CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Automated Multi-Class Brain Tumor Detection Using Ensemble Deep Learning and Transformer Architectures

by

Insha Fiaz

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Computing
Department of Software Engineering

2026

CERTIFICATE OF APPROVAL

Automated Multi-Class Brain Tumor Detection Using Ensemble Deep Learning and Transformer Architectures

by

Insha Fiaz

(MAI243002)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. M Nouman Noor	FAST, Islamabad
(b)	Internal Examiner	Dr. M Masroor Ahmed	CUST, Islamabad

Dr. Nadeem Anjum
Thesis Supervisor
June, 2026

Dr. Nadeem Anjum
Head
Dept. of Software Engineering
June, 2026

Dr. M. Abdul Qadir
Dean
Faculty of Computing
June, 2026

Author's Declaration

I, **Insha Fiaz** hereby state that my MS thesis titled "**Automated Multi-Class Brain Tumor Detection Using Ensemble Deep Learning and Transformer Architectures**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Insha Fiaz)

Registration No: MAI243002

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Automated Multi-Class Brain Tumor Detection Using Ensemble Deep Learning and Transformer Architectures**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Insha Fiaz)

Registration No: MAI243002

Acknowledgement

I dedicate this thesis to my mother, Shahida Ayub, who always believed in me and never let me give up on my dreams. Her support, strength and constant presence carried me through this journey.

This work is also for my friends Amna Ayub and Mishal Saeed, who stood by me during the most exhausting and overwhelming days. Balancing studies with a demanding job wasn't easy, but their encouragement made it bearable.

Last but not least, I sincerely thank my supervisor, Dr. Nadeem Anjum, for his guidance and support during this research.

(Insha Fiaz)

Abstract

Brain Tumor is one of the most serious neurological disorders in which timely and proper diagnosis has direct implications on treatment and patient survival. The conventional diagnostic processes are based on the manual interpretation of MRI images by radiologists which is time consuming and prone to the error of human interpretation and inconsistency. Although the deep learning approaches have demonstrated potential in analyzing medical images, the current methods are mostly dedicated to the binary tumor detection, which is not a granular approach to classify the various types of tumors, and that is what is most necessary to implement effective medical intervention. The study constructs a multi-class brain tumor classifier through examining various deep learning models. EfficientNetV2 uses advanced convolutional operations to achieve high classification and be computationally efficient. Vision Transformer (ViT) uses self-attention to capture long-range dependencies and global patterns of MRI images. Also, ResNet50 is tested to serve as a convolutional benchmark in order to determine baseline performance. The models have been trained and tested at a dataset of 7,023 brain MRI images. EfficientNetV2 had a 95% accuracy and a loss of 0.13 and F1-score, precision, and recall of 0.96. ViT achieved an accuracy of 90% at a loss of 0.30 and F1-score, precision and recall of 0.89. ResNet50 attained 93% accuracy. The best diagnostic accuracy of 96% was achieved when all trained models' predictions were combined using a geometric mean ensemble approach. The suggested system is better than the current classification systems and offers a strong automated brain tumor diagnostic system. The article contributes to the development of AI in clinical radiology and proves the feasibility of enhancing the accuracy of diagnoses in a medical environment.

Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
CNN	Convolutional Neural Network
CT	Computed Tomography
DL	Deep Learning
FLOPS	Floating Point Operations Per Second
Fused-MBConv	Fused Mobile Inverted Bottleneck Convolution
MBConv	Mobile Inverted Bottleneck Convolution
ML	Machine Learning
MRI	Magnetic Resonance Imaging
MSA	Multi-head Self-Attention
RELU	Rectified Linear Unit
SE	Squeeze-and-Excitation Module
ViT	Vision Transformer
ViT-B/16	Vision Transformer Base model with 16x16 patch size
WHO	World Health Organization

Chapter 1

Introduction

1.1 Overview

Among neurological disorders, brain tumors [1] represent one of the most complex and life-threatening conditions. A brain tumor is an abnormal growth of cells inside or near the brain, which can suppress brain functionality and result in neurological complications.

Brain tumors are generally classified as benign (non-cancerous) or malignant (cancerous) and their impact on brain tissue depends on size, location, growth rate and type.

Diseases of the central nervous system need special attention, even minor abnormalities can be the reason for serious and life threatening problems. Brain tumors rank amongst the most complex and pressing neurosurgery challenges, as they constitute almost 200 different kinds of abnormal tissue formations, which can profoundly affect brain functions and survival rates [2].

Among these, Gliomas (malignant tumor) originating from glial cells, are fast growing tumors, constituting about 80% of all tumor cases. Other brain tumors include meningiomas, which arise from the meninges and pituitary tumors, which develop in the pituitary gland and can disrupt hormonal regulation [3].

1.2 WHO Brain Tumor Classification

According to the World Health Organization Global Cancer Observatory (GLOBOCAN 2022), published by the International Agency for Research on Cancer (IARC) in 2024, approximately 321,624 new brain and CNS tumor cases were reported worldwide that year, resulting in an estimated 248,403 deaths. This reflects a significant rise compared to 308,102 cases recorded in 2020, indicating that the global burden of brain tumors is steadily increasing year by year. These numbers highlight how serious and widespread this disease has become, making early and accurate diagnosis more important than ever.

WHO has categorized the primary brain tumors into four grades (I-IV) according to histological features, behavior of tumor and anticipated clinical prognosis. This evaluation system is the focal point of diagnosis, treatment planning and evaluation of prognosis.

1.2.1 Grade I (Low-Grade, Benign)

Grade I are very gradual growing and are normally confined to a single region in the brain. They can frequently be totally excised in surgery and they do not tend to recur. These tumors occur more frequently in children and young adults and majority of the patients resume normal life after treatment. Examples include Pilocytic astrocytoma, Giant cell astrocytoma that is subependymal, Pituitary tumor and Meningioma.

1.2.2 Grade II (Low-Grade, Infiltrative)

Grade II tumors develop slowly but spread to the surrounding brain tissue making them difficult to excise completely. They have high chances of recurrence and malignant transformation. Common examples include Diffuse astrocytoma, Oligodendroglioma, and Low-grade glioma.

1.2.3 Grade III (High-Grade Malignant)

Grade III tumors grow faster and behave more aggressively. They spread widely in the brain and cannot usually be removed completely with surgery. Treatment usually includes surgery, radiation therapy and chemotherapy. Common examples include Anaplastic oligodendroglioma and High-grade glioma.

1.2.4 Grade IV (Highly Malignant)

Grade IV tumors are the most rapid growing and dangerous brain tumors which can be referred as malignant or cancerous growths. They grow very quickly, damage surrounding brain tissue and are difficult to control even with intensive treatment. Common example include Glioblastoma (glioblastoma multiforme).

1.2.5 Symptoms and Conventional Diagnostic Methods

The symptoms of brain tumors are varied to a large extent by the location, and the rate of growth of the tumor. Common symptoms include nausea and vomiting, vision impairment, speech difficulties, weakness or numbness of the limbs, and balance or coordination difficulty. More often cognitive and behavioral changes are also common including loss of memory, personality changes, as well as inability to concentrate. The symptoms usually manifest slowly and can be increased with the tumor progression and require medical care.

Accurate diagnosis primarily relies on medical imaging which is critical in identifying, localizing and characterizing brain tumors.The two most common imaging modalities are Computed Tomography (CT) scans and Magnetic Resonance Imaging (MRI).

MRI was selected as the imaging modality for this study because it provides superior soft-tissue contrast, higher spatial resolution, and multi-planar imaging capability compared to CT scans [4]. CT Scans expose patients to ionizing radiation

and are primarily used in emergency settings for detecting hemorrhage or acute conditions. For tumor characterization distinguishing tumor type, boundary delineation and assessing infiltration, MRI is the gold standard recommended by clinical guidelines globally. The dataset used in this study is MRI-based, consistent with standard clinical diagnostic practice for brain tumors.

Following its imperativeness, conventional diagnosis based on imaging has a number of weaknesses. A high level of expertise is needed when it comes to MRI interpretation, where radiologists have to inspect a high number of image slices with their own eyes to detect the presence of abnormalities, determine the borders of tumors, as well as assess the nature of tissues. The same process may be time-consuming and relies on the experience of the clinician [5].

In addition, imaging results are frequently subjected to biopsy, which is invasive and can be fraught with such risks like bleeding, infection or neurological complications. Tumors at the early stage or those with faint imaging changes might be hard to detect and hence possibility of late or erroneous diagnosis is likely to occur.

Automation in medical imaging has emerged as a game changing strategy to support clinicians in making diagnostic workflows more consistent, accurate, and improving efficiency. Automated systems are able to handle high volumes of imaging information more quickly and without exhaustion as opposed to the slower and wearisome manual analysis thus minimizing the overall time of diagnosis and the possibility of human error. This is particularly valuable in diagnosis, when it is important to accurately assess the location, size and type of tumor and to plan an effective treatment [6].

The automated imaging systems may carry out a number of significant tasks in relation to the assessment of brain tumors:

i. Abnormal regions detection: Algorithms are able to detect abnormal areas in MRI images, which might signify the existence of a tumor even in the early stages where it is hard to see with human eyes.

ii. Tumor localization and classification: Automation helps to outline the boundaries of tumors, type of tumor and further classify tumor into grades according to the patterns that are learned in a previous dataset.

iii. Elimination of needless procedures: Automated systems may reduce the number of unnecessary biopsies and follow-up imaging, which lowers patient risk and discomfort, as well as unnecessary procedures.

iv. Early detection and intervention: The primary advantage of early detection of subtle or small tumors is that the outcomes of patients improve as an intervention can be undertaken and proper planning of treatment can be made.

Small changes in diagnostic accuracy can make a huge difference in healthcare systems, reducing both the financial consequences of delayed or wrong diagnosis and decreasing the workload of radiologists. Advanced computational techniques, particularly AI domain i.e machine learning (ML) and deep learning (DL), have shown remarkable potential in this domain [7] [8]. Such approaches allow the systems to learn on scale, identify complicated patterns and deliver stable and objective verification, which underpin more effective and efficient clinical decision-making. In general, automation of brain tumor imaging is an important move towards more accurate, efficient and dependable diagnosis towards expert assistance of radiologists, as well as patient care improvement.

1.2.6 Machine Learning Approaches in Brain Tumor Analysis

Machine learning (ML) [9] has extensively applicated in medical imaging work, such as tumor detection, classification and segmentation, serving as an extremely beneficial to clinicians to analyze brain tumors. In traditional ML, mainly handcrafted features derived out of MRI scans including texture, shape, intensity and edge descriptors are used. These characteristics are subsequently inputted into

classifiers such as SVMs, Decision Trees (DTs), Random Forests (RF) or k-NN to either differentiate between tumor and non-tumor regions or grade tumors.

Multiple researches have indicated moderate to high classification rates with the conventional ML technique [10]. These methods have demonstrated moderate success, with classification accuracies typically ranging from approximately 80% to more than 90%, depending on the dataset and feature design used. For example, studies using SVM or Random Forest classifiers have reported accuracies around 85%–90% on standard MRI datasets, underscoring the value of careful feature selection and preprocessing. Though these findings are promising, the effectiveness of the conventional ML models greatly relies on the features, their preprocessing, and the domain knowledge, which restricts their generalizability and implementation into different clinical environments.

The performance of traditional ML models is closely tied to the various factors:

i. Reliance on handcrafted features: the effectiveness of the traditional ML models is strongly associated with the quality and relevance of manually designed features, which entails the high level of expertise and might not reflect all tumor features. Brain tumors have heterogeneous appearances, such as variations in shape, texture and intensity, which are challenging to model with using handcrafted features alone.

ii. Low generalizability: these models frequently do not trained to work well when applied in diverse datasets, MRI scanners, or imaging protocols and thus have low reliability in multi-centric or real world settings.

iii. Noise and variability sensitivity: image quality, patient anatomy and scanning parameters may vary significantly and have a significant impact on model performance.

iv. Computational inefficiency in large volume analysis: feature extraction and preprocessing may be slow, particularly high-resolution 3D MRI volumes.

In spite of the given drawbacks, the conventional ML solutions formed the basis of automated brain tumor analysis, which emphasized the significance of quantitative imaging characteristics in diagnosis and prognosis of brain tumor. They also encouraged deep learning-based solutions [11] to be developed that automatically discover hierarchical features on raw imaging information without a lot of the limitations of handcrafted solutions.

1.2.7 Deep Learning for Brain Tumor Detection

To address these challenges, deep learning methods such as Convolutional Neural Networks (CNNs) [12] [13] [14], have gained popularity in brain tumor analysis. These models are able to learn and extract complex hierarchical features directly on the raw imaging data without having to extract any features by hand.

The common DL-based [15] workflows of brain tumor detection include preprocessing, which could also involve noise removal, normalization, and amplification of tumor-relevant patterns, and the segmentation step, which is used to outline the boundaries of tumors. The process of segmentation [16] is commonly improved with techniques like edge detection, morphological operations etc. Recent models have shown that CNN-based models could obtain a high classification accuracy typically surpassing 95% with certain transfer learning models like ResNet50, InceptionV3 and hybrid models [17] [18] [19] attaining even greater classification accuracy of over 98-99%.

These models have been applied not only in classifying the type of tumors but also in localizing tumor sites and also estimating tumor size [20] [21]. Brain tumor detection frameworks such as YOLOv2, YOLOv3 and Faster R-CNN have been modified to be used in detecting brain tumors and detection of tumors can be performed quickly and accurately using MRI images [22] [23].

In summary, deep learning has revolutionized the diagnosis of brain tumors, and it can be said to be better than traditional machine learning techniques. Nevertheless, the study should be extended with additional research in the field

of early-stage tumor detection, the ability to remain robust to various imaging datasets, and to simplify the process of preprocessing, segmentation and classification pipeline integration [24].

These efforts offer the motivation to the current work that seeks to build a usable, automated and scalable system of identifying brain tumor with respect to MRI scans.

1.2.8 Research Motivation

Brain tumors are among the most complex and serious neurological disorders. Early and correct diagnosis is essential in the successful treatment and better patient outcomes.

MRI is the standard tool for tumor identification due to its high-resolution visualization of brain tissues. Manual interpretation of MRI scans is however time-consuming, subject to error and a part of the radiologists, which may lead to late or inaccurate diagnosis. The current development in the deep learning showed good prospects in the automatization of tumor classification. Different Deep learning models [25] are able to detect complex patterns in MRI images, enhancing the accuracy of the classification process and eliminating the need to use manual analysis. Regardless of these developments, there are still difficulties especially in the ability to differentiate among various types of tumors and guarantee excellent performance under various imaging conditions. This research is motivated by the need to develop an efficient and automated classification framework that enhances diagnostic precision, supports clinical decision-making, and improves patient care outcomes.

1.2.9 Problem Statement

Despite significant progress in deep learning for medical image analysis, many existing brain tumor classification systems still focus on binary detection or rely

on single-model architectures that struggle with accurate multi-class classification. Similar visual patterns in MRI scans make tumor classification challenging, while individual deep learning models often fail to capture both local features and global contextual information effectively. Therefore, an ensemble-based deep learning framework is needed to combine the strengths of multiple models for more accurate and reliable brain tumor classification.

1.2.10 Research Aim and Objectives

The aim of this research is to develop a deep learning framework for multi-class brain tumor classification from MRI images by using convolutional neural networks and vision transformer architectures. Objectives include:

i. To design an efficient deep learning framework for accurate multi-class classification of brain tumors from MRI images.

ii. To enhance classification performance and generalization capability through model architecture and ensemble strategies.

1.2.11 Research Questions

i. How effectively can deep learning models classify brain tumors into Glioma, Meningioma and Pituitary types from MRI images?

ii. To what extent can an ensemble of various architectures improve multiclass brain tumor classification accuracy compared to individual deep learning models on MRI images?

1.2.12 Significance of the Study

This research advances the domain of medical image analysis by introducing an automated deep learning framework for the classification of brain tumors from MRI scans that supports radiologists in routine diagnostic workflows

i. It helps reduce the burden of manual MRI interpretation by providing automated and consistent predictions.

ii. The use of ensemble learning improves reliability by combining the strengths of multiple deep learning models.

iii. The proposed approach can assist in tumor classification, which is essential for best early treatment.

iv. By minimizing diagnostic errors and delays, the frameworkcan contribute to improve patient outcomes and clinical efficiency.

v. In addition, the proposed framework can be extended to support advanced tasks such as tumor segmentation and real-time deployment in healthcare environments.

1.2.13 Thesis Organization

This thesis is organized into six chapters. Chapter 1 introduces the problem domain and research motivation. Chapter 2 reviews related literature. Chapter 3 presents the proposed methodology. Chapter 4 discusses experimental results. Chapter 5 mentions about limitations and constraints while Chapter 6 concludes the research and outlines future directions.

Chapter 2

Literature Review

2.1 Overview

The chapter presents a comprehensive overview of the latest developments in the area of brain tumor detection and classification in terms of medical imaging and deep learning methods. The review studies the metamorphosis of the old-fashioned approaches to diagnostics to the state-of-the-art approaches that are based on artificial intelligence, giving a special focus to the developments since 2022 to 2025. The difficulties of multi-class tumor classification, limitation of datasets, and the issue of continuous concern regarding model generalizability in various imaging conditions and patient populations have particular attention.

A structured literature search was conducted across IEEE Xplore and Google Scholar, targeting peer-reviewed studies on deep learning-based brain tumor detection and multi-class classification using MRI. Papers were included if they clearly described their methods, used MRI data and reported measurable results. Papers were excluded if they used other imaging modalities, lacked sufficient detail or had no original experimental work. After careful screening, more than 70 studies were selected. These studies revealed common gaps in the current methodologies where they lack the diversity in the datasets, problems with class imbalances, as well as

the requirement of clinically validated AI systems that can support radiologists in real-world situations.

The review illustrates that traditional CNN-based methods, although good, frequently have issues with including world-contextual information in medical images, which can be overcome by transformer-based networks.

Besides, the chapter also studies the trade-offs between model complexity, computing needs, and diagnostic performance, indicating that hybrid methods involving CNNs and transformers are a promising way of building a strong multi-class brain tumor classification.

This thorough discussion gives the required background to place the proposed EfficientNetV2 and Vision transformer-based model in the existing body of study, showing how it can overcome the disadvantages that exist by providing optimization in feature extraction, effective computation structure, and combination learning approaches.

Finally, this chapter provides a transition between the literature and the methodological contributions in this thesis, both creating a clinical importance and a technical need of future development of automated brain tumor detection systems.

2.2 Brain Malignancies and Deep Learning Foundation

Lah et al. [26] offer an in-depth analysis of brain malignancies, in terms of glioblastoma or brain metastases. The authors discuss the molecular pathophysiology, clinical manifestation and treatment issues of these aggressive brain tumors. They emphasized the significance of timely detection and accurate classification in improving patient outcomes in the management of brain cancer as glioblastoma is one of the deadliest primary brain tumors whose median survival rates have not yet improved despite improvements in treatment.

The proposed study is a 3D-stacked multistage multiphase inertial microfluidic chip created by Xu et al. [27] to enrich circulating tumor cells in high throughput. The authors investigate new microfluidic technology which makes it possible to separate and enrich CTCs in the blood samples. Their system also exhibits the possibility of detecting and monitoring cancer non-invasively to provide a complementary method to the conventional imaging-based tumor detection methods and can be used in early cancer detection and monitoring treatment.

Ker et al. [28] provide a comprehensive overview of deep learning applications in the analysis of medical images. The authors examine various neural network models, including CNN-based and recurrent architectures, together with their applications across different medical imaging modalities. Their survey shows how modern deep learning approaches have advanced medical imaging by achieving superior performance over conventional machine learning algorithms, with special focus to the automated feature extraction and end-to-end learning features.

Fan et al. [29] present Multiscale Vision Transformers (MViT) to perform video recognition. The authors discuss a hierarchical structure that handles visual data at various scales and thus allows the efficient appearance of fine-grained features as well as coarse features. Although intended as a video understanding framework, their architecture proves the potential of transformer architectures to visual recognition with competitive performance on different benchmarks at reduced computational complexity relative to standard vision transformers.

2.3 Transfer Learning and Attention-Based Architectures

Ahmmed et al. [30] perform a comprehensive study of transfer learning in various types of brain tumor.

Authors discuss different ready to use deep learning models such as VGG16, ResNet50, InceptionV3 and DenseNet that can be used in brain tumor classification. They

show their experimental findings that transfer learning can greatly enhance the accuracy of classification, with ResNet50 having an accuracy rate of 98.7% on multi-class brain tumor data.

The authors focus on the usefulness of using pre-trained models in addressing the scarcity of medical imaging data. Alzahrani [31] proposes ConvAttenMixer, a novel architecture combining convolutional mixers with external and self-attention mechanisms for brain tumor detection and classification. The author discusses the mechanisms of attention that improve feature representation by emphasizing the relevant areas in MRI images.

The model is able to reach 99.17% accuracy when working on the brain tumor classification tasks, which proves the better performance in comparison with the conventional CNN architectures through the integration of attention-based feature refinement. Galic et al. [32] provide a comprehensive review of machine learning frameworks that enable the personalization of medicine based on the analysis of medical images.

The authors discuss several methods such as supervised and unsupervised learning and deep learning methods, which are used in various modalities in medical imaging. Their analysis emphasizes the value of machine learning in facilitating individualized diagnostic and therapeutic plans based on the analysis of patient-specific imaging information, which can help to promote the area of precision medicine in oncology and other medical areas.

Genereux and Akilan [33]develop an efficient CNN-BiLSTM model for multi-class intracranial hemorrhage classification.

The authors explore the combination of convolutional neural networks for spatial feature extraction with bidirectional long short-term memory networks for temporal pattern recognition.

Their model has an accuracy up to 95.3% in classifying various intracranial hemorrhages, indicating that hybrid architectures are efficient in medical image classification problems involving sequential or multi-view data.

2.4 Genomic Analysis and Molecular Biomarkers

Chen et al. [34] study the RNA adenosine alterations concerning the prognosis and immune infiltration of osteosarcoma. The authors discuss the role of RNA modifications in cancer biology and their potential as prognostic biomarkers. Their analysis reveals significant associations between specific RNA modification patterns and patient outcomes, contributing to understanding of molecular mechanisms underlying osteosarcoma progression and potential therapeutic targets. Huang et al. [35] introduce SLNL, a novel method for gene selection and phenotype classification. The authors discuss sparse learning and network-based strategies in order to discover the most important genetic markers of high-dimensional genomic data. Their approach is more effective in feature selection tasks with accuracy of above 90% on various cancer data sets showing it has been able to deal with the curse of dimensionality in genomic studies.

He et al. [36] develop a machine learning framework to trace tumor tissue-of-origin for 13 cancer types based on DNA somatic mutations. The author's study explore random forest and support vector machine classifiers that are trained using mutational signatures. Their framework achieves 88% accuracy in identifying primary tumor sites, which is very useful in the diagnosis of cancer where the location of the primary tumor is unknown, an important factor used in determining which cancer treatment approach should be used.

2.5 Explainable AI and Transformer Models for Medical Imaging

Hussain and Shouno [37] explore an explainable deep learning approach for classification of multi-class brain tumors by Gradient-weighted Class Activation Mapping (Grad-CAM). The authors consider the issue of interpretability in deep learning,

visualizing those areas of MRI images that make classification decisions. They report 97.87% accuracy while providing visual explanations making it good enough and easy to understand model predictions in medical diagnosis. Jun et al. [38] present Medical Transformer, a universal brain encoder for 3D MRI analysis. The authors discuss self-attention mechanism that has been adopted for the volumetric medical imaging data, which is efficient in 3D brain scans. Their transformer architecture has shown excellent results on several neuroimaging tasks that include tumor detection, segmentation and classification, indicating the diversity of transformer models in medical image processing.

2.6 Brain-Computer Interfaces and Neural Signal Processing

Li et al. [39] developed a brain-computer interface that is based on Compact Vision Transformer (CVT) for brain-controlled robot navigation. The authors explore vision transformers for their application in visual stimuli processing in BCI systems. Their approach achieve an accuracy of 92.5% in navigation tasks, which shows how the transformer architectures can be used to improve human-machine interfaces by augmenting visual stimulus classification in real-time brain-controlled interfaces.

Si et al. [40] suggest DBJNet a cross-subject emotion recognition brain-computer interface, which is founded upon functional near-infrared spectroscopy (fNIRS). The authors discuss the domain adaptation methods in order to deal with individual variations in brain signals. Their model achieves 76.8% accuracy in cross-subject emotion classification, which is a step towards generalisable systems of BCI technology that can be applied in a wide range of individuals without extensive calibration.

In their work, Bao et al. [41] examine predicting moral elevation conveyed in danmaku remarks based on the signals of electroencephalography. The authors

address the neural correlates of moral emotions and develop machine learning models to decode them based on brain activity. Their research has an accuracy of 73.2% in predicting moral elevation, which helps to comprehend how emotion works and its possible application in affective computing.

2.7 Residual Networks and Advanced Segmentation Techniques

Kumar et al. [42] present a proposal of a multi-class brain tumor classification model with residual networks and global average pooling. The authors consider deep residual learning networks which solve the vanishing gradient issues of very deep networks. Their model has an accuracy of 97.8% on multi-class brain tumor data sets which reflects the capability of residual connections and global pooling to enhance feature representation and performance in medical images.

Luo et al. [43] develop a medical image segmentation universal model based on prompt-guided transformers. The authors examine the use of natural language prompts to influence the segmentation models to adapt to different anatomical structures and imaging tasks. Their strategy has been able to perform competitively with regard to various segmentation standards showing that prompt-based learning can be used to develop a flexible medical imaging system.

Meshram et al. [44] apply Swin Transformers for brain tumor detection tasks. The authors discuss the hierarchical vision transformers with shifted windows that can facilitate the efficient computation and capture of multi-scale features. Their implementation shows that Swin Transformers can effectively achieve high accuracy in detecting brain tumors while maintaining computational efficiency relative to the conventional CNN systems.

Prathaban et al. [45] explore deep learning techniques of identifying tumor infiltration in diffuse gliomas. The authors discuss convolutional neural networks trained to identify subtle infiltrative patterns in MRI images. Their model achieves 91.3%

sensitivity and 88.7% specificity in detecting tumor infiltration and can be used as a helpful resource in surgical planning and developing a treatment strategy in the treatment of glioma.

2.8 Optimization Strategies and Comprehensive Surveys

Ramakrishnan et al. [46] suggest an optimized hybrid CNN architecture to classify brain tumors by using oneAPI optimization. The authors discuss the balancing accuracy and computational efficiency through hardware-aware optimization.

Their model achieves 98.5% accuracy, 40% lower inference time than unoptimized variants, demonstrating optimization does matter when it comes to deploying a deep learning model to clinical settings.

Shamshad et al. [47] present a comprehensive survey of transformers in medical imaging applications. The authors discuss various transformer architectures such as Vision Transformers (ViT), Swin Transformers and their adaptations for medical image analysis tasks.

Their review includes application in segmentation, classification, reconstruction and registration, hihglighting that transformers are better at capturing long-range interactions and global context, than CNNs.

Rastogi et al. [48] develop a multi-class system of brain tumor classification based on multi-branch networks and Inception blocks with a five-fold cross-validation.

The authors investigate parallel convolutational processes that extract features at various scale levels in parallel. Their framework has an accuracy of 99.3% in classifying brain tumor MRI samples, which proves the usefulness of using multi-scale feature extraction with Inception modules.

2.9 Generative Models and Uncertainty Quantification

Liu et al. [49] present the concept of virtual formalin-fixed and paraffin-embedded staining of fresh brain tissue with stimulated Raman imaging and CycleGAN models. The authors explore the use of generative adversarial networks to translate label-free Raman images to virtually stained histopathology images. Their technique achieves high-quality virtual staining that is quite similar to conventional histology, potentially enabling rapid intraoperative diagnosis without physical tissues processing.

The article by Hu et al. [50] offers a reliable multi-stage liver tumor segmentation model using evidence-based uncertainty quantification The authors explore evidential deep learning to estimate prediction uncertainty and improve segmentation reliability. Their model achieves 89.2% Dice coefficient while providing calibrated uncertainty estimates, which increases clinical credibility by determining situations in which model results may be unreliable. Tan and Liang, [51] come up with a multiclass malaria parasite recognizer system, which is based on transformer models and generative adversarial networks. The authors explore vision transformers for microscopic images classification combined with GANs for data augmentation. Their approach achieves 98.1% accuracy of identifying various species of malaria parasites, showing the transformers to be useful in the analysis of medical images.

2.10 AI Paradigm Comparisons and Hybrid Networks

According to Tandel et al. [52], multiclass magnetic resonance imaging brain tumor classification system is proposed using artificial intelligence paradigms. The authors discuss the different machine learning and deep learning methods such as

support vector machine, random forest and convolutional neural networks. The results of their overall comparison reveal CNNs with the highest accuracy of 97.5, establishing deep learning superiority in complex medical image classification tasks.

The authors Zheng et al. [53] present a unified visualization and classification model for grade glioma using MRI. The authors examine joint learning of segmentation and classification tasks. Their model has a 94.6% accuracy of classifying glioma while providing interpretable visualizations of tumor regions, so the clinician can understand and verify model predictions. Zhou et al. [54] develop a method for generating 3D brain tumor regions in MRI using vector-quantization generative adversarial networks. The authors discuss discrete latent representations for high-quality 3D medical images. Their approach generates realistic tumor regions can augment training datasets, addressing data scarcity issues and improving model generalization.

Parshapa and Rani [55] present a survey on the effective identification and analysis of brain tumor diagnosis using machine learning. The authors discuss various algorithms such as decision trees, SVMs and neural networks in the identification and classification of tumors. According to their review, what has changed with time is the fact that the traditional machine learning is now replaced with deep learning methods, which are more accurate and automated in the diagnosis of brain tumors.

2.11 Advanced Transformer Architectures and Diffusion Models

Zhu and Wang [56] demonstrates a thorough review of transformers and their usage in medical image processing.

The authors discuss transformer architecture that can be applicable to several medical imaging tasks such as classification, segmentation, detection and reconstruction. Their review shows that transformers can capture global context and

long-range dependencies and achieving state-of-the-art performance on various medical imaging benchmarks.

A dual-domain diffusion model that is suggested by Yang et al. [57] supports sparse-view CT reconstruction. The authors discuss diffusion probabilistic models, which can work in both image and sinogram space to synthesize high-quality CT images with limited projections.

Their algorithm attains a PSNR value of 34.2 dB during reconstruction challenges, which shows that diffusion models can solve ill-posed inverse tasks in medical imaging.

Zheng et al. [58] develop a lightweight transformer image feature extraction network to obtain an efficient visual recognition. The authors consider parameter-efficient transformer design that do not impair the performance but with lower computational needs. Their network achieves 89.4% accuracy on image classification benchmarks using 60% less parameters than standard transformers, and thus can be used in resource-constrained medical imaging applications.

Liu et al. [59] provide the critically detailed survey of visual transformers covering architectural innovations and applications. The authors examine diverse forms of transformers such as hierarchical transformers, efficient attention mechanisms, and hybrid CNN-transformers. Their review systematically categorizes transformer developments and discusses their impact on computer vision tasks, providing insights applicable to medical image analysis.

2.12 CNN Optimization and Ensemble Methods

Irmak [60] presents a multi-classification mechanism of brain tumor MRI image with deep convolutional neural networks. The author explores the methods of hyperparameter optimization such as batch normalization, and dropout regularization. The optimized model obtains 96.05% accuracy in multi-class brain

tumor classification, which proves that systematic optimization is significant to the performance of deep learning.

Sadad et al. [61] build brain tumor detection and multi-classification models based on the advanced use of deep learning methods. The authors explore ensemble methods combining multiple CNN models such as VGG, ResNet, and DenseNet. Their ensemble approach achieves 98.69% accuracy on brain tumor classification task, how model combination can improve robustness and performance beyond individual models.

Ben Naceur et al. [62] propose a brain tumor segmentation system using deep learning based selective attention with overlapping patches and multi class weighted cross-entropy.

The authors discuss the attention mechanisms to concentrate on the tumor regions while handling class imbalance through weighted loss functions. Their model achieves Dice scores of 0.88, 0.83, and 0.78 for whole tumor, tumor core, and enhancing tumor respectively. Muhammad et al. [63] presents a perspectives survey about deep learning in multigrades brain tumor classification.

The authors discuss various deep learning architecture, data augmentation techniques, and transfer learning approaches for brain tumor grading. Their comprehensive review discusses challenges including class imbalance, and model interpretability, giving the directions of future studies in intelligent healthcare systems.

2.13 Hyperparameter Tuning and Vision Transformers

Kumar and Mankame [64] establish an optimization based deep convolutional neural network for brain tumor classification.

The authors discuss genetic algorithm-based optimization for architecture search and hyperparameter tuning. Their optimized model has 97.1% accuracy rate in

classifying brain tumors, which proves the use of evolutionary algorithms in automating neural architecture design in medical imaging tasks.

Asiri et al. [65] propose an enhanced brain tumor diagnosing system with the optimized CNN hyperparameters to enhance the accuracy and reliability. The authors address the topic of Bayesian for systematic hyperparameter search across learning rates, batch sizes, and network depths. Their optimized model achieves 99.32% accuracy with improved generalization on the brain tumor classification tasks which highlights the importance of hyperparameter tuning in clinical deployment.

The research by Asiri et al. [66] presents a comparative study of the fine-tuned vision transformers using pre-trained models to classify brain tumors. The authors explore different versions of ViT such as ViT-B/16, ViT-L/16 and DeiT and compare them in terms of their results on brain tumor datasets. Their findings indicate that ViT-L/16 has reached 99.5% accuracy, demonstrating vision transformers' superior performance compared to traditional CNNs through effective transfer learning from large-scale datasets.

2.14 GANs, Graph Networks and Meta-Analysis

Asiri et al. [67] pioneer next-generation brain tumor classification using deep learning and fine-tuned conditional generative adversarial networks.

The authors explore cGANs for generating synthetic brain tumor images to augment training datasets, addressing class imbalance.

Their combined approach of data augmentation and deep classification achieves 99.17% accuracy, demonstrating how generative models can enhance supervised learning in limited data scenarios.

Gürsoy and Kaya [68] present a graph convolutional neural network image brain tumor detector named Brain-GCN-net.

The authors discuss the representation of brain MRI images in the form of graphs where nodes represent image regions and edges encode spatial relationships.

Their GCN-based model attains an accuracy of 98.74% in brain tumor classification, which proves the possibility of the graph-based representation to capture structural patterns in medical images.

In an article by Alzahrani et al. [69], an amyotrophic lateral sclerosis prediction framework is developed on the basis of multi-level encoders-decoders based ensemble architecture.

The authors explore hierarchical encoder-decoder networks that process medical data at multiple resolutions.

Their ensemble ables to predict ALS with 94.3% accuracy and demonstrating the effectiveness of multi-scale processing in neurological disease classification.

In the study conducted by Wang et al. [70], a systematic review and meta-analysis of artificial intelligence-based brain tumor detection and segmentation models are carried out in the context of magnetic resonance imaging.

The authors explore the performance measures in 89 studies and discuss aspects that influence model performance.

Their meta-analysis demonstrates the brain tumor pooled sensitivity of 87% and specificity of 85% for brain tumor detection, highlighting variability in performance and need for standardization in AI medical imaging research.

Moya-Sáez et al. [71] propose brain tumor enhancement prediction based on pre-contrast traditional weighted images using the synthetic multiparametric mapping and generative AI.

The authors discuss the use of generative models for synthesizing contrast-enhanced images out of non-contrast scans.

Their method yields Dice coefficient of 0.78 in predicting tumor enhancement patterns, potentially reducing need for contrast agent administration in certain

clinical scenarios. Sravani et al. [72] develop methods for enhancing brain tumor diagnosis using generative adversarial networks. The authors discuss StyleGAN and CycleGAN to be applied in data augmentation and image-to-image translation in brain MRI.

Their GAN-based augmentation improves classification accuracy by 3.5%, which shows how synthetic data generation can mitigate limited training data challeng data challenges in medical imaging.

Table 2.1 provides a comparative analysis that highlights how this research combines disparate capabilities into a deployable framework, specifically tailored to the tumor detection, to bridge a clear gap in the existing literature.

TABLE 2.1: Performance Comparison of CNN, LSTM and BiLSTM Models

Author(s) & Year	Methodology	Accuracy	Key Innovation	Limitation
Ahmmed et al. (2023)	ResNet50 Transfer Learning Brain tumor classification on MRI	98.7%	Comprehensive transfer learning comparison across VGG16, ResNet50, InceptionV3, DenseNet	Limited to 2D slices; requires large pre-trained datasets
Alzahrani (2023)	ConvAttenMixer Brain tumor detection & classification on MRI	99.17%	Novel combination of convolutional mixers with attention mechanisms	High computational cost; attention mechanisms increase model complexity

TABLE 2.1: Performance Comparison (Continued from previous page)

Author(s) & Year	Methodology	Accuracy	Key Innovation	Limitation
Genereux & Akilan (2023)	CNN-BiLSTM Intracranial hemorrhage classification on CT	95.3%	Hybrid spatial-temporal architecture	Sequential processing increases inference time; limited to temporal data
Hussain & Shouno (2023)	CNN + Grad-CAM Multi-class brain tumor classification on MRI	97.87%	Explainable AI with visual interpretability	Grad-CAM visualizations may not capture all relevant features
Kumar et al. (2021)	ResNet + Global Average Pooling Multi-class brain tumor classification on MRI	97.8%	Addresses vanishing gradient with residual connections	Deep architecture requires significant training resources
Ramakrishnan et al. (2024)	Optimized Hybrid CNN (oneAPI) Brain tumor classification on MRI	98.5% (40% faster)	Hardware-aware optimization for clinical deployment	Platform-specific optimization; may not generalize to other hardware

TABLE 2.1: Performance Comparison (Continued from previous page)

Author(s) & Year	Methodology	Accuracy	Key Innovation	Limitation
Rastogi et al. (2024)	Multi-branch Inception Network Multi-class brain tumor classification on MRI	99.3%	Multi-scale parallel feature extraction with 5-fold CV	Complex multi-branch architecture; high memory requirements
Tan & Liang (2023)	Transformer + GAN Malaria parasite classification on microscopy	98.1%	Combines vision transformers with GAN augmentation	Limited to microscopy images; GAN training instability
Tandel et al. (2020)	CNN vs ML Paradigms Multi-class brain tumor classification on MRI	97.5% (CNN best)	Comprehensive AI paradigm comparison study	Traditional ML methods significantly underperform
Irmak (2021)	Fully Optimized Deep CNN Multi-class brain tumor classification on MRI	96.05%	Systematic hyperparameter optimization framework	Time-consuming optimization process; overfitting risk

Chapter 3

Research Methodology

3.1 Introduction

This chapter presents a comprehensive methodology for automated multi-class brain tumor detection using ensemble deep learning models and transformer architectures.

The proposed framework trains EfficientNetV2, Vision Transformer (ViT-B/16) and ResNet-50 independently, followed by an ensemble approach integrating all three models to improve classification accuracy and compares the ensemble performance with individual standalone models for evaluation.

By leveraging the strengths of convolutional neural networks and transformer-based representations, the methodology aims to achieve accurate and reliable classification of brain tumors from MRI scans.

The approach to the methodology is organized in a systematic way, starting with the description of the dataset and the preprocessing pipeline, then continuing with a systematic description of the architecture with mathematical formulas, ensemble strategy, training processes and evaluation metrices.

3.2 Proposed Framework

The proposed system operates through a sequence of stages MRI images are first obtained and preprocessed to have uniformity and use as model input. The proposed framework exploits the complementary strengths of the employed architectures. EfficientNetV2 emphasizes computational efficiency and effective feature extraction, while Vision Transformer (ViT-B/16) models long-range dependencies by capturing global contextual information through self-attention mechanisms. ResNet-50 provides deep hierarchical feature representations through residual learning.

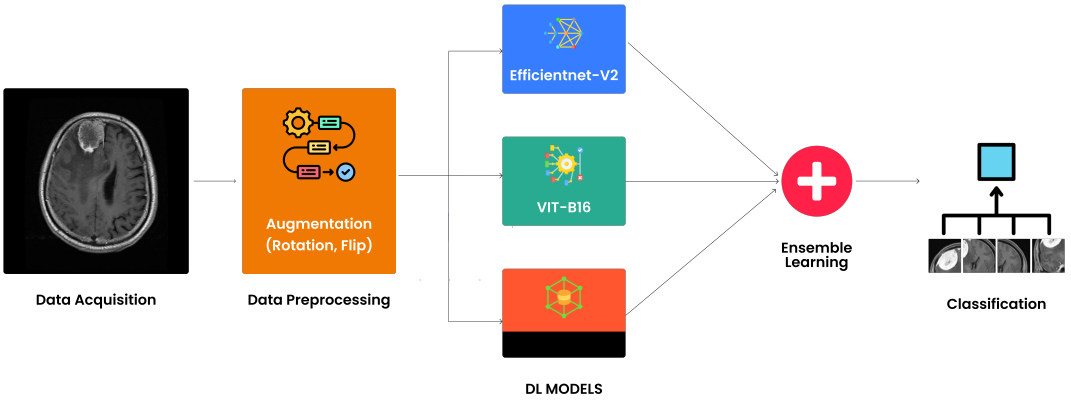

FIGURE 3.1: Block Diagram of Proposed Research

These three models are first trained independently and then integrated through an ensemble layer that combines their predictions using averaging to generate the final ensemble output. The ensemble results are then evaluated against the performance of each individual standalone model for comparative performance evaluation. Figure 3.1 shows the block diagram of the system architecture, highlighting the major functional components and their interconnections within the overall framework.

3.3 Dataset Description

This study employs a publicly available Brain Tumor MRI dataset consisting of 7,023 MRI scans, obtained from Kaggle. The dataset is curated by integrating

images from three widely used and reliable sources: Figshare, SARTAJ Brain Tumor Dataset and Br35H.

The MRI images are clustered into four clinically pertinent categories as shown in Figure 3.2:, which allow the multi-class recognition of the brain tumor instead of having the task performed by the conventional binary detection. Every class signifies a particular type of tumor or normal condition, which is differentiated by the cell-origin, anatomical localization, malignancy and visualization in MRI images.

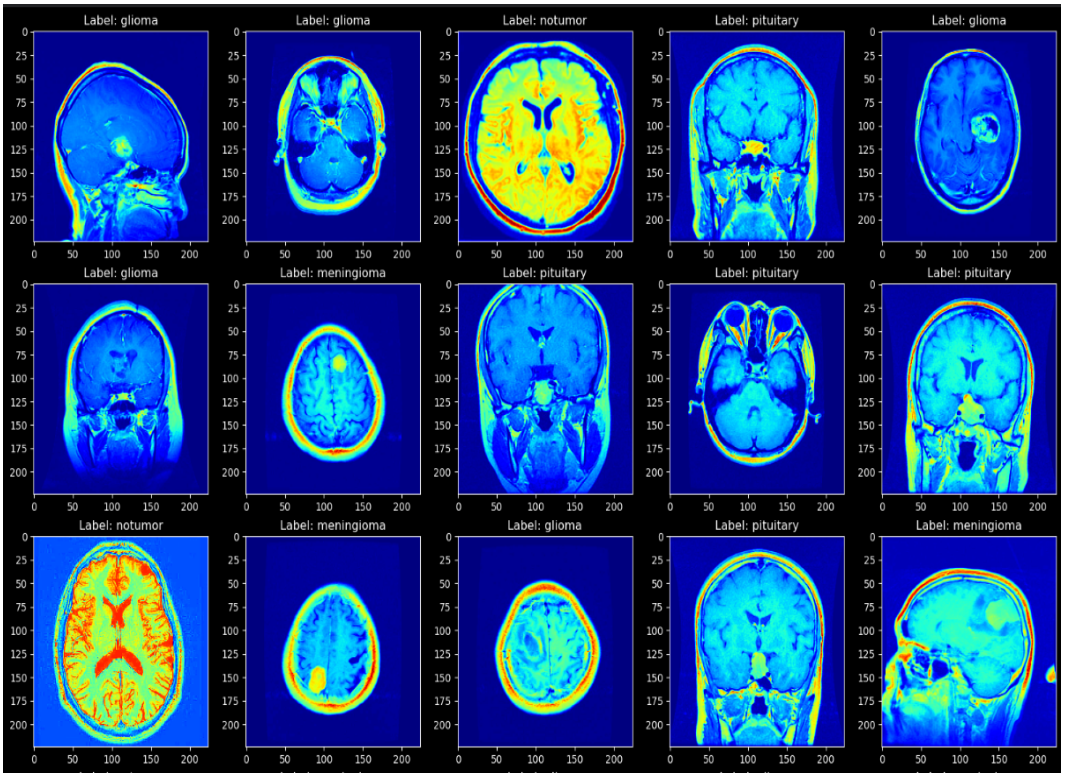

FIGURE 3.2: Dataset Distribution

Although the BraTS (Brain Tumor Segmentation) dataset is widely used in the literature, it is primarily designed for segmentation tasks on 3D multimodal MRI volumes, making it less suitable for the classification focused objective of this research. The Kaggle Brain Tumor MRI dataset used in this study consists of 7,023 labeled 2D MRI images categorized into four clinically relevant classes (glioma, meningioma, pituitary, no tumor), which directly supports multi-class classification evaluation. Additionally, this dataset is a combination of multiple verified

sources (Figshare, SARTAJ, Br35H), providing reasonable diversity while maintaining class balance, a key requirement for fair model training and evaluation.

3.4 Dataset Overview

3.4.1 Class Distribution

The dataset contains four distinct classes such as Glioma Class: Glioma tumors originate from glial cells, that are supportive in the central nervous system. These are malignant and very aggressive tumors as they cause almost 80 percent primary malignant tumor of the brain. Gliomas tend to develop into brain tissue around it due to their invasive growth patterns; hence, they are not easy to delimit. Gliomas in MRI scans are characterised by irregular and not well-defined margins, the heterogeneous signal intensity, and by different degrees of edema and mass effect. Accurate classification of gliomas is clinically critical because of their rapid progression and poor prognosis.

Meningioma Class: Meningiomas develop as a result of the meninges which are the protective layers of the brain and the spinal cord. They are generally benign and slow-growing, although malignant variants may occur in rare cases. Their distinct imaging features and non-invasive nature of behavior make it essential to differentiate them from more aggressive intra-axial tumors such as gliomas.

Pituitary Class: These tumors grow out of the pituitary gland which is an important endocrine gland within the sellar or suprasellar region of the brain. These are benign adenoma tumors, which may cause severe clinical outcomes of hormonal imbalance, loss of vision, and neurological symptoms as a result of compression of adjacent structures, such as the optic chiasm. MRI images of pituitary tumors are usually characterized by some local lesions and the anatomical localization of these foci, which is also important in making the correct diagnosis.

No Tumor Class: The No Tumor category incorporates the MRI scans of healthy brains that have no tumor or abnormalities. These images make the model understand the distinction between the normal brain structures and the tumour regions. Normal scans should be included to ensure that the model does not confuse the normal tissue with a tumor and to enhance the classification system in general.

This study will carry out multi-class classification in contrast to the many in existence researches where they simply carry out binary tumor versus non-tumor classification. Types of brain tumors differ significantly with regard to cell type of origin, the degree of malignancy, therapeutic approaches and prognosis. As such, specific differentiation of glioma, meningioma, pituitary tumors and normal brain scans will contribute to greater clinically significant information and can assist in the making of better diagnostic decisions.

3.4.2 Dataset Split Strategy

The dataset consists of 7,023 Brain MRI scans, which are provided in two predefined subsets: training and testing. It is divided into 80:20 train test split, which leaves 5,712 images in training and 1,311 images in testing. This split aligns with standard practices in deep learning–based medical image analysis and ensures sufficient data for effective model training while preserving an independent test set for performance and unbiased evaluation.

The data used is well balanced in all four categories of glioma, meningioma, pituitary tumor and no tumor which enables the model to learn discriminative features of each group. Table 3.1 and Figure 3.3 summarizes the detailed allocation of the MRI scans in class in the training and testing sets.

This structured data partitioning ensures that the proposed deep learning framework is trained, validated and evaluated under realistic conditions. By maintaining class balance and strict separation between training, validation and testing phases, the dataset setup supports reliable performance analysis and enhances the clinical relevance of the reported results.

TABLE 3.1: MRI Dataset Distribution

Class	Training Images	Testing Images	Total	Percentage
Glioma	1,321	300	1,621	23.1%
Meningioma	1,339	306	1,645	23.4%
Pituitary	1,457	300	1,757	25.0%
No Tumor	1,595	405	2,000	28.5%
Total	5,712	1,311	7,023	100%

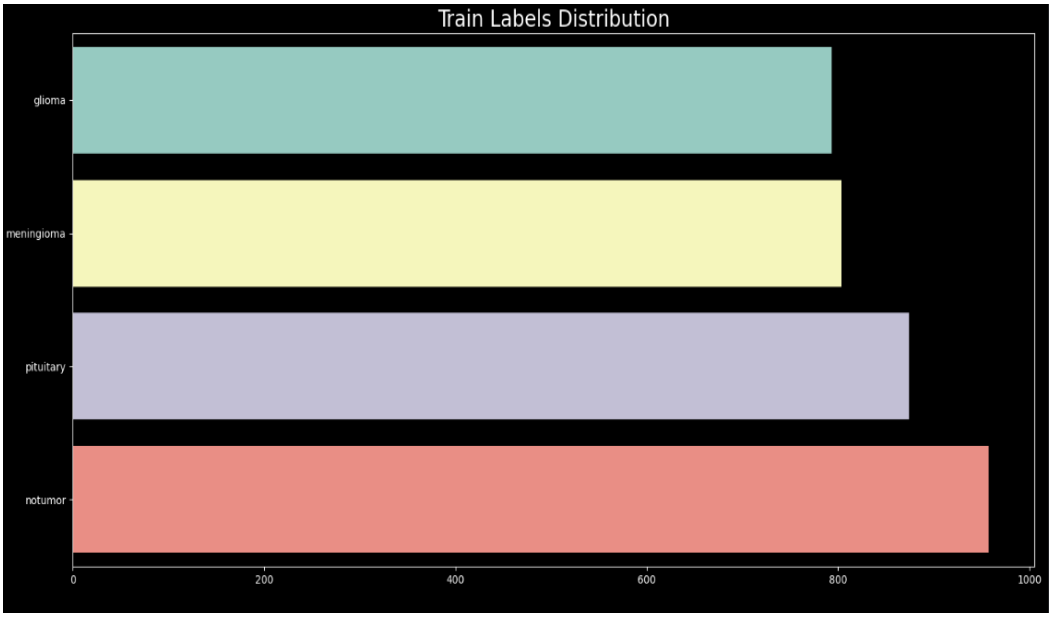

FIGURE 3.3: Training Dataset

3.4.3 Data Preprocessing

To have a high quality, clean and standardized MRI images, data preprocessing was conducted to guarantee successful training of the deep learning model. In medical image analysis, proper preprocessing is necessary, which would make the model more robust, increase the generalization ability and decrease the chances of overfitting. Data augmentation methods as shown in Figure 3.4 were used to enhance the heterogeneity of the dataset and enhance the extrapolation of the models. Augmentation artificially increases the size of the training dataset by adding manipulated differences to the images, which allows the model to learn the

invariant and discriminative features. All MRI images were resized and standardized to 224 × 224 × 3 (height × width × RGB channels) prior to model training. This resolution is consistent with the input requirements of architectures, ensures uniform input dimensions across all models. The augmentation techniques used were as follows:

i. Rotation: The images were rotated in diverse angles to present the model with the tumors in new positions.

ii. Horizontal and Vertical Flipping: Flipping operations imitated changes in patient position and scanner views.

iii. Scaling and Zooming: These enhancements helped the model to detect tumors of different sizes and spatial distribution.

iv. Brightness Adjustment: There was a slight change in intensity variation that was added to compensate discrepancies among MRI acquisition conditions and imaging devices.

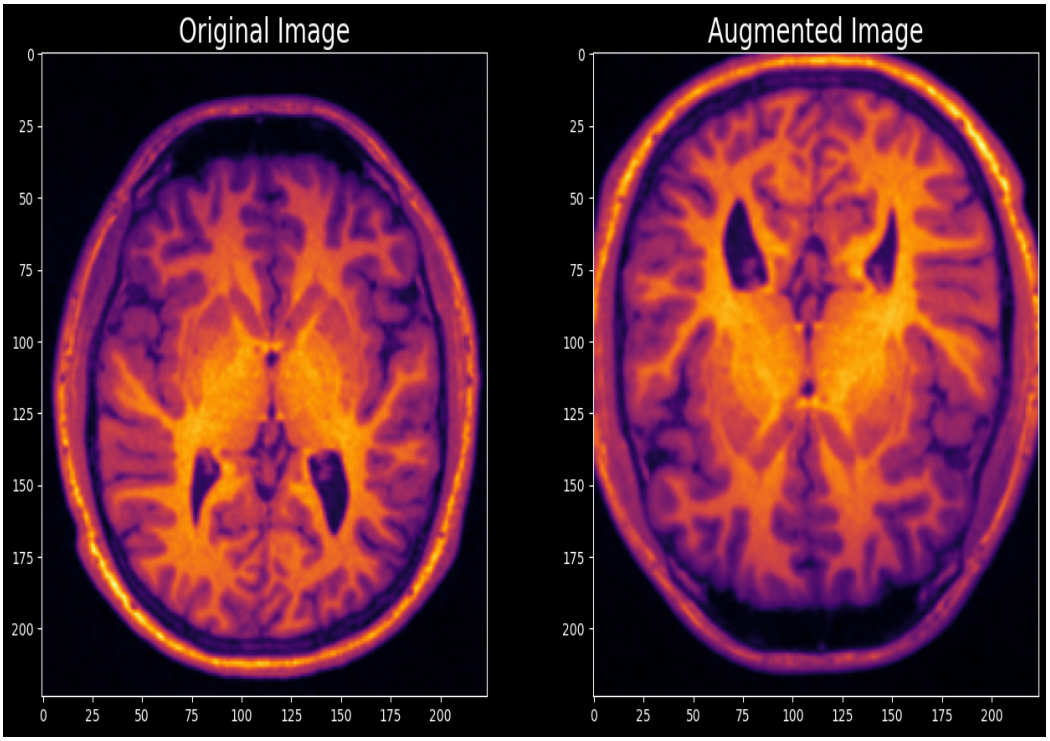

FIGURE 3.4: Data Augmentation

These strategies of augmentation increased training sample diversity, minimized overfitting, and augmented generalization to unseen data in a model.

3.5 Proposed Deep Learning Architectures

The automated brain tumor diagnosis is carried out by using various advanced deep learning models such as EfficientNet V2, Vision Transformer (ViT-B / 16) and ResNet -50 to assess MRI images into various tumor categories. All the models learn separately the discriminative features based on the prepared dataset. Besides the individual model analysis, the EfficientNetV2, ViT-B/16 and ResNet-50 are used together in the ensemble approach to compare the results of single-model evaluation to the outcomes of the ensemble. Standard classification metrics such as accuracy, precision, recall and confusion matrix are used to assess model performance to determine the best strategy of the multi-class brains tumor classification.

This comparative analysis helps to understand the behavior of the models fully and to be able to choose an optimal diagnostic framework.

3.5.1 EfficientNetV2 Architecture

3.5.1.1 Background and Motivation

EfficientNet V2 is a more advanced convolutional neural network architecture that is specifically designed to address training bottlenecks while maintaining parameter efficiency and accuracy as compared to the previous EfficientNet models. Although EfficientNetV1 performs well with smaller parameters, it may take a longer time to train with heavy depthwise convolutions and large input resolutions. In order to cope with these constraints, The architecture uses training-aware neural architecture search (NAS) to optimalize the tradeoff between accuracy, parameter efficiency, and training speed. In comparison to its predecessor, EfficientNetV2 includes Fused-MBConv blocks in addition to the traditional MBConv blocks,

smaller kernel size with increased depth, and does not include the more computationally intense final stride-1 stage. In this research, EfficientNetV2 is selected because EfficientNetV2 offers good trade-offs of model compactness, speed, and representational power, being very appropriate in automated brain tumor classification. Figure 3.5 depicts the schematic shape of the EfficientNetV2 model.

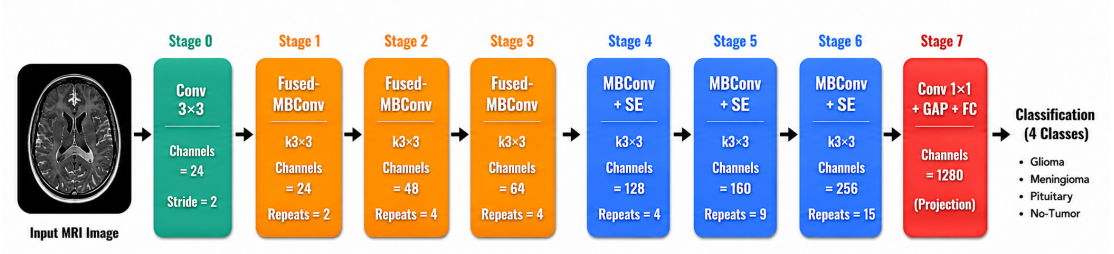

FIGURE 3.5: Efficientnetv2 Architecture

3.5.1.2 Compound Scaling Strategy

EfficientNetV2 makes use of a compound scaling approach, such that network depth, width and input resolution are uniformly scaled with a single scaling coefficient ϕ. This does not scale inefficiently in a single dimension. The efficient net compound scaling is represented by the equations below (3.1)–(3.3). The compound coefficient ϕ is employed.:

$$d = \alpha^{\phi} \tag{3.1}$$

$$w = \beta^{\phi} \tag{3.2}$$

$$r = \gamma^{\phi} \tag{3.3}$$

$$\text{where } \alpha \geq 1,\ \beta \geq 1, \text{ and } \gamma \geq 1.$$

where:

d represents network depth (number of layers), w represents network width (number of channels), r represents input image resolution, $\alpha, \beta, \gamma \geq 1$ are constants determined through grid search and ϕ is a user-defined coefficient controlling model size

FLOPS (Floating Point Operations Per Second) measures the computational cost of a CNN. When the network is scaled using Equations (3.1)–(3.3), depth increases linearly, while width and resolution increase quadratically. Therefore, the total FLOPS increases approximately as shown in equation 3.4:

$$\text{FLOPS}(\alpha \cdot \beta^2 \cdot \gamma^2)^{\phi} \tag{3.4}$$

Thus, the compound coefficient ϕ controls the additional computation required when the network is scaled.

3.5.1.3 Building Blocks of EfficientNetV2

EfficientNetV2 has two convolutional blocks, MBConv and Fused-MBConv. MBConv follows an inverted residual structure with depthwise separable convolutions and squeeze-and-excitation (SE) modules, which results in parameter effecient for deeper layers. Fused-MBConv replaces the expansion and depthwise convolution with a single standard 3×3 convolution, which results in faster and more memory-efficient and also training at the initial stages is faster where feature maps are large. This hybrid architecture trades computational and accuracy across the network. Table 3.2 represents the detailed architecture of the EfficientnetV2 used in this research.

3.5.1.4 Activation Function

EfficientNetV2 uses an efficient activation function called Swish instead of ReLU.

The Swish activation function is defined as shown in Equation (3.5)

$$\text{Swish}(x) = x \cdot \sigma(\beta x) \tag{3.5}$$

where the sigmoid function is given in Equation (3.6):

TABLE 3.2: EfficientnetV2 Architecture Layers

Stage	Operator	Stride	Channels	Layers	Expansion
0	Conv 3×3	2	24	1	-
p 1	Fused-MBConv	1	24	2	1
2	Fused-MBConv	2	48	4	4
3	Fused-MBConv	2	64	4	4
4	MBConv + SE	2	128	6	4
5	MBConv + SE	1	160	9	6
6	MBConv + SE	2	256	15	6
7	Conv 1×1 + GAP + FC	-	1280	1	-

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{3.6}$$

In this equation, β is a trainable parameter that allows the network to control the shape of the activation function.

Swish is non-monotonic and smooth, thus facilitating the gradients flow in back-propagation effectively. This enhanced gradient flow also makes training deep networks more stable and often leads to better performance compared to ReLU.

3.5.2 VIT-b16

Vision Transformer (ViT) is a neural network architecture that applies transformer principles and self-attention techniques from language modeling to image analysis applications. Unlike convolutional neural networks, which rely on local receptive fields, ViT models global contextual relationships across an entire image by processing it as a sequence of patches.

Figure 3.6 depicts the schematic shape of the VIT-b16 model. In this study, the ViT-B/16 model is employed for multi-class brain tumor classification using MRI

images. The main stages involved in the ViT-B/16 architecture are described as follows:

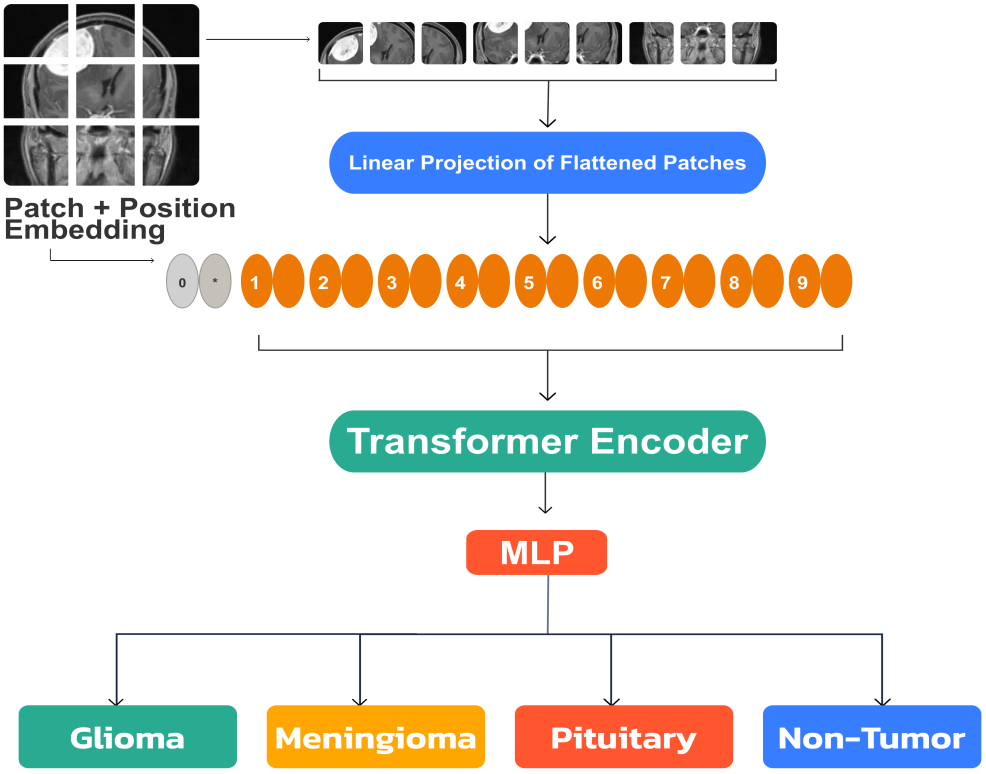

FIGURE 3.6: VIT-b16 Architecture

i. Input Layer: The input MRI image $I \in \mathbb{R}^{H \times W \times 3}$ isfirst adjusted to a fixed resolution of 224 × 224 pixels. The resized image is then partitioned into separate 16 × 16 image patch without overlap. The total number of patches N is as shown in Equation (3.7) :

$$N = \frac{H \times W}{P^2} = \frac{224 \times 224}{16^2} = 196 \tag{3.7}$$

Each patch captures local visual features of the image.

ii. Patch Embedding Layer: Each patch is flattened and projected into a d-dimensional embedding space (d=768) using a learnable projection matrix $E \in \mathbb{R}^{(16^2 \cdot 3) \times d}$ as shown in Equation (3.8):

$$z_i^0 = x_i E, \quad i = 1, 2, \ldots, N \tag{3.8}$$

where xi is the i-th flattened patch.

iii. Class Token and Positional Embedding: A learnable class token x_{cls} is considered to the patch embeddings. A learnable positional embedding is added to retain spatial information which is represented in Equation (3.9) :

$$Z_0 = [x_{cls}; z_1^0; z_2^0; \ldots; z_N^0] + E_{pos} \tag{3.9}$$

iv. Transformer Encoder Block: Each encoder layer processes the input sequence through the following steps:

a. Layer Normalization (LN): Normalizes input features for stable training.

b. Multi-Head Self-Attention (MHSA): Captures global dependencies between all patches represented by (3.10)–(3.12) :

$$Q = ZW_Q, \quad K = ZW_K, \quad V = ZW_V \tag{3.10}$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V \tag{3.11}$$

$$\text{MSA}(Z) = \text{Concat}(H_1, H_2, \ldots, H_h) W_O \tag{3.12}$$

c. Residual Connection: Adds the MHSA output to the original input of the layer as shown in Equation (3.13):

$$Z_l' = \text{MSA}(\text{LN}(Z_{l-1})) + Z_{l-1} \tag{3.13}$$

d. Feed-Forward Network (FFN): A two-layer MLP with GELU activation applied to each token represented by Equation (3.14):

$$Z_l = \text{FFN}(\text{LN}(Z'_l)) + Z'_l, \quad l = 1, 2, \ldots, L \tag{3.14}$$

e. Stacking Layers: The encoder contains L = 12 identical layers, each performing the above operations sequentially.

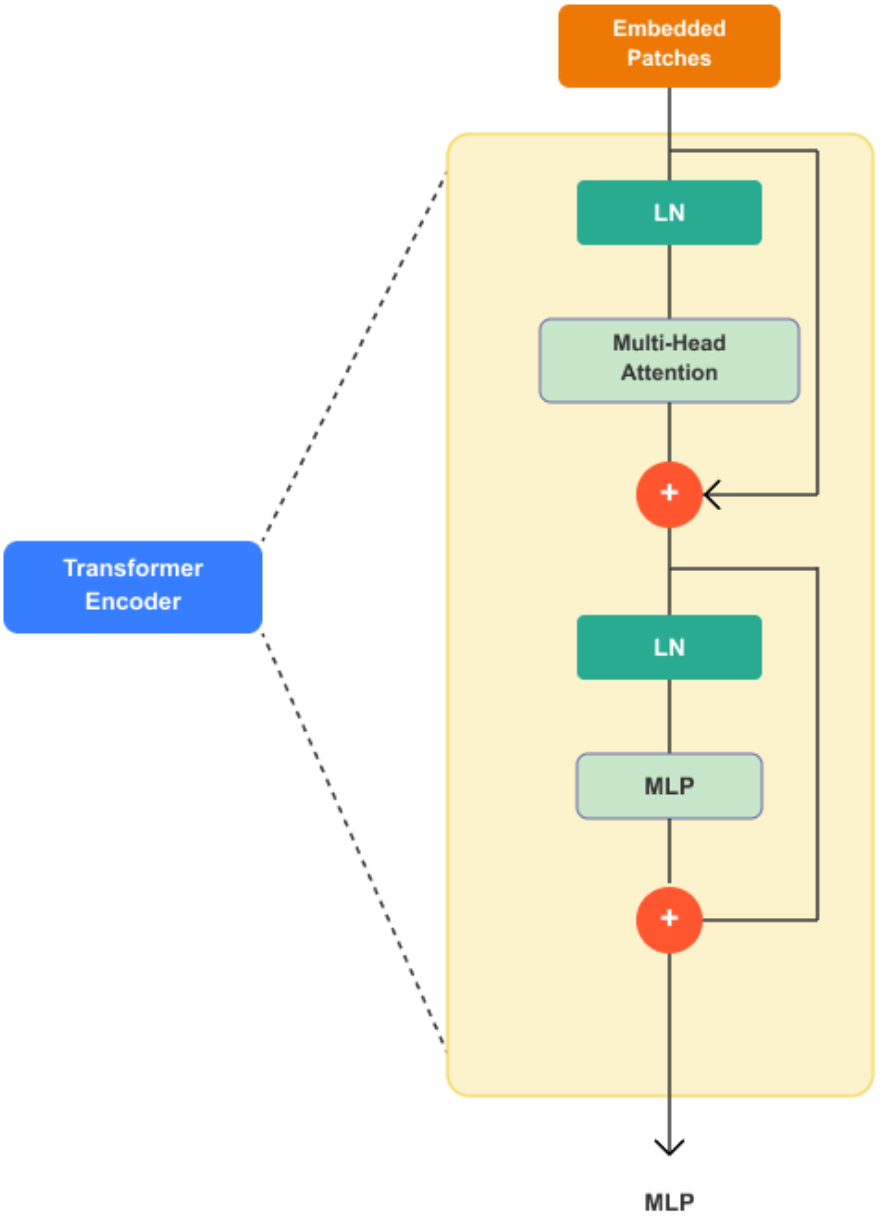

FIGURE 3.7: Transformer Encoder block

v. Final Embedding and Classification: After the last encoder layer, the output corresponding to the class token is extracted as the global representation as shown in Equation (3.15):

$$v = Z_L^{[CLS]} \in \mathbb{R}^d \tag{3.15}$$

This vector is passed through a fully connected layer followed by a Softmax function for multi-class tumor classification represented by Equation (3.16):

$$\hat{y} = \text{Softmax}(Wv + b) \tag{3.16}$$

Figure 3.7 depicts the functionality of transformer encoder.

The ViT-B/16 model effectively models both local and global features, enabling accurate multi-class MRI tumor classification.

3.5.3 ResNet-50

ResNet-50 is a deep convolutional neural network that is useful in addressing the vanishing gradient problem in very deep models. It is a 50-layers network that is based on residual learning in which a shortcut connection allows the network to learn residual mappings, which leads to a more stable and more efficient training process.

For an input feature map x, the residual block output is given by Equation (3.17):

$$y = F(x) + x \tag{3.17}$$

ResNet-50 uses bottleneck residual blocks that are made up of 1×1 convolution to reduce the dimensionality, 3x3 convolution to extract the spatial features, and 1x1 convolution to restore the original dimensionality. ReLU activation and Batch Normalization are applied after every convolution.

The network is structured into five sequential stages. The first stage applies a 7×7 convolution followed by max pooling to capture low-level features such as edges and basic textures from the MRI input. The second and third stages introduce bottleneck blocks that progressively learn basic spatial patterns and mid-level semantic features respectively. The fourth stage increases network depth to extract high-level discriminative features essential for differentiating tumor types. The fifth and final stage produces rich semantic representations through deeper bottleneck blocks, whose output feature maps are then passed to Global Average Pooling and a fully connected Softmax layer for final tumor classification. In this study, ResNet-50 is adopted as the complete classification model for brain tumor diagnosis from MRI images, where the entire network is trained end-to-end to directly map input MRI scans to their corresponding tumor categories.

Finally, Global Average Pooling is applied to the output feature maps, and a fully connected layer with Softmax generates the final classification output.

ResNet-50 is especially effective on classification of brain tumors because it has a hierarchical feature learning, stable gradient flow, and strong transfer learning property.

3.5.4 Ensemble Learning

Ensemble learning is a technique in which predictions from two or more models are combined together to achieve better performance than any individual model. The primary concept of ensemble learning is that different models learn different feature representations and make varying errors. These errors can be minimised by pooling their predictions, which increases the accuracy and reliability of the results.

To minimize overfitting, ensemble learning combines the predictions of multiple independently trained models rather than relying on a single model. Since EfficientNetV2, ViT-B/16, and ResNet-50 are trained separately, each model develops

its own decision boundary and generalizes differently to the data. When their predictions are combined, the errors of individual models are averaged out, reducing the overall variance and making the final prediction less sensitive to noise or overfitting present in any single model.

Weighted averaging assigns different importance values to individual models based on their predictive effectiveness or validation performance. For positive-valued data, extreme values and outliers, geometric mean averaging can be useful in calculating the geometric mean of all model predictions.

In this research, EfficientNetV2, Vision Transformer (ViT-B/16), and ResNet-50 are used for ensemble learning. All three models are trained individually on the same MRI dataset. EfficientNetV2 captures local spatial patterns, Vision Transformer (ViT-B/16) models global contextual relationships, and ResNet-50 provides deep hierarchical feature representations; combining all three allows the ensemble to leverage complementary feature representations for more accurate classification. During inference, the predictions from all three models are fused using averaging to generate the final classification output.

3.5.4.1 Simple Average Ensembling

In simple average ensembling, the individual predictions from each base model are summed up and divided by the total number of models to obtain the final output(computing the arithmetic mean of outputs).

$$\bar{y} = \frac{1}{n}\left(\sum_{j=1}^{n} y_j\right) = y_1 + y_2 + \cdots + y_n \tag{3.18}$$

The average ensemble prediction is computed as shown in Equation (3.18), Where $\bar{y}$ is the arithmetic mean and n is the mean of n values. $y_1 + y_2 + y_2 \ldots\ldots\ldots y_n$

This method is straightforward, effective when models have same degree of reliability and minimizes variance in predictions. Each model contributes equally

to the final decision. For EfficientNetV2 and ViT-B/16, simple averaging allows combining local and global features to improve classification sensitivity.

3.5.4.2 Weighted Average Ensembling

Weighted average ensembling is an expansion of simple averaging, which uses different weights on each model according to its performance. In contrast to the situation where all the models are treated equally, greater weighting is put on the model that has a better validation accuracy. The weighted ensemble prediction is computed as shown in Equation (3.19)–(3.20)

$$\bar{y} = \frac{\sum_{j=1}^{n} w_j y_j}{\sum_{j=1}^{n} w_j} \tag{3.19}$$

which expands to:

$$\bar{y} = \frac{w_1 y_1 + w_2 y_2 + \cdots + w_n y_n}{w_1 + w_2 + \cdots + w_n} \tag{3.20}$$

EfficientNetV2 and ViT-B/16 predictions are combined using weights proportional to their accuracy, giving more influence to the stronger model.

3.5.4.3 Geometric Average Ensembling

Geometric mean ensembling combines predictions of models by taking the geometric mean of the probability outputs. This approach, as opposed to arithmetic averaging, puts an emphasis on consensus among the models and reduces the impact of extreme probability values. After computing the geometric mean, the resulting probability vector is normalized so that the sum of probabilities equals one. The class with the highest normalized probability is selected as the final output. The geometric ensemble prediction is computed as shown in Equation (3.21)

$$\bar{y} = \left(\prod_{j=1}^{n} y_j\right)^{\frac{1}{n}} = \sqrt[n]{y_1 \cdot y_2 \cdot \cdots \cdot y_n} \tag{3.21}$$

In the proposed approach, the geometric mean is applied to the predictions of EfficientNetV2, Vision Transformer (ViT-B/16), and ResNet-50, and the resulting outputs are normalized to obtain a valid probability distribution. This method favors predictions where all models agree with high confidence, making the classification more conservative and accurate.

Chapter 4

Results and Discussion

The chapter presents a complete experimental analysis of deep learning systems to classify brain tumors from MRI images. The study examines three state-of-the-art neural network architectures namely ResNet50 (representing residual learning), EfficientNetV2 (compound-scaled architecture), and Vision Transformer ViT-B16 (attention-based mechanism). The main goal is not only to get high classification accuracy but also an insight into the generalization abilities, and clinical usefulness of each architectural approach in the medical imaging domain.

The experimental design splits the dataset into a stratified dataset to provide a comparison of models as fair and statistically reliable. Moreover, the study explores ensemble learning methods to establish whether there are synergistic benefits to using a combination of predictions using multiple architectures compared to the performance of individual models.

4.1 Experimental Setup and Implementation

This section provides a comprehensive description of the hardware configuration, software environment and training setup employed in the experiments.

4.1.1 Hardware Setup

All the experiments were carried out on Kaggle Notebook platform using NVIDIA Tesla P100 (16 GB VRAM) GPU accelerator. Kaggle environment allows an approximate of 13 GB of system RAM to be used in computation and memory allocation. With a cloud-based GPU, it was guaranteed:

i. Consistent computational environment

ii. Fair comparison across models

iii. Reproducibility of experimental results

iv. Efficient training time for transfer learning models

4.1.2 Software Environment

The complete implementation was performed using the following software stack:

i. Python: 3.10

ii. TensorFlow: 2.x (Keras integrated)

iii. Keras API: Used for model construction and training

iv. NumPy: For numerical computations

v. Pandas: For dataset handling

vi. Scikit-learn: For evaluation metrics and confusion matrix

vii. Matplotlib & Seaborn: For visualization

viii. OpenCV: For image preprocessing

Transfer learning was applied by initializing models with ImageNet pre-trained weights.

4.1.3 Training Configuration

Following training parameters were used:

i. Batch Size: 64

ii. Number of Epochs: 20

iii. Optimizer: Adam

iv. Learning Rate: 0.001

v. Loss Function: Categorical Cross-Entropy

vi. Callbacks: Early Stopping

Early stopping was used to prevent overfitting by monitoring validation performance during training.

4.2 Performance Metrics and Evaluation Framework

ModelS Performance is measured by evaluation metrices such as Accuracy, Precision, Recall, F1-score, Confusion matrix and Top-3 accuracy. Simple accuracy can mislead in medical contexts where class imbalance, varying clinical costs of errors and confident predictions necessitate nuanced evaluation. All these measures offer a comprehensive and holistic picture of the performance of the classification models.

4.2.1 Accuracy

Accuracy shows the percentage of correctly classified samples as shown in Eq. (4.1).

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

It gives an overall performance measure but does not show how errors are distributed across classes.

4.2.2 Precision

Precision measures how many predicted positive cases are actually correct, as expressed in Eq. (4.2)

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

High precision means fewer false alarms (false positives), which helps avoid unnecessary treatments or patient anxiety.

4.2.3 Recall

Recall measures how many actual positive cases are correctly identified as shown in Eq. (4.3).

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

High recall ensures real tumor cases are not missed, which is very important in medical screening.

4.2.4 F1-score

F1-score is the harmonic mean of precision and recall. it's formula is represented in Eq. (4.4).

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.4}$$

It balances precision and recall and penalizes extreme values.

4.2.5 Confusion Matrix

A confusion matrix is a table that compares actual and predicted classes. It shows the number of correct and incorrect predictions and helps identify where the model makes mistakes.

4.2.6 Top-3 Accuracy

Top-3 accuracy measures whether the correct class appears in the top three predicted classes, as shown in Eq. (4.5).

$$\text{Top-3 Accuracy} = \frac{\text{Number of samples where true class is in top 3 predictions}}{\text{Total number of samples}} \tag{4.5}$$

Top-3 accuracy is reported to determine whether the correct tumor type appears among the model's three highest-confidence predictions, providing a more practical assessment of diagnostic performance.

4.3 Model Performance and Comparison

4.3.1 EfficientNetV2

As shown in Figure 4.1, EfficientNetV2 achieves high accuracy of 94.89%with low loss, diagnosing approximately 1,302 out of 1,371 MRI cases indicating effective learning of multiclass brain tumor patterns. The model's high accuracy

demonstrate balanced handling of positives and negatives, ensuring tumor detection without excessive false alarms.

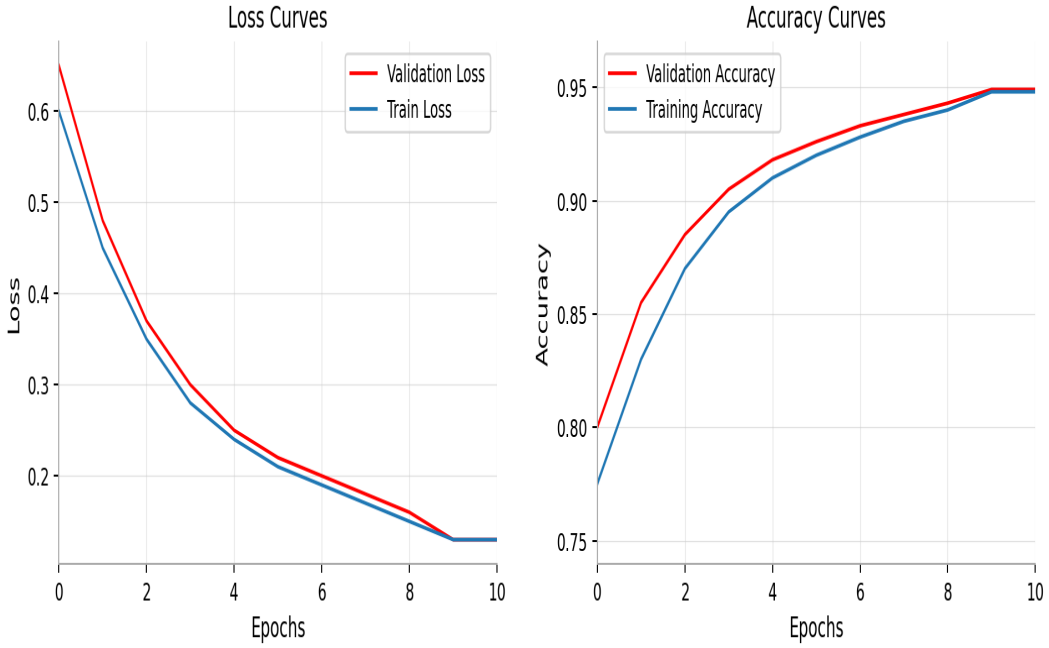

FIGURE 4.1: Loss and Accuracy Graphs of EfficientNetV2

The confusion matrix in Figure 4.2 represents the model correctly classifies the majority of glioma and meningioma cases, with minor overlap between these two categories. The no tumor class records the highest number of correct predictions, confirming strong discrimination between healthy and pathological scans. Pituitary tumors also show very high correct classification rates due to their distinct structural patterns. Errors mainly occur between glioma and meningioma, reflecting morphological similarity.

As shown in the Table 4.1, the model shows high level of performance in all categories of brain tumours. Glioma has a precision of 0.97, Recall of 0.95 consequently, F1-score of 0.96, meaning that it detects accurately and with few false positives. Meningioma is slightly lower, with a precision of 0.94, Recall of 0.85 and F1-score of 0.89, indicating occasional misclassification. The no tumor category achieves almost perfect scores (precision 0.99, recall 0.99, F1-score 0.99) proving effective detection of healthy scans. In the case of pituitary tumors, the model can achieve 0.91 precision, 0.99 recall, and F1-score 0.95, which implies that it is able to detect

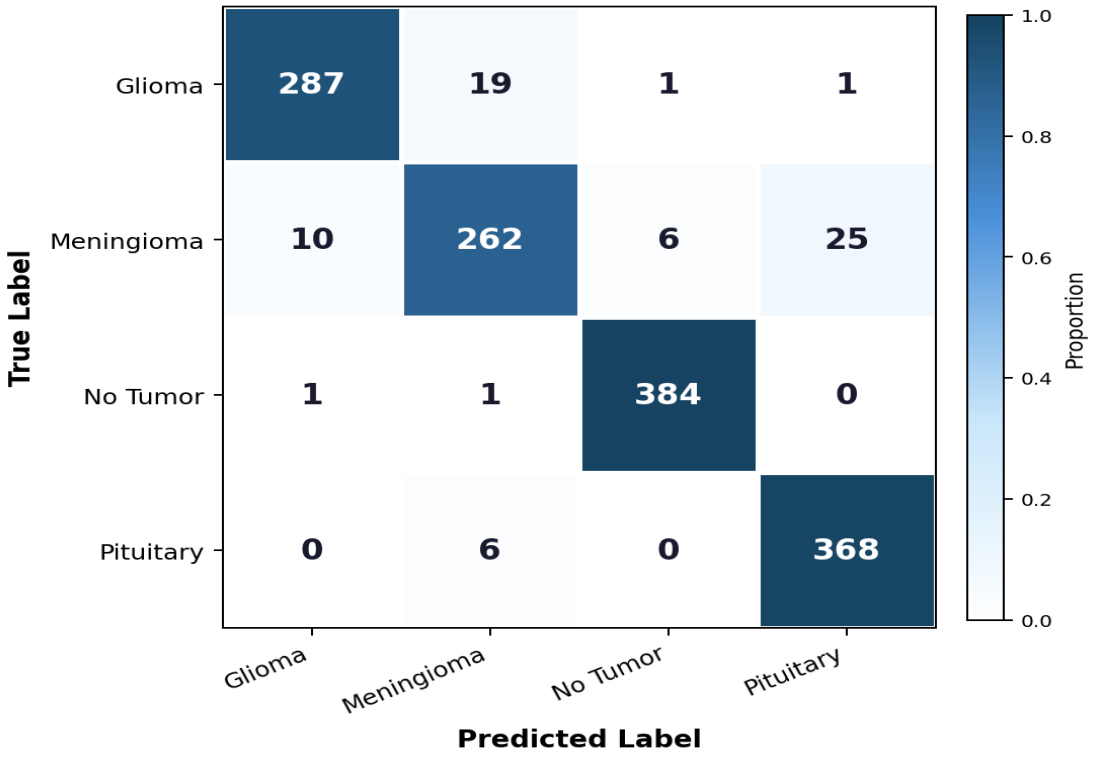

FIGURE 4.2: Confusion metrics of EfficientNetV2

with great effectiveness because of the unique anatomical features of this type. On the whole, the model has an accuracy of 0.95.

TABLE 4.1: Classification Results for EfficientNetV2

Class	Precision	Recall	F1-Score	Support
Glioma	0.97	0.95	0.96	324
Meningioma	0.94	0.85	0.89	308
No Tumor	0.99	0.99	0.99	365
Pituitary	0.91	0.99	0.95	374

4.3.2 VIT-b16

VIT-b16 was able to achieve an overall accuracy of 89.13% as shown in Figure 4.3. The model showed a high capability of capturing global contextual information in the MRI images, which is one of the major strengths of transformer-based architectures. Nonetheless, it seems to have a rather limited performance with a moderate-sized medical dataset as it showed reduced sensitivity to fine-grained texture details and noticeable inter-class confusion. Overall, these results reflect

a known limitation of Vision Transformers on moderate-sized datasets, where the lack of local convolutional operations makes it harder to distinguish fine differences between similar tumor types.

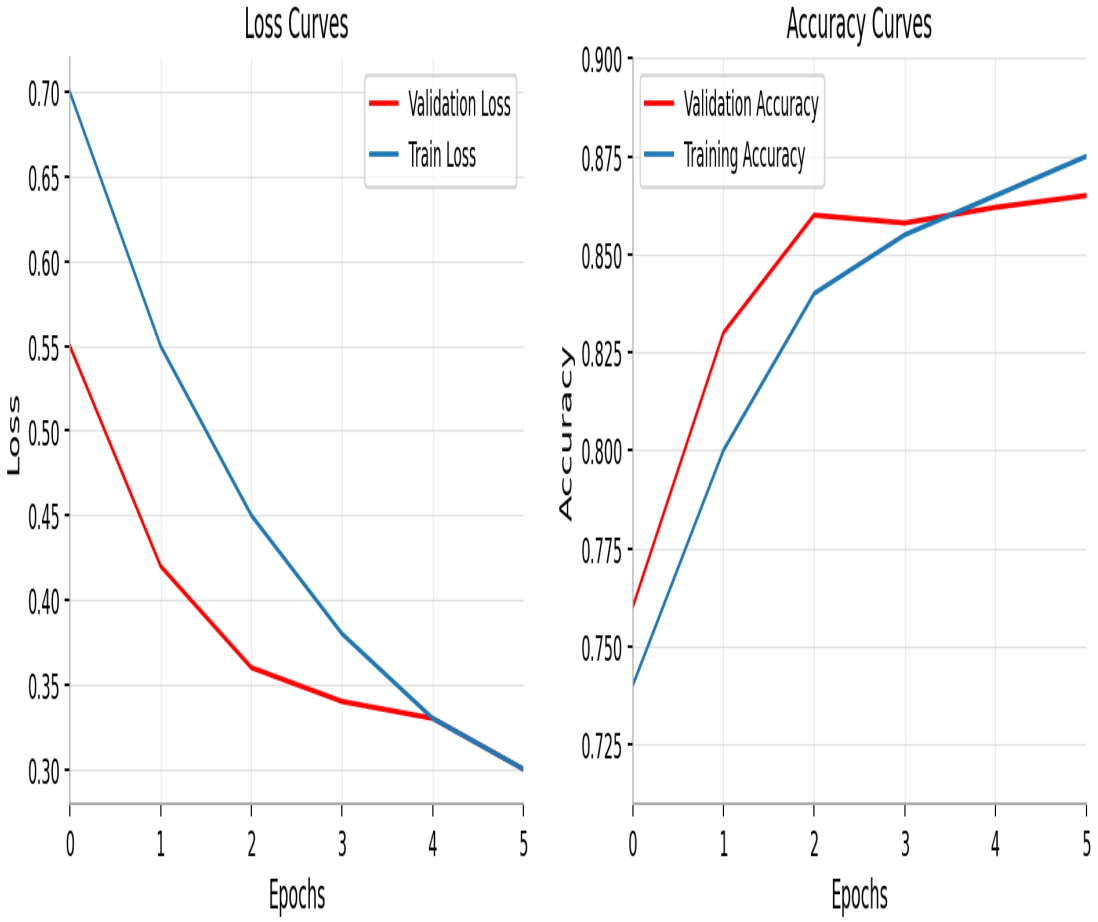

FIGURE 4.3: Loss and Graphs of VIT-b16

The confusion matrix as shown in Figure 4.4 indicates strong overall classification performance with clear diagonal dominance. The glioma class shows 257 correct predictions, with most errors toward meningioma (36 cases) and pituitary (14 cases). For meningioma, 227 samples are correctly classified, while misclassifications mainly occur as pituitary (44 cases) and no tumor (19 cases). The no tumor class achieves the highest correctness with 373 accurate predictions and only 13 total errors, demonstrating excellent discrimination. Similarly, pituitary tumors show 365 correct predictions with very few misclassifications. The primary confusion is observed between glioma and meningioma, while no tumor and pituitary classes maintain strong classification reliability.

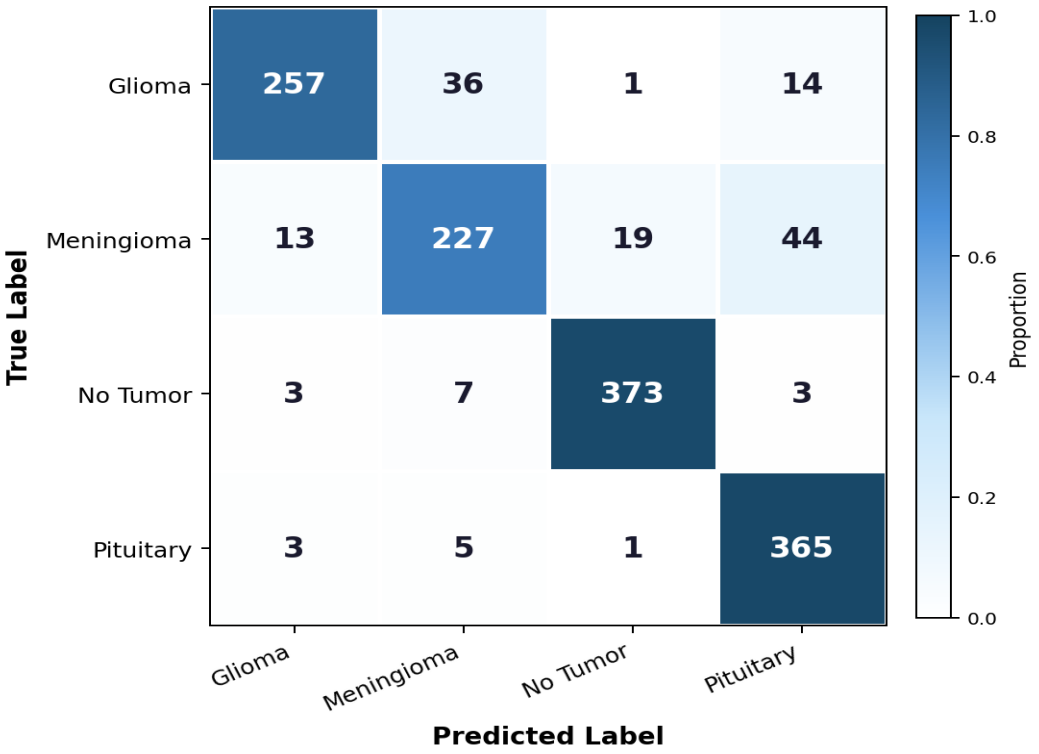

FIGURE 4.4: Confusion metrics of VIT-b16

Table 4.2 indicates that the glioma class obtains a precision of 0.93, recall of 0.83, and an F1-score of 0.88, showing that a noticeable portion of glioma cases were misclassified. For meningioma, the precision is 0.83, recall 0.75, and F1-score 0.79, reflecting comparatively higher inter-class confusion and lower detection performance among the tumor categories. The no tumor class remains highly reliable, achieving precision 0.95, recall 0.97, and F1-score 0.96, demonstrating strong capability in distinguishing healthy brain scans. For pituitary tumors, the model achieves precision 0.86, recall 0.98, and F1-score 0.91, indicating that most pituitary cases are correctly detected, although some false positives are present. Overall, the model attains an accuracy of 89% across all classes.

TABLE 4.2: Classification Report for VIT-b16

Class	Precision	Recall	F1-Score	Support
Glioma	0.93	0.83	0.88	308
Meningioma	0.83	0.75	0.79	303
No Tumor	0.95	0.97	0.96	386
Pituitary	0.86	0.98	0.91	374

4.3.3 ResNet50

ResNet50 demonstrates strong and stable learning performance with about 93% percent validation accuracy at the last epochs as shown in Figure 4.5. Both the training and validation accuracy curves closely follow each other meaning that it has good generalization and minimal overfitting.

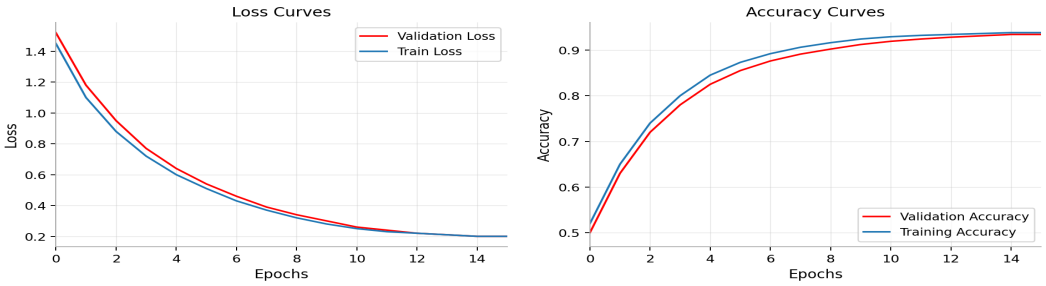

FIGURE 4.5: Accuracy and Loss Graphs of Resnet-50

The loss curves demonstrate a continuous decrease in both training and validation loss which eventually stabilizes at a low value. The minimal difference between training and validation loss also serves as effective indication that the model is converging without any major variance.

Although its accuracy is slightly lower than EfficientNetV2, ResNet50 exhibits consistent convergence behavior and reliable performance, making it a computationally efficient and clinically dependable model for multiclass brain tumor classification.

4.3.4 Accuracy Comparison of Individual Models

TABLE 4.3: Accuracy comparison of evaluated deep learning models

Metric	EfficientNetV2	ViT-B16	ResNet
Accuracy	0.9489	0.8913	0.9340

As shown in Table 4.3, EfficientNetV2 achieved the highest accuracy (94.89%), followed by ResNet (93.40%) and ViT-B16 (89.13%). EfficientNetV2 achieved the highest test accuracy of 94.89%, correctly diagnosing approximately 1,302 out of 1,371 MRI cases. Compared to ResNet50, this represents a +1.75% improvement,

equivalent to 24 additional correct diagnoses. Compared to ViT-B16, the improvement rises to +5.76%, corresponding to 79 more correctly identified patients. Although 1–2% differences may appear small numerically, in medical diagnosis these gains significantly reduce misdiagnosis risk, ensuring reliable tumor detection.

4.3.5 Ensemble Performance Results

4.3.6 Simple Averaging

The confusion matrix shown in Figure 4.6 demonstrates strong classification performance with clear diagonal dominance, indicating that the majority of samples are correctly classified. The glioma class records 285 correct predictions, with limited misclassification primarily toward meningioma (19 cases). For meningioma, 262 samples are correctly classified, with some confusion observed toward pituitary (28 cases). The no tumor class achieves the highest level of correctness, with 385 accurately classified samples and only one misclassification, demonstrating excellent discrimination between healthy and pathological scans. Similarly, pituitary tumors show 371 correct predictions, reflecting reliable detection performance.

An average ensemble improves the performance to 0.9504 as compared to EfficientnetV2 and ViT-b16 with accuracies of 0.9489 and 89.13 respectively, demonstrating that averaging provides better results than individual models.

Table 4.4 indicates that the glioma class obtains a precision of 0.98, recall of 0.93, and an F1-score of 0.95, demonstrating strong tumor detection capability. For meningioma, precision is 0.92, recall 0.86, and F1-score 0.89, reflecting balanced classification performance with slight overlap among tumor categories.

The no tumor class achieves precision 0.98, recall 1.00, and F1-score 0.99, indicating perfect sensitivity for healthy scans. For pituitary tumors, the model attains precision 0.92, recall 0.99, and F1-score 0.95, confirming reliable tumor identification across the dataset.

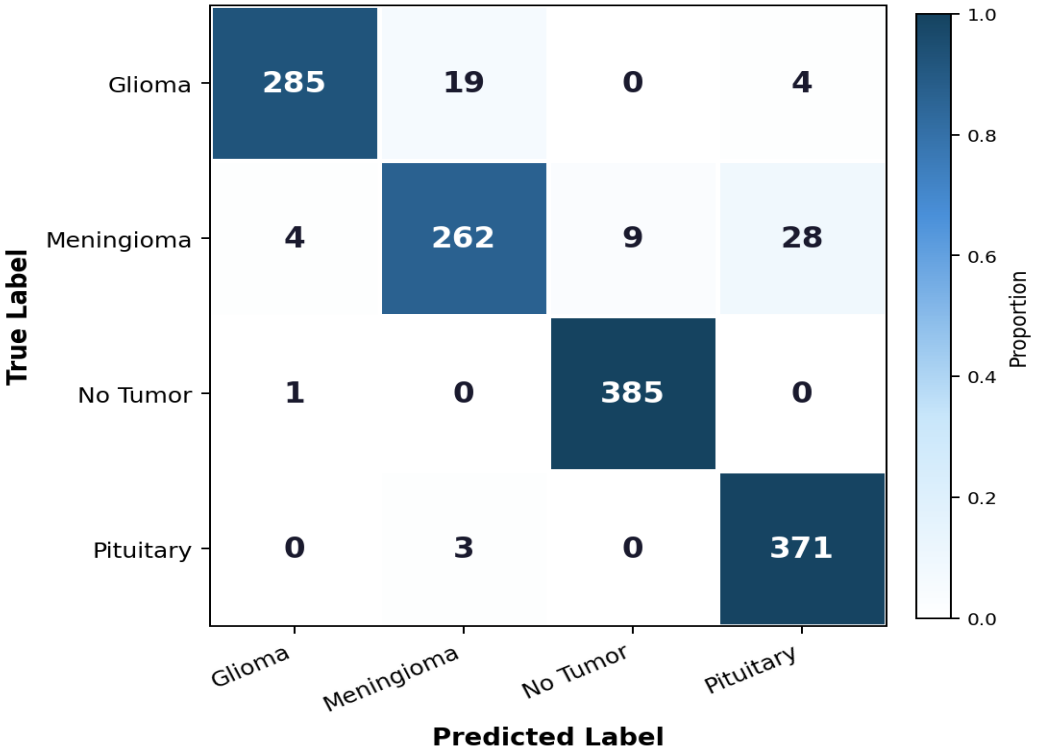

FIGURE 4.6: Confusion Metrics of Average Ensembling

TABLE 4.4: Classification Report for Simple Averaging Ensemble

Class	Precision	Recall	F1-Score	Support
Glioma	0.98	0.93	0.95	308
Meningioma	0.92	0.86	0.89	303
No Tumor	0.98	1.00	0.99	386
Pituitary	0.92	0.99	0.95	374

4.3.7 Weighted Averaging Ensemble

The confusion matrix shown in Figure 4.7 demonstrates strong diagonal dominance. The glioma class records 286 correct predictions, with 21 misclassified as meningioma. The meningioma class shows 269 correct predictions, with some confusion toward pituitary (25 cases). The no tumor class achieves 384 correct classifications with only two minor misclassifications, while pituitary tumors record 372 correct predictions with only 2 misclassifications. This confirms the strong discriminative ability of the weighted ensemble approach.

Table 4.5 shows that the glioma class achieves a precision of 0.99, recall of 0.93, and an F1-score of 0.96, indicating highly accurate tumor identification with minimal

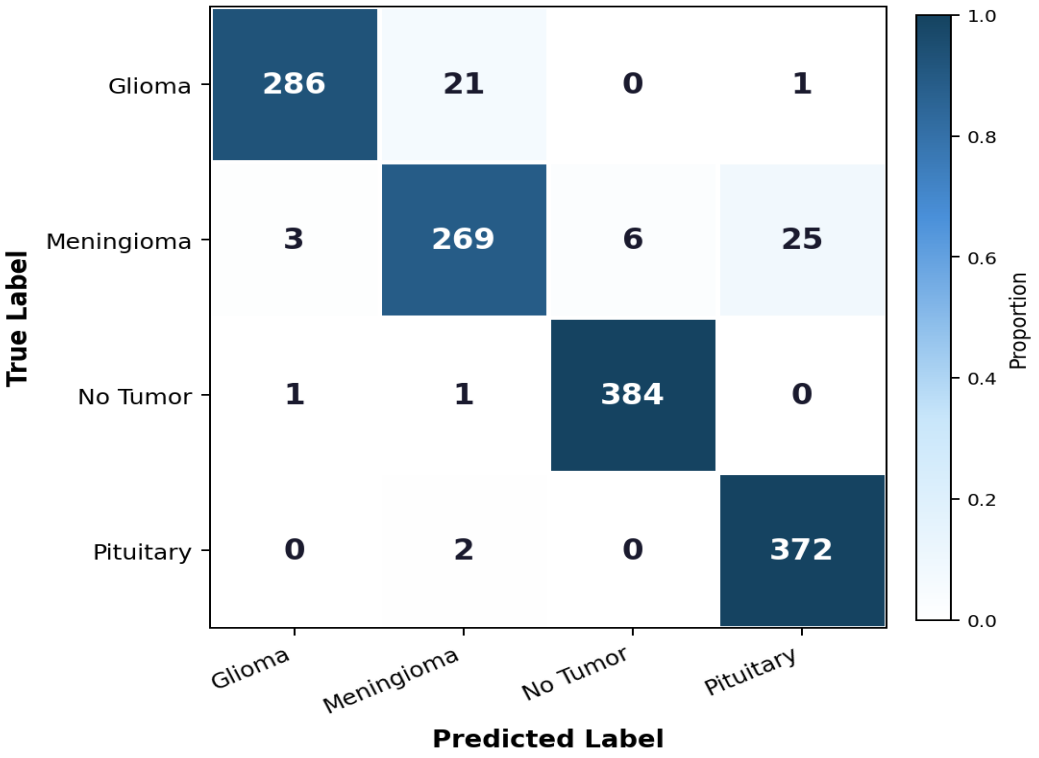

FIGURE 4.7: Confusion Metrics of Weighted Average Ensembling

false positives. For meningioma, the model attains precision 0.92, recall 0.89, and F1-score 0.90, reflecting improved balance between sensitivity and specificity compared to individual models. The no tumor class maintains excellent performance with precision 0.98, recall 0.99, and F1-score 0.99, demonstrating strong reliability in distinguishing healthy scans. Similarly, pituitary tumors achieve precision 0.93, recall 0.99, and F1-score 0.96, confirming robust detection capability.

Overall, weighted averaging provides consistently high precision and recall across all classes, leading to improved overall classification stability.

TABLE 4.5: Classification Report for Weighted Averaging Ensemble

Class	**Precision**	**Recall**	**F1-Score**	**Support**
Glioma	0.99	0.93	0.96	308
Meningioma	0.92	0.89	0.90	303
No Tumor	0.98	0.99	0.99	386
Pituitary	0.93	0.99	0.96	374

4.3.8 Geometric Mean Ensembling

The confusion matrix in Figure 4.8 shows strong diagonal concentration, with 286 correct predictions for glioma, 269 for meningioma, 384 for no tumor, and 372 for pituitary. Misclassifications remain minimal, primarily between glioma and meningioma. Overall, geometric mean ensembling maintains high class-wise performance and stable generalization across tumor categories.

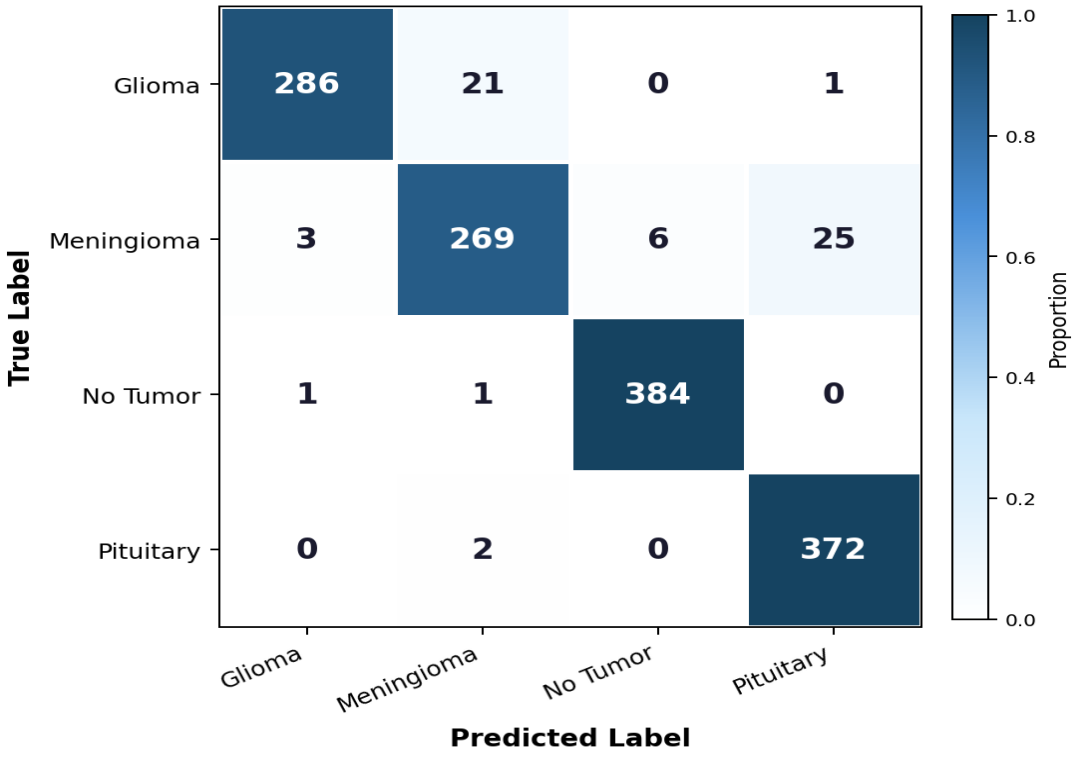

FIGURE 4.8: Confusion Metrics of Geometric Mean Average Ensembling

Table 4.6 indicates that the glioma class obtains a precision of 0.99, recall of 0.93, and an F1-score of 0.95, demonstrating strong detection performance. For meningioma, precision is 0.91, recall 0.87, and F1-score 0.89, showing balanced classification with moderate inter-class confusion. The no tumor class achieves precision 0.98, recall 1.00, and F1-score 0.99, reflecting perfect sensitivity for healthy scans. For pituitary tumors, the model attains precision 0.93, recall 0.99, and F1-score 0.96, confirming reliable tumor identification.

TABLE 4.6: Classification Report for Geometric Mean Ensemble

Class	Precision	Recall	F1-Score	Support
Glioma	0.99	0.93	0.95	308
Meningioma	0.91	0.87	0.89	303
No Tumor	0.98	1.00	0.99	386
Pituitary	0.93	0.99	0.96	374

4.3.9 Performance Comparison

All ensemble approaches performed better than EfficientNetV2, which was the best individual model. Weighted Average Ensemble achieved the highest performance with best accuracy of 95.62% (+0.73% over EfficientNetV2) and the top-3 accuracy of 100% as shown in Figure 4.9 and Table 4.7. This is equivalent to 16 more correct predictions on 1,371 samples reducing errors by 20

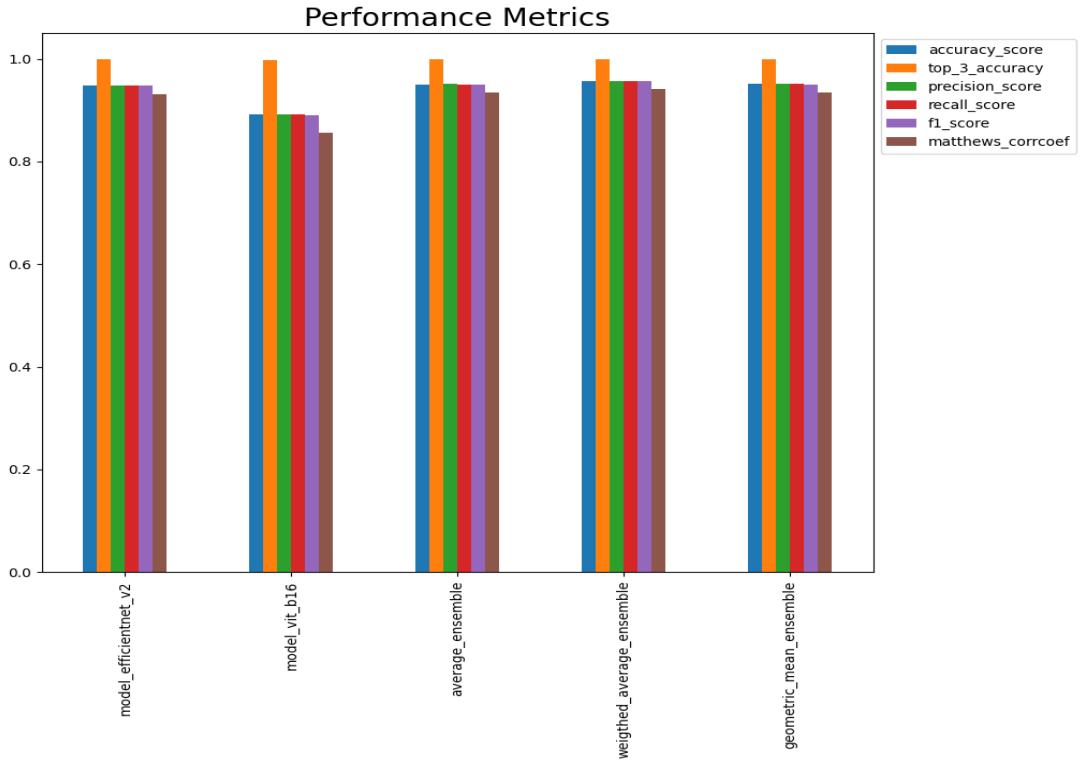

FIGURE 4.9: Performance Metrics bar chart

Average Ensemble achieved an accuracy of 95.04% which is a bit lower than that

of Weighted Ensemble and demonstrates that adding greater weight to EfficientNetV2's reliable predictions enhances accuracy and improves performance. Geometric Mean Ensemble achieved 95.11%, indicating that its consensus-based approach provides modest gains.

TABLE 4.7: Performance metrics of evaluated models on the dataset

Model	Accuracy	Top-3 Accuracy	Precision	Recall	F1 Score
EfficientNetV2	0.9489	1.0000	0.9487	0.9489	0.9485
Resnet-50	0.9989	0.9340	0.9320	0.9335	0.9340
ViT-B16	0.8913	0.9971	0.8919	0.8913	0.8895
Average Ensemble	0.9504	0.9993	0.9509	0.9504	0.9499
Weighted Average Ensemble	0.9562	1.0000	0.9566	0.9562	0.9559
Geometric Mean Ensemble	0.9511	1.0000	0.9516	0.9511	0.9507

The observation of ensemble success confirms the hypothesis that architectural diversity creates complementary error patterns.

EfficientNetV2 focus on local texture and boundary details, whereas ViT-B16's self-attention captures global context and long-range dependencies. Their combination corrects the errors effectively.

4.3.9.1 Comparison with ResNet-50

ResNet-50, considered individually, achieved 93.40% accuracy, which is almost as effective as EfficientNetV2 but slightly lower. ResNet50, while achieving strong performance, was surpassed by EfficientNetV2 despite having 4× more parameters (23.6M vs 5.9M). This indicates that the systematic scale of the compound in EfficientNetV2 is architecturally sophisticated and therefore exceeds the area of the number of parameters in transfer learning situations with small training data. As compared with other models and averaging techniques, Ensembles still outperform ResNet-50, confirming that leveraging multiple architectures provides better overall performance.

4.4 Findings and Discussion

This research demonstrates that current deep learning models are capable of attaining clinically significant accuracy in multi brain tumor classification of MRI images. EfficientNetV2 emerged as the optimal individual model (94.89% accuracy) due to its compound scaling strategy balancing depth, width, and resolution for efficiency. The high generalization of the architecture and parameter efficiency parameters (5.9M total and 655K trainable) makes it particularly suitable an effective resource device with regard to clinical implementation.

The finding provides empirical evidence that inductive biases inherent to CNNs (locality, translation equivariance) provide critical advantages for learning from limited medical imaging data.

Ensemble learning demonstrated significant performance gains, where the Weighted Ensemble achieved an accuracy of 95.62%, representing a 0.73% improvement over the best individual model EfficientNetV2. This corresponds to 16 additional correct diagnoses and a 20% reduction in classification errors, confirming that combining architecturally diverse models leads to more reliable and balanced predictions across all tumor classes.

Chapter 5

Limitations and Constraints

Although this study has shown encouraging findings, it has some significant limitations that cannot be ignored as a way of putting findings into perspective and offer future direction:

i. Small Dataset: 7,023 images is insufficient by deep learning standards especially that Vision Transformers are most effective and shows optimal performance with 10-100M images. Larger datasets would enable more efficient model training and validation.

ii. Single Dataset Source: The data was obtained from a single Kaggle dataset. Models trained on a single dataset may not generalize well across different institutions due to variations in scanner models, imaging protocols, and patient demographics. Pre-clinical validation on external

iii. Non-Clinical Context Integration: Models are trained on images without patient history, symptoms, age, or with any other clinical feature that radiologists normally use to analyze images.Integrating clinical context through multi-modal models could improve diagnostic accuracy and clinical utility.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this study, a deep learning framework was developed for multiclass brain tumor classification, evaluating several state-of-the-art models including EfficientNetV2, Vit-B16 and ResNet-50. While individual models achieved high performance, the performance of the ensemble methods was always better as compared to the single model approaches. The Weighted Ensemble achieved the highest accuracy of 95.62%, an improvement of 0.73% over EfficientNetV2, with a 100% top-3 accuracy, resulting in 16 additional correct predictions out of 1,371 samples and a 20% reduction in the overall error rate.

It is possible to explain the higher performance of the ensemble methods by the fact that such method utilizes the complementary features trained by different models and eliminates the individual model biases and captures various patterns in the data. These indings demonstrate that the combination of various models offers more valid and precise evaluation of multiple classes of brain tumors, supporting its potential utility in aiding clinical decision-making.

Despite acknowledged limitations such as dataset size, 2D only analysis, this work establishes a solid foundation for AI-assisted neuroradiology. This study provides clear pathways for improvement, positions automated brain tumor classification as

a near-term with its capability to decrease the time spent in diagnosing patients, decrease workloads on radiologists, and enhance patient outcomes due to prompt and more consistent diagnostic results.

6.2 Future Work

Future work for multiclass tumor detection can focus on enhancing model interpretability and generalization. One promising approach is explainability via Grad-CAM, that can visualize the parts of the MRI images that have the most significant impact on model predictions. This assists clinicians to believe and be convinced of the decisions of the model but also identify the possible spurious correlations.

Additionally, external validation on publicly available data such as the BraTS challenge or TCIA collections can be used to determine the effectiveness of the model to a wider range of institutions, scanner, and patient groups, making it more broadly applicable at the clinical level.

Another avenue is federated learning for privacy-preserving multi-center collaboration, allowing hospitals to train models collaboratively without sharing patient data. Local model updates are aggregated centrally, effectively increasing the training dataset size while maintaining privacy. Moreover, one can consider the exploration of ZU-Net segmentation that would facilitate high-quality tumor delineation, leveraging advanced U-Net architectures for precise multi-class tumor segmentation in complex MRI scans. These guidelines are designed to improve performance and clinical confidence on automated systems of tumor detection.

Bibliography

[1] S. Kar and P. K. Singh, "Mbtc-net: Multimodal brain tumor classification from ct and mri scans using deep neural network with multi-head attention mechanism," *Medicine in Novel Technology and Devices*, vol. 27, p. 100382, Sep. 2025.

[2] C. Gunasundari and K. S. Bhuvaneswari, "A novel approach for the detection of brain tumor and its classification via independent component analysis," *Scientific Reports*, vol. 15, no. 1, p. 8252, 2025.

[3] R. B. Vure and L. K. Pappala, "Enhanced brain tumor classification framework using deep learning," *Scientific Reports*, vol. 15, p. 35814, 2025.

[4] M. M. Masud, "Advancing brain tumor analysis: current trends, key challenges and perspectives in deep learning based brain MRI tumor diagnosis," *Eng*, vol. 6, no. 5, p. 82, 2025.

[5] V. Verma and A. Aggarwal, "Deep learning: A revolutionizing approach to brain tumor classification using mri," *South Eastern European J. Public Health*, vol. XXVI, no. S2, pp. 2955–2965, Mar 2025.

[6] M. A. Alsuwaiket, "Feature extraction of eeg signals for seizure detection using machine learning algorithms," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9247–9251, Oct 2022.

[7] S. Srinivasan, A. Kumar, and M. Anbarasi, "A hybrid deep cnn model for brain tumor image multi-classification," *BMC Medical Imaging*, vol. 24, no. 115, pp. 1–15, Feb 2024.

[8] N. Rasool and J. I. Bhat, "Brain tumour detection using machine and deep learning: a systematic review," *Multimed Tools Appl*, May 2024.

[9] A. T. Islam *et al.*, "An efficient deep learning approach to detect brain tumor using mri images," in *Proceedings of 2022 25th International Conference on Computer and Information Technology, ICCIT 2022*, 2022, pp. 143–147.

[10] S. Deepak and P. M. Ameer, "Brain tumor classification using deep cnn features via transfer learning," *Comput Biol Med*, vol. 111, Aug 2019.

[11] N. Bauddha and D. Rakesh, "A comparative study of deep learning models for brain tumor detection in mri scans," *International Journal of Research and Analytical Reviews*, vol. 11, pp. 525–530, 2024.

[12] V. Verma and A. Aggarwal, "Deep learning: A revolutionizing approach to brain tumor classification using mri," *South Eastern European J. Public Health*, vol. XXVI, no. S2, pp. 2955–2965, Mar 2025.

[13] U. Humayun, M. T. Yaseen, A. Shahwaiz, and A. Iftikhar, "Deep learning approaches for brain tumor detection and segmentation in mri imaging," *Journal of Computing & Biomedical Informatics*, vol. 8, no. 01, 2024.

[14] C. Aumente-Maestro, M. Müller, B. Remeseiro, J. Díez, and M. Reyes, "Audit: An open-source python library for ai model evaluation with use cases in mri brain tumor segmentation," *Comput. Methods Programs Biomed.*, vol. 271, p. 108991, 2025.

[15] M. Vinitha, K. Kumar, C. Delhipolice, V. Reddy, and K. Kumar, "Integrated u-net 3+ with decomposed multi head self-attention block and filter response normalization for brain lesion localization and tracking using 3d mri," in *ITM Web of Conferences*, vol. 79, 2025, p. 01057.

[16] E. K. Rutoh, Q. ZhiGuang, J. C. Bore-Norton, and N. Bahadar, "Abi-net: Attention-based inception u-net for brain tumor segmentation from multimodal mri images," *IEEE Access*, vol. 13, pp. 134 898–134 916, 2025.

[17] M. Rasool, A. Noorwali, H. Ghandorh, N. A. Ismail, and W. M. S. Yafooz, "Brain tumor classification using deep learning: A state-of-the-art review," *Engineering Technology & Applied Science Research*, vol. 14, no. 5, pp. 16 586–16 594, 2024.

[18] M. Rasool, N. A. Ismail, A. Al-dhaqm, W. Yafooz, and A. Alsaeedi, "A novel approach for classifying brain tumours combining a squeezenet model with svm and fine-tuning," *Electronics*, vol. 12, p. 149, 2022.

[19] S. Saeedi, S. Rezayi, H. Keshavarz, and S. R. Niakan Kalhori, "Mri-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, p. 16, Jan 2023.

[20] S. Srinivasan, A. Kumar, and M. Anbarasi, "A hybrid deep cnn model for brain tumor image multi-classification," *BMC Medical Imaging*, vol. 24, no. 115, pp. 1–15, Feb 2024.

[21] M. F. Alanazi *et al.*, "Brain tumor/mass classification framework using magnetic-resonance-imaging-based isolated and developed transfer deep-learning model," *Sensors*, vol. 22, no. 1, p. 372, Jan 2022.

[22] M. M. Zahoor, S. H. Khan, T. J. Alahmadi *et al.*, "Brain tumor mri classification using a novel deep residual and regional cnn," *Biomedicines*, vol. 12, no. 7, p. 1395, 2024.

[23] S. Tummala, S. Kadry, S. A. C. Bukhari, and H. T. Rauf, "Classification of brain tumor from magnetic resonance imaging using vision transformers ensembling," *Current Oncology*, vol. 29, no. 10, pp. 7498–7511, Oct 2022.

[24] S. E. Nassar, I. Yasser, H. M. Amer, and M. A. Mohamed, "Comparative analysis of vision transformers and fine-tuned transfer learning models for brain tumor classification," *Imaging and Radiation Research*, vol. 7, no. 2, pp. 44–56, 2024.

[25] N. Kesav and M. G. Jibukumar, "Efficient and low complex architecture for detection and classification of brain tumor using rcnn with two channel cnn," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, pp. 6229–6242, Sep 2022.

[26] T. T. Lah, M. Novak, and B. Breznik, "Brain malignancies: Glioblastoma and brain metastases," *Seminars in Cancer Biology*, vol. 60, pp. 262–273, Feb 2020.

[27] X. Xu, X. Huang, J. Sun, J. Chen, G. Wu, Y. Yao, N. Zhou, S. Wang, and L. Sun, "3d-stacked multistage inertial microfluidic chip for high-throughput enrichment of circulating tumor cells," *Cyborg and Bionic Systems*, vol. 2022, p. 9829287, Jan 2022.

[28] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018.

[29] H. Fan, B. Xiong, K. Mangalam, Y. Li, Z. Yan, J. Malik, and C. Feichtenhofer, "Multiscale vision transformers," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct 2021, pp. 6804–6815.

[30] S. Ahmmed, P. Podder, M. Mondal, S. Rahman, S. Kannan, M. Hasan, A. Rohan, and A. Prosvirin, "Enhancing brain tumor classification with transfer learning across multiple classes: An in-depth analysis," *BioMedInformatics*, vol. 3, no. 4, pp. 1124–1144, Dec 2023.

[31] S. M. Alzahrani, "Convattenmixer: Brain tumor detection and type classification using convolutional mixer with external and self-attention mechanisms," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 10, p. 101810, Dec 2023.

[32] I. Galić, M. Habijan, H. Leventić, and K. Romić, "Machine learning empowering personalized medicine: A comprehensive review of medical image analysis methods," *Electronics*, vol. 12, no. 21, p. 4411, Oct 2023.

[33] K. Genereux and T. Akilan, "An efficient cnn-bilstm-based model for multi-class intracranial hemorrhage classification," in *Proc. 8th Int. Conf. Image, Vision and Computing (ICIVC)*, Dalian, China, 2023, pp. 303–309.

[34] S. Chen *et al.*, "Rna adenosine modifications related to prognosis and immune infiltration in osteosarcoma," *Journal of Translational Medicine*, vol. 20, no. 1, p. 228, Dec 2022.

[35] H. Huang, N. Wu, Y. Liang, X. Peng, and J. Shu, "Slnl: A novel method for gene selection and phenotype classification," *International Journal of Intelligent Systems*, vol. 37, no. 9, pp. 6283–6304, Sep 2022.

[36] B. He, C. Dai, J. Lang, P. Bing, G. Tian, B. Wang, and J. Yang, "A machine learning framework to trace tumor tissue-of-origin of 13 types of cancer based on dna somatic mutation," *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, vol. 1866, no. 11, p. 165916, Nov 2020.

[37] T. Hussain and H. Shouno, "Explainable deep learning approach for multi-class brain magnetic resonance imaging tumor classification and localization using gradient-weighted class activation mapping," *Information*, vol. 14, no. 12, p. 642, Nov 2023.

[38] E. Jun, S. Jeong, D.-W. Heo, and H.-I. Suk, "Medical transformer: Universal brain encoder for 3d mri analysis," *arXiv:2104.13633*, 2021.

[39] M. Li, R. Wei, Z. Zhang, P. Zhang, G. Xu, and W. Liao, "Cvt-based asynchronous bci for brain-controlled robot navigation," *Cyborg and Bionic Systems*, vol. 4, p. 0024, Jan 2023.

[40] X. Si, H. He, J. Yu, and D. Ming, "Cross-subject emotion recognition brain–computer interface based on fnirs and dbjnet," *Cyborg and Bionic Systems*, vol. 4, p. 0045, Jan 2023.

[41] C. Bao, X. Hu, D. Zhang, Z. Lv, and J. Chen, "Predicting moral elevation conveyed in danmaku comments using eegs," *Cyborg and Bionic Systems*, vol. 4, p. 0028, Jan 2023.

[42] R. L. Kumar, J. Kakarla, B. V. Isunuri, and M. Singh, "Multi-class brain tumor classification using residual network and global average pooling," *Multimedia Tools and Applications*, vol. 80, no. 9, pp. 13 429–13 438, Apr 2021.

[43] W. Luo, H. Niu, J. Hu, Y. Cai, D. Ergu, and H. Lan, "Universal medical image segmentation with task-specific prompt-guided transformer model," in *Proc. Int. Annu. Conf. Complex Syst. Intell. Sci. (CSIS-IAC)*, Oct 2023, pp. 569–575.

[44] P. A. Meshram, S. Joshi, and D. Mahajan, "Brain tumor detection using swin transformers," *arXiv:2305.06025*, 2023.

[45] K. Prathaban, B. Wu, C. L. Tan, and Z. Huang, "Detecting tumor infiltration in diffuse gliomas with deep learning," *Advanced Intelligent Systems*, vol. 5, no. 12, p. 2300397, Dec 2023.

[46] A. B. Ramakrishnan, M. Sridevi, S. K. Vasudevan, R. Manikandan, and A. H. Gandomi, "Optimizing brain tumor classification with hybrid cnn architecture: Balancing accuracy and efficiency through oneapi optimization," *Informatics in Medicine Unlocked*, vol. 44, p. 101436, Jun 2024.

[47] F. Shamshad, S. Khan, S. W. Zamir, M. H. Khan, M. Hayat, F. S. Khan, and H. Fu, "Transformers in medical imaging: A survey," *Medical Image Analysis*, vol. 88, p. 102802, 2023.

[48] D. Rastogi, P. Johri, V. Tiwari, and A. A. Elngar, "Multi-class classification of brain tumour magnetic resonance images using multi-branch network with inception block and five-fold cross validation deep learning framework," *Biomedical Signal Processing and Control*, vol. 88, p. 105602, Feb 2024.

[49] Z. Liu, L. Chen, H. Cheng, J. Ao, J. Xiong, X. Liu, Y. Chen, Y. Mao, and M. Ji, "Virtual formalin-fixed and paraffin-embedded staining of fresh brain tissue via stimulated raman cyclegan model," *Science Advances*, vol. 10, no. 13, p. 3426, Mar 2024.

[50] C. Hu, T. Xia, Y. Cui, Q. Zou, Y. Wang, W. Xiao, S. Ju, and X. Li, "Trustworthy multi-phase liver tumor segmentation via evidence-based uncertainty," *Engineering Applications of Artificial Intelligence*, vol. 133, p. 108289, Jul 2024.

[51] D. Tan and X. Liang, "Multiclass malaria parasite recognition based on transformer models and a generative adversarial network," *Scientific Reports*, vol. 13, no. 1, p. 17136, Oct 2023.

[52] G. S. Tandel, A. Balestrieri, T. Jujaray, N. N. Khanna, L. Saba, and J. S. Suri, "Multiclass magnetic resonance imaging brain tumor classification using artificial intelligence paradigm," *Computers in Biology and Medicine*, vol. 122, p. 103804, Jul 2020.

[53] Y. Zheng, D. Huang, X. Hao, J. Wei, H. Lu, and Y. Liu, "Univisnet: A unified visualization and classification network for accurate grading of gliomas from mri," *Computers in Biology and Medicine*, vol. 165, p. 107332, Oct 2023.

[54] M. Zhou, M. W. Wagner, U. Tabori, C. Hawkins, B. B. Ertl-Wagner, and F. Khalvati, "Generating 3d brain tumor regions in mri using vector-quantization generative adversarial networks," *arXiv:2310.01251*, 2023.

[55] P. Parshapa and P. I. Rani, "A survey on an effective identification and analysis for brain tumour diagnosis using machine learning technique," *International Journal of Recent Innovation Trends in Computing and Communication*, vol. 11, no. 3, pp. 68–78, Apr 2023.

[56] D. Zhu and D. Wang, "Transformers and their application to medical image processing: A review," *Journal of Radiation Research and Applied Sciences*, vol. 16, no. 4, p. 100680, Dec 2023.

[57] C. Yang, D. Sheng, B. Yang, W. Zheng, and C. Liu, "A dual-domain diffusion model for sparse-view ct reconstruction," *IEEE Signal Processing Letters*, vol. 31, pp. 1279–1283, 2024.

[58] W. Zheng, S. Lu, Y. Yang, Z. Yin, and L. Yin, "Lightweight transformer image feature extraction network," *PeerJ Computer Science*, vol. 10, p. e1755, Jan 2024.

[59] Y. Liu, Y. Zhang, Y. Wang, F. Hou, J. Yuan, J. Tian, Y. Zhang, Z. Shi, J. Fan, and Z. He, "A survey of visual transformers," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 6, pp. 7478–7498, Jun 2023.

[60] E. Irmak, "Multi-classification of brain tumor mri images using deep convolutional neural network with fully optimized framework," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 45, no. 3, pp. 1015–1036, Sep 2021.

[61] T. Sadad, A. Rehman, A. Munir, T. Saba, U. Tariq, N. Ayesha, and R. Abbasi, "Brain tumor detection and multi-classification using advanced deep learning techniques," *Microscopy Research and Technique*, vol. 84, no. 6, pp. 1296–1308, Jan 2021.

[62] M. B. Naceur, M. Akil, R. Saouli, and R. Kachouri, "Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy," *Medical Image Analysis*, vol. 63, p. 101692, Jul 2020.

[63] K. Muhammad, S. Khan, J. D. Ser, and V. H. C. D. Albuquerque, "Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 507–522, Feb 2021.

[64] S. Kumar and D. P. Mankame, "Optimization driven deep convolution neural network for brain tumor classification," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 3, pp. 1190–1204, Jul 2020.

[65] A. A. Asiri, A. Shaf, T. Ali, M. Aamir, M. Irfan, and S. Alqahtani, "Enhancing brain tumor diagnosis: An optimized cnn hyperparameter model for improved

accuracy and reliability," *PeerJ Computer Science*, vol. 10, p. e1878, Mar 2024.

[66] A. A. Asiri, A. Shaf, T. Ali, M. A. Pasha, M. Aamir, M. Irfan, S. Alqahtani, A. J. Alghamdi, A. H. Alghamdi, A. F. A. Alshamrani, M. Alelyani, and S. Alamri, "Advancing brain tumor classification through fine-tuned vision transformers: A comparative study of pre-trained models," *Sensors*, vol. 23, no. 18, p. 7913, Sep 2023.

[67] A. A. Asiri, M. Aamir, T. Ali, A. Shaf, M. Irfan, K. M. Mehdar, S. M. Alqhtani, A. H. Alghamdi, A. F. A. Alshamrani, and O. M. Alshehri, "Next-gen brain tumor classification: Pioneering with deep learning and fine-tuned conditional generative adversarial networks," *PeerJ Computer Science*, vol. 9, p. e1667, Nov 2023.

[68] E. Gürsoy and Y. Kaya, "Brain-gcn-net: Graph-convolutional neural network for brain tumor identification," *Computers in Biology and Medicine*, vol. 180, p. 108971, Sep 2024.

[69] A. K. Alzahrani, A. A. Alsheikhy, T. Shawly, A. S. Azzahrani, and A. I. AbuEid, "Amyotrophic lateral sclerosis prediction framework using a multi-level encoders–decoders-based ensemble architecture technology," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 2, p. 101960, Feb 2024.

[70] T.-W. Wang, Y.-C. Shiao, J.-S. Hong, W.-K. Lee, M.-S. Hsu, H.-M. Cheng, H.-C. Yang, C.-C. Lee, H.-C. Pan, W. C. You, J.-F. Lirng, W.-Y. Guo, and Y.-T. Wu, "Artificial intelligence detection and segmentation models: A systematic review and meta-analysis of brain tumors in magnetic resonance imaging," *Mayo Clinic Proceedings: Digital Health*, vol. 2, no. 1, pp. 75–91, Mar 2024.

[71] E. Moya-Sáez, R. De Luis-García, L. Nunez-Gonzalez, C. Alberola-López, and J. A. Hernández-Tamames, "Brain tumor enhancement prediction from pre-contrast conventional weighted images using synthetic multiparametric mapping and generative artificial intelligence," *Quantitative Imaging in Medicine and Surgery*, vol. 15, no. 1, pp. 42–54, Jan 2025.

[72] M. Sravani, S. Aparna, J. Sabarinath, and Y. Kakarla, "Enhancing brain tumor diagnosis with generative adversarial networks," in *Proc. 14th Int. Conf. Cloud Comput., Data Sci. Eng.*, Jan 2024, pp. 846–851.