, including bias towards a single category, as mentioned in the previous section, and poor generalization. It was unable to differentiate between types and did not learn enough features from the images to classify them correctly. This model leads to misdiagnosis, making it unsuitable for clinical application without further improvement.

Table 4.1 : Models performance average

Approaches	Models	Accuracy	Precision	Recall	F1-score
Pre-trained CNN	MobileNetV2	82%	81%	79%	79%
	EfficientNetB0	68%	14%	21%	15%
	Inception v3	68%	69%	70%	68%

Table 4.1 clearly compares the average performance of the three models. The table includes four evaluation metrics: Accuracy, Precision, Recall, and F1-score.

The results show that MobileNetV2 is the best and most suitable choice due to its high scores. Inception v3 performed acceptably, while EfficientNetB0 was the worst and requires future configuration adjustments.

Table 4.2: Training Time for each Model

Approaches	Models	Number of Epoch	Training time	¹ Epoch
Pre-trained CNN	MobileNetV2	4	65	16.25
	EfficientNetB0	16	472	25
	Inception v3	8	663	82.88

Table 4.2 presents a comparison of the training times for each model, as well as the number of different training cycles for each model. This analysis helps in evaluating the computational efficiency of each model and selecting the best, most suitable, and most efficient one. The MobileNetV2 model recorded the shortest training time and demonstrated its efficiency and superiority, followed by the EfficientNetB0 model, where each epoch took approximately 25 minutes, thus requiring slightly more training time than the previous model due to the larger number of epochs. This indicates that this model needs more time to converge. On the other hand, the Inception v3 model recorded the longest training time, indicating that its complex architecture increases the number of computations, and therefore significantly increases the training time.

Table 4.3: Models performance in each class

Approaches	Models	Classes	Precision	Recall	F1-score
Pre-trained CNN	MobileNetV2	Glioma	88%	57%	69%
		Meningioma	75%	79%	77%
		No tumor	76%	97%	85%
		pituitary	85%	82%	84%
	EfficientNetB0	Glioma	10%	4%	6%
		Meningioma	28%	72%	40%
		No tumor	17%	10%	12%

		pituitary	0%	0%	0%
	Inception v3	Glioma	64%	47%	54%
		Meningioma	77%	67%	72%
		No tumor	64%	80%	71%
		pituitary	70%	85%	77%

Chapter 5

Conclusion and Future Work

5.1 Conclusion:

In conclusion, early detection of brain tumors can accelerate treatment and reduce mortality. In this study, we used several models, and the results varied. Based on previous tests, the MobileNetV2 model proved to be the most effective for this task, achieving the highest accuracy on the data. The EfficientNetB0 model demonstrated excellent stability during training, but its lower test accuracy suggests that stability alone does not guarantee better generalization. The InceptionV3 model performed the worst overall, exhibiting instability and weak predictive performance. Its accuracy was comparable to EfficientNetB0. Furthermore, the results were enhanced across all accuracy metrics and uncertainty matrices. The use of data augmentation techniques was crucial in overcoming the difficulties posed by unbalanced datasets, allowing for the development of more versatile models that can be generalized across a wide range of imaging scenarios. Given the increasing incidence of brain tumors worldwide, this strategy aligns with the urgent need for effective diagnostic technologies.

5.2 Future Work:

Future research should focus on improving the performance of EfficientNetB0 and InceptionV3 by fine-tuning hyper parameters, scaling up data, and increasing data volume to fully leverage their architecture advantages and balance tumor types. It should also include testing and comparing other models to determine the best fit, as well as conducting clinical trials with appropriate privacy considerations. This testing will ensure the models can be generalized across different devices, such as CT and MRI, and across different patient populations, ages, and ethnicities. These studies could greatly benefit low-income healthcare facilities and remote areas, reducing the workload for healthcare practitioners and improving the speed and accuracy of results.

References

- [1] Ben Brahim, S., Dardouri, S., & Bouallegue, R. (2024). Brain tumor detection using a deep CNN model. *Applied Computational Intelligence and Soft Computing*, 2024
- [2] Abdusalomov, A., Mukhiddinov, M., & Whangbo, T. (2023). Brain tumor detection based on deep learning approaches and magnetic resonance imaging.
- [3] Khan, A. B., Abbas, S., Khan, M., Farooq, U., Khan, W., Siddiqui, S. H., & Ahmad, A. (2022). Intelligent model for brain tumor identification using deep learning. *Applied Computational Intelligence and Soft Computing*, 2022.
- [4] Kadam, A., Bhuvaji, S., & Deshpande, S. (2021). Brain tumor classification using deep learning algorithms. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*.
- [5] Iftikhar, S. H., Anjum, N., Siddiqui, A. B., Ur Rehman, M., & Ramzan, N. (2025). Explainable CNN for brain tumor detection and classification through XAI-based key features identification. *Brain Informatics*.
- [6] AlTahhan, F., Khouqeer, G. H., Saadi, S., Elgarayhi, A., & Sallah, M. (2023). Refined automatic brain tumor classification using hybrid convolutional neural networks for MRI scans.
- [7] Brain Tumor Classification (MRI). (n.d.). Kaggle. [Brain Tumor Classification \(MRI\)](#)
- [8] Brain Tumor Dataset. (n.d.). Kaggle. [Brain Tumor Dataset](#)
- [9] Ahmad, S., Eali, S. N. J., Nakka, T. R., Chihaoui, N., Alsubait, T., & Khanam, H. (2026). A scalable and reliable deep learning framework for enhanced brain tumor detection and diagnosis using AI-based medical imaging. *Frontiers in Neuroscience*, 13.
- [10] Gasmi, K., Ben Aoun, N., Alsalem, K., Ben Ltaifa, I., Alrashdi, I., Ben Ammar, L., Mrabet, M., & Shehab, A. (2024). Enhanced brain tumor diagnosis using combined deep learning models and weight selection technique. *Frontiers in Neuroinformatics*, 26 November 2024.
- [11] Hasan, M. R., Rudra, S. M., Karmoker, N., Yousuf, M. A., Akhter, J., Al-Moisheer, A. S., Alyami, S. A., & Moni, M. A. (2026, June 10). *3D MRI reconstruction and brain tumor diagnosis using deep learning with explainable AI*.

[12] Khan, R., Taj, S., Khan, Z. U., Khan, S. U., Khan, J., Arshad, T., & Ayouni, S. (2025, December). High-precision brain tumor classification from MRI images using an advanced hybrid deep learning method with minimal radiation exposure. *ScienceDirect*.

[13] [14] Rahman, M. A., Miah, M. B. A., Hossain, M. A., & Hosen, A. S. M. S. (2025, June 5). *Enhanced brain tumor classification using MobileNetV2: A comprehensive preprocessing and fine-tuning approach*.

[15] Naseer, M. M., Veerabhadraiah, V., Ramji, B. R., Lokeshappa, S. K. K., Nagendra, S. C., Ramesh, Y. G., & Puttegowda, K. (2025, October 21). *An optimized EfficientNet-B0 framework for multi-class brain tumour detection and classification from MRI images*.

[16] Raihan, J., Mahmud, Z., Ridita, S. A., & Bhattacharya, A. (2023). *PhytoCare: A hybrid approach for identifying rice, potato and corn diseases*. Department of Computer Science and Engineering, University of Liberal Arts Bangladesh.

[17] Alhassan, A. M., & Altmami, N. I. (2025). IV3TM: Inception V3 enabled bidirectional long short-term memory network for brain tumor classification. *PLOS ONE*, 20(10), e0335397.

[18] Rahman, M. A., Miah, M. B. A., Hossain, M. A., & Hosen, A. S. M. S. (2025). Enhanced brain tumor classification using MobileNetV2: A comprehensive preprocessing and fine-tuning approach. *Applied*