Democratic and Popular Algerian Republic
Ministry of Higher Education and Scientific Research



Yahia Fares University of Medea
Faculty of Sciences
Department of Mathematics and Computer Science

Final Master Project

Submitted for Obtaining The Master's Degree in Computer Science

Speciality: Applied Artificial Intelligence

Detection of brain cancer from MRI images using Deep Learning

---

Prepared and Presented by:
Sadik Benyahia

Supervisor:
Dr. Chergui Nadjah

2025-2026

Dedication

Alhamdulillah, the Most Gracious, the Most Merciful, who granted me the strength, patience, and clarity of mind to carry out this work. It is by His grace and guidance that this journey has been brought to a successful completion.

I would like to express my sincere gratitude to my supervisor for her valuable guidance, continuous support, and encouragement throughout this work.

I am deeply grateful to my father and mother for their endless love, prayers, sacrifices, and unwavering support. Special thanks also go to my wife for her patience, understanding, and encouragement during this journey.

I would like to dedicate my heartfelt thanks to my unborn son, who has been a source of hope and motivation. Finally, I extend my sincere appreciation to all my friends for their kindness, support, and encouragement.

Abstract

The diagnosis of brain tumors from Magnetic Resonance Imaging (MRI) images is a challenging clinical task, due to the visual similarity of the tumor subtypes and the variability of the medical imaging data. Deep learning has demonstrated great potential in automating this process, but most of the existing frameworks treat detection and classification as a one-stage task, which may result in inter-class confusion and hinder the diagnostic reliability. This work proposes an automatic two-stage hybrid deep learning framework for brain tumor analysis from MRI images to overcome these limitations. In the first stage, a hybrid CNN-VGG19 model performs binary classification to differentiate normal brain scans from tumor-positive cases. In the second stage, the detected tumors are classified into three clinically relevant subtypes, glioma, meningioma, and pituitary tumor, by a fine-tuned ResNet50 model incorporating dual pooling and progressive fine-tuning strategy. All images are pre-processed by CLAHE-based contrast enhancement and data augmentation before training to improve the input quality and generalization of the model. Experimental evaluation on two publicly available MRI datasets shows that the proposed framework achieves 99% accuracy in Stage 1 and 99% accuracy in Stage 2 with consistently high precision, recall and F1-score across all classes. The two-stage architecture reduces the inter-class confusion and provides a clinically interpretable decision making process, making it a promising computer-aided diagnosis tool to assist radiologists in brain tumor detection and classification.

Keywords: brain tumor, detection cancer, Artificial Intelligence, Machine Learning, MRI, deep learning, CNN, transfer learning, classification, ResNet50, VGG19.

Résumé

Le diagnostic des tumeurs cérébrales à partir d'images d'imagerie par résonance magnétique (IRM) constitue un défi clinique majeur, en raison de la similitude visuelle entre les différents sous-types de tumeurs et de la variabilité des données d'imagerie médicale. L'apprentissage profond a démontré un grand potentiel pour automatiser ce processus, mais la plupart des cadres existants traitent la détection et la classification comme une tâche en une seule étape, ce qui peut entraîner une confusion entre les classes et nuire à la fiabilité du diagnostic. Ce travail propose un cadre d'apprentissage profond hybride automatique en deux étapes pour l'analyse des tumeurs cérébrales à partir d'images IRM afin de surmonter ces limites. Dans la première étape, un modèle hybride CNN-VGG19 effectue une classification binaire pour différencier les scans cérébraux normaux des cas positifs pour une tumeur. Au cours de la deuxième étape, les tumeurs détectées sont classées en trois sous-types cliniquement pertinents (gliome, méningiome et tumeur hypophysaire) par un modèle ResNet50 finement ajusté intégrant une stratégie de double regroupement et d'ajustement progressif. Toutes les images sont prétraitées par amélioration du contraste basée sur CLAHE et augmentation des données avant l'entraînement afin d'améliorer la qualité des entrées et la généralisation du modèle. Une évaluation expérimentale sur deux ensembles de données IRM accessibles au public montre que le cadre proposé atteint une précision de 99% à l'étape 1 et de 99% à l'étape 2, avec une précision, un rappel et un score F1 élevés et constants pour toutes les classes. Cette architecture en deux étapes réduit la confusion entre les classes et offre un processus décisionnel cliniquement interprétable, ce qui en fait un outil de diagnostic assisté par ordinateur prometteur pour aider les radiologues à détecter et à classer les tumeurs cérébrales.

Mots-clés: tumeur cérébrale, détection du cancer, intelligence artificielle, apprentissage automatique, IRM, apprentissage profond, CNN, apprentissage par transfert, classification, ResNet50, VGG19

الملخص

يُعد تشخيص أورام الدماغ من خلال صور التصوير بالرنين المغناطيسي (MRI) مهمة سريرية صعبة، وذلك بسبب التشابه البصري بين أنواع الأورام المختلفة وتباين بيانات التصوير الطبي. وقد أظهر التعلم العميق إمكانات كبيرة في أتمتة هذه العملية، لكن معظم الأطر الحالية تعامل الكشف والتصنيف كمهمة من مرحلة واحدة، مما قد يؤدي إلى الخلط بين الفئات ويعيق موثوقية التشخيص. يقترح هذا العمل إطارًا هجينًا للتعلم العميق من مرحلتين لتحليل أورام الدماغ من صور التصوير بالرنين المغناطيسي للتغلب على هذه القيود. المرحلة الأولى هي مشكلة تصنيف ثنائية بين فحوصات الدماغ الطبيعية والحالات الإيجابية للأورام باستخدام نموذج هجين CNN-VGG19. في المرحلة الثانية، يتم تصنيف الأورام المكتشفة إلى ثلاثة أنواع فرعية ذات صلة سريرية، وهي الورم الدبقي، والورم السحائي، وورم الغدة النخامية، وذلك بواسطة نموذج ResNet50 مُحسّن بدقة يتضمن استراتيجية التجميع المزدوج والتحسين التدريجي. تتم معالجة جميع الصور مسبقًا عن طريق تحسين التباين باستخدام CLAHE وزيادة البيانات قبل التدريب لتحسين جودة المدخلات وتعميم النموذج. يُظهر التقييم التجريبي على مجموعتي بيانات التصوير بالرنين المغناطيسي المتاحتين للجمهور أن الإطار المقترح يحقق دقة بنسبة 99% في المرحلة الأولى و99% في المرحلة الثانية مع دقة واسترجاع ودرجة F1 عالية باستمرار عبر جميع الفئات. تقلل البنية المكونة من مرحلتين من الخلط بين الفئات وتوفر عملية اتخاذ قرار قابلة للتفسير سريريًا، مما يجعلها أداة تشخيص مدعومة بالحاسوب واعدة لمساعدة أخصائيي الأشعة في الكشف عن أورام الدماغ وتصنيفها.

الكلمات المفتاحية: الذكاء الإصطناعي, تحليل ورم في الدماغ, السرطان, الكشف عن السرطان, التعلم الآلي, ResNet50, VGG19, التعلم بالنقل, التصنيف.

Contents

Contents i

List of Tables iv

List of Figures v

List of Abbreviations vii

General Introduction 1
Problem Statement 2
Research Objectives 2

1 The Human Brain And Central Nervous System 5
1.1 Introduction 5
1.2 Brain Tumors 5
1.2.1 Definition and general classification (benign/malignant, primary/secondary) 6
1.2.2 Main Tumor Types Considered in This Work . . . 7
1.3 Medical Imaging of the Brain 9
1.3.1 Overview of Brain Imaging Modalities (CT, PET, MRI) 9
1.3.2 Role of MRI Compared to Other Modalities 9
1.4 Brain Magnetic Resonance Imaging (MRI) 9
1.4.1 Basic Principle of MRI 10
1.4.2 MRI Sequences for Brain Tumor Imaging 10
1.4.3 Specific Characteristics of Brain MRI Images . . . 10
1.4.4 Common MRI Artefacts 11
1.5 Conclusion 13

2 Deep Learning 15
2.1 Introduction 15

2.2 Machine learning 15
2.3 Deep learning 15
2.3.1 How Does Deep Learning Work? 16
2.4 Deep neural network architectures 17
2.4.1 Recurrent neural network (RNN, recurrent neural network) 17
2.4.2 Convolutional Neural Networks (CNN) 18
2.5 Different layers of CNN 19
2.5.1 Convolutional layers 19
2.5.2 Pooling layer 20
2.5.3 ReLU Activation Function 21
2.5.4 Fully Connected Layer 22
2.5.5 Flatten Layer 22
2.6 Different Types of CNN Models 23
2.6.1 AlexNet: 23
2.6.2 Resnet 24
2.6.3 VGG 24
2.7 Transfer learning 25
2.8 Deep learning applications 26
2.9 Related work 26
2.9.1 Binary Classification 26
2.9.2 Multi-class Classification 27
2.10 Conclusion 28

3 Proposed Approach 30
3.1 Introduction 30
3.2 Proposed Model 30
3.2.1 Datasets 32
3.2.2 Data Preprocessing and Augmentation 33
3.2.3 Stage 1 Model Architecture 35
3.2.4 Stage 2 Model Architecture 39
3.3 Conclusion 42

4 Results & Discussion 44
4.1 Introduction 44
4.2 Evaluation Metrics of Models 44
4.3 Split Dataset 45
4.4 Software Environment 46
4.4.1 Deep Learning Framework 46

4.4.2 Development Environment 46
4.4.3 Image Processing Libraries 47
4.4.4 Supporting Libraries 47
4.5 Experiments results 47
4.5.1 stage 1 Results 47
4.5.2 stage 2 Results 49
4.6 Conclusion 54

General Conclusion **55**

Bibliography **57**

List of Tables

1.1 Comparison of Brain Imaging Modalities 9
1.2 MRI Sequences for Brain Tumor Characterization 10

2.1 Recent comparative studies 27

3.1 Architecture of the proposed Stage 1 Hybrid VGG19-CNN model 38
3.2 Stage 2 Progressive Fine-Tuning Strategy 41
3.3 Stage 2: Hybrid ResNet50 Architecture 41

4.1 Stage 1: Classification Report (Test Set) 49
4.2 Stage 2 Classification Report (Test Set) 51
4.3 Classification Report of the CNN-VGG19 Model 53

List of Figures

1.1 Anatomical structure of the brain and central nervous system[3]. 6
1.2 Pituitary Tumors[5] 7
1.3 Meningioma Tumors[5] 8
1.4 Gliomas Tumors[5] 8
1.5 Artifacts caused by head movement on MRI[9]. 11
1.6 Periodic Motion Noise Artifact in Brain MRI 12
1.7 Partial Volume Effect (PVE) in Brain MRI[10]. 12

2.1 Multilayer neural network 16
2.2 Example of how deep learning works 17
2.3 Basic architecture of a simple RNN network[15] 18
2.4 Architecture of a traditional CNN Convolutional[17] 18
2.5 Example of a convolution product[18] 20
2.6 Example of pooling (Max pooling and Average pooling)[19] 21
2.7 RelU Activation Function[15] 22
2.8 fully connected layer[20] 22
2.9 flatten layer[21]. 23
2.10 AlexNet architecture.[23] 23
2.11 ResNet architecture[25] 24
2.12 VGG-16 Architecture.[23] 25
2.13 VGG-19 Architecture.[27] 25
2.14 Difference between transfer learning and machine learning.[29] 26

3.1 Architecture Overview of the Proposed Brain Tumor Detection and Classification System. 31
3.2 Examples of MRI images of the first base in the ”No” class 32
3.3 Examples of MRI images of the first base from the ”Yes” class. 33
3.4 Examples of MRI images from the second class in a multi-class dataset. 33
3.5 Illustration of the Data Preprocessing and Augmentation. 34

3.6 Mind Map of the MRI Preprocessing and Augmentation Workflow. 34
3.7 architecture of the proposed Stage 1 Hybrid VGG19-CNN model 38
3.8 Stage 2 Hybrid ResNet50 Architecture 42

4.1 Splitting the image database. 46
4.2 Accuracy and Loss Curves of the Stage 1 Hybrid VGG19-CNN. 48
4.3 Confusion Matrix of the Stage 1 Hybrid VGG19-CNN. . . 48
4.4 Accuracy and Loss Curves of the Stage 2 ResNet50 Model (Phase 1: Head Training; Phase 2: Fine-Tuning). 50
4.5 Confusion Matrix of the Stage 2 ResNet50 Multi-Class Classifier. 50
4.6 Accuracy and Loss Curves of the model Multi-Class Classifier (VGG19-CNN). 52
4.7 Confusion Matrix of the model Multi-Class Classifier (VGG19-CNN). 52
4.8 Accuracy comparison between prior methods and the proposed framework. 54

List of Abbreviations

CNN	Convolutional Neural Network
VGG16	Visual Geometry Group 16
VGG19	Visual Geometry Group 19
ResNet50	Residual Network 50
CLAHE	Contrast Limited Adaptive Histogram Equalization
MRI	Magnetic Resonance Imaging
GAN	Generative Adversarial Network
PSO	Particle Swarm Optimization
DMFN	Deep Multi-Feature Network
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
GWO	Grey Wolf Optimization
YOLOv5	You Only Look Once version 5
ViT	Vision Transformer
FCN	Fully Convolutional Network
U-Net	U-Net Architecture

General Introduction

Brain tumors are life-threatening conditions, and early diagnosis is essential for improving patient survival and treatment outcomes. Because of their severity, complicated location, and direct effect on essential brain functions, Early detection is essential, as it can significantly improve treatment outcomes and increase the chances of recovery. One of the most popular methods for examining the brain and identifying anomalies is magnetic resonance imaging, or MRI. It helps physicians find potential cancers and produces precise images of brain structures.

However, manually interpreting MRI scans requires a great deal of expertise, takes a lot of time, and can occasionally result in mistakes or inconsistent diagnoses among specialists. The development of deep learning and artificial intelligence has led to the proposal of new automated techniques for medical image interpretation. These techniques can highlight intricate details and identify patterns that the human eye might miss.

Convolutional neural networks have thus been effectively applied to the classification of images and the identification of brain tumors from MRI data. This thesis aims to investigate and develop a novel method for brain tumor identification. The primary objective is to develop a system for automatically analyzing MRI images in order to improve the speed and accuracy of medical diagnosis.

This work proposes an automatic two-stage MRI-based framework for brain tumor analysis. In the first stage, a hybrid CNN-VGG19 model is used for binary tumor detection, and in the second stage, a fine-tuned ResNet50 model is applied for multi-class classification of glioma, meningioma, and pituitary tumor. We also use CLAHE preprocessing and data augmentation to enhance the quality of images and the generalization ability of the model. The suggested framework not only performs well but also minimizes confusion between classes and offers a decision-making process that can be interpreted in a clinical context.

A solution like this could reduce the workload of radiologists and enhance patient care. There are multiple chapters in the thesis. General ideas

about the brain, brain tumors, and medical imaging are covered in Chapter 1.

Problem Statement

Gliomas, meningiomas and pituitary tumors are among the most dangerous neurological diseases and early and accurate diagnosis is essential for successful treatment. At present, diagnosis is based on the manual examination of MRI scans by specialized radiologists, which is a time-consuming process that is vulnerable to human error and heavily dependent on individual expertise. Such variability in diagnosis can lead to delays or misdiagnosis in treatment decisions which can directly affect patient outcomes. Technically, most of the existing automatic techniques treat the detection and classification of tumor as a single stage problem, which leads to inter-class confusion especially for tumors with similar visual appearance. Furthermore, the existence of artifacts such as motion blur and noise in MRI images reduces the reliability of automated systems.
This work therefore addresses the following question: how to build a two-stage automatic deep learning framework that first separates normal from abnormal MRI scans and then correctly classifies the identified tumor into its proper subtype while being robust to common image quality degradations.

Research Objectives

The proposed system aims to:

- Automatic detection and classification of brain tumors from multi-sequence MRI (T1, T2, FLAIR, post-contrast).
- High accuracy with parallel processing and dual feature extraction (edges/regions).
- Quicker and more reliable diagnostic decisions for radiologists.
- The effectiveness and competitiveness of the proposed framework in comparison to the recent related studies.

This thesis is organized into four chapters:

- **Chapter 1: The Human Brain And Central Nervous System:** This chapter presents the basics of brain tumor detection and classification. It gives medical and technical background for the MRI based analysis. It introduces the major types of brain tumors, including gliomas, meningiomas and pituitary tumors. It describes the basics of MRI imaging, including T1, T2 and FLAIR sequences, and common artifacts that might interfere with image interpretation. The chapter concludes with a summary of the basics of deep learning, transfer learning, and the impact of automated systems on the work of radiologists.
- **Chapter 2: Deep Learning:** In this chapter, we provide the basic ideas of deep learning for medical imaging with major focus on cnn architectures like AlexNet, ResNet, VGG for brain tumor analysis. It talks about the evolution of machine learning, basic layers of neural networks like convolutional, pooling, ReLU and fully connected layers, recurrent neural networks and transfer learning techniques. The chapter ends with a review of recent works on brain tumor detection and their drawbacks such as overfitting and imbalanced datasets and small datasets.
- **Chapter 3: Proposed Model:** The suggested two-stage MRI image-based brain tumor detection framework is presented in this chapter. The Stage 1 VGG19-CNN model and the Stage 2 ResNet50 classifier's architectures, datasets, and preprocessing and augmentation techniques are all covered.
- **Chapter 4: Results and Discussion:** In this chapter we show the experiment results of the proposed framework. The experiment results include accuracy, loss and confusion matrix analysis. It also compares the performance of the VGG19 and ResNet50 models and discusses the effect of transfer learning and fine-tuning on performance of classification.

Chapter 1: The Human Brain And Central Nervous System

Chapter 1

The Human Brain And Central Nervous System

1.1 Introduction

This chapter gives the medical and technical background required to understand automated brain tumor analysis with MRI. We start with an anatomical review of the human brain and central nervous system, followed by the 2021 World Health Organization (WHO) classification of brain tumors. The chapter then introduces the three tumor types that are the main focus of this work: glioma, meningioma and pituitary tumor and describes their MRI signatures and clinical relevance. We then discuss the major brain imaging modalities, with an emphasis on MRI principles, acquisition sequences, and common artifacts that can affect both manual and automated analysis. The ideas introduced in this chapter form the basis of the deep learning framework discussed in the next chapters[1].

1.2 Brain Tumors

Brain tumors are abnormal cell masses which develop in brain or central nervous system. They are divided into two categories primary tumors which arise directly in the brain and metastatic tumors derived from other organs. As per the classification by WHO (2021), there are four grades of tumors I to IV. The higher the grade, the more aggressive and metastasizing the tumor is Primary brain tumors include gliomas (arising from glial cells), meningiomas (arising from meninges), and pituitary tumors In adults. however, brain metastasis is more prevalent. Tumors cause a number of complications by way of mass effect, leading to swelling, bleeding or infiltration of neighboring tissue. This is why early diagnosis is essential in order to adapt treatment. Magnetic resonance imaging (MRI)

remains the gold standard for diagnosis, thanks to its ability to provide excellent contrast of soft tissues and images in multiple planes. The Figure 1.1 shows the principal anatomical components of the human brain and surrounding structures, which are essential for understanding brain MRI interpretation[2].

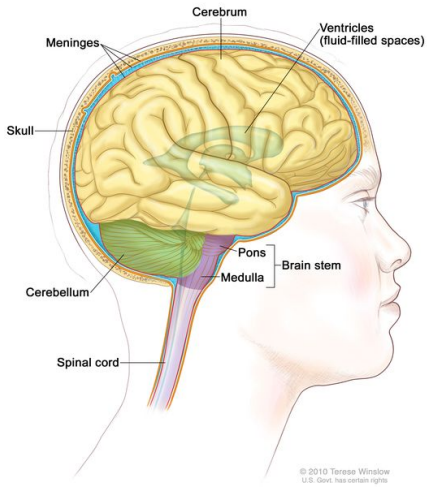


Figure 1.1: Anatomical structure of the brain and central nervous system[3].

1.2.1 Definition and general classification (benign/malignant, primary/secondary)

Brain tumors are abnormal masses of cells that are formed by uncontrolled cell division inside the skull. These tumors are classified according to two main criteria:

- **Primary vs Secondary (By Origin)** Primary tumors arise from the cells of the brain itself or from surrounding tissues. Secondary/Metastatic Tumors Cancer cells that have spread to the brain from other organs in the body (such as lung or breast).
- **Benign vs Malignant (By Behavior)** Benign tumors : Slow growing cells with clear margins that do not invade surrounding tissue ; often present as well-defined masses on magnetic resonance imaging (MRI). Malignant tumors: Cancerous cells that rapidly multiply and invade and destroy surrounding healthy tissue. There are no clear boundaries between malignant tumors and healthy tissue[1].

1.2.2 Main Tumor Types Considered in This Work

The way brain tumors appear on magnetic resonance imaging (MRI) is determined by where they come from, what they are made of histologically, and where they are located anatomically. This study concentrates on three frequently encountered categories employed in multi-class classification tasks: pituitary tumors, meningiomas, and gliomas. Each of these tumor types poses unique difficulties for automated detection and classification methodologies[4].

a) Pituitary Tumors

Pituitary adenomas arise from the pituitary gland, which is located at the base of the brain. On MRI, they are seen as circumscribed rounded lesions which generally have a relatively uniform signal. Hypo-to iso-intense on T1 weighted images and hyper intense on T2 weighted images are their usual appearance, contrast enhancement will vary after administration of gadolinium. Extension suprasellarly and compression of surrounding anatomy are other changes that MRI may show[4].

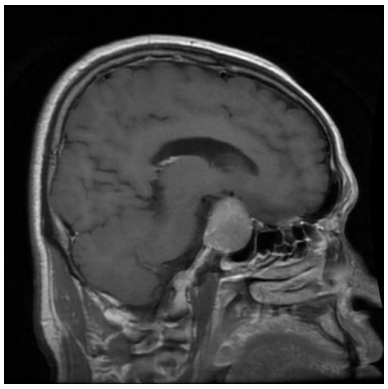


Figure 1.2: Pituitary Tumors[5]

b) Meningioma

Meningioma are neoplasms of arachnoid cap cells. They commonly appear as extra-axial, sharply demarcated, homogenously enhancing mass lesions which are characteristically hyperintense in comparison to gray matter, though appear isointense on contrast enhanced MRI T1 weighted sequences, in relation to gray matter (where they have a dural tail sign), with a heterogeneous enhancement on T2 weighted sequence, due to the presence of necrotic components in larger lesions. MRI appearance and

location are diagnostically significant, and they often have a hyperostosis of the adjacent bone[4].

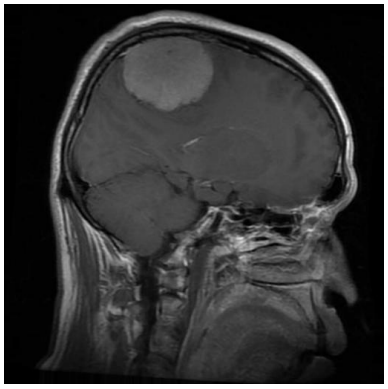


Figure 1.3: Meningioma Tumors[5]

c) Gliomas

Gliomas originate from glia and can range in grade from low-grade (grade II) to high-grade (grade IV-glioblastoma). They usually are intra-axial lesions with indistinct margins and heterogeneous signal on T1 imaging. On T2 and FLAIR images the edema and infiltrative component is seen as hyperintense. The enhancement patterns on contrast are widely variable from ring-enhancement seen in glioblastomas to minimal-enhancement seen in low grade tumors. Areas of necrosis and hemorrhage are often seen with high grade tumors[4].

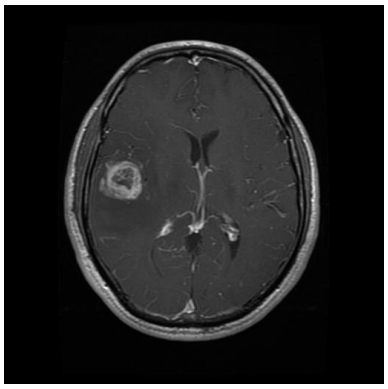


Figure 1.4: Gliomas Tumors[5]

These tumor forms were selected for their clinical significance, unique MRI characteristics, and high frequency in public datasets utilized for training deep learning models. They retain similarities with non-tumoral lesions and imaging artifacts that hinder their classification.

1.3 Medical Imaging of the Brain

Medical imaging modalities for the brain, including MRI and CT scans, offer intricate and non-invasive insights into the internal structures of the brain. These can proficiently identify cancers, strokes, and neurological diseases.

1.3.1 Overview of Brain Imaging Modalities (CT, PET, MRI)

Several imaging modalities are used to evaluate the brain, each with its own strengths:

modality	**Principle**	**strengths**	**Limitation**
Ct	X-ray attenuation	X-ray attenuation bone	Poor soft tissue contrast, radiation
Pet	Radioactive tracer uptake	Metabolic activity, functional	Expensive, radiation, low spatial resolution
Mri	Magnetic field + radiofrequency	Excellent soft tissue contrast, multisequence	Long acquisition, contraindications (pacemakers)

Table 1.1: Comparison of Brain Imaging Modalities

1.3.2 Role of MRI Compared to Other Modalities

MRI is the gold standard for brain tumor imagingdue to:

- **S**uperior soft tissue contrast without ionizing radiation.
- **Multi** Multi-planar imaging capability (axial, sagittal, coronal).
- **Multiple** Multiple sequences providing complementary information.
- **Ability** Ability to assess tumor anatomy, vascularity, and invasion[6].

1.4 Brain Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging (MRI) of the brain is a noninvasive imaging modality that uses powerful magnets and radio waves to create detailed images of the brain's structures. MRI is excellent for detection of tumors, stroke, and soft tissue abnormalities and does not involve radiation.

1.4.1 Basic Principle of MRI

In MRI, a powerful static magnetic field B is applied to the body and it results in the alignment of atomic nuclei (mainly water protons) present in the body. Radio frequency waves are then projected to disturb this balance, and when the excitation ceases, the protons slowly reorient sensitive back to their original state by releasing a measurable signal. The device's antennas pick up this signal, which is converted into data and processed by the software to produce images. Different types of tissue contribute as variations in signal that helps discriminating between structures and lesions[6].

1.4.2 MRI Sequences for Brain Tumor Imaging

MRI provides multiple complementary views of brain tumors by using different sequences such as T1, T2, FLAIR, and post-contrast T1. Each sequence highlights different tissue properties like anatomy, edema, necrosis, and tumor enhancement, helping radiologists to make more accurate diagnoses and assess the extent of disease. These sequences are combined as multimodal inputs in deep learning to improve the detection, segmentation, and grading of brain tumors[7].

Sequence	**Purpose**	**Brain Tumor Appearance**
T1	Anatomy	Hypointense tumor core, normal gray matter reference
T1-CE (contrast)	Blood-brain barrier breakdown	Ring enhancement (glioblastoma), homogeneous (meningioma)
T2	Edema, CSF	Hyperintense peritumoral edema
FLAIR	Infiltration beyond contrast	Tumor extent, suppresses CSF signal

Table 1.2: MRI Sequences for Brain Tumor Characterization

1.4.3 Specific Characteristics of Brain MRI Images

MRI images are unique in that their contrast depends on acquisition parameters and tissue properties, resulting in different types of weighting (T1, T2, proton density). Technical parameters such as TR (repetition time) and TE (echo time) directly influence this contrast and must be considered during interpretation and automatic image analysis design[7].

1.4.4 Common MRI Artefacts

a) Motion artifacts:

Caused by periodic (heartbeat, breathing) or non-periodic (head movement, swallowing) motions during acquisition. They create blurring and ghosting. Fast imaging sequences (echo-planar) reduce these effects[8]. The figure 1.5 demonstrates that patient motion during acquisition can lead to motion artifacts causing blurring and ghosting patterns which obscure anatomical detail in brain MRI.

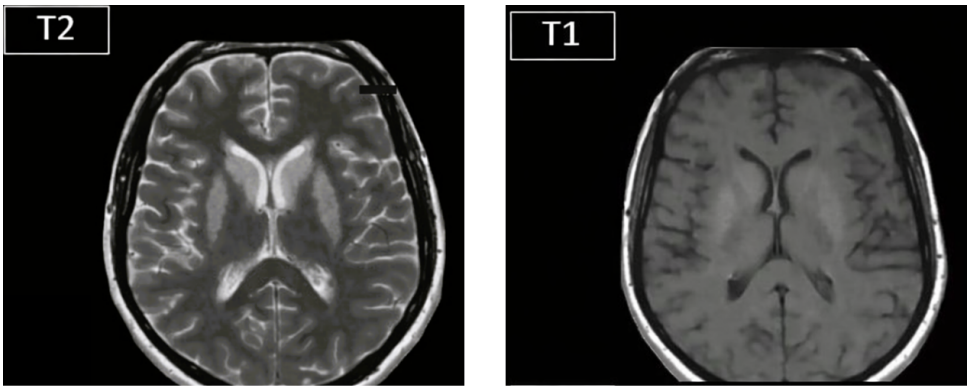


Figure 1.5: Artifacts caused by head movement on MRI[9].

b) Random noise:

Thermal noise from protons and electronics. In terms of signal-to-noise ratio (SNR), higher field strength (3T vs 1.5T) and averaging increase SNR but also increase acquisition time[8], This figure 1.6 shows a brain MRI image and its corrupted version, which is corrupted by periodic noise and introduces visible grid-like artifacts and degrades the quality of the image.

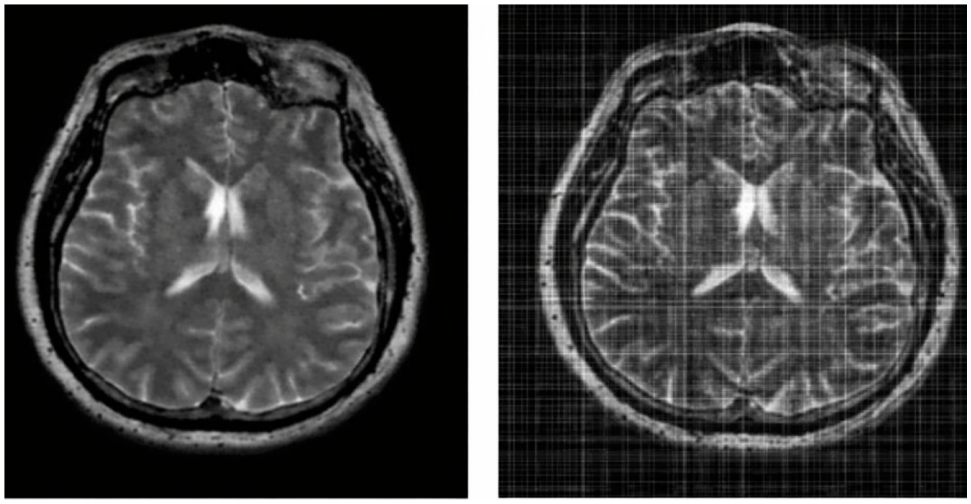


Figure 1.6: Periodic Motion Noise Artifact in Brain MRI

c) Partial volume effect:

This issue occurs when the slice thickness is larger than the size of the structure being imaged. A voxel contains several tissue types and an intermediate signal intensity. This information is crucial for detecting small tumors and delineating their boundaries[8]. the figure 1.7 illustrates the partial volume effect in brain MRI, where different tissue types contribute to the same voxel, resulting in ambiguous signal intensities that may affect image interpretation.

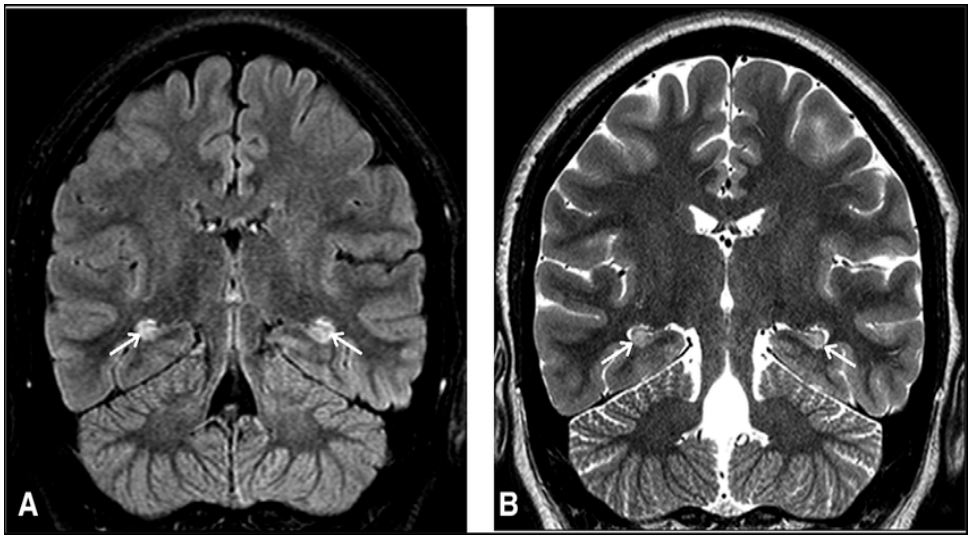


Figure 1.7: Partial Volume Effect (PVE) in Brain MRI[10].

1.5 Conclusion

In this chapter we have provided the necessary medical and technical background for automatic detection of brain tumor by MRI. We reviewed brain anatomy, WHO classification of tumors (gliomas, meningiomas, pituitary adenomas - grades I-IV) and modalities of imaging, emphasizing on superiority of MRI because of soft tissue contrast and multiple sequences (T1, T2, FLAIR, post-contrast). MRI is still the gold standard for accurate characterization of these tumors and detection of common artifacts (motion, thermal noise, partial volume effect) that hinder automatic analysis. These elements provide the necessary basis for the development of powerful artificial intelligence algorithms for multi-class classification revealing promising prospects for early diagnosis and treatment planning.

Chapter 2: Deep Learning

Chapter 2

Deep Learning

2.1 Introduction

In this Chapter, we present the fundamental concepts of deep learning that are useful in brain tumor classification. It introduces the main CNN components, transfer learning and most widely used architectures such as AlexNet, ResNet and VGG which form the basis of the proposed framework. The chapter also reviews related work providing the necessary background for understanding the motivation and design of the present study.

2.2 Machine learning

Machine learning (ML) is a sub-field of artificial intelligence (AI) that focuses on the development of systems that can learn or enhance their performance from experience or data. Artificial intelligence is a generic term for systems or machines mimicking some aspect of human intelligence. ML and AI are often talked about together and these words are sometimes used interchangeably but they do not mean exactly the same thing. It is important to note that ML is a subset of AI, but AI is not only ML. ML is used in all fields these days. We use ML algorithms to optimize, facilitate, and secure our experience when we are dealing with banks, shopping online, or using social media. ML and the technology around it are moving fast, and we're just beginning to see what it can do[11].

2.3 Deep learning

Deep learning is a subfield of machine learning that uses artificial neural networks with multiple layers to extract deeper insights from data. Each layer takes the easier patterns from the previous layer, transforming raw

MRI pixels into edges and shapes and then into tumor shapes via a series of math operations (see Figure 2.1). Unlike traditional ML where you hand select features. Deep learning automatically learns relevant representations from data . It teaches itself to identify significant patterns automatically from the noisy images leading to reliable diagnoses by adjusting its own settings by gradient descent with no intermediate steps. It's perfect for end-to-end processing of medical scans. This approach has transformed computer vision and healthcare, and has attained state-of-the-art results in tasks such as tumor detection and segmentation. It is driven by large-scale data and GPUs and continues to improve in handling difficult pattern recognition in medicine and beyond[12].

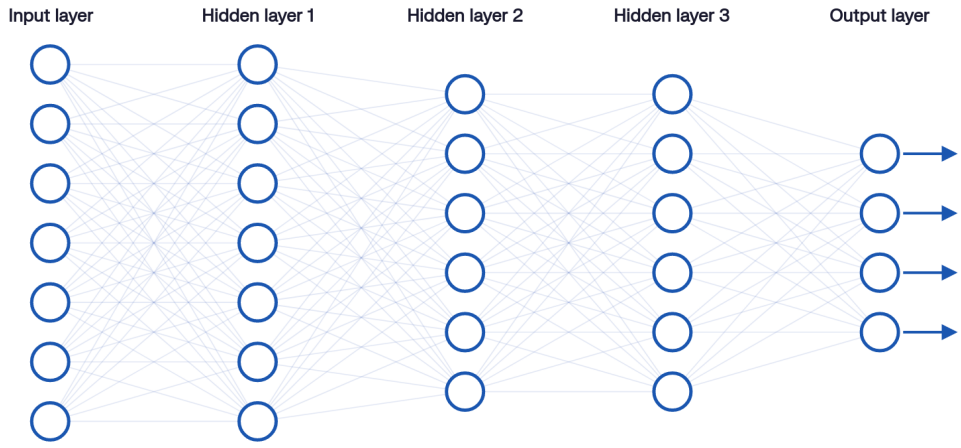


Figure 2.1: Multilayer neural network

2.3.1 How Does Deep Learning Work?

Deep learning is powered by layers of neural networks, which are algorithms loosely modeled on the way human brains work. Training with large amounts of data is what configures the neurons in the neural network. The result is a deep learning model which, once trained, processes new data. Deep learning models take in information from multiple data sources and analyze that data in real time, without the need for human intervention. In deep learning, graphics processing units (GPUs) are optimized for training models because they can process multiple computations simultaneously [13].

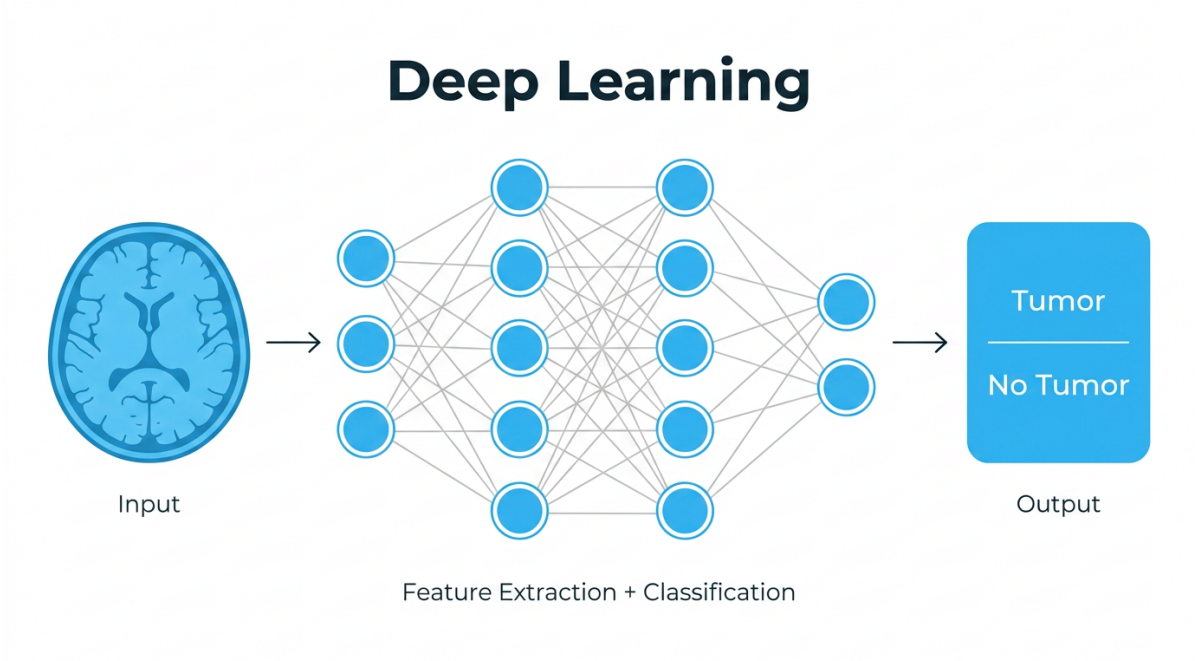


Figure 2.2: Example of how deep learning works

Deep neural networks need input and output layers to function. The input layer is where data is input into the network for processing. The output layer is where the final prediction or classification is made. Figure 2.2 The workflow of deep learning. The brain image is processed by several hidden layers for feature extraction and the final classification result.

2.4 Deep neural network architectures

Deep neural network architectures such as CNNs, RNNs, and ResNets are constructed from multiple layers stacked on top of each other, which learn hierarchical features, allowing them to perform well in tasks like image analysis, sequential data processing, and complex pattern recognition.

2.4.1 Recurrent neural network (RNN, recurrent neural network)

Recurrent Neural Networks (RNNs) are a type of neural networks that can represent sequential data and recall the state information from the previous phase. RNNs perform something similar, by "remembering" past information and using all of the previous words to determine the most likely next word. This memory of past steps helps the network recognize context and make better estimates. They are particularly useful for tasks where the context and order matter[14].

- Designed for sequential and temporal data.
- Maintains memory of past inputs.
- Widely used in NLP, forecasting and speech tasks.

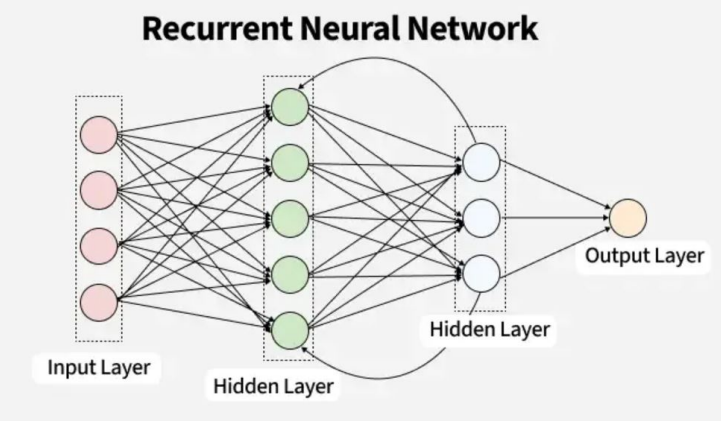


Figure 2.3: Basic architecture of a simple RNN network[15]

2.4.2 Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) are a special type of neural network that are used for processing data that has a known, grid-like topology. An example of this is image data. In at least one of their layers CNNs use a mathematical operation called convolution instead of general matrix multiplication. Inspired by the human visual system, CNNs have three key properties: sparse interactions, parameter sharing, and equivariant representations that reduce the number of parameters and improve generalization. They learn hierarchical features from simple edges to complex objects, capture spatial patterns using convolution operations, detect objects regardless of their position in the image and reduce computation by focusing on local regions instead of the whole image[16].

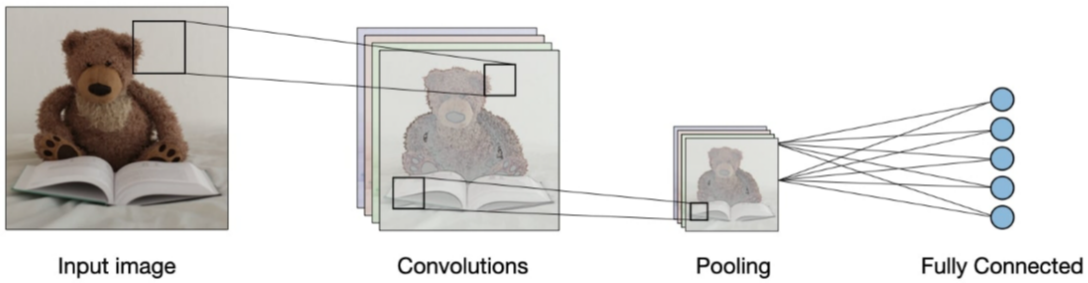


Figure 2.4: Architecture of a traditional CNN Convolutional[17]

The convolution layer and the pooling layer can be fine-tuned with respect to hyperparameter that are described in the next sections.

2.5 Different layers of CNN

In CNNs, the convolutional, pooling, ReLU, fully connected, and flatten layers work together to extract features from images efficiently, reducing dimensions while maintaining significant patterns for precise classification.

2.5.1 Convolutional layers

Convolutional neural networks are constructed from convolutional layers. Their primary function is to extract features from the input data (usually images). Their main operations are as follows:

- **Convolutional filtering:** The convolution layers apply a filter (or kernel) to the input image. A filter is a small matrix, usually 3x3 or 5x5, that passes over (or "convolves") the image.
- **Feature identification:** Each filter searches for different types of features, such as certain edges, textures or patterns. For example, one filter may look for horizontal edges and the other for vertical edges.
- **Feature maps:** The convolution layers apply a set of filters to the input that can be learnt, and generate a set of feature-maps. The number of output feature maps equals the number of filters, and each map encodes a specific set of learned spatial features.
- **Nonlinearity:** Then, a nonlinear activation function such as ReLU (Rectified Linear Unit) is often used after the filter to add nonlinearity to the model. This aids in capturing more complex relationships in the data [16].

The figure 2.5 below shows how convolution extracts features from an image.

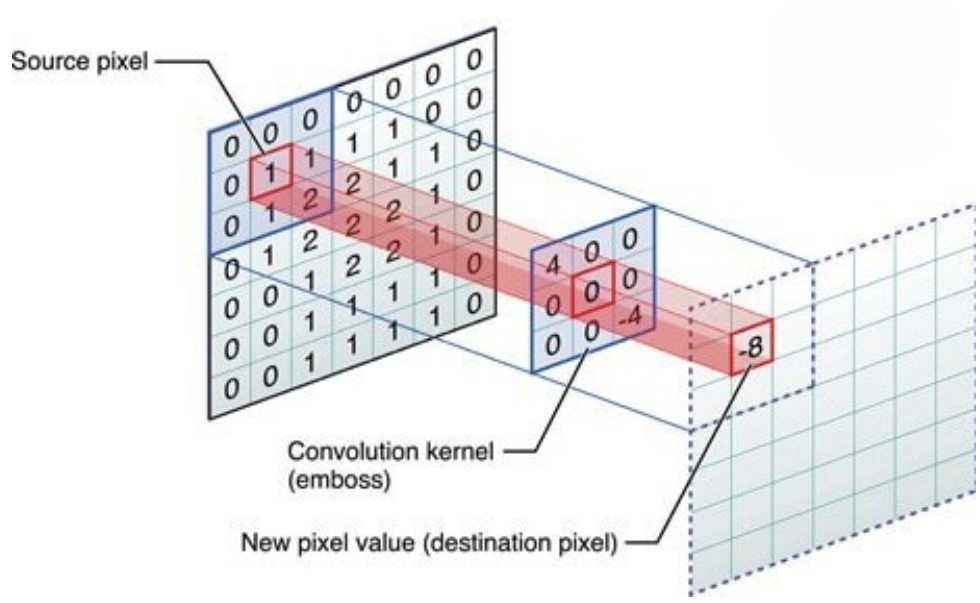


Figure 2.5: Example of a convolution product[18]

2.5.2 Pooling layer

Pooling or sub-sampling layers help to reduce the spatial dimensions of the feature maps, while preserving the important information. They greatly reduce overfitting by retaining only the most salient activations, and therefore greatly reduce the number of trainable parameters and computational cost. Two common pooling strategies are:

- **Max-Pooling:** Max Pooling is the most common type of pooling. It divides the feature map into non-overlapping sub-regions, and keeps the maximum activation value from each sub-region. For instance, in a 2×2 pooling window, the Max Pooling operation selects the maximum value among the four activation values.
- **Average-Pooling:** Another popular technique is Average Pooling in which we compute the mean of the activation value for each non-overlapping sub-region. This strategy is more conservative than Max Pooling, as it takes all activations in the region, but tends to maintain less spatial detail.

Pooling reduces the size of feature maps. This decreases the number of parameters and computations in the network. This can help to improve the model efficiency and reduce the chance of overfitting. (see figure 2.6)[19].

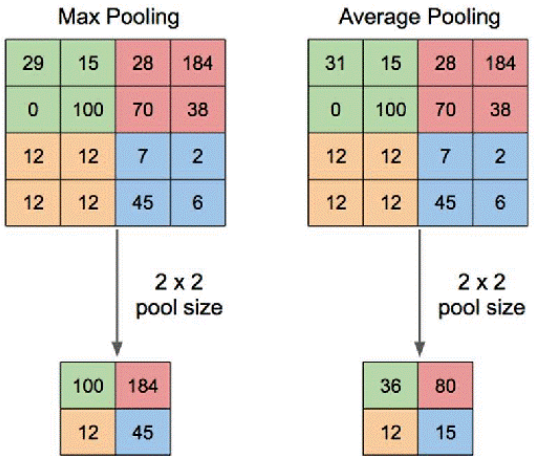


Figure 2.6: Example of pooling (Max pooling and Average pooling)[19]

2.5.3 ReLU Activation Function

ReLU, the go-to activation function in deep learning, is positive values pass through untouched, negatives get zeroed out. Its piecewise linear nature keeps networks efficient and sidesteps vanishing gradient headaches that plague sigmoid/tanh. Default choice for a reason. The ReLU function can be described mathematically as follows:

f(x)=max(0,x)

Where:

- x is the input to the neuron.
- The function returns x if x is greater than 0.
- If x is less than or equal to 0, the function returns 0.

The formula can also be written as:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

ReLU is notably effective for training deep neural networks since it preserves nonlinearity without complex transforms and enables more efficient training[15].

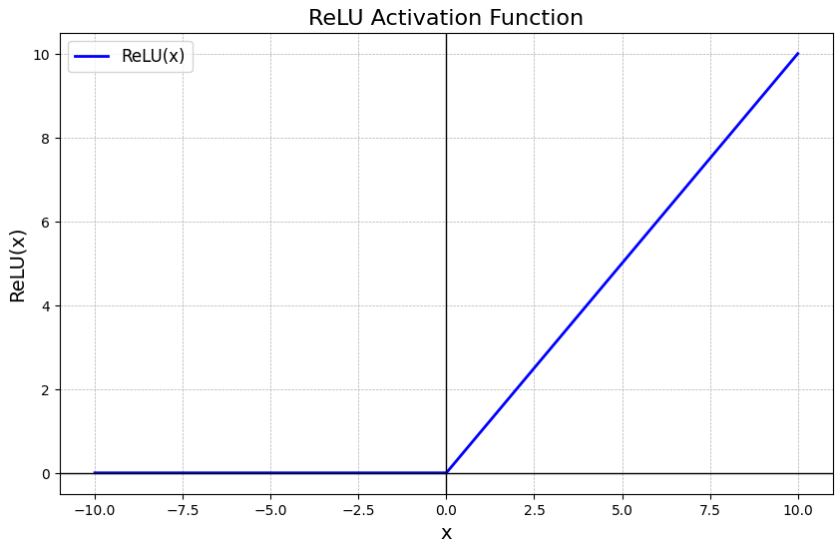


Figure 2.7: RelU Activation Function[15]

2.5.4 Fully Connected Layer

A fully connected layer (or dense layer) is a fundamental building block of artificial neural networks in which each neuron receives input from every other neuron in the previous layer. Its primary function is to combine the features extracted by previous layers to make final, complex predictions[16].

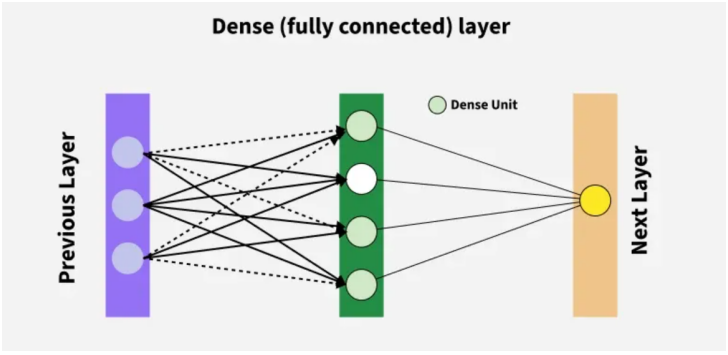


Figure 2.8: fully connected layer[20]

2.5.5 Flatten Layer

Flattening is one of the basic operations in neural networks, specifically when going from convolution layers to fully connected layers Another important layer of Convolutional Neural Network (CNN) is flatten layer(figure 2.9, which transforms the output from convolution and pooling layers to a single-dimensional or vector. This vector is then passed to other fully connected (dense) layers for class or regression objectives[16].

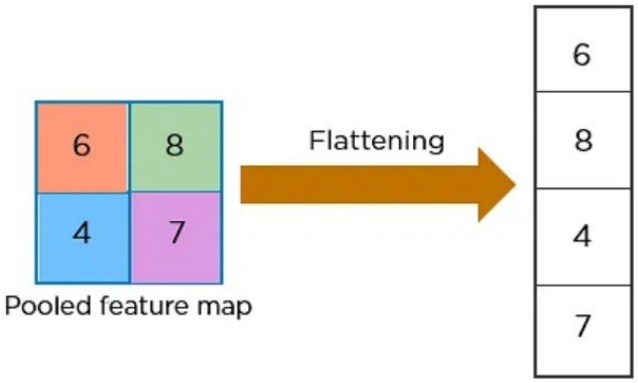


Figure 2.9: flatten layer[21].

2.6 Different Types of CNN Models

This section presents key CNN architectures (AlexNet, ResNet, VGG) pivotal for brain tumor classification on MRI images.

2.6.1 AlexNet:

AlexNet is a CNN architecture that was developed by Alex Krizhevsky, Ilya Sutskever and Geoffrey Hinton in 2012. It was the first CNN to win the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) a major image recognition competition. AlexNet processes input images measuring 227×227×3. The first convolutional layer applies 96 filters measuring 11×11 (stride=4), producing a volume of 55×55×96. A MaxPooling layer (3×3, stride=2) follows immediately. The architecture progresses in this way to a Dense(9216) layer, then two Dense(4096) layers, ending with Softmax(1000 classes). In total, AlexNet contains approximately 60 million trainable parameters[22], see figure 2.10.

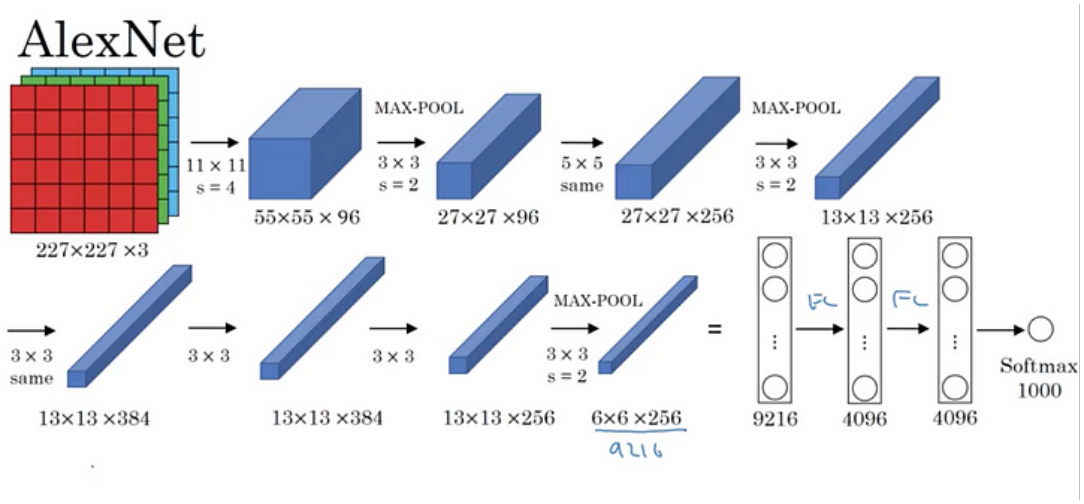


Figure 2.10: AlexNet architecture.[23]

2.6.2 Resnet

Residual Networks (ResNets) are deep learning architectures designed to facilitate the training of very deep neural networks. It incorporates skip (shortcut) connections that enable the model to learn residual mappings instead of direct mappings.

- Helps prevent vanishing gradient problems in very deep models.
- Allows information to flow directly across layers using skip connections.
- Enables building networks with hundreds or even thousands of layers[24], See figure 2.11.

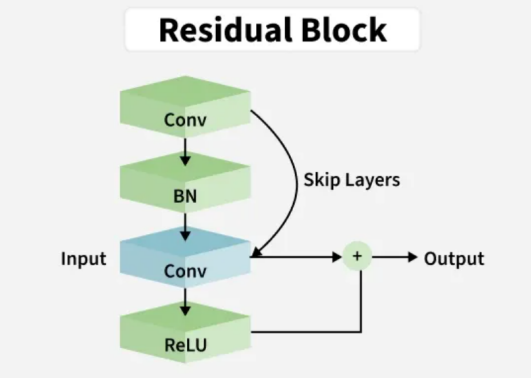


Figure 2.11: ResNet architecture[25]

2.6.3 VGG

The VGG architecture is developed by the Visual Geometry Group at the University of Oxford. It is characterized by the use of small 3 X 3 convolutional filters that are stacked in multiple layers. It is well suited to tasks of image classification. The simplicity and uniformity of the design of VGG results in high performance on visual recognition problems. VGG-16 and VGG-19 models improved state-of-the-art results on ImageNet, showing the importance of network depth in CNNs[26].

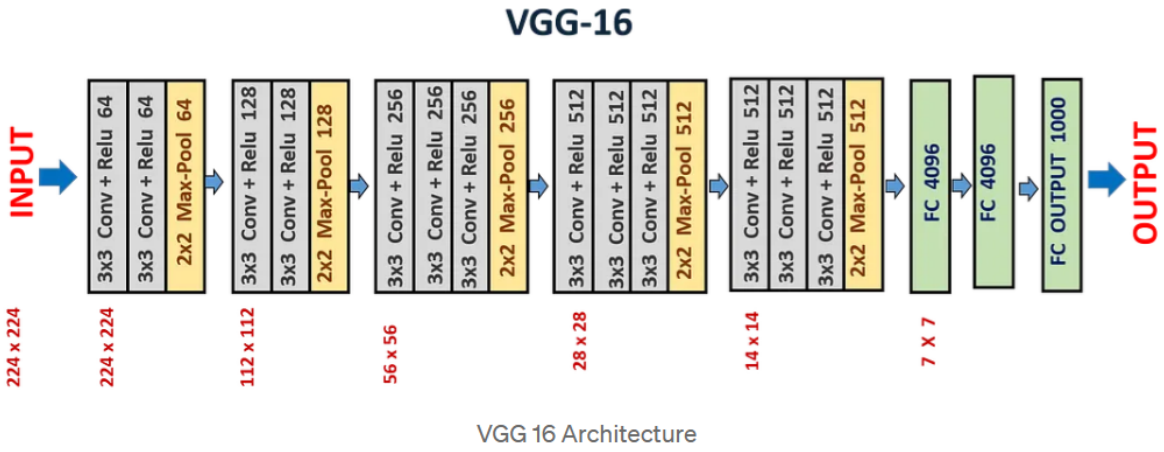


Figure 2.12: VGG-16 Architecture.[23]

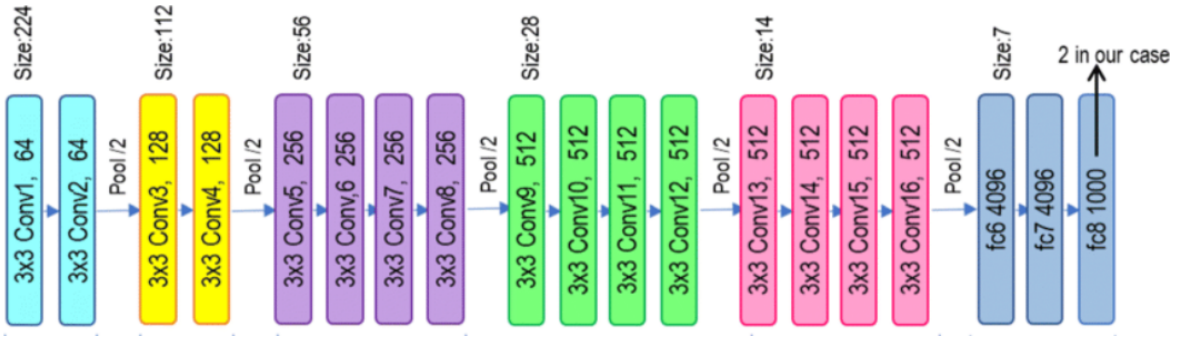


Figure 2.13: VGG-19 Architecture.[27]

2.7 Transfer learning

Transfer learning is the process of taking a previously trained model on one task and reapplying it to a similar task, this is particularly useful if the new task contains limited data. This helps in the following ways:

- Uses learned features from the first task.
- Reduces training time for the new task.
- Improves accuracy with less data.
- Uses general features that work across tasks[28].

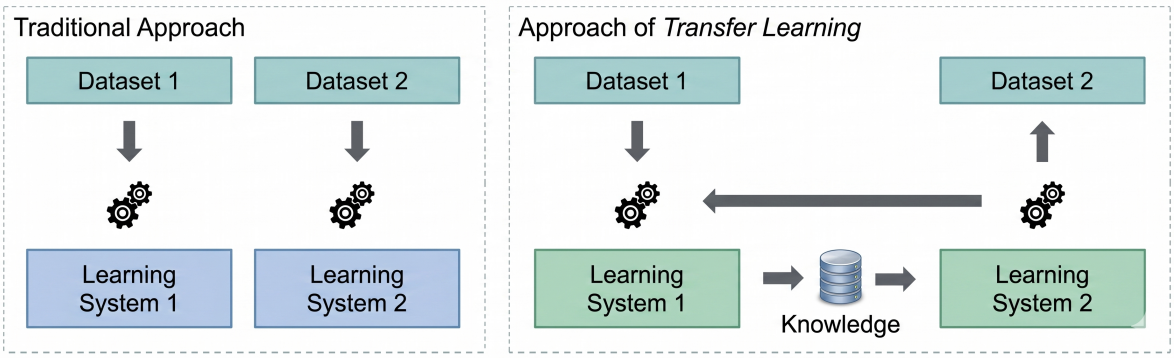


Figure 2.14: Difference between transfer learning and machine learning.[29]

2.8 Deep learning applications

Deep learning shines in complex recognition tasks, especially medical imaging where it spots brain tumors (gliomas, meningiomas, pituitary) on MRIs with higher accuracy than humans catching subtle signs like ring enhancement or edema that eyes miss. Industrial side, it flags production defects in real-time. Beyond images, it classifies millions of scans, auto-generates radiology reports via NLP, translates diagnoses globally, and predicts disease progression from image time series for early treatment tracking.

2.9 Related work

Previous works on brain tumor analysis are grouped into three classes binary classification, multi-class classification, and hybrid techniques. These works have different model designs, different sizes of datasets, and different classification targets. However, they aim to improve the automatic identification and diagnosis of brain cancers from MRI scans.

2.9.1 Binary Classification

Binary classification studies aim to distinguish between two categories, namely tumor and no tumor. Previous studies in this area include Garg et al[30], who proposed a hybrid framework combining Otsu segmentation, SWT, PCA, GLCM feature extraction, and a KNN-RF-DT majority-voting classifier, achieving an accuracy of 97.305% for benign/malignant classification. Mathivanan et al[31]. also presented a lightweight MRI-based brain tumor detection framework, reporting up to 99.68% accuracy on one of the evaluated datasets after preprocessing and data augmentation. These works show that binary classification can achieve high per-

formance, but many methods still rely on handcrafted features, complex preprocessing, or limited dataset diversity.

2.9.2 Multi-class Classification

Multi-class classification focuses on identifying multiple tumor types, usually including glioma, meningioma, pituitary, and in some cases a no-tumor class. Rahman et al[32]. proposed a multi-class model based on a hybrid deep learning architecture and reported accuracies ranging from 97.8% to 100% across datasets. Similarly, ZainEldin et al[33]. presented a four-class YOLOv5-based model enhanced with a Sine-Cosine Grey Wolf Optimization strategy, achieving 99.3% mAP, 99.1% precision, and 99.5% recall. In addition, Ullah et al[34]. introduced TumorDetNet, which achieved 99.27% accuracy for three-class classification. Although these studies report strong results, their limitations include high computational cost, model complexity, and the absence of clinical validation on diverse real-world MRI data. Table 2.1 Summary of selected studies related to brain tumor classification We present the models used, the classification tasks and the main limitations of each work. This comparison aids in identifying the research gap which is addressed in this thesis.

Study	**Accuracy**	**Model**	**Challenge**
Bauer et al. (2013)	~88%	FCN, U-Net, AlexNet/VGG	Outdated, small datasets, no modern architectures
Garg et al. (2021)	97.305%	KNN-RF-DT	Manual feature engineering, limited clinical generalization
Amin et al. (2022)	97.8%–100%	23-layer CNN + VGG16	Very small dataset, overfitting
ZainEldin et al. (2022)	99.3%	YOLOv5 + GWO	Small dataset, high complexity, poor interpretability
Solanki et al. (2023)	98.5%	CNN-SVM + Grey Wolf Optimization	Manual feature engineering, compute-intensive, limited data
Rahman et al. (2023)	98.7%	ViT + CNN ensemble	High GPU cost, no clinical validation
Ullah et al. (2023)	99.83%	TumorDetNet	High complexity, long training, no clinical validation
Agarwal et al. (2024)	98.36%	GAN + ResNet18 + PSO + DMFN	High computational cost, complex pipeline
Mathivanan et al. (2025)	99.68%	BTDN + SecureNet	Heavy preprocessing, limited datasets, validation needed

Table 2.1: Recent comparative studies

The main problems across these studies are:

- **Model-related problems:** Transfer learning is the process of taking a previously trained model on one task and reapplying it to a similar task, this is particularly useful if the new task contains limited data. This helps in the following ways:
- **Accuracy-related problems:** Some articles report very high accuracy, but the results might be dataset-specific and may not be totally dependable in real-world applications without multi-center validation.
- **Image-related problems:** Datasets can sometimes be short, imbalanced, or lacking in diversity, which may lead the model to learn patterns specific to the dataset rather than authentic traits of tumors.
- **Overfitting / underfitting:** When the dataset is too tiny, overfitting frequently occurs. When the model is unable to adequately represent intricate tumor patterns, underfitting may result. In conclusion, the primary drawbacks include a lack of data, poor generalization, overfitting danger, and expensive computation.

It should be noted that the selected studies were conducted on different datasets, so the reported performance values reflect the characteristics of each individual dataset.

2.10 Conclusion

The chapter presented the literature review. It included the basic components of CNNs, the common architectures like AlexNet, VGG and ResNet, and the importance of transfer learning in the field of medical imaging. The chapter also presented review of existing studies and common issues such as overfitting, high computational cost and poor clinical generalization that have influenced design choices made in this work.

Chapter 3: Proposed Approach

Chapter 3

Proposed Approach

3.1 Introduction

In contrast to the bulk of past studies that immediately use multi-class classification or binary tumor identification in one stage, in this chapter, the proposed two-step approach is given. The first stage is a binary classification of whether an MRI picture has a tumor or not, and the second stage is triggered only for positive samples to determine the tumor type. This two-step approach avoids redundant calculation, promotes interpretability, and is more analogous to the diagnostic pathway utilized in clinics. Stage 1 is a hybrid CNN-VGG19 architecture based on binary classification (tumor and no tumor), and Stage 2 is a fine-tuned ResNet50 model based on multi-class tumor type classification (glioma, meningioma, and pituitary). This chapter contains the full detailed implementation of the two steps.

3.2 Proposed Model

We propose a cascaded framework for brain tumor MRI images analysis using deep learning methods. The system is two stages, rather than a single model for the entire task. The first phase locates whether a tumor is present in the MRI image, and the second phase classifies the detected tumor into one of the known classes. The motivation of such a two stage design is that brain tumor classification is challenging if all classes are considered in a single stage. Dividing the process allows each model to concentrate on a more specialized task, potentially leading to better overall quality on predictions and less classification errors. In addition, this approach prevents normal MRI images from being transferred to the classification step, the entire framework is more clean and more efficient. The first stage is the binary classification task of deciding whether the given

MRI image includes a brain tumor or not, serving as a filtering step to pass only tumor-positive cases to the second stage. All images are pre-processed with CLAHE for local contrast enhancement and then are input to the model trained with binary cross-entropy loss function and optimized using the Adam optimizer. Adam optimizer was used because it provides fast convergence, adaptive learning rates, and stable training performance for deep neural networks. Stage 2 is the first fine-grained classification, which takes as input only tumor positive images that have been identified in Stage 1, subdividing these into three subtypes: Glioma, Meningioma, and Pituitary tumor As shown in Figure 3.1 the training strategy of the model is based on a two stage process by freezing the backbone for the head training and then partially unfreezing for finetuning with a smaller learning rate.

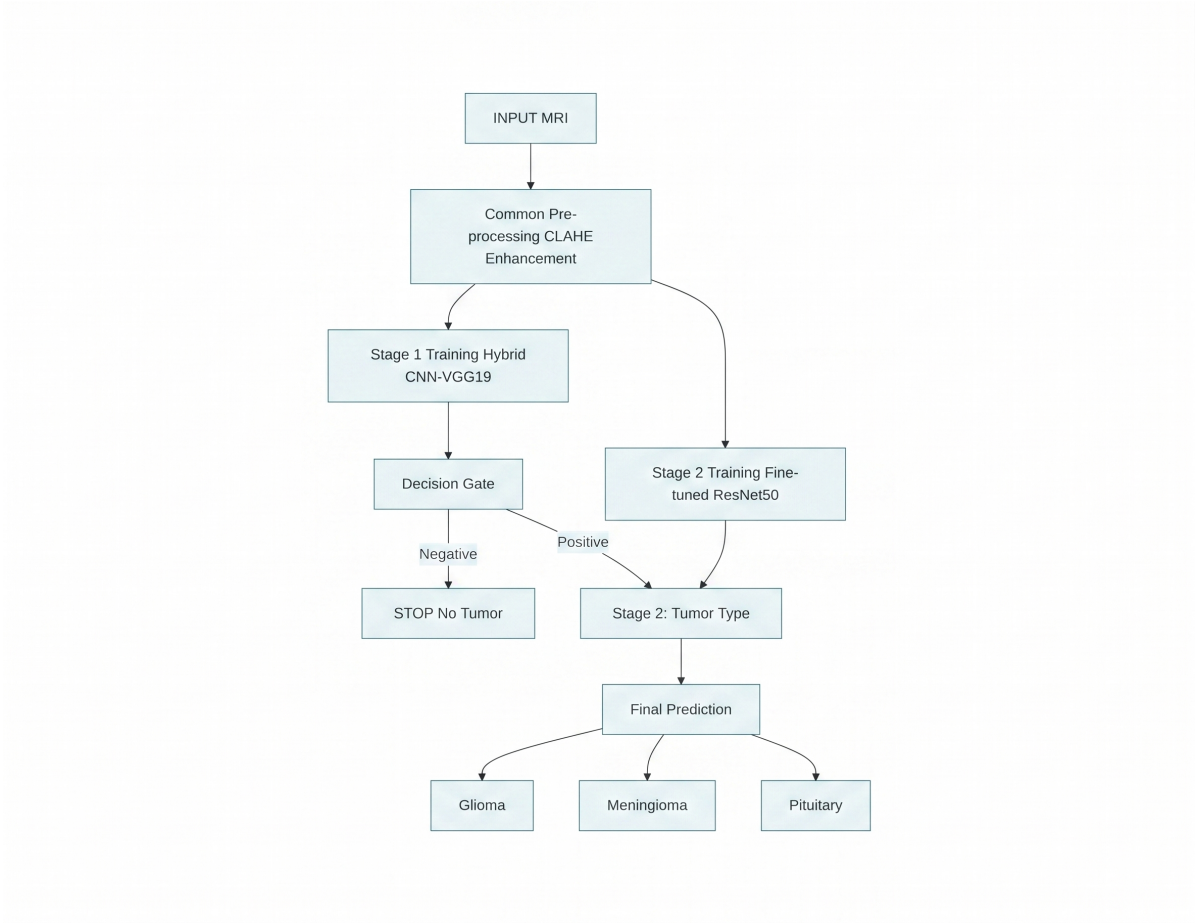


Figure 3.1: Architecture Overview of the Proposed Brain Tumor Detection and Classification System.

3.2.1 Datasets

The images used in this study were collected from two publicly available datasets hosted on the kaggle platform, The first dataset is meant for binary brain tumor detection consisting of 2 classes (yes (tumor present) and no (no tumor)).
The first dataset contains 3000 images divided into two classes[35] :

- **class 1**: presence of a brain tumor, labeled "Yes", comprising 1,500 images.
- **class 2**: absence of a brain tumor, labeled "No", comprising 1,500 images.

The second dataset contains 4200 images and is organized into three classes [5], each corresponding to a distinct tumor type :

- **class 1**: 1400 images Glioma tumor.
- **class 2**: 1400 images Meningioma tumor.
- **class 3**: 1400 images Pituitary tumor.

All images are in JPEG/PNG format, obtained through Magnetic Resonance Imaging (MRI), making the data optimal for training deep learning techniques for brain tumor detection and classification.



Figure 3.2: Examples of MRI images of the first base in the "No" class



Figure 3.3: Examples of MRI images of the first base from the "Yes" class.

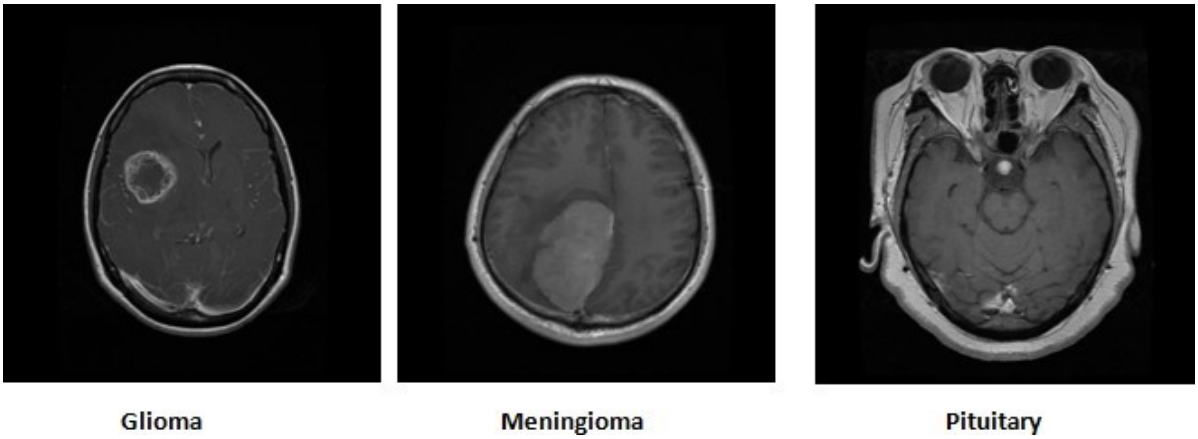


Figure 3.4: Examples of MRI images from the second class in a multiclass dataset.

3.2.2 Data Preprocessing and Augmentation

The MRI scans were initially gathered as raw data images and were then subjected to the preprocessing stage prior to being fed to the model. The first thing we need to do is to reshape the images to $224 \times 224 \times 3$. This dimension is not randomly selected the training shape is consistently applied to all images in the dataset for easier process. The 224×224 is the image size (height and width) and the 3 is for the three color channels. Even if MRI images are inherently grayscale, they are typically duplicated across three channels in order to conform to the input format of pre-trained packages like VGG19, ResNet50, which expect 3 input channels. Then, CLAHE[36] is applied after resizing this method is adopted to enhance the contrast of the MRI image. In layman terms, it allows to see more details in the image, more specifically at tumor region if exists. Instead of performing global contrast enhancement as in the case of ordinary histogram equalization, CLAHE contrast enhancement operates on small

data regions which are called tiles, and it enhances the contrast of each tile artificially without affecting the rest of the tiles, having the effect of making the overall image appear brighter without making it look too dark or too bright. It is very beneficial in medical images,as some regions of the image may have poor contrast and cannot be differentiated easily. Figure 3.5 presents a Illustration of the Data Preprocessing and Augmentation.

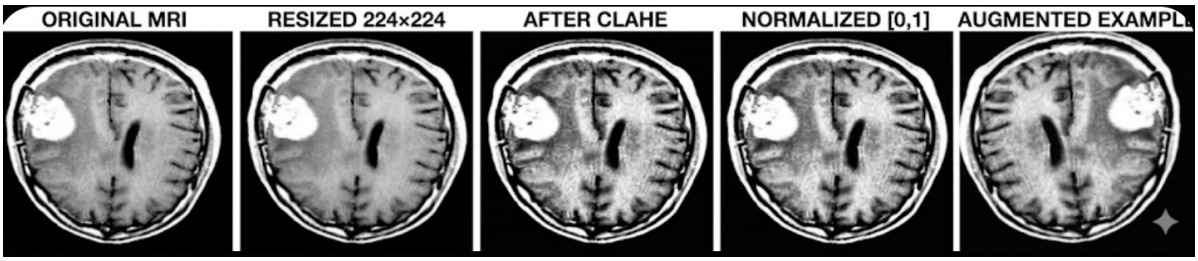


Figure 3.5: Illustration of the Data Preprocessing and Augmentation.

Figure 3.6 illustrates the overall MRI preprocessing workflow, highlighting the role of data augmentation as a final step applied exclusively to the training set to improve generalization and reduce overfitting.

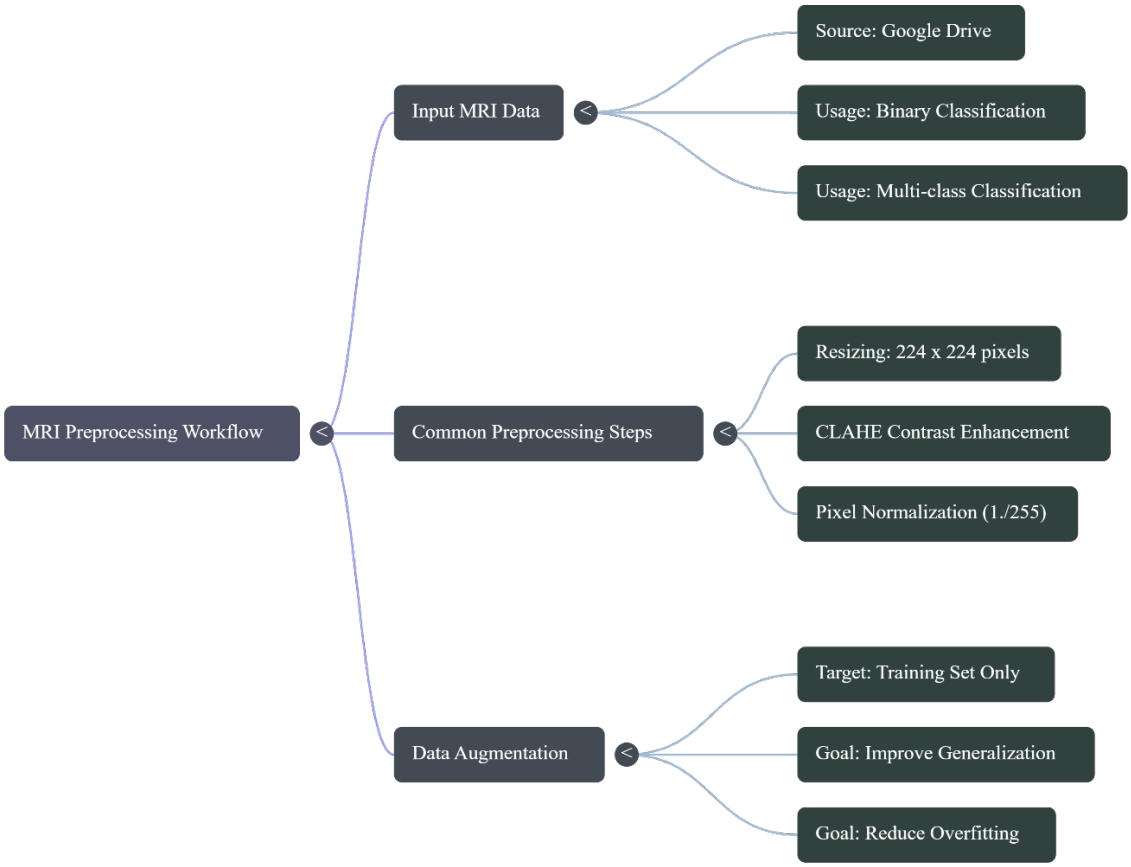


Figure 3.6: Mind Map of the MRI Preprocessing and Augmentation Workflow.

3.2.3 Stage 1 Model Architecture

The first step is to detect whether the brain MRI image has a tumor or not. as it serves as a gating mechanism which filters out the negative tumor cases before any further analyses. This is important to avoid false inputs to the second stage, which helps to reduce the classification error and increase the overall reliability of the system. For this purpose, the binary dataset "mydataset-yes-or-no" is split into training, validation and testing subsets of 70 %, 15% and 15% respectively. The proposed hybrid architecture has two parts a pretrained VGG19 branch and a custom CNN branch, where the VGG19 part learns general visual features from ImageNet-pretrained weights and the CNN branch learns task-specific MRI representations. The last output layer has a sigmoid activation function to give a binary prediction Tumor or No Tumor.

Input

All the MRI scans are resized to a fixed resolution of 224×224 pixels with 3 RGB channels to be consistent with the inputs of the pretrained VGG19 network. This uniform size ensures the compatibility to the pretrained weights and provides the same spatial information to both processing branches.

Branch 1 : VGG19 (Transfer Learning)

The first branch utilizes VGG19, a deep convolutional neural network pretrained on ImageNet a large scale dataset featuring over one million natural images across one thousand categories. By loading the network without its final classification layers (include-top=False), the architecture retains only the feature extraction backbone, leveraging rich, prelearned visual patterns such as edges,textures, and shapes. To prevent severe over fitting a critical risk given that medical imaging datasets are relatively small compared to ImageNet. the backbone weights are explicitly frozen during training and remain unaffected by back propagation. This intentional process retains the general knowledge in VGG19, while offering a stable and computationally-efficient training procedure. This VGG19 is followed by a Global Average Pooling (GAP) layer. This results in a compact and informative feature vector. Instead of flattening the whole feature map into a long vector (which can be millions of parameters and over-fitting), GAP computes the spatial average of each feature map. This implicitly regularizes the model, and improves the generalization ability of the model.

Branch 2 Custom Convolutional Neural Network:

The second branch is a lightweight CNN built from scratch specifically for this task. Its purpose is to learn low level and mid level visual patterns that are unique to brain MRI images and that a general purpose network like VGG19 may not have encountered during its ImageNet training. The branch is composed of three sequential convolutional blocks, each with progressively increasing capacity :

- **Block 1 (32 filters):** Detects basic features such as intensity gradients, edges, and early structural patterns in the MRI scan.
- **Block 2 (64 filters):**Captures more abstract representations such as regional textures and the early shape characteristics of potential tumor regions.
- **Block 3 (128 filters):**Learns high level, task specific features that distinguish tumor from healthy tissue.

The All blocks have the same structure 3x3 convolution to detect local patterns Batch normalization to stabilize training and accelerate convergence by normalizing the activations within each mini batch Max-pooling to reduce the spatial dimensions by half while preserving the most prominent features. The computational costs are reduced by successively halving the sampling rate, and the model becomes insensitive to the precise location of the features, achieving spatial invariance. After the third block, a flattening layer reshapes the multi dimensional feature maps into one dimensional vector to enable seamless concatenation with the output of branch 1.
The two branches process the input image independently. Their feature vectors are concatenated into a unified representation. We prefer concatenation to addition as it keeps the contribution of each branch, instead of forcing an early fusion of the two feature spaces. This allows the classifier to leverage both generic visual features learned from ImageNet-pretrained VGG19 and task-specific MRI features learned by the custom CNN, thus providing a more informative and discriminative representation.

Classifier

The computational costs are reduced by successively halving the sampling rate, and the model becomes insensitive to the precise location of the features, achieving spatial invariance.

Output layer

The neural network for medical image tumor detection is a binary classification problem, with the output layer having one neuron with sigmoid activation function. This neuron predicts the probability of the presence of a tumor and produces values in the range of 0 to 1. The threshold value for classification is 0.5 values above the threshold are labeled as Tumor and values below the threshold are labeled as No Tumor. In this case the sigmoid function is appropriate as the posterior probability of the positive class is modeled directly, without requiring a second output neuron.

Decision Gate

After the first stage prediction is generated predict (image), a decision gate determines whether the image should proceed to the second stage. if the predicted probability is below a predefined threshold, the final output is directly assigned as no tumor and the pipeline stops. However, if the probability is greater than or equal to the threshold, the MRI image is considered tumor-positive and is forwarded to the next classification stage. This logic is consistent with classifier cascade designs, in which only selected cases are passed to later stages for more specialized analysis. The following table 3.1 and figure 3.7 illustrate the architecture of Stage 1.

Component	Layer / Operation	Details
Input	InputLayer	Shape: $(224 \times 224 \times 3)$
Branch 1	VGG19 backbone	ImageNet weights, frozen, `include_top=False`
	GlobalAveragePooling2D	Spatial reduction
Branch 2	Conv2D + BN + MaxPool	32 filters, 3×3, ReLU
	Conv2D + BN + MaxPool	64 filters, 3×3, ReLU
	Conv2D + BN + MaxPool + Flatten	128 filters, 3×3, ReLU
Fusion	Concatenate	Merges Branch 1 + Branch 2
Classifier	Dense + BN + Dropout	256 units, ReLU, Dropout = 0.5
	Dense + Dropout	64 units, ReLU, Dropout = 0.3
Output	Dense	1 unit, Sigmoid
Training	Optimizer	Adam, $lr = 1 \times 10^{-4}$
	Loss	Binary Cross-Entropy
	Metrics	Accuracy, AUC

Table 3.1: Architecture of the proposed Stage 1 Hybrid VGG19-CNN model

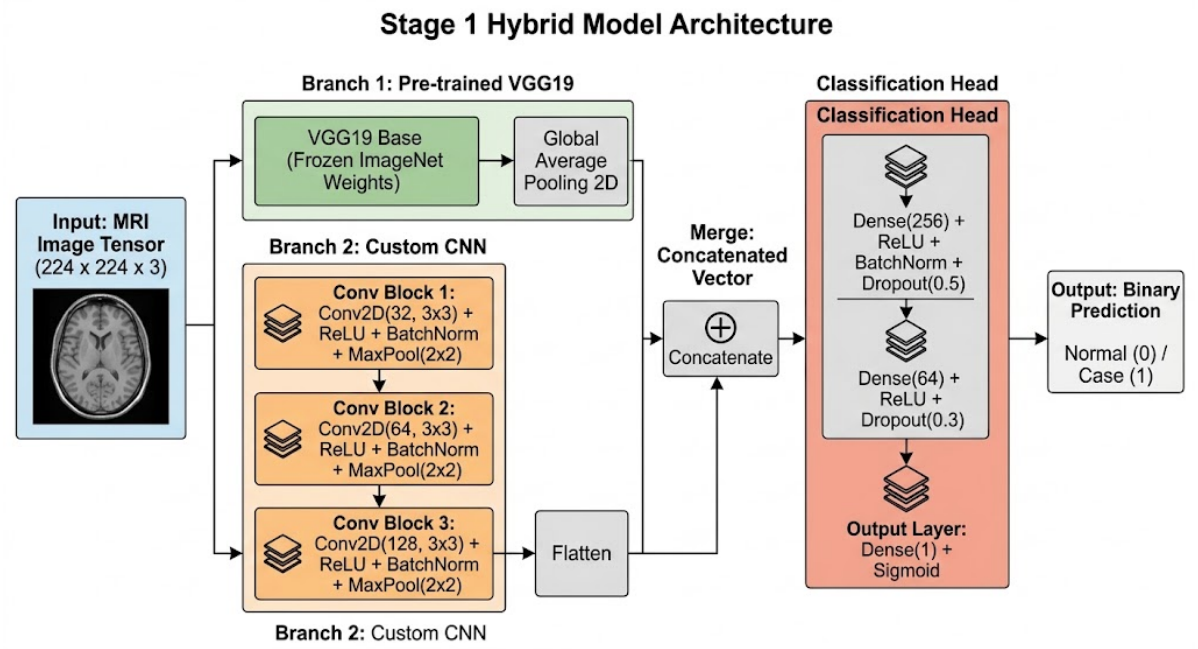


Figure 3.7: architecture of the proposed Stage 1 Hybrid VGG19-CNN model

3.2.4 Stage 2 Model Architecture

Stage 2 determines what kind of tumor it is once Stage 1 has determined that it is a tumor. In this stage the model does not decide anymore if the MRI shows a tumor or not, but rather it classifies the type of tumor. Hence, it is formulated as a multi-class classification problem with three outputs Glioma, Meningioma and Pituitary tumor. The backbone of this stage is ResNet50, a deep convolutional neural network previously trained on ImageNet. We selected ResNet50 as the target network because the residual connection alleviates the vanishing gradient problem in the network, allowing very deep models to train effectively. As a result, the model can learn complex visual patterns more efficiently.

input

The input to the model is an RGB MRI image resized to 224×224×3. This size is used because it is compatible with ResNet50 and preserves enough visual detail for tumor type classification

Resnet50 Backbone

In this study, ResNet50 is used as a backbone feature extractor with the top classification layers removed. The network is pre-trained on ImageNet and weights are frozen during the first stage of training. This allows the new classification head to learn first without changing the pretrained representations. The training=False is also important as it keeps the internal Batch Normalization layers in inference mode, preserving the pretrained statistics.

Dual Pooling

Once the backbone has finished processing the image, two pooling operations are performed in parallel global average pooling summarizes each feature map by calculating its average value, while global max pooling captures the strongest response within it. The two outputs are then merged into a single feature vector. The benefit of this design is that mean pooling captures the overall pattern of the image, while max pooling highlights the most significant response by combining them, the model obtains a richer representation of the MRI image.

Classification head

The fused feature vector is passed through a small fully connected classifier. This network first goes through a dense layer with 512 neurons. It uses ReLU activation, Batch Normalization, Dropout with 0.4 and L2 for weight regularization.
The second layer is also a dense layer. It has 256 units and it uses the same structure. The only difference is that the dropout of 0.4 has been decreased to 0.3 in order to prevent overfitting of the model on the validation set.
The last layer has 3 neurons with Softmax activation for the multi-class classification task to predict one of the three different types of tumors for a given image. Softmax activation is appropriate for this classification task because it is a multi-class problem, where the model is attempting to classify an image into one of three possible tumor types . The use of Batch Normalization helps in training, while Dropout reduces overfitting. The L2 regularization of the weights penalizes high weights. Thus, it increases the models ability to generalize well.

Training Strategy

The Stage 2 model is trained in three phases.

- **Phase 1: Head Training** The first step is to train the new head with the ResNet50 backbone frozen. This allows the classifier layers to adapt and depend on the tumor dataset without updating or changing the pretrained backbone. For this step I used a learning rate (1e-3) so that the new layers can learn fast.

- **Phase 2: Fine-Tuning conv5** The backbone is unfrozen and trained with conv5 block only after the simple decision boundaries are learnt by classification head This step is called fine-tuning and is very specific to personalize the deepest features of ResNet to the medical domain. A lower learning rate of 1e-4 with Cosine Decay is used to make the updates smoother and more controlled.

- **Phase 3: Fine-Tuning conv4 + conv5** Finally, in the last stage both conv4 and conv5 are fine-tuned. This lets the model fine-tune a few earlier layers of features, but retains most of its pretrained knowledge. The learning rate is lowered again to 5e-5 in order to avoid overfitting and allow convergence. The table shows the progressive fine-tuning scheme in Stage 2 with three successive freezing/unfreezing phases.

Phase	**Frozen Layers**	**Trainable Layers**	**Learning Rate**
Phase 1	Full ResNet50	Classifier Head only	10^{-3}
Phase 2	Up to conv4	conv5 + Classifier	10^{-4}
Phase 3	Up to conv3	conv4 + conv5 + Classifier	10^{-5}

Table 3.2: Stage 2 Progressive Fine-Tuning Strategy

The following table 3.3 and figure 3.8 illustrate the architecture of Stage 2

Component	**Layer / Operation**	**Configuration**
Input	RGB MRI Image	$224 \times 224 \times 3$
Backbone	ResNet50	Pre-trained on ImageNet, top removed
Pooling 1	Global Average Pooling	Output: 2048-dim vector
Pooling 2	Global Max Pooling	Output: 2048-dim vector
Fusion	Concatenate (GAP+GMP)	Output: 4096-dim vector
Classifier	Dense(512)	BatchNorm + Dropout(0.4)
	Dense(256)	BatchNorm + Dropout(0.3)
	Dense(3)	Softmax $\rightarrow$ Glioma / Meningioma / Pituitary
Optimizer	**Adam**	**LR** $= 10^{-3} \rightarrow 10^{-4} \rightarrow 10^{-5}$
Loss	**Categorical Cross-Entropy**	**Label smoothing = 0.1**
Metrics	**Accuracy, AUC**	—

Table 3.3: Stage 2: Hybrid ResNet50 Architecture

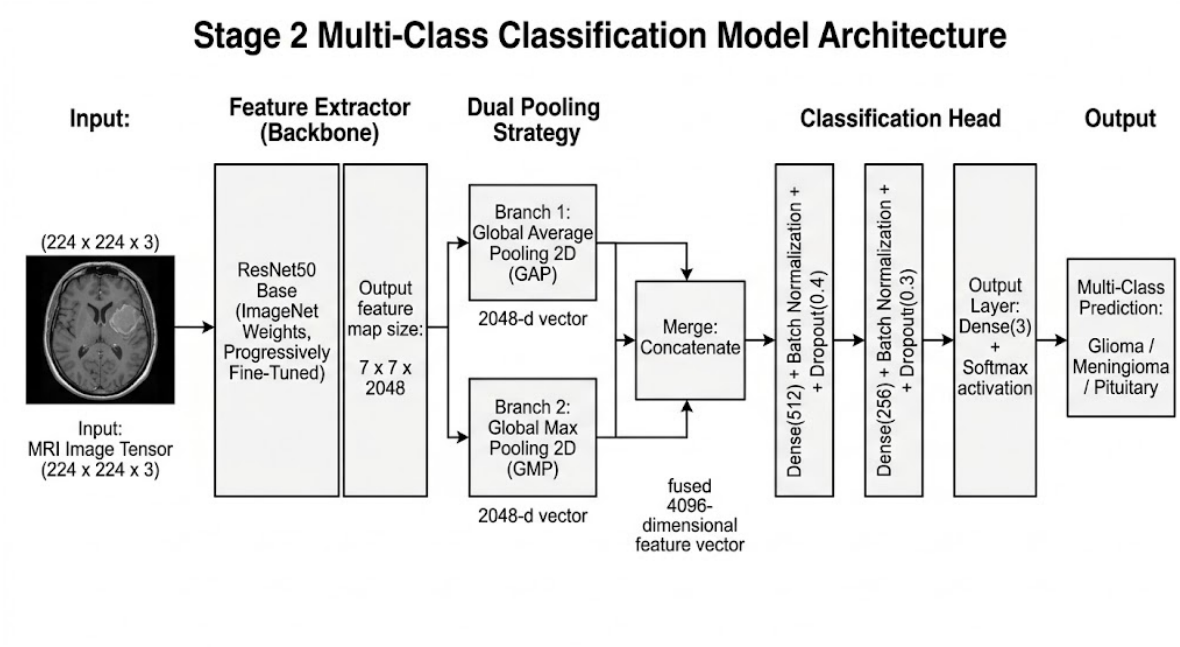


Figure 3.8: Stage 2 Hybrid ResNet50 Architecture

3.3 Conclusion

This chapter presented the datasets, the preprocessing pipeline and the proposed two-stage hybrid framework have been presented. The first stage used a hybrid VGG19-CNN model for binary tumor detection, and the second stage used a progressively fine-tuned ResNet50 model with dual pooling for multi-class tumor subtype classification. Both stages were trained on MRI images resized to 224x224 and augmented with CLAHE and data augmentation for robustness. The first stage was a good detection gate, and the second stage was an accurate tumor subtype classifier. The following chapter describes the experimental setup and model evaluation.

Chapter 4: Results & Discussion

Chapter 4

Results & Discussion

4.1 Introduction

This chapter contains a detailed analysis of the experimental results of the proposed deep learning based brain tumor detection and classification system. The chapter starts with Evaluation Metrics of Models It then describes the dataset splitting strategy flowed by The hardware and software specifications used during the experimentation are discussed. The next section describes the proposed hybrid model which consists of a frozen VGG19 backbone with a custom CNN branch along with CLAHE preprocessing applied to MRI images. This chapter discusses two stages of training namely binary classification (tumor vs no tumor) and multi-class tumor type classification using transfer learning and fine tuning. The performance of the models is evaluated in terms of accuracy, AUC and confusion matrix analysis. The chapter ends with the key findings, limitations identified and future research directions.

4.2 Evaluation Metrics of Models

Model evaluation metrics are quantitative tools that enable researchers to measure the predictive capability, generalization ability, and overall quality of a machine learning model, Selecting appropriate metrics is crucial, as each type of learning task whether classification or regression requires different evaluation criteria. Furthermore, performance metrics facilitate systematic comparison across competing models, allowing practitioners to discard under performing approaches and concentrate optimization efforts on the most promising ones[37].

Accuracy is the ratio of correctly classified instances, positive and negative to total instances in a dataset. It gives a general idea about the

predictive performance of the model[37].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

- **Precision** is the proportion of true positives among all the predicted positives. This is indicative of the model's capacity to minimize false positive errors[37].

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4.2}$$

- **Recall** (Sensitivity) is the ratio of actual positive instances that the model predicted correctly. It measures how well the model can identify all the positive cases[37].

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4.3}$$

- **F1-Score** is the harmonic mean between precision and recall. It provides a single balanced measure that penalizes extremes in either measure. It is especially useful for imbalanced data sets[37].

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.4}$$

- **Specificity** Specificity is the proportion of true negatives that are correctly identified as negative. This measures the ability of the model to correctly reject non-tumor cases[37].

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{4.5}$$

4.3 Split Dataset

The dataset was divided into 3 non-overlapping subsets of 70%, 15% and 15% for training, validation and testing, respectively, using a stratified splitting strategy.This ratio was chosen to have sufficient training data and to reliably evaluate on unseen samples, as is standard practice in deep learning research. This method preserves the class distribution among all subsets and mitigates the bias towards the majority class.

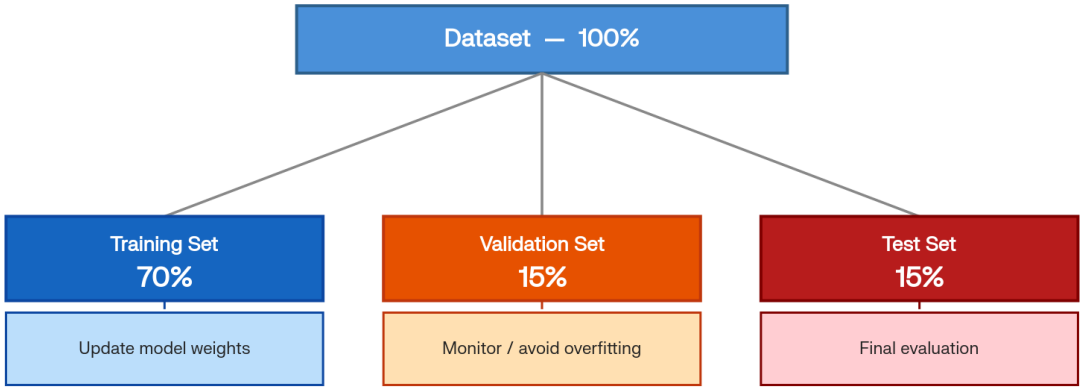


Figure 4.1: Splitting the image database.

4.4 Software Environment

All experiments were done in the **Python** programming language (version 3.10). Python has become extremely popular in the deep learning community owing to its rich ecosystem of scientific and machine learning libraries.

4.4.1 Deep Learning Framework

The models were built and trained using **TensorFlow 2.x** and the **Keras** high level API. TensorFlow is an open source deep learning framework developed by Google to make it easier to build, train and deploy neural network models on different types of hardware including CPUs and GPUs. The model architectures were implemented as an interface in Keras. compiling training configurations, and handling callbacks such as `EarlyStopping`, `ModelCheckpoint`, and `ReduceLROnPlateau`.

4.4.2 Development Environment

All the experiments were carried out on **Google Colaboratory** (Colab), which is a cloud-based Jupyter notebook that provides free GPU resources, including NVIDIA Tesla T4, which significantly accelerates the training of deep convolutional neural networks. Model weights and datasets were stored and managed in **Google Drive**.

4.4.3 Image Processing Libraries

The **OpenCV** (cv2) library was used for image preprocessing (image enhancement) to apply the CLAHE enhancement algorithm to each MRI image before training. **Keras ImageDataGenerator** was used for data loading, augmentation and pipeline management. We were able to do augmentation in real time during training with no additional memory overhead.

4.4.4 Supporting Libraries

Additional scientific libraries used throughout the project include:

- **NumPy:** numerical array operations and preprocessing
- **Matplotlib:** visualization of training curves and evaluation metrics
- **scikit learn:** computation of the confusion matrix, classification report, and ROC-AUC score
- **split folders:** stratified partitioning of the dataset into training, validation, and test subsets

4.5 Experiments results

The Stage 1 model was trained with binary cross entropy loss and accuracy and AUC as evaluation metrics. We used Adam optimizer with learning rate 1e-4, and trained the model with ModelCheckpoint, EarlyStopping and ReduceLROnPlateau. ”best-stage1-model-clahe.h5” was saved as the best performing model.

4.5.1 stage 1 Results

- **Training and Validation Curves:** As shown in Figure 4.2, the training accuracy increases from 73% to 98% while the validation accuracy stabilizes from epoch 30 at 97-98%. The smooth decrease of both loss curves and the lack of any divergence between training and validation loss curves imply that the model generalizes well without overfitting.

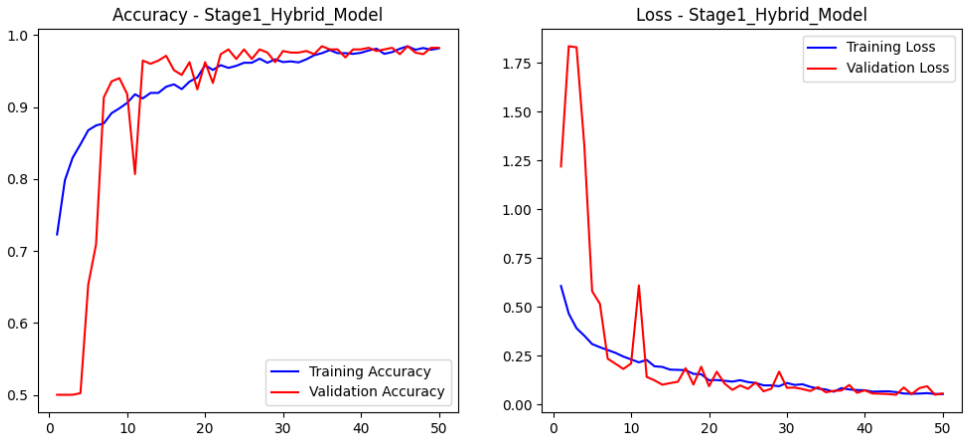


Figure 4.2: Accuracy and Loss Curves of the Stage 1 Hybrid VGG19-CNN.

- **Confusion Matrix:** As shown in figure 4.3 the model correctly classified 224 out of 225 Normal images and 220 out of 225 Case images with only 1 false positive and 5 false negatives. From a clinical point of view a particularly important aspect is the low false negative rate (2.2%) which indicates high sensitivity for tumor detection.

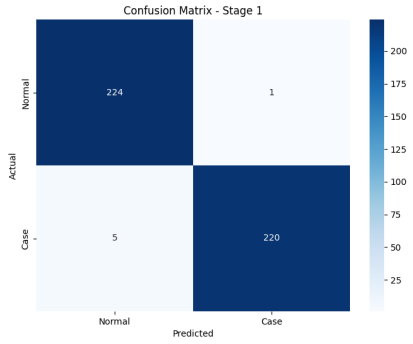


Figure 4.3: Confusion Matrix of the Stage 1 Hybrid VGG19-CNN.

- **Classification Report** as reported in Table 4.1, The overall accuracy of the Stage 1 model was 99% . The perfect recall for Normal (1.00) and perfect precision for Case (1.00) demonstrate its reliability as a binary detection gate, only forwarding confirmed tumor cases to Stage 2 for tumor type classification.

Class	Precision	Recall	F1-Score	Support
Normal	0.98	1.00	0.99	225
Case	1.00	0.98	0.99	225
Overall (Macro Avg)	**0.99**	**0.99**	**0.99**	**450**

Table 4.1: Stage 1: Classification Report (Test Set)

The training of Stage 2 is carried out in three successive phases using a progressive fine-tuning strategy. In the first stage, the ResNet50 backbone is fully frozen and only the newly added classification head is trained with the adam optimize with a learning rate of 10^{-3}.
In the second phase, the last convolutional block (conv5) of ResNet50 is unfrozen and fine-tuned with a learning rate of 10^{-4} to adapt high level features to MRI data. In the third phase, we unfreeze another block (conv4) as well as conv5, and we reduce the learning rate again to 10^{-5} to fine-tune the weights carefully. The loss function is categorical cross-entropy with label smoothing of 0.1 to avoid overconfidence. The best model weights are saved as `best_stage2_resnet_clahe.keras` based on validation accuracy.

4.5.2 stage 2 Results

- **Training and Validation Curves:** The progressive fine-tuning of the Stage 2 model in two steps separated by the dashed vertical line is shown in Figure 4.4. In Phase 1 (Head Training, Epochs 1-28) the training accuracy steadily increases from ∼78% to ∼95%, while the validation accuracy shows marked oscillations, as is expected with the frozen backbone. Transitioning to Phase 2 (Fine-Tuning, Epochs 28–46), both curves become much more stable: training accuracy converges to ∼99.8% and validation accuracy stabilizes at around ∼97–98%. The loss curves show a clear decreasing trend from 1.4 to around 0.4 with no separation between training and validation loss which is a sign of good generalization without overfitting.

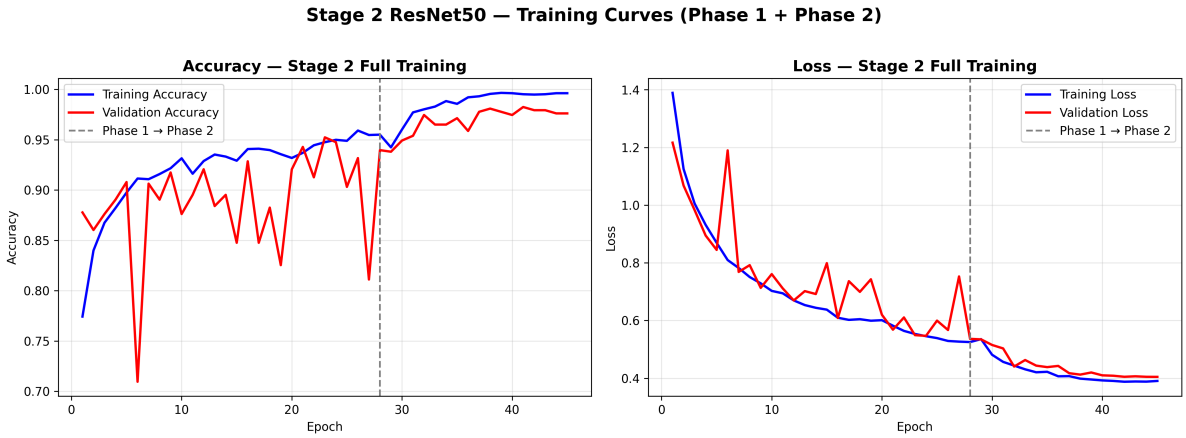


Figure 4.4: Accuracy and Loss Curves of the Stage 2 ResNet50 Model (Phase 1: Head Training; Phase 2: Fine-Tuning).

- **Confusion Matrix:** Figure 4.5, indicates that a significant proportion of the test samples in all three tumor categories are correctly classified by Stage 2 model. In particular, 210 out of 211 Glioma images are classified correctly (1 is misclassified as Meningioma), 205 out of 211 Meningioma images are correctly classified (3 are misclassified as Glioma and 3 as Pituitary). 209 out of 211 Pituitary images are correctly classified (2 are misclassified as Meningioma). The most interesting aspect is the lack of confusion between Glioma and Pituitary, the two most clinically distinct categories, indicating a high level of discrimination.

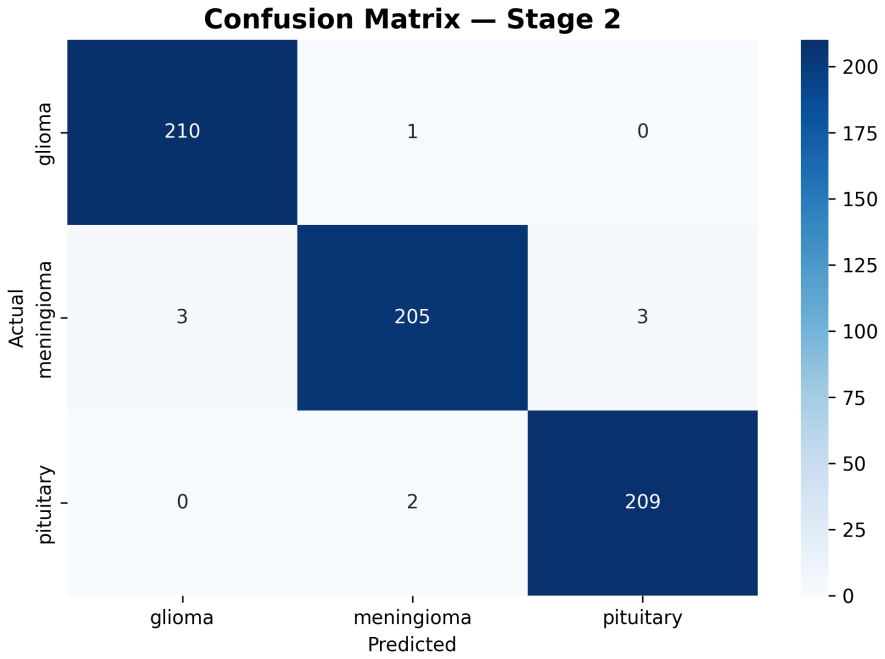


Figure 4.5: Confusion Matrix of the Stage 2 ResNet50 Multi-Class Classifier.

- **Classification Report:** As seen in Figure 4.5, the Stage 2 model correctly classifies a large majority of the test samples in the three tumor categories. Specifically, 210 images of Glioma are correctly identified out of 211 (1 as Meningioma), 205 images of Meningioma are correctly classified out of 211 (3 as Glioma and 3 as Pituitary) and 209 images of Pituitary are correctly identified out of 211 (2 as Meningioma). What is especially interesting is the lack of any confusion between Glioma and Pituitary, the two most clinically distinct categories, indicating a high level of discrimination.

Class	**Precision**	**Recall**	**F1-Score**	**Support**
Glioma	0.99	1.00	0.99	211
Meningioma	0.99	0.97	0.98	211
Pituitary	0.99	0.99	0.99	211
Accuracy	0.99	0.99	0.99	0.99
Macro Avg	0.99	0.99	0.99	633

Table 4.2: Stage 2 Classification Report (Test Set)

The training and validation curves are quite close to each other which means that there is no significant gap in the accuracy or loss. This means that the model has learned meaningful patterns and not memorized the training data. This is further corroborated by the confusion matrix and classification report which show consistently high precision, recall and F1-score for all three tumor classes on the held-out test set confirming that the proposed model generalizes well to unseen data.

To verify the effectiveness of the proposed two-stage pipeline, we train a baseline model under the same hybrid architecture (CNN + VGG19) for a straightforward four-class classification problem, where the network directly predicts glioma, meningioma, no-tumor,and pituitary at once . As illustrated in Figure 4.6 the baseline attained a 91% accuracy in total, with a macro average precision, recall and F1-score of 0.91. The model was also confused between glioma and meningioma, mistaking 42 glioma for meningioma figure see figure 4.7, which made the recall of the glioma a relatively low 0.77. In contrast , the two stage framework proposed here addresses that concern by initially eliminating no-tumor samples, and subsequently with a specific classifier for tumor-type differentiation, which should mitigate inter-class confusion and potentially elevate recall for the utmost

medically relevant classes.

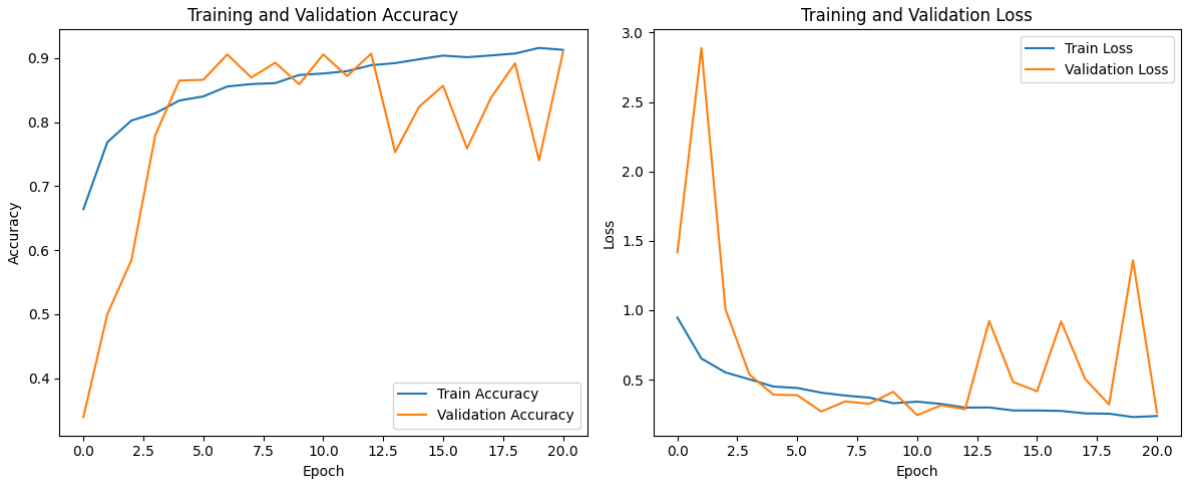


Figure 4.6: Accuracy and Loss Curves of the model Multi-Class Classifier (VGG19-CNN).

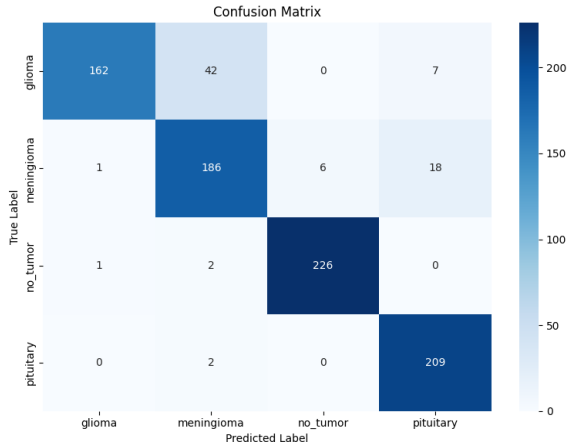


Figure 4.7: Confusion Matrix of the model Multi-Class Classifier (VGG19-CNN).

Table 4.3 summarizes the classification performance of the CNN-VGG19 model across the four brain tumor classes.

Class	Precision	Recall	F1-score	Support
Glioma	0.99	0.77	0.86	211
Meningioma	0.80	0.88	0.84	211
No_tumor	0.97	0.99	0.98	229
Pituitary	0.89	0.99	0.94	211
Accuracy		0.91		862
Macro avg	0.91	0.91	0.91	862

Table 4.3: Classification Report of the CNN-VGG19 Model

The proposed framework is compared with some of the state-of-the-art studies of brain tumor classification in terms of methodology and performance. Nazir et al [38] used the hybrid CNN-VGG16 model and got good performance on the small MRI datasets. Agarwal et al[39] used ResNet18 for feature extraction, GAN for augmentation, and PCA-PSO for selection, achieving a validation accuracy of 98.36%. Rahman et al[32] proposed a parallel CNN structure based on ResNet50 and DenseNet121 and obtained 98.7% accuracy. Ullah et al[34] proposed TumorDetNet, a 48-layer CNN and achieved 99.27% accuracy for three class classification. ZainEldin et al[33] proposed a YOLOv5-based framework, which achieved 99.3% mAP on Figshare dataset. Different from single-stage ResNet50 and VGG19-CNN based methods, our framework follows a two-stage design that first separates abnormal images from normal tissue before tumor sub-typing. This way the confusion between different classes is reduced and the second stage is more concentrated. In Figure 4.8, the proposed framework is compared with studies that use similar architectures to the one used in this work.

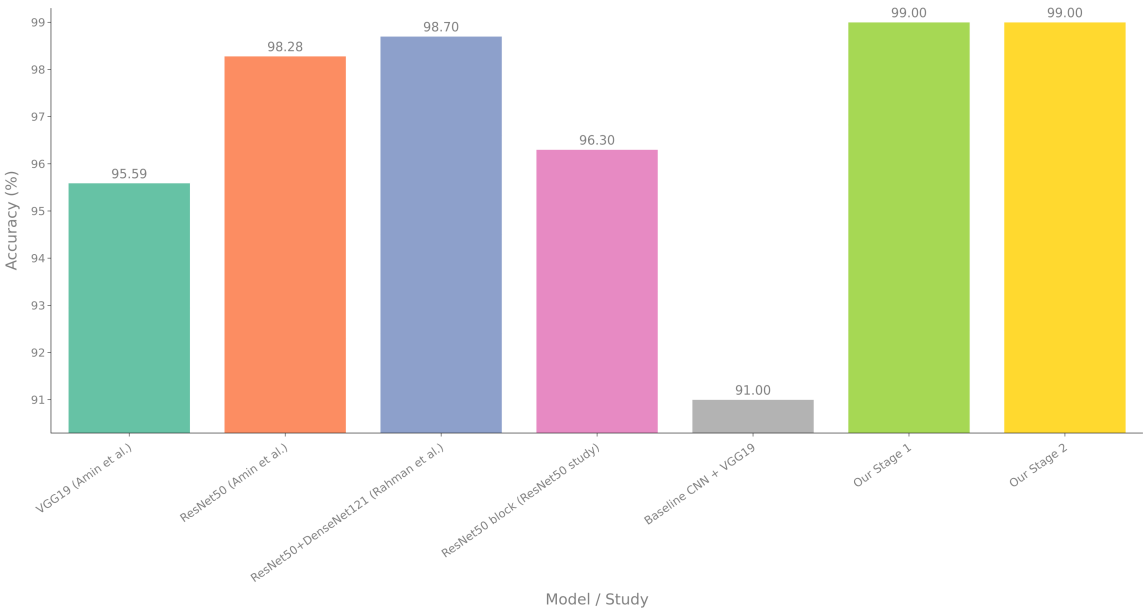


Figure 4.8: Accuracy comparison between prior methods and the proposed framework.

It is worth noting that the methods compared in Figure 4.8 were trained and tested on different datasets with different sizes and acquisition protocols. So this comparison is a general indicator of competitive performance, not a strict quantitative benchmark.

4.6 Conclusion

In this chapter the experimental evaluation of the proposed framework in terms of training behavior, performance metrics and comparison with related studies was provided. The results show that the two-stage design improves the classification process in the first stage to find out the abnormal image and in the second stage to find out the tumor type, thus reducing the inter-class confusion. The proposed model outperformed recent hybrid and transfer-learning approaches providing competitive results especially in the second stage of classification. In summary, the proposed model performs competitively and provides a clinically relevant two-stage decision-making procedure. However, the model's accuracy relies on the quality and heterogeneity of the training data, and further external validation is required to confirm its robustness in real-world clinical settings.

General Conclusion

In this thesis, the problem of automatic brain tumor detection and classification from MRI images with deep learning techniques was researched. In this work, we were motivated by the limitations of manual MRI interpretation, which is time-consuming, subject to inter-observer variability, and highly dependent on clinical expertise.
The main goal was to implement a robust, accurate and clinically interpretable framework able to support radiologists in the diagnostic process. To this end, we proposed a two-stage hybrid deep learning framework.
In the first phase consisted of a binary classification of normal brain Scans and Tumor positive cases with the hybrid CNN-VGG19 model. In the second stage, a fine-tuned ResNet50 model with dual pooling and progressive fine-tuning strategy was used to classify the detected tumors into three subtypes: glioma, meningioma and pituitary tumor. We used CLAHE-based contrast enhancement and data augmentation throughout for better image quality and model generalization.
The experimental results demonstrated that the proposed framework achieved 99% accuracy in stage 1 and 99% in stage 2 with high precision, recall, and F1-score for all tumor classes. The two-stage design performed well in reducing the inter-class confusion with respect to single-stage methods, and the progressive fine-tuning strategy allowed the ResNet50 backbone to be successfully adapted to the medical imaging domain.
These results show that the proposed framework meets the goals established at the start of this work and is a competitive contribution to the field of AI-assisted brain tumor diagnosis.
However, like any other research study, this study has some limitations. The models were trained and tested on publicly available datasets which may not represent the full variability of real clinical settings and further external validation on larger and more diverse data is needed to determine the generalizability of the proposed framework. In addition, this study provided useful practical and scientific learning experiences such as a better understanding of types of brain tumors, MRI interpretation, medical

image analysis using deep learning, and the significance of transfer learning in improving classification performance.
It also helped identify several challenges in implementation such as limited computing resources, hardware constraints and the difficulty of maintaining an efficient working environment for training and experimentation. Future work may include multi-sequence MRI fusion, tumor segmentation and system integration with explainability techniques for real-world clinical deployment.

Bibliography

[1] David N Louis, Arie Perry, Pieter Wesseling, Daniel J Brat, Ian A Cree, Dominique Figarella-Branger, Cynthia Hawkins, Ho-Keung Ng, Stefan M Pfister, Guido Reifenberger, et al. The 2021 world health organization classification of tumors of the central nervous system: a summary. *Neuro-oncology*, 23(8):1231–1251, 2021.

[2] M. F. Alanazi et al. Brain tumor detection using deep learning approaches. *Biomedicines*, 11(2):364, 2023.

[3] National Cancer Institute. Central nervous system – nci dictionary of cancer terms. `https://www.cancer.gov/publications/dictionaries/cancer-terms/def/central-nervous-system`, 2024. Accessed: June 16, 2026.

[4] Radiopaedia. Brain tumors, 2024. Accessed June 15, 2026.

[5] Masoud Nickparvar. Brain tumor MRI dataset. `https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset`, 2021. Accessed: June 2026.

[6] Felix Lugauer and Jens Wetzl. Magnetic resonance imaging. In A. Maier, S. Steidl, V. Christlein, et al., editors, *Medical Imaging Systems: An Introductory Guide [Internet]*, chapter 6. Springer, Cham (CH), Aug 2018. NCBI Bookshelf ID: NBK546152.

[7] B. H. Menze et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024, 2015.

[8] Adrien Lepogam. *Correction des effets de volume partiel en tomographie d'émission*. PhD thesis, Université François-Rabelais de Tours, 2010.

[9] Radiology Masterclass. Mri interpretation: T1 vs t2 images, 2024.

[10] ResearchGate. Partial-volume artifact obscuring hippocampal internal architecture, 2021. Accessed: 2026-06-14.

[11] Tom M. Mitchell. *Machine Learning*. McGraw-Hill, 1997.

[12] Dinggang Shen, Guorong Wu, and Heung-Il Suk. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19:221–248, 2017.

[13] Oracle. What is deep learning?, 2022. Retrieved April 11, 2026.

[14] Alex Sherstinsky. Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. *Physica d: Nonlinear phenomena*, 404:132306, 2020.

[15] GeeksforGeeks. Recurrent neural networks in r, 2024. Accessed: 2026-05-16.

[16] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[17] Shervine Amidi. Cs 230 - convolutional neural networks cheatsheet, 2018. Accessed: 2026-05-16.

[18] MRI Questions. What is convolution?, 2022. Accessed: 2026-05-16.

[19] Giulio Russo. Building smarter neural networks: A guide to pooling, dropout, and normalization techniques, 2024. Accessed: 2026-05-16.

[20] GeeksforGeeks. What is fully connected layer in deep learning. `https://www.geeksforgeeks.org/deep-learning/what-is-fully-connected-layer-in-deep-learning/`, May 2024. Accessed: 2026-05-18.

[21] PhD Elhousieny. Understanding convolutional neural networks (cnns), Feb 2024. Accessed: 2026-05-18.

[22] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *NeurIPS*, 2012.

[23] Siddhesh Bangar. Siddhesh bangar - medium, 2024. Accessed: 2026-05-18.

[24] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

[25] GeeksforGeeks. Residual networks (resnet) - deep learning, May 2026. Accessed: 2026-05-18.

[26] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[27] Karen Simonyan and Andrew Zisserman. Vgg-19 architecture, 2024. Accessed: 2026-05-18.

[28] Chuanqi Tan, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang, and Chunfang Liu. A survey on deep transfer learning. *International Conference on Artificial Neural Networks*, pages 270–279, 2018.

[29] Liora. Transfer learning: What is it? `https://liora.io/en/transfer-learning-what-is-it`, 2024. Accessed: 2026-05-18.

[30] Ginni Garg and Ritu Garg. Brain tumor detection and classification based on hybrid ensemble classifier. *arXiv preprint arXiv:2101.00216*, 2021.

[31] Sandeep Kumar Mathivanan, Saravanan Srinivasan, Manjula Sanjay Koti, Virendra Singh Kushwah, Rose Bindu Joseph, and Mohd Asif Shah. A secure hybrid deep learning framework for brain tumor detection and classification. *Journal of Big Data*, 12(1):72, 2025.

[32] Takowa Rahman and Md Saiful Islam. Mri brain tumor detection and classification using parallel deep convolutional neural networks. *Measurement: Sensors*, 26:100694, 2023.

[33] Hanaa ZainEldin, Samah A Gamel, El-Sayed M El-Kenawy, Amal H Alharbi, Doaa Sami Khafaga, Abdelhameed Ibrahim, and Fatma M Talaat. Brain tumor detection and classification using deep learning and sine-cosine fitness grey wolf optimization. *Bioengineering*, 10(1):18, 2022.

[34] Naeem Ullah, Ali Javed, Ali Alhazmi, Syed M Hasnain, Ali Tahir, and Rehan Ashraf. Tumordetnet: A unified deep learning model for brain tumor detection and classification. *Plos one*, 18(9):e0291200, 2023.

[35] C.V. Singh. Brain tumor mri dataset: 3000 images (1500 bt + 1500 no bt). `https://www.kaggle.com/datasets/cvsingh007/brain-tumor`, 2022. Accessed: June 2026.

[36] C. Rubini, N. Pavithra, Electrical Engineering, M.Kumarasamy, and A. Input Image. Contrast enhancement of mri images using ahe and clahe techniques. *International Journal of Innovative Technology and Exploring Engineering*, 2019.

[37] Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 45(4):427–437, 2009.

[38] Maria Nazir, Sadia Shakil, and Khurram Khurshid. Role of deep learning in brain tumor detection and classification (2015 to 2020): A review. *Computerized medical imaging and graphics*, 91:101940, 2021.

[39] Monika Agarwal, Geeta Rani, Ambeshwar Kumar, Pradeep Kumar, R Manikandan, and Amir H Gandomi. Deep learning for enhanced brain tumor detection and classification. *Results in Engineering*, 22:102117, 2024.