
**Arab American University of Palestine**

**Faculty of Graduate Studies**

# Diagnosis and Classification of Brain Tumors Based on Deep Learning Techniques

Prepared By

**Oroob Abdullah Farhan Yassin**

Supervisor By

**Prof. Dr. Mohammed Awad**

**This Thesis Was Submitted in Partial Fulfillment of the Requirements for the Master's Degree in Computer Science**

**January / 2023**

I

# Thesis Approval

## Diagnosis and Classification of Brain Tumors Based on Deep Learning Techniques

By

**Oroob Abdullah Farhan Yassin**

This Thesis was Defended Successfully on **25/01/2023** and Approved By:

| Committee Members: | Signature |
|---|---|
| **1.** Supervisor: **Prof. Mohammed Awad** |  |
| **2.** Internal Examiner: **Dr. Yousef Daraghmi** |  |
| **3.** External Examiner: **Dr. Ahmad Ewais** |  |

# Declaration

I declare that the work presented in this thesis is based on my own work and has not been submitted for any degree in any University, institution, or other colleges of higher education than Arab American University – Palestine (AAUP) for academic credit.

Name: Oroob Abdullah Yassin

Signature  ……..

Date: 24/5/2023

# Dedication

“I say there’s not more happiness for me than the freedom of my homeland.”

To the land of martyrs and heroes, to my homeland and its people, to Palestine, and the steadfast Palestinian people, I dedicate this thesis.

To those who are not matched by anyone in the universe, to whom God has commanded us to honor them, to those who have made a great deal, and have given what cannot be returned, to you these words, my dear mother and father, I dedicate this research to you; You have been my best supporter throughout my academic career.

I dedicate my thesis to my siblings who have always been there for me even in my tough days. I’m so thankful for having you by my side. I also dedicate this thesis to my supervisor, Prof. Dr. Mohammed Awad, thanking him for all the support and assistance

For all of my friends and colleagues, I appreciate all of you and your support and encouragement for me during my years of study. All that you do for me never goes unnoticed.

# Acknowledgments

I would also like to seize the opportunity and express my deep gratitude to Prof. Mohammed Awad for his support, great pieces of advice, and valuable suggestions. It all helped me overcome the obstacles I faced while working on my thesis.

# Diagnosis and Classification of Brain Tumors Based on Deep Learning Techniques

**By: Oroob Abdullah Farhan Yassin**

**Supervisor By: Prof. Dr. Mohammed Awad**

## Abstract

Brain tumor diagnosis through imaging tests is a crucial aspect of medical treatment planning. The most common imaging test for diagnosing brain tumors is Magnetic Resonance Imaging (MRI). Despite the expertise of medical practitioners, accurately classifying the type of tumor can be challenging due to the complexity of medical images. This is why the study of using deep learning algorithms to classify brain tumors is crucial.

In this study, a dataset of MRI images of patients diagnosed with brain tumors was used to train deep learning models such as Convolutional Neural Networks (CNN), Visual Geometry Group-16 Model (VGG16), and Visual Geometry Group-19 Model (VGG19). The results of the first phase showed an accuracy of 95.40% for the CNN model and 90-93% for the other deep-learning models. To improve the accuracy, the second phase applied transfer learning to the VGG16 model, resulting in a 96.27% accuracy improvement. In the third phase, data augmentation techniques were used to balance the data and prevent over fitting, resulting in an accuracy improvement of 98.45%.

In conclusion, this study highlights the potential of deep learning algorithms in accurately classifying brain tumors and supporting medical practitioners in determining the best treatment plan for patients. The results demonstrate the significance of AI in the medical field and the potential for further advancements in the future.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviation

| | |
|---|---|
| AI | Artificial Inelegance |
| AUC | Area Under Curve |
| CNN | Convolutional Neural Networks |
| Conv | Convolution |
| ConvNet | Convolution Network |
| DenseNet121 | Densely Connected Convolutional Networks |
| DL | Deep Learning |
| FC | Fully Connected |
| FN | False-Negative |
| FP | False-Positive |
| ML | Machine Learning |
| MRI | Magnetic resonance imaging |
| PC | Personal Computer |
| ReLU | Rectified Linear Unit |
| ResNet50 | Residual learning Network |
| ROC | Receiver Operating Characteristic Curve |
| TN | True Negative |
| TNR | True Negative Rate |
| TP | True Positive |
| TPR | True Positive Rate |
| VGG-16 | Visual Geometry Group-16 Model |
| VGG-19 | Visual Geometry Group-19 Model |

# Chapter One

# Introduction

## 1.1 Introduction

Brain tumors are one of the most prevalent forms of cancer that can kill a person, and the number of people infected with this tumor grows every year. Therefore, the importance of early detection of brain cancer is to preserve human life. Early detection techniques must be capable of detecting a tumor or alerting to the existence of a tumor in the brain, as well as determining the kind of tumor, which is a large part of determining the correct stage of treatment. Recently, treatments that fight cancer, especially those in the early stages of tumor detection, have been developed through receiving treatment in early stages to preserve the patient's life and increase the likelihood of survival, compared to patients diagnosed in the late stages (DeAngelis, 2001). The brain is an essential organ in the human body since it organizes all of the body's operations. Brain tumors pose a threat to human life due to the fact that these tumors are a mass resulting from the abnormal growth of cells that grow in size and accumulate as a mass, which constitutes either cancerous or non-cancerous (benign) masses (Tom et al., 1998).

This increased growth of solid cells inside the skull affects the brain cells in terms of increasing pressure on them, which leads to serious problems that lead to damage that affect the patient's life (Jain & Mishra, 2013). According to the tumor's genesis, tumors are divided into main and secondary categories. The brain is where primary cancers start to grow, which can be benign tumors (non-cancerous cells) or malignant tumors (cancerous cells), while secondary (metastatic) tumors arise in one of the body’s organs

and spread to the brain through the blood or lymph (Thompson et al., 1998). Tumors are divided according to the degree of brain tissue abnormality into four levels, each level showing the rate of growth of cancer cells (Glass-Macenka et al., 2013). Grades 1 and 2 represent lower grades of the tumor and are less severe (least malignant), and grades 3 and 4 (more malignant) constitute high-risk tumors (Louis et al., 2007). The methods of treating brain tumors are surgery, radiotherapy, or chemotherapy, depending on the type, grade, and size of the tumor (Al-Tamimi & Sulong, 2014).

Artificial intelligence (AI) refers to robots' capacity to mimic human behavior. This notion was a watershed moment in human history in many industries, businesses, education, health, and other spheres (Battineni et al., 2020); (Topol, 2019). Machine learning is described as a computer's capacity to learn from experience, that is, to alter processing depending on the newly obtained information. This procedure can be built on a simple decision-making tree. While the concept of deep learning that simulates the human brain through neural networks appears in the processing of data and the creation of patterns for use in the decision-making process (LeCun et al., 2015). As a result of the above, machine learning and deep learning have received great interest in the medical field, especially in dealing with medical images (Yuan & Wang, 2018); (Hu et al., 2016). By applying computer programs to some algorithms, these programs can gain more experience in a short period compared to humans gaining experiences in their lives. For example, a specialist needs approximately 40 years to research approximately 225,000 MRI / CT scans, in contrast, artificial intelligence accomplishes this task within a short period of up to millions of scans, and the result is an improvement in the accuracy of Diagnosis. It is necessary that the speed of reading magnetic images or CT

images and their diagnosis using AI be faster than reading the average human and also have a higher accuracy (Yokota et al., 2017).

From a clinical standpoint, the process of identifying a brain tumor entails the correct identification of tumor tissue, to its determination from magnetic resonance imaging using well-established clinical information and diagnostic criteria. This process should lead to the treatment of the tumor in the correct manner and at the appropriate time. To achieve this process, there must be clinical knowledge and a database that represents the information through which the decision-making and diagnosis process can be made. The difficulty appears in the manual tumor diagnosis process due to the diversity of tumor shapes and types, and the diagnosis process can be less accurate (Gordillo et al., 2013).

There is no doubt about the doctor’s experience in diagnosing the tumor, but sometimes there are some mistakes in classifying the type of tumor that negatively affect the patient’s health due to the difficulty of distinguishing between these tumors through medical images that show the characteristics of the tumors. The characteristics of tumors allow for determining the type of tumor, whether it is tumor benign or malignant, and even how it will progress. As a result, knowing these traits is critical in deciding the medical team's treatment approach. And here is the purpose of providing a set of medical images from the magnetic resonance images of the previous patients diagnosed with a brain tumor and benefiting from it in building a system using automated deep learning techniques (DL) such as (CNN, VGG16, or VGG19, ... etc.) (Litjens et al., 2017); (Pan et al., 2015); (Ravì et al., 2016). These techniques contribute significantly and accurately to help the doctor diagnose the type of tumor with high

accuracy. This study will concentrate on these methods for precisely diagnosing tumors using medical images.

### 1.1.1 Types of Brain Tumors

According to the World Health Organization (WHO), there are many types of brain and central nervous system tumors, with more than 120 types (Glass-Macenka et al., 2015). Where the most common types are (Glass-Macenka et al., 2015):

1. Glioma: classified as a primary tumor. A tumor is formed from the abnormal growth of glial cells located around nerve cells in the brain and spread to surrounding tissues (Menze et al., 2015). This type of tumor presents as an area of heterogeneous texture, low signal intensity, and bright tumor borders, and is widely diagnosed in adults (Bauer et al., 2013).
2. Meningioma: classified as a benign tumor that arises from the covering of the brain, as it is located under the skull, and the tumor usually grows slowly within the brain (Glass-Macenka et al., 2015). A meningioma appears as extra-axial masses with a homogeneous region, and increased signal intensity. That is, it appears brighter than the surrounding tissues (Ismael, 2018).
3. Pituitary tumor: The responsibility of the pituitary gland lies in controlling the production of hormones that control other glands in the body such as the ovaries, adrenal glands, thyroid gland, and others. A pituitary tumor is a benign tumor that targets the pituitary gland, causing other glands' functions to fail. Tumor growth can be reduced through a series of treatments that prevent it from spreading to other brain tissues (Glass-Macenka et al., 2015). A pituitary tumor

is a large mass of glands on the side of a micro adenoma, presenting with a heterogeneous brown spot with low to high density (Weerakkody and Gaillard).

### 1.1.2 Tumor Diagnosis Techniques

Brain tumors are diagnosed using different types of scans; the most famous methods are magnetic resonance imaging, computed tomography, angiography, and others. However, magnetic resonance imaging is one of the best and most popular techniques currently in the process of diagnosing a brain tumor, and it produces more detailed images that help doctors in the diagnosis process by looking at the images without any surgical intervention (Gopal & Karnan, 2011); (Chauhan & Sharma, 2014). It also provides preliminary information about the tumor and its location (Cancer.Net, 2021).

The diagnosis process at the beginning of the disease is the key stone of the patient treatment process at the beginning of the disease to reduce the impact and symptoms of the disease on the patient's health. As a result, it is critical to pay close attention to the process of early diagnosis of any disease, especially if the condition is a brain tumor. The procedure of detecting and diagnosing the type of brain tumor boosts the patient's chances of recovery by 90%. Due to the location of the complicated tumor, its shape, and its kind, early discovery of the tumor necessitates actual experience in the patient assessment process, which may be expensive both financially and physically due to the likelihood of a flawed evaluation procedure that might kill the patient. Because of the disease's sensitivity, modern procedures must be used in the process of diagnosing brain tumors. Using magnetic resonance pictures on a computer develops a program that is responsible for the process of diagnosing brain cancers. This will contribute to assisting the human expert in determining the correct treatment stage (Amsaveni et al., 2013).

Doctors gain experience and knowledge from their careers in caring for and treating patients. This is the concept that underpins artificial intelligence, the more data, and experience, the better the ability to make knowledge-based decisions. In the current era, large amounts of data have emerged which are a treasure trove of gaining experience and the mainstay of artificial intelligence (Turing, 1950).

### 1.1.3 Artificial Intelligence in Healthcare

Artificial intelligence (AI) refers to robots' capacity to mimic human behavior. This notion was a watershed moment in human history in many industries, businesses, education, health, and other spheres (Battineni et al., 2020); (Topol, 2019). Machine learning is described as a computer's capacity to learn from experience, that is, to alter processing depending on the newly obtained information. This procedure can be built on a simple decision-making tree. While the concept of deep learning that simulates the human brain through neural networks appears in the processing of data and the creation of patterns for use in the decision-making process (LeCun et al., 2015). As a result of the above, machine learning and deep learning have received great interest in the medical field, especially in dealing with medical images (Yuan & Wang, 2018); (Hu et al., 2016). By applying computer programs to some algorithms, these programs can gain more experience in a short period compared to humans gaining experiences in their lives. For example, a specialist needs approximately 40 years to research approximately 225,000 MRI / CT scans, in contrast, artificial intelligence accomplishes this task within a short period of up to millions of scans, and the result is an improvement in the accuracy of Diagnosis. It is necessary that the speed of reading magnetic images or CT images and their diagnosis using AI be faster than reading the average human and also have a higher accuracy (Yokota et al., 2017).

### 1.1.4 Tumor Diagnosis Problems

From a clinical standpoint, the process of identifying a brain tumor entails the correct identification of tumor tissue, to its determination from magnetic resonance imaging using well-established clinical information and diagnostic criteria. This process should lead to the treatment of the tumor in the correct manner and at the appropriate time. To achieve this process, there must be clinical knowledge and a database that represents the information through which the decision-making and diagnosis process can be made. The difficulty appears in the manual tumor diagnosis process due to the diversity of tumor shapes and types, and the diagnosis process can be less accurate (Gordillo etal., 2013).

There is no doubt about the doctor's experience in diagnosing the tumor, but sometimes there are some mistakes in classifying the type of tumor that negatively affect the patient's health due to the difficulty of distinguishing between these tumors through medical images that show the characteristics of the tumors. The characteristics of tumors allow for determining the type of tumor, whether it is tumor benign or malignant, and even how it will progress. As a result, knowing these traits is critical in deciding the medical team's treatment approach.

### 1.1.5 Brain Tumor Diagnosis through Deep Learning Techniques

Brain tumors are a significant health concern and early detection is critical to successful treatment. The traditional methods of diagnosing brain tumors have limitations, and this study aims to explore the potential of deep learning in improving the accuracy of diagnosis. This thesis will investigate the use of automated deep learning techniques, such as Convolutional Neural Networks (CNNs), VGG16, VGG19, Xception, ResNet50, and DenseNet121(Litjens et al., 2017); (Pan et al., 2015); (Ravì et al., 2016),

in the diagnosis of brain tumors using magnetic resonance imaging (MRI) scans of previous patients. The study will focus on evaluating the performance of these algorithms in accurately identifying the type of brain tumor present. The results of this research will contribute to the field of medical diagnosis and have the potential to improve the lives of patients with brain tumors. The thesis will provide a comprehensive analysis of the use of deep learning in brain tumor diagnosis, highlighting its strengths and limitations.

## 1.2 Objectives

The primary goal is to increase the accuracy of the categorization process for the three most prevalent brain cancers in persons with one of these tumors: glioma, meningioma, and pituitary tumor. The classification process uses artificial intelligence techniques, and processes are applied to extensive global data collection. These techniques are applied in analyzing the data set, automatically selecting features, and applying them to the classification process to find a more accurate method in the tumor diagnosis process using deep learning. The study's primary goal is to develop a more accurate method of diagnosis. Reducing the time of the diagnosis process and increasing accuracy using artificial intelligence contributes significantly to the treatment plan. It supports the decision of doctors and radiologists in the diagnosis process. This study contributes to achieving this through artificial intelligence techniques.

**Other Specific Objectives of the Study are the Following:**

- Several DL techniques were used to classify brain cancers using an international collection of MRI medical Images. To find the best technique with the highest classification accuracy.

- The VGG16 model will be used to extract features from the dataset and increase the accuracy of brain tumor classification by implementing the transfer learning approach.
- We are applying the concept of data augmentation to improve the model training process to avoid over fitting and balance the data to improve accuracy.

## 1.3 Contribution

Research has proposed methods for identifying patients' most prevalent brain tumors using machine learning and deep learning approaches. Deep learning has proven successful and capable of solving problems that include pattern recognition or shape detection.

This study utilizes magnetic resonance imaging (MRI) and deep learning algorithms to detect brain tumors. Transitional learning has been shown to initially reduce the hectic work of training models. As a result, transfer learning is employed with the VGG16 trained model, which was previously trained using the Image Net database containing over 1000 different target groups. As part of the tumor diagnosis process, a data set that consists of a set of magnetic resonance images is pre-processed to make it suitable for the tumor diagnosis process. The notion of data augmentation, on the other hand, is used to improve the model training process to minimize over fitting and balance the data to increase accuracy.

## 1.4 Overview

In this section of the chapter introduction, a summary will be presented for the rest of the upcoming chapters that will be discussed in this thesis, which seeks to achieve the goals and objectives of the work. The structure of the thesis is as follows: The second

chapter gives context for the worldwide data set that was used to carry out the research, which consists of a collection of magnetic resonance images of brain tumors. The literature on the research topic is then examined, as well as several other techniques related to diagnosing brain tumors. The third chapter will shed light on the methodologies used in this study, from the beginning of data collection and data handling through pre-processing, which is the mechanism of improving the images of the data set. In addition to covering deep learning models for classifying pre-trained images, they are represented models VGG16, VGG19, Xception, ResNet50, and DenseNet121, in addition to building a CNN network, in addition to building a CNN, as well as the VGG-16 model with a learning transfer mechanism. Finally, different performance metrics have been introduced to evaluate deep learning models. The fourth chapter will discuss some experiments and the results of each data set with different models. A comparison will be provided between the models that show the best model for each data set, as well as a comparison between the results of global data sets. Chapter 5 offers a conclusion and suggestions for further development.

# Chapter Two

# Background

## 2.1 Background

The rapid development in information and engineering has contributed to the development of healthcare technology and to providing quick solutions to recover from healthcare problems for patients. 12,764 people are affected by brain tumors, according to reports from the National Cancer Institute and the World Health Organization (Alqazzaz et al., 2019). Since modern healthcare technologies are used in early detection of the examination and detection of brain tumors (Kong et al., 2014). Recently, much research has focused on studying the automatic detection of tumors for different types of tumors and encouraging researchers to work on increasing the efficiency of the tumor diagnosis process (Dahab et al., 2012). One of the challenging issues facing the medical personnel in trying to diagnose the tumor accurately is brain tumor diagnosis.

In light of artificial intelligence (AI) methods, automated methods have improved the accuracy and reliability of diagnosis and helped doctors and radiologists analyze and detect medical images effectively (Hao et al., 2016). With the help of artificial intelligence, machine learning and deep learning contribute to diagnosing medical images of the brain and brain tumors (Abd-Ellah et al., 2018). Deep learning has proven its success in diagnosing tumors with accuracy comparable to human performance (Teare et al., 2017; Bejnordi et al., 2017). Deep learning has emerged as a field of research. It has a vital role in tumor detection and classification, where deep architecture displays complex interactions without the need for a vast number of nodes effectively,

unlike SVM and KNN architectures (Aslam et al., 2015). According to the results of several studies (Soltaninejad et al., 2017); (Kaur & Gandhi, 2020); (Pravitasari et al., 2020), these studies showed the success of the ability of deep learning to predict and diagnose diseases according to the factors that play a role in diagnosing the disease.

Magnetic resonance imaging (MRI), as well as ultrasound and computed tomography, are technologies utilized in the identification and study of brain malignancies. Magnetic imaging is primarily used in brain tumor analysis, which provides a three-dimensional evaluation in “axial, coronal, and sagittal” directions (Sazzad et al., 2019). MRI scans the brain region, which in turn scans healthy and unhealthy areas. MRI images of the brain are categorized into weighted T1 and T2 (empty fluid area in the brain image) based on the assessment of contrast in image areas (Gurunathan & Krishnan, 2021). Magnetic resonance imaging is utilized in the diagnostic process and is the most common in diagnosing tumors, where the diagnosis process is manual. However, the manual method’s drawbacks reduce its use in clinical applications. The disadvantages are in analyzing and reading MRI images, which carry a quantity of data that makes visual interpretation expensive, inaccurate, and expensive.

Many studies in the field of information technology have suggested that many algorithms, machine learning, and deep learning techniques are used in diagnosing brain tumors (Gurunathan & Krishnan, 2021); (Nayak et al., 2022); (Gupta et al., 2022); (Demir, 2022); (Ranjbarzadeh et al., 2021); (Deb & Roy, 2021).

In this study, to identify the most critical features (tumor characteristics) used in the process of classifying brain tumors, feature extraction techniques are used to do this process, this utilized deep learning (DL) techniques, such as CNN, VGG16, VGG19,

Xception, ResNet50, DenseNet121. These techniques contribute significantly and accurately to help the doctor diagnose the type of tumor with high accuracy. This work will concentrate on these strategies for effectively diagnosing cancers using medical imaging. In addition to applying the concept of learning transfer on the VGG16 model to work on classifying tumors. In addition to using a data augmentation strategy and a technique to enhance the volume of data used to train the model to generate credible predictions. Deep learning methods and deep learning models need a large amount of training data, which is not always accessible. As a result, we enhance our previous data with this strategy to develop a more generalized model.

## 2.2 Dataset Description

The brain tumor dataset used in this study consists of 3064 MRI slices from 233 patients, (Cheng et al., 2015); (Cheng et al., 2016). The dataset was obtained from figshar, a website with an open-access dataset for scientific research purposes. It consists of three types of brain tumors: pituitary tumor, glioma, and meningioma, in addition to three different views: sagittal, axial, and coronal view. The dataset is in ".mat" format and includes patient ID, 512x512 image data in uint16 format, tumor type tag, vector of tumor boundaries with coordinate points, and binary mask image. Given the nature of the convolutional neural network that requires image data, the 512x512 pixel images will be resized to 224x224 based on prior research for utilization in the study (Cheng, 2017).

The dataset consists of meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices) for training deep learning models. Figure (2.2-1) shows the

distribution of images in the dataset between images of tumor types, and this was implemented using Python code.

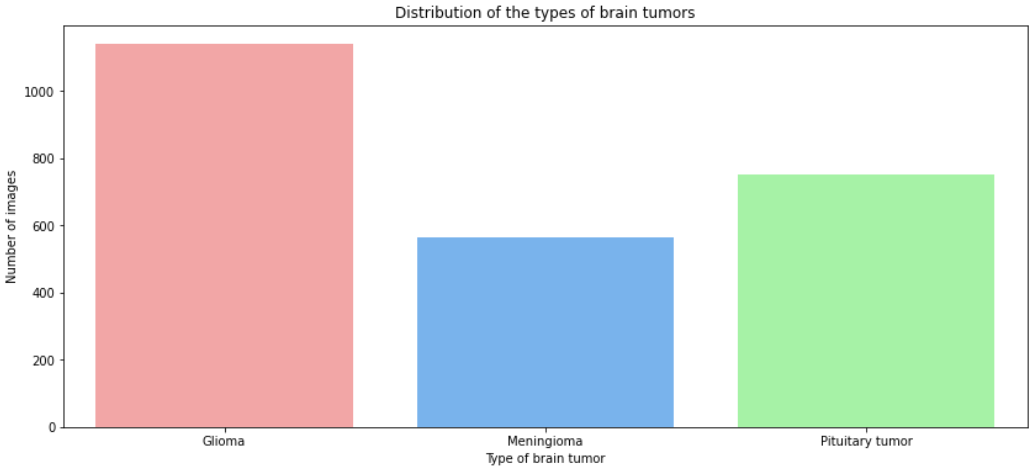



Figure 2.2-1: Distribution of Tumor Types

Figure (2.2-2) shows a sample from the training set of medical MRI images (Cheng, 2017).

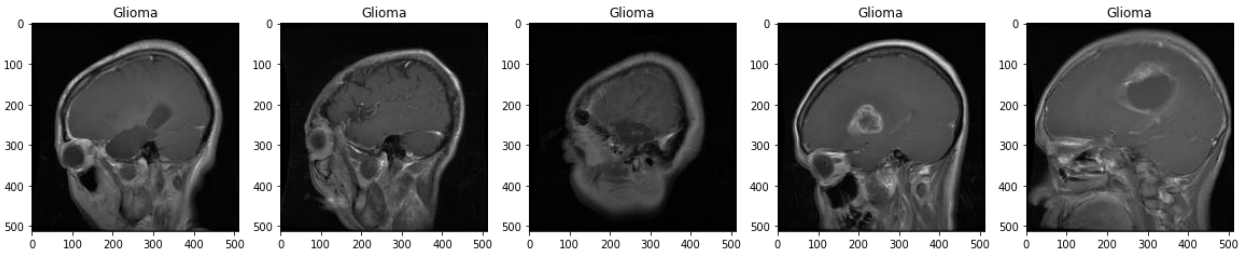



Figure 2.2-2: Sample brain tumor dataset diagnosed with glioma

Figure (2.2-3) shows a sample from the training set of medical MRI images diagnosed with a meningioma.

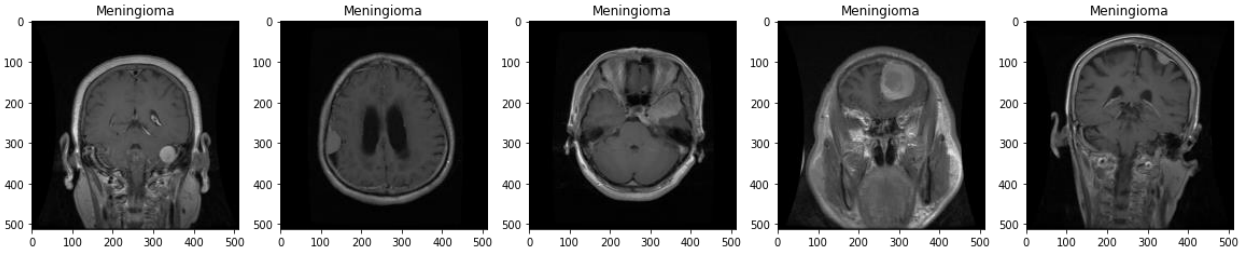



Figure 2.2-3: Sample brain tumor dataset with a meningioma

Figure (2.2-4) shows a sample from the training set of medical MRI images diagnosed with a pituitary tumor.

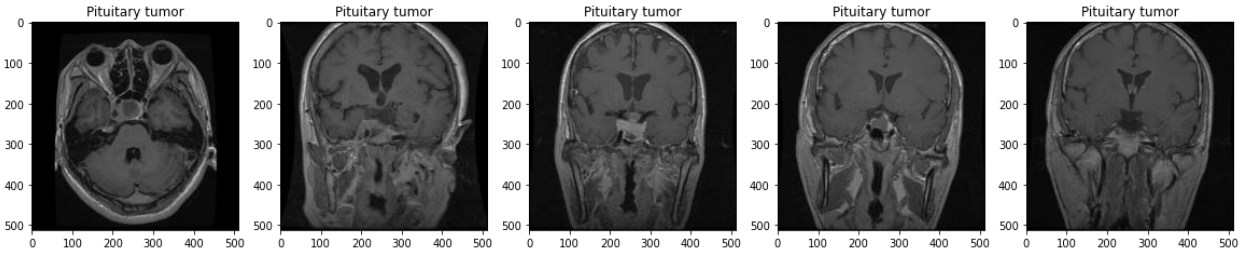



Figure 2.2-4: Sample brain tumor dataset with a pituitary tumor

## 2.3 Related Work

The classification and diagnosis of brain tumors greatly assist the medical field, especially doctors, in making appropriate decisions in patients' treatment plans and follow-up treatment. Recently in the literature, machine learning (ML) and deep learning (DL) have demonstrated their superiority in the accuracy and detection of disease diagnosis processes and preparation of treatment plans, which is the main objective of automated and semi-automated assessment of patients (Anaraki et al., 2019). Previous studies focused on applying deep learning techniques to the data set, which was global in some studies. Some studies focused on the researcher's community data set at the local level. Brain tumor diagnosis studies have also been applied to deep

learning and machine learning techniques as a hybrid model in the classification of brain tumors. Some studies also compared the types of deep learning techniques to prove the accuracy of a particular technique on a global data set. This review will present the relevant studies from the most recent to the oldest.

### 2.3.1 Deep Learning in Related Work

In this section, we delve into the body of research that has utilized deep learning as a tool for diagnosing brain tumors. We examine the innovative approaches and techniques employed in this field and assess the impact they have had on improving the accuracy and efficiency of tumor diagnosis.

Kesav and Jibukumar (2022) proposed a structure using RCNN technology for brain tumor classification and tumor-type object detection. They applied it to a dataset of magnetic imaging with dimensions of 512*512 pixels. It was processed by changing its size to become 227*227 pixels in size, and it was repeated three times to get three input channels. 80% were allocated for training and 20% for testing. The goal of the proposal is to reduce the implementation time of the traditional RCNN architecture by using a low-complex framework and to propose a system for brain tumor analysis. Initially, a low-complexity structure was used to achieve the goal of the proposal, a dual-channel CNN to classify MRI samples between glioma and MRI samples for tumors, with success with an accuracy of 98.21 percent. The proposed structure was also used as a feature extractor for RCNN to detect the tumor areas of the Glioma MRI sample that were classified from the previous stage. The tumor area is restricted using bounding boxes. This method has been applied to include two other types of meningioma and

pituitary adenoma. The methodology achieved an average confidence level of 98.83% in achieving a low implementation time compared to other current structures.

Fayyadh and Ibrahim (2020) proposed a convolutional neural network to classify brain tumours according to the unique structure of this algorithm. The implementation of a neural network is based on blocks, with each block containing many types of layers. First, the input layer, the convolution layer and then the activation function used, the Rectified Linear Units (ReLU), the normalization layer and the pooling layer. In addition to the fully connected classification layer and the softmax layer. The proposal was evaluated, and the results showed the overall accuracy rate (98.029%) in the testing phase and (98.29%) in the training phase of the dataset. Before evaluating the model, some image pre-processing and segmentation are performed to streamline the model evaluation process better. At this stage, the image size is changed from 256*256 to 128*128 to reduce the image size, increase the speed of the transmission method, and calculate the test pattern.

Díaz-Pernas et al., 2021 proposal was based on a brain tumor classification model using a deep convolutional neural network that includes a multiscale approach. The proposal focused on processing the input images at three spatial scales along different processing pathways. The model analyzes the MRI images that contain three types of tumors: meningioma, glioma, and pituitary tumor, over sagittal, coronal, and axial views, and therefore there is no need for pretreatment of the insertion images to remove parts of the skull or spine in advance. To increase the data, the use of elastic transformation was relied upon to increase the training data set and prevent overfitting. The results of the evaluation showed an accuracy of tumor classification of 0.973.

Bhanothu et al. (2020), the R-CNN deep learning algorithm has been proposed for tumor detection by marking its region of occurrence with a region suggestion network (RPN). The proposal was based on the VGG-16 architecture as the base layer for the region proposal network and the classifier network. The proposed results showed a mean average accuracy of 77.60% for all glioma, meningioma, and pituitary tumor classes, with a mean accuracy of 75.18%, 89.45%, and 68.18%, respectively. As a data preprocessing step, each image in the MRI dataset is normalized using the min-max method.

Siar & Teshnehlab (2022) proposed a method to detect a convolutional neural network in three main stages. The first is feature extraction generated by various methods, followed by feature selection. The final stage involves combining the results and determining weight factors for each chosen method. This approach showed an accuracy of 99.76% and improved the original method by 2%. Additionally, a third category called "I do not know" was added to increase the reliability of the tumor detection process.

### 2.3.2 Hybrid Methods in Related Work

Deepak & Ameer (2021) proposed a Siamese neural network (SNN) consisting of three fully connected layers, which is designed to have lower complexity and fewer parameters than deep transport-learned convolutional neural networks (CNN) to extract features from magnetic resonance imaging, and perform nearest neighborhood analysis of the k-nearest model (k-NN), using Euclidean and Mahalanobis spaces on a 2-D encoded SNN feature space, so that it is computationally less dense. After being applied to the Radiopaedia, Harvard, and Figshare datasets, the proposed method proved

respective classification accuracy on cross-validation of 92.6%, 98.5%, and 92.6%. Preprocessing the data includes normalizing all images to zero mean and unit variance and resizing images to 128 x 128 so that the image size equals the size of the input layer in the SNN. As for the validation, 5-fold validation was followed. In addition, the SNN model is applied to all three modalities provided by the dataset separately, and images corresponding to the three tumor classes for a given modality are provided for each SNN-kNN arrangement. Combine the performance of each classifier to get the overall classification performance.

Sadad et al. (2021) proposal contributed to the brain tumor classification process by introducing the concept of segmentation using the Unit architecture with ResNet50 as a backbone in the Figshare dataset and achieved a level of 0.9504 of the intersection over union (IoU). They provided a framework called NASNet, which uses evolutionary algorithms (EAs) and reinforcement learning (RL) for the optimization task. In addition to applying other deep transfer learning models such as ResNet50, DenseNet201, MobileNet V2, and InceptionV3 to classify brain tumors and compare them with the NASNet model to show the efficiency of the NASNet model. In the preprocessing process, the contrast stretching algorithm is used to produce high-resolution images, and the data augmentation process is used to achieve better results in the classification process. The results are shown by applying different CNN models for tumor classification, such as MobileNet V2 with 91.8 accuracies, Inception V3 with 92.8 accuracies, ResNet50 with 92.9 accuracies, DenseNet201 with 93.1 accuracies, and NASNet with the highest accuracy. 99.6%. To extract essential features from the magnetic resonance imaging slides, the proposal focuses on the implementation of freezing and fine-tuning transfer learning. Brain tumor multiple classifications are also

performed using Transformational Learning and ResNet50-UNet, and the NASNet architecture.

Munira & Islam (2022) convolutional neural networks (CNN) and machine learning (ML) classifiers have been proposed as hybrid deep learning techniques for classifying brain cancers. The proposal developed a 23-layer CNN architecture to extract deep brain features from magnetic resonance imaging (MRI). Using random forest (RF) classifiers and support vector machine (SVM), the CNN model's in-depth features were evaluated from the flat layer. The results showed that the accuracy rate of the CNN-RF model is 96.52%, while the accuracy rate of the CNN-SVM model is 95.41%.

In this research, a global dataset of Brain tumors is used. Different pre-trained deep learning models such as CNN, VGG16, VGG19, Xception, ResNet50, and DenseNet121 were applied to classify and diagnose Brain tumors. On the other hand, transfer learning was applied to the VGG16 model to extract features from the dataset and to improve the results of brain tumor classification accuracy. An augmentation model was applied to the dataset to improve the model training process; this step improved the results of the accuracy in the model testing process forVGG16 model.

# Chapter Three

# Methodology

## 3.1 Proposed Method

This chapter highlights the infrastructure used in this thesis and explores and defines the available pre-trained networks based on their construction, architectural design, and trainable standards. The networks used in this thesis are also considered.

Pre-trained networks are referred to as the networks trained on Image Net, which contains many images, exceeds 15 high-resolution images available, and the number of categories exceeds 22000. The networks are trained on 1.2 million images with 1,000 defined classes (Deng et al., 2009), which are trained for several days in many classes, and, given their ability to identify specific features, must generalize well. Hence the reason for preferring the use of pre-trained networks. The result of the initial learning process shows its importance in reducing the computational costs of training a new network and overcoming the problem of having a relatively small data set. The previously trained networks described in this chapter are available through the Keras library in python.

This chapter of the thesis will analyze the utilization of pre-trained deep learning (DL) models in the context of brain tumor classification. To evaluate the performance of various models, the following pre-trained models will be compared: VGG16, VGG19, Xception, ResNet50, and DenseNet121. A custom Convolutional Neural Network (CNN) will also be developed to compare its performance with the pre-trained models

and determine the best model for improved classification accuracy. The methodology and results of the DL models will be thoroughly described in this chapter.

To enrich the thesis, we will use a pre-trained model's architecture, such as VGG16, and apply Transfer Learning approaches to classify brain tumor image datasets. In addition to describing the pre-processing steps on the data set. Finally, we will explain the performance metrics used to evaluate the model's performance.

## 3.2 Data Preprocessing Phase

The study dataset included unbalanced samples from three types of brain tumors, meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices), with significantly fewer standard images than meningioma images in both the training and also in the test group and a relatively small data volume. This can lead to poor validation and overfitting after training. Therefore, we apply data expansion to the original datasets and generate new images through horizontal inversion and histogram equalization to compensate for the missing data size. In addition to normalizing image pixels, input scaling affecting images contributes equally to overall loss.

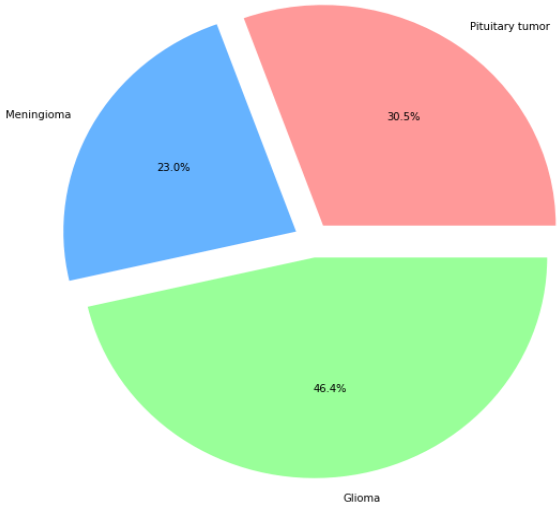



Figure 3.2-1: The percentage of the data set

### 3.2.1 Image Preprocessing: Scaling Image Pixels

From the grayscale image in Figure (3.2-2), each digital image is formed by pixels with values in the range from 0 to 255. 0 is black, and 255 is white. Since 255 is the maximum pixel value. Rescale to 1. /255 is also called the input normalization process in which pixels in the range [0,255] are converted to the range [0,1].

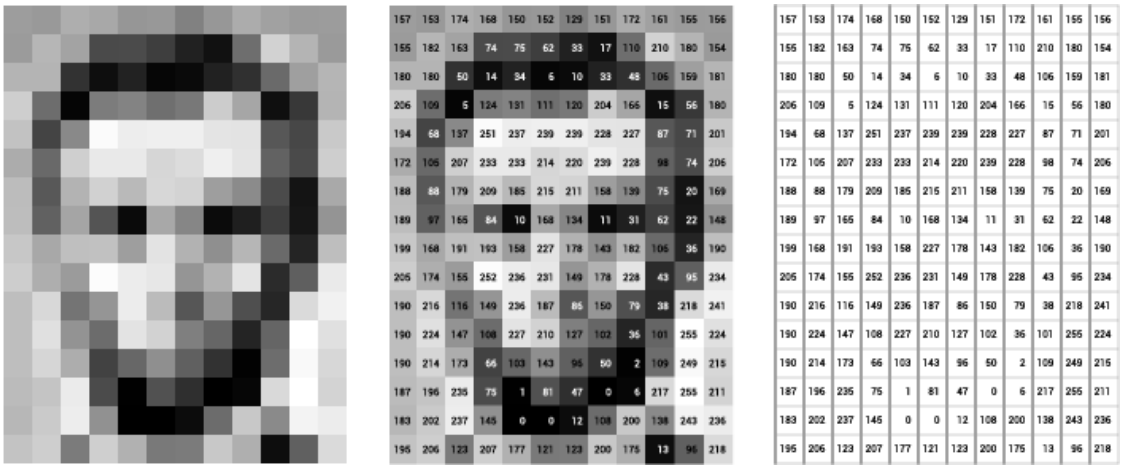


Figure 3.2-2: Digital image as a matrix of numerical values. Numerical values represent the pixel's brightness in this example (Levin, 2022)

All images are treated the same in terms of subscribing to the same model, same weight, and learning rate, although the pixel range is different for each image which can be high or low pixel range. Without scaling, on the one hand, high-range images tend to create stronger loss due to their significant role in determining how the weights are updated. On the one hand, low-range images tend to have a weak loss, and the sum of the two contributes to the back propagation process. Therefore, scaling each image to the same range [0, 1] makes the images contribute equally to the total loss. It also affects the chance of convergence of the neural network because it makes the coefficients in the range of [0,1] vs. [0,255], and thus affects the model to process the input faster. In addition to the above, it affects the typical learning rate in terms of the amount of the learning rate, since images with a higher pixel range produce a higher loss and therefore a smaller learning rate should be used, otherwise, images with a lower pixel range should use a larger learning rate (Chollet, 2016).

### 3.2.2 Data Augmentation

Data augmentation is artificially increasing the amount of data by creating new data points from existing data and, in other words, expanding the data by creating more data for training deep learning models. This helps the model overcome the problem of over fitting resulting from the CNN learning network learning well in the training data, but its performance could be better in the test data. This technique helps the process of increasing the data without the need to collect it manually (Khamparia et al., 2021). There are many data augmentation techniques, in this research; we will use horizontal flipping and histogram equalization. Data augmentation increased the dataset from 3064 to 5022 images. Figure (3.2-3) shows the distribution of images in the dataset between images of tumor types after the data augmentation process.

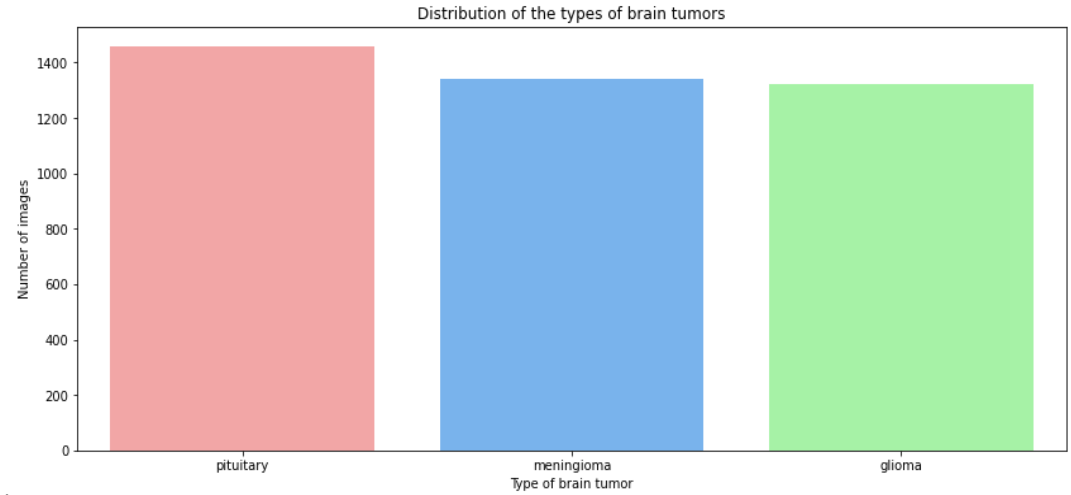



Figure 3.2-3: Data distribution dataset after data augmentation

### 3.2.2.1 Horizontal Flip Augmentation

Flipping means that the image is rotated on the horizontal or vertical axis. A horizontal flip increment is called flipping entire rows and columns of an image pixel horizontally.

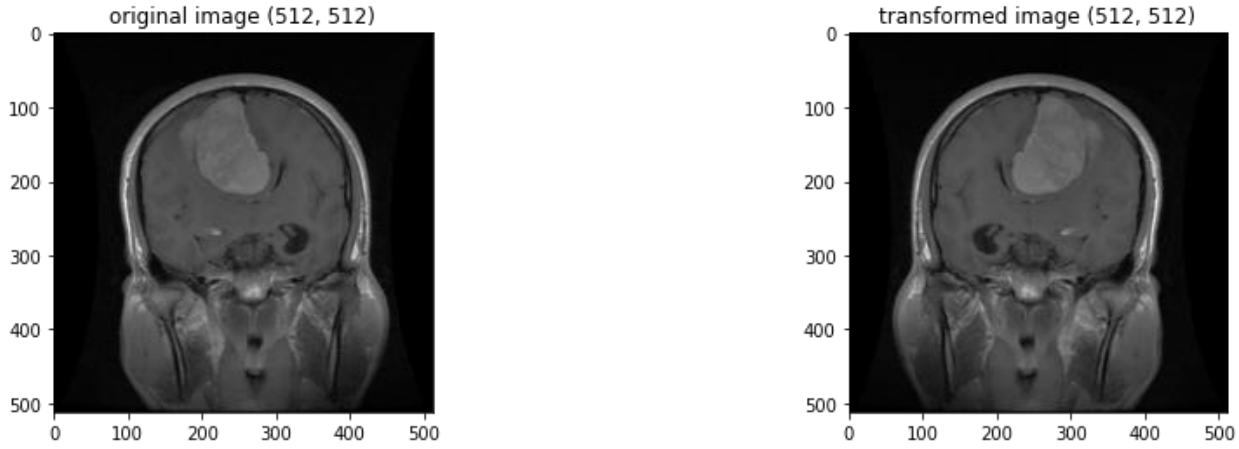



Figure 3.2-4: shows the horizontal flipping application in the form of an MRI

### 3.2.2.2 Histogram Equalization Augmentation

An image histogram is a graphical representation of the intensity distribution of an image, specifying the number of pixels for each intensity value considered. The contrast in the image is improved by what is known as histogram equalization to extend the

intensity range. In other words, the image (2.2-5) pixels appear homogeneous around the middle of the available range of intensities. After applying the equation, the graph equation extends this range.

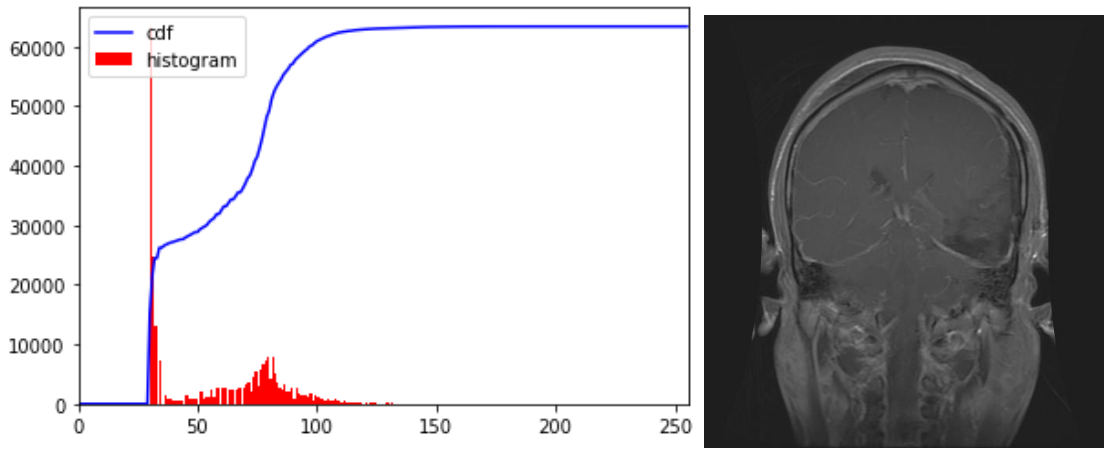



Figure 3.2-5:Histogram for image MRI

After applying the equalization, we get a histogram like a figure (3.2-6) the resulting image is shown in the picture at right.

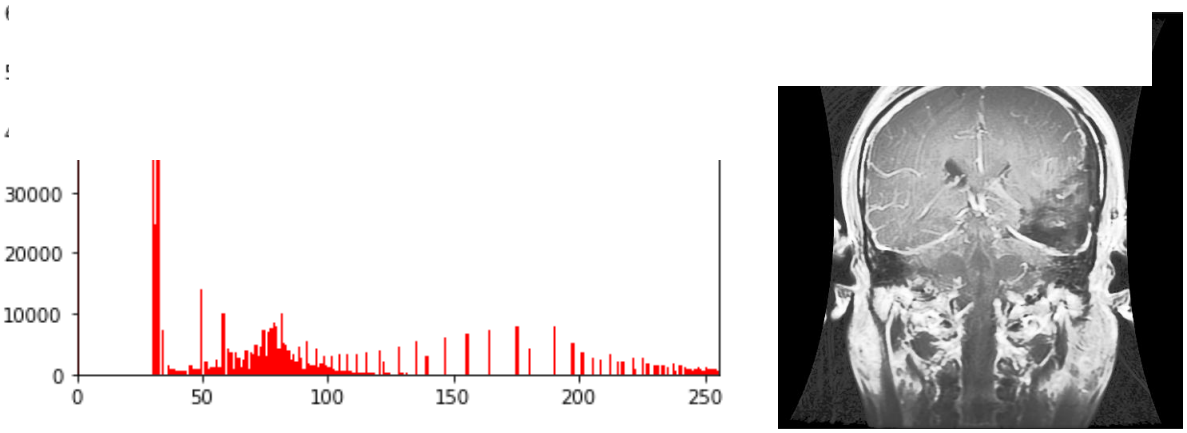



Figure 3.2-6: Histogram the equalization for image MRI

## 3.3 Building Models Phase

In this part of the methodological chapter, we will highlight the different models applied to the data set used in this thesis. Pre-trained deep learning models were applied to the data in several parts. The first part: A cable was applied from the models VGG16, VGG19, Xception, ResNet50, and DenseNet121, in addition to building a CNN network and a CNN, and then a comparison was made between the classification accuracy between the models that were applied. The second part includes applying the learning transfer mechanism on the VGG16 model with the process of tuning the model. The third part includes applying the data augmentation process to the data set to improve the model training process in classifying brain tumor types' magnetic resonance image sets.

- **First Part: Building Models Phase**

In this part, we discuss the deep learning models that we used to classify global data; the figure (3-7) shows the scheme of the algorithms in general.

The research scheme for classifying brain tumors in MRI brain images.

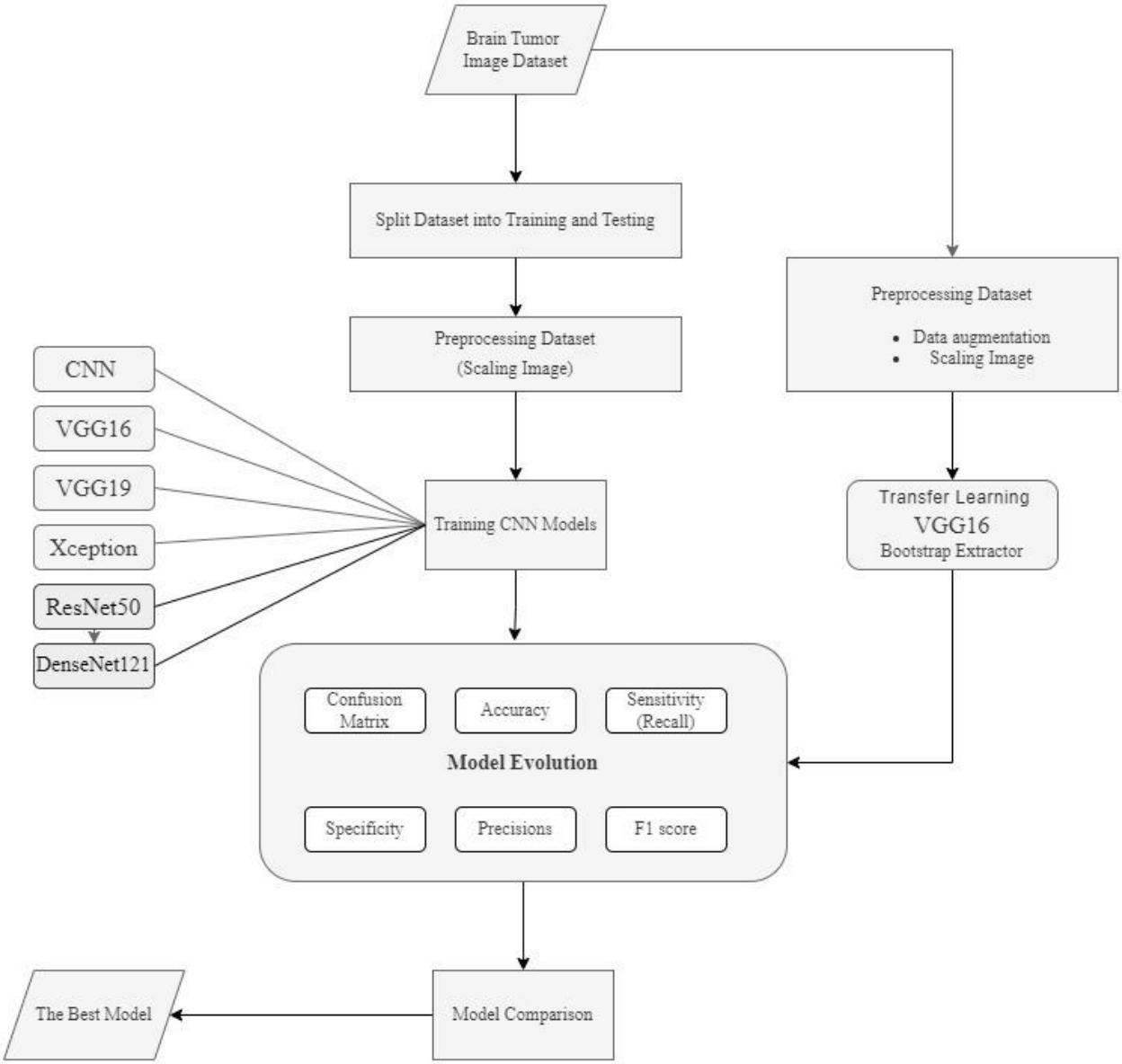



Figure 3.3-1: Research Scheme

### 3.3.1 Convolutional Neural Network (CNN)

A convolutional neural network (CNN) is a deep learning model inspired by the organization of the animal visual cortex for processing data (such as images) containing a reticular pattern. CNN design aims to deal with 2D and 3D structures, which are achieved through local receive field connections and standard weights. In addition to,

learning spatial hierarchy automatically and adaptively for features, according to models from low to high levels (Hubel & Wiesel, 1968). From one of the various deep learning architectures comes the multi-layered neural network model of spatial data, the Convolutional Neural Network (CNN or ConvNet). The main advantage of CNNs is that they contain much fewer trainable parameters due to the weights allocation feature of CNN (Kim, 2016), which affects the processing process in terms of reducing overfitting and optimizing the generalization process. The strength of CNN lies in its ability to output the model in a more structured and feature-based output because the classification layer and the feature extraction layer learn together (Ghosh et al., 2020).

CNN architecture consists of a mathematical structure consisting of a model consisting of three types of layers or what is known as blocks, namely: convolution, pooling, and fully connected layers. The function of the first layer aims to extract the features as a result of learning the spatial and temporal features. The function of the second layer is to downsample or reduces the spatial size of the input elements, such as the dimensions of the input image. In contrast, the third layer aims to assign the extracted features in the final outputs, such as classification (Goyal et al., 2019); (Albawi et al., 2017). A convolution layer consists of a stack of mathematical operations, an example of which is the convolution layer, a specialized linear operation type. Digital images are treated as input to the network based on pixel values and stored in a two-dimensional (2D) grid, an array of numbers. Then it uses the kernel/filters, which form a small network of parameters to extract the features. The network is defined as an optimizable feature extractor at each image position, which gives CNN high image processing efficiency, which may happen feature emergence anywhere from the picture. As layers in a

network pass their output to the next layer, the extracted features can become hierarchical and increasingly complex (Benzebouchi et al., 2019).

Figure (3.3-2) shows an example of CNN structure.

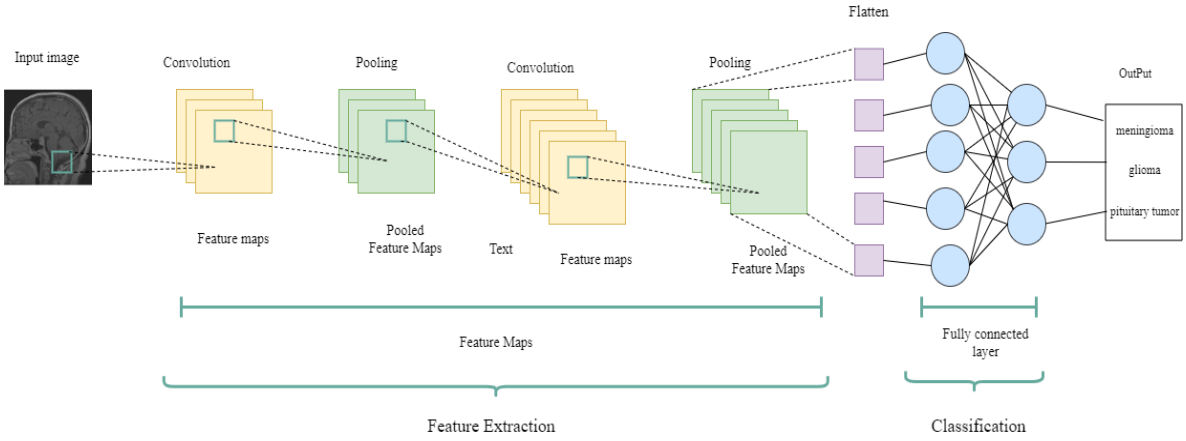



Figure 3.3-2: Convolutional neural network

### 1.1.1.1 Extraction Feature

This part consists of input layers, Conv layers, activation functions, and pooling layers.

1- Input Layer: Input layers contain the raw data fed to the network, multi-array data. First, the images are loaded, the data is stored for the processing stage in the network, and the input data is specified in the form of the image's width, height, and depth.

2- Convolution layer: known as the convolutional layer (CONV), which is the primary building block of the CNN algorithm with most of the arithmetic operations that transform the input data from the previous layer to the next layer (Shaha & Pawar, 2018).

3- Activation functions: In a convolutional neural network, it is necessary to determine which neurons are to be activated or not from a given set of input

nodes. The activation function is to achieve the last goal. The most common activation systems in convolutional neural networks are sigmoid, ReLU, tanh, and softmax functions (Murphy, 2016); (Mukhopadhyay, 2018); (Ying et al., 2019).

#### 1.1.1.2 Convolution Operation

Convolutional neural networks perform an arithmetic operation to calculate the correlation between the weights of the kernel and the pixels or output of the previous layer, known as a convolution operation. To produce the feature map (F(i,j)), some kernel/filters (K(i,j)) wrap with the input image (I(i,j)).

$$F(i,j) = (I * K) * (i,j) = \sum_{\mathrm{m}} \sum_{n} I(m,n)K(i-m, j-n) \qquad (3.1)$$

The input image is represented as I and the convolutional kernel is represented as K. The indexes of rows and columns of the result matrix are marked with I and J, respectively, while those of M and N deal with that of the kernel. In the conv layer, to detect the local features at different positions, the raster product between the input image and the kernel is computed by moving the kernel across the image by Conv. layer. Mathematically, a significant value is obtained in feature maps since the filter template is located in the input image. This enables the capture of similar image features of the nucleus. In CNN, filters are U-weights set during network training via the back propagation algorithm (Giri et al., 2020). Kernel weights are learned, which enables it to extract relevant features from the input, as they are spatially shared with all locations in the image. This is because the feature extracted by the kernel in a specific part of the image is also associated with other parts of the image. This is what distinguishes CNNs

by reducing the number of parameters. The operation is done through the array rather than the multiplication and individual grouping of elements in all widely used deep learning libraries (Zabihollahy et al., 2021).

#### 1.1.1.3 ReLU

The Rectified Linear Unit (ReLU) is defined as a non-linear activation function widely used in deep neural networks. What distinguishes it is that there is no parameter inside the ReLU layer, so there is no need to learn the parameter in this layer. As shown in Figure (3.3-3). In the case of images, the idea of the ReLU function is to allow passing through features that result in positive patterns for specific patterns, which causes other negative patterns to be equal to zero. This method reduces complexity and memory consumption due to parallaxes in the model by making the convolutional neural network converge faster because the activation functions perform simple threshold operations instead of complex operations. On the other hand, other common activation functions, such as the sigmoid and the tanh, seek to maximize saturation (Glorot et al. 2011).

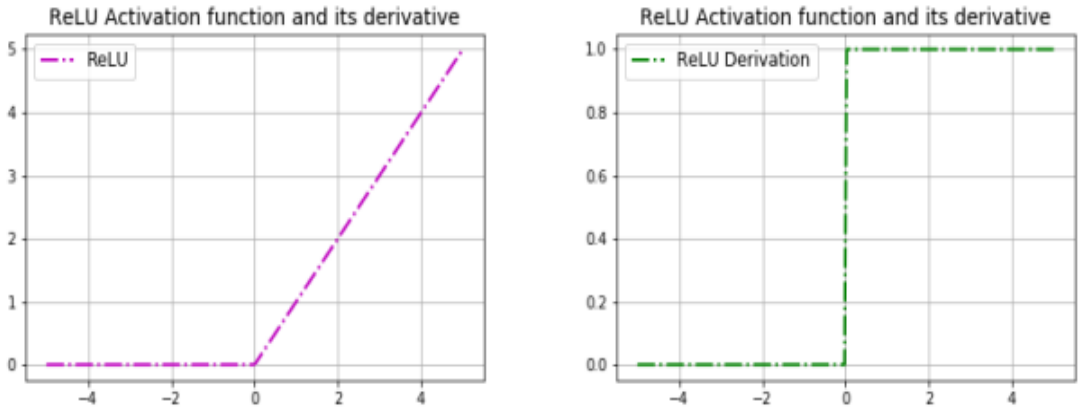



Figure 3.3-3:ReLU activation function

According to equation (3.1), the maximum value between zero and the input value is the ReLU output. The value of the input is zero if the input is a negative value, while the input value is if the input is positive. According to the following equation:

$$f(x) = \max(0, x) \qquad (3.1)$$

The derivative of ReLU is:

$$f'_{(x)} = \begin{cases} 1, & if x > 0 \\ 0, & otherwises \end{cases} \qquad (3.2)$$

The input to the Rectified Linear Unit (ReLU) function is denoted as 'x' and its output as 'f(x)'.

#### 1.1.1.4 Softmax Function

The activation functions on the last connected layer differ entirely from the other functions. The appropriate activation function must be selected according to each task. In the model representation, the activation function known as the soft max function is used, one of the non-linear activation functions (Cheng, 2014), as it is used in the multi-category classification task. The activation function adjusts the actual values of the outputs from the last fully connected layer to the probabilities of the used class (Ju et al., 2018). Each value ranges from 0 to 1, and all values are summed to 1. The equation below represents that the output layers receive the Zi input vector and the Yi value output for every single neuron (Rguibi et al., 2020).

$$Y_i = \frac{e^{z_i}}{\sum_{i=0}^{m} \mathrm{e}^{z_i}} \qquad (3.3)$$

The elements of the input vector to the softmax function are represented by 'Zi' and 'M' denotes the number of classes in the multi-class classification problem.

#### 1.1.1.5 Pooling Layer

The functionality of the aggregation layer is known as subsampling by its ability to reduce the dimensionality of data while maintaining the stability of the network structure (Raja, 2019). More specifically, it aims to reduce the dimensions of the feature maps produced by the convolution layer using downsampling, which aims to reduce the number of learnable parameters for subsequent layers, i.e., it combines the features of neighboring pixels into a single feature value. The goal behind this layer is to reduce the computational power required to process data by reducing dimensionality. In addition to extracting dominant traits that are fixed in rotation and position, which are helpful in the practical training process of the model, they also play a significant role in controlling overfitting. The pooling layer performs its function independently on each depth slice of the input and changes its size spatially, as it does not affect the depth dimensions of the input size. There are two most popular pool types, Max Pooling and Average Pooling.

**Max pooling:** Aims to extract corrections in the feature map then return the maximum value in each patch of the feature map and discard all remaining values. Average Pooling: Returns the average of all values from the feature map (Zhou & Sun, 2020).

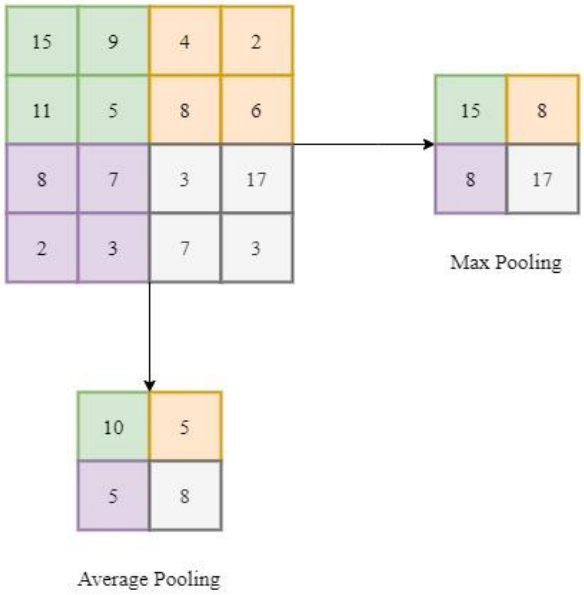



Figure 3.3-4: Difference between Average and Max pooling

The most common use in most ConvNet architectures, Maxpooling, is used due to its ability to reduce computational cost (Raja, 2019).

$$a_j = max_{(p,q)\in R_{ij}} a_{kpq} \qquad (3.4)$$

The pooling operator output connected with the feature map kij is represented by $a_{kij}$ The elements located at the position '(p,q)' within the pooling area 'ij' are denoted by $a_{kpq}$, which represents a receptive field centered at the position '(i,j)'.

#### 1.1.1.6 Flatten Layer

A pooled feature map is produced after the pooling layer. After pooling, the flattened layer is one of the most critical layers since it is necessary for processing to convert the whole matrix representing the input pictures into a single-column vector. The last dense layer that is fully linked provides a classification based on the fed-in inputs received

from this flattened layer (Rawat & Wang, 2017). Figure (3.3-5) shows the process of flattening.

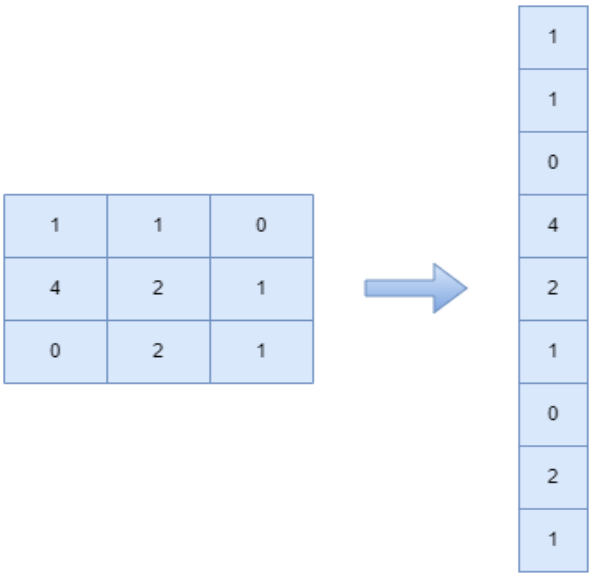



Figure 3.3-5: Flatting in CNN

#### 1.1.1.7 Fully Connected Layer

During the training process, the feature maps of the final convolution or aggregation layer are transformed into a one-dimensional (1D) array of numbers (or vectors), which are connected to one or more fully connected layers, which are known as dense layers, where all the inputs are connected to the outputs Learning weight. After feature extraction by convolution layers and sampling by pooling layers, they are translated to the network's final output, such as the probabilities for each class in classification tasks, via a selection of fully connected layers. The number of output nodes in the final fully linked layer typically equals the number of classes.

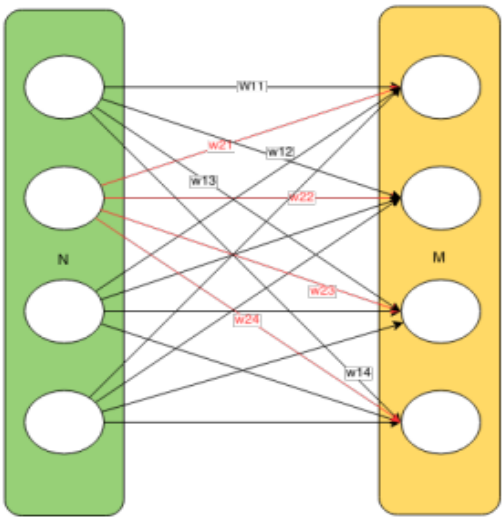



Figure 3.3-6: Fully Connected Layer of CNN (Bengio & LeCun, 2007)

Shows the connection between two layers where the right one is the fully connected layer.

The fully connected layer is formed by connecting each neuron to its previous neural layer such as a neural network and is usually a column vector. As for its activation, it is calculated by multiplying the matrix by its weight followed by the bias like a neural network. The most famous activation function used is softmax at the top of the deep convolutional neural network (Patterson & Gibson, 2017). All of the activation levels from the preceding layers will be fully connected to this layer by a single-dimension vector. These layers' objective is to retrieve the crucial information currently present, classify the class using the SoftMax classifier, and determine the loss function using cross-entropy.

### 3.3.2 Visual Geometry Group-16 (VGG16) Model

VGG-16 is one of the most popular CNN architectures trained using (Image Net), which constitutes a sizeable visual database in the research of visual object recognition

software. VGG-16 is an acronym for Visual Geometry Group, the network's creators (Simonyan & Zisserman, 2014). The VGG-16 network was released in 2014, which Simonyan and Zisserman introduced (Simonyan & Zisserman, 2014). It was considered at that time a deep network. The VGG-16 architecture shows that network depth mainly contributes to better classification accuracy or classification in CNN.

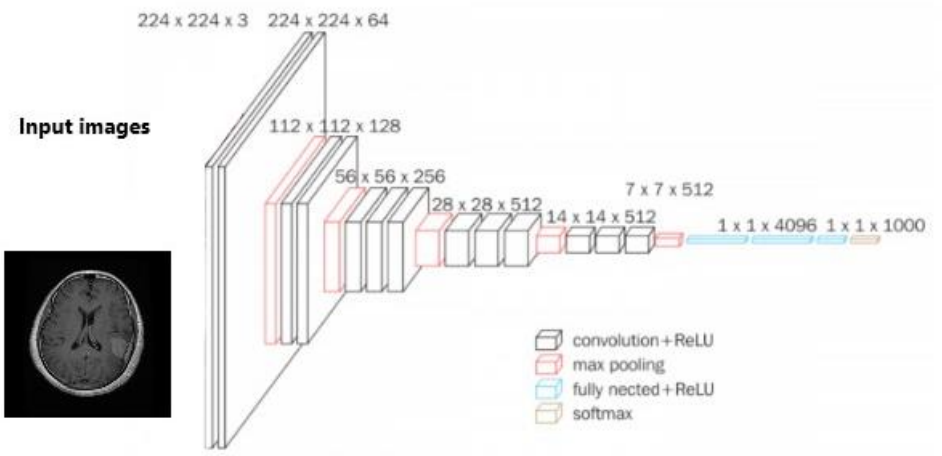



Figure 3.3-7: VGG-16 model architecture.

The VGG16 model has 13 convolutional layers, 3 Fully connected layers, 1 SoftMax classifier, 5 pooling layers, and it uses ReLU. It also has 138 million parameters. The total consists of 16 trained layers, which are used to display filters and feature maps because they have a simple, unified structure consisting of convolution, maxing, and fully connected layers arranged sequentially. The VGG-16 network has smaller filter sizes, which gives it the advantage of making it less computationally expensive than some networks (Basaveswara, 2019).

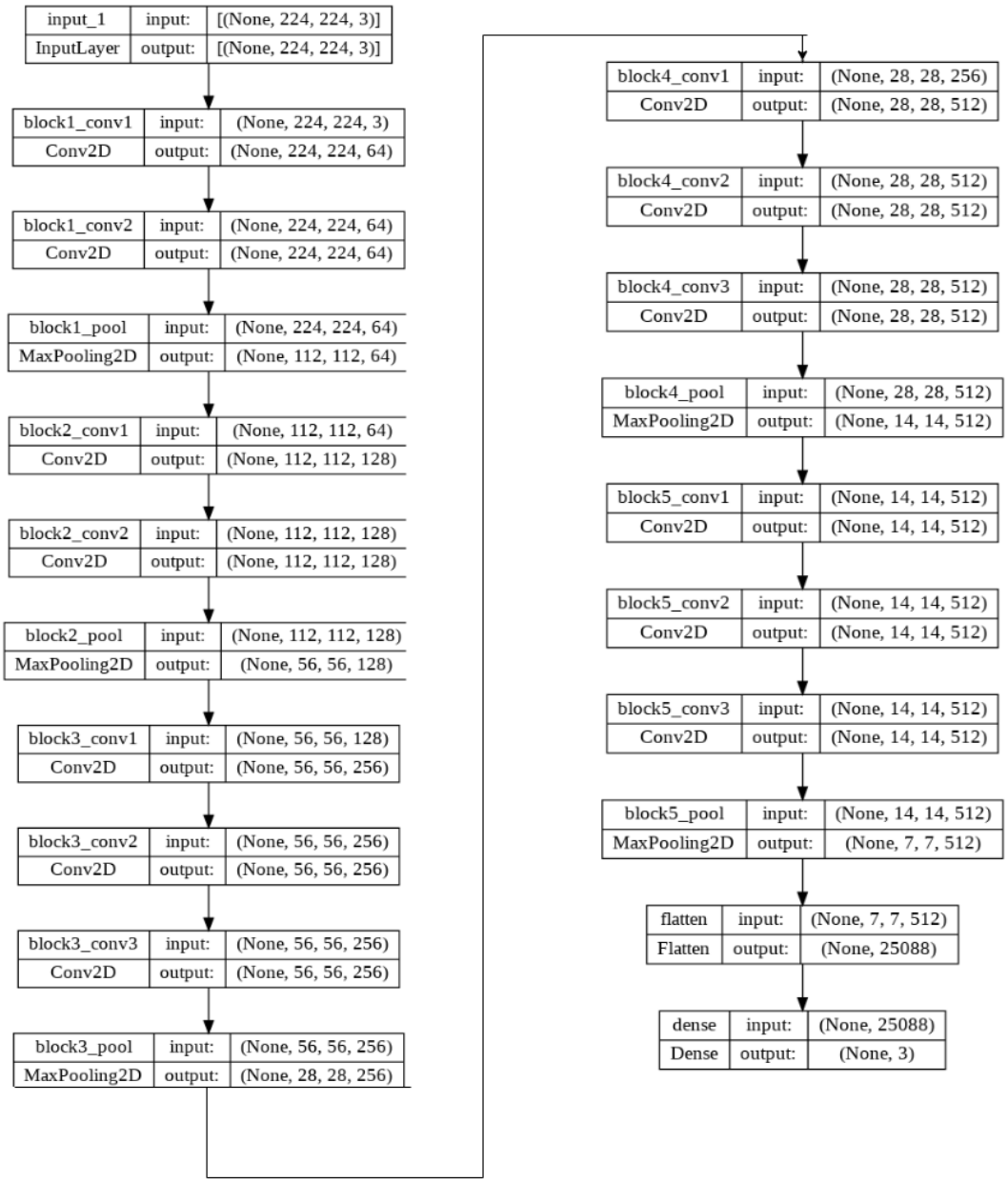



Figure 3.3-8: Stages Implementation of VGG-16 model.

### 3.3.3 Visual Geometry Group-19 (VGG19) Model

VGG19 and VGG16 are similar in network architecture. The network contains 19 layers proposed by (Simonyan & Zisserman, 2014). The VGG19 network consists of 19 trainable layers and 144 million trainable parameters. Which the network consists of 16 convolutional layers and three fully connected layers; it deals with filters with smaller

receptive fields size 3 x 3 to capture the notion of (left/right, up/down, and center), with stride and pad of 1, along with 2x2 max-pooling layers with stride 2, called the VGG-19 model.

There are 41 layers, including Max pool, a fully connected layer; it uses ReLU as the activation function layer in all hidden layers, the Dropout layer, and the Soft max layer. Similar to Vgg16, the input layer has a size of 224x224×3. The last layer is the classification layer.


Figure 3.3-9: Graphical representation of VGG19 (Özyurt, 2020)

### 3.3.4 Residual Networks (ResNet50) Model

The architecture of the remaining networks is different from that of the traditional serial CNN. The goal of ResNet is to solve the degradation problems that occur with the various connections in the CNNs it uses, which arise when deep networks begin to appear (Polat & Güngen, 2021). In this context, the ResNet50 architecture is used, one of the ResNet models that have emerged. The ResNet50 network consists of 50 layers of residual networks. The structure of ResNet50 contains different combinations of identical lavers, as shown in the colors in the figure in Figure (3.3-10). As for the curve lines, they represent the used identification blocks, which indicate the use of the previous layers in the following layer.

ResNet50 is distinguished from the others by the counterfeit problem of vanishing or exploding gradients and the degradation (accuracy to saturation first and then degradation) problem in training intense networks.

The ResNet50 network shown in Figure 6 consists of a set of layers. The first layer comprises three identical blocks, as shown in grey. The first layer contains 64 filters with a kernel size of 77, followed by a max-pooling layer of 3-3. In the same way, the second group, the third group, and the fourth group are formed with 4 indented blocks, 4 identical blocks, and 3 indented blocks, respectively. The curves between some groups are highlighted in blue. Different sizes represent the identity block that connects the two phases. In addition to the above, these blocks are followed by 38 fully connected layers responsible for the classification task (Mukti & Biswas, 2019).

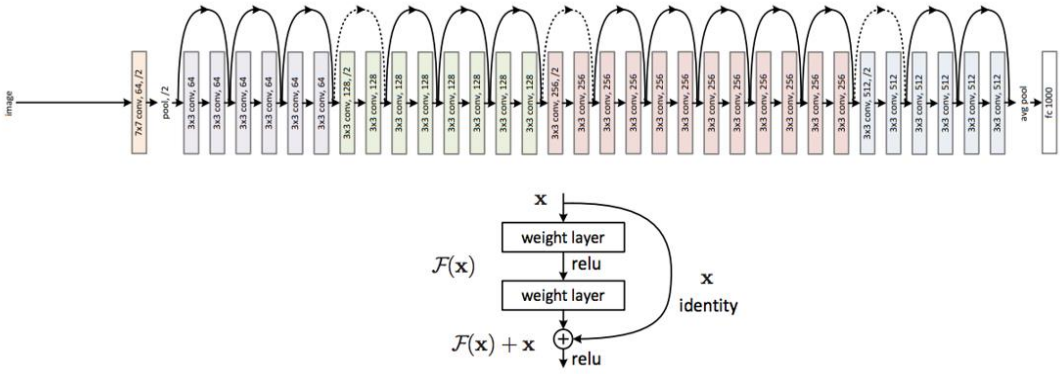



Figure 3.3-10:: Resnet-50 architecture (Mukti & Biswas, 2019)

### 3.3.5 Densely Connected CNN (DenseNet121) Model

The architecture of Deep Convolutional Networks (DCNNs) has the advantage of being the most productive framework for image recognition due to the presence of particular types of convolutional and pooling layers. As the network goes more profound, the input information or gradient that passes through most layers will vanish when it reaches the network's last layer.

DenseNets comes to overcome the problem of gradient vanishing, which is by working to connect all layers with feature sizes equal to each other directly. DenseNets is distinguished by its ability to extract the network's most general and profound features. The pre-trained Connectivity-Densified Convolutional Neural Network consists of 121 layers, which use this network for the feature extraction process proposed by (Huang et al. 2016).

DenseNet-121's network architecture consists of 120 Convolutions, 4 AvgPools, and 1 Full Connected Layer. The architecture of the DenseNet network initially consists of both a basic convolution layer and a pooling layer, 3 transitional layers, and 4 dense blocks. After these layers comes the final layer, which is the classification layer. The performance of the first convolutional layer is 7 × 7 convolutional with stride 2, followed by a max pooling of 3 × 3 with stride 2. Then the network consists of a dense block, followed by 3 sets, each consisting of a transition layer, followed by a dense block.

As Huang et al. (2017) suggested in DenseNets, dense connectivity is received by fetching direct connections to and from any other layer in the network. The network feature maps of all previous layers are received by the lth layer, thus ameliorating

gradient flow across the entire network. This requires the concatenation of feature maps of previous layers, which, unless all feature maps are of the same size, cannot be performed. However, convolutional neural networks primarily aim to sample the size of the feature maps. It is a DenseNets architecture divided into multiple dense blocks connected by density. The layers between the dense blocks are referred to as transitional layers. So that each transition layer in the network consists of a batch normalization layer and a 1×1 convolutional layer followed by an average pooling layer 2-2 that uses a stride of 2. As noted earlier, there are 4 dense blocks. Each block contains 2 convolution layers. The first is 1x1, followed by 3x3. All 4 dense blocks in the previously tested DenseNet169 architecture pre-trained on ImageNet are 6, 12 and 24, and 16. The final layer is the classification layer that performs the global average pooling of 7 x 7, followed by a fully connected final layer that uses “softmax” as the activation (Hao et al., 2016).

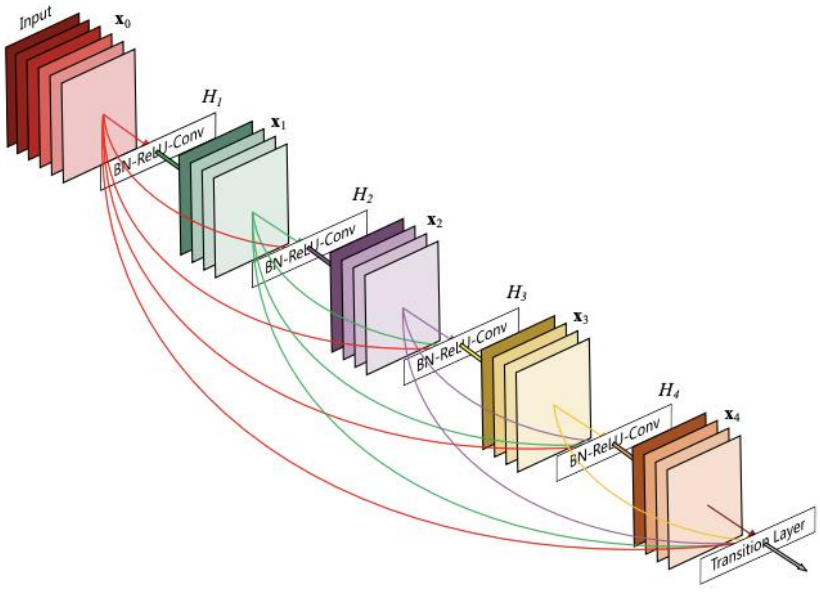



Figure 3.3-11: DenseNet Architecture (Huang et al., 2017)

### 3.3.6 Depthwise Separable Convolutions (Xception)

Depthwise separable convolutions are utilized in the Xception deep convolutional neural network design. Google researchers developed it. According to Google, inception modules in convolutional neural networks serve as a transitional stage between depthwise separable convolution and regular convolution (a depthwise convolution followed by a pointwise convolution). A depthwise separable convolution may be considered an Inception module with a maximum number of towers.

Based on this discovery, they propose a novel deep convolutional neural network design with Inception modules substituted with depthwise separable convolutions. The data enters the model through the entry flow, then the middle flow, repeated eight times, and finally, the exit flow. Furthermore, batch normalization follows each Convolution and Separable Convolution layer (Chollet, 2017).

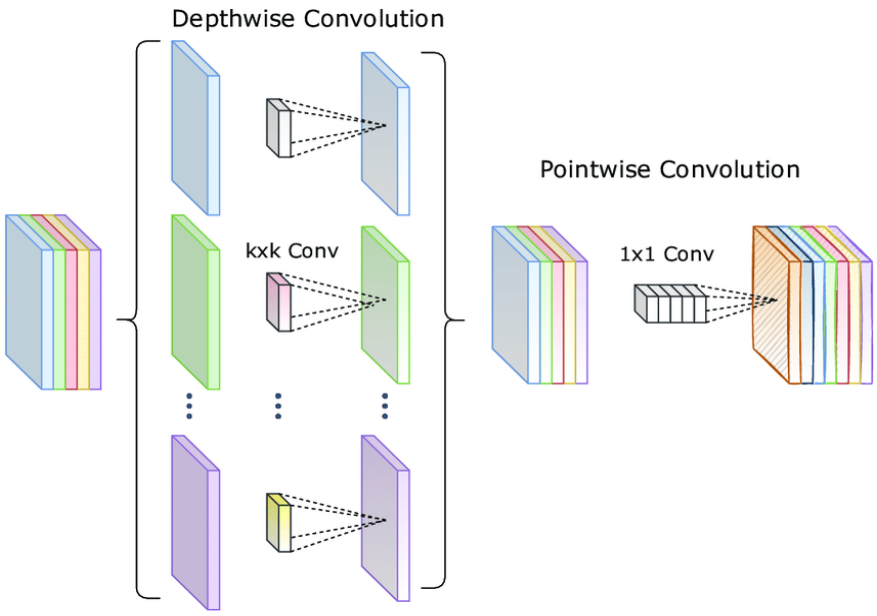



Figure 3.3-12: Depthwise Separable Convolutions architecture (Sultonov et al., 2022)

### 3.4 Transfer Learning

The convolutional neural network for the training process requires extensive data. This is difficult to achieve due to the difficulty of collecting extensive data for classification. It is also sometimes not possible to match training and test data. The concept of transfer learning is one of the advanced machine learning methods that use the concept of learning the required knowledge in solving a problem and reusing it to solve another problem in any related field.

The concept of transfer learning refers to a method of knowledge exchange that reduces the volume of training data, time, and computational costs when building deep learning models. The network is trained using an appropriate data set for the specific task. Then the task is transferred to a specific target that is supposed to be trained by a target data set. That is, it works to transfer one task to another to learn specific features in the neural network in identifying Images for a specific data set. It develops models for solving tasks and transfers model knowledge to related tasks to solve the same problems in an image recognition task. The goal behind the transfer learning concept is to transfer the learning of a pre-trained model to a new model due to the difficulty of training a convolutional neural network from scratch due to the large amount of image data required (Shaha & Pawar, 2018).

Figure (3.3-13) illustrates the concept of transfer learning by using the model pre-trained on a large standard dataset such as ImageNet (Deng t al., 2009) as a feature extractor for a different task to train a relatively small dataset such as an MRI dataset. The results of using transfer learning lead to reducing the long training time required by

neural networks and replacing the requirement for an extensive data set by using a smaller size to train the model (Tajbakhsh et al., 2016); (Pan & Yang, 2010).

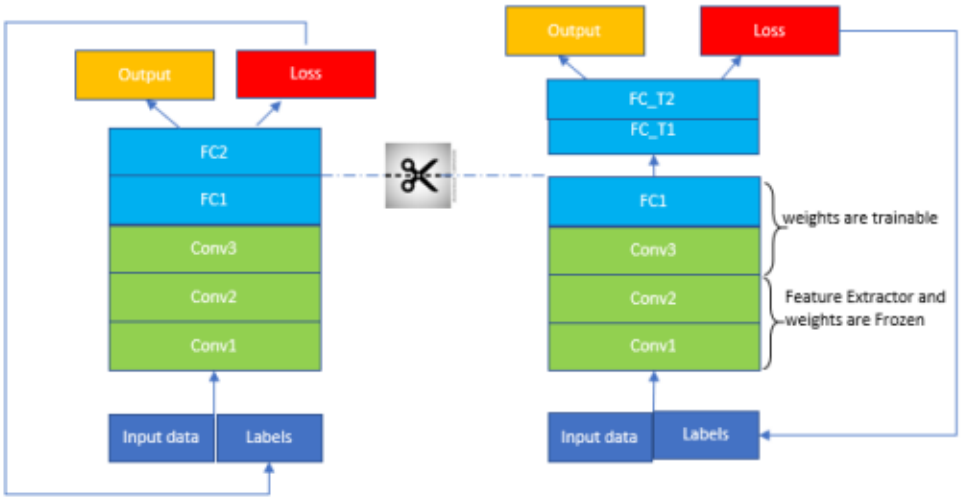



Figure 3.3-13: Implementation of Transfer Learning Method (Tahamid, 2020)

Figure (3.3-13) shows the process of transfer learning to extract high-level features by using the first convolution layers and working to change the fully connected layers to the new data set. The scenarios of the learning transfer process are Convolution Network as a static feature extractor, ConvNet fine-tuning, as the static feature extraction takes ConvNet pre-trained on ImageNet, and work to change the last classification layers, which is that the output of this layer is 1000 class degrees for a different task Like ImageNet, for the new dataset, ConvNet is used as a static feature extractor. ConvNet is configured according to the new dataset for training the network (Shu, 2019) by fine-tuning the weight and modifying the fully connected layers. To train the network, the weight and bias initialization contained in the respective model is used (Vasilev et al., 2019).

### 3.4.1 Deep Feature Extraction Using Pre-Trained CNN Models

We suggest using pre-trained models to determine the classification of brain tumors. There are cases to take advantage of the benefits of transfer learning. Among them, the first case lies in transferring learning as a classifier, in which the classification process is done by using the model before training directly by changing the last layer, i.e., the softmax layer is changed so that it contains three neurons to classify brain tumors into one of the three types, meningioma, glioma, and pituitary tumor. The second case lies in transfer learning as a feature extractor using deep learning models. Deep learning models consist of multi-layered architectures that learn different features at different levels. To do the final task of getting the output, these layers are eventually connected to a set of fully connected layers. This architecture allows using pre-trained networks such as VGG without including its final layer for feature extraction and other tasks.

A stand-alone feature extractor, the pre-trained model, or a subset of the model, is used to preprocess pictures and extract pertinent features. Bootstrap or integrated feature extractor is a new model prefixed with the previously trained model or a desired section of the model. In this scenario, the R weights of the pre-trained model may be frozen to prevent them from changing while the new model is trained.

In another case, transfer learning as a feature extractor and fine-tuning. In this case, we retrain some of the previous layers and replace the final layer to classify which is more active. It involves fine-tuning the pre-testing model on Bootstrap, fully connecting new layers and outputs, freezing the previously tested layers, and then unfreezing and retraining the previous few pre-tested layers. Frozen layers modify features. The non-frozen, i.e., trainable, layers will be trained on the assigned data set and updated based

on the predictions of the fully connected layers (Subramanian & VE, 2022). In our work, we trained five deep learning models to classify brain tumors as a transfer learning as a classifier process in addition to building a CNN model, in addition to training the VGG16 network as a transfer learning model as a feature extractor and fine-tuning the model to improve the classification process of brain tumor dataset. The precise structure of the VGG-16 network is shown in Figure (3.3-14) as follows:

- The first and second convolutional layers comprise 64 feature kernel filters with a filter size of 3×3. The input picture size (RGB image with depth 3) changed to 224x224x64 and passed into the first and second convolutional layers. The output is then sent to the max pooling layer with a stride of 2.
- The third and fourth convolutional layers comprise 124 feature kernel filters with a filter size of 3×3. Following these two layers is a max pooling layer with stride 2, and the resultant output is 56x56x128.
- Convolutional layers with a kernel size of 3×3 are used in the fifth, sixth, and seventh levels. All three of them make use of 256 feature maps. A stride 2 max pooling layer follows these layers.
- Eighth through thirteenth are two groups of convolutional layers with kernel size 3×3. All of these convolutional layer sets contain 512 kernel filters. Following these layers is a max pooling layer with a stride of 1.
- Fourteen and fifteen layers are fully connected hidden layers of 1024 and 512, respectively units, followed by a softmax output layer (Sixteenth layer) of 3 units. Before and after each fully connected layer, a dropout layer is added.

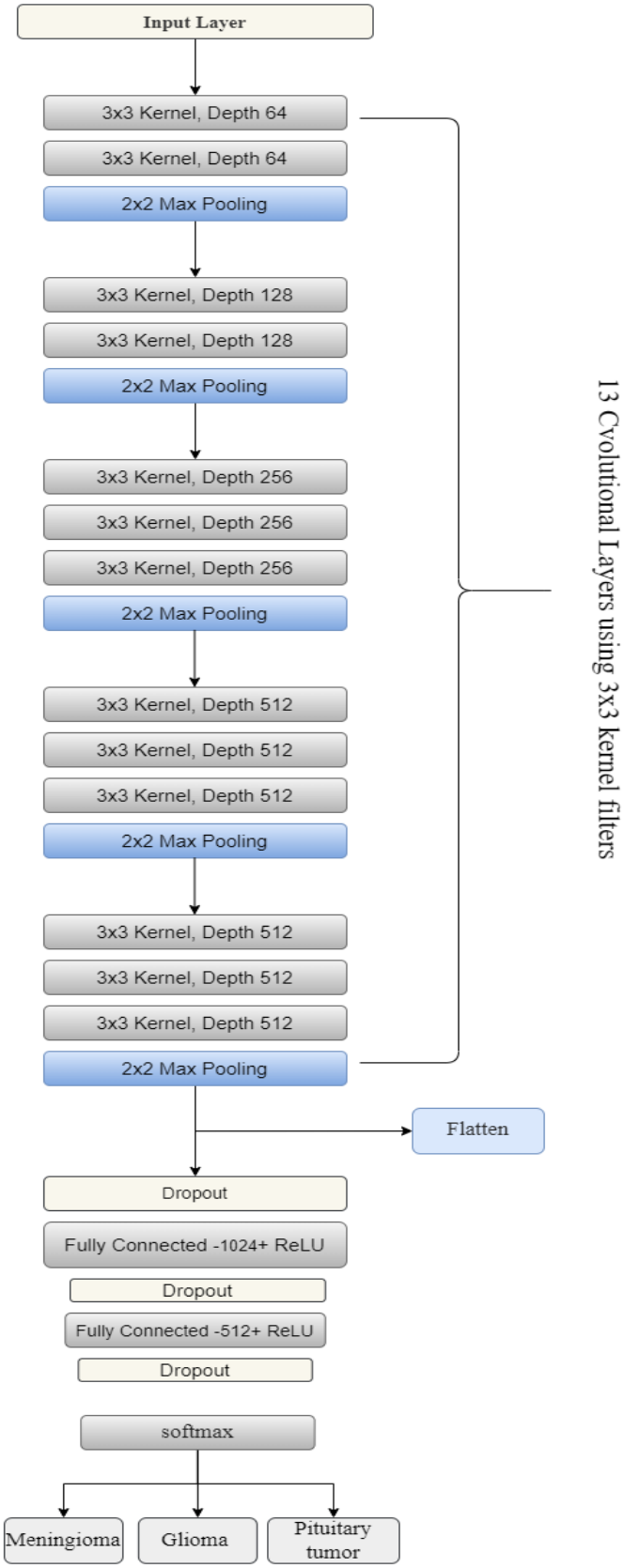



Figure 3.3-14: The architecture of VGG16 as a feature extractor and fine tuning

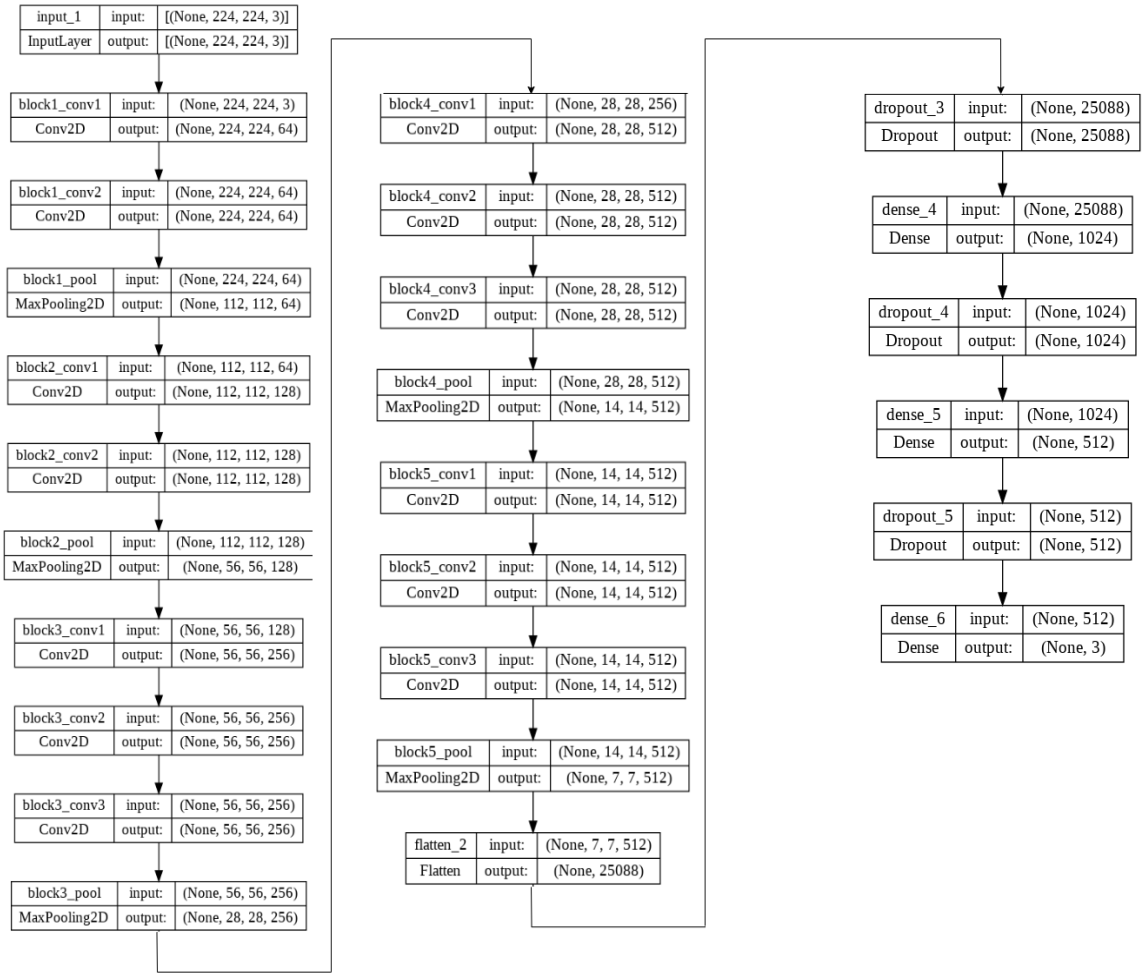



Figure 3.3-15: Stages Implementation of transfer learning VGG-16 model.

## 3.5 Performance Metrics Selection

After developing and training deep learning models, the next step is to determine their effectiveness using performance metrics of test datasets. These models are evaluated using a variety of performance measures. Measures for evaluating the effectiveness of the brain tumor classification system are based on the expected and observed outputs of the data set. The obtained results are the actual outputs, and the observed outputs are accurate labels for the data. There are several techniques to assess a convolutional neural network's performance, but the three most common ones are accuracy, precision, and recall (sensitivity). Accuracy describes the number of accurate classifications

throughout the whole pool of data. While some research favor using accuracy as the only parameter to assess the model, accuracy overlooks the balance between false positive and false negative percentages. If the false negatives and positives are different, two networks with distinct true negatives and true positives can produce outputs with identical accuracy. Therefore, that might give a false impression of the importance of the class outcome. Some main performance metrics like; the accuracy, confusion matrix, and area under curve "AUC" were shown for all models VGG16, VGG19, Xception, ResNet50, DenseNet121, and CNN.

We need to discuss the performance matrices to determine how correctly our model performs. This part will discuss the performance metric we use to assess our model. In this section, we will discuss the performance indicator that we employ to rate our model. We need to familiarize ourselves with a few terms connected to performance metrics. We focused on those utilized for classification tasks. Therefore, we employed the F1 score, recall, accuracy, and precision.

- **Accuracy:** is used to determine the model's ability to assign the test samples to the appropriate classes. The best-known performance matrix measures how often the classifier makes the correct prediction. Mathematically, accuracy is the ratio of correctly predicted frames to the total number of symbolically represented structures.

$$\text{Accuracy} = \frac{Correct\ Expectations}{Total\ Expectations} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

It should be noted that there are limitations if there is an imbalance in the studied data, which affects the uneven distribution of the categories samples. At the same time,

precision considers all data correctly classified (TP + TN) and finds the general ratio of these samples to the entire data set.

- **Confusion Matrix:** One of the most intuitive measures for determining the accuracy and correctness of the model is the confusion matrix. It is a special tabular arrangement that enables viewing of an algorithm's performance and is also referred to as an error matrix. The matrix's rows correspond to the predicted class states, while its columns correspond to the actual class states (or vice versa). The name refers to how simple it is to determine whether the system conflates the two groups, i.e., mistaking one for the other.

The confusion matrix is a useful tool in the evaluation of multi-class classification problems. It consists of four parameters, which are:

- ✓ True positive (TP): The true class of the sample is a positive example, and the result that the model classified correctly is also a positive example.
- ✓ True Negative (TN): The negative class of the sample is a negative example, and the result correctly classified by the model is also a negative example.
- ✓ False Positive (FP): The true class of the sample is negative, but the model classifies it as positive.
- ✓ False-negative (FN): The true class of the sample is a positive example, but the model classifies it as a negative example.

The model must make accurate predictions to be considered clever. This implies that while simultaneously reducing the mistakes, which will result in a reduction in the false positives and negatives, the true positives and true negatives should be as high as

feasible. The TP and TN ratios should be extremely high, while your FP and FN ratios should be extremely low (Markoulidakis et al., 2021).

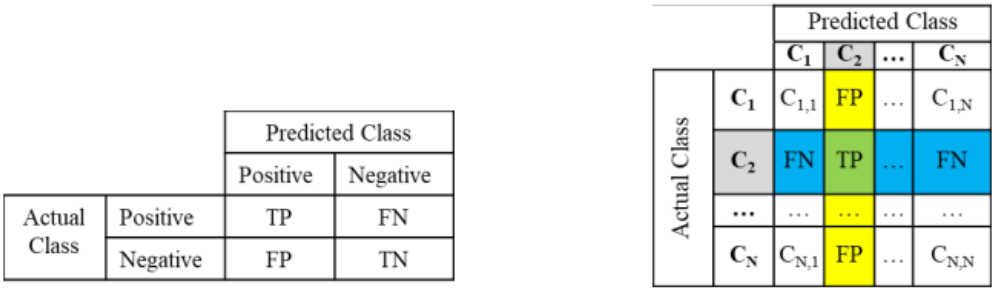



Figure 3.5-1: Confusion matrix examples.

(a) Binary classification problem confusion matrix. (b) Multiclass classification problem confusion matrix.

- **Sensitivity or Recall:** measures how capable the system can classify types of brain tumors, and it is described as a measure of accurately detected positive proportions. Sensitivity is computed from the proportion of true positives using the relation:

$$Sensitivity = \frac{TP}{TP+FN} \tag{3.23}$$

- **Specificity**: is the True Negative Rate (TNR) of the model. The ability of the model to correctly classify the specific type of brain tumor is calculated a:

$$Specificity = \frac{TN}{TN+FP} \tag{3.24}$$

- **Precision:** The information that was retrieved is what the model needs to know. Precision is the ratio of the number of correctly identified tumor images (TP) to

the total number of pictures that are identified as tumors (TP + FP). The Precision increases as the FP decreases. In cases where the accuracy rate is higher, the model is more useful.

$$Precision = \frac{TP}{TP+FP}$$

(3.25)

- **F-Score:** It serves as a gauge of test accuracy and is the harmonic mean of Precision and Recall. The best and worst values for the F-score are 1 (100% accuracy and recall), and 0. F-Score is characterized as:

$$\mathrm{F} - \mathrm{Score} = 2 * \frac{\mathrm{P\,recision} * \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} = \frac{2\mathrm{TP}}{2\mathrm{T\,P} + \mathrm{F\,P} + \mathrm{FN}}$$

- **Receiver Operating Characteristic Curve (ROC):** is curve is a performance metric for classification problems. Essentially, it separates the "signal" from the "noise" by plotting the TPR against the FPR at different threshold levels. Each point on the roc curve represents sensitivity to the same signal stimulus, making it a complete indicator reflecting the continuous sensitivity and specificity variables.

In figure (3.5-2): "TPR" on the y-axis and False Positive Rate "FPR" on the x-axis fig where:

$$TPR = \frac{TP}{TP+FN}$$

(3.26)

$$FPR = \frac{FP}{FP+TN}$$

(3.27)

**The Area under Curve (AUC):** The area under the ROC may be used as a gauge to assess the classification method since it is obvious that when the FPR is relatively low and the TPR is relatively high, there will be a better model and a bigger area under the corresponding curve. As a result, the AUC stands for the areas under the ROC curve, or the model's capacity to categorize classes. Figure (3.5-2) shows AUC with the blue area and ROC curve.

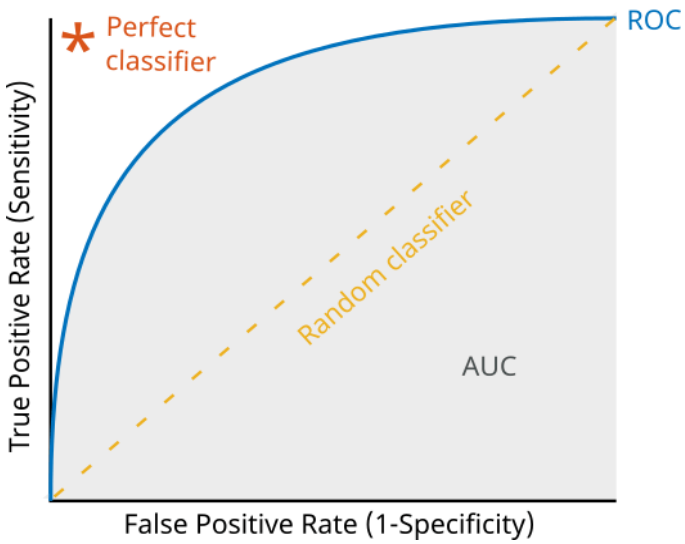



Figure 3.5-2: AUC and ROC

The usual statistic for models incorporating unbalanced data is the receiver operating characteristic (ROC) area under the curve (AUC) score. A ROC curve, based on sensitivity and specificity rates, depicts the TP rate vs. FP rate. Due to taking into account the sensitivity and specificity rates, the area under the curve, which estimates the area under this curve, is scale-invariant. As a result, how many instances or photos there are is no longer critical, and an ongoing evaluation of the model's general detection capabilities may be done.

# Chapter Four

# Experiments and Results

## 4.1 Experiments and Results

In this chapter, after addressing the deep learning models presented in the previous chapter, the global data set will be applied, and the application results will be reviewed and discussed. The testing process is divided into three phases. This is represented in the first stage by applying the global data set to each of the CNN, VGG16, VGG19, Xception, and ResNet50 DenseNet121 models, and the accuracy of each model will be measured. In addition to calculating each model's sensitivity, specificity, and accuracy, calculating the area under the curve (AUC) curves and receiver operating curves (ROC) for the best classification model. Based on the results of the previous measurements, the best model will be determined from the pre-trained models for the global data set classification. Second, based on the preceding, in the second experiment, the transfer learning strategy will be applied to the VGG16 model to improve the learning process of the pre-trained model to achieve higher accuracy results, represented by adding two layers of fully connected traditional neural networks and adjusting parameters. Based on the experiment's results, the model's accuracy will be measured, in addition to calculating the sensitivity, specificity, accuracy, Area Under Curve (AUC), and Receiver Operation Curves (ROC) curves of the improved model. Moreover, compare the results of previous models with the improved model and determine the best model in the classification process for the global data set. The previous two phases will include the application of data augmentation technology as the third phase on each of the

previous models and measuring the accuracy of each of the models, in addition to measuring the sensitivity and specificity of each model.

- **Computing Environment**

The Computing Environment in this work was done by HP Laptop 15-bs1xx, Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz 1.80 GHz, RAM 16.0 GB, HDD: 1 TB with Windows 10 Pro. For applying experiments Jupyter Notebook was used.

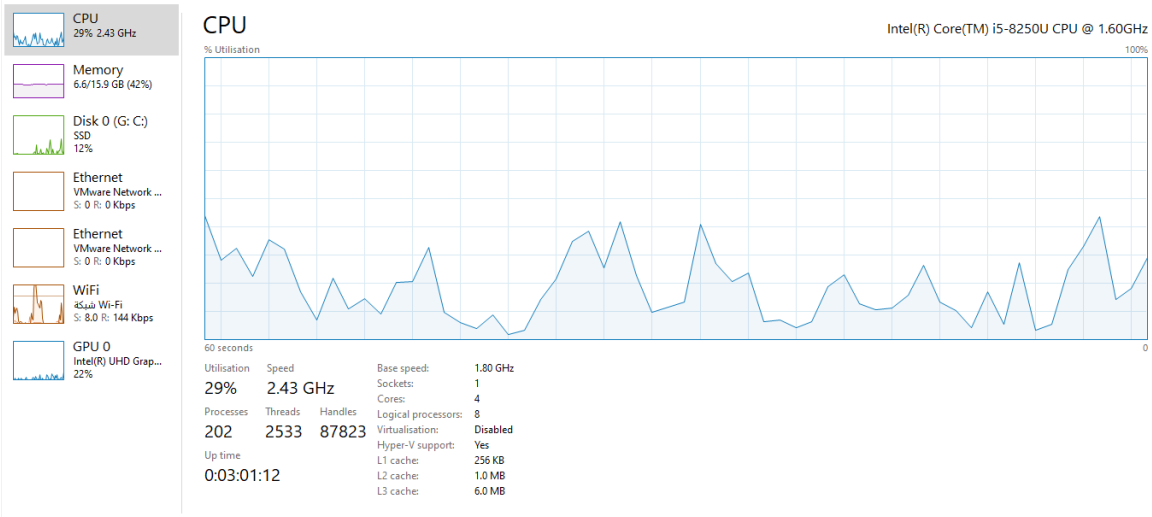



Figure 4.1-1: The performance of HP device Jupyter Notebook

## 4.2 Deep Learning Practical Experiment

### 4.2.1 Classification Results for Global Dataset

*First Experiment:* In this part, we apply five pre-trained deep learning algorithms to classify brain tumors on a global data set. They are VGG16, VGG19, Xception, ResNet50, and DenseNet121, in addition to building a CNN network. In this experiment, the data were divided into a test group and a training group, with a percentage of 70% for training the data, 2456 images, and 30% for the test 616 images, processed to evaluate the model. Table (4.2.1) also shows the results of applying the

mentioned models to the global data set. The table shows four performance measures: overall accuracy, sensitivity, specificity, and accuracy.

Table 4.2-1: Classification results for different model on all variable global dataset

| Model Type | Accuracy | Sensitivity | Specificity | precision | F-Score |
|---|---|---|---|---|---|
| CNN | 95.40% | 95% | 98.97% | 96.80% | 95% |
| VGG16 | 90.30% | 88.00% | 74.21% | 88.50% | 90% |
| VGG19 | 86.36 % | 85.10% | 95.05% | 85% | 85% |
| Xception, | 91.74% | 98.74% | 96.58% | 90.30% | 90% |
| ResNet50 | 86.85% | 86.85% | 93.42% | 85.85% | 85% |
| DenseNet121 | 93.18% | 92.33% | 96.59% | 93.18% | 92% |

In this experiment, we build a convolutional neural network consisting of 6 convolutional layers to classify brain tumors. We also use transfer learning as a classifier using pre-trained deep learning models by changing the softmax layer into three neurons to classify brain tumors, namely meningioma, glioma, and pituitary tumor. The results show that the CNN model obtained the highest test accuracy of 95.40% in tumor types. In applying the pre-trained deep learning models, the DenseNet121 model obtained a test accuracy of 93.18% in the classification process. The Xception model obtained an accuracy of 91.74% in the classification process. As for the VGG16 model, it obtained an accuracy value of 90.30%. The VGG19 and ResNet50 models obtained 86.36% and 86.85%, respectively.

Figure (4.2-1): Illustrates the Comparison between performance metrics for all Classification Models.

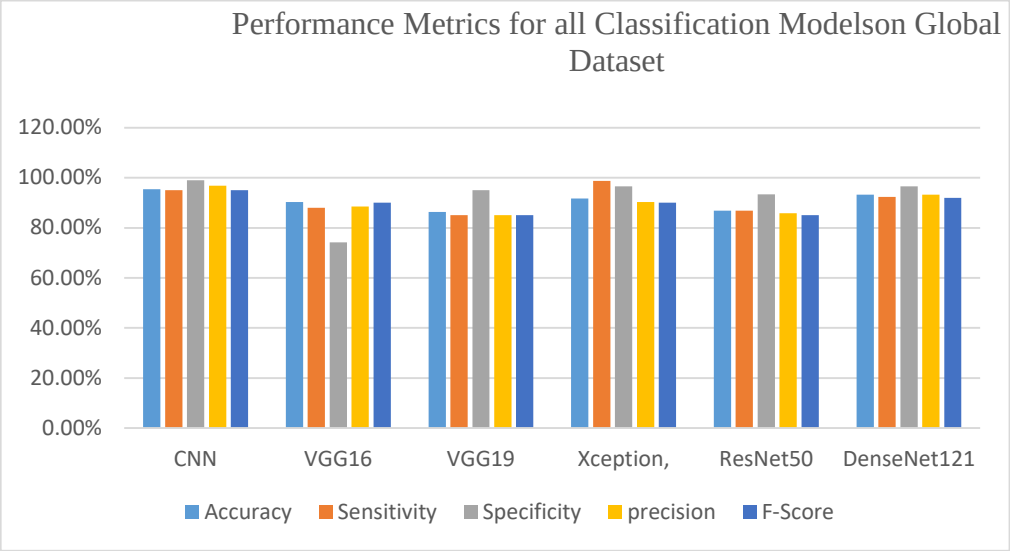



*Figure 4.2-1: Chart of Comparison between performance metrics for all Classification Models*

Table 4.2-2: Table summarizing of the overall accuracy of all classification models for the global dataset

| Model Type | Accuracy |
|---|---|
| CNN | 95.40% |
| VGG16 | 92.30% |
| VGG19 | 86.36% |
| Xception | 91.74% |
| ResNet50 | 86.85% |
| DenseNet121 | 93.18% |

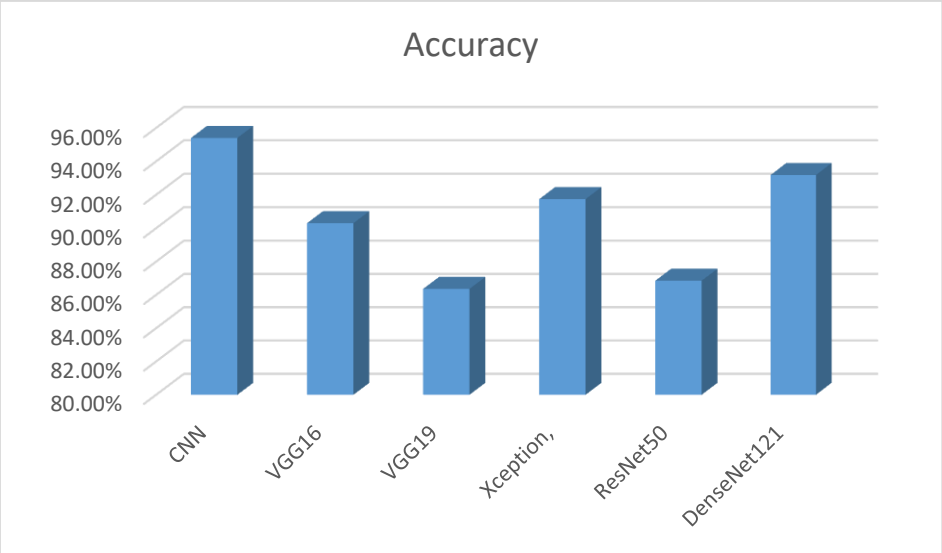



Figure 4.2-2: Chart of summary of the overall accuracy of all classification models for the global dataset

### 4.2.2 Results of Architecture CNN

*First Experiment:* Our trained CNN architecture consists of six convolutional layers with alternating batch normalization and max pooling layers and two fully connected dense layers (1024 x 512) that make up the classification layers. It achieved 95.45% accuracy, 95,45% sensitivity(recall), 95.58% specificity, and 100% of ROC AUC score by some calculations from these results based on these parameters (TP, TN, FP, and TN). After training the model, the results came as follows:

Figure (4.-2-3) shows the model training process curve. The figure shows the change in training accuracy and validation accuracy with an increasing number of epochs (time)

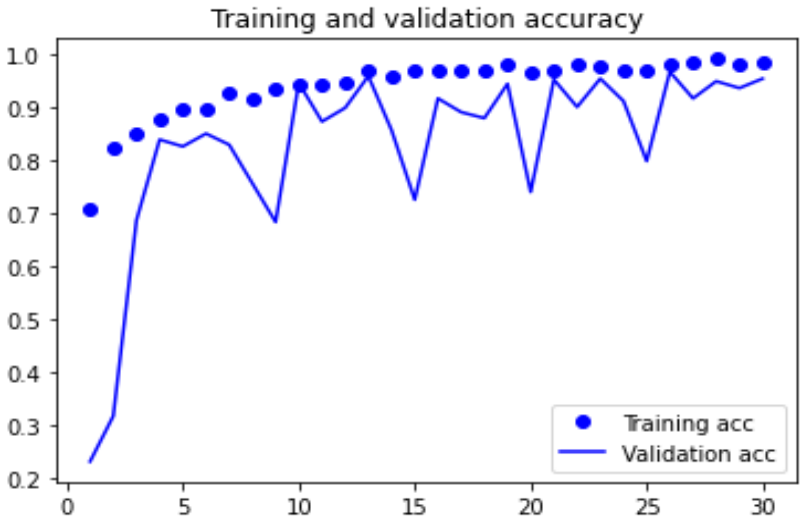



Figure 4.2-3: Performance learning curves for training & validation accuracy CNN model

As seen from the curve, the training process of the CNN grows, affecting the model's performance. That is, its performance improves with each learning process. However, as noted with time, the model does not improve its performance. That is, it has reached the stability stage, which means it cannot learn over time. Figure (4.2-4) shows the confusion matrix for the model dataset, where Glioma: 0, Meningioma: 1, and Pituitary tumor: 2.

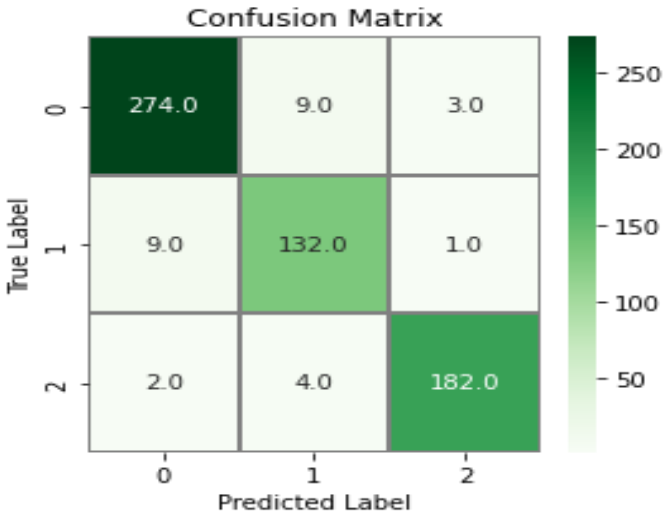



Figure 4.2-4 :Confusion matrix for CNN model

Table 4.2-3: shows a summary of the results of the CNN model

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.96 | 0.96 | 0.96 | 286 |
| 1 | 0.91 | 0.93 | 0.92 | 142 |
| 2 | 0.98 | 0.97 | 0.97 | 188 |
| accuracy | | | 0.95 | 616 |
| macro avg | 0.95 | 0.95 | 0.95 | 616 |
| weighted avg | 0.95 | 0.95 | 0.95 | 616 |

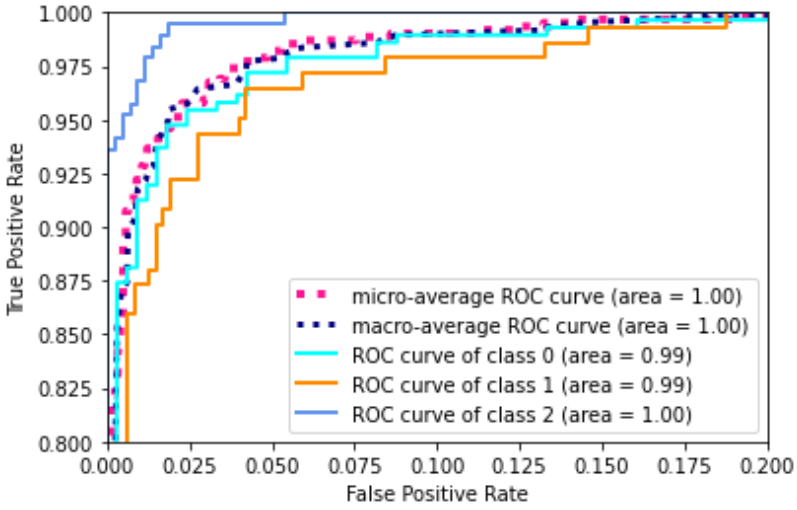



Figure 4.2-5: Receiver operating characteristic (ROC) to Multi-Class Validation Set CNN

Figure (4.2-5) represents connected lines of the AUC and ROC curves for CNN model results [best accuracy] and represents sensitivity for three classes (class 0 is Glioma, class 1 is Meningioma, and class 2 is Pituitary tumor); sensitivity was 95%. The specificity was 98.97%. Moreover, the precision was 96.80% for classes 0,1 and 2, respectively. These results were good results when the points were in the upper-left corner; here area under the curve occupies 100% of these curves for three classes, and the current classifier equals (0.99), (0.99), and (1.00) for class 0,1 and 2 respectively.

### 4.2.3 Results of Architecture VGG16

In this experiment, we use the Visual Geometry Group 16 (VGG-16) model to classify brain tumors based on a convolutional neural network (CNN) pre-trained on ImageNet. The experiment was carried out on Jupyter Notebook to write Python code using the Keras library. The data was divided into two sets, a test set, and a training set, and the input size is 244 * 224. Figure (4.2-6) shows the curve resulting from the model training process, which shows the learning and test validation process with an increase in the number of epochs (time).

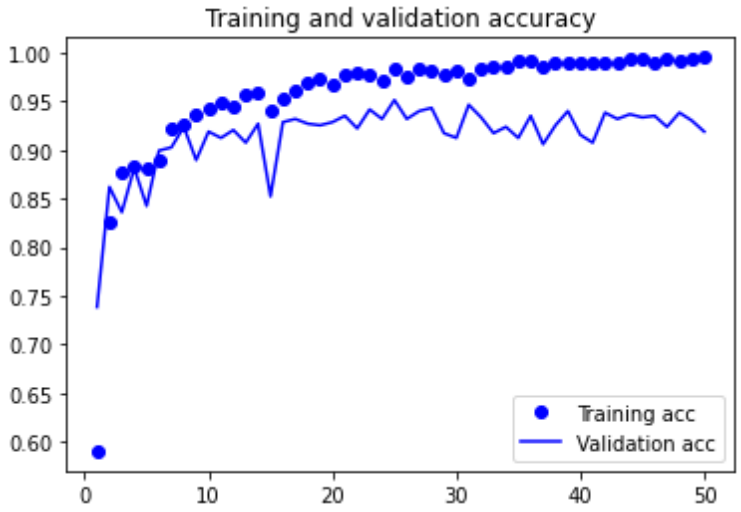



Figure 4.2-6: Performance learning curves for training & validation accuracy VGG16

By looking at the model, the accuracy of the model test increases with the increase in the training process, but with time the training process stops at a stage in which the learning process is proven. That is, there is no increase in model learning. The results show that the model gave a test accuracy of 96.27%.

Figure (4.2-7) shows the confusion matrix for the dataset of the model, where Glioma: 0, Meningioma: 1, and Pituitary tumor: 2.

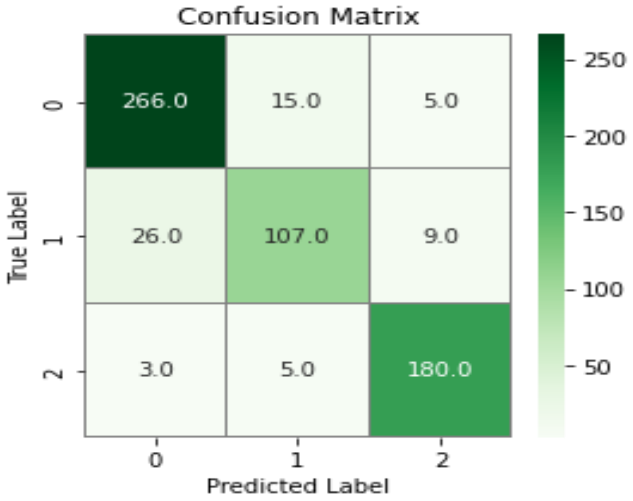



Figure 4.2-7: Confusion matrix for VGG16 model

Table 4.2-4: shows a summary of the results of the VGG16 model

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.90 | 0.93 | 0.92 | 286 |
| 1 | 0.84 | 0.75 | 0.80 | 142 |
| 2 | 0.93 | 0.96 | 0.94 | 188 |
| accuracy | | | 0.90 | 616 |
| macro avg | 0.89 | 0.88 | 0.88 | 616 |
| weighted avg | 0.90 | 0.90 | 0.90 | 616 |

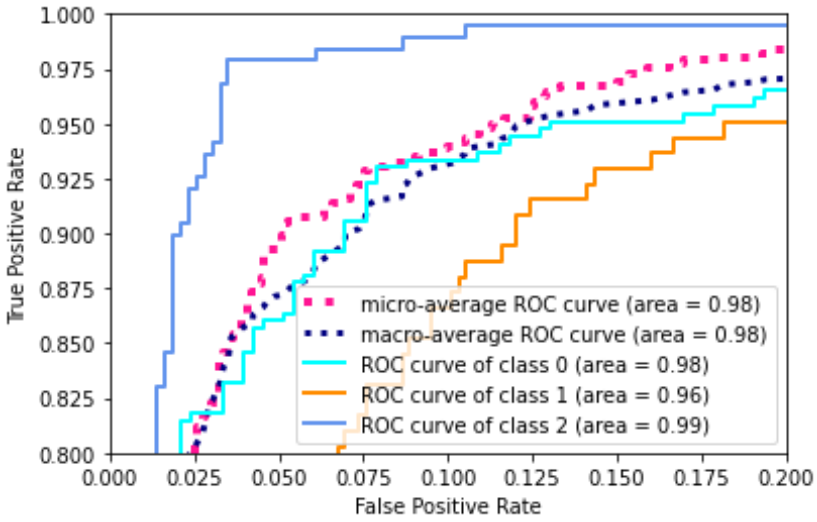



Figure 4.2-8: Receiver operating characteristic (ROC) to Multi-Class Validation Set VGG16

Figure (4.2-8) represents connected lines of the AUC and ROC curves for VGG16 model results [best accuracy], which represents sensitivity for three classes (class 0 is Glioma, class 1 is Meningioma, and class 2 is Pituitary tumor); the sensitivity was 88.00% the specificity was 74.21%. Moreover, the precision was 89.50% for classes 0,1 and 2, respectively. These results were good results when the points are in the upper-left corner; here area under the curve occupies 98% of these curves for three classes, and the current classifier equals (0.98), (0.96), and (0.99) for class 0,1 and 2 respectively.

### 4.2.4 Results of Architecture VGG19

The VGG19 model was used in this experiment, a pre-trained model consisting of an advanced CNN network. VGG19 is a deep form developed from the VGG16 model. As in the case of the previous models, the data was divided into two groups, training and testing. The training results on obtaining the model showed an accuracy of 94% in testing the model. Curve (4.2-9) shows the model training process on the dataset, which shows the increase in the learning process with each epoch (time).

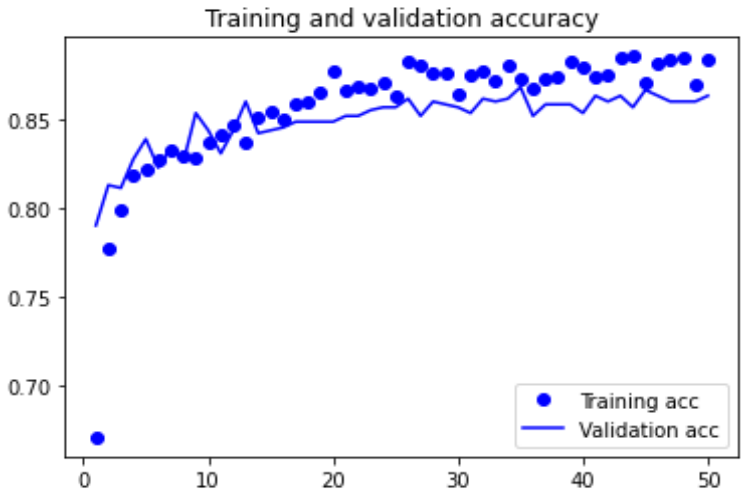



Figure 4.2-9: Performance learning curves for training & validation accuracy VGG19

Figure (4.2-10) shows the confusion matrix for the dataset of the model, where Glioma: 0, Meningioma: 1, and Pituitary tumor: 2.

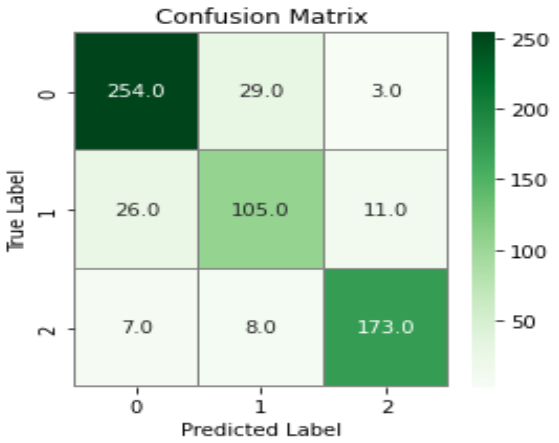



Figure 4.2-10: Confusion matrix for VGG-19 model

Table 4.2-5: shows a summary of the results of the VGG19 model

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.89 | 0.89 | 0.89 | 286 |
| 1 | 0.74 | 0.74 | 0.74 | 142 |
| 2 | 0.93 | 0.92 | 0.92 | 188 |
| accuracy | | | 0.86 | 616 |
| macro avg | 0.85 | 0.85 | 0.85 | 616 |
| weighted avg | 0.86 | 0.86 | 0.86 | 616 |

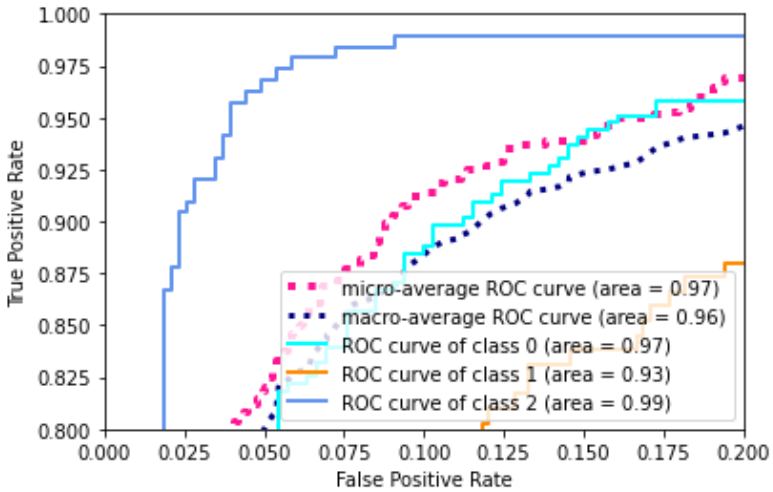



Figure 4.2-11: Receiver operating characteristic (ROC) to Multi-Class Validation Set VGG19

Figure (4.2-11) represents connected lines of the AUC and ROC curves for VGG19 model results [best accuracy], which represents sensitivity for three classes (class 0 is Glioma, class 1 is Meningioma, and class 2 is Pituitary tumor); the sensitivity was 85.00% the specificity was 95%. Moreover, the precision was 85% for classes 0,1 and 2, respectively. These results were good results when the points were in the upper-left corner; here area under the curve occupies 97% of these curves for three classes, and the current classifier equals (0.97), (0.93), and (0.99) for class 0,1 and 2 respectively.

### 4.2.5 Results of Architecture Xception

In this experiment, the Xception model was used, a convolutional neural network pre-trained on a database containing more than one million images. The model was used to train the model to classify tumors. As in the case of previous models, the data was divided into two groups, training and testing. The training results on obtaining the model showed an accuracy of 91% in testing the model. Curve (4.2-12) shows the model training process on the dataset, which shows the increase in the learning process with each epoch (time).

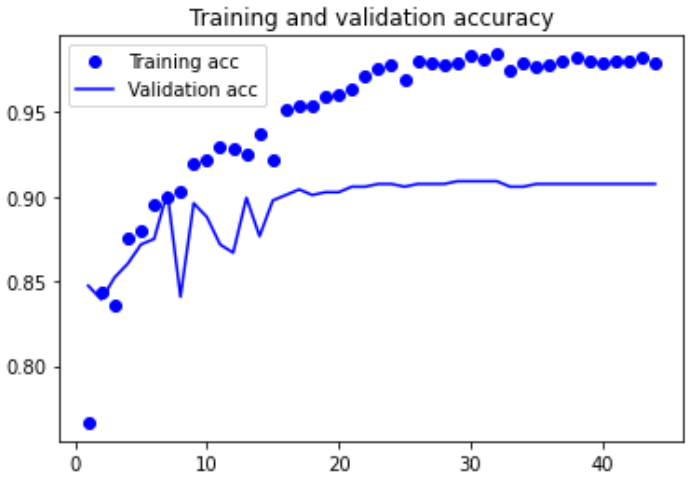



Figure 4.2-12: Performance learning curves for training & validation accuracy exception model

Figure (4.2-13) shows the confusion matrix for the dataset of the model, where Glioma: 0, Meningioma: 1, Pituitary tumor: 2

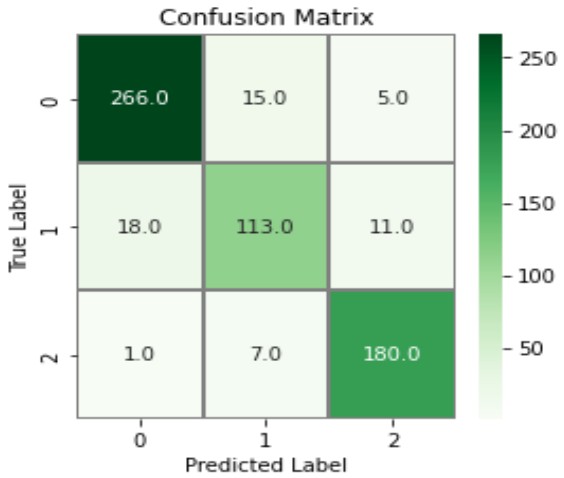



Figure 4.2-13: Confusion matrix for exception model

Table 4.2-6 : shows a summary of the results of the exception model

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.93 | 0.93 | 0.93 | 286 |
| 1 | 0.84 | 0.80 | 0.82 | 142 |
| 2 | 0.92 | 0.96 | 0.94 | 188 |
| accuracy | | | 0.91 | 616 |
| macro avg | 0.90 | 0.89 | 0.90 | 616 |
| weighted avg | 0.91 | 0.91 | 0.91 | 616 |

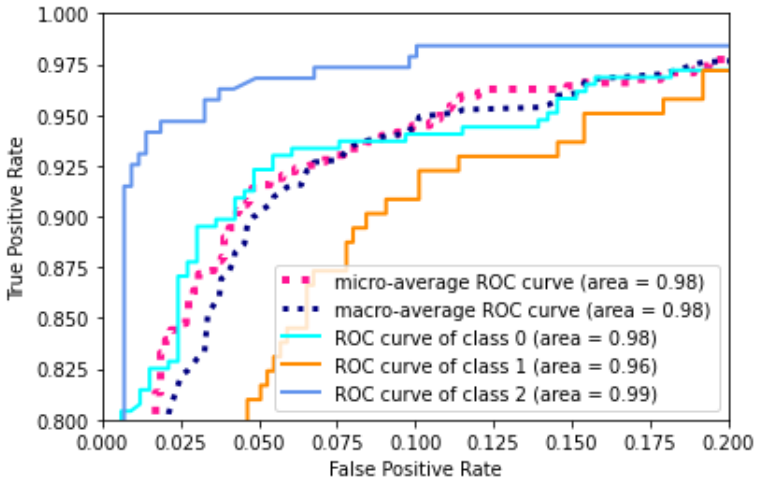



Figure 4.2-14: Receiver operating characteristic (ROC) to Multi-Class Validation Set exception model

Figure (4.2-14) represents connected lines of the AUC and ROC curves for Xception, model results [best accuracy], which represents sensitivity for three classes (class 0 is Glioma, class 1 is Meningioma, and class 2 is Pituitary tumor); the sensitivity was 98.00% the specificity was 96%. Moreover, the precision was 90% for classes 0, 1, and 2, respectively. These results were good results when the points were in the upper-left corner; here area under the curve occupies 98% of these curves for three classes, and the current classifier equals (0.98), (0.96), and (0.99) for class 0,1 and 2 respectively.

### 4.2.6 Results of Architecture ResNet50

In this experiment, the ResNet50 model was used, a 50-layer deep convolutional neural network pre-trained on a database containing more than one million images. The model was used to train the model to classify tumors. As in the case of the previous models, the data was divided into two groups, training and testing. The training results on obtaining the model showed an accuracy of 87% in testing the model. Curve (4.2-15) shows the model training process on the dataset, which shows the increase in the learning process with each epoch (time). The curve shows that the learning process stops at a level where the model stops learning.

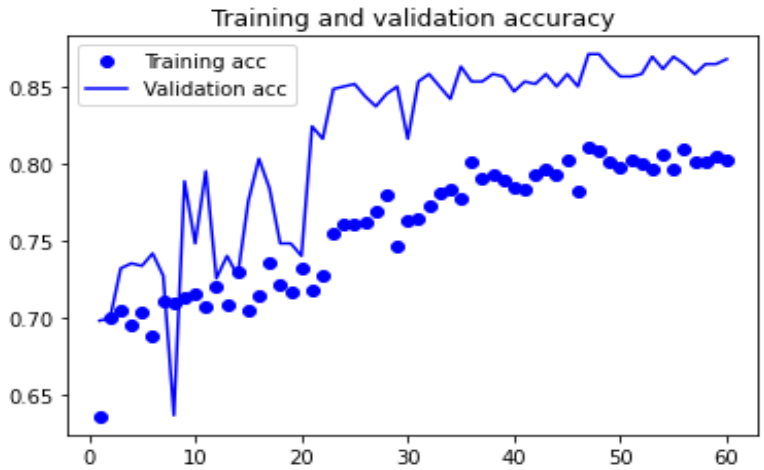



Figure 4.2-15: Performance learning curves for training & validation accuracy ResNet50 model

Figure (4.2-16) shows the confusion matrix for the dataset of the model, where Glioma: 0, Meningioma: 1, Pituitary tumor: 2

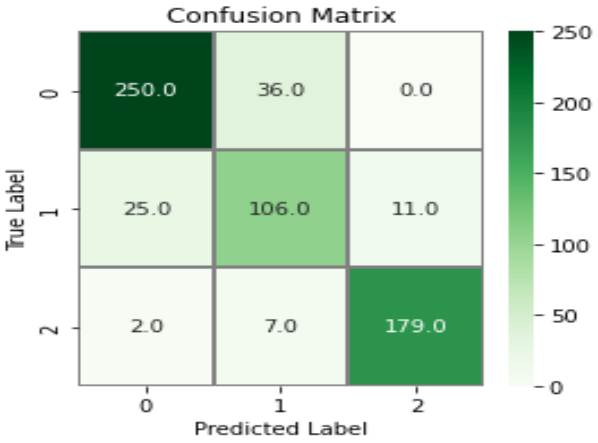



Figure 4.2-16: Confusion matrix for ResNet50 model

Table 4.2-7 shows a summary of the results of the ResNet50 model

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.90 | 0.87 | 0.89 | 286 |
| 1 | 0.71 | 0.75 | 0.73 | 142 |
| 2 | 0.94 | 0.95 | 0.95 | 188 |
| accuracy | | | 0.87 | 616 |
| macro avg | 0.85 | 0.86 | 0.85 | 616 |
| weighted avg | 0.87 | 0.87 | 0.87 | 616 |

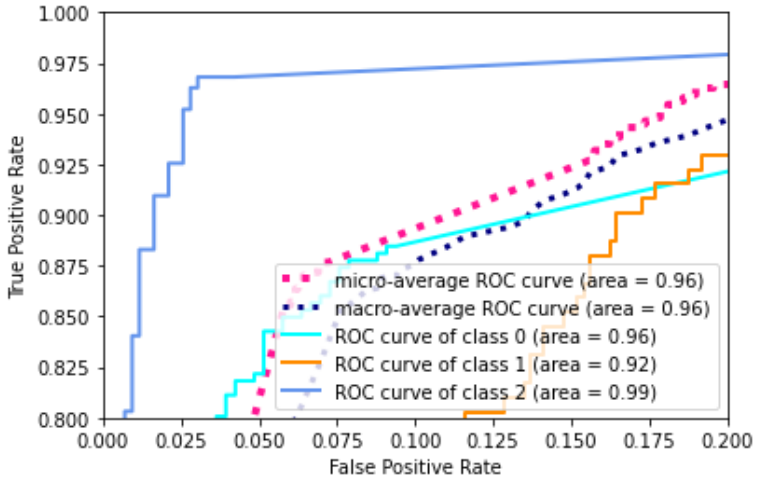



Figure 4.2-17: Receiver operating characteristic (ROC) to Multi-Class Validation Set ResNet50 model

Figure (4.2-17) represents connected lines of the AUC and ROC curves for ResNet50, model results [best accuracy], which represents sensitivity for three classes (class 0 is Glioma, class 1 is Meningioma, and class 2 is Pituitary tumor); the sensitivity was 86.00% the specificity was 93%. Moreover, the precision was 85% for classes 0,1 and 2, respectively. These results were good results when the points were in the upper-left corner; here area under the curve occupies 96% of these curves for three classes, and the current classifier equals (0.96), (0.92), and (0.99) for class 0,1 and 2 respectively.

### 4.2.7 Results of Architecture DenseNet121

In this experiment, the DenseNet121 model was used, a convolutional neural network with 120 convolutions, and was pre-trained on a database containing more than one million images. Since DenseNets requires fewer parameters and allows features to be reused, the model was used to train the model to classify tumors. As in the case of the previous models, the data was divided into two groups, training and testing. The training results on obtaining the model showed an accuracy of 93% in testing the model.

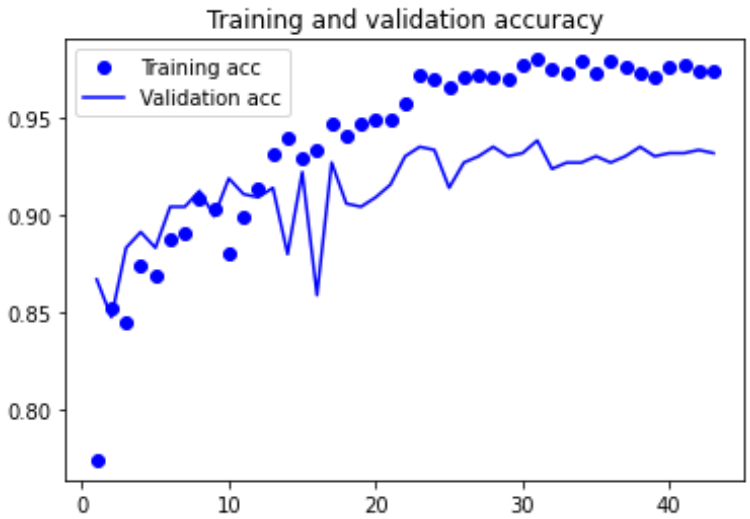



Figure 4.2-18: Performance learning curves for training & validation accuracy DenseNet121 model

Curve (4.2-18) shows the model training process on the dataset, which shows the increase in the learning process with each epoch (time). The curve shows the stability of the learning process at a level where the model cannot learn.

Figure (4.2-19) shows the confusion matrix for the dataset of the model, where Glioma: 0, Meningioma: 1, Pituitary tumor: 2

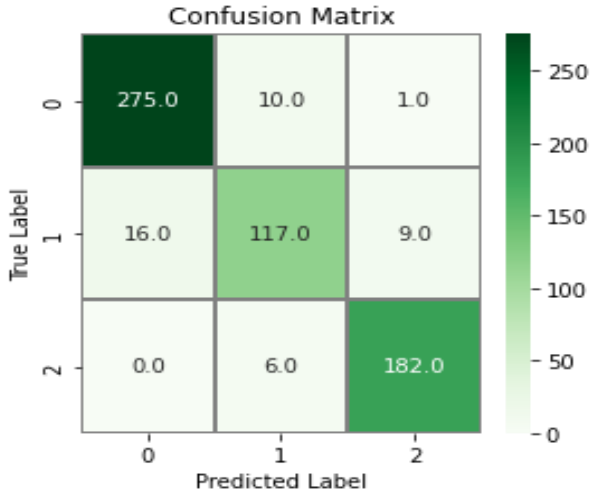



Figure 4.2-19: Confusion matrix for DenseNet121 model

Table 4.2-8: shows a summary of the results of the DenseNet121 model

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.95 | 0.96 | 0.95 | 286 |
| 1 | 0.88 | 0.82 | 0.85 | 142 |
| 2 | 0.95 | 0.97 | 0.96 | 188 |
| accuracy | | | 0.93 | 616 |
| macro avg | 0.92 | 0.92 | 0.92 | 616 |
| weighted avg | 0.93 | 0.93 | 0.93 | 616 |

Figure (4.2-20) represents connected lines of the AUC and ROC curves for DenseNet121, model results [best accuracy], which represents sensitivity for three classes (class 0 is Glioma, class 1 is Meningioma, and class 2 is Pituitary tumor); the

sensitivity was 92.00% the specificity was 96%. Moreover, the precision was 93% for classes 0,1 and 2, respectively. These results were good results when the points are in the upper-left corner; here area under the curve occupies 99% of these curves for three classes, and the current classifier equals (0.99), (0.98), and (1.00) for class 0,1 and 2 respectively.

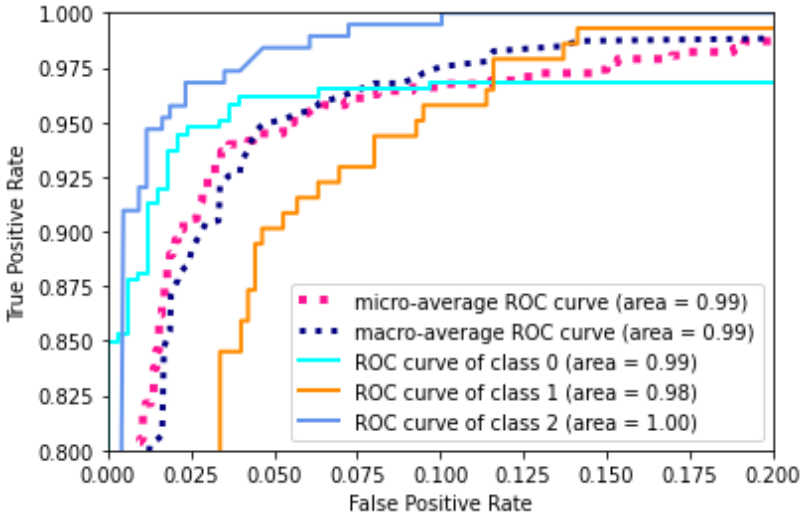



Figure 4.2-20: Receiver operating characteristic (ROC) to Multi-Class Validation Set DenseNet121 model

## 4.3 Transfer Learning the VGG16 Model

In this section, we use what is known as transfer learning. It is a method in which one of the pre-trained deep learning models is used on a task to solve a target task. We will use the VGG-16 model as a starting point for feature extraction to classify a brain tumor dataset. In this experiment, the model is fine-tuned, and its effect on the model's accuracy is determined. First, the VGG16 model is imported to create our input set. We will import the convolutional and aggregation layers and cut the fully connected layer called the upper model.

Our data set will be passed through layers to extract the convoluted visual features, i.e., the resulting data set will be 3D. It entails fine-tuning and freezing the pre-trained convolutional layers and leaving them untrainable. Otherwise, the new fully connected layer will be selected, which will be trained using backpropagation in addition to including dropout to reduce over-fitting.

Looking at the curve (4.3-1) of the model training process indicates the growth of the model training process with each increase in the epoch. That is, it notes the increase in the model experience. The model reaches learning constancy at a level where it stops learning.

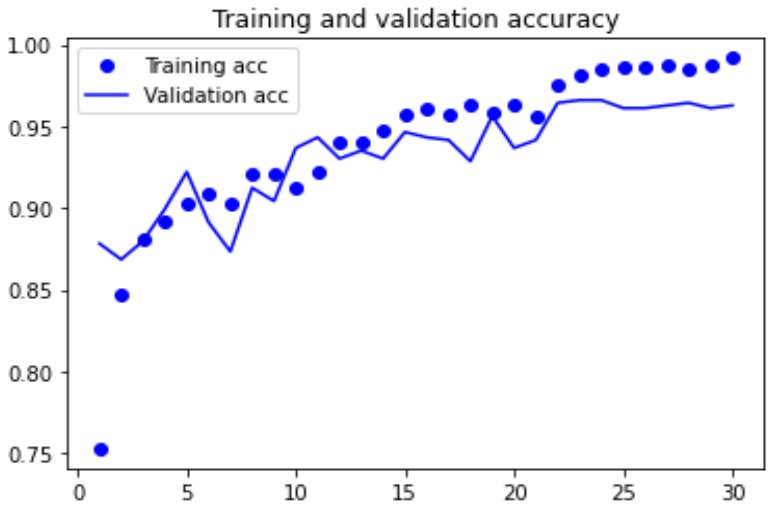



Figure 4.3-1: Performance learning curves for training & validation accuracy Transfer Learning the VGG16 Model

The curve results show that the model obtained an accuracy of 96% in the model testing process. Compared with the standard VGG16 model, an improvement in the accuracy of the model test was shown by 4%. That is, there was an improvement in training the model on the data set.

Table 4.3-1: Classification results for accuracy Transfer Learning the VGG16 Model on global dataset

| Model Type | Accuracy | Sensitivity | Specificity | precision | F-Score |
|---|---|---|---|---|---|
| VGG16 | 96.27% | 96.26% | 98.13% | 96.26% | 96.26% |

Figure (4.3-2) shows the confusion matrix for the dataset of the model, where Glioma: 0, Meningioma: 1, Pituitary tumor: 2

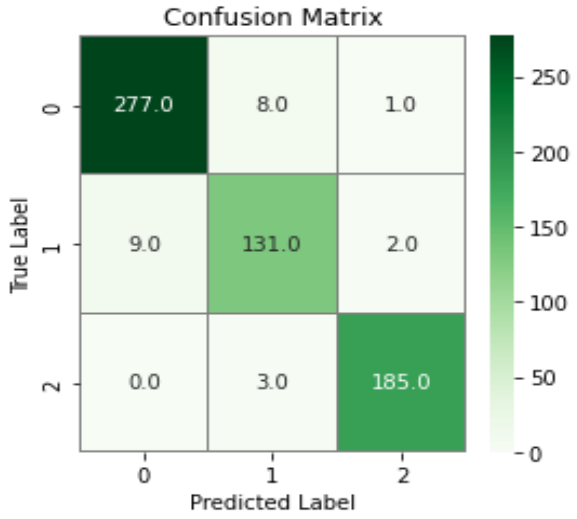



Figure 4.3-2: Confusion matrix for accuracy Transfer Learning the VGG16 Model

Table 4.3-2: shows a summary of the results of Transfer Learning the VGG16 model

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.97 | 0.97 | 0.97 | 286 |
| 1 | 0.92 | 0.92 | 0.92 | 142 |
| 2 | 0.98 | 0.98 | 0.98 | 188 |
| accuracy | | | 0.96 | 616 |
| macro avg | 0.96 | 0.96 | 0.96 | 616 |
| weighted avg | 0.96 | 0.96 | 0.96 | 616 |

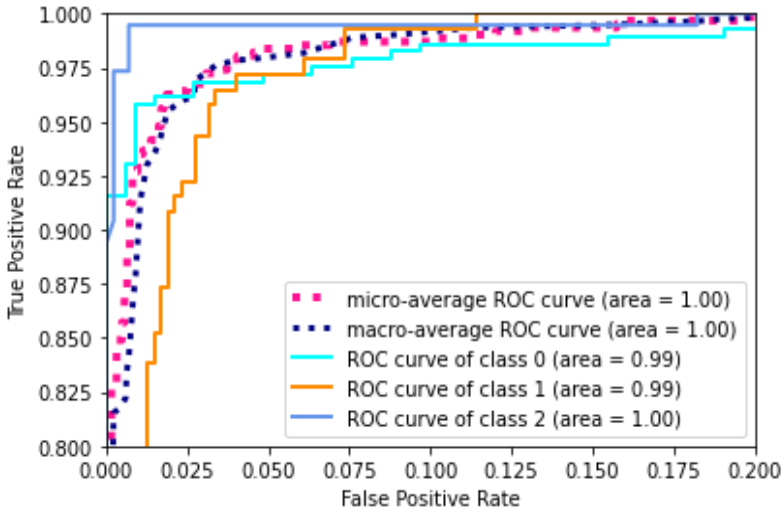



Figure 4.3-3: Receiver operating characteristic (ROC) to Multi-Class Validation Set

Transfer Learning the VGG16 Model

Figure (4.3-3) represents connected lines of the AUC and ROC curves for Transfer Learning the VGG16 Model, model results [best accuracy], which represents sensitivity for three classes (class 0 is Glioma, class 1 is Meningioma, and class 2 is Pituitary tumor); the sensitivity was 96.26% The specificity was 98.13%. And the precision of 96.26% for classes 0,1, and 2, respectively. These results were good results when the points were in the upper-left corner; here area under the curve occupies 100% of these curves for three classes, and the current classifier equals (0.99), (0.99), and (1.00) for class 0,1 and 2 respectively.

## 4.4 Transfer Learning the VGG16 Model and Data Augmentation Technique

In this stage, as the last stage in the process of improving the learning of the model, we will use the data augmentation technique due to the small number of data available to us, which is represented in 3064 images from 233 patients with three kinds of brain tumor: meningioma (708 slices), glioma (1426 slices), and pituitary tumor (930 slices).

On the other hand, deep learning requires a large amount of data, which is always difficult to provide. From here, we will use data augmentation technology to improve the diversity of data available to us. On the one hand, we will provide several data that fit the requirements of deep learning in a large amount of data it needs in the learning process. Here we will use a combination of online and offline data augmentation. We will use the default data augmentation via the Internet, where images are randomly taken from the training data and apply data augmentation techniques, and then train the model on the original data set and the default augmentation. Here we will use the technique of increasing the horizontal flip, which means flipping entire rows and columns of pixels in the image horizontally. In addition to the above, data augmentation technology will be used offline. In this case, the data will be increased before training the model and mixing it with the original data set. We will use change image contrast technology to increase the images, store the resulting images, and mix them with the original data. This method enables us to overcome our lack of data and adjust our data balance.

Table 4.4-1: Table summarizing of the overall accuracy of all classification models for the global dataset

| Model Type | Accuracy | Sensitivity | Specificity | precision | F-Score |
|---|---|---|---|---|---|
| CNN | 95.40% | 95% | 98.97% | 96.80% | 95% |
| VGG16 | 90.30% | 88.00% | 74.21% | 88.50% | 90% |
| VGG19 | 86.36% | 85.10% | 95.05% | 85% | 85% |
| Xception, | 91.74% | 98.74% | 96.58% | 90.30% | 90% |
| ResNet50 | 86.85% | 86.85% | 93.42% | 85.85% | 85% |
| DenseNet121 | 93.18% | 92.33% | 96.59% | 93.18% | 92% |
| VGG16 transfer learning | 96.27% | 96.26% | 98.13% | 96.26% | 96.26% |
| VGG16 transfer learning &data augmentation | 98.45% | 98.23% | 99.60% | 98.99% | 98.72% |

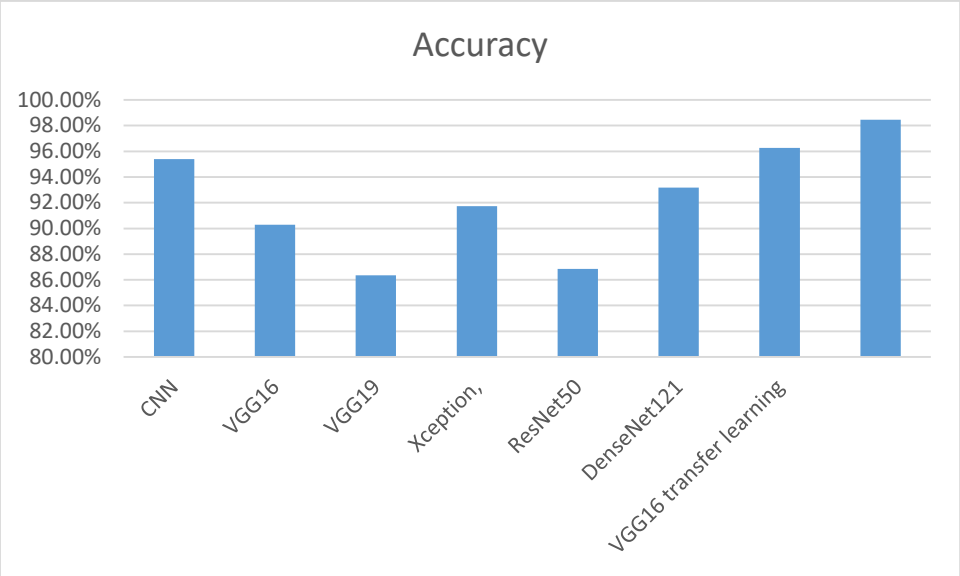



Figure 4.4-1: Chart of summary of the overall accuracy of all classification models for the global dataset

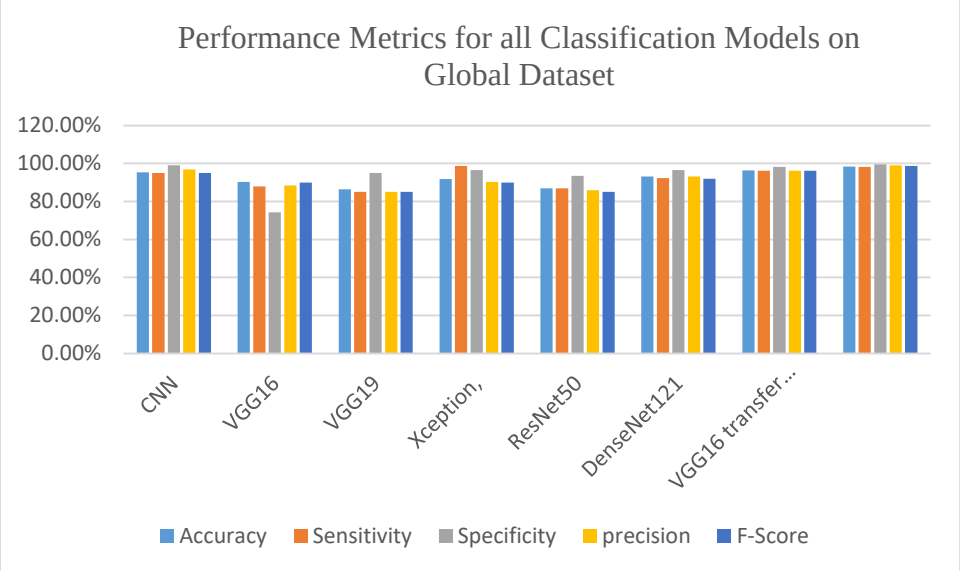
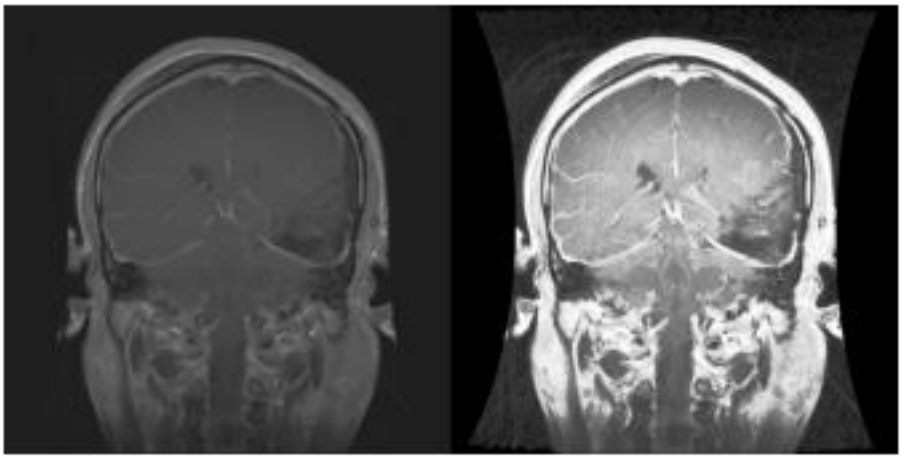
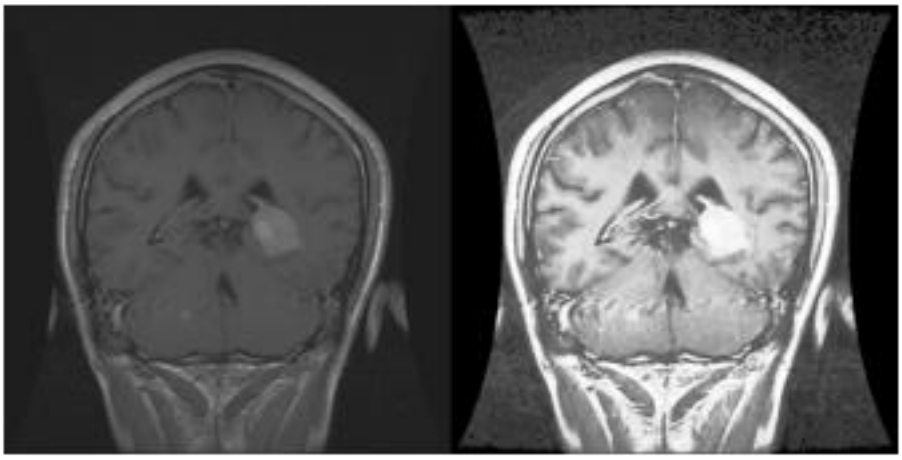



Figure 4.4-2: Chart of Comparison between performance metrics for all Classification Models

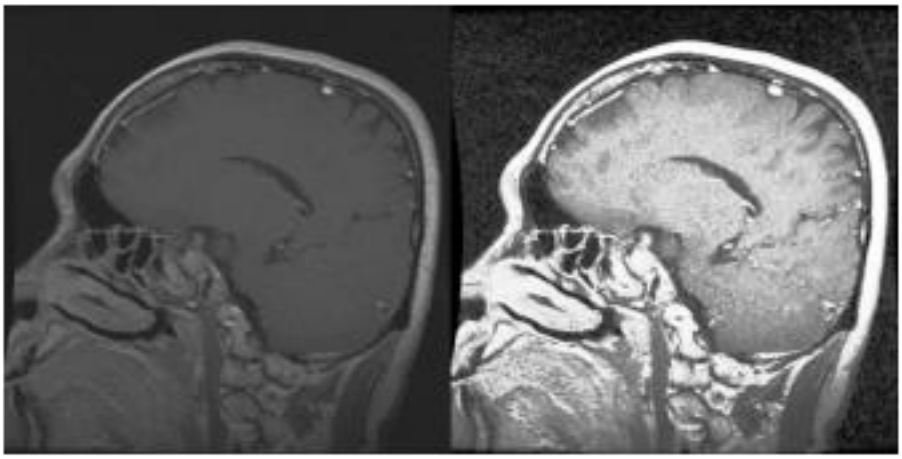
Figure (4.4-3) shows some of the images that applied the data augmentation technique using histogram equalization, which is a technique used in image processing to adjust the contrast using the image histogram.



(a) Original Image Glioma (b) Histogram equalization Image Glioma



(c) Original Image Meningioma (d) Histogram equalization Image Meningioma



(a) Original Image Pituitary tumor (b) Histogram equalization Image Pituitary tumor

Figure 4.4-3: sample images that applied the data augmentation technique

Figure (4.4-4) shows the model training process curve. The figure shows the training and validation accuracy change with an increasing number of epochs (time). As shown by the curve, significant progress is shown in training the model and the validation process of the model. We conclude that increasing the data affected the performance of the presented model and improved its performance. That is, it reached a stage of stability in the training process and stability in the model validation process, which gave an accuracy of 98%. In contrast, the model's results without increasing the data were 96%. Figure (4-29) shows the confusion matrix for the dataset of the model, where Glioma: 0, Meningioma: 1, and Pituitary tumor: 2.

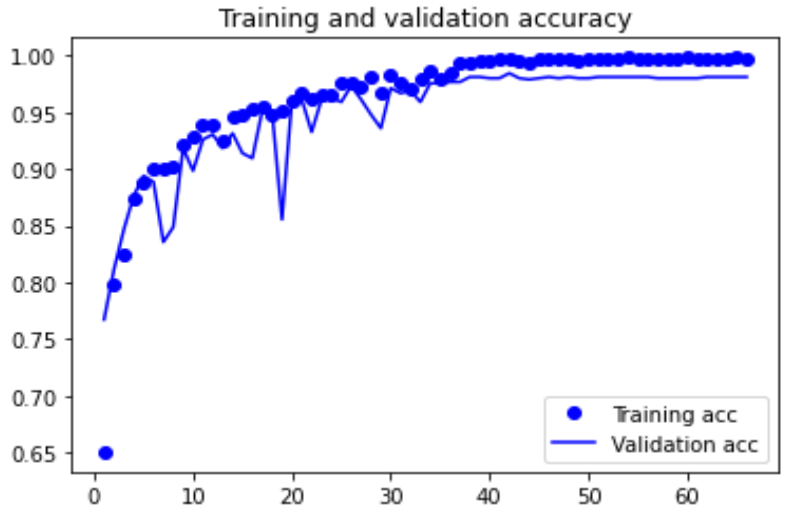



Figure 4.4-4: Performance learning curves for training & validation accuracy for Transfer Learning the VGG16 Model after data augmentation technique

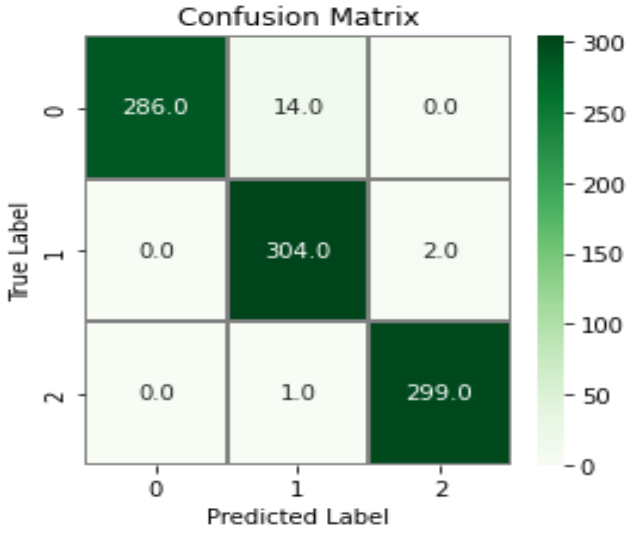



Figure 4.4-5: Confusion matrix for Transfer Learning the VGG16 Model after data augmentation Technique

Table 4.4-2: Classification results for accuracy Transfer Learning the VGG16 Model after data augmentation

| Model Type | Accuracy | Sensitivity | Specificity | precision | F-Score |
|---|---|---|---|---|---|
| VGG16 | 98.45% | 98.23% | 99.60% | 98.99% | 98.72% |

Table 4.4-3: Summary of result for the VGG16 Model after the data augmentation technique

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 1.00 | 0.95 | 0.98 | 300 |
| 1 | 0.95 | 0.99 | 0.97 | 306 |
| 2 | 0.99 | 1.00 | 1.00 | 300 |
| accuracy | | | 0.98 | 906 |
| macro avg | 0.98 | 0.98 | 0.98 | 906 |
| weighted avg | 0.98 | 0.98 | 0.98 | 906 |

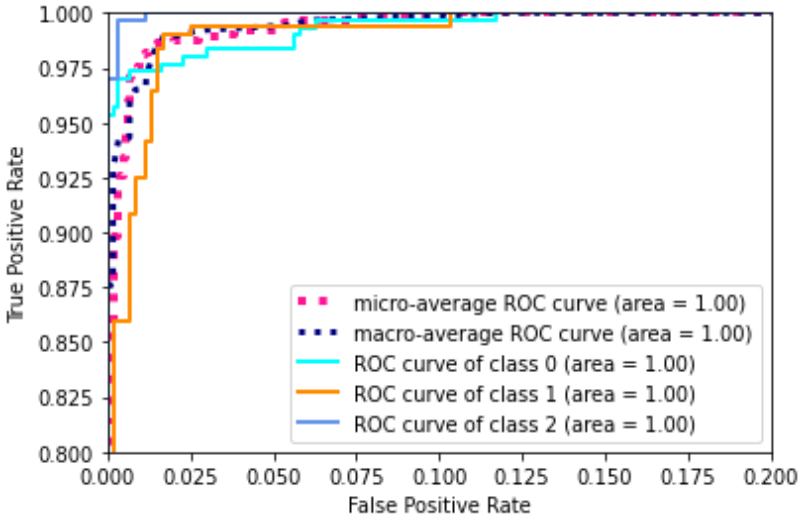



Figure 4.4-6: Receiver operating characteristic (ROC) to Multi-Class Validation Set for Transfer Learning the VGG16 Model after the data augmentation technique

Figure (4.4-6) represents connected lines of the AUC and ROC curves for Transfer Learning the VGG16 Model after data augmentation technique, model results [best accuracy], which represents sensitivity for three classes (class 0 is Glioma, class 1 is Meningioma, and class 2 is Pituitary tumor); the sensitivity was 98.23% The specificity was 99.0%. And the precision of 98.99% for classes 0,1, and 2, respectively. These results were good results when the points were in the upper-left corner; here area under

the curve occupies 100% of these curves for three classes, and the current classifier equals (1.00), (1.00), and (1.00) for class 0,1 and 2 respectively.

## 4.5 Challenges and Limitation

We encountered some limitations in our thesis work, which we will mention in this section. One of the hurdles we faced was the need for a sufficient amount of images for the MRI dataset of brain tumors. This reduces the accuracy. The models need a long training time on the graphical data, in addition to the need for a processor with powerful specifications, and this was not available, as all the training operations were done on a personal computer (PC) with simple specifications that we mentioned in the introduction to this chapter. Better results can be obtained if a computer with higher specifications is used. To generalize this research, there is a need to collect brain tumor MRI data to overcome the need for more data to train these models on accurate data.

On the other hand, during the proposal submission period, we planned to use a local dataset, but unfortunately, we did not find this data locally.

# Chapter Five

# Conclusion and Future Work

## 5.1 Conclusion

The fatal illness, brain tumors, has an extremely low life expectancy in its most severe form. Brain tumor misdiagnosis will lead to ineffective medical intervention and lower patient survival rates. Getting an effective treatment plan to cure and prolong the lives of people with brain tumor illness depends on making an accurate diagnosis of the condition (Rehman et al., 2020).

The intracranial tumor, commonly referred to as a brain tumor, is an unwelcome, abnormal mass of tissue in which many cells proliferate and expand out of control. Various tumors might have a diverse characteristic, which means they can also have different treatment methods, the most famous of which are meningioma, glioma, and Pituitary tumor (Rehman et al., 2020). Convolutional neural networks and computer-aided tumor detection systems have significantly advanced machine learning and offered success stories. In contrast to standard previous neural network layers, deep convolutional layers automatically extract significant and reliable features from the input space (Abd El Kader et al., 2021). Therefore, adequately trained convolutional neural networks may considerably aid physicians in making a better diagnosis or aid surgeons in thoroughly studying a surgical plan in less time, ultimately leading to improved health care.

In this work, we covered deep learning (DL) models. We took advantage of their advantages in classification and diagnosis in the medical field, especially in their ability

to achieve qualitative and quantitative results from traditional image processing techniques so that they can make identification and evaluation based on these systems more accessible and more transparent when compared to other methods. On the other hand, Medical data such as MRI images can be used to train different deep learning models and build expert systems from them in classifying brain tumors; These enrich physicians' diagnoses of brain tumors and give them a new perspective.

In this thesis, we conduct research utilizing five convolutional neural network architectures (VGG16, VGG19, Xception, ResNet50, DenseNet121) and create a CNN network to classify types of brain tumors meningioma, glioma, and pituitary. It uses the MRI slices of the brain tumor dataset from Figshare. We also introduce the concept of learning transfer approaches, such as fine-tuning and freezing and apply it to deep learning models to study the effect on the model training process to increase the accuracy in the model testing process. The data augmentation techniques are applied to the MRI slices to generalize the findings, expand the dataset samples, and reduce the chance of overfitting. The fine-tuned VGG16 architecture achieved the best accuracy, up to 98.45%, in the planned process of diagnosing brain tumors to be achieved.

## 5.2 Future Work and Recommendations

In future work, we will create a local data set of MRI images of brain tumors, use other models to diagnose brain tumors, and make comparisons in terms of accuracy in more than one model. Also, it is possible to work on using pre-trained deep learning models as a feature extractor and using machine learning models in classifying brain tumors and making comparisons in accuracy in the diagnosis process. This work can also be developed, and an easy-to-use user interface application can be created so that it is an

easy way to help medical specialists in the process of diagnosing brain tumors. In addition to working on establishing a joint medical base for all medical laboratories in Palestine to form a digital medical repository to collect the most significant number of diseases and use them in practical research in machine learning and deep learning.

On the other hand, a future model that depends on the deep learning model could also consider the tumor's severity. It is known that tumors can be classified as malignant or benign, and they have a classification where they can be type 1, type 2, type 3, or type 4, depending on the severity. Thus, one could have a model that classifies the type of tumor and its severity.

## Bibliography

Abd El Kader, I., Xu, G., Shuai, Z., Saminu, S., Javaid, I., & Salim Ahmad, I. (2021). Differential deep convolutional neural network model for brain tumor classification. *Brain Sciences*, 11(3), 352.

Abd-Ellah, M. K., Awad, A. I., Khalaf, A. A., & Hamed, H. F. (2018). Two-phase multi-model automatic brain tumour diagnosis system from magnetic resonance images using convolutional neural networks. *EURASIP Journal on Image and Video Processing*, (1), 1-10.

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. *In 2017 international conference on engineering and technology (ICET)*, (pp. 1-6).

Alqazzaz, S., Sun, X., Yang, X., & Nokes, L. (2019). Automated brain tumor segmentation on multi-modal MR image using SegNet. *Computational Visual Media*, 5, 209-219..

Al-Tamimi, M. & Sulong, G. (2014). Tumor Brain Detection through MR Images: A Review of Literature. *Journal of Theoretical and Applied Information Technology*. 62.

Amsaveni, V., Singh, N. A., & Dheeba, J. (2013). Computer aided detection of tumor in MRI brain images using cascaded correlation neural network. *In IET Chennai Fourth International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2013) (pp. 527-532). IET*.

Anaraki, A. K., Ayati, M., & Kazemi, F. (2019). Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. *biocybernetics and biomedical engineering*, 39(1), 63-74

Aslam, A., Khan, E., & Beg, M. S. (2015). Improved edge detection algorithm for brain tumor segmentation. *Procedia Computer Science*, 58, 430-437.

Basaveswara, S. K. (2019). CNN Architectures, a Deep-dive, *Medium*, [Online]. Available: https://towardsdatascience.com/cnn-architectures-a-deep-dive-a99441d18049 (Accessed: 22-Mar-2020).

Battineni, G., Sagaro, G. G., Chinatalapudi, N., & Amenta, F. (2020). Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis. *Journal of personalized medicine*, 10(2), 21. https://doi.org/10.3390 /jpm10020021

Bauer, S., Wiest, R., Nolte, L. P., & Reyes, M. (2013). A survey of MRI-based medical image analysis for brain tumor studies. *Physics in medicine and biology*, 58(13), R97–R129. https://doi.org/10.1088/0031-9155/58/13/R97

Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., & Litjens, G.. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. Jama, 318(22), 2199-2210.

Bengio, Y., & LeCun, Y. (2007). Scaling learning algorithms towards AI. *Large-scale kernel machines*, 34(5), 1-41.

Benzebouchi, N. E., Azizi, N., & Ayadi, K. (2019). A computer-aided diagnosis system for breast cancer using deep convolutional neural networks. *In Computational intelligence in data mining* (pp. 583-593). Springer, Singapore.

Bhanothu, Y., Kamalakannan, A., & Rajamanickam, G. (2020). Detection and classification of brain tumor in MRI images using deep convolutional network. *In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 248-252)*. IEEE.

Cancer.Net, (2021). Brain Tumor: Diagnosis, [Online]. Available: https://www.cancer.net/cancer-types/brain-tumor/diagnosis.

Chauhan, S., & Sharma, E. N. (2014). Brain tumor detection and segmentation using artificial neural network techniques. *International Journal of Engineering, Science and Technology (IJEST)*, 3(8), 288-93.

Cheng, J. (2017). Brain Tumor Dataset. *figshare*. Dataset. https://doi.org/10.6084/m9.figshare.1512427.v5

Cheng, J., Huang, W., Cao, S., Yang, R., Yang, W., Yun, Z. & Feng, Q. (2015). Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PloS one*, 10(10), e0140381.

Cheng, J., Yang, W., Huang, M., Huang, W., Jiang, J., Zhou, Y., & Chen, W. (2016). Retrieval of brain tumors by adaptive spatial pooling and fisher vector representation. *PloS one*, 11(6), e0157112.

Cheng, P. (2014). Deep Learning with Convolutional Neural Networks for Radiologic Image Classification. Keck school of Medicine of USC- USA.

Chollet, F. (2016). Building powerful image classification models using very little data. *Keras Blog*, 5.

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, (pp. 1251-1258).

D. Y. Weerakkody and A. P. F. Gaillard, "No Title." [Online]. Available: https://radiopaedia.org/articles/pituitary-tumours

Dahab, D. A., Ghoniemy, S. S., & Selim, G. M. (2012). Automated brain tumor detection and identification using image processing and probabilistic neural network techniques. *International journal of image processing and visual communication*, 1(2), 1-8.

DeAngelis L. M. (2001). Brain tumors. The New England journal of medicine, 344(2), 114–123. https://doi.org/10.1056/NEJM200101113440207

Deb, D., & Roy, S. (2021). Brain tumor detection based on hybrid deep neural network in MRI by adaptive squirrel search optimization. *Multimedia tools and applications*, 80(2), 2621-2645.

Deepak, S., & Ameer, P. M. (2021). Brain tumour classification using siamese neural network and neighbourhood analysis in embedded feature space. *International Journal of Imaging Systems and Technology*, 31(3), 1655-1669.

Demir, F. (2022). Deep autoencoder-based automated brain tumor detection from MRI data. In *Artificial Intelligence-Based Brain-Computer Interface* (pp. 317-351). Academic Press.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. *In 2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255).

Díaz-Pernas, F. J., Martínez-Zarzuela, M., Antón-Rodríguez, M., & González-Ortega, D. (2021). A deep learning approach for brain tumor classification and segmentation using a multiscale convolutional neural network. *In Healthcare* (Vol. 9, No. 2, p. 153). Multidisciplinary Digital Publishing Institute.

Fayyadh, S. B., & Ibrahim, A. A. (2020). Brain Tumor Detection and Classification Using CNN Algorithm and Deep Learning Techniques. In 2020 *International Conference on Advanced Science and Engineering (ICOASE) (pp. 157-161).*

Ghosh, A., Sufian, A., Sultana, F., Chakrabarti, A., & De, D. (2020). Fundamental concepts of convolutional neural network. *In Recent trends and advances in artificial intelligence and Internet of Things (pp. 519-567)*. Springer, Cham.

Giri, C., Goodwin, M., & Oppedal, K. (2020). Deep 3D Convolution Neural Network for Alzheimer's Detection. *In International Conference on Machine Learning, Optimization, and Data Science* (pp. 347-358). Springer, Cham.

Glass-Macenka, L. Hays, A. Varner, E. and Wen, P. (2013). Tumor Types: Understanding Brain Tumors," in Frankly Speaking about Cancer series, Cancer Support Community.

Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In Proceedings of the fourteenth international conference on artificial intelligence and statistics (pp. 315-323). *JMLR Workshop and Conference Proceedings*.

Gopal, N. & Karnan, M.. (2011). Diagnose brain tumor through MRI using image processing clustering algorithms such as Fuzzy C Means along with intelligent optimization techniques. *2010 IEEE International Conference on Computational Intelligence and Computing Research*, ICCIC 2010. 1 - 4. 10.1109/ICCIC.20 10.5705890.

Gordillo, N., Montseny, E., & Sobrevilla, P. (2013). State of the art survey on MRI brain tumor segmentation. *Magnetic resonance imaging*, 31(8), 1426–1438. https://doi.org/10.1016/j.mri.2013.05.002

Goyal, M. Goyal, R. Lall, B. (2019). Learning Activation Functions: A New Paradigm of Understanding Neural Networks. arXiv 2019, arXiv:1906.09529.

Gupta, R. K., Bharti, S., Kunhare, N., Sahu, Y., & Pathik, N. (2022). Brain Tumor Detection and Classification Using Cycle Generative Adversarial Networks. *Interdisciplinary Sciences: Computational Life Sciences*, 1-18

Gurunathan, A., & Krishnan, B. (2021). Detection and diagnosis of brain tumors using deep learning convolutional neural networks. *International Journal of Imaging Systems and Technology*, 31(3), 1174-1184.

Hao, H., Wang, Q., Li, P., & Zhang, L. (2016). Evaluation of ground distances and features in EMD-based GMM matching for texture classification. *Pattern Recognition*, 57, 152-163.

Hu, J., Zhou, J., & Wu, X. (2016). Non-local MRI denoising using random sampling. *Magnetic resonance imaging*, 34(7), 990–999. https://doi.org/10.1016/j .mri.2016.04.008

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *In Proceedings of the IEEE conference on computer vision and pattern recognition*, (pp. 4700-4708).

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, 195(1), 215-243.

Ismael, M. (2018). Hybrid Model - Statistical Features and Deep Neural Network for Brain Tumor Classification in MRI Images. Dissertations. 3291.

Jain, S., & Mishra, S. (2013). ANN Approach Based On Back Propagation Network and Probabilistic Neural Network to Classify Brain Cancer. International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 3.

Ju, C., Bibaut, A., & van der Laan, M. (2018). The relative performance of ensemble methods with deep convolutional neural networks for image classification. *Journal of Applied Statistics*, 45(15), 2800-2818.

Kaur, T., & Gandhi, T. K. (2020). Deep convolutional neural networks with transfer learning for automated brain image classification. *Machine Vision and Applications*, 31(3), 20.

Kesav, N., & Jibukumar, M. G. (2022). Efficient and low complex architecture for detection and classification of Brain Tumor using RCNN with Two Channel CNN. *Journal of King Saud University-Computer and Information Sciences*, 34(8), 6229-6242.

Khamparia, A., Singh, P. K., Rani, P., Samanta, D., Khanna, A., & Bhushan, B. (2021). An internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning. T*ransactions on Emerging Telecommunications Technologies*, 32(7), e3963.

Kim, K. G. (2016). Book Review: Deep Learning. *Healthcare Informatics Research*, 22(4), 351. https://doi.org/10.4258/hir.2016.22.4.351

Kong, Y., Deng, Y., & Dai, Q. (2014). Discriminative clustering and feature selection for brain MRI segmentation. *IEEE Signal Processing Letters*, 22(5), 573-577.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Levin, G. (2022). Image Processing and Computer Vision, ofBook (work in progress), https://openframeworks.cc/ofBook/chapters/image_processing_comput er_vision.html

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005

Louis, D. N., Ohgaki, H., Wiestler, O. D., Cavenee, W. K., Burger, P. C., Jouvet, A., Scheithauer, B. W., & Kleihues, P. (2007). The 2007 WHO classification of tumours of the central nervous system. Acta neuropathologica, 114(2), 97–109. https://doi.org/10.1007/s00401-007-0243-4.

Markoulidakis, I., Rallis, I., Georgoulas, I., Kopsiaftis, G., Doulamis, A., & Doulamis, N. (2021). Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. *Technologies*, 9(4), 81.

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M. A., Arbel, T., Avants, B. B., Ayache, N., Buendia, P., Collins, D. L., Cordier, N., Corso, J. Van Leemput, K. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE transactions on medical imaging*, 34(10), 1993–2024. https://doi.org/10.1109/TMI.2014.2377694

Mukhopadhyay, S. (2018). Deep learning and neural networks. *In Advanced data analytics using python* (pp. 99-119). Apress, Berkeley, CA..

Mukti, I. Z., & Biswas, D. (2019). He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Munira, H. & Islam, M. (2022). Hybrid Deep Learning Models for Multi-classification of Tumour from Brain MRI. *Journal of Information Systems Engineering and Business Intelligence*. 8. 162-174. 10.20473/jisebi.8.2.162-174.

Murphy, J. (2016). An overview of convolutional neural network architectures for deep learning. *Microway Inc*, 1-22.

Nayak, D. R., Padhy, N., Mallick, P. K., Zymbler, M., & Kumar, S. (2022). Brain Tumor Classification Using Dense Efficient-Net. *Axioms*, 11(1), 34.

Özyurt, F. (2020). Efficient deep feature selection for remote sensing image recognition with fused deep learning architectures. *The Journal of Supercomputing*, 76(11), 8413-8431.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.

Pan, Y., Huang, W., Lin, Z., Zhu, W., Zhou, J., Wong, J., & Ding, Z. (2015). Brain tumor grading based on neural networks and convolutional neural networks. *In 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 699-702).

Patterson, J., & Gibson, A. (2017). Deep learning: A practitioner's approach. O'Reilly Media, Inc.

Polat, Ö., & Güngen, C. (2021). Classification of brain tumors from MR images using deep transfer learning. *The Journal of Supercomputing*, 77(7), 7236-7252.

Pravitasari, A. A., Iriawan, N., Almuhayar, M., Azmi, T., Irhamah, I., Fithriasari, K. & Ferriastuti, W. (2020). UNet-VGG16 with transfer learning for MRI-based brain tumor

segmentation. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 18(3), 1310-1318.

Raja, S. S. (2019). Deep learning based image classification and abnormalities analysis of MRI brain images. *In 2019 TEQIP III Sponsored International Conference on Microwave Integrated Circuits, Photonics and Wireless Networks (IMICPW)*, (pp. 427-431).

Ranjbarzadeh, R., Bagherian Kasgari, A., Jafarzadeh Ghoushchi, S., Anari, S., Naseri, M., & Bendechache, M. (2021). Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. *Scientific Reports*, 11(1), 1-17.

Ravì, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2016). Deep learning for health informatics. *IEEE journal of biomedical and health informatics*, 21(1), 4-21.

Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 29(9), 2352-2449.

Rehman, A., Naz, S., Razzak, M. I., Akram, F., & Imran, M. (2020). A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circuits, Systems, and Signal Processing*, 39(2), 757-775.

Rguibi, Z. Hajami, A. & Zitouni, D. (2020). Brain tumor segmentation by using Convolutional Networks. 10.13140/RG.2.2.33510.14405.

Sadad, T., Rehman, A., Munir, A., Saba, T., Tariq, U., Ayesha, N., & Abbasi, R. (2021). Brain tumor detection and multi-classification using advanced deep learning techniques. *Microscopy Research and Technique*, 84(6), 1296-1308.

Sazzad, T. S., Ahmmed, K. T., Hoque, M. U., & Rahman, M. (2019). Development of automated brain tumor identification using MRI images. In 2019 *International conference on electrical, computer and communication engineering (ECCE)* (pp. 1-4).

Shaha, M., & Pawar, M. (2018). Transfer learning for image classification. *In 2018 second international conference on electronics, communication and aerospace technology (ICECA)*, (pp. 656-660). IEEE.

Shaha, M., & Pawar, M. (2018). Transfer learning for image classification. *In 2018 second international conference on electronics, communication and aerospace technology (ICECA)*, (pp. 656-660).

Shu, M. (2019). Deep learning for image classification on very small datasets using transfer learning. Creating. *Components*, Jan.

Siar, M., & Teshnehlab, M. (2022). A combination of feature extraction methods and deep learning for brain tumour classification. *IET Image Processing*, 16(2), 416-441.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Soltaninejad, M., Yang, G., Lambrou, T., Allinson, N., Jones, T. L., Barrick, T. R. & Ye, X. (2017). Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI. *International journal of computer assisted radiology and surgery*, 12, 183-203.

Subramanian, M., Lv, N. P., & VE, S. (2022). Hyperparameter optimization for transfer learning of VGG16 for disease identification in corn leaves using Bayesian optimization. *Big Data*, 10(3), 215-229.

Sultonov, F., Park, J. H., Yun, S., Lim, D. W., & Kang, J. M. (2022). Mixer U-Net: An Improved Automatic Road Extraction from UAV Imagery. *Applied Sciences*, 12(4), 1953.

Tahamid, A. (2020). Tomato leaf disease detection using Resnet-50 and MobileNet Architecture (Doctoral dissertation, Brac University).

Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning?. *IEEE transactions on medical imaging*, 35(5), 1299-1312.

Teare, P., Fishman, M., Benzaquen, O., Toledano, E., & Elnekave, E. (2017). Malignancy detection on mammography using dual deep convolutional neural networks and genetically discovered false color input enhancement. *Journal of digital imaging*, 30(4), 499-505.

Thompson P.M., Moussai J., Zohoori S., Goldkorn A., Khan A.A., Mega M.S., Small G.W., Cummings J.L., Toga A.W., (1998). Cortical variability and asymmetry in normal aging and Alzheimer's disease. Cereb Cortex. 8(6), 492-509, 1998.

Tom M., Rolf B., Ole D.L., and Mark L.R., (1998). Brain tumor invasion: biological, clinical, and therapeutic considerations, Copyright 1998 by Wiley-Liss, Inc.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine*, 25(1), 44–56. https://doi.org/10.1038/s 41591-018-0300-7

Turing, A. (1950). Computing machinery and intelligence. Mind, Vol. 16, Issue 236, 433-460.

Vasilev, I., Slater, D., Spacagna, G., Roelants, P., & Zocca, V. (2019). Python Deep Learning: Exploring deep learning techniques and neural network architectures with Pytorch, Keras, and TensorFlow. *Packt Publishing Ltd*.

Ying, Y., Su, J., Shan, P., Miao, L., Wang, X., & Peng, S. (2019). Rectified exponential units for convolutional neural networks. *IEEE Access*, 7, 101633-101640.

Yokota, H., Goto, M., Bamba, C., Kiba, M., & Yamada, K. (2017). Reading efficiency can be improved by minor modification of assigned duties; a pilot study on a small team of general radiologists. *Japanese journal of radiology*, 35(5), 262–268. https://doi.org/10.1007/s11604-017-0629-8

Yuan, J., & Wang, J. (2018). Compressive sensing based on L1 and Hessian regularizations for MRI denoising. *Magnetic resonance imaging*, 51, 79–86. https://doi.org/10.1016/j.mri.2018.04.015

Zabihollahy, F., Ukwatta, E., & Schieda, N. (2021). Computer-aided diagnosis of renal masses. *In State of the Art in Neural Networks and their Applications* (pp. 179-195). Academic Press.

Zhou, H., & Sun, Q. (2020). Research on Principle and Application of Convolutional Neural Networks. *In IOP Conference Series: Earth and Environmental Science*, (Vol. 440, No. 4, p. 042055). IOP Publishing.

# الملخص

عادة ما يتم تشخيص أورام الدماغ من خلال اختبارات التصوير، حيث يعتبر التصوير بالرنين المغناطيسي (MRI) هو الأكثر شيوعًا لتشخيص أورام الدماغ. يتم حقن صبغة من خلال أحد أوردة الذراع أثناء التصوير بالرنين المغناطيسي. حيث يساعد ذلك الطبيب في التشخيص وتقييم الورم. تكمن أهمية التشخيص الدقيق للورم في تحديد خطة العلاج اللازمة لمحاربة الورم والقضاء عليه دون التأثير سلبًا على صحة المريض. لا يوجد شك في خبرة الطبيب في تشخيص الورم، ولكن في بعض الأحيان يكون هناك بعض الأخطاء في تصنيف نوع الورم التي تؤثر سلباً على صحة المريض بسبب صعوبة التمييز بين هذه الأورام من خلال الصور الطبية التي توضح خصائص الأورام.

تسمح خصائص الأورام التي تظهر في التصوير بالرنين المغناطيسي بتحديد نوع الورم، سواء كان الورم حميدًا أم خبيثًا أو حتى كيفية تطويره. بناءً على ذلك، فإن تحديد هذه الخصائص يلعب دورًا مهمًا في تحديد خطة العلاج من قبل الفريق الطبي. أثبت الذكاء الاصطناعي نجاحه في العديد من المجالات، لا سيما التعلم الآلي وتقنيات التعلم العميق لا سيما في المجال الطبي. لذلك تعتمد هذه الدراسة على استخدام خوارزميات التعلم العميق لتصنيف أورام الدماغ. يتم استخدام مجموعة بيانات من صور التصوير بالرنين المغناطيسي الطبية التي تم جمعها من المرضى السابقين الذين تم تشخيص إصابتهم بورم في المخ والاستفادة منها في بناء نظام باستخدام تقنيات التعلم العميق (DL) مثل الشبكات العصبية التلافيفية (CNN)\، ( Visual Geometry Group-16 Model (VGG16)وVisual Geometry Group-19 Model (VGG19) ...إلخ. تساهم هذه التقنيات بشكل كبير ودقيق في مساعدة الأطباء على تشخيص نوع الورم بدقة عالية.

يعتمد البحث على مجموعة البيانات الخاصة بمجموعة من أنواع الأورام المتوفرة عالميًا على موقع figshare الإلكتروني. المرحلة الأولى من الدراسة تكون حول مجموعة من نماذج التعلم العميق المدربة مسبقًا على أساس CNN و VGG16 و VGG19 و Xception وResNet50 و DenseNet121، لتحديد الأنماط في مجموعة بيانات ورم الدماغ واستخراج الميزات لعملية التصنيف إضافة إلى بناء شبكة (CNN) لتصنيف الأورام. أظهرت نتائج دقة

التصنيف لكل من CNN دقة 95.40٪ ، ولكل من نماذج التعلم العميق، كانت نتائج دقة التصنيف على التوالي 90.30٪ ، 86.36٪ ، 91.74٪ ، 86.85٪ ، 93.18٪.

في المرحلة الثانية، تم تطبيق مفهوم التعلم الانتقالي على نموذج VGG16 في عملية استخراج الميزات من مجموعة البيانات لتحسين نتائج دقة تصنيف أورام المخ، وقد ظهر التحسن بنسبة 96.27٪. في المرحلة الثالثة، تم تطبيق مفهوم تعزيز البيانات على مجموعة البيانات لتحسين عملية التدريب النموذجية. تم استخدام تقنيتين في ذلك: معادلة الرسم البياني، وتقنية القلب الأفقي لزيادة عملية تدريب النموذج على المزيد من الأنماط لتجنب التداخل وموازنة البيانات لتحسين الدقة.أظهرت نتائج الدقة في عملية اختبار النموذج تحسنًا في العملية التدريبية بنسبة 98.45٪.