Deep Learning-Based MRI Analysis for Brain Tumor Classification

By Ying Wu

B.S. in Chemistry, June 2010, University of Science and Technology of China
Ph.D. in Chemistry, October 2015, Columbia University

A Praxis submitted to

The Faculty of
The School of Engineering and Applied Science
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Engineering

August 31, 2025

Praxis directed by

Shahryar Sarkani
Adjunct Professor of Engineering Management and Systems Engineering

The School of Engineering and Applied Science of The George Washington University certifies that Ying Wu has passed the Final Examination for the degree of Doctor of Engineering as of July 1, 2025. This is the final and approved form of the Praxis.

Deep Learning-Based MRI Analysis for Brain Tumor Classification

Ying Wu

Praxis Research Committee:

Shahryar Sarkani, Adjunct Professor of Engineering Management and Systems Engineering, Praxis Director

John Fassaceca, Professorial Lecturer of Engineering Management and Systems Engineering, Praxis Committee Chair

Vijay Raghavan, Professorial Lecturer in Engineering and Applied Science, Committee Member

© Copyright 2025 by Ying Wu
All rights reserved

Dedication

I would like to dedicate this praxis work to my husband, Zhang, and my son, Grant, whose love and support have made this journey possible. Thank you for the many weekends you spent exploring the world together- on car shows, airplane watching, and other adventures- so that I could focus on studying. Your patience, understanding, and encouragement gave me the space to grow. Your laughter and presence reminded me why I strive.

This work is as much yours as it is mine.

Acknowledgement

I would like to thank my advisor, Dr. Sarkani, who provided sharp criticism, thoughtful suggestions and constant encouragement, guiding me through every step of this research. Your guidance has been invaluable to my understanding of artificial intelligence, and I am especially grateful that you have believed in my work.

I would also like to thank Dr. Fossaceca for providing thorough academic guidance and encouraging a supportive program, which have equipped me with the tools to navigate this rapidly evolving era of AI and machine learning. Your guidance and focus on excellence have profoundly influenced me.

I would like to express my sincere gratitude to all the faculty members. Your guidance and teaching laid a solid foundation for this praxis project and also enabled me to conduct in-depth explorations in the latest AI/ML technology. Your contribution made the whole D.Eng in AI/ML program truly enriching.

I thank the committee members, Dr. Sarkani, Dr. Fossaceca, and Dr. Raghavan, for your time, critical evaluation, and constructive comments.

Finally, I am deeply thankful for my fellow cohort classmates. Artificial intelligence is evolving rapidly, and it's great to be able to study with such talented and motivated people. All the curiosity, collaboration, and resilience from our cohort made this experience not only inspiring but also memorable.

Abstract of Praxis

Deep Learning-Based MRI Analysis for Brain Tumor Classification

Brain tumor diagnosis demands rapid and reliable analysis of medical image scans, yet routine magnetic-resonance imaging (MRI) analysis is labor-intensive, error-prone, and constrained by shortages of expert neuroradiologists. This praxis developed an end-to-end deep-learning model that automatically classifies brain MRI slices into four categories—no-tumor, meningioma, glioma, and pituitary adenoma and additionally explains the image regions that drive each decision with saliency map analysis.

Two public datasets were merged (3064 brain tumor MRI scans and 1162 healthy brain MRI scans), followed by pre-processing, data cleaning, and image augmentation with geometric and intensity transformations. A ResNet-101 backbone, pretrained on ImageNet, was fine-tuned by unfreezing the deepest residual block and attaching a four-node soft-max head. Hyper-parameter search and regularization techniques yielded a model with 99.1 % overall accuracy, macro-precision 98.9 %, and macro-recall 99.2 % on a 20 % patient-wise hold-out test set- surpassing the 95 % target. Vision-transformer (Swin) and two-stage transformer hybrids were evaluated but under-performed (< 75 % tumor-type accuracy), underscoring the efficiency of residual CNNs for modest medical datasets.

Interpretability of the fine-tuned ResNet101 model was assessed with saliency map as well as through a dual-model strategy: final-layer saliency maps from the classifier were over-laid on tumor masks generated by a pretrained U-Net segmentation

model. Histograms of intersection-over-union scores revealed that salient pixels clustered predominantly at tumor margins, mirroring radiologic practice and supporting Hypothesis 1 that tumors' boundary features dominate classification results.

For the fine-tuned ResNet101 model, inference time on an A100 GPU was < 3 s for a 60-slice study. The low requirement on computational resources enables real-time triage, particularly as a rapid first-pass screening tool in the clinical workflow as the model achieved excellent results in classifying non-tumor scans.

The praxis contributes by developing (i) a deep learning model for margin-aware four-class brain-MRI image classification. In particular, the model has great potential to serve as a rapid and automated first-pass screening tool in the clinical workflow; (ii) a transparent "explained-as-you-screen" workflow that can visualize the most important image pixels for the model's classification decisions; and (iii) an analysis of transformer limitations in the context of limited image data setting.

Table of Contents

Dedication ..... **iv**

Acknowledgement ..... **v**

Abstract of Praxis ..... **vi**

List of Figures ..... **xi**

List of Tables ..... Error! Bookmark not defined.

Chapter 1- Introduction **1**

1.1 Background 1

1.2 Research Motivation 3

1.3 Problem Statement 4

1.4 Thesis Statement 5

1.5 Research Objectives 7

1.6 Research Questions and Hypotheses 8

1.7 Scope of Research 8

1.8 Research Limitations 9

1.9 Organization of Praxis 10

Chapter 2- Literature Review and Background **11**

2.1 Introduction 11

2.2 Evolution of Deep Learning Models for Classifying Brain Tumors 12

2.3 Datasets and Data Processing 22

2.4 Current Limitations and Challenges 25

2.5 Research Gaps and Opportunities 27

2.6 Summary 28

Chapter 3- Methodology 30

3.1 Introduction 30

3.2 Collect Data and Perform EDA 32

3.3 Data Preprocessing 34

3.4 Model Development 36

3.5 Training Configuration of the ResNet101-based Model 45

3.6 Feature Importance Analysis and Visualization for the ResNet101-based Model 47

3.7 Limitations of Methodology 53

Chapter 4- Results 55

4.1 Introduction 55

4.2 ResNet101 Model Performance 56

4.3 Feature Importance Analysis for Model Interpretation of the ResNet101-based Model 60

4.4 SWIN Transformer Experiments 66

4.5 Summary of Findings 67

Chapter 5 — Discussion and Conclusions 69

5.1 Discussion 69

5.2 Conclusions and Practical Implications 70

5.3 Recognized Limitations 72

5.4 Contributions of Research..... 73

References..... 75

List of Figures

Figure 1. CNN-based UNet architecture. 14

Figure 2. Comparison of different deep learning approaches for brain MRI analysis, highlighting SWIN transformer architecture with shifted window. 17

Figure 3. Hybrid architecture of UNet- Swin transformer. 20

Figure 4. Graphical Model of Research- level 0 summary 31

Figure 5. Graphical Model of Research- Level 1 breakdown for MRI dataset. 32

Figure 6. Class label counts in the combined dataset. 33

Figure 7. Samples of MRI from 4 classes. 34

Figure 8. Graphical Model of Research- Level 1 breakdown for data preprocessing. 35

Figure 9. Graphical Model of Research- Level 1 breakdown for model development. 37

Figure 10. Custom modifications on ResNet101 architecture in this praxis. 39

Figure 11. Graphical Model of Research- Level 1 breakdown for hyperparameter tuning. 46

Figure 12. Graphical Model of Research- Level 1 breakdown for model interpretability with feature importance analysis. 48

Figure 13. Confusion matrix from the held-out test data. 57

Figure 14. An example of the extracted saliency maps from final, middle and early layers of the ResNet101-based model for a meningioma MRI test image. 62

Figure 15. IoU calculation for two areas, U-Net segmented tumor area and saliency map area for important features in an MRI scan. 63

Figure 16. Histograms that show the distribution of IoU (U-Net segmentation mask and saliency map) in four classes. 63

Figure 17. Three representative scenarios for IoU values calculated from the saliency map and the U-Net tumor segmentation mask. 65

List of Tables

Table 1: Detailed Performance Metrics by Class..... 59

Chapter 1- Introduction

1.1 Background

Brain tumors have become a major healthcare challenge, especially in the aging population. There are diverse brain tumor types ranging from benign to malignant. The most common brain tumors are gliomas, meningiomas, and pituitary adenomas. These tumor types have different tumor growth rates, response to treatment, and progression to late stage. (Chatterjee et al., 2022). For example, high-grade gliomas generally require aggressive treatment. Meningiomas can be effectively controlled with surgery alone, and pituitary adenomas may respond to targeted therapies (Ostrom et al., 2020). It is critical to identify brain tumor types accurately and timely for selecting appropriate treatment options and improving patient outcomes. If the diagnosis is inaccurate or delayed, the patient may not receive the best treatment plan, which reduces patient's chance of survival. Therefore, accurate and effective diagnostic tools are required to help medical professionals offer the best treatment for patients.

Magnetic resonance imaging (MRI) is the main imaging technique for the diagnosis of brain tumors. It is preferred because MRI does not need invasive measures such as surgery (Chatterjee et al., 2022). MRI shows anatomical details of the brain, allowing medical professionals to see and assess tumor types and conditions and thereby providing medical appropriate treatment to patients based on MRI analysis. MRI is more superior at soft-tissue contrast than other imaging techniques, and that helps to see details and abnormalities in the brain with clearer visualization. Though MRI technology has

made great advancement, it remains a difficult and time-consuming task for radiologists to interpret MRI and provide diagnosis (Bhatia et al., 2019; Najjar, 2023). Manually reviewing various MRI slices to identify brain tumor types requires deep expertise and is prone to error. The current manual process may delay critical diagnosis and potentially impact patient treatments.

Artificial intelligence has made significant achievements in improving the analysis and augmenting the effectiveness of various medical imaging techniques. The application of AI technology in medical treatment has been approved by the Food and Drug Administration (FDA). For example, AI-powered tools have been FDA-approved for detecting diabetic retinopathy with retinal images (Ting et al., 2019; FDA news, 2018) and for screening breast cancers with mammograms (McKinney et al., 2020; news, 2023). Current deep learning technology has proven that in the field of medical imaging, it can help doctors enhance the accuracy of diagnosis and make examinations faster and more efficient. There have been many recent advances in computer vision and deep learning, especially convolutional neural networks (CNNs) such as ResNet and U-Net (Huang et al., 2020; Cao et al., 2022), and transformer-based visual transformers such as the Swin transformer (Liu et al., 2021), which have shown outstanding capabilities in completing medical image analysis tasks more efficiently and accurately. These advancements in computer vision offer an opportunity to enhance the process of brain tumor classification from MRI scans, potentially revolutionizing the field of brain tumor diagnosis by improving accuracy, reducing inter-observer variability, and accelerating the diagnosis process.

1.2 Research Motivation

The motivation for this praxis research lies in the need for more efficient and consistent MRI analysis to diagnose brain tumors. The research aims to address the lack of such automatic diagnostic tools in the field, with great potential of integrating such deep learning models into the clinical workflow of MRI-based brain tumor diagnosis. By leveraging deep learning techniques, there is a great potential to improve the speed, accuracy, and consistency of brain tumor diagnoses, leading to faster diagnosis, more timely treatments, and ultimately better patient outcomes.

In recent years, the U.S. Food and Drug Administration (FDA) has approved AI-powered software for brain MRI analysis for neuroinflammatory or neurodegenerative diseases (Pixyl news, 2023; VUNO news, 2023), demonstrating the growing potential as well as acceptance of deep learning models in clinical practice relating to brain imaging. Despite these advancements, a notable gap remains: there is no FDA-approved AI software specifically designed for brain tumor detection and classification using MRI. This gap serves a primary motivation to conduct research into development of a deep learning-driven MRI analysis for accurate and efficient brain tumor classification.

In the publicly available literature, several publications attempted to develop deep learning models for brain tumor classification. Unfortunately, the publications either failed to include non-tumor as a class in the classification model (Sultan et al., 2019; Amou et al., 2022; Deepa et al., 2023) or only focused on classifying tumor and non-tumor without the depth of further analyzing brain tumor types (Khan et al., 2020). This gap in the research presented a significant opportunity for improvement. A

comprehensive model that could accurately distinguish between non-tumor cases and various types of brain tumors would be more beneficial in clinical application.

This research aims to contribute to the growing field of AI-assisted medical imaging analysis by researching on available datasets and experimenting with established and emerging deep learning architectures to develop an efficient and accurate deep learning model for MRI-based classification of major brain tumor types and differentiation of healthy brain tissue.

1.3 Problem Statement

Diagnosing brain tumors from MRI images requires deep expertise and is highly error-prone, which leads to misdiagnosis and/or diagnostic delays. These limitations in diagnosis accuracy and speed can directly delay treatment for patients, who may miss the best time for intervention. The variability in inter-observer assessments of brain MRI scans further emphasizes the need for an automated method to enhance the consistency and reliability of brain tumor diagnoses.

As the population ages, the growing cases of brain tumors worldwide lead to an increased demand for more efficient diagnostic tools of brain tumors. Data published by the World Health Organization estimates that central nervous system tumors have 300,000 new cases diagnosed each year with a high mortality rate (Brain tumor statistics news, 2024). The climbing cases of brain tumor, plus complex healthcare needs of an aging population increase challenges to healthcare systems and highlight the need for advanced, automated diagnostic solutions to help alleviate the healthcare burden.

Furthermore, equitable access to healthcare is a gap in resource-limited regions, where there is a significant shortage of specialized radiologists. Patients in these areas often find it difficult to get a correct diagnosis in a timely manner because of insufficient medical resources and limited number of specialists, often leading to poorer survival rates. The high cost and scarcity of specialized neuroradiology expertise present a significant barrier to accessing high-quality diagnostic capabilities. The reliance on manual interpretation by a limited number of experts creates bottlenecks in the diagnostic process.

1.4 Thesis Statement

A deep learning-based predictive model for the analysis of brain MRI is needed to improve the speed and accuracy of brain tumor diagnosis. Such an AI model can help achieve more timely, equitable, and effective patient care. To achieve healthcare-access equity in the field of brain tumor diagnosis, more accessible and efficient diagnostic tools are urgently needed to help narrow the gap. An AI-based diagnostic solution can improve patient access to high-quality diagnostic capabilities by reducing the reliance on high-demand expertise and improve the efficiency of radiologists; thereby improving healthcare-access equity.

The development of an efficient and accurate AI-based solution for brain tumor classification is expected to address these issues and lead to faster diagnoses, better healthcare access, and improved patient outcomes, as demonstrated by an increasing trend of AI-powered imaging analysis tools approved by the FDA.

This praxis research focuses on developing a deep learning-based model for classifying brain tumors and non-tumorous conditions from MRI images. The model can be deployed in the clinic and facilitate healthcare providers to optimize their current diagnostic workflows to meet the increasing demand of brain tumor patients.

This praxis research aims to develop a predictive model for classifying three types of brain tumors (meningioma, glioma, and pituitary tumors) as well as identifying non-tumorous conditions. The goal of this AI model is to provide a fast, consistent and automated analysis for brain tumor classification, directly addressing the issue of high reliance on expertise and error-prone manual interpretation as identified in the problem statement. In summary, the model aims to address these challenges: existing diagnostic approaches require high expertise, are prone to human error, and more efficient tools were needed in the face of the rising number of brain tumor cases.

AI models can reduce inter-observer variations via uniform automation, decrease reliance on the scarce expertise by handling most of the easier analysis tasks, and allow human expertise to focus on providing timely diagnoses for the more complicated cases. Therefore, integrating this AI-driven approach into clinical workflows will not replace radiologists but rather make their work easier and augment their capabilities. It will allow them to focus on more complex cases and provide timely interpretations. This collaboration between AI and human experts will have a great potential to improve overall diagnostic accuracy and efficiency of brain tumor diagnosis.

The models in this praxis would utilize computer vision algorithms such as ResNet-based model, which had shown great promise in the field of medical image

analysis. Leveraging fine-tuning techniques and techniques for optimizing models' generalizability, the praxis research aims to develop a model that achieves high accuracy while maintaining computational efficiency, making it suitable for integration into clinical workflows.

1.5 Research Objectives

The main objectives of this research are:

1. To use publicly available datasets of brain MRI images, to design and implement deep learning models for classifying different types of brain tumors and identifying non-tumor cases.

2. To use a CNN architecture, such as ResNet, with additional training through fine-tuning to achieve optimal model performance and generalizability.

3. To evaluate the performance of the developed model with classification metrics such as accuracy, precision, sensitivity, and F1 score.

4. To extract feature importance from ResNet-based model, which helps interpret which part of the tumor scan the model relies on for the classification decision.

5. To add to the growing body of research on AI-enabled analysis of medical imaging, particularly for classifying brain tumor types and non-tumorous conditions from MRI scans.

1.6 Research Questions and Hypotheses

Q1: Does the model's feature importance indicate greater reliance on tumor margin rather than tumor core for brain tumor classification?

Q2: Can a predictive model for brain MRI analysis be developed to improve the speed and accuracy of tumor diagnosis, helping to achieve timely and effective patient care?

Q3: Which deep learning algorithm leads to the most improvement in speed and accuracy of tumor diagnosis in brain MRI analysis?

H1: Features from the tumor margin will contribute more significantly to brain tumor classification decisions than the tumor core features.

H2: A predictive model for brain MRI analysis will correctly classify brain tumors with accuracy of 95% or greater.

H3: Fine-tuning ResNet will perform 10% better than Swin transformer in brain tumor classification.

1.7 Scope of Research

This research focuses on the development and evaluation of deep learning models for brain tumor classification using MRI images. The scope includes:

- Analysis of existing deep learning architectures suitable for medical image classification.

- Fine-tune and implement a ResNet-based model for classifying brain tumor types and non-tumorous conditions from brain MRI images and explore a vision transformer model based on SWIN transformer for brain tumor classification.
- Training and validation of the model using publicly available brain MRI datasets
- Performance evaluation of the model using metrics such as accuracy, precision, sensitivity and speed.
- Extract feature importance from the ResNet-based model to understand which part of the tumor the model relies on for brain tumor classification.
- Comparison of the model performance for brain tumor classification between ResNet-based model and SWIN-transformer-based model.

1.8 Research Limitations

This research aims to address challenges in brain tumor diagnosis, and it is important to acknowledge certain limitations:

- The study only considers the analysis of MRI images and does not consider other imaging techniques.
- The performance of the model is dependent on the quality and diversity of training datasets that were available; thus, only the major brain tumor types are considered.
- The research focuses on the technical aspects of model development and does not include clinical trials or integration into existing hospital systems.

1.9 Organization of Praxis

This praxis is organized as follows:

Chapter 1 introduces the background of brain tumor diagnosis, MRI technology and its current use for diagnosing brain tumors, and the potential of employing AI and deep learning models to improve the MRI-based brain tumor diagnosis.

Chapter 2 provides a literature review of current brain tumor classification methods, deep learning techniques in medical imaging, and relevant CNN architectures.

Chapter 3 describes the methodology used in this praxis research, including data preprocessing, data augmentation, model design, model training and model evaluation.

Chapter 4 presents the results of the study, including model performance metrics, feature importance analysis, and comparative analyses.

Chapter 5 discusses the implications of the findings, model explainability, addresses the research objectives, and suggests areas for future research.

This praxis concludes with references.

Chapter 2- Literature Review and Background

2.1 Introduction

In the last decade, deep learning models have made many advancements in medical image analysis, such as for brain tumor classification. This chapter reviews the latest research on the development of deep learning models to classify brain tumors with MRI scans from three components relating to the model development- deep learning architectures, MRI medical image datasets and preprocessing methods, and current challenges and limitations in the field of brain tumor classification from medical images.

Deep learning models are potentially powerful automated tools for brain tumor classification in clinical practice by reducing inter-observer variability, accelerating diagnosis processes, and improving the accuracy of tumor diagnosis (Tandel et al., 2019). The progression from traditional machine learning models to more powerful deep learning architectures has improved the way medical images are analyzed and interpreted (Akkus et al., 2017). Unlike traditional machine learning methods that rely heavily on hand-crafted features, deep learning models can automatically learn hierarchical representations directly from raw imaging data, making deep learning models more suitable and more useful for the application in medical image analysis as features in medical images are not easily hand-crafted.

For the automated analysis of medical images, deep learning has shown good performance in identifying and segmenting abnormalities such as tumors, as well as classifying and diagnosing diseases (Litjens et al., 2017; Zhou et al., 2021). There are two

main driving forces behind this success: one is the technological breakthroughs in computational capabilities and the other is the increasing availability of medical imaging datasets for model training, similar to the contributors for the advancements of deep learning models in other areas such as large language models (LLM). This chapter provides a review of the current state of research in automated brain tumor classification with deep learning, by reviewing three key areas:

- Deep Learning Algorithms: the chapter reviews algorithms, including Convolutional Neural Networks (CNNs), transformer-based architectures such as Vision Transformers (ViT) and hybrid models by combining the strength of both CNN and ViT, and review their advantages and disadvantages for medical image analysis (Liu et al., 2021; Cao et al., 2022).
- Data Collection and Preprocessing Methods: The chapter also reviews publicly available datasets and preprocessing techniques which were designed to optimize model performance (Cheng et al., 2015; Mzoughi et al., 2020).
- Current Challenges and Limitations: The chapter also reviews the challenges in the field, such as insufficient data, class imbalance, and the need for interpretability of AI frameworks in clinical settings (Pashaei et al., 2018).

2.2 Deep Learning Models for Classifying Brain Tumors

2.2.1 Convolutional Neural Networks (CNNs)

CNNs have emerged as a key architecture for medical image analysis by revolutionizing the way we process and analyze medical imaging data. This section

focuses on reviewing CNN algorithms for the application of brain tumor classification. The key achievements of CNN models are described as follows.

Basic CNN Implementations: Sultan et al. used a basic CNN to identify three different types of brain tumors with an accuracy of 96.13%, and Abiwinanda et al. reported 84.19% validation accuracy. These implementations demonstrated the potential of CNNs for classifying brain tumors from medical images, though the publications primarily focused on classifying tumor types and missed the non-tumorous class in the model development (Sultan et al., 2019; Abiwinanda et al., 2019).

Pre-trained CNN with transfer-learning: a recent focus of deep learning model developments is to leverage transfer learning to address specific applications such as for brain tumor classification. Ullah et al. employed transfer learning approaches using pre-trained CNNs, demonstrating how to adapt the pre-training system to classify brain tumors (Ullah et al., 2020). More cutting-edge CNN variant models based on transfer-learning have also emerged in this field. For example, Deepak and Ameer (2019) reported using the features from GoogleNet with the help of transfer learning (Deepak et al., 2019).

ResNet Architecture Evolution: The introduction of Residual Networks (ResNet) made major architectural progress in the field of deep learning-based computer vision models. ResNet solved the degradation problem encountered by deep networks by using skip connections, which allows the gradient to flow smoothly for training much deeper networks (He et al., 2016). For example, Çinar, et al. achieved 97.2% accuracy using a modified ResNet-50 architecture for classifying tumor and non-tumor classes

from brain MRI scans (Çinar et al., 2020). Residual connections helped preserve fine-grained details in medical imaging data that are critical for accurately detecting and classifying brain tumors.

EfficientNet's Scaled Architecture: EfficientNet introduced a novel compound scaling method that uniformly adjusted network depth, width, and resolution by using a compound coefficient (Tan et al., 2019). This balanced scaling was particularly useful in computationally demanding tasks such as brain tumor classification. EfficientNet models achieved state-of-the-art accuracy though using significantly fewer parameters compared to previous architecture. Fast processing speed is advantageous for the model deployment in clinical settings where computational resources may be limited.

Development of Specialized Architectures: Specialized frameworks such as U-Net (see Figure 1), have made great advancements in medical image analysis.

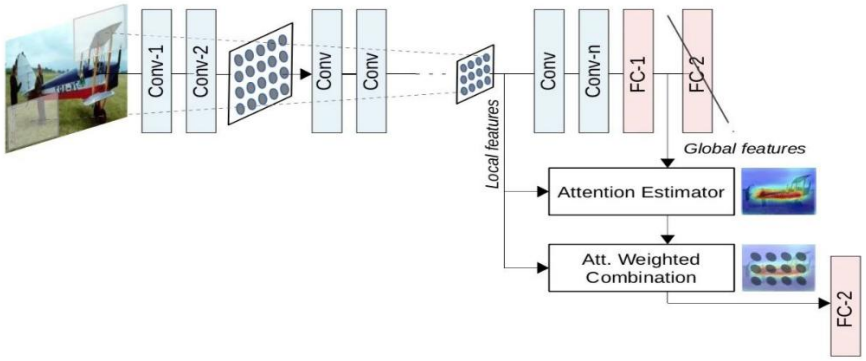


Figure 1. **CNN-based UNet architecture.**

The deep network consists of multiple convolutional layers (Conv), fully connected layers (FC) and an attention estimator module that combines local and global features through attention-weighted combination to improve classification performance.

Credit: Reproduced from GitHub repo: https://github.com/sfczekalski/attention_unet

The encoder-decoder structure and skip connections in the U-Net was effective to improve model performance because they preserve spatial data and capture hierarchical features. Huang et al. improved the design of U-Net by adding dense skip-connections, which allowed the reuse of features and smooth flow of gradients (Huang et al., 2020).

Summary of Technical Improvements with CNN:

- Residual connections addressed the degradation problem in training deep networks.
- Transfer-learning successfully leveraged pre-trained CNN models to improve performance of the brain tumor classification task.
- Combining multiple CNN frameworks showed stronger robustness.

Some challenges remain despite the advances:

- The need for large, well-annotated datasets
- Computational resource requirements
- Model interpretability concerns

The advancements show the strengths of CNN architectures for brain tumor classification.

2.2.2 Vision Transformers

Vision Transformers has brought new capabilities to the automated analysis of medical images. Transformers have the capability to capture long-range dependencies through self-attention mechanisms, which enables more comprehensive feature learning across entire images, an advantage over the CNNs.

Fundamental Transformer Architecture: Transformers were primarily used for natural language processing, but were later modified by Dosovitskiy and others to become a model that can process digital vision (Dosovitskiy, et al.; 2020). The basic ViT architecture cuts images into patches, treats these patches as tokens such as words in text, and then processes them through several transformer encoder layers. This approach allows: 1) Global context modeling by using self-attention, 2) Position-aware feature learning, 3) Parallel task processing capabilities for image recognition under high throughput conditions.

Swin Transformer (Figure 2): The Swin Transformer architecture (Liu et al., 2021) solved the computational limitations of standard ViTs via the following measures: 1) Introducing hierarchical feature representation, 2) Using shifted windows to reduce the required resources of computing attention. 3) Using linear computational complexity for image size 4) Better handling of different scales in medical images.

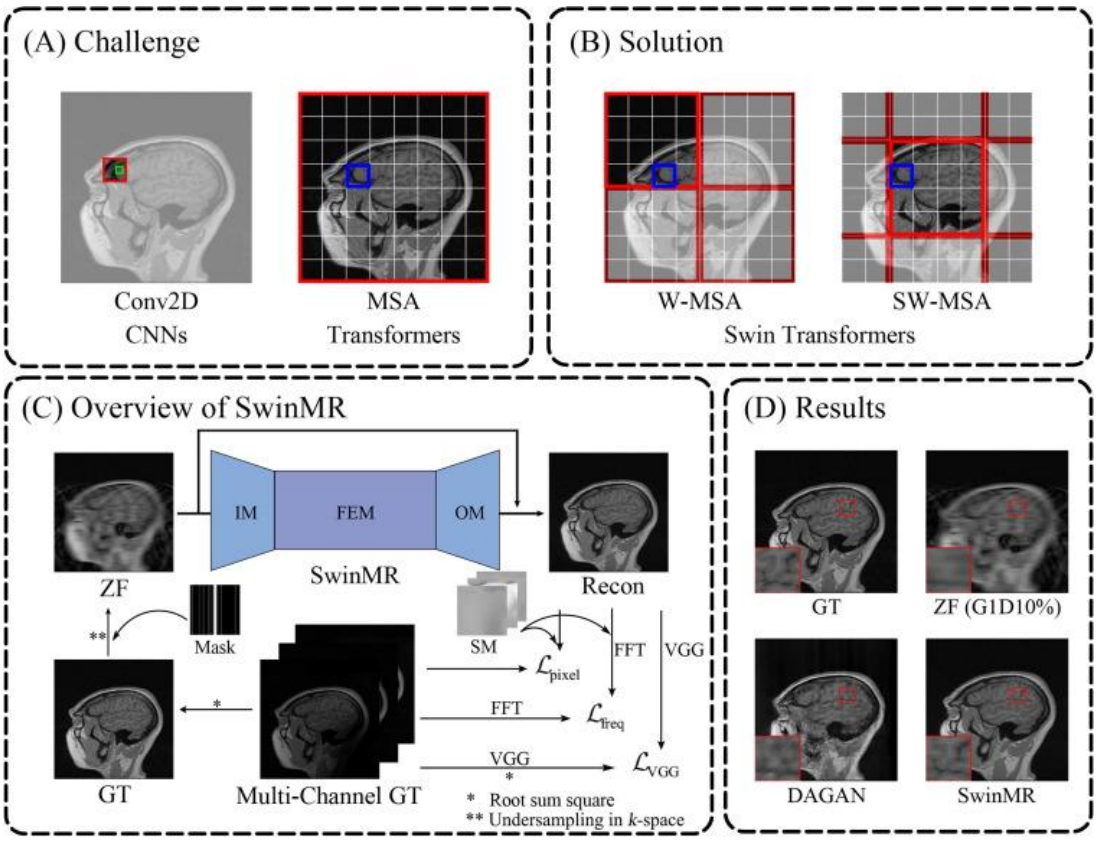


Figure 2. **Comparison of different deep learning approaches for brain MRI analysis, highlighting SWIN transformer architecture with shifted window.**

(A) Traditional CNN and transformer limitations in processing medical images. (B) Swin transformer's solution using window-based attention mechanisms. (C) Overview of the SwinMR architecture showing the image processing pipeline with input, feature extraction modules (IM, FEM, OM), and reconstruction components. (D) Results comparing ground truth (GT), zero-filled (ZF), DAGAN, and SwinMR reconstructions of brain MRI images. Credit

Credit: Reproduced from Fig.1, Huang et al., 2022.

Application of Visual Transformer Models in Medical Imaging:: Several studies have shown that transformers are effective architectures in medical image analysis: Cao

et al. (2022) launched Swin-Unet, which combined the advantages of a transformer and the structure of a U-Net. The Swin-Unet achieved excellent segmentation performance for medical images. Hatamizadeh et al. (2022) developed UNETR, a transformer-based architecture which is specifically designed for medical image segmentation (Hatamizadeh et al., 2022).

The Vision Transformer has many advantages in brain tumor classification, including the ability of capturing the correlation between near and far areas in medical images, and more clearly understanding how tumors interact with the surrounding brain tissue. This enhanced ability for identifying subtle patterns in a medical image is useful for making accurate diagnosis (Cao et al., 2022). Vision Transformers also excelled at feature learning from medical images. The Vision transformer performs well at multi-scale feature representation, and it also handles tumors of various morphologies. The architecture of Vision Transformers naturally deals with long-range dependencies that are present in medical images, which is critical for understanding complex medical imaging data (Cao et al., 2022).

Vision transformers' scalability is another key advantage in medical imaging applications such as MRI-based brain tumor diagnosis. The size of vision transformer models can be adjusted and scaled according to actual applications. The different sizes of models handle medical images of various resolutions effectively, and the models are also flexible to correctly predict MRI images taken from different viewing angles. Compared with previous transformer architecture, the patch-based vision transformers require less computational resources by operating on patches instead of tokenizing every pixel. Therefore, easier computation make vision transformers more suitable for clinical

applications where fast inferencing is required and computational resources may be limited (Cao et al., 2022).

Future research on vision transformers is actively exploring several promising directions, such as designing more powerful attention mechanisms, developing specialized pre-training methods for medical images, and designing lightweight transformers. (Cao et al., 2022).

For example, The Shifted WINdow (SWIN) transformers, which include SWIN-TINY and SWIN-SMALL versions with less number of parameters, is a breakthrough in the field of visual transformers and has significant potential for medical imaging applications. SWIN-TINY has approximately 28M parameters and SWIN-SMALL has approximately 50M parameters, a lighter alternative to the original SWIN structure, though they are still powerful and effective (Cao et al., 2022). The hierarchical architecture with shifted windows in the SWIN models enabled efficient modeling of both local details and global dependencies in medical images. The shifted window strategy helped capture fine details in tumor regions without increasing the burden of computational resources. SWIN models demonstrated promising results in the application of medical image segmentation and classification tasks. For clinical deployment, if computing resources are limited, SWIN-TINY can achieve a good balance between model size and accuracy. If computing resources are sufficient, SWIN-SMALL can offer stronger feature extraction capabilities (Cao et al., 2022).

Vsion transformers still face challenges in the implementation of medical imaging despite their advantages. The challenges include relatively high computational

requirements during the training phase, the need for large, well-annotated datasets, and concerns about model interpretability. Therefore, there has been active advancement in improving vision transformers on medical-focused applications.

2.2.3 Hybrid Architectures Combining CNN and ViT

Recent research has worked on combining the advantages of Convolutional Neural Networks (CNNs) and vision transformers (ViT) to simultaneously leverage the localized feature detection and the global context understanding capabilities. These hybrid architectures (e.g. Figure 3) aim to address the limitations of a pure CNN or a pure transformer approach while keeping their respective advantages.

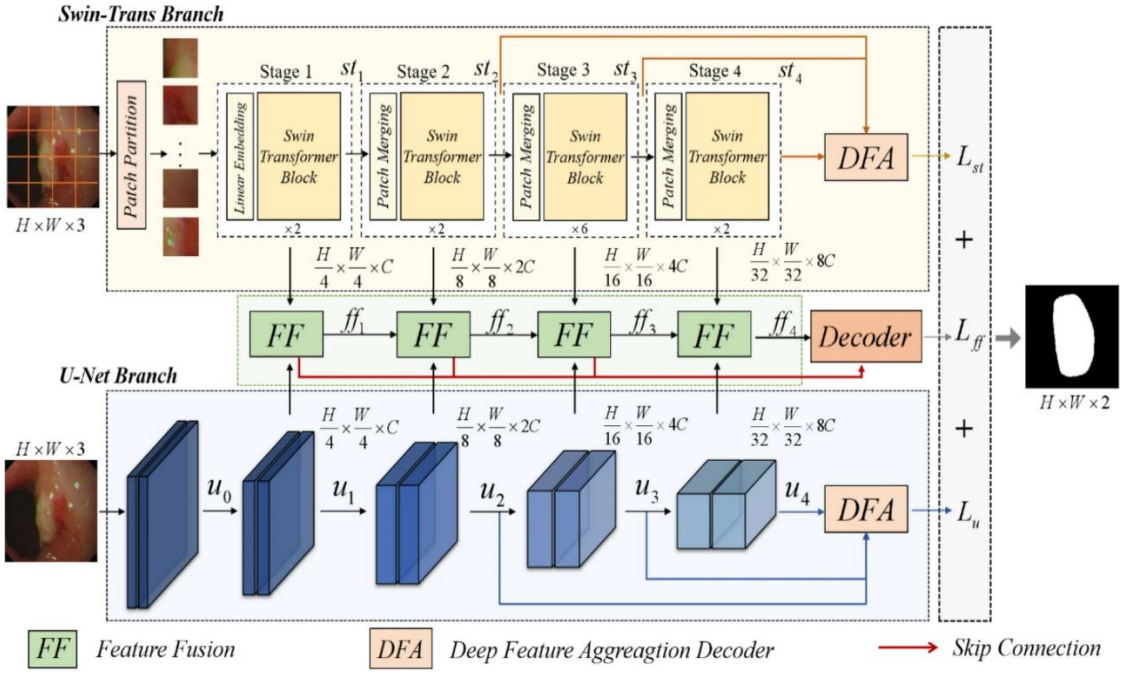


Figure 3. **Hybrid architecture of UNet- Swin transformer.**

The model simultaneously feeds the input da54a into the transformer channel and the convolution channel for processing data, then it fuses the features obtained from the two trajectory together, and finally outputs the segmentation result.

Credit: Reproduced from Fig 1, He et al., 2023.

Below is a brief review of notable examples:

Swin-UNet: Swin-UNet was a model developed by Cao et al. (2022) and it incorporated the Swin transformer module into the U-Net framework. The hybrid Swin-Unet model leveraged both hierarchical feature transformation and shifted window-based self-attention to efficiently process high-resolution medical images. The shifted window mechanism helped capture cross-window connections for maintaining computational efficiency. The authors have experimentally demonstrated that Swin-UNet achieves superior performance in medical image segmentation tasks compared to traditional U-Net or pure transformer approaches.

TransUNet: Chen et al. designed TransUNet that combined both the CNN encoder and transformer layers. The CNN encoder was responsible for extracting low-level features, while the transformer layer was used to process long-distance correlations (Chen et al., 2021). With efficient combination of both local and global feature representations, the model demonstrated performance improvements in various medical segmentation tasks (Chen et al., 2021).

Attention-enhanced CNNs: Several deep learning models were reported to add transformer-style attention mechanisms to traditional CNNs. For example: ConvNeXt

(Liu et al., 2022) introduced transformer-like architectures into CNNs, but the model still preserved the benefits of CNNs' inductive biases. TransBTS (Wenxuan et al., 2021) was reported to use a hybrid CNN-Transformer architecture for the application of brain tumor segmentation.

The success of these hybrid methods were due to their ability to:

- Capture both fine-grained and local features using convolutions
- Model long-range dependencies using self-attention mechanisms
- Maintain computational efficiency by hierarchical feature processing
- Adapt to various image scales by using the multi-resolution feature analysis

Recent studies demonstrated that hybrid architecture outperformed pure CNN or pure transformer models in medical image analysis tasks (Zhou et al., 2021, Lin et al., 2022). For example, the hybrid model that combined local texture features and global contextual information not only identified tumor edges but also distinguished different tumor types, improving the accuracy of predicting brain tumors from MRI images (Isensee et al., 2021).

2.3 Datasets and Data Processing

2.3.1 Available Datasets

The quality of MRI datasets is critically important in developing and validating brain tumor classification models. Several publicly available brain MRI datasets have become standard benchmarks in this field. The Figshare Brain Tumor Dataset (Cheng et al., 2015), which was collected from the Nanfang Hospital and the General Hospital of

Tianjin Medical University, contained 3064 T1-weighted contrast-enhanced MRI images from 233 patients. The brain MRI scans of these patients were divided into three brain tumor categories: meningioma (708 slices), glioma (1426 slices) and pituitary tumor (930 slices). These MRI images were taken in axial, coronal, and sagittal directions. Though this dataset was widely used in research, it lacked non-tumor samples, which made the dataset alone not sufficient for developing a comprehensive diagnostic tool that can be served as a first-pass screening tool for healthy MRI scans.

The Brain Tumor Segmentation (BraTS) challenge dataset is another important resource of brain tumor MRI scans for developing deep learning models. BraTS has become one of the most comprehensive collections for brain tumor imaging datasets since it was established in 2012 (Menze et al., 2014). The dataset provides segmentation masks annotated by experts, making it valuable for training both classification and segmentation models (Bakas et al., 2017).

2.3.2 Data Preprocessing

Data preprocessing is a critical step for achieving optimal performance in training deep learning models. The typical data preprocessing begins with intensity normalization to standardize pixel values to a uniform rage of intensity across different MRI scans from the dataset. Because the intensity of MRI scans may vary greatly in raw data among different scanning sessions and equipment, this intensity normalization step is essential to ensure the success of training models (Nyúl et al., 1999). Another example of medical image data processing related to brain tumor diagnosis, skull stripping is a technique that

removes irrelevant brain tissue from the MRI image, allowing analysis to focus on critical brain areas and thereby making calculations more efficient (Kleesiek et al., 2016).

Noise reduction techniques for image data preprocessing, such as Gaussian filtering or wavelet denoising, were used to improve the image quality for better model training. These noise reduction methods helped remove random variations in pixel intensities from image acquisition, leading to more generalizable and robust models (Manjón et al., 2010). Data preprocessing methods for improving image contrast, such as histogram equalization or adaptive contrast improvement, were also frequently used to optimize the clarity of tumor regions and boundaries (Manjón et al., 2010).

2.3.3 Image Augmentation

Data augmentation is also an important step for improving the generalization ability of models, especially in cases when access to large quantities of labeled medical imaging data is limited or not possible. Image augmentation techniques are selected to balance the introduction of meaningful variability by image augmentation and preserving critical diagnostic information from the original data.

Medical image augmentation techniques used in developing deep learning model include geometric transformations such as rotation with typically ±10-15 degree, ±10-15% scaling and flipping along various axes. Small rotations within a range of ±10-15 degrees are used to simulate the variations of patient positioning during MRI acquisition in the clinic. Scaling transformations of ±10-15% can help the model learn natural fluctuation of tumor size and position within the brain of patients. The translation

operation offsets the image in the x and y directions in small degrees to simulate the slight shaking of the patient that happens during the scan (Shorten et al., 2019).

More advanced augmentation techniques for developing deep learning models include: 1. elastic deformations that simulate anatomical variations in tissues, 2. Gaussian noise which improves the model robustness to variation of image quality, and 3. intensity augmentations that help models adapt to variations under different imaging conditions, such as brightness and contrast adjustments (Perez, 2017).

2.4 Current Limitations and Challenges

2.4.1 Dataset Limitations

Most of the publicly available brain tumor datasets for developing deep learning models focused exclusively on brain tumor cases, lacking the class of normal brain MRI scans (Swati et al., 2019). This dataset limitation restricted the development of comprehensive models that could be capable of distinguishing between healthy and tumor cases.

Issues for the consistency of MRI data quality posed additional challenges for the development of the deep learning model. For example, fluctuation in MRI acquisition parameters from different institutions and/or scanning protocols resulted in inconsistent image features (Pereira et al., 2016). Studies showed that differences in magnetic field strength, slice thickness, or contrast enhancement protocols greatly affected the image features that would be used for developing deep learning models (Deepak et al., 2019). Additionally, inconsistent annotation standards and the inter-observer variability in

ground truth labeling further complicated the development of reliable training datasets (Isunuri et al., 2022).

Class imbalance is a major challenge in the dataset for training deep learning models, including brain tumor classification models. In the real world of clinical practice, some types of tumors have more data than other types of tumors, which occur naturally among patients. This natural distribution results in an uneven distribution of samples in the available dataset for model training. The class imbalance can affect the final tumor classification performance by biasing the model training process, particularly impacting the minor brain tumor class that have limited samples in the training dataset. Research by Cheng et al. demonstrated that class imbalance significantly impacted the model performance, for example, the accuracy between the main class and the minority class differed by up to 15% (Cheng et al., 2015).

2.4.2 Model Limitations

Generalization is a common challenge faced by deep learning models, including application in brain tumor classification. Though many models achieved high performance on specific datasets that the models were trained on, they sometimes failed to perform well when the model was applied to data from different institutions or populations. Research by Sultan et al. revealed significant performance degradation when models were tested on external datasets, with accuracy dropping by about 20% (Sultan et al., 2019).

The interpretability of a deep learning model is a significant concern in clinical applications. The deep neural networks are complex and it is often difficult to understand

how these models arrived at their predictions, which led to a "black box" problem (Ait et al., 2022). The lack of model transparency hindered regulatory approval and clinical adoption of the advanced deep learning models. Recent studies attempted to address the interpretability challenge by using techniques such as attention mechanisms and feature visualization, but significant challenges are still present regarding fully interpreting the decisions made by deep learning models (Trebing et al., 2021).

Intensive requirements on computational resources pose practical challenges for both training and deploying deep learning models. Training advanced deep learning models requires substantial computational resources, including many high-performance GPUs and large amounts of memory (Deepa et al., 2023). Furthermore, the time and computational cost required for model inferencing also has an influence on the application of integrating deep learning models into clinical workflows, especially when real-time analysis is needed.

2.5 Research Gaps and Opportunities

The literature review identified several significant research gaps and opportunities in the field of deep learning-based brain tumor classification. A main identified gap was developing comprehensive classification models that can simultaneously handle both tumor/non-tumor differentiation and brain tumor type classification. Though studies by Sajjad et al. (2019) and Sultan et al. (2019) have achieved high accuracy in classifying tumor types, they excluded non-tumor cases, limiting the application of these models in practical clinical settings. The gap of lacking normal tissue prediction is particularly significant as the ability to distinguish between normal and abnormal cases is crucial in

clinical practice and the model may serve as a first-pass automated screening tool for healthy cases if this gap is filled.

Since clinical deployment prefers transparent and trustworthy models, improving interpretability of deep learning models is a significant challenge in medical imaging applications. Most current deep learning models function as "black boxes," making it difficult for healthcare providers to understand and trust decisions made by models. Though studies such as the one by Deepak and Ameer (2019) attempted to address interpretability through visualization techniques, the field still need to have more comprehensive methods to explain model interpretability in a way that can provide clinically relevant explanations of model decisions.

Model efficiency regarding the requirement of computational resources is another critical gap in current research. Many existing deep learning architectures are computationally intensive and require significant resources for achieving high accuracy in deployment. The work by Abiwinanda et al. demonstrated that simpler architectures were possible to achieve reasonable accuracy (Abiwinanda et al., 2019). The balance between the high accuracy and the computational efficiency is crucial for practical clinical implementation, such as in resource-constrained settings.

2.6 Summary

The literature review summarized the advancement of deep learning models in brain tumor classification from MRI scans and identified research gaps and opportunities. In more detail, significant progress was made in achieving high classification accuracy for just brain tumor prediction, several crucial gaps remain to be addressed. These

research gaps include the need for a comprehensive classification system that can predict both normal cases and brain tumor cases, and the need for efficient and interpretable models that can be more readily adopted in the clinic.

The research gaps identified from reviewing the literature guide the direction of this praxis research. The methodology presented in Chapter 3 addresses these gaps by developing a deep learning model that has comprehensive classification capabilities for both brain tumor and non-tumor cases and by exploring the model interpretability. More specifically, the approach in this praxis will focus on developing a deep learning model that can efficiently handle both brain tumor/non-tumor classification and brain tumor type differentiation, potentially with clinical relevance, efficiency and practicality. Furthermore, this praxis studies the feature importance extracted from the top performing model, which is the RestNet101-based model, to contribute to the field with interpretability assessment.

Chapter 3- Methodology

3.1 Introduction

This chapter describes the methodology employed to develop and evaluate deep learning models for brain tumor classification based on MRI images. The methodology addresses the challenges identified in Chapters 1 and 2, which is the need for efficient, accurate, and interpretable automated systems for brain tumor diagnosis from MRI images. As discussed in Chapter 2, several deep learning models showed promises in medical image analysis, but significant gaps remain in developing models that can effectively distinguish healthy brain tissue and various tumor types at the same time. Gaps were also identified for model interpretability in clinical settings as deep learning models are often complex and non-obvious for its decision making.

The organization of this chapter is depicted in Graphical Model of Research (Figure 4). The main objective is to create a robust predictive model that is capable of classifying brain MRI images in four classes: no tumor, meningioma, glioma, and pituitary tumor. This chapter details the technical implementation of fine-tuning deep learning architectures, including ResNet, and SWIN transformer models, the fine-tuning techniques to leverage these pre-trained computer vision models and the training configurations. The model development incorporates image augmentation to ensure the model's generalizability. The methodology also addresses model interpretability by visualizing and statistically analyzing the extracted feature importance from the ResNet-

based model. Classification metrics are described to evaluate the model's performance for the classification task as well as the potential for clinical use.

This chapter concludes with limitations of the methodology in data preprocessing, image augmentation and model architecture. Moreover, the model has not yet been evaluated under diverse MRI acquisition settings used at different institutions, so its robustness to such variability is still unknown.

These limitations will be considered in the interpretation of results in Chapter 4.

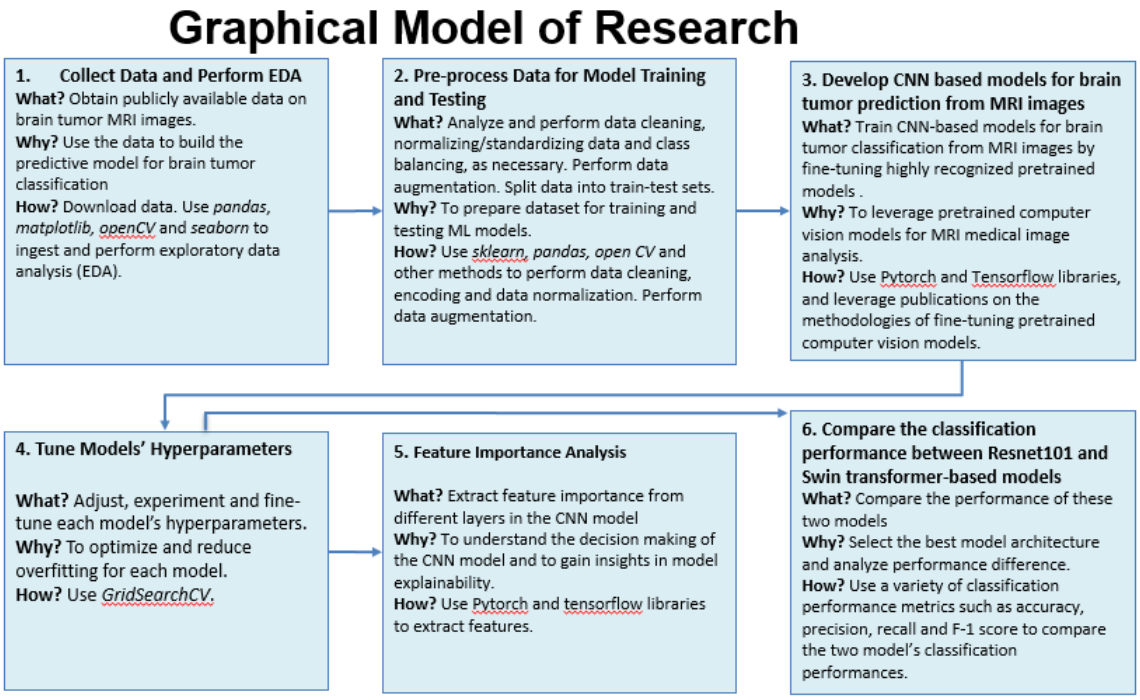


Figure 4. **Graphical Model of Research- level 0 summary**

3.2 Collect Data and Perform EDA

Two public datasets were merged to make the custom dataset for training the four-class model. Exploratory data analysis (EDA) was performed with this combined dataset to prepare for the model training. (Figure 5)

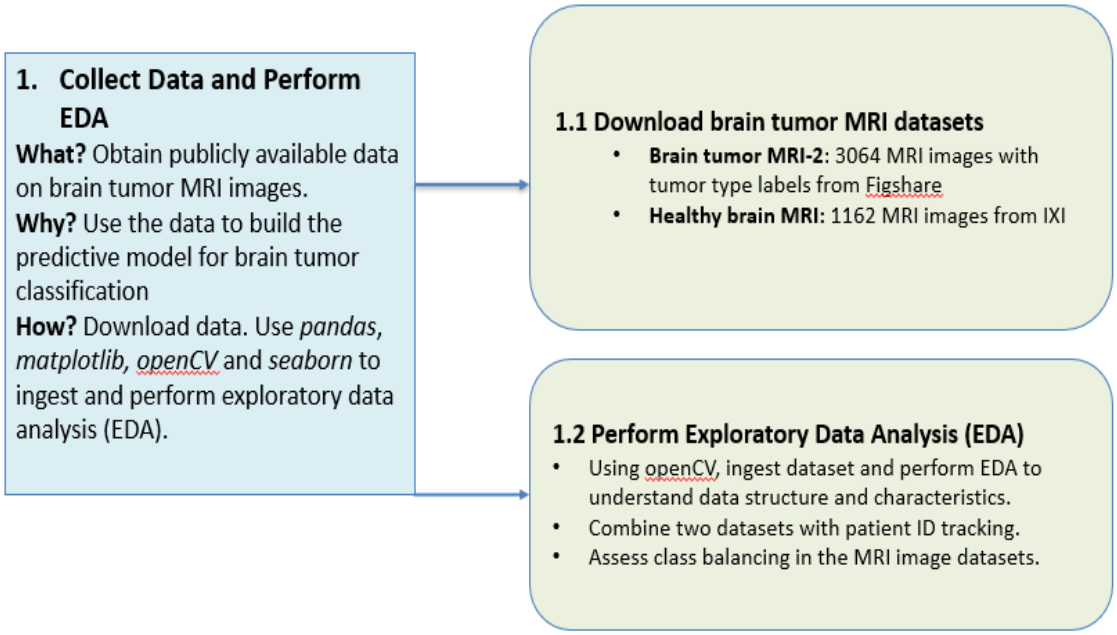


Figure 5. Graphical Model of Research- Level 1 breakdown for MRI dataset.

3.2.1 Dataset Description

As shown in 1.1 in Figure 5, this study used two MRI datasets to address previous limitations which is missing non-tumor cases in brain tumor classification research, as identified in Chapter 2. The primary dataset, obtained from the Figshare repository (https://figshare.com/articles/dataset/brain_tumor_dataset/1512427), comprises 3,064 T1-weighted contrast-enhanced MRI images collected from 233 patients. T1-weighted imaging is a basic pulse sequence of MRI, which is used to show differences in T1

relaxation time of tissues. This dataset includes three brain tumor classes: meningioma (708 slices), glioma (1,426 slices), and pituitary tumor (930 slices) (see data distribution in Figure 6, and image samples in Figure 7). The study incorporates a second dataset from the IXI database, a project assembled one of the largest open collections of in-vivo MRI scans of healthy adults. The selected healthy brain scans for this praxis study consist of 1,162 T1-weighted images of healthy brains. This addition is critical as it enables the development of a more comprehensive four-class classification model, capable of predicting brain tumor types as well as non-tumor cases, addressing a notable gap in the existing research as identified in Chapter 2. The dataset combination uses patient-wise partitioning strategy, where all slices from an individual patient are assigned to the same train, validation or test fold. The combination of these two datasets creates a custom dataset for developing a clinically relevant classification model that can be potentially incorporated in the brain tumor diagnosis workflow.

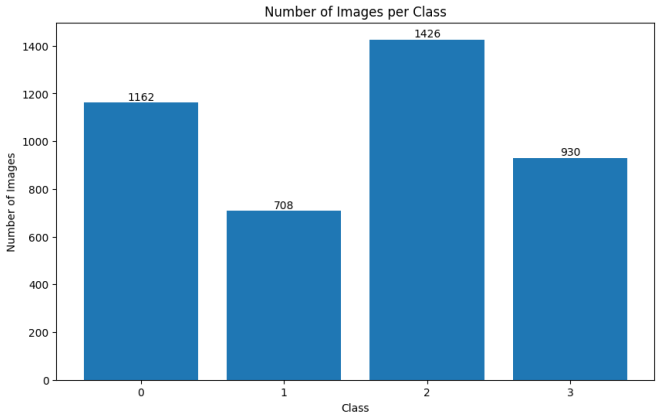


Figure 6. **Class label counts in the combined dataset.**

Class labels: 0- no tumor, 1- meningioma, 2- glioma, and 3- pituitary tumor.

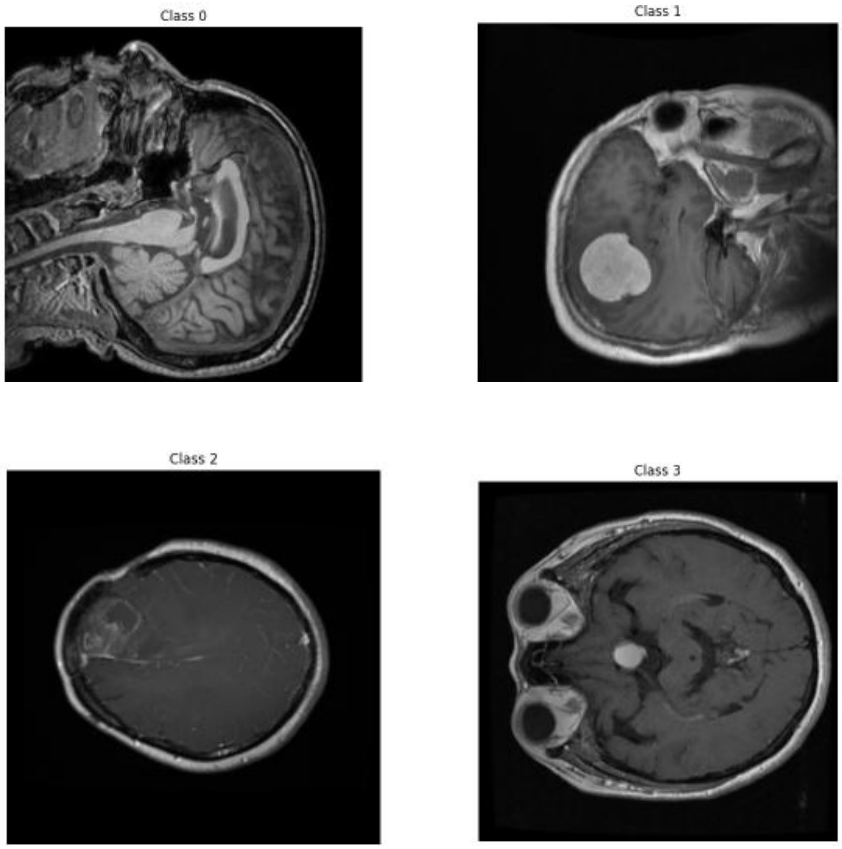


Figure 7. **Samples of MRI from 4 classes.**

Class labels: 0- no tumor, 1- meningioma, 2- glioma, and 3- pituitary tumor.

3.3 Data Preprocessing

After data cleaning and train-validate-test split (2.1 and 2.2 in Figure 8), the data preprocessing pipeline systematically prepares the imaging data for deep learning models. First, all images are resized to 224×224 pixels to ensure dimensional consistency across the dataset. Next, this image size standardization is followed by pixel intensity normalization to the range [0,1], helping stabilize the training process and ensure consistent feature extraction across the dataset. (Figure 6)

Quality control measures include the implementation of class-specific Laplacian variance thresholds to identify and remove images that are below standards. This process eliminated approximately 5-10% of images that are either too blurry or too sharp, ensuring that the training dataset maintained consistent quality.

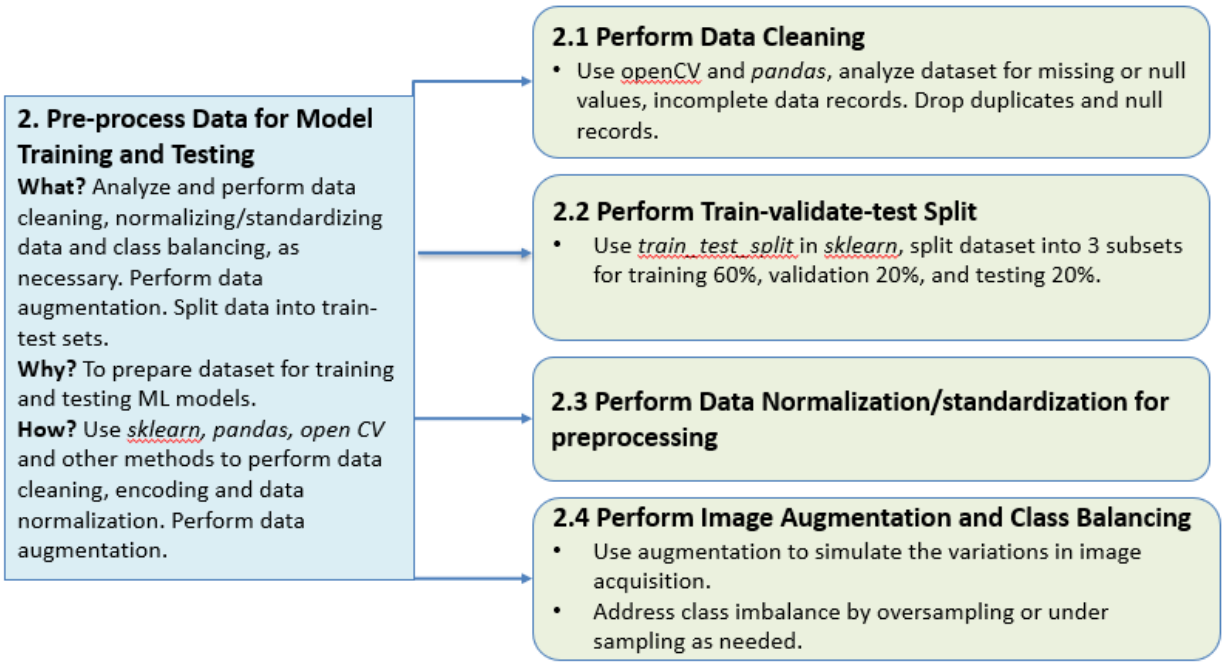


Figure 8. **Graphical Model of Research- Level 1 breakdown for data preprocessing.**

3.3.1 Data Augmentation

The data augmentation protocol has been designed to optimize the model robustness (2.4 in Figure 6). Based on recent work on medical image improvement approach (Islam et al., 2024; Goceri, 2023), the study implements transformations that simulate realistic variations in MRI acquisition conditions for image data augmentation. The transformation for data augmentation include: random rotations within ±15 degrees

to simulate patient's positioning variations; width and height shifts of 5% to simulate patient's movement; shearing and deforming the image by 5% to simulate a slightly different brain anatomical structure; zoom variations of 10%-15% to account for acquisition distance differences between body and instrument; and subtle changes in brightness and contrast to simulate variations in image acquisition environments.

3.4 Model Development

The model development started with the selection of base architecture, between CNN-based model and vision transformer-based model. The model development then focused on implementation of ResNet101-based model with a fine-tuning strategy. The vision-transformer architecture was also explored for comparison. (Figure 9)

For the ResNet101-based model, custom modifications were made to fine-tune the model for the classification of brain MRI images under four categories- three tumor classes and the non-tumorous condition. Hyperparameters were tuned to develop the model, using metrics that evaluate the classification performance. The performance metrics include accuracy, precision, recall and F1 score. The model was tested on the 20% held-out test data and evaluated with the confusion matrix and classification performance metrics.

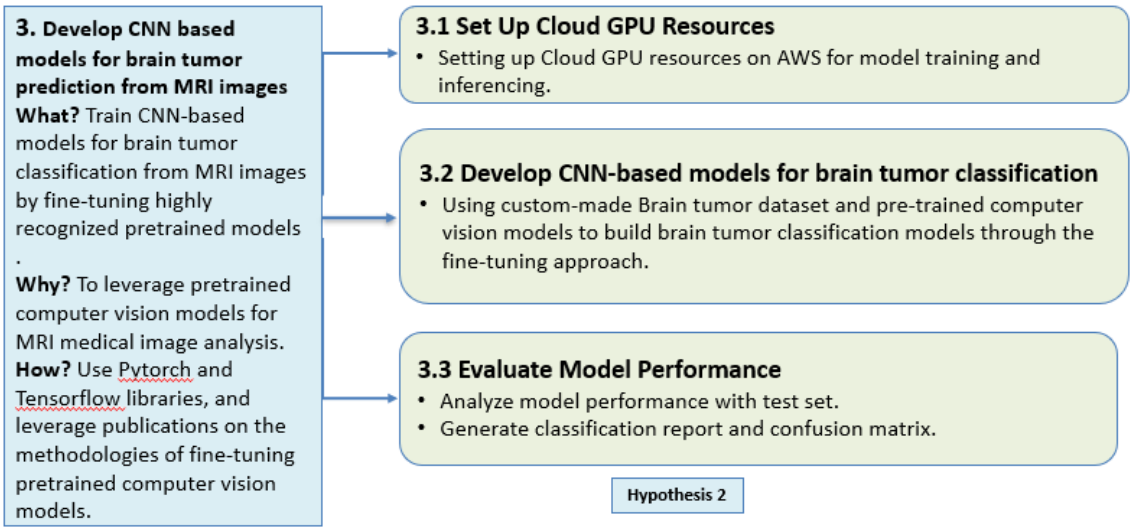


Figure 9. **Graphical Model of Research- Level 1 breakdown for model development.**

The graph focused on the CNN-based model, though SWIN transformer model was also explored in this praxis.

3.4.1 Base Architecture Selection

This praxis explored deep learning architectures for brain tumor classification, including both traditional convolutional neural networks (CNNs) and the more recent vision transformer-based models. Literature showed that ResNet101 and SWIN transformers were both promising algorithms to achieve high performance for brain tumor classification tasks. Each algorithm has its own advantages: the deep residual learning structure of ResNet101 performs reliably in health image inspection, and the SWIN transformer captures the global image context by using a hierarchical attention mechanism.

This section describes the implementation and fine-tunning of these architectures including any modifications made. First, the ResNet101-based model with appropriate fine-tuning demonstrated high performance in brain tumor classification. Next, experiments with SWIN transformer variants offered valuable insights into the challenges of applying transformer architectures to medical image classification tasks. The methodology described in this section would support the Hypothesis 3 for selecting the best-performing model.

3.4.1.1 Implementation of ResNet101-based Model

This praxis research implements a deep learning model based on the ResNet101 architecture, leveraging fine-tuning the pre-trained model. ResNet101 was selected as the backbone architecture due to its demonstrated effectiveness in medical image analysis tasks and its ability to handle the complexity of features in brain MRI images through its deep residual learning structure (Ghosal et al., 2019; Çınar et al., 2022). The deep residual learning configuration is particularly useful for tackling challenging medical imaging tasks, as it allows very deep networks to be trained, as well as preserving gradient flow, as demonstrated in various successful medical imaging applications (Zhou et al., 2021). The ResNet base model accepts input images of size 224x224 pixels with three color channels, which meets the standard medical imaging preprocessing requirements.

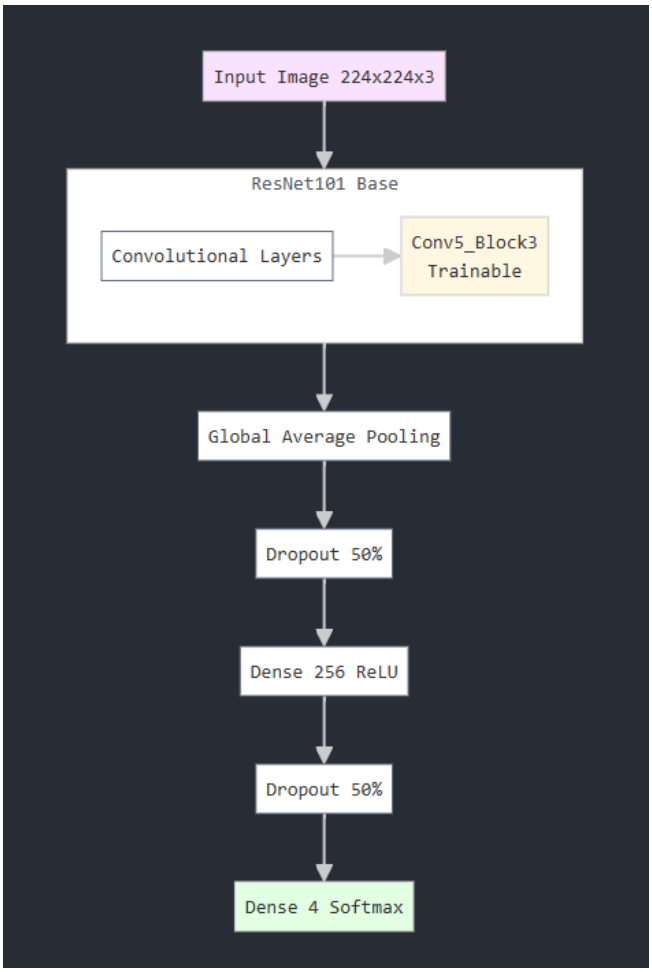


Figure 10. **Custom modifications on ResNet101 architecture in this praxis.**

The architecture of the ResNet101 has an initial convolutional layer with 64 filters, followed by five main convolutional stages and residual connections and each convolution is followed by batch normalization and then a ReLU activation function. The convolution stage 5- block 3, was made trainable for brain tumor classification.

The architecture (Figure 10) used in this praxis incorporated several custom modifications to adapt it specifically for brain tumor classification. On top of the ResNet101 base, the model implemented a Global Average Pooling layer to reduce spatial dimensions and still preserved important feature information. Next was a dropout layer

with a dropout rate of 50% to prevent overfitting. The architecture then included a fully connected layer with 256 units using ReLU as activation function, which provided an intermediate feature space before the final classification layer. The output layer consisted of 4 units with softmax activation function, corresponding to the four classes: no tumor, meningioma, glioma, and pituitary tumor.

3.4.1.2 Implementation of Transformer-based SWIN Model

This research also experimented with a SWIN transformer-based model for classifying brain tumors. Two fine-tuning approaches were tested. The first approach involved fine-tuning a pre-trained SWIN transformer model, focusing on adjusting the head, last layer, and second-to-last layer to classify four categories: non-tumor cases and three tumor types (meningioma, glioma, and pituitary tumor). Regardless of extensive hyperparameter tuning and optimization of both loss functions and learning strategies, this approach did not achieve good performance for all four classes. While the model achieved an accuracy of 97% for just the non-tumor class, it did not work when it came to distinguishing three different types of tumors.

In response to these limitations, another step-wise three-phase approach was experimented:

- Initial classification between tumor and non-tumor cases
- Secondary classification to differentiate among the three brain tumor types
- Implementation of contrastive learning to further increase the differentiation among brain tumor types

This praxis research attempted the three-phase approach to improve the model's performance, particularly aiming to differentiate between the three brain tumor types. Nevertheless, the poor performance with the SWIN transformer approach led the research to focus on a ResNet-based fine tuning model which achieved better results to classify all four classes.

This finding is consistent with previous research about the challenges of applying transformer architectures to medical imaging tasks. Though the vision transformer models are in theory advantageous in capturing global context, however the models sometimes struggle with fine-grained feature discrimination necessary for medical diagnosis that requires attention to details (Chen et al., 2021). The results suggest that hybrid approaches combining CNN and vision transformer, or traditional CNN architectures may work for this medical image classification task in future investigation. This praxis focused on fine-tuning a CNN architecture for brain tumor classification.

3.4.2 Fine Tuning Strategy with ResNet101

After selecting ResNet101 as the base architecture, the implementation of the ResNet101-based model for brain tumor classification follows a fine-tuning approach designed to maximize the benefit of pre-trained weights by adapting the RseNet model to the specific requirements of the brain tumor classification task. The Imagenet pre-trained weights of our base model, ResNet101, was downloaded using TensorFlow library. Furthermore, custom layers were added to the base model to meet the multiclass classification needs. These layers are the Global Average Pooling Layer, Dropout Layers

to reduce overfitting, Dense Layers as a fully connected layer, and a final classification layer that used the SoftMax activation function (Figure 10).

In the initial phase of training, all layers of the ResNet101 base model remained frozen, preserving the valuable feature extraction capabilities learned from ImageNet. In this way, the model maintained its ability to recognize fundamental image features, and the newly added custom classification layers are responsible for learning the interpretation of these features in the context of brain tumor classification.

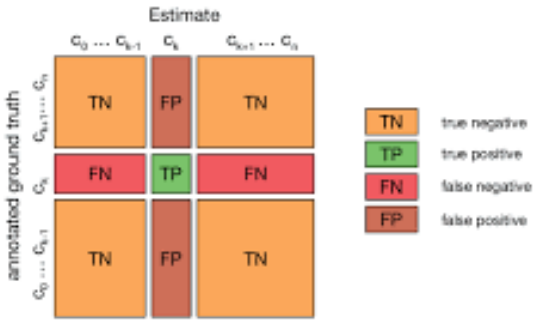
During the fine-tuning phase, the model selectively unfreezes the conv5_block3 layers while keeping weights in earlier layers frozen. This approach allows the model to retain the powerful low-level feature recognition capabilities obtained via ImageNet training, and specifically refining its high-level feature extraction capabilities for brain tumor classification through fine-tuning. The model contains 42,561,028 trainable parameters and 105,344 non-trainable parameters, maintaining a balance between adaptability and stability.

3.4.3 Performance Metrics

After the best-performing model was selected and saved in the training phase, the evaluation of the model was carried out on the 20% held-out test data. The model testing and evaluation employed a set of performance metrics to assess the effectiveness of brain tumor classification.

3.4.3.1 Metrics for Evaluating Classification Performance

Accuracy (Equation 1) was the fundamental metric for evaluating classification performance. It represents the overall proportion of correct predictions across all four classes (no tumor, meningioma, glioma, and pituitary tumor). The targeting accuracy was 95% or greater as specified in Hypothesis 2 of this praxis.



- Accuracy = $(TP_0+TP_1+TP_2+TP_3)/total$ (1)
- Precision $_i = TP_i/(TP_i+FP_i)$; i=class 0, 1, 2, 3 (2)
- Recall $_i = TP_i/(TP_i+FN_i)$; i=class 0, 1, 2, 3 (3)
- F1 Score $_i = 2*$Precision $_i*$ Recall $_i/($Precision $_i +$Recall $_i)$;
  i=class 0, 1, 2, 3 (4)

Furthermore, because the application of the model is in the medical field and the potential consequence of misdiagnosis can be life-threatening, accuracy alone was not sufficient. Therefore, precision in Equation 2 and recall in Equation 3 were calculated for each class individually. Precision calculates the proportion of correct predictions in each class. It is a crucial metric in determining whether the model can avoid making incorrect

predictions or false positives. For the no-tumor class, the goal is to achieve 100% precision for this class as the model should not predict any tumor case as no-tumor. Recall, which is also known as sensitivity, quantifies the model's ability to correctly identify actual positive cases corresponding to the class. The F1 score in Equation 4 is used to assess the performance of the model from a balanced perspective. It is the harmonic mean of precision and recall. These metrics were calculated using tensorflow's classification report functionality, which generated model performance statistics for each class and calculated weighted averages (Equations 5, 6 and 7) across all classes.

Weighted Precision:

$$\text{Precision}_{\text{weighted}} = \frac{\sum_{i=1}^{n}(\text{Precision}_i \times \text{Support}_i)}{\sum_{i=1}^{n}\text{Support}_i} \quad (5)$$

Weighted Recall:

$$\text{Recall}_{\text{weighted}} = \frac{\sum_{i=1}^{n}(\text{Recall}_i \times \text{Support}_i)}{\sum_{i=1}^{n}\text{Support}_i} \quad (6)$$

Weighted F1 Score:

$$\text{F1}_{\text{weighted}} = \frac{\sum_{i=1}^{n}(\text{F1}_i \times \text{Support}_i)}{\sum_{i=1}^{n}\text{Support}_i} \quad (7)$$

Macro Precision:

$$\text{Precision}_{\text{macro}} = \frac{1}{n}\sum_{i=1}^{n}\text{Precision}_i \quad (8)$$

where n is the number of classes.

Macro Recall:

$$\text{Recall}_{\text{macro}} = \frac{1}{n}\sum_{i=1}^{n}\text{Recall}_i \quad (9)$$

Macro F1 Score:

$$\text{F1}_{\text{macro}} = \frac{1}{n}\sum_{i=1}^{n}\text{F1}_i \quad (10)$$

In the multi-class classification report, the macro and weighted averages provide important insights into the model's overall performance across classes. The macro averages (Equations 8, 9 and 10) calculated the unweighted mean of each metric, which is precision, recall, and F1-score, across all four classes (no tumor, meningioma, glioma, and pituitary tumor). The macro averages treat each class equally regardless of its size. This method can help assess the performance of the model on predicting the minority class as the macro averages give equal weight (unweighted) to the performance of each class. On the other hand, the weighted averages computed the mean of each metric weighted by the number of samples in each class, which better reflected the model's performance on the actual data distribution.

For visualization, confusion matrix was generated to show how many times the model correctly and incorrectly predicts data in the four classes. The matrix was also color-coded for easier interpretation. The confusion matrix broke down the number of predicted class instances against the number of ground truth instances of a specific class. The confusion matrix helped identify specific patterns in misclassifications, which was valuable for understanding which classes of data were most often misclassified and thereby showed where the model needed improvement.

3.5 Training Configuration of the ResNet101-based Model

The model training implemented multiple regularization techniques to increase its generalization ability. Two dropout layers, each with a 50% dropout rate, were strategically placed in the architecture to prevent over-fitting. The model retained the batch normalization layers inherited from the ResNet101 base architecture, which helped

stabilize the model training and speed up convergence. Data augmentation as described in 3.2.3 was conducted to expand the size of the effective training dataset, with the goal to help train the model to be more reliable for predicting variable image inputs.

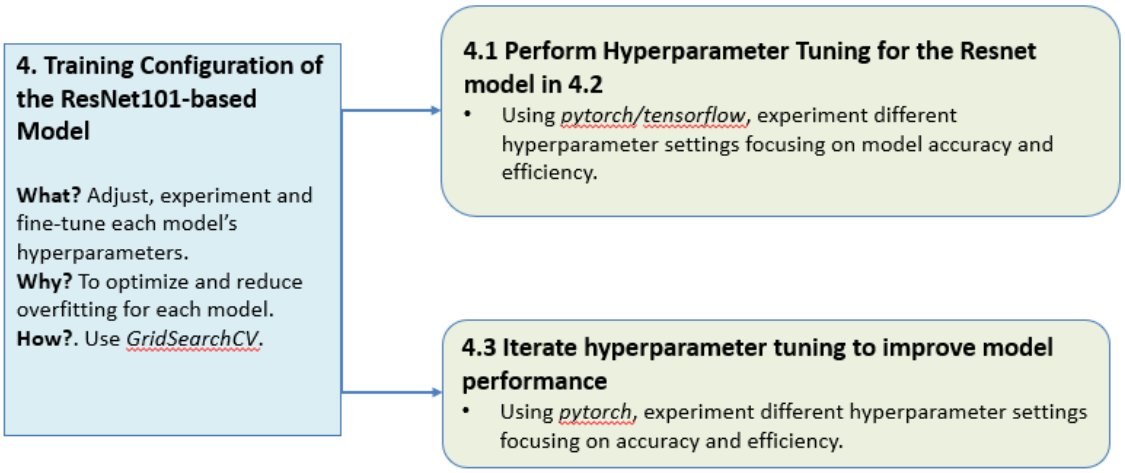


Figure 11. **Graphical Model of Research- Level 1 breakdown for hyperparameter tuning.**

Figure 11 outlines the optimization for training configuration of the ResNet101-based model to improve model robustness and to prevent overfitting. Hyperparameters were optimized with the grid search method. As a result of hyperparameter tuning, the ResNet101-based model used the Adam optimizer and adjusted the learning rate to 5e-6. This parameter is selected to help the RestNet101-based model to converge stably, as well as adapting to the specific features in brain MRI images. A training batch size of 32 was selected to balance between computational efficiency and the training stability. During training, the categorical cross entropy loss function guided the optimization process and provided a suitable gradient for the multi-class classification task. Therefore, the model was compiled with categorical cross-entropy as the loss function, Adam as the

optimizer and accuracy as the evaluation criterion. Next, the model was evaluated from calculating metrics such as precision, recall and F1-Score for each class on the test datasaet. The optimization for training configuration improved the model's generalization ability and overall accuracy across all four classes.

The training progress was monitored by the training/ validation loss and the training/ validation accuracy over training epochs. The weights of the best-performing model based on the validation accuracy were saved. The training process employed validation-based model selection, which used classification metrics calculated from a 20% held-out validation for guiding model selection and preventing overfitting to the training data.

3.6 Analysis and Visualization of Feature Importance extracting from the ResNet101-based Model

To better understand the model interpretability, feature importance was extracted from different layers of the ResNet101-based model using saliency maps to visualize the pixels that the model relied on to make classification decisions. In particular, the final layer of feature importance was used for the analysis. Next, the tumor region on the MRI scans was segmented with an orthogonal Unet-based model. Finally, these two masks (saliency map and Unet masks) were analyzed with the intersection over union ration (IoU). A large IoU indicate the model relied on tumor cores for classification, while a small IoU suggests the model used tumor margin features for the classification decisions. Figure 12 shows this analysis process for understanding model interpretability.

This section starts by describing the clinical relevance of tumor core and margin features on MRI scans, followed by describing the three steps as illustrated in Figure 12.

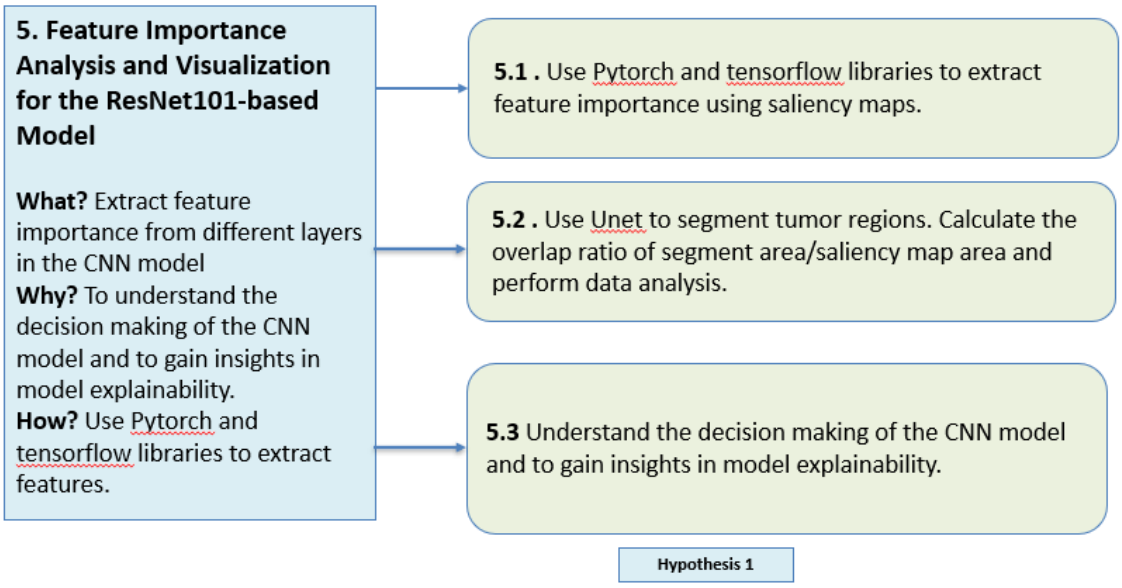


Figure 12. **Graphical Model of Research- Level 1 breakdown for model interpretability using feature importance analysis.**

For understanding the relationship between the extracted feature importance and anatomical tumor features.

3.6.1 Clinical Relevance of Tumor Core and Margin Features

Radiologists use specific imaging features of tumor cores and tumor margins for making brain tumor diagnosis, including diagnosis of tumor presence and classification of tumor types, which impact on selecting the best treatment option. The tumor core in the MRI scan provides essential diagnostic information through patterns of signal intensity and enhancement. For instance, meningiomas are reported to show a consistent and uniform signal intensity on enhanced MRI. But gliomas usually have variable

internal textures and irregular improvement patterns on enhanced MRI (Watts et al., 2014).

The characteristics in the tumor margins are critical for diagnosis, as the boundary between tumor and surrounding tissue offers important diagnostic information. Meningiomas typically have well-defined margins, which has limited swelling and a distinct boundary. Gliomas often demonstrate infiltrative growth pattern, which has poorly defined borders and/or extensive, blurred swelling patterns that extend beyond the apparent tumor margins. (Joo et al., 2021; Padmaja et al., 2024)

These established clinical criteria provide critical basis for assessing whether deep learning models can effectively learn and utilize similar diagnostic features that are used by radiologists. The alignment between feature importance extracted from the Resnet101-based model and segmented tumor regions reveal whether the model relied on tumor core or tumor margin for classification. By analyzing such alignment for model's decision making, it provides information for the model's interpretability for clinical applications.

3.6.2 Feature Importance Extraction Using Saliency Map

This praxis research employed saliency maps to extract and visualize the model's feature importance and understand the model's decision-making processes. The saliency maps were applied in different layers of the ResNet model to analyze and visualize how features were processed throughout the network.

Saliency maps were employed to provide a visualization of pixel-level insight into the model's feature importance patterns when making output predictions. By computing

the gradients of the model's prediction with respect to the input image, a saliency visualizes the importance of each input pixel for the model's prediction. Saliency maps offer precise spatial information about feature importance.

To implement saliency maps, this praxis research utilized PyTorch and TensorFlow libraries to extract feature importance from different layers of the ResNet model. The visualization process employed a color-coding scheme ranging from red (indicating high importance) through green to blue (indicating decreasing relevance) to indicate the relative importance of different image regions in the model's processing and decision-making. As discussed in detail in 3.5.4, this saliency map visualization technique is valuable for validating whether the model's feature importance patterns aligned with clinically relevant tumor characteristics. The visualization of feature importance can provide a foundation for explaining the model's classification processes to medical professionals and may help increase users' confidence in the model's capability to focus on clinically relevant aspects of brain MRI scans.

Using saliency maps, feature importance were extracted from three stages of the ResNet model:

- Final layers: The saliency map visualization focused on examining how the model processes detailed, class-specific features. The saliency maps highlighted the regions that most directly influenced the final classification decision. The saliency maps from the final layers were used for analysis in section 3.5.4.

- Middle layers: The saliency maps examined how intermediate layers process and refine feature representations. These layers were analyzed to understand the transition from general to more specific feature detection.
- Early layers: The visualization with saliency map showed how the model processed basic input image features and began to construct its understanding of the input image. In early layers, feature importance examined the initial feature importance patterns across the image.

This research approach using saliency maps to visualize feature importance is designed to explore whether the feature importance extracted from ResNet model can correspond to the clinical features that radiologists usually pay attention to when making diagnoses. The saliency maps extracted from the final layers of the ResNet model would be used to compare with the tumor regions segmented using a U-Net model (described in 3.5.3) and provide insights about the model's decision-making process and improve the model's interpretability.

3.6.3 Tumor Segmentation with U-Net Model

As discussed in Chapter 2, U-Net is a well-known convolutional neural network architecture that is widely used for medical image segmentation, such as for tumor segmentation in MRI scans. The U-Net model localizes and segments regions such as tumors from the whole image, which can help clinicians and researchers find and explain the specific conditions of the lesions. For this section, a pretrained U-Net model was employed to inference on the test dataset of MRI images for calculating the brain tumor areas.

In brief summary, the segmentation model used a U-Net architecture with an encoder-decoder structure. The encoder extracted spatial properties via convolutional layers, batch normalization, and max pooling, and the decoder reconstructed the segmentation mask using transposed convolutions combined with the encoder's features. The final tumor region mask was generated by a 1×1 convolutional layer with a sigmoid activation function. Each image and mask was resized to 224×224 pixels, normalized to a 0-1 range. They were then binarized by setting pixel values higher than 0.5 to 1 (tumor) and the others to 0 (background). Using the U-Net model to inference on the test dataset of the model developed in section 3.3.3, the tumor areas were thus identified and calculated for further analysis in 3.5.4.

3.6.4 Analysis of Overlap Ratio (Intersection over Union or IoU)

The overlap ratio is a metric calculated between the segmentation mask produced by the U-Net model (section 3.5.3) and a saliency mask generated by the classification model (section 3.5.2). This ratio quantifies the similarity between these two masks by measuring the ratio of their intersection over their union (IoU), which essentially assessed how much the areas from the two masks aligned.

Next, the overlap ratios (IoU of the segmentation mask and the saliency mask) were plotted in histograms for each tumor type or non-tumorous condition to assess the distribution of the ratio. In the histograms, x axis represented a range of the ratios, or commonly known as a histogram bin, and y axis represented the frequency. If the model relied on tumor cores for classification, frequent high overlap ratios would be shown on

the histogram. On the other hand, if the model relied on tumor margins for classification, the histograms were anticipated to show frequent small overlap ratios.

3.7 Limitations of Methodology

The research methodology exhibits several limitations in its approach to brain tumor classification, described as follows.

3.7.1 Limitations in Data Preprocessing and Augmentation

The data preprocessing methodology had limitations in image standardization and image cleaning. The image preprocessing pipeline in this praxis included resizing all images to 224x224 pixels and applying class-specific Laplacian variance thresholds, This method may cause the loss of certain critical diagnostic information. Though image standardization was necessary for consistent model input, uniform resizing may result in loss of resolution for larger tumors or detail distortion for smaller ones. For the Laplacian variance-based image cleaning, removing 5-10% of images based on this blur metric was intended to improve data quality, but it also has the risk of screening out a small number of clinically relevant borderline cases that could be important for model training.

The data augmentation methodology also has constraints in its current form. The augmentation techniques in this praxis attempted to simulate realistic variations in MRI acquisition, but the relatively conservative parameters (rotation ±10-15 degrees, scaling ±10-15%, small translations) may not fully capture the range of variations of MRI scans encountered in clinical settings. The data augmentation methods can not represent all the complex artifacts encountered in real-world MRI scans.

3.7.2 Limitations in Model Development

In the current ResNet101-based model, only the conv5_block3 layers were trainable and earlier layers remained frozen with ImageNet weights. Chapter 4 describes that this method was effective for brain tumor classification, but it might still has the limitation for the model to learn specific features of MRIs that showed only subtle differences, which were needed to identify tumors. Freezing the early layers of the network met the goal for computational efficiency, but it might have the risk of preventing the model from fully adapting to the unique features of medical images.

Chapter 4 shows the model's learning configuration, such as a learning rate of 5e-6 and a batch size of 32, was effective for achieving good brain tumor classification results. But it is unknown if other configurations may provide the most efficient path to convergence or the best final performance.

If this praxis's model performs well on current datasets, its ability to adapt to different MRI scan parameters and image quality fluctuation was not verified. This model was not tested on varied MRI scans from different institutions or scanning protocols.

Chapter 4- Results

4.1 Introduction

This chapter describes the results of the deep learning-based brain tumor classification models developed in this praxis research, which employed the methodology described in Chapter 3. This chapter presents the research findings and test experimentation, aiming to find the significance of the generated data in relation to the research objectives described in Chapter 1. The results addressed three research hypotheses presented in Chapter 1: (1) the alignment of the extracted feature importance to the tumor regions for model interpretability, (2) the achievement of 95% or greater classification accuracy, and (3) the comparative performance of two different deep learning architectures. The results encompassed ResNet101-based model's classification performance metrics, visualization and statistical analysis on extracted feature importance, and comparative performance evaluations of two different architectures.

The analysis of the results focused on three key areas: (1) the classification performance of the ResNet101-based transfer learning model, which emerged as the most effective architecture during experimentation. The classification results were assessed and compared with the objective of 95% accuracy. (2) feature importance visualized by saliency maps, which were extracted from the ResNet101-based model. And the alignment of saliency map to the clinically relevant tumor regions, which were

segmented from an orthogonal model using a U-Net-based architecture. (3) the challenges of transformer-based architectures.

4.2 Performance of the ResNet101-based Model

4.2.2 Model Performance by Confusion Matrix Analysis

ResNet101 was fine-tuned on the combined MRI dataset (comprising normal brain scans plus three tumor subtypes: meningioma, glioma, and pituitary tumor), using the methodology as described in Chapter 3. After data augmentation, the ResNet101-based model was fine-tuned by unfreezing the Conv5_block3 block and adding a 4-class classification head. After model training, the model was evaluated on the held-out test set. The correct predictions in the diagonal cells of the confusion matrix in Figure 13 demonstrated the model's high performance across the three tumor classes and the no-tumor class.

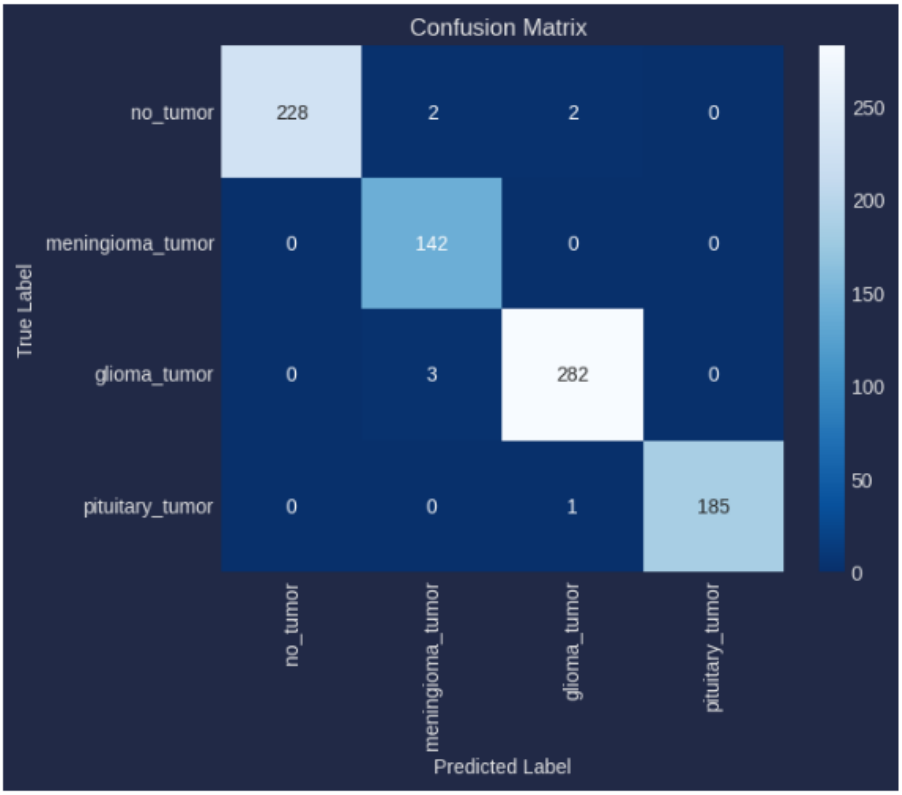


Figure 13. **Confusion matrix from the held-out test data.**

The confusion matrix shows excellent classification performance and minimal misclassifications:

- False positive rates are below 1% for all classes
- Most misclassifications occurred between glioma and meningioma cases, which may originate from similar MRI textural or marginal features of these two types of brain tumors.
- Importantly, the model showed strength in identifying normal (no tumor) cases, underlining the suitability of this model as a first-pass screening tool.

4.2.1 Model Performance by Classification Metrics

Based on the confusion matrix, classification metrics, which were described in detail in Chapter 3, were calculated and used to evaluate the performance of the model. In brief:

- **Accuracy**: The overall proportion of correctly classified cases.
- **Precision**: The ability of the model to avoid false positives.
- **Recall (Sensitivity)**: The ability of the model to detect positive cases correctly.
- **F1-score**: The harmonic mean of precision and recall.

The ResNet101-based transfer learning model achieved excellent performance in classifying brain tumors across all four categories (no-tumor, glioma, meningioma, and pituitary tumor). The model demonstrated the following key performance metrics:

- Overall Accuracy: 99.1%
- Macro-average Precision: 98.9%
- Macro-average Recall: 99.2%
- Macro-average F1-Score: 99.0%

These results exceeded the target accuracy of 95% specified in Hypothesis 2, demonstrating the model's robust capability in distinguishing among different tumor types, as well as normal brain tissue.

The final performance metrics achieved for each class were listed in Table 1.

Table 1: Detailed Performance Metrics by Class

Class	Precision	Recall	F1-Score
No Tumor	1.000	0.983	0.991
Meningioma	0.966	1.000	0.983
Glioma	0.989	0.989	0.989
Pituitary	1.000	0.995	0.997

No-Tumor classification yielded near-perfect performance (Precision = 1.000, Recall = 0.983), which suggests the model is reliable at screening out normal brain scans. Meningioma, glioma, and pituitary tumor classification also achieved excellent metrics, with F1-scores ranging from 0.983 to 0.997.

The brain tumor classification model achieved high performance across all metrics and classes because the brain tumor classification model was successfully developed through Resnet101-based transfer learning with effective training strategies and data augmentation techniques. The fine-tuned ResNet101 model learned both general and subtle features of each class to perform the classification task. The strong classification performance indicates that the data augmentation strategy including random

rotations, translations, scaling, brightness/contrast changes was successful in improving the model's robustness to real-world MRI variability.

4.3 Feature Importance Analysis for Model Interpretation of the ResNet101-based Model

4.3.1 Saliency Map Visualization of Feature Importance

One aim of this praxis was to address **Hypothesis 1**, which states that Features from the tumor margin contribute more significantly to brain tumor classification than the tumor core features. To understand which part of the tumor on MRI scans contributed to the classification decision, feature importance analysis using saliency map was conducted on test dataset using methodology as describe in 3.5.2. The saliency map complemented the purely numerical metrics (accuracy, precision, recall, etc.), showing visually which parts of the MRI scan the model deemed most relevant. Below are key aspects of saliency map generation and interpretation:

- Per-pixel gradients of the classification scores were computed with respect to the input MRI image. The scores were shown in a color-coded saliency map.
- Red, orange, or "hotter" tones in the heatmap indicate areas with higher influence on the final classification, while cooler colors (greens or blues) reflect lower importance.

The results of feature importance extracted from different layers of the ResNet101 model provided insights into how the model processed and interpreted MRI images. The visualization techniques revealed the following patterns in early, middle and final layers.

After comparing the saliency map results extracted from early, middle and final layers of the fine-tuned ResNet101 model, the analysis for model interpretation in section 4.3.2 used the saliency map results from the final layers (Figure 14).

Early Layers:

- Focused on basic image features such as edges and textures
- Showed broad and diffused attention across the entire image
- Demonstrated sensitivity to general tissue boundaries

Middle Layers:

- Developed more specialized feature detectors
- Showed increased focus on tumor-relevant regions
- Began differentiating between tumor and normal tissue characteristics

Final Layers (used for further analysis in 4.3.2):

- Exhibited highly specific attention to tumor-related regions
- Demonstrated strong alignment with clinically relevant features
- Showed clear differentiation between tumor types based on their characteristic appearances

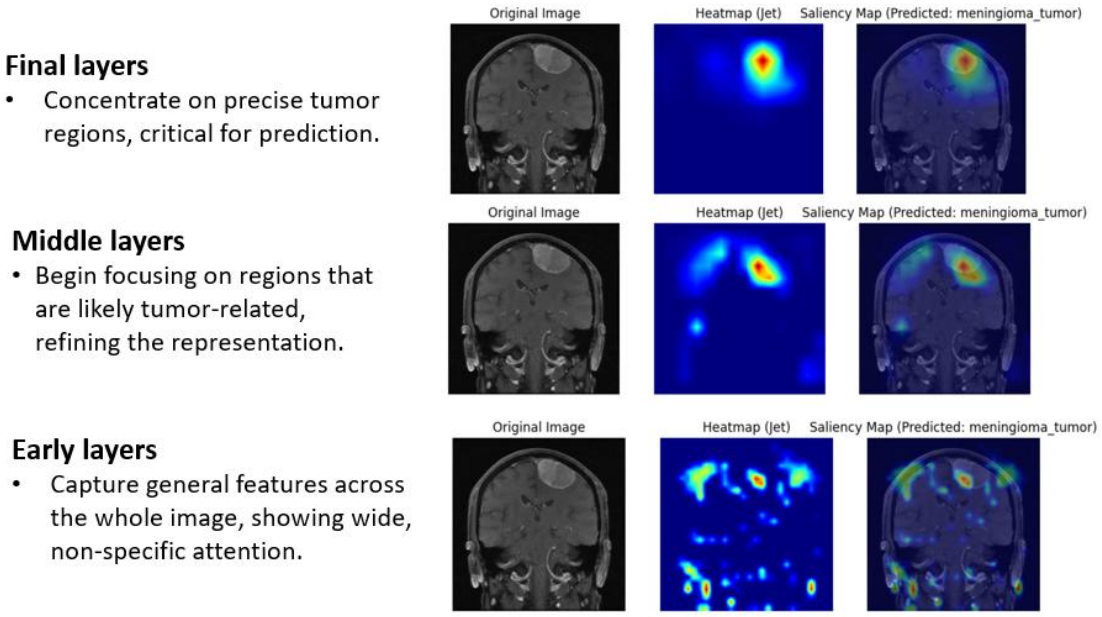


Figure 14. **An example of the extracted saliency maps from final, middle and early layers of the ResNet101-based model for a meningioma MRI test image.**

4.3.2 Model Interpretation: Analysis of Overlap Between Final-layer Saliency Map and U-Net Segmented Tumor Regions (ResNet101 Model)

Beyond generating saliency maps, a quantitative overlap analysis between saliency map and segmented tumor mask was conducted to confirm whether the ResNet101-based model relied more on tumor margin or tumor core for classification decisions. As described in Chapter 3, a pre-trained U-Net–based segmentation model was employed to generate binary masks for tumor regions. Next, the overlap ratio that quantifies the similarity between segmented area and saliency map area was computed by measuring the percentage of their intersection over their union (IoU, Figure 15). For statistical analysis, the overlap ratios (IoUs) were plotted in histograms for each tumor or non-tumor class in Figure 16.

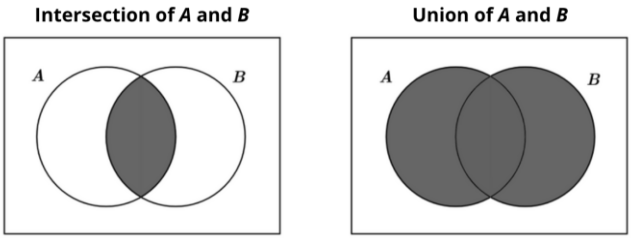


Figure 15. **IoU calculation for two areas, U-Net segmented tumor area and saliency map area for important features in an MRI scan.**

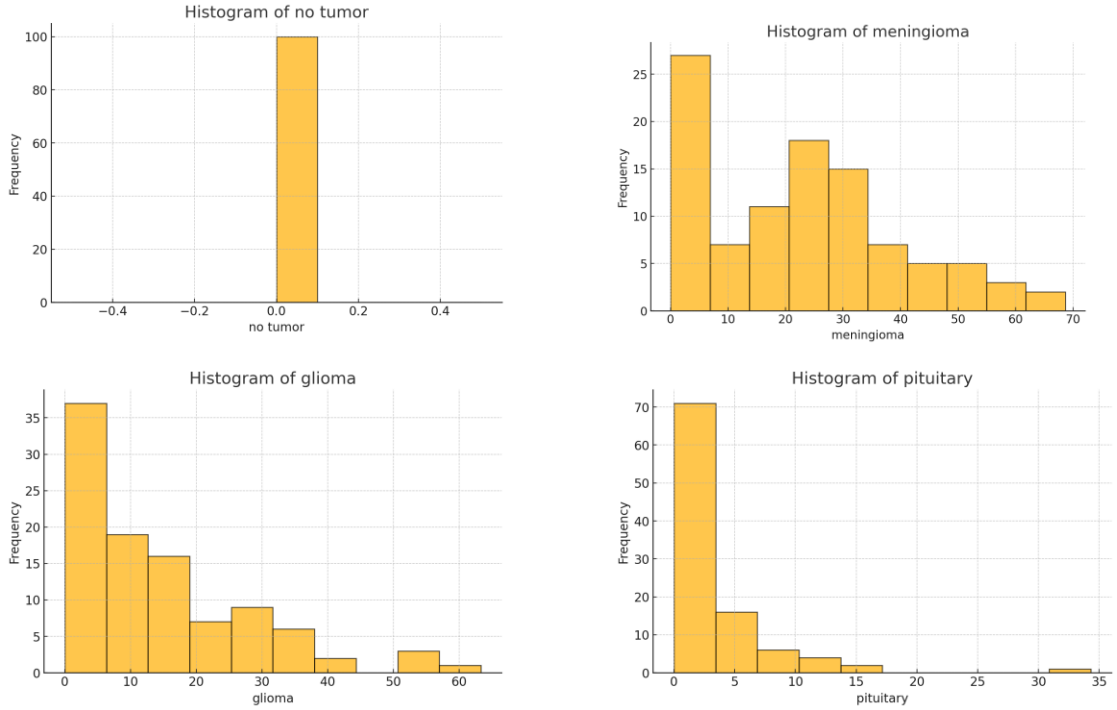


Figure 16. **Histograms that show the distribution of IoU (U-Net segmentation mask and saliency map) in four classes.**

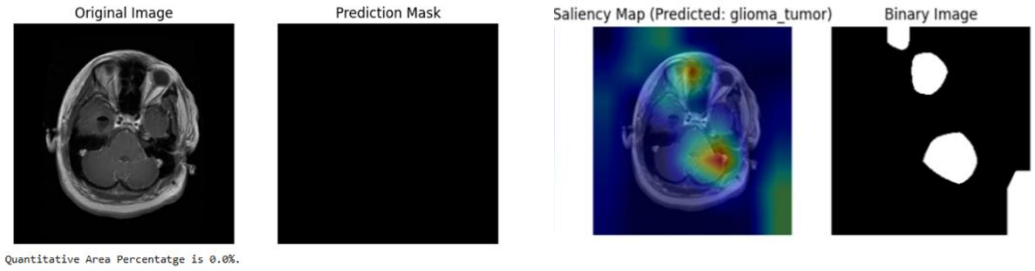
The histograms of the IoU highlighted tumor margin are important for the model's brain tumor classification. Figure 17 illustrates how the IoU values correspond to this analysis. On the histograms, several images resulted in 0% ratios, which was due to the U-Net did not find a tumor, yet the ResNet101-based model still worked (Figure 17A).

Low IoU values, which were common on the histograms, showed that the saliency map overlapped at the tumor margin of the segmentation mask (Figure 17B). This indicates that the ResNet101-based model relied on tumor margin to make classification decisions.

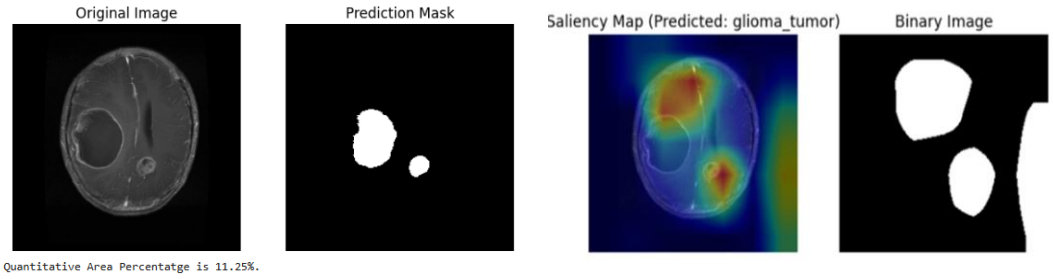
Some medium to high IoU values were present on the histograms. A closer look comparing the saliency map and tumor segmentation mask revealed that high IoUs were due to the ResNet101-based model capturing two sides of tumor margins with dual-hot zones; thus, the binary saliency map mask covered more tumor regions than a single hot zone (Figure 17C).

In conclusion, the IoU histograms and the comparative analysis of saliency maps and segmentation masks collectively underscored the model's reliance on tumor margins for accurate classification, as low IoU values were most frequent on histograms for all tumor classes. Dual-hot zone saliency patterns contributed to higher IoUs, highlighting how the ResNet101-based model robustly captures these pivotal boundary features to enhance its performance.

A. 0% ratio results from no tumor area found from segmentation model, yet ResNet101-based classification still worked.



B. Low ratio result showed the Resnet101-based model captured tumor margin effectively.



C. High ratio result showed the ResNet101-based model captured two sides of tumor margins with dual-hot zones.

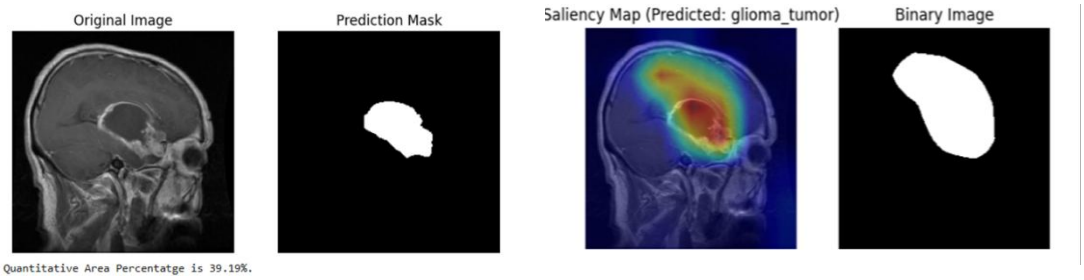


Figure 17. **Three representative scenarios of IoU values calculating from the saliency map and the U-Net tumor segmentation mask.**

(A) 0% IoU corresponds to when the U-Net fails to locate the tumor, yet the ResNet101-based classifier still makes correct predictions. (B) Low IoU values indicate that the ResNet101-based model relies on the tumor margin for classification. (C) High IoU values occur when dual-hot zones capture both sides of the tumor margin, resulting in

greater coverage of the tumor region that overlaps the important pixels in model's decision making.

4.4 Experimentation on SWIN Transformer

This praxis also explored a Vision Transformer–based architecture- a SWIN Transformer, for the same four-class brain tumor classification as in 4.3. The SWIN-based model did not match the classification performance of the ResNet101-based model, despite hyperparameter tuning and experimenting with two training strategies, which are as follows:

- **Four-Class Direct Classification**
    - The final head layer was fine-tuned and second-to-last layers was from a pre-trained SWIN Transformer.
    - While achieving 97% accuracy for identifying non-tumor scans, it struggled to distinguish among the three tumor types which are meningioma, glioma, pituitary tumors. The tumor-type accuracy remained well below the 95% threshold.
- **Two-Stage Classification**
    - First stage: Distinguish "tumor vs. no tumor."
    - Second stage: Classify tumor type (glioma, meningioma, pituitary).
    - Despite these adjustments, final tumor-type accuracy remained well below the 95% threshold.

4.4.1 Possible Reasons for SWIN Underperformance

Several factors likely contributed to the underperformance of SWIN Transformer:

- **Fine-Grained Tumor Details:** Transformers generally excel at capturing global context, but subtle edge details and fine-grained tissue contrasts in MRI may favor deep CNN residual blocks (like in ResNet) that are already well-tuned for such localized features.
- **Data Complexity:** Each tumor type's morphological variability can be high. SWIN's partitioned window attention might not sufficiently localize the complex patterns that differentiate meningioma vs. glioma margins.
- **Training Requirements:** Transformers often require very large datasets to outperform CNNs. Our dataset, although augmented, may not be large enough to unleash SWIN's full potential in a specialized medical domain.

As a result, the SWIN approach, while promising in some image analysis tasks, did not outperform ResNet101 on this specific four-class brain tumor dataset.

4.5 Summary of Findings

This chapter detailed the experimental outcomes associated with the proposed MRI-based brain tumor classification pipeline. Key points include:

1. **ResNet101-based Transfer Learning Model**
    - Achieved 98% overall accuracy, surpassing the ≥95% benchmark in Hypothesis 2.

- Demonstrated high precision and recall across all four classes.
- Saliency maps, confusion matrix, and overlap with a separate U-Net segmentation confirm the ResNet-based model relied on tumor margins, rather than tumor cores, for the classification decisions. (supporting Hypothesis 1).

2. **SWIN Transformer Challenges**
   - Although SWIN performed well in separating tumor vs. no tumor, it struggled to differentiate the three tumor subtypes.
   - Multiple training strategies, including a two-phase training, did not yield accuracy above ~65–75% for multi-class classification.
3. **Interpretability and Efficiency**
   - Feature-importance analysis showed the CNN paying close attention to tumor margins.
   - Inference time and moderate resource usage make ResNet101 suitable for real-time or near-real-time clinical scenarios.

Overall, the ResNet101-based model emerged as the best-performing and most reliable method for four-class brain tumor classification on the given MRI dataset. Chapter 5 further discusses the implications for clinical practice, the limitations of the current approach, and potential avenues for future enhancements—especially pertaining to data diversity, advanced interpretability methods, and further exploration of hybrid CNN–Transformer architectures.

Chapter 5 — Discussion and Conclusions

5.1 Discussion

This praxis set out to develop and interpret a deep learning model for automated brain MRI analysis. The developed CNN-based model can discriminate normal brain tissue from tumors and sub-classify meningioma, glioma and pituitary tumors with high accuracy, precision and recall. This praxis also worked on interpreting the model's MRI analysis through a dual-model approach, further improving the feasibility of the model being potentially deployed in a clinical setting. In detail:

- **Fine-tuning ResNet101 and clinical workflow.** Fine-tuning a ResNet-101 backbone—while freezing early layers and retraining the deepest residual block—achieved 99 % overall test accuracy, outperforming the ≤ 75 % obtained with a windowed-attention Swin Transformer. The fine-tuned ResNet101 model demonstrated potential utility in a clinical workflow as an automated tool to initially screen normal brain MRI scans out of tumor brain scans.

- **Interpretability through dual-model analysis.** Saliency maps from the classifier and masks from an independent U-Net segmentation model showed that the network's highest-weight pixels clustered at tumor margins rather than cores (IoU histograms peaked at ≤ 0.20). This aligns with radiological practice in which infiltrative borders govern brain tumor grading and surgical planning.

- **Robustness from data strategy.** Conservative spatial augmentations (±15° rotation, ≤ 15 % zoom/shift, elastic deformations) and class-balanced sampling neutralized acquisition variability across the Figshare and IXI cohorts, letting the model generalize to three orthogonal MRI planes.

- **Transformer shortcomings.** The Swin architecture generalized poorly to the three tumor sub-types despite extensive hyperparameter search were experimented. The window partition limited its receptive field for small, heterogeneous lesions, and the dataset was an order of magnitude smaller than Vision-Transformer benchmarks—both factors may have limited model convergence.

- **Clinical realism.** Inference time on an A100 GPU allows a typical 60-slice study in < 3 s, permitting near-real-time flagging during a radiologist workflow.

5.2 Conclusions and Practical Implications

By fine-tuning a ResNet101 model, this praxis developed a four-class MRI classifier that can automatedly (i) screen out normal brain scans with near-perfect confidence, (ii) reliably distinguish among glioma, meningioma and pituitary tumors and (iii) provide interpretability analysis that showed tumor-margin awareness. Leveraging data-augmentation and a fine-tuned ResNet-101 backbone, the study achieved a 99 % overall accuracy—well above the 95 % benchmark declared in Chapter 1—and demonstrated that the network's most salient cues reside at the tumor margin, aligning with the radiologic methods used in practice. When compared with a SWIN-Transformer model, the CNN model, such as a fine-tuned ResNet101, not only converged faster on a

modest dataset but also generalized better, underscoring the continued relevance of lightweight, residual architectures when data volumes are <10 k images.

Beyond the model performance metrics, this praxis research adds model interpretability in the workflow. U-Net–assisted saliency map validation can be embedded in the workflows together with brain tumor classification. The resulting "explained-as-you-screen" workflow provides both a triage mechanism for over-burdened radiology services and a transparent decision trail that can potentially be audited for model interpretability. In detail:

- For practical deployment of the model, the model can be used as a first pass high-throughput tool as perfect precision for the "no-tumor" class means normal scans can be auto-cleared, shortening time-to-diagnosis for abnormal cases in resource-limited centers.
- The model can also be potentially deployed in a clinical setting with low to moderate requirement for computational resources.
- Interpretable AI improves trust. Overlaying saliency heat-maps on U-Net tumor masks offers radiologists an intuitive quality check, easing the trust concern for integration of the model into clinical decision-making.
- The SWIN-Transformer likely required more data and compute, yet it under-performed on brain tumor classification; CNNs with a residual network (ResNet101) is preferable when datasets are small.

5.3 Recognized Limitations

Despite meeting the target classification performance thresholds, the current four-class MRI classifier is constrained by several factors that may limit its generalizability and long-term clinical impact. Limitations originated from data preparation and augmentation techniques that were used to cope with limited and heterogeneous public datasets and also to keep training feasible on a single A100 GPU. As a result, several methodological shortcuts—uniform image-resizing by downsizing, conservative data-augmentation, and partial layer freezing—may have introduced hidden biases or reduced the model's capacity to learn subtle tumor signatures across wider patient populations.

The experiments relied on single-sequence T1-weighted scans sourced from two repositories; multi-sequence data, institution-specific scanning protocols, and rare tumor sub-types were absent. Consequently, the system's robustness to unseen artefacts, scanner vendors, or demographic shifts has yet to be demonstrated. The saliency-based interpretability strategy—while promising—still falls short of a full clinical audit trail. In detail:

- Public data collections used in this praxis research may under-represent the diversity of healthy brains and do not include rare brain tumor variants, only the three main brain tumor types.
- Training and validation were limited to contrast-enhanced T1 images of MRI scans; hence the omission of T2, FLAIR, or diffusion sequences in the model restricts applicability in multi-modal workflows for brain tumor classification.

- Conservative augmentation. Rotations (±10–15°) and mild scaling used in this praxis research may not capture the full range of motion artefacts, coil inhomogeneity, or vendor-specific contrasts seen in practice.
- This praxis research used a partially frozen backbone. Freezing all but the last ResNet block accelerated training but may limit the network's ability to internalize MRI-specific low-level features, potentially capping the model's ultimate accuracy.
- Model performance on MRI scans from external hospitals, different field strengths, or other acquisition protocols remain untested.
- Alternative backbone architecture such as the SWIN-Transformer performed markedly worse, underscoring that limitations in dataset may not yield a high-performing model with state-of-the-art model architectures.
- Saliency-overlap metrics is an initial step to mitigate the "black-box" concern, but it still lacks robust validation by neuroradiologists or regulatory agencies, leaving final decision accountability unresolved.

Collectively, these recognized limitations set the boundaries within which the current system can be deployed. The limitations also highlight priority areas for subsequent research improvement, such as adding data diversity, domain-shift benchmarking, and clinician-verified explainability,

5.4 Contributions of Research

This praxis developed a fine-tuned ResNet101 model for brain tumor MRI analysis and delivered a four-class classification model for the automated analysis of

brain MRI scans that can both triage normal brain scans and differentiate glioma, meningioma and pituitary tumors with 99.1 % overall accuracy and class-wise F1 scores ≥ 0.983, surpassing the ≥ 95 % target set in Chapter 1. Beyond classification performance metrics, the study implements a margin-aware interpretability protocol in which ResNet-derived saliency maps are quantitatively overlapped with U-Net tumor masks, demonstrating that the network's decisions rely on tumors' marginal features that are also used by neuroradiologists. By embedding this "explained-as-you-screen" workflow together with the fine-tuned ResNet101 model, the research demonstrates that deep learning tools can be both transparent and clinically practical with low to moderate requirement for computational resources.

- **Dataset enrichment for clinical utility.** Integrating non-tumor images into tumor-only classification model, which are commonly reported in the literature, enables first-pass screening use and addresses a key literature gap.

- **High-fidelity four-class classifier.** 99.1 % accuracy and ≤ 1 % false-positive rate across all classes.

- **Margin-centric interpretability.** Saliency-to-mask IoU analysis demonstrated the model relied on tumor-margin features rather than tumor core, mirroring radiologic diagnosis methods.

- **Resource-efficient transfer learning and inferencing.** Fine-tuning final block with ResNet101 backbone architecture worked on an A100 GPU. Inference allows a typical 60-slice study in < 3 s on an A100 GPU, permitting near-real-time use in a radiologist workflow.

References

Chatterjee, S., Nizamani, F. A., Nürnberger, A., & Speck, O. (2022). Classification of brain tumours in MR images using deep spatiospatial models. Scientific Reports, 12(1), 1505.

Ostrom, Q. T., Patil, N., Cioffi, G., Waite, K., Kruchko, C., & Barnholtz-Sloan, J. S. (2020). CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2013–2017. Neuro-oncology, 22(Supplement_1), iv1-iv96.

Bhatia, A., Birger, M., Veeraraghavan, H., Um, H., Tixier, F., McKenney, A. S., ... & Young, R. J. (2019). MRI radiomic features are associated with survival in melanoma brain metastases treated with immune checkpoint inhibitors. Neuro-oncology, 21(12), 1578-1586.

Najjar, R. (2023). Redefining radiology: a review of artificial intelligence integration in medical imaging. Diagnostics, 13(17), 2760.

Ting, D. S. W., Pasquale, L. R., Peng, L., Campbell, J. P., Lee, A. Y., Raman, R., ... & Wong, T. Y. (2019). Artificial intelligence and deep learning in ophthalmology. British Journal of Ophthalmology, 103(2), 167-175.

https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., ... & Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. Nature, 577(7788), 89-94.

https://www.diagnosticimaging.com/view/mammography-news-fda-clears-ai-software-for-breast-arterial-calcification-detection

Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., ... & Wu, J. (2020, May). Unet 3+: A full-scale connected unet for medical image segmentation. In ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 1055-1059). IEEE.

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2022, October). Swin-unet: Unet-like pure transformer for medical image segmentation. In European conference on computer vision (pp. 205-218). Cham: Springer Nature Switzerland.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).

https://www.neurologylive.com/view/fda-clears-ai-powered-brain-mri-software-pixyl-neuro

https://www.diagnosticimaging.com/view/fda-clears-emerging-ai-software-for-brain-mri

Sultan, H.H.; Salem, N.M.; Al-Atabany, W. Multi-Classification of Brain Tumor Images Using Deep Neural Network. IEEE Access 2019, 7, 69215–69225.

Amou, M.A.; Xia, K.; Kamhi, S.; Mouhafid, M. A Novel MRI Diagnosis Method for Brain Tumor Classification Based on CNN and Bayesian Optimization. Healthcare 2022, 10, 494

Deepa, S.; Janet, J.; Sumathi, S.; Ananth, J.P. Hybrid Optimization Algorithm Enabled Deep Learning Approach Brain Tumor Segmentation and Classification Using MRI. J. Digit. Imaging 2023, 36, 847–868.

Khan, H.A.; Jue, W.; Mushtaq, M.; Mushtaq, M.U. Brain tumor classification in MRI image using convolutional neural network. Math. Biosci. Eng. 2020, 17, 6203–6216

https://www.aaroncohen-gadol.com/en/patients/brain-tumor/types/statistics#:~:text=The%20overall%20incidence%20rate%20of,7%20per%20100%2C000%20individuals%20annually.&text=35.,tumor%20type%2C%20and%20other%20factors.

Tandel, G. S., Biswas, M., Kakde, O. G., Tiwari, A., Suri, H. S., Turk, M., ... & Suri, J. S. (2019). A review on a deep learning perspective in brain cancer classification. Cancers, 11(1), 111.

Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. (2017). Deep learning for brain MRI segmentation: state of the art and future directions. Journal of digital imaging, 30, 449-459.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical image analysis, 42, 60-88.

Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., Van Ginneken, B., Madabhushi, A., ... & Summers, R. M. (2021). A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. Proceedings of the IEEE, 109(5), 820-838.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2022, October). Swin-unet: Unet-like pure transformer for medical image segmentation. In European conference on computer vision (pp. 205-218). Cham: Springer Nature Switzerland.

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

Mzoughi, H., Njeh, I., Wali, A., Slima, M. B., BenHamida, A., Mhiri, C., & Mahfoudhe, K. B. (2020). Deep multi-scale 3D convolutional neural network (CNN) for MRI gliomas brain tumor classification. Journal of Digital Imaging, 33, 903-915.

Pashaei, A., Sajedi, H., & Jazayeri, N. (2018, October). Brain tumor classification via convolutional neural network and extreme learning machines. In 2018 8th International conference on computer and knowledge engineering (ICCKE) (pp. 314-319). IEEE.

Sultan, H.H.; Salem, N.M.; Al-Atabany, W. Multi-Classification of Brain Tumor Images Using Deep Neural Network. IEEE Access 2019, 7, 69215–69225.

Abiwinanda, N., Hanif, M., Hesaputra, S. T., Handayani, A., & Mengko, T. R. (2019). Brain tumor classification using convolutional neural network. In World Congress on Medical Physics and Biomedical Engineering 2018: June 3-8, 2018, Prague, Czech Republic (Vol. 1) (pp. 183-189). Springer Singapore.

Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A., & Baik, S. W. (2019). Multi-grade brain tumor classification using deep CNN with extensive data augmentation. Journal of computational science, 30, 174-182.

Ullah, N., Khan, J. A., Khan, M. S., Khan, W., Hassan, I., Obayya, M., ... & Salama, A. S. (2022). An effective approach to detect and identify brain tumors using transfer learning. Applied Sciences, 12(11), 5645.

Deepak, S., & Ameer, P. M. (2019). Brain tumor classification using deep CNN features via transfer learning. Computers in biology and medicine, 111, 103345.

Huang, D., Guo, H., & Zhang, Y. (2021). ADD-Net: Attention U-Net with Dilated Skip Connection and Dense Connected Decoder for Retinal Vessel Segmentation. In Advances in Computer Graphics: 38th Computer Graphics International Conference, CGI 2021, Virtual Event, September 6–10, 2021, Proceedings 38 (pp. 327-338). Springer International Publishing.

https://jinglescode.github.io/2019/12/08/biomedical-image-segmentation-u-net-attention/

Alexey, D. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv: 2010.11929.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).

Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., ... & Xu, D. (2022). Unetr: Transformers for 3d medical image segmentation. In Proceedings of the IEEE/CVF winter conference on applications of computer vision (pp. 574-584).

Huang, J., Fang, Y., Wu, Y., Wu, H., Gao, Z., Li, Y., ... & Yang, G. (2022). Swin transformer for fast MRI. Neurocomputing, 493, 281-304.

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., ... & Zhou, Y. (2021). Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306.

Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A convnet for the 2020s. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11976-11986).

Wenxuan, W., Chen, C., Meng, D., Hong, Y., Sen, Z., & Jiangyun, L. (2021). Transbts: Multimodal brain tumor segmentation using transformer. In International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer (pp. 109-119).

Zhou, H. Y., Guo, J., Zhang, Y., Yu, L., Wang, L., & Yu, Y. (2021). nnformer: Interleaved transformer for volumetric segmentation. arXiv preprint arXiv:2109.03201.

Lin, A., Chen, B., Xu, J., Zhang, Z., Lu, G., & Zhang, D. (2022). Ds-transunet: Dual swin transformer u-net for medical image segmentation. IEEE Transactions on Instrumentation and Measurement, 71, 1-15.

Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods, 18(2), 203-211.

He, D., Zhang, Y., Huang, H., Si, Y., Wang, Z., & Li, Y. (2023). Dual-branch hybrid network for lesion segmentation in gastric cancer images. Scientific Reports, 13(1), 6377

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

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., ... & Van Leemput, K. (2014). The multimodal brain tumor image segmentation benchmark (BRATS). IEEE transactions on medical imaging, 34(10), 1993-2024.

Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J. S., ... & Davatzikos, C. (2017). Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. Scientific data, 4(1), 1-13.

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., ... & Prior, F. (2013). The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. Journal of digital imaging, 26, 1045-1057.

Nyúl, L. G., & Udupa, J. K. (1999). On standardizing the MR image intensity scale. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 42(6), 1072-1081.

Kleesiek, J., Urban, G., Hubert, A., Schwarz, D., Maier-Hein, K., Bendszus, M., & Biller, A. (2016). Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. NeuroImage, 129, 460-469.

Manjón, J. V., Coupé, P., Martí-Bonmatí, L., Collins, D. L., & Robles, M. (2010). Adaptive non-local means denoising of MR images with spatially varying noise levels. Journal of Magnetic Resonance Imaging, 31(1), 192-203.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of big data, 6(1), 1-48.

Perez, L. (2017). The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621.

Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4401-4410).

Swati, Z. N. K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., & Lu, J. (2019). Brain tumor classification for MR images using transfer learning and fine-tuning. Computerized Medical Imaging and Graphics, 75, 34-46.

Alanazi, M. F., Ali, M. U., Hussain, S. J., Zafar, A., Mohatram, M., Irfan, M., ... & Albarrak, A. M. (2022). Brain tumor/mass classification framework using magnetic-resonance-imaging-based isolated and developed transfer deep-learning model. Sensors, 22(1), 372.

Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. IEEE transactions on medical imaging, 35(5), 1240-1251.

Isunuri, B. V., & Kakarla, J. (2022). Three‐class brain tumor classification from magnetic resonance images using separable convolution based neural network. Concurrency and Computation: Practice and Experience, 34(1), e6541.

Ghosal, P., Nandanwar, L., Kanchan, S., Bhadra, A., Chakraborty, J., & Nandi, D. (2019, February). Brain tumor classification using ResNet-101 based squeeze and excitation deep neural network. In *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)* (pp. 1-6). IEEE.

Çınar, N., Kaya, B., & Kaya, M. (2022, March). Comparison of deep learning models for brain tumor classification using MRI images. In *2022 International conference on decision aid sciences and applications (DASA)* (pp. 1382-1385). IEEE.

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

Sultan, H. H., Salem, N. M., & Al-Atabany, W. (2019). Multi-classification of brain tumor images using deep neural network. IEEE access, 7, 69215-69225.

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

Ait Amou, M., Xia, K., Kamhi, S., & Mouhafid, M. (2022, March). A novel MRI diagnosis method for brain tumor classification based on CNN and Bayesian Optimization. In Healthcare (Vol. 10, No. 3, p. 494). MDPI.

Trebing, K., Stańczyk, T., & Mehrkanoon, S. (2021). SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture. Pattern Recognition Letters, 145, 178-186.

Deepa, S., Janet, J., Sumathi, S., & Ananth, J. P. (2023). Hybrid optimization algorithm enabled deep learning approach brain tumor segmentation and classification using MRI. Journal of Digital Imaging, 36(3), 847-868.

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

Çinar, A., & Yildirim, M. (2020). Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. Medical hypotheses, 139, 109684.

Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.

Swati, Z. N. K., Zhao, Q., Kabir, M., Ali, F., Ali, Z., Ahmed, S., & Lu, J. (2019). Brain tumor classification for MR images using transfer learning and fine-tuning. Computerized Medical Imaging and Graphics, 75, 34-46.

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

Natekar, P., Kori, A., & Krishnamurthi, G. (2020). Demystifying brain tumor segmentation networks: interpretability and uncertainty analysis. Frontiers in computational neuroscience, 14, 6.

Islam, T., Hafiz, M. S., Jim, J. R., Kabir, M. M., & Mridha, M. F. (2024). A systematic review of deep learning data augmentation in medical imaging: Recent advances and future research directions. *Healthcare Analytics*, 100340.

Zhou, S. K., Greenspan, H., Davatzikos, C., Duncan, J. S., Van Ginneken, B., Madabhushi, A., ... & Summers, R. M. (2021). A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. *Proceedings of the IEEE*, *109*(5), 820-838.

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., ... & Zhou, Y. (2021). Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*.

Watts, J., Box, G., Galvin, A., Brotchie, P., Trost, N., & Sutherland, T. (2014). Magnetic resonance imaging of meningiomas: a pictorial review. *Insights into imaging*, *5*, 113-122.

Prager, A. J., Martinez, N., Beal, K., Omuro, A., Zhang, Z., & Young, R. J. (2015). Diffusion and perfusion MRI to differentiate treatment-related changes including

pseudoprogression from recurrent tumors in high-grade gliomas with histopathologic evidence. American Journal of Neuroradiology, 36(5), 877-885.

Joo, L., Park, J. E., Park, S. Y., Nam, S. J., Kim, Y. H., Kim, J. H., & Kim, H. S. (2021). Extensive peritumoral edema and brain-to-tumor interface MRI features enable prediction of brain invasion in meningioma: Development and validation. *Neuro-oncology*, *23*(2), 324-333.

Padmaja, S., Jakkani, A., Vijay, K. K., Moorthy, N. L. N., & Ramesh, G. (2024). Magnetic resonance imaging spectrum of intracranial meningiomas: An institutional review of common and uncommon imaging appearances. *Journal of Clinical Imaging Science*, *14*.