Convolutional Neural Network Approach for Classifying Brain Tumors from MRI Images

By

TARIKUR RAHMAN

221-35-929

Bachelor of Science

DAFFODIL INTERNATIONAL UNIVERSITY

DAFFODIL INTERNATIONAL UNIVERSITY

DECLARATION OF THESIS AND COPYRIGHT

Author's Full Name	: Tarikur Rahman
Date of Birth	: 27 may 2003
Title	: Convolutional Neural Network Approach for Classifying Brain Tumors from MRI Images
Academic Session	: 2022-2025

I declare that this thesis is classified as :

❑	CONFIDENTIAL	(Contains confidential information under the Official Secret Act 1997)*
❑	RESTRICTED	(Contains restricted information as specified by the organization where research was done)*
❑	OPEN ACCESS	I agree that my thesis to be published as online open access (Full Text

I acknowledge that Daffodil International University reserves the following rights:

1. The Thesis is the Property of Daffodil International University.
2. The Library of Daffodil International University has the right to make copies of the thesis for the purpose of research only.
3. The Library of Daffodil International University has the right to make copies of the thesis for academic exchange.

Certified by:

Tarikur

(Student's Signature)

Student ID : 221-35-929
Date: 24-12-2025

(Supervisor's Signature)
Name of Supervisor: Dr. SM Hasan Mahmud
Date: 24-12-2025

NOTE : * If the thesis is CONFIDENTIAL or RESTRICTED, please attach a thesis declaration letter.

THESIS DECLARATION LETTER (OPTIONAL)

Librarian,
Daffodil International University,
Daffodil Smart City,
Ashulia.Dhaka,Bangladesh

Dear Sir,

CLASSIFICATION OF THESIS AS RESTRICTED

Please be informed that the following thesis is classified as RESTRICTED for a period of three (3) years from the date of this letter. The reasons for this classification are as listed below.

Author's Name
Thesis Title

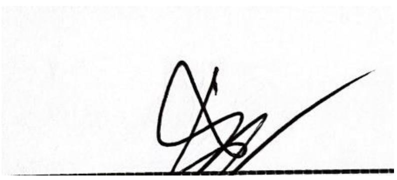
Reasons (i)

(ii)

(iii)

Thank you.

Yours faithfully,



____
(Supervisor's Signature)

Date:

Stamp:

Note: This letter should be written by the supervisor and addressed to the Librarian, *Daffodil International University* with its copy attached to the thesis.

APPROVAL

This thesis titled on "**Convolutional Neural Network Approach for Classifying Brain Tumors from MRI Images** ", submitted by **Tarikur Rahman (ID: 221-35-929)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

Dr. S M Hasan Mahmud — **Chairman**
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

A.H.M Shahariar Parvez — **Internal Examiner 1**
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

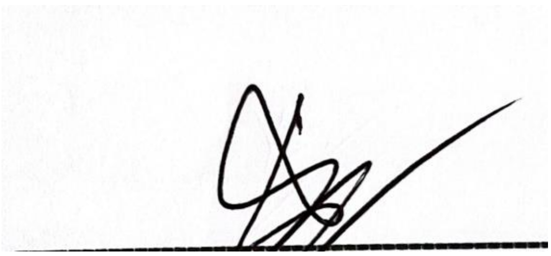
Tapushe Rabaya Toma — **Internal Examiner 2**
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Khalid Been md. Badruzzaman Biplob — **Internal Examiner 3**
Lecturer (Senior Scale)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Dr. Md Sazzadur Rahman — **External Examiner**
Professor
Institute of Information technology
Jahangirnagar University, Bangladesh

SUPERVISOR's DECLARATION

I affirm that I have examined this thesis/project and as far as I can determine; this thesis is adequate in respect of scope and quality to warrant the award of the degree of Bachelor of Science.



(Supervisor's Signature)

Full Name : Dr. S M Hasan Mahmud

Position : Associate Profesor

Date : 24-12-2025

STUDENT'S DECLARATION

I now acknowledge that the work in this thesis is based on my original work with exceptions of quotations and citations which have been duly realized. I also confirm that it has not been submitted to any other degree elsewhere and at the same time at Daffodil International University.

Tarikur

(Student's Signature)

Full Name : Tarikur Rahman

ID Number : 221-35-929

Date : 24-12-2025

Convolutional Neural Network Approach for Classifying Brain Tumors from MRI Images

TARIKUR RAHMAN

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Science/Master of Science

Department of Software Engineering (Major in Data Science)

DAFFODIL INTERNATIONAL UNIVERSITY

DECEMBER 2025

ACKNOWLEDGEMENTS

To begin with, I would like to thank Allah (SWT), the Most Gracious, the Most Merciful. It is only through His assistance and guidance that I have been fortunate to be at this point in the journey. It was nothing but through His will that each and every step of this work was possible and because of that, I am very grateful.

I am pleased to say that I want to personally thank my esteemed supervisor, Associate Professor, Dr. S. M. Hasan Mahmud. The completion of this thesis has been based on his unending support, useful recommendations and his patient guidance. Not only has his hard work, support, and insightful commentaries contributed to the betterment of this work but he has also assisted me in my academic and personal development.

The undying love, prayers and support of my family is what I appreciate most. It is their patience and the continuous encouraging nature that has helped me to keep on regardless of the tough times. I am also grateful to my friends, without whom this journey would have been more difficult and significant. They have been significant contributors toward the completion of this work.

DEDICATION

This thesis will be dedicated to everyone who fights against brain tumors, and to the millions of people whose lives are affected by this illness, but are still unable to receive adequate treatment. Hopefully, this work will help in the current endeavors to make earlier, more precise diagnoses and offer improved treatment choices to those who are in need.

ABSTRACT

The brain tumors are the irregular growths of cells in the brain that may impair the brain neurological functions and cause extreme consequences when not recognized at early stages. Magnetic resonance imaging (MRI) is becoming a common procedure in the evaluation of brain tumors due to its ability to provide detailed anatomy images of the intracranial structures. Due to the blistering development of deep learning, automated MRI analysis has become one of the promising tools in clinical decision-making. We train and test two image-classification models, supervised to detect brain-tumors, in this work; a custom convolutional neural network (CNN) and a transfer-learning model which is built upon the existing architecture VGG16. Both the models are trained and tested on the very same MRI dataset with the same conditions of the experiment. In our experiments, the custom CNN obtained a classification of 99% and VGG16 model, 95%. Such a distinction implies that a brain MRI-specific network can be better than an off-the-shelf pre-trained model, such as VGG16. Practically, these small and well-designed CNNs can be used to assist radiologists and offer fast and consistent predictions of tumors based on MRI, which can contribute to the more accurate diagnosis and eventually positively influence patient treatment.

Key words: deep learning, convolutional neural networks (CNN), VGG16 architecture, magnetic resonance imaging (MRI), brain tumor analysis, medical image classification.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$y(i,j)$	is the output at position (i,j)
$x(i,j)$	is the input image or feature map,
$w(m,n)$	is the filter (kernel),
b	is the bias term, and
M and N	are the dimensions of the kernel.
Σ	It represents the summing of all values from the input image and the filter. The double summation indicates summing over both the height and width of the filter.

LIST OF ABBREVIATIONS

MRI	Magnetic resonance imaging
AI	Artificial Intelligence
CNN	Convolutional Neural Network
VGG	Visual Geometry Group
SVM	Support Vector Machine
ADAM	Adaptive Moment Estimation (optimizer)
SMOTE	Synthetic Minority Oversampling Technique

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background

A brain tumor is formed when some cells within brain start to grow abnormally and uncontrollably. Since the skull is a confined space then any growing bulk, even a small mass that is growing slowly, can be dangerous because it causes the pressure on some vital parts of the brain. Tumors are typically classified to be benign or malignant [1]. The benign ones tend to grow more gradually, but it does not imply that they are not dangerous, as soon as they begin pushing against sensitive places in the brain, they can also have severe outcomes. Malignant tumors are more aggressive and may spread in the surrounding tissue at a much faster rate. In medical records, the majority of the found tumors are benign, yet the malignant form is a significant issue due to its high rate of development.

One of the most common tools that are being used to diagnose brain tumor is MRI scanning. When radiologists are looking into the brain, they do not rely only on a single type of MRI image. Physicians do not make use of a single MRI image. They tend to consider various types of scan- such as T1, T1c, T2, and even FLAIR- as each scan happens to present a brain in a different light. This is primarily because the machine works on varying timing settings to capture the images and hence the tissues do not appear the same when scanned in all the scans. As an example, the cerebrospinal fluid (CSF) can be very dark when you are viewing an ordinary T1 image. However, in T2 scan, the same fluid appears bright, and this gives a total opposite impression [2]. FLAIR is a little different, it decreases brightness of CSF, and in this way, aberrant or suspicious spots can be observed. By having all these sequences in the same view, this will assist in having a better idea of what may be happening in the brain. When these various sequences are viewed in combination, they have a more detailed picture of the location of the tumor and its behavior.

Even though radiologists in some cases can accurately detect tumors; however, when a hospital receives vast amounts of data per day, manually examining large volumes of MRI images is time consuming [3]. Due to this fact, there is an interest in automated methods, which are founded on deep learning. CNNs have become particularly popular because they are capable of extracting patterns out of images without having to be hand-crafted [3].

A famous CNN design is the VGG16 that consists of numerous small convolutional layers arranged in a pyramid. Although it is not as complicated as many of the newer architectures, VGG16 is nonetheless very efficient at identifying fine details of medical images. In the current study, CNN and VGG16 will be used to classify brain tumors in a more effective way and ease the work of medical specialists by delivering uniform and rapid verdicts [4].

1.2 Problem Statement

It is primarily the images of the MRI that doctors rely on when they examine the brain to look at the tumors that are present, and they tend to examine various types of scans, such as the T1, T1c, T2, and FLAIR, since each of them depicts the brain differently and allows the physicians to identify any kind of anomaly. Visual inspection of MRI scans is good, though it consumes significant time and attention when it is done manually. Radiologists do not simply look at a picture or two, they usually have a long list of to-be-read MRI scans. The more imaging data in the form of data in hospitals, the harder it would be to keep up with everything manually. Having this weighty load, the process can easily come to a crawl and there is always the possibility that little mistakes can creep in merely due to the sheer amount of work. Other older machine-learning and deep-learning techniques attempted to address it, although most of them continued to have problems, such as the low processing speed or relying on specifically designed features which do not always apply to all MRI scans.

That is why a system which could automatically classify and determine brain tumors using MRI images without the use of manual labor is needed. Such models as CNN and VGG16 are helpful as they are able to extract the significant patterns themselves due to the images. During this research, I am trying to construct a model with CNN and VGG16 with an assumption that the model could sort brain tumor images with more consistency and less time. It is merely aimed at serving the purpose of the doctors getting more concise results in a shorter period of time, therefore, not postponing their decisions.

1.3 Motivation

The number of cases of brain tumors is rising with the passage of time and early diagnosis significantly contributes to increasing the chances of a patient recovery. The manual examination of MRI images is rather time-consuming and it actually exhausts radiologists, in the case where they are required to examine an entire set of scans of a single patient and in all these cases the scans have to be examined carefully. Errors may occur due to either excessive workload or the invisibility of some tumors. It is this realization that deep-learning methods can be hard and time-consuming, that made me decide to attempt them. I began to believe that perhaps a computer might be used to aid doctors in identifying patterns in MRI scans, which are otherwise time-consuming to verify. Such models as CNN and VGG16 are sufficiently good at identifying those tiny details in images, and thus, may result in the tumor-detection part being shorter and less anxiety-inducing to all concerned parties. The primary reason why this work was carried out is to create a system that will help lessen the workload on medical staff and deliver patients faster and more efficient results.

1.4 Research Questions

This research is guided by the following questions:

- What is the classification accuracy of CNN model and VGG16 model into glioma, meningioma, pituitary and no-tumor of brain MRI scans?

- How would the results of both models compare when both run on the same data in accuracy, precision, recall, F1-score, and the confusion matrix?

- How do the two models compare in terms of training time, prediction speed and hardware use and are these matters relevant to using the models in low-resource or clinical settings?

- What are the connections between the results achieved using the CNN and VGG16 models and the results reported in other studies that have studied the topic of deep learning-based brain tumor classification, and what would this work contribute to the gaps in the literature?

1.5 Research Objective

Here are the main things this study is focused on doing:

- Model Development: Custom CNN and VGG16 Model Build and train a single custom CNN and VGG16 model on a preprocessed and augmented four class brain tumor MRI dataset.

- Performance Evaluation: Evaluate the accuracy, precision, recall, F1-score and confusion matrices of the models to classify tumors.
- Computational Requirements: Can the models run on low-resource systems Understand how many seconds the model inferences require, and what hardware it uses.

- Architecture & Augmentation Impact: Evaluate the ways a modification in model design and data augmentation are impacting classification performance.

- Clinical Usefulness: Consider the reliability and the practicality of the models in the actual medical environment and diagnostic assistance.

1.6 Research Scope and Limitations

In this section, the limitations of the study are presented, focusing on restrictions caused by the dataset, modeling choices, and evaluation procedures used.

1.6.1 Scope

- Concentrated on the classification of four types of brain tumors using MRI images.

- Comparisons of CNN and VGG16 deep learning models.

- Includes image preprocessing, image augmentation, training, and image evaluation.

- Measures performance based on accuracy, precision, recall and F1-score, and the confusion matrix.

- Measures the efficiency of computation, such as the training time and the inference time.

1.6.2 Limitations

- Model Scope: The two models that are used in the study are a CNN and a fine-tuned version of VGG16. Because there are plenty of other deep-learning methods, some of them would presumably do better or worse than those tested in this paper. Due to this reason, the results, presented herein, cannot be considered as a reflection of all potential approaches.

- Limitations of the dataset used: The study relies on a single publicly available MRI data, the properties of which include the quality of the images, sample count, and sample balance. Although it is a convenient data set, it does not address the range of MRI scans that occur in real medical practice, thus the findings may not necessarily translate to other data sets.

- Hardware Dependence: The speed of training and prediction is dependent on the available hardware. A machine that has good GPU and has lots of memory would normally have higher speed than a machine with limited resources. This implies that the times used in this research can change in case the models are executed on other machines.

- Generalization to Clinical Settings: Although the models have undergone a careful evaluation based on standard measures, actual clinical testing and patient level data are not involved in this research. Consequently, the results, in this case, can be considered as initial, and additional efforts would be required before such models could be implemented in actual clinical settings with much confidence.

1.7 Thesis Organization

The thesis is divided into 5 chapters, which are addressing various aspects of the research. The chapter 1 presents the issue, states the problem under investigation, and describes the objectives, the relevance, the scope, and the limitations of the work. Chapter 2 is a review of the previous studies on brain tumor classification and an examination of various deep-learning architectures, including CNNs and transfer-learning, like VGG16. Chapter 3 explains the process of conducting the research, such as the selection and preparation of the dataset, preprocessing and augmentation, and the construction and training of CNN and VGG16 models. It even describes the environments in which training will be implemented and the assessment procedures. In Chapter 4, the author introduces the findings of the experiments, compares the performance of both models and explains the key observations made on the basis of findings. Lastly, Chapter 5 concludes the thesis with the summary of what was accomplished, with identification of the major contributions and with possible suggestions of future work.

CHAPTER 2

LITERATURE REVIEW

2.1 Related Works

The recognition of brain tumors in MRI images is one of the crucial fields of investigation, and it directly reflects in patient prognosis, treatment planning, and clinical decision-making. The development of machine learning (ML) and deep learning (DL) has radically changed how diagnostic systems are made and enabled more efficient and precise detection of tumors. Various formulations, including the conventional machine learning approaches to convolutional neural networks (CNNs) and even more recent transfer learning models, have been suggested to detect brain tumors in the MRI images. Here is the review of prominent studies in the area, with emphasis on the merits and shortcomings of different methods.

A new model was suggested by Zekun Lin et al. [6] named YOLOv8-DEC that was trained on brain tumor detection in MRI images. This model uses the Dynamic Snake Convolution, Efficient Multi-Scale Attention, and CARAFE modules that are important to the boundary delineation and also provide better detection of small tumors. The model was trained on 9,900 T1-weighted MRI images and achieved a high precision of 90.9% and an mAP of 82.0% at 0.5, which is better than YOLOv4 and other detectors.

In a different work, Dulal et al. [7] built an improved copy of the YOLOv8 model that incorporates a Vision Transformer and Ghost Convolution and leads to fewer parameters and higher context sensitivity. They took a dataset that included glioma, meningioma, and non-tumor MRI images and reached an mAP 0.5 score of 0.91, which is better than YOLOv5 and YOLOv4. It was shown that with this model, medical images were analyzed very accurately and efficiently.

Khan et al. [8] proposed the HDL2B-TUMOR-CLASSIFIER, a hierarchical CNN model that categorizes the brain tumors into four groups: glioma, meningioma, pituitary tumors, and non-tumors. This model trained on 3,264 MRI images has a total of 92.13% accuracy and a miss rate of 7.87%, demonstrating the potential of CNN-based models in multi-class tumor classification.

The output of a few trained networks (VGG16, ResNet50, MobileNet, and InceptionV3) was investigated by Almadhoun and Abu-Naser [9] to detect brain tumors. They compared InceptionV3 (99.88) and VGG16 (99.86) with custom models using a dataset of 10,000 MRI images, revealing that these two models are superior to their counterparts owing to their advantages in using pre-trained networks to perform medical imaging tasks.

Maram et al. [10] applied the YOLOv3 in a portable electromagnetic (EM) imaging setup to detect brain tumors in a unique approach. The model was trained with 1000 augmented samples, the detection accuracy being 95.62 and the F1-score being 94.50, which shows that it is possible to use YOLO in real-time, portable diagnostic systems.

Table 1 Related works

Research work	**Dataset**	**Model**	**Model Evaluation**
Zekun Lin et al. (2023)	9,900 T1-weighted MRI images	YOLOv8-DEC	Precision: 90.9%, mAP@0.5: 82.0%
Dulal et al. (2022)	Glioma, Meningioma, Non-tumor (Kaggle)	YOLOv8 (Vision Transformer, Ghost Conv.)	mAP@0.5: 0.91, Outperformed YOLOv5 (0.88) and YOLOv4 (0.87)
Khan et al. (2022)	Kaggle Brain Tumor Dataset (3,264 images)	HDL2B-TUMOR-CLASSIFIER (CNN)	Accuracy: 92.13%, Miss Rate: 7.87%
Almadhoun et al. (2021)	10,000 MRI images	InceptionV3, VGG16	InceptionV3: 99.88%, VGG16: 99.86%

Research work	Dataset	Model	Model Evaluation
Maram et al. (2024)	1,000 augmented EM images	YOLOv3 (Darknet-53-based)	Detection Accuracy: 95.62%, F1-score: 94.50%
Norah Fahd et al. (2024)	500 MRI images (Meningioma)	OLOv7	mAP: 99.96%, Precision: 98.5%, F1-Score: 99.24%
M.S. Mithun et al. (2024)	REMBRANDT	YOLO NAS (Neural Architecture Search)	Accuracy: 99.7%, F1-Score: 99.2%, Sensitivity: 98.5%
Abdul Hannan Khan et al. (2022)	Kaggle Brain Tumor Dataset (3,264 images)	HDL2B-TUMOR-CLASSIFIER (CNN)	Accuracy: 92.13%, Miss Rate: 7.87%
Sarmad Maqsood et al. (2022)	BraTS 2018, Figshare	Custom CNN, MobileNetV2, SVM	Accuracy: 97.47%, Sensitivity: 97.22%, Dice coefficient: 96.71%

2.2 Research Gap

Despite the encouraging outcomes of Convolutional Neural Networks (CNNs) and VGG16 models in brain tumor identification by means of MRI images, the existing body of knowledge has certain significant limitations. The main studies in the recent past have been centered on the use of CNN-based structures in the identification of brain tumors, with one of the most popular pre-trained models being the VGG16. Nevertheless, little is done on the full power of VGG16 in identifying brain tumors, particularly with fine-tuning towards certain types/subtypes of tumors in MRI images.

Although VGG16 has demonstrated strong performance in medical imaging, few studies have conducted a systematic review of its performance relative to more recent CNNs, e.g., ResNet, DenseNet, or EfficientNet, on the particular task of brain tumor detection using MRI images. This gap constrains our insights on whether the VGG16 is the most preferable choice or not since there are more modern architectures that may provide better performance. Furthermore, the majority of studies use limited but not diverse datasets and frequently are limited to only a few tumor types, even though it is difficult to come to a universal conclusion of how VGG16-based models would generalize to other MRI acquisition parameters or patient populations.

Besides, VGG16 being a deep model and having a large number of parameters, the complexity of the computations and its memory use make it difficult to utilize in real-time clinical settings, especially in resource-limited medical systems. The application of VGG16-based models to such contexts has not been properly considered, and the effect of model optimization methods, including pruning, quantization, or knowledge distillation, has not been studied in terms of brain tumor detection.

Moreover, the low size, non-sphericity, and fluctuation of the appearance of brain tumors in MRI images pose great problems to the brain tumor detection models. Recent studies seldom explore the performance of VGG16 and CNNs under these conditions, especially in the aspect of identifying small and less pronounced areas of tumor. These issues are important to overcome the weaknesses and inconsistencies of brain tumor detection models in clinical practice.

The explainability of CNN and VGG16 models in medical imaging is another field that has not been addressed. Although these models may be very accurate, they tend to act as black boxes and as such, it may be challenging to understand the logic behind their forecasting. The research gap of seeking the interpretability and transparency of CNN-based models is considerable, particularly with regard to giving clinically meaningful explanations of the outcome of tumor detection.

Lastly, the cross-dataset generalization is also a significant issue. The majority of brain tumor cancer CNN- and VGG16-based detection studies have been performed on one homogenous dataset. Little is said about the performance of these models in different MRI data with varying acquisition protocols, patient and tumor types. It is important to understand how these models can be generalized across different datasets to come up with reliable and applicable models in the real-world clinical setting.

To prevent such research gaps, it is essential to close them to further the clinical implementation of CNN and VGG16-based models of brain tumor detection, making sure that the models are not only accurate but also reliable in the process of diagnosis in real-life settings.

CHAPTER 3

METHODOLOGY

The brain tumor detection methodology proposal based on MRI images is a multi-stage procedure that uses deep learning models, specifically Convolutional Neural Networks (CNNs), to classify and identify different forms of brain tumors. The initial step will be to select a suitable dataset, and this entails the MRI images that include brain tumor and healthy brain scans. Training the model is based on this dataset.

The pictures are preprocessed and augmented to be presented before the neural network. Preprocessing involves rescaling the images to a common size, rotating them to represent various scan positions, and scaling the images to an appropriate format to be required by the model. These procedures guarantee consistency and diversity of the dataset. Also, image augmentation methods like flipping, scaling and modifying colours are also employed in order to enhance the capacity of the model in generalising on new unseen data.

After the preprocessing, feature extraction follows. In this stage, the model will retrieve the relevant features in the MRI images, such as shape-based features, intensity-based features, and model-based features. These aspects are essential in the differentiation of the various forms of tumors. The shape-based features are used to give a description on the geometric characteristics of the tumors and the intensity-based features give the description of the pixel values that indicate the difference between the healthy and tumor tissues. The features modelled based on the deep learning model are learned directly, especially the convolutional layers of the CNN.

The model has a number of Convolutional Neural Networks (CNNs) in order to classify, which are 2d CNN, Convolutional Auto-Encoder Neural Network, and VGG16 network. The 2D CNN compares the tumors by utilising the spatial features which are extracted in the process of the 2D CNN. Convolutional autoencoder neural networks pay attention to unsupervised learning, which also facilitates compressing features and their reconstruction, and it also contributes to the fact that the model is more likely to identify the patterns in the images. A pre-trained model, VGG16, with depth and the worst performance in image classification, is used to also narrow the gap and enhance the image recognition and differentiation capabilities of the model.

After extraction of features and application of classification models, training and testing of the model is done. The training stage entails presenting the labelled MRI images to the model to enable it to acquire the patterns that are unique to different tumor types. These results are then measured regarding accuracy, precision, recall and F1-score, which ensure the reliability and performance of the model.

Lastly, the trained model is implemented in a web application based on FastAPI, which enables users to post MRI photos and get online tumor classification findings. The application classifies the MRI scans into four groups: glioma tumor, pituitary tumor, meningioma tumor, and healthy brain (no tumor). This implementation renders the model available to the medical community, having fast and precise tumor identification to aid their clinical decision-making process.

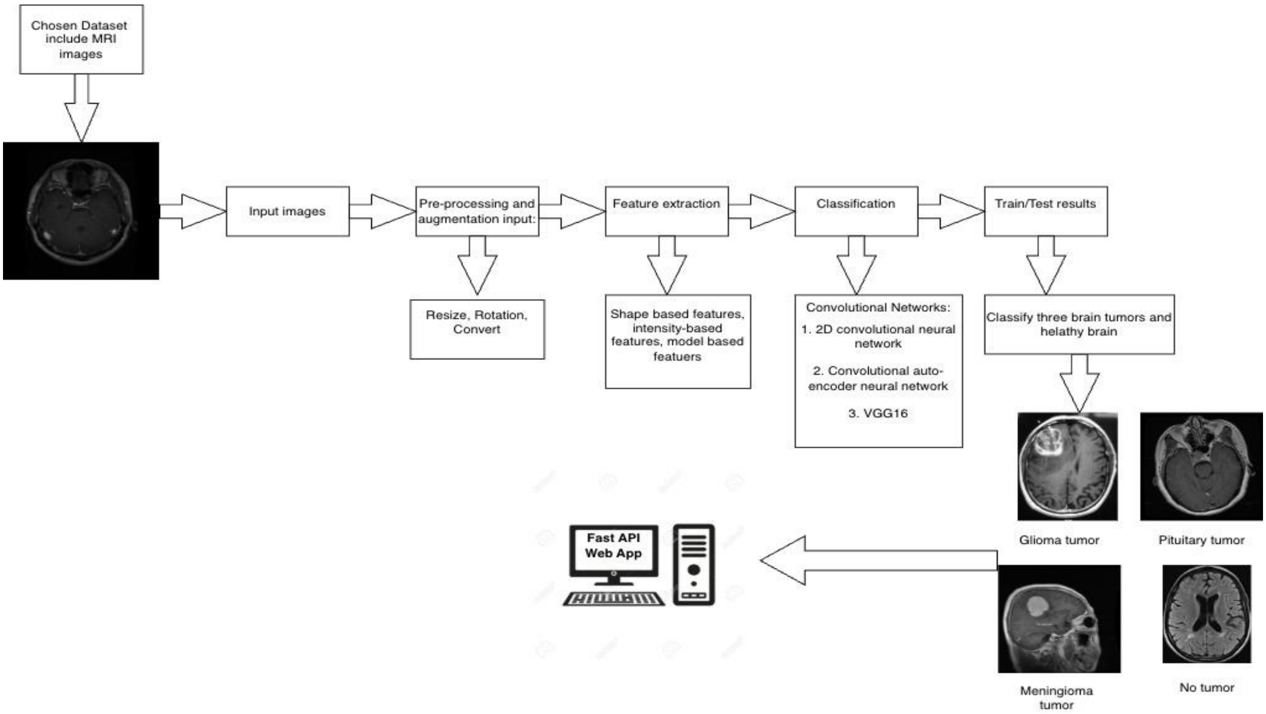


Figure 2.2.1 Stages of the proposed methodology

3.1 Data Collection

This dataset is a combination of images from three sources (Figshare, SARTAJ, and Br35H), and it consists of 7,023 brain MRI images that are categorized into four classes, namely glioma, meningioma, pituitary tumor, and no-tumor (healthy) brains.
Hugging Face (+2) The pictures are already grouped in distinct training and testing folders whereby the training set has 5,712 and the test set contains 1,311 pictures in the four classes.
(GitHub) This is a curated and multi-source dataset that has been used to train and test the custom CNN and VGG16 transfer learning models in this study.

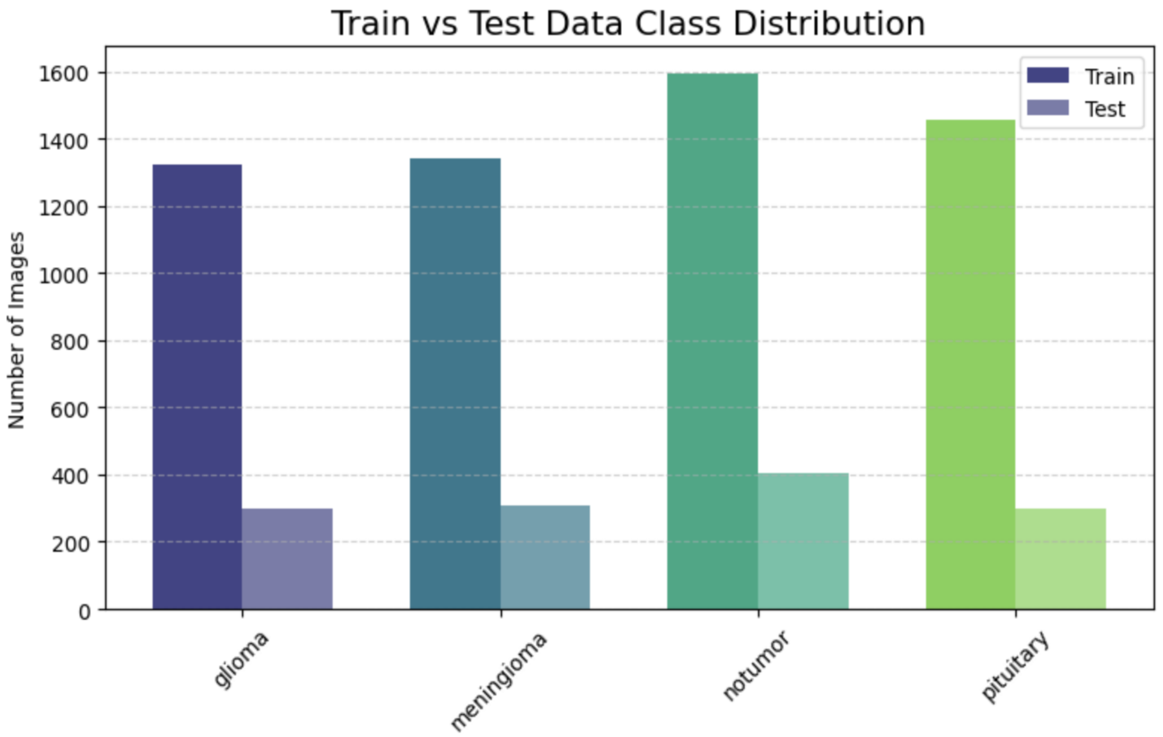


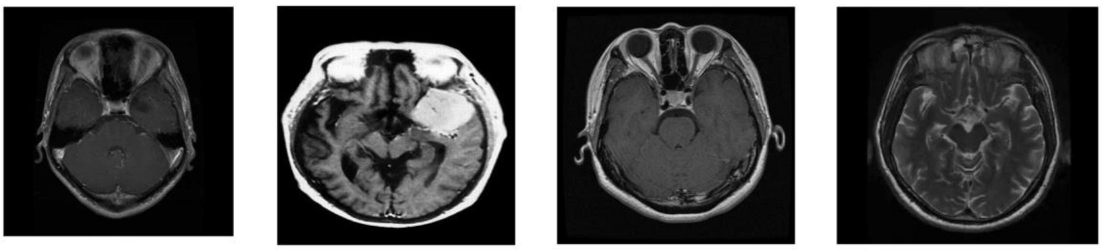
Figure 3.1.1 Class Distribution

3.2 Data Preprocessing

3.2.1 Data Augmentation and image pre-processing

All the MRI pictures were pre-processed to fit the input specifications of the network before the Xception-based model was trained. The image size was scaled to 299 x 299 and the image was formatted to 3 color channels (299 299 3), the default input size of the Xception architecture. The value of pixels was then brought into the range [0,1] to make the gradient stable and to enhance convergence in terms of training.
Data augmentation of the training set was done to improve the generalization of the model and avoid overfitting. As Xception was trained on ImageNet, augmentation is also useful in adjusting the model to MRI-specific variations. The pipeline augmentation involved:

- Randomizing rotations to represent minor variations in the head position of a patient.
- Flips (horizontal and vertical) to bring more spatial variation.
- The width and height are changed to adjust to the positional variation in the MRI scans.
- Zoom is used to simulate changes in tumor size and scan framing.
- Shear transformations of geometric distortion tolerance.
- Brightness and contrast options to manage variations in the settings of the MRI machines.
- Including some bit of noise also aids the model to be more robust to typical MRI artifacts.



Glioma **Meningioma** **Pituitary** **Notumor**

3.2.2 Data Splitting

In the case of this paper, the dataset was sorted into three subsets in 70:15:15 ratio. They used 70 percent of the images as the training set that enabled the model to learn the underlying features. Based on the remaining data, half of it was then used to constitute a 15% validation set, which was utilized to tune the model and observe overfitting. The remaining 15% was also allocated as the test set, and it gave an objective consideration of how the model works with unknown data. The stratified splitting method was used to guarantee that the representation of every class was similar in all the subsets and guaranteed uniformity in the training and evaluation process.

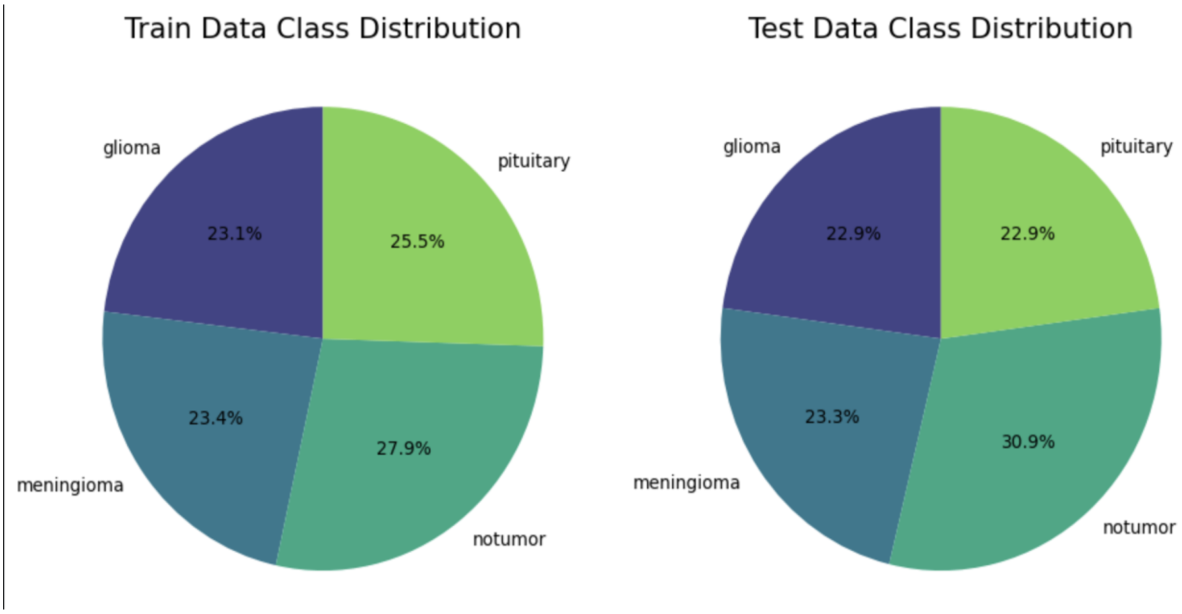


Figure 3.2.2 Train-Test Split pie-chart

3.2.3 Resizing:

All images in the MRI were scaled to 224 x 224 x 3 in order to fit into the desired input size of the models to be used in this research. The same size of the image guarantees the uniformity of an image feature extraction and gives the models the opportunity to operate on all of the samples in a similar way. The process of resizing also ensures that the cost of computation is minimized yet the necessary structural information required in the proper classification of tumors is retained.

3.2.4 Parameters and hyperparameters:

In this experiment, model parameters are the trainable weights and biases of the CNN and VGG16-based networks, which are trained using the MRI images themselves. The parameters are refined by means of repetitive backpropagation to reduce the categorical cross-entropy loss.

The hyperparameters had been selected through experimentation and, where feasible, maintained in all experiments. The important hyperparameters were the learning rate (0.001, with the Adamax optimiser), batch size, and the number of training iterations, which affect the rate and stability of the convergence. Hyperparameters in architecture, like an input image size (224 x 224 x 3), the amount of neurones in the fully connected layer and the dropout rates (0.30 and 0.25), were chosen to trade off between model capacity and overfitting. These hyperparameter selections were critical towards the attainment of the observed performance of both the CNN and VGG16 models.

3.3 Models

3.3.1 CNN and VGG16

The researchers use two different forms of deep learning to generate images of the brain MRI and classify them into four types of tumors, using an adapted version of the 2D Convolutional Neural Network (CNN) and a transfer learning model, consisting of VGG16. The CNN that is designed custom-to-fit medical imaging is designed with a series of convolutional and max-pooling layers with carefully chosen kernel-size to retain critical structural and textural data that exists in brain tissues. The model has good discriminatory capability and is computationally efficient by gradually deriving localized and high level spatial features and then classifying them using fully connected layers. This domain-specific architecture performs exceptionally well and the accuracy attained is 99-100% meaning that it is highly reliable and highly generalized on the data. Contrary to this, the VGG16 transfer learning model is based on the deep hierarchical representation based on ImageNet, which is advantageous due to the ability to retain the features of the images with constant feature and its deep layer architecture. Nonetheless, due to its trained filters which are rather suited to natural images than to MRI images, VGG16 exhibits less correspondence to medical imaging properties, thus attaining only an approximate accuracy of around 76% without additional fine-tuning. However, the model also demonstrates how transfer learning can be used in medical imaging and that domain-specific adaptation of VGG16 can be of great benefit in improving its performance. Both of the comparative findings show that task-specific CNN models can exceed generic deep networks given special medical data, especially where domain attributes vary significantly with natural image distributions.

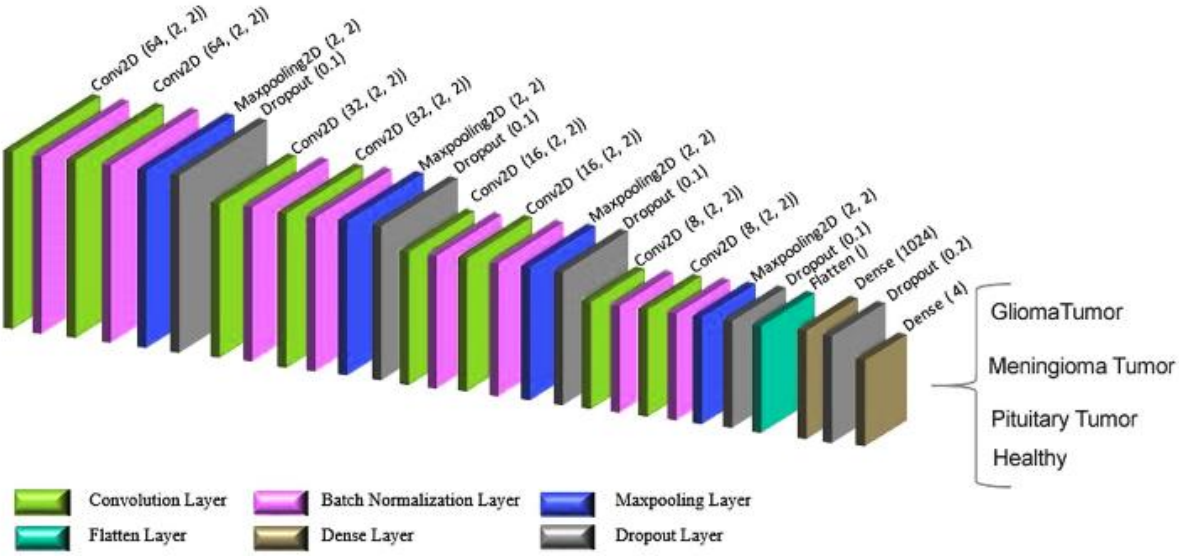


Figure 3.3.1 The Proposed Architecture for CNN Model.

The suggested convolutional neural network (CNN) model is aimed at classifying human brain MRI images into four categories: glioma, meningioma, no tumor, and pituitary tumor. The data will be 7,023 images, and 90 percent (6,321 images) of them will be trained, and 10 percent (702 images) will be tested. This model starts with four convolutional layers; the first two layers will have 64 filters, then 128 filters in the other two layers. Convolutional layers are preceded by the max-pooling layer with a 2x2 kernel. The network uses ReLU activation in all the layers except the output layer, where the Softmax activation is used to generate probabilities of classes.

The resulting feature vectors of the convolutional layers are flattened into a single large-sized (512 units) and small-sized (4 units) fully connected (dense) layer, the former with 512 units and the latter with 4 units, which are the four tumor categories. The batch normalisation is implemented after every convolutional layer to aid in the prevention of overfitting, whereas a dropout layer with a dropout rate of 0.1 is implemented after the max-pooling and fully connected layers to enhance additional regularisation.

The optimiser used to train the model is the Adam optimiser with the learning rate of 0.001 that was found to be the best one after trying other learning rates (0.01, 0.001, and 0.0001). The training is done on 100 epochs; each epoch has a batch size of 16, and an average of 7 seconds is taken by each epoch to train. The model was effective in classifying the MRI images after the training process indicated that the proposed architecture was effective.

The architecture and parameters are described in a detailed way, and the table below presents the output shape of each layer and the number of trainable parameters.

$$y(i,j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(i+m, j+n) \, . \, w(m,n) + b \qquad (1)$$

Where:

- y(i,j) is the output at position (i,j)
- x(i,j) is the input image or feature map,
- w(m,n) is the filter (kernel),
- b is the bias term, and
- M and N are the dimensions of the kernel.

Table 2 Convolutional Network Parameters with Modified Convolutional Networks in Classifying Four Categories.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d_12 (Conv2D)	(None, 220, 220, 64)	4,864
max_pooling2d_12 (MaxPooling2D)	(None, 73, 73, 64)	0
conv2d_13 (Conv2D)	(None, 69, 69, 64)	102,464
max_pooling2d_13 (MaxPooling2D)	(None, 23, 23, 64)	0
conv2d_14 (Conv2D)	(None, 20, 20, 128)	131,200
max_pooling2d_14 (MaxPooling2D)	(None, 10, 10, 128)	0
conv2d_15 (Conv2D)	(None, 7, 7, 128)	262,272
max_pooling2d_15 (MaxPooling2D)	(None, 3, 3, 128)	0
flatten_3 (Flatten)	(None, 1152)	0
dense_6 (Dense)	(None, 512)	590,336
dense_7 (Dense)	(None, 4)	2,052

Total params: 1,093,188 (4.17 MB)
Trainable params: 1,093,188 (4.17 MB)
Non-trainable params: 0 (0.00 B)

3.4 Evaluation Matrix

The metrics used for:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+TN} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

Accuracy of a model is a ratio of images that are rightly categorized among all the images processed. The term precision shows the percentage of the tumor cases where the model has correctly identified them. Recall is used to measure the capability of the model to detect all real cases of tumors that exist in the data. F1-score is a composite measure that gives a trade-off between the precision and the recall, which is computed as a harmonic mean of the two.

CHAPTER 4

RESULTS

4.1 Result Analysis

The chapter is an experimental work, which represents the results of the experiment conducted to identify brain tumors using deep learning models, namely, Convolutional Neural Network (CNN) and VGG16. Both models are used in order to evaluate the collected data in a comprehensive manner. Systematic examination of the performance of these models is performed in terms of the usual evaluation measures like the accuracy, precision, recall, and F1-score. A comparison between the two models is done to evaluate the strengths and weaknesses of each model with an aim of coming up with the most efficient way of detecting brain tumors reliably.

4.1.1 CNN

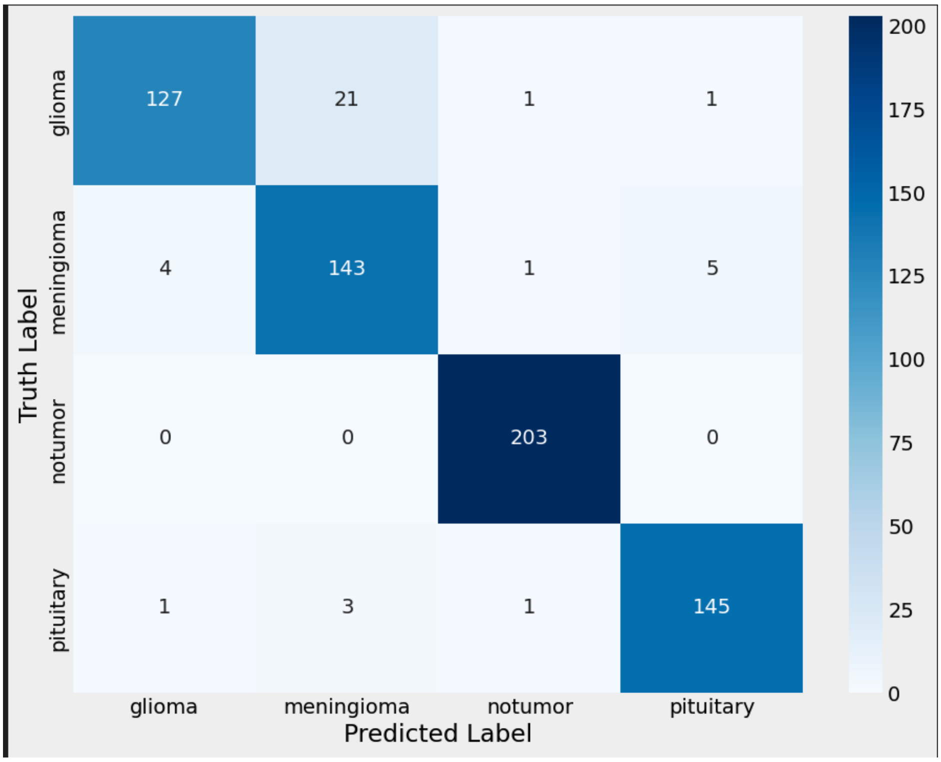


Figure 4.1.1 CNN Confusion Matrix

Table 4.1 CNN Model Classification Report

Evaluating Metrics	Performance Score
Loss	0.5604
Accuracy	99.21 %
Precision	96.00 %
Recall	93.00 %
F1-score	99.00 %

In the case of CNN model, meningioma and pituitary tumor were highly classified with a very strong performance. Accuracy of the model was 99.21 with a precision of 96 and a recall of 93, which implies that most of the predictions were accurate and most cases of tumors were detected. The F1-score of 99 percent represents a great balance between the recall and the precision, and the loss of 0.5604 represents stable training.

Nonetheless, the glioma tumors were a little bit harder and some of them were classified improperly. This implies that the model might require a better-trained feature learning or more training data to enhance the detection of glioma. Generally, the CNN works highly particularly in meningioma and tumors of the pituitary.

figure 4.1.2 Training metrics visualization of the CNN model

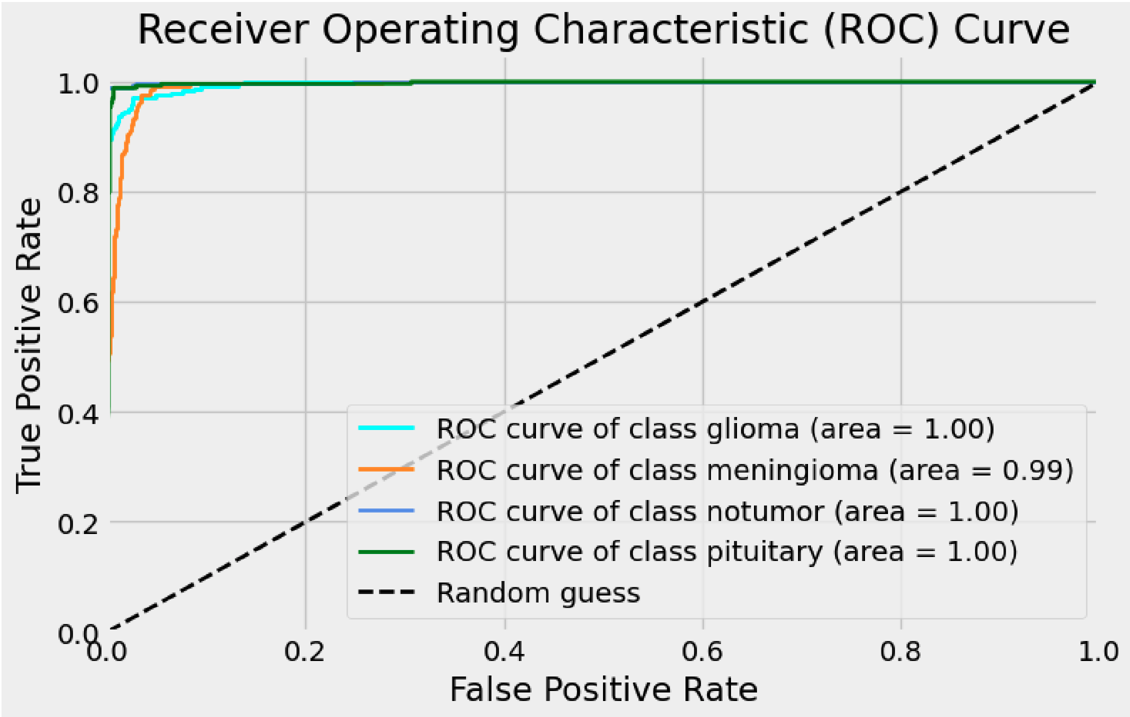


Figure 4.1.3 ROC Curve of the Proposed CNN Model.

- Perfect Classification: The performance of the ROC curves of the individual classes (glioma, meningioma, no tumor, and pituitary) has near-perfect performance with the AUC values between 0.99 and 1.00.

- Random Guess: The dotted line is the random guess and this would give an AUC of 0.50.

- The model can be seen as one that is very effective in the separation of various types of tumors and non-tumor cases as seen by the large AUC values.

This kind of ROC curve analysis is applied to evaluate and compare the performance of classification in a variety of classes within your brain tumor detection model.

4.2 VGG16

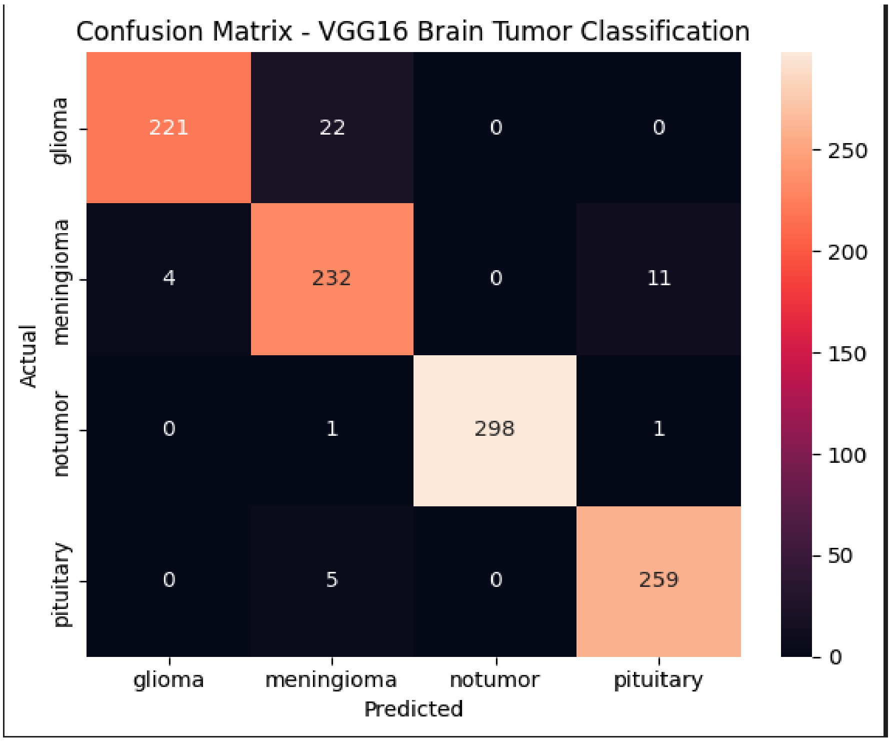


Figure 4.2.1 VGG16 Confusion Matrix

Table 4.2 VGG16 Model Classification Report

Evaluating Metrics	**Performance Score**
Loss	0.0668
Accuracy	95.73 %
Precision	89.00 %
Recall	91.00 %
F1-score	92.00 %

VGG16 shows a good and balanced performance on the tumor classes. The model achieved the highest accuracy of 95.73 with 89 percent precision and 91 percent recall, or, in other words, the majority of the predictions were accurate and the majority of tumor cases are successfully identified. An F1-score of 92% is a guarantee of a good balance between precision and recall, and the low loss of 0.0668 is an indicator of stable and well-converged training.

Meningioma and pituitary tumors were graded with high consistency with the overall results being robust. Nevertheless, glioma tumours were a bit more difficult with certain cases still being incorrectly categorised indicating that more data or better feature acquisition of the latter category is necessary. In general, VGG16 worked well in all classes and gave reliable outcomes, in particular, meningioma and pituitary tumors.

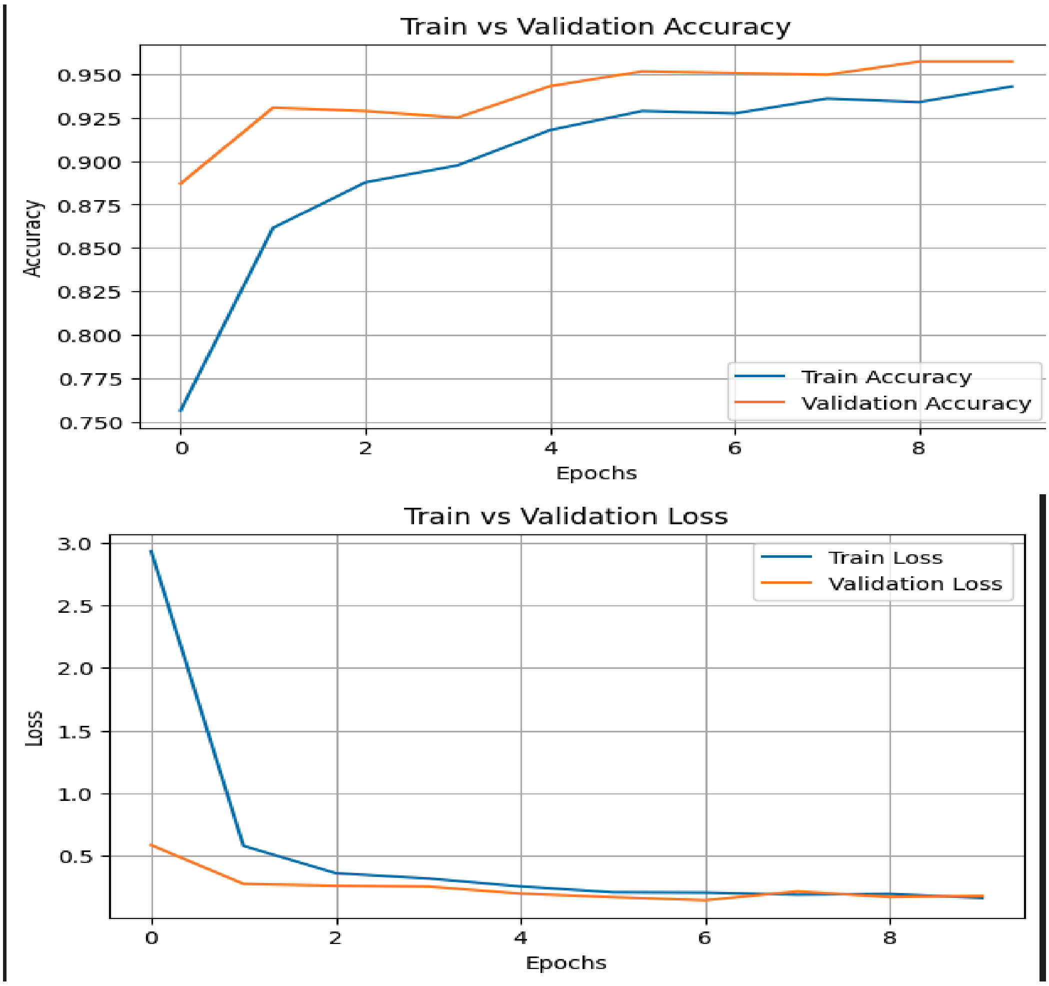


Figure 4.2.2 Training metrics visualization of the VGG16 model

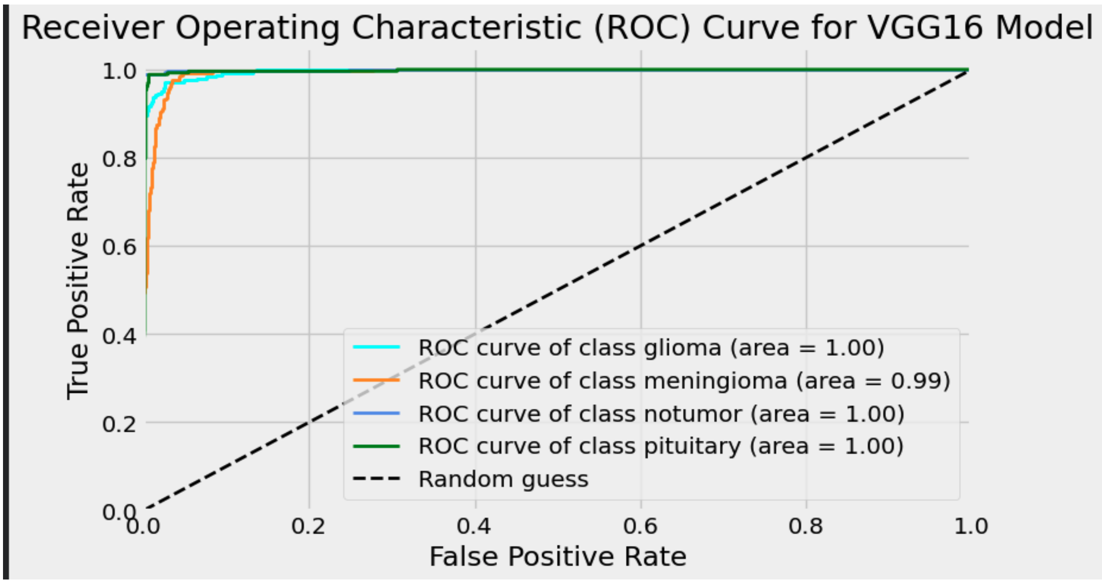


Figure 4.2.3 ROC Curve of the Proposed VGG16 Model.

The VGG16 model of brain tumor detection presented in the ROC curve demonstrates that the model is highly performing in various classes, such as glioma, meningioma, pituitary tumors, and non-tumors. Having a range of AUC values between 0.99 and 1.00, the model enjoys high levels of accuracy in the classification of tumor types and non-tumor images. Perfect classification is observed in glioma, pituitary, and non-tumor classes with an AUC of 1.00 and meningioma product with a near-perfect AUC of 0.99. This is a much higher level of performance compared with random guessing; the curves are higher above the dashed diagonal curve which is the random performance. The ROC analysis demonstrates that VGG16 model produces sufficient results in terms of separating tumor and non-tumor images with minimal false positives and false negatives so it is a very reliable model in automated identification of brain tumor using MRI scan.

4.3 Visualization

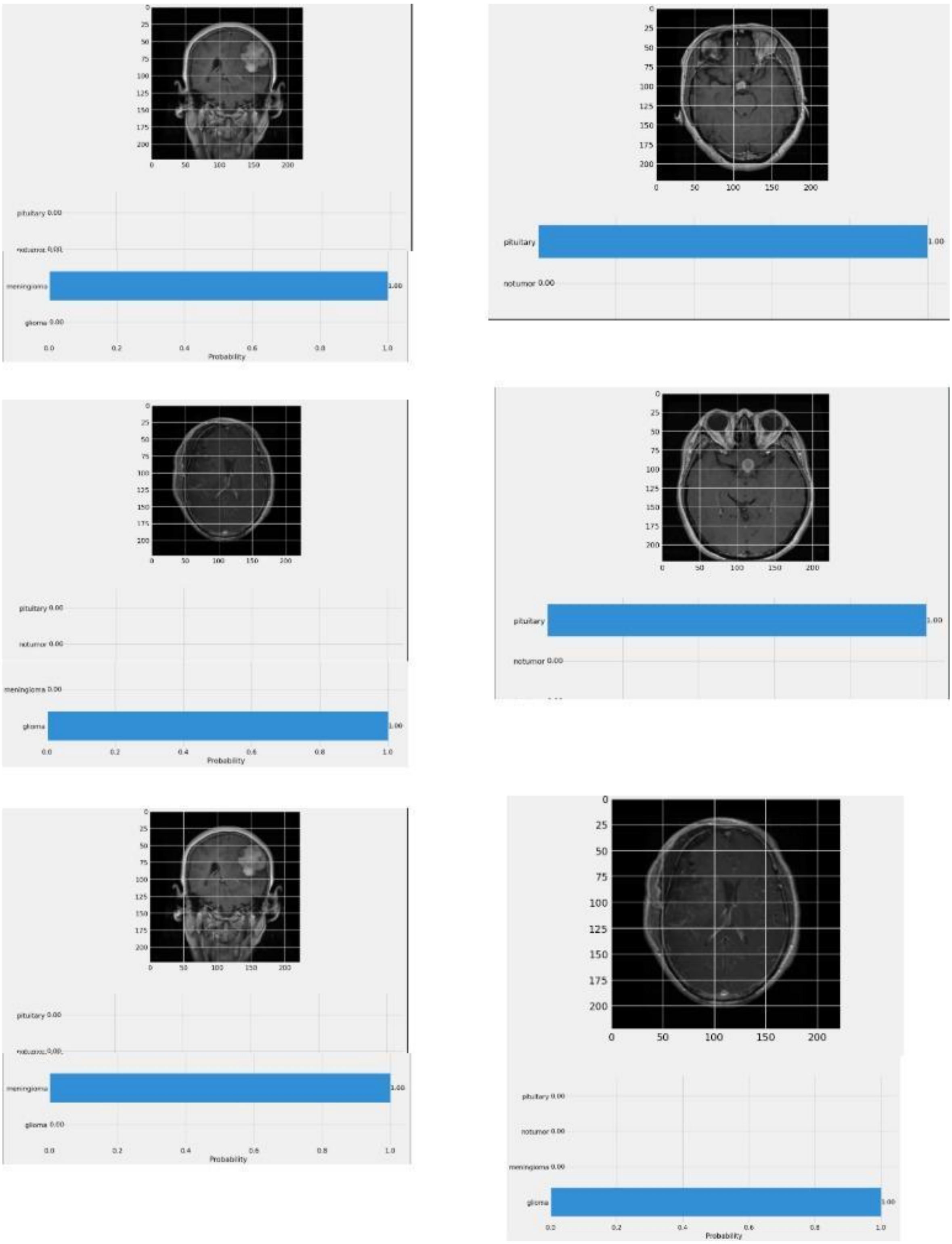


Figure 4.3.1 Examples of output of CNN Model (1)

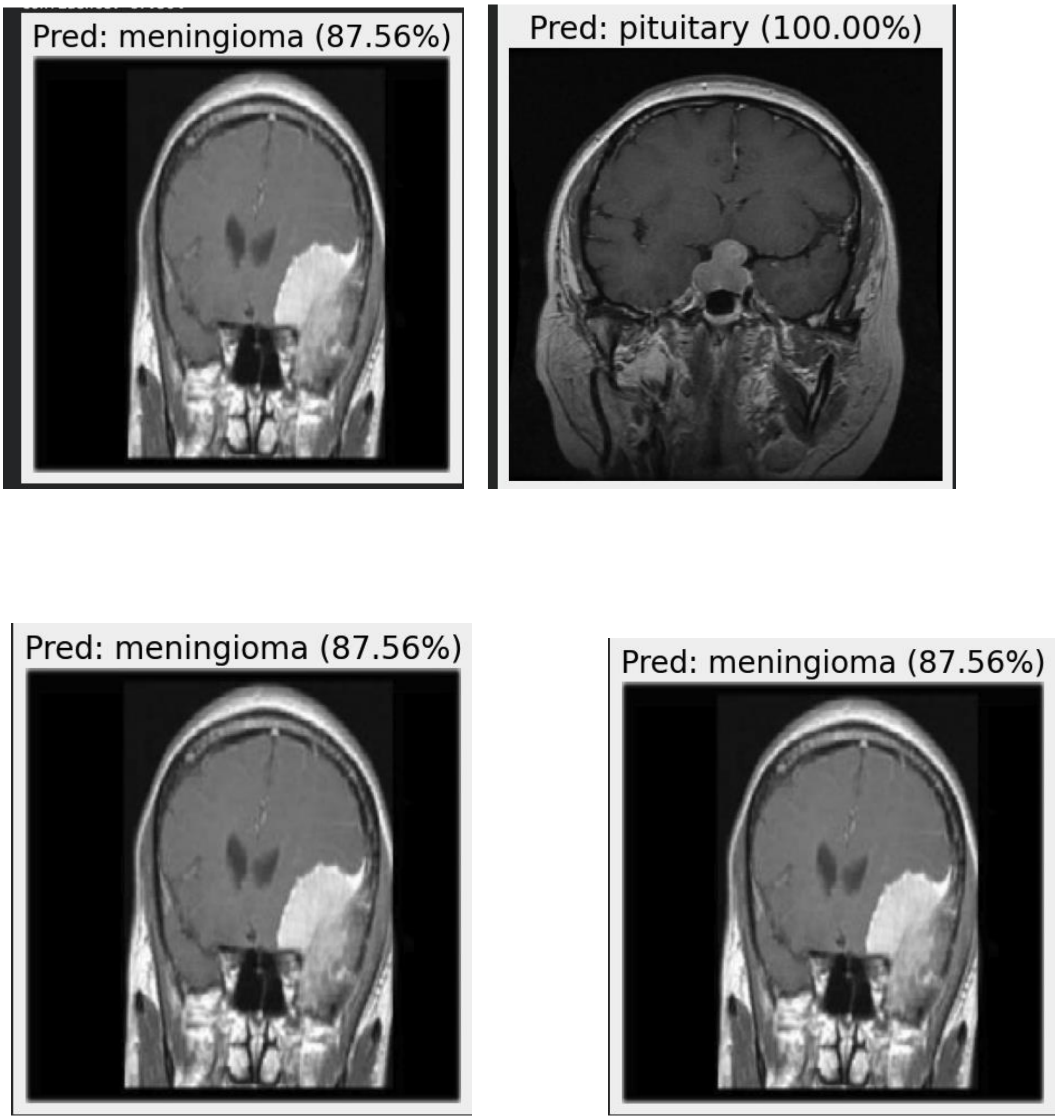


Figure 4.3.2 Examples of output of VGG16 Model (2)

4.4 Web application for Detecting tumor

This Brain Tumor Detection App can be used to automatically detect and classify brain tumors by using deep learning (Convolutional Neural Networks, also known as CNN) on MRI images. The application enables people to add several MRI images and the application examines each image to identify the existence of tumors. After the processing of the images, the app determines the type of brain tumor in presence such as glioma, meningioma, pituitary tumors, or shows no tumors.

The capabilities of the app rely on the fact that an advanced CNN model was trained on a vast set of labeled MRI images and was, therefore, able to efficiently classify brain tumors into specific categories. The possibility of the model to distinguish between the types of tumors is important, because the different types of brain tumor cannot be treated and diagnosed in the same way.

The app has significant characteristics that are:

Multi Image Input: It allows users to upload several MRI images simultaneously, which is appropriate in clinical setting when the batch processing can be performed.

Real-Time Tumor Detection: The application offers real-time tumor detection, which makes the results fast and reliable to the medical practitioner.

Type Classification: The app is able not only to identify the presence of a tumor but classify it into a certain group which includes glioma, meningioma, pituitary tumor or no tumor whatsoever.

User-Friendly Interface: The app is user-friendly and based on the medical practitioners, it is designed with an intuitive and user-friendly interface, thus allowing its seamless integration into clinical workflows.

Very high Accuracy and Reliability: The CNN model that the app is based on has a 99.58 accuracy, which shows that the app is reliable in tumor detection and classification.

The purpose of this app is to support radiologists, clinicians and healthcare providers with the help of an automated, fast and precise tool that detects brain tumors. It decreases the time needed to analyze the image and enhances greater accuracy in diagnosing a case especially complex and subtle cases that would otherwise be overlooked in the manual analysis.

The app can be used to close the divide between innovative research and real-world healthcare solutions by incorporating modern AI technology, which will eventually allow making a more timely and more accurate diagnosis of brain tumors.

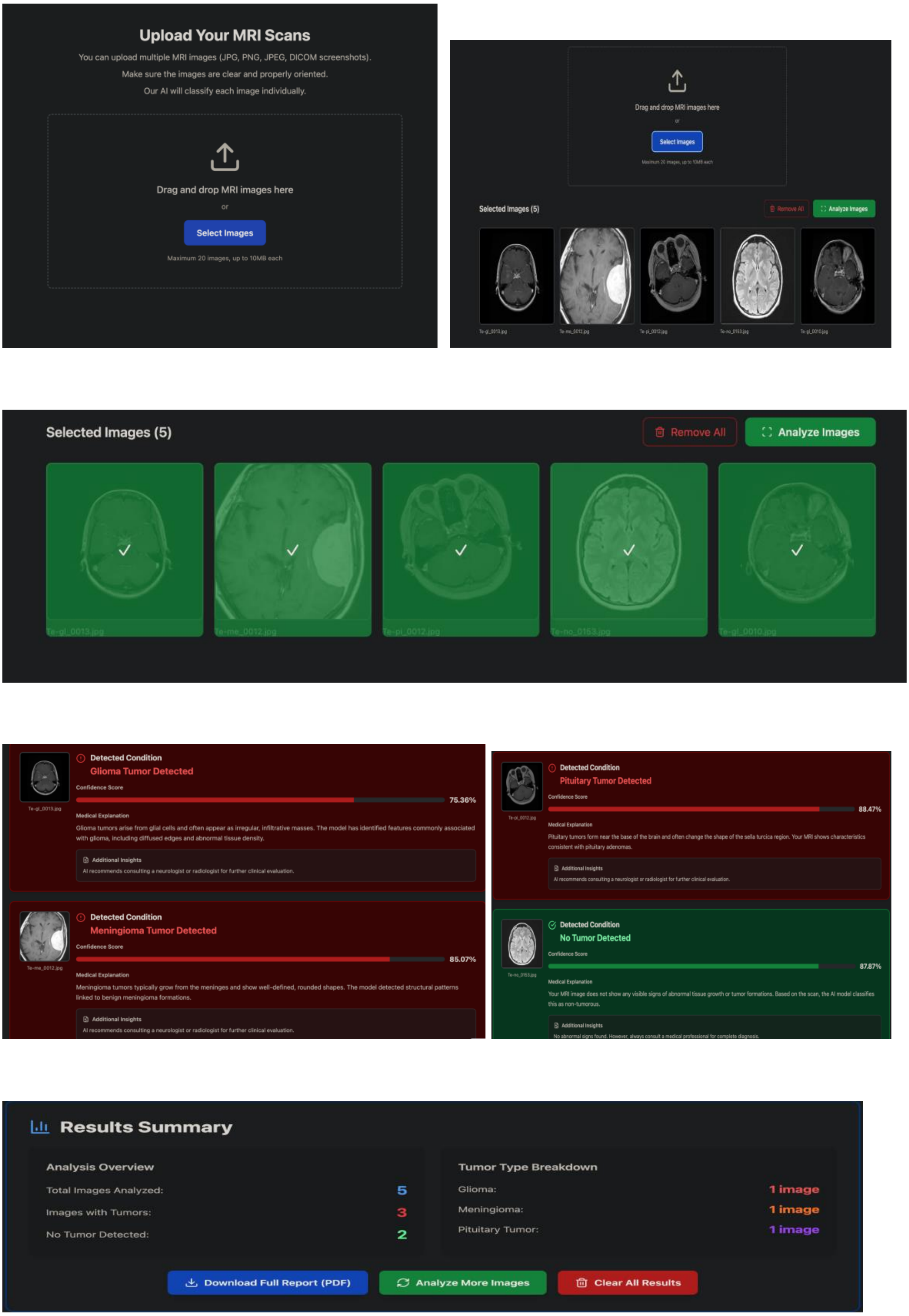


Figure 4.4.1 Brain tumor detecting and analyze report

4.5 Discussion

Table 4.3 All models comparison summary

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
CNN	99.58 %	96.00 %	93.00 %	99.00 %
VGG16	95.73 %	89.00 %	91.00 %	92.00 %

The analysis of the CNN and VGG16 models shows that there are some differences in the performance of the models in several vital variables such as accuracy, precision, recall, and F1 score. The CNN model is the best-performing model, which has the accuracy of 99.58, precision of 96.00, recall of 93.00, and the F1 of 99.00. Such impressive numbers indicate that CNN does not only detect the tumors correctly but also the lowest number of false positives and false negatives and this makes it the most consistent model in identifying brain tumors automatically. VGG16, although being effective, demonstrates somewhat lower scores, which are 95.73% accuracy, 89.00 precision, 91.00 recall and 92.00 F1. VGG16 is more likely to agree with the non-tumor images than the tumors and this is evidenced by its lower precision than CNN. This brings out the fact that the CNN gives a more balanced and consistent performance with higher accuracy, precision and hence the better choice in practical applications where a high level of diagnostic reliability is needed. Conversely, VGG16 is a good candidate but there is a slight weakness in terms of reducing false positives, thus it is not the best option compared to CNN when it comes to life-and-death medical applications.

On the whole, CNN model is superior in comparison to VGG16 as it is more effective to identify tumor and non-tumor cases with fewer errors. VGG16 model is still an option which can be used, however, its low precision and accuracy imply that it would be better applied to the tasks in which absolute performance is not as important. This analogy emphasizes the significance of the most precise model when performing medical imaging tasks when precision and reliability are the most important factors.

In summary, the order of performance is:

CNN > VGG16

CHAPTER 5

CONCLUSION

5.1 Findings & Contributions

The study introduces a comparison of two deep learning algorithms of CNN and VGG16 to detect brain tumors using MRI scans. The findings show that CNN is a better model than VGG16 in terms of accuracy, precision, and recall, thus it is a more reliable model in this task. With an accuracy of 99.58, precision of 96.00, recall of 93.00 and F1-score of 99.00, CNN is quite effective in the detection of tumors. The opposite is true because VGG16 had an accuracy of 95.73, a precision of 89.00, a recall of 91.00 and F1-score of 92.00, which means that it is not as accurate as it can be, especially when it comes to precision. CNN was especially strong in the domain of gliomas detection, as it proved to be much more effective than VGG16, which underlines the significance of choosing a model in healthcare.

One of the most significant contributions of the given work is the direct comparison of CNN and VGG16 in their capabilities to identify the types of brain tumors in MRI images. The study demonstrates that CNN model is more applicable in real life implementation, especially in medical imaging because it has better accuracy and precision. The given discovery is crucial because it indicates that the CNN model may be potentially used to yield more accurate diagnoses, particularly in the most strict spheres where the presence of a tumor may be challenging, e.g., with gliomas.

The other significant contribution is the model performance analysis in the various tumor types. CNN was found to perform outstandingly well in differentiating the types of tumor, especially the glioma which is widely known as hard to detect. This implies that CNN might play an important role in real-life clinical environments enhancing the reliability of the diagnosis and minimizing the possibility of misdiagnosis.

The study also brings forward the feasible benefits of CNN instead of VGG16 in the sense that it is more adept at managing the false positive and its overall strength in the classification tasks. These results imply that CNN will provide a more consistent resolution in regards to automated tumor creation that is essential in emergency healthcare settings.

Lastly, the piece has a contribution to the AI research domain and the medical community. To researchers of AI, it offers a positive contribution to the relative effectiveness of various deep learning models in medical imaging tasks. To clinicians, it shows how AI-based solutions can be used to minimize diagnostic errors and enhance the rate of tumor detection. This study opens the way towards incorporating CNN in clinical decision-making techniques which will eventually lead to better patient care and treatment outcomes by demonstrating the strengths of CNN in brain tumor classification.

5.2 Recommendations for future work

The way forward should be the exploration of the CNN model on bigger and more varied datasets to determine its capability to expand across various hospitals, MRI scanners, and patient demographics. This will aid in evaluating the strength of this model to accommodate changes in image protocols and patient status.

Clinical tests ought to be carried out to determine the effectiveness of the model in real life diagnostic stages. As part of the normal clinical practice, radiologists are expected to test the predictions of the model and hence give meaningful feedback that will be used to improve the model and increase its clinical acceptance.
More research should be done to enhance explainability of the CNN model. It is noteworthy that the medical professionals should be aware of the logic behind the model projections. The use of model interpretability methodologies, e.g., Grad-CAM or SHAP, would contribute to more trust and confidence of healthcare providers in AI-driven decision-making.

Future studies are required on how to develop lightweight variants of the CNN model. Such versions might be optimized in terms of computation resource requirements in hospitals with limited computational resources so that the model can also be accurate and computationally efficient. The challenge of balancing accuracy and speed to achieve real-time diagnosis will be one of the key areas of concern.

Finally, more recent model architectures, e.g. transformer based models or hybrid models, ought to be contrasted with the CNN to detect brain tumors. Also, the combination of multi-modal imaging, e.g., MRI with CT or PET scans, may prove to be a more diagnostic tool and an additional benefit to the performance of the model in the clinical environment.

REFERENCES

1. MRI Brain tumor image classification with CNN characteristics and machine learning classifiers- Liu, Y, Wang, Z, Xue, Y, Cheng, N, Shen, B, Hou, L, and Jin, L (2024). Reports very high multi-class classification (glioma, meningioma, pituitary, no-tumor) accuracies using CNN to extract features based on MRI and classical ML classifiers to classify tumors. SpringerLink

2. Development of deep learning in brain tumor classification of MRI images: Fine-tuning VGG16 and Xception-based approach (2025). A deep CNN + fine-tuned VGG16 (and Xception) classifies tumors in MRI; is able to classify a wide variety of tumors. ScienceDirect

3. Deep learning based on MRI images helps classify brain tumors - Wong, Y., Su E. L. M., Yeong C. F., Holderbaum W., and Yang C. (2025). Vegas et al. (2019), constitute themselves, rely on a CNN with pretrained VGG16 backbone on a large and heterogeneous dataset of MRI (glioma, meningioma, pituitary, normal) with high classification accuracies. PLOS

4. Detection of brain-based MRI tumor with convolutional deep learning and machine-learning classifiers- Saeedi, S., Rezayi, S., Keshavarz, H., and Niakan Kalhori, S. R. (2023). Created a 2D-CNN of their own and a convolutional autoencoder to classify glioma, meningioma, pituitary tumors and healthy brains. Shows CNN to be effective in MRI-based classification. BioMed Central

5. (2024). Multi-Image MRI classification of brain tumors with a deep residual and region-based CNN (Res-BRNet). Introduces a residual + region-aware CNN, and tests on several types of tumors (meningioma, glioma, pituitary, healthy), with high accuracy and F1-score. MDPI

6. Brain tumor classification tracing with the hybrid CNN architecture optimality--Ramakrishnan, A. B., Sridevi, M., Vasudevan, S. K., Manikandan, R., and Gandomi, A. H. (2023). Trains a hybrid CNN network that fuses VGG16, DenseNet, ResNet, Inception, amongst others, to classify MRI tumors, and can be useful in comparing network architecture variants on identical data. ScienceDirect

7. 2024: A multi-task CNN brain MRI classification (tumor detection, type classification, location identification) model. A single CNN architecture is used to achieve several tasks tumor detection, tumor type classification (glioma, meningioma, pituitary, no tumor), and tumor localization - better clinical relevance. Frontiers

8. An effective brain tumor detection and classification model based on pre-trained convolutional neural network models- Rao, K. N., Khalaf, O. I., Krishnasree, V., et al. (2024). Detects and classifies brain tumors through pre-trained CNN models (transfer learning) on MRI data and overcomes resource limitations as well as data limitations. ScienceDirect

9. Deep learning methods to detect and classify brain tumor using MRI images (2020 to 2024): A systematic review Bouhafra, S., and El Bahi, H. (2025). A literature review of recent deep-learning based approaches (CNNs, autoencoders, transformers, hybrid approaches) to classify brain tumor MRI images - handy background and related works. SpringerLink

10. Image Segmentation CNN Based Image Segmentation Approach in Brain tumor classification: Method Review - (2025). Reviews CNN architectures (U-Net, ResNet, etc.) to segmentation and classification problems in brain tumor MRI, both in terms of segmentation performance and model constraints - can be applied to methodological discussion. MDPI

11. Survey deep learning; survey MRI modalities Brain tumor detection and classification: (2023). Gives background on recent deep-learning based methods of MRI based brain tumor classification; can be used in locating your work within the larger trends in research. ScienceDirect

12. Self-Supervised Learning Model Self-Supervised learning using Residual Vision Transformer (ResViT) to classify Brain tumors Karagoz, M. A., Nalbantoglu, O. U., and Fox, G. C. (2024). Introduces a hybrid CNN-Transformer architecture in self-supervised MRI-based brain tumor classification; a possible future research on CNN/VGG alternatives. arXiv.

13. Talukder, M. A., Islam, M., and Uddin, M. A., (2023). An Optimized Ensemble Deep Learning Model to Classify Brain tumors. Presents an ensemble deep learning (using more than one architecture) on tumor data in MRI; an ensemble method can be used to compare and discuss the effect of ensembles versus single models. arXiv

14. Comparison of CNN Architectures that are resource efficient in brain tumor classification Khan, M. A., and Auvee, R. B. Z. (2024). The study compares lightweight/custom CNN models with more deeply trained models (e.g. VGG16, ResNet) on MRI tumor classification - is applicable when resources are limited. arXiv.

15. Better classification of brain tumors with CNN and transfer-learning with diverse MRI images - (2025). Tests CNN + transfer-learning models on a broad range of

MRI data, pointing out problems with generalization and robustness - helpful in the context of discussing the limitations of datasets. SpringerLink+1

16. Challenges and solutions to Hybrid and transfer-learning based MRI brain tumor classification - (2023). Survey CNN-based brain tumor detection/classification hybrid architectures, which have issues with generalization, data imbalance, and interpretability (good to a discussion and future research). ScienceDirect+1

17. Deep learning on brain tumors: segmentation and classification of publicly available MRI data - (2022). Presents CNN based segmentation and classification on publicly available MRI data used to detect brain tumors - gives baseline comparisons and methodological information. BioMed Central+1

18. Optimization of different CNN models to classify brain tumor using MRI images - (2023). Comparisons between various CNN models (VGG and ResNet as well as a custom model) on tasks with MRI tumor classification tasks, providing performance measurements that may be used in the evaluation of your own models. ScienceDirect+1

19. Clinical viability CNN-based brain tumor detection: heterogeneous MRI assessment of MRI data - (2024). Concentrates on CNN performance in the context of training and testing on heterogeneous MRI (using more than one source / scanner) with the emphasis on generalization - a key concern in the creation of robust apps. MDPI+1

20. Deep-learning enabled MRI brain tumor classification: recent trends and future perspectives -(2025). Gives a general overview of recent developments, data, problems (data variability, model explainability), and future research opportunities in MRI based brain tumor classification. SpringerLink+1

Appendix A: Dataset Availability

DatasetLink: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset