People's Democratic Republic of Algeria
Ministery of Higher Education and Scientific Research
Mohamed El Bachir El Ibrahimi University of Borj Bou Arréridj
Faculty of Mathematics and Computer Science
Department of Computer Science

DISSERTATION
Presented in fulfillment of the requirements of obtaining the degree
Master in Computer Science
Specialty: Networking & Multimedia

THEME

Brain MRI Image Segmentation Using YOLO and U-Net Deep Learning Models

Presented by:
Dhikra YOUSFI

Publicly defended on: 10/06/2025
In front of the jury composed of:
President: Dr.Mouhoub BELAZOUG
Examiner: Dr.Farid NOUIOUA
Supervisor: Dr.Foudhil BELHADJ

2024/2025

Dedication

To my mother

whose unconditional love, countless sacrifices, and endless encouragement have been the greatest inspiration throughout my journey.

To my father

whose pride in me brings joy and whose unwavering support has always pushed me forward.

To my little sister Douaa

whose lighthearted spirit and presence bring happiness to our family every day.

To my brothers Aymen and Sohaib

for their constant love, support, and motivation.

To my grandmother

whose gentle presence and heartfelt prayers have always blessed my path.

To my friends

who have stood by me with encouragement, patience, and loyalty throughout this academic journey.

To each of you, this accomplishment is a reflection of your love, support, and belief in me. I am deeply and sincerely thankful.

Acknowledgment

Foremost of all, I am very thankful to God Almighty for the strength, tenacity, and sound judgment to be able to complete this work.

I would like to thank my supervisor, Mr. Belhadj Foudil, for his guidance, support, and constructive feedback that were precious at all levels of this thesis. His patience, encouragement, and trust have had a lasting effect on my academic and personal development.

My sincere gratitude also goes to all faculty who have shared their knowledge with passion and dedication and to the university community, which provided a supportive learning and development environment.

I am also grateful to all those who, in some direct or indirect way, provided support, motivation, or inspiration throughout the process.

Last but not least, I would like to thank my family and friends from the bottom of my heart for their unwavering support, love, and motivation throughout this time. They have been a strength in having faith in me.

Abstract

The precise detection of brain tumors is crucial in medical practice, and the use of automated segmentation techniques has the potential to improve this significantly. The paper presents a comparative evaluation of U-Net and YOLOv8 deep learning models for automatic brain tumor segmentation from magnetic resonance imaging (MRI). U-Net, with its ability to achieve pixel-level accuracy, performed superiorly in terms of Intersection over Union (IoU) and Dice coefficient, indicating its robustness in boundary delineation. In contrast, YOLOv8 had better precision and recall, thus being more appropriate for fast segmentation. The comparison was made with the standard metrics of precision, recall, IoU, Dice coefficient, and F1 score, providing a well-rounded evaluation. The findings show that U-Net is better at generating precise segmentation boundaries, whereas YOLOv8 performs better when it comes to detecting tumors quickly and accurately. This comparison provides important insights for the determination of the most appropriate model depending on specific application requirements.

Keywords: Brain Tumor Segmentation, U-Net, YOLOv8, Deep Learning, MRI, Medical Imaging.

Résumé

La détection précise des tumeurs cérébrales est cruciale dans la pratique médicale, et l'utilisation de techniques de segmentation automatique peut améliorer considérablement ce processus. Cet article présente une évaluation comparative des modèles d'apprentissage profond U-Net et YOLOv8 pour la segmentation automatique des tumeurs cérébrales à partir d'images de résonance magnétique (IRM). U-Net, avec sa capacité à atteindre une précision au niveau des pixels, a montré des performances supérieures en termes d'intersection sur union (IoU) et de coefficient de Dice, ce qui témoigne de sa robustesse dans la délimitation des contours. En revanche, YOLOv8 a obtenu de meilleurs résultats en termes de précision et de rappel, le rendant plus adapté à une segmentation rapide. La comparaison a été réalisée à l'aide des métriques standard de précision, de rappel, de IoU, de coefficient de Dice et de score F1, offrant une évaluation complète. Les résultats montrent que U-Net est plus performant pour générer des contours de segmentation précis, tandis que YOLOv8 excelle dans la détection rapide et précise des tumeurs. Cette comparaison fournit des informations importantes pour déterminer le modèle le plus approprié en fonction des exigences de l'application.

Mots-clés : Segmentation des tumeurs cérébrales, U-Net, YOLOv8, Apprentissage profond, IRM, Imagerie médicale.

ملخص

الكشف الدقيق عن أورام الدماغ أمر بالغ الأهمية في الممارسة الطبية، ويمكن أن تُحسن تقنيات التقسيم الآلي من هذه العملية بشكل كبير. تقدم هذه الدراسة تقييماً مقارناً بين نموذجين للتعلم العميق، (U-Net) و (YOLOv8)، لتقسيم أورام الدماغ تلقائياًا من صور الرنين المغناطيسي (MRI) . أظهر نموذج (U-Net)، بقدرته على تحقيق دقة على مستوى البكسل، أداءً متفوقاًا من حيث معامل التقاطع (IoU) ومعامل (Dice)، مما يعكس قوته في تحديد حدود الأورام بدقة. في المقابل، حقق نموذج (YOLOv8) دقة واسترجاع أعلى، مما يجعله أكثر كفاءة في عمليات التقسيم السريعة. تم إجراء المقارنة باستخدام المقاييس القياسية للدقة (Precision)، الاسترجاع (Recall)، (IoU)، معامل (Dice)، ومعامل (F1)، مما يضمن تقييماًا شاملاً. أظهرت النتائج أن (U-Net) يتفوق في إنتاج حدود تقسيم دقيقة، في حين أن (YOLOv8) أفضل في اكتشاف الأورام بسرعة ودقة. توفر هذه الدراسة رؤى مهمة لاختيار النموذج الأنسب وفقاًا لمتطلبات التطبيق.

الكلمات المفتاحية: تقسيم أورام الدماغ، التعلم العميق، الرنين المغناطيسي، التصوير الطبي.

Table of Contents

List of Abbreviations

AI Artificial Intelligence

CNN Convolutional Neural Network

YOLO You Only Look Once

MRI Magnetic Resonance Imaging

CT Computed Tomography

FLAIR Fluid-Attenuated Inversion Recovery

DWI Diffusion-Weighted Imaging

PWI Perfusion-Weighted Imaging

MRS Magnetic Resonance Spectroscopy

IoU Intersection over Union

BBox Bounding Box

JSON JavaScript Object Notation

COCO Common Objects in Context

ReLU Rectified Linear Unit

PANet Path Aggregation Network

BiFPN Bidirectional Feature Pyramid Network

BCE Loss Binary Cross Entropy Loss

NCE Loss Negative Cross Entropy Loss

List of Figures

List of Tables

List of Algorithms

General Introduction

The fast evolution of deep learning and the artificial intelligence field has transformed numerous disciplines, and medical imaging is no exception. One of the most hopeful applications of deep learning in the medical domain is detecting and segmenting brain tumors based on magnetic resonance imaging (MRI). Brain tumors are severe health issues that need early and correct detection to ensure proper medical interventions. MRI is routinely used for brain tumor diagnosis since it provides high-resolution images and is able to image soft tissues accurately.

This thesis tries to utilize deep learning techniques to achieve precise segmentation of brain tumors in MRI images. To this end, two state-of-the-art models, U-Net and YOLOv8, were designed and tested for the task. U-Net, which has high pixel-level precision, was selected due to its capacity to generate intricate segmentation maps. Comparatively, we selected the efficient and swift YOLOv8 model due to its superior detection performance. In this comparison, we hope to enlighten the most suitable approach for different medical applications.

For this purpose, we employed a huge dataset of brain MRI images, both images of tumor and healthy tissue, for training our models.

By utilizing the deep learning approach, these models have learned on their own to recognize salient features of brain tumors, enabling the precise delineation of tumor-affected regions. This study seeks to enhance the ability of computational systems to detect brain tumors, raising the prospects for early detection and improving healthcare outcomes for patients afflicted with this serious condition.

By providing precise and effective means for the detection of brain tumors, we wish to contribute positively to medical practice and patient care.

This dissertation has three chapters, each of which sheds light on various facets of our

research.

Chapter 1 provides an overview of the human brain, citing its structural and functional complexity and posing a daunting challenge to clinical diagnosis. It also discusses the classification, symptoms, and etiology of brain tumors. The chapter ends with a discourse on medical imaging, and specifically Magnetic Resonance Imaging (MRI), and its invaluable role in the precise detection and diagnosis of brain tumors.

Chapter 2 covers the theoretical foundation of deep learning, in particular convolutional neural networks (CNNs), which have been widely used for computer vision and medical image processing. It also talks about U-Net and YOLOv8 architecture and training.

Chapter 3 describes the execution of our brain tumor segmentation approach through U-Net and YOLOv8, discusses the outcome of our experimental procedures, and gives a comparative evaluation of the models' performance.

We conclude with a general conclusion and outlook, stating the conclusions and suggesting future work. This thesis has tried to contribute meaningfully to brain tumor segmentation using deep learning, proposing a comprehensive methodology, in-depth analyses, and strong experimental results.

Chapter 1

Fundamentals of Brain Tumors and Medical Imaging

1.1 Introduction

Brain cancer is a serious illness that can affect the body in various ways, depending on the location where it takes place. When detected early, treatment becomes more effective and the prognosis for patients becomes better. Brain tumors can be primary (developing in the brain) or secondary (developing in other parts), which can metastasize to other organs like the lungs or breasts. They can be benign or malignant; malignant tumors are more invasive. MRI is widely used for fast and accurate tumor detection, but a correct diagnosis needs accurate technology. This chapter explains tumor classification, symptoms, diagnosis, and segmentation issues, leading to automated detection methods.

1.2 Overview of the Human Brain

One of the most complex and amazing organs in the human body, the brain regulates thought, movement, emotion, and a host of vital physiological processes[3]. A very sophisticated mechanism for processing sensory information and coordinating responses, it has about 100,000 kilometers of interconnectivity and some 100 billion neurons[4].

The brain also undergoes a complex development process beginning in the third week of

pregnancy and extending to late adolescence, with its structural development extending well into adulthood [5]. It is composed of many interrelated regions, each specialized for particular functions, that all work together for higher-order cognitive functions, problem-solving ability, and overall intelligence [6].

1.2.1 Basic Structure of the Brain

The brain is structurally divided into several main parts, as illustrated in Figure 1.1:

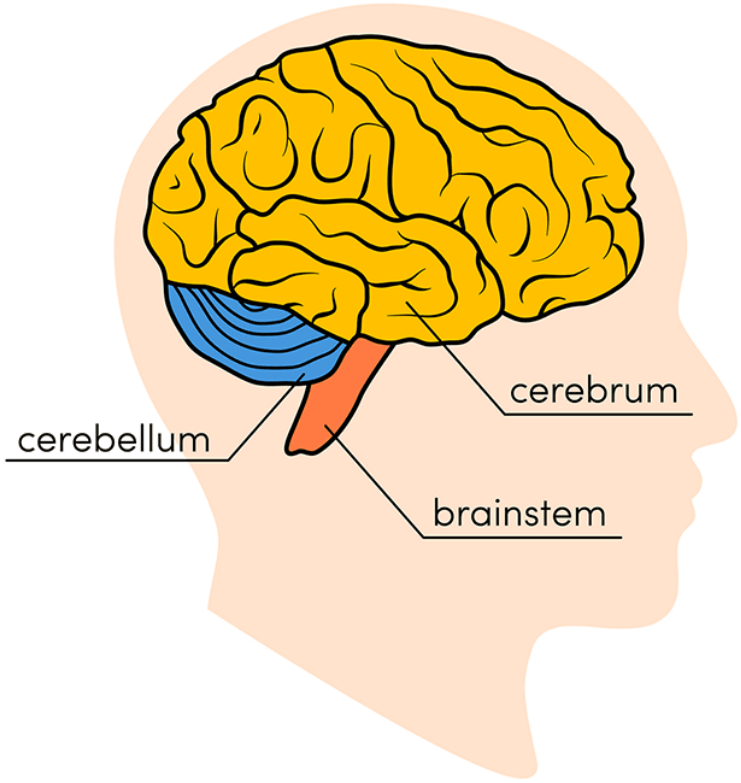

Figure 1.1: Basic structure of the brain
[7]

- **Cerebrum:** The largest part, responsible for higher cognitive functions, sensory processing, voluntary movements, and language [3].
- **Cerebellum:** Located at the back of the brain, this region is crucial for balance, coordination, and fine motor control [4].

- **Brainstem:** Connects the brain to the spinal cord and regulates vital functions such as breathing, heart rate, and digestion [5].

1.2.1.1 Lobes of the Brain

The cerebrum is divided into four lobes [8] as shown in Figure 1.2:

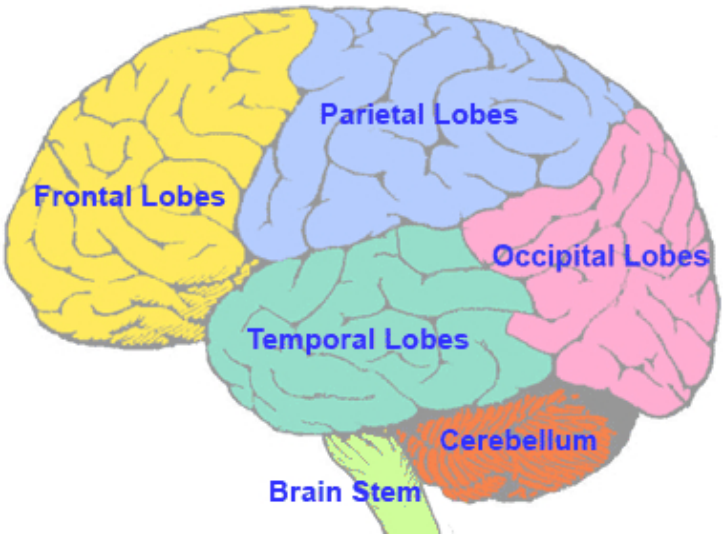

Figure 1.2: Major lobes and functional regions of the human brain.
[9]

1. **Frontal lobe:** engaged in emotions, voluntary movement, problem-solving, and reasoning.
2. **Parietal lobe:** processes touch, temperature, and spatial awareness.
3. **Temporal lobe:** essential for auditory processing and memory.
4. **Occipital lobe:** primarily responsible for vision.

Synaptic connections between neurons create complex networks that facilitate learning and adaptability [5].Additionally, the brain has white matter, which promotes communication between areas, and gray matter, where processing takes place [5].

1.2.2 Functional Regions of the Brain

Functionally, the brain is separated into several processing areas, each of which is tailored to a certain task:

- **Sensory and Motor Areas:**
  - The **primary motor** cortex controls voluntary movements.
  - The **somatosensory cortex** processes touch, pain, and temperature [8].
- **Visual and Auditory Processing:**
  - The **occipital lobe** contains the primary **visual cortex (V1)**, which processes images.
  - The **temporal lobe** houses **the auditory cortex**, responsible for sound perception [8].
- **Memory and Learning:**
  - The **hippocampus** plays a key role in memory formation and spatial navigation [5].
- **Emotion and Decision-Making:**
  - The **prefrontal cortex** is crucial for decision-making, personality, and impulse control.
  - The **amygdala** regulates emotions like fear and pleasure [4].
- **Language Processing:**
  - **Broca's area** (frontal lobe) is essential for speech production.
  - **Wernicke's area** (temporal lobe) helps in language comprehension [8].

Because the brain relies on networks that integrate many regions, its activities are highly interrelated [6] as Figure 1.3 shows.

1.2.3 Brain Function and Its Complexity in Diagnosis

Analyzing both anatomical and functional connections is necessary to comprehend how the brain works. Complex neuronal circuits in the brain allow many regions to work together to produce perception, movement, and thought [6].

- **Neural Network Complexity:** The brain is a highly interconnected system where mental

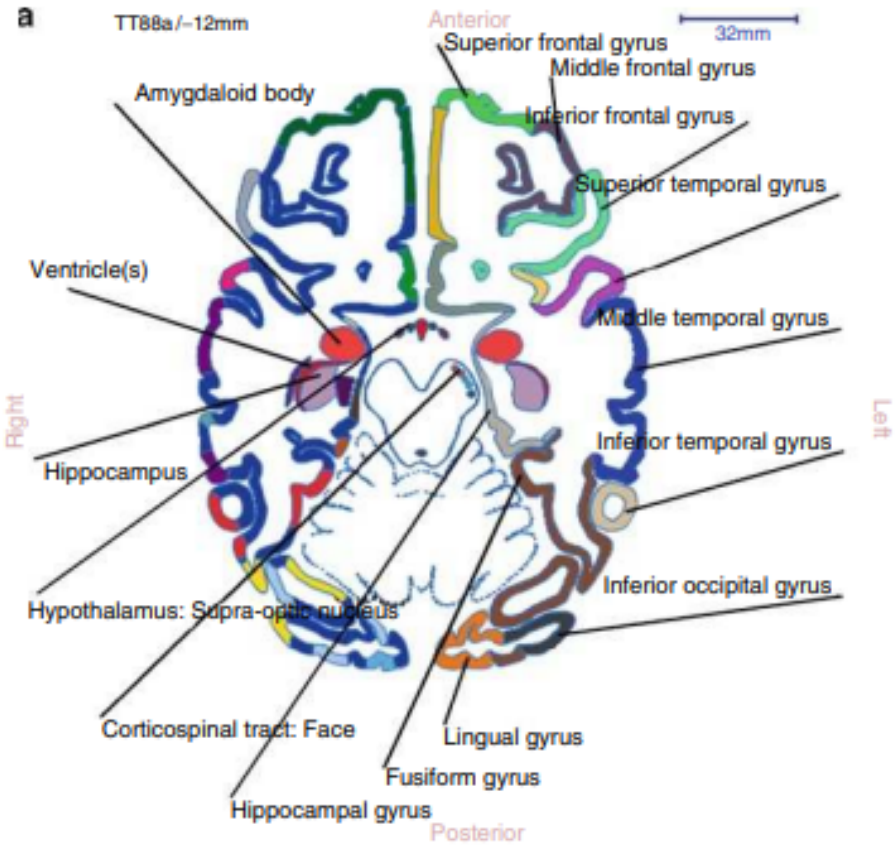

Figure 1.3: Functional Regions of the Human Brain.
[10]

states arise from interactions at multiple levels of processing.Recent advances in complex network theory offer new perspectives on how brain regions interact [6].

- **Cognitive and Behavioral Implications:** Functional brain regions do not operate singly; they interact dynamically and influence each other to control behavior and adjust to novel stimuli. Breakdowns in brain connectivity can lead to neurological disorders like Alzheimer's, schizophrenia, or traumatic brain injury [4].

- **Challenges in Diagnosis:** Diagnosis of diseases related to the brain remains challenging because of interpersonal heterogeneity of neural structure and superimposition of the function of various areas. Current imaging methods such as MRI and functional MRI (fMRI) enable an understanding of such complexity [8].

1.3 Brain Tumors

Brain tumors are a general group of neoplasms that happen in the central nervous system (CNS). They may be benign, slow-growing tumors or malignant, fast-growing tumors[11][12]. The tumors may arise in the brain itself primarily (primary tumors) or may be metastatic tumors from cancer spread from any other location in the body [13].

1.3.1 Definition

A brain tumor is an abnormal cell growth that can be in the brain or from another location in the body. The tumors vary in their severity, from benign and slow-growing types to very malignant and aggressive types [14].

1.3.2 Classification of Brain Tumors

Brain tumors are categorized according to their nature (benign or malignant) and origin (primary or secondary).

1.3.2.1 Based on Nature

Brain tumors can be benign or malignant depending on their biologic behavior.

- **Benign Tumors:** Grow slowly and do not invade nearby brain tissue like meningiomas and pituitary adenomas [15].
- **Malignant Tumors:** Aggressive, malignant, and able to spread to other brain regions. Glioblastoma Multiforme (GBM) is the most malignant primary brain tumor [14].

1.3.2.2 Based on Origin

Brain tumors are also classified based on their origin

- **Primary Brain Tumors:** Gliomas, meningiomas, and medulloblastomas are all brain-derived tumors [15].
- **Secondary (Metastatic) Tumors:** Start in other organs (lungs, breast, melanoma) and spread to the brain. These tumors are more prevalent than primary brain tumors [14].

1.3.2.3 Common Types of Brain Tumors

Tumor Type	Origin	Characteristics
Gliomas	Glial cells	- **Astrocytomas**: Range from low-grade (slow-growing) to high-grade (glioblastomas, highly aggressive). - **Oligodendrogliomas**: Arise from oligodendrocytes; typically low to intermediate grade. - **Ependymomas**: Originate from ependymal cells lining brain ventricles.
Meningiomas	Meninges (brain/spinal cord coverings)	Generally benign and slow-growing; can cause symptoms by pressing on the brain or spinal cord.
Pituitary Adenomas	Pituitary gland	Usually benign; can affect hormone production, leading to various systemic symptoms.
Medulloblastomas	Cerebellum	Highly malignant; more common in children; can spread to other parts of the central nervous system.
Schwannomas	Schwann cells (nerve sheath)	Typically benign; commonly affect the vestibular nerve, leading to hearing loss and balance issues.
Primary CNS Lymphomas	Lymphatic tissue of the central nervous system	Aggressive tumors; often associated with immunocompromised individuals.

Table 1.1: Classification of Brain Tumors Based on Origin and Characteristics [1]

1.3.3 Causes and Risk Factors

Several factors may contribute to brain tumor development, including genetics, environment, and age-related risks.

1.3.3.1 Genetic Predisposition

Brain tumors may be associated with genetic mutations like TP53, NF1, and IDH1 that are associated with high susceptibility [12].

Certain genetic conditions, such as Li-Fraumeni Syndrome, greatly raise the risk of developing astrocytomas and gliomas [15].

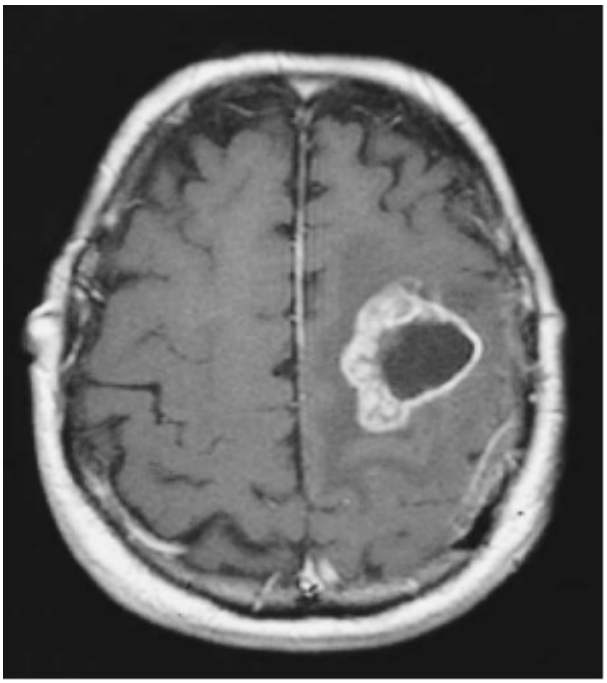

Figure 1.4: MRI of Glioblastoma Multiforme in the Left Frontal Lobe
[12]

1.3.3.2 Environmental Factors

Radiation exposure, such as from radiation therapy and nuclear accidents, is a significant risk factor [12]. Chemical exposure (e.g., industrial chemicals, pesticides) is also thought to play a role, but there is no clear evidence [14].

1.3.3.3 Age and Lifestyle Factors

Though medulloblastomas and ependymomas are more prevalent among children, brain tumors are more prevalent in older people [15]. Alcohol, tobacco, and obesity can all increase the risk of tumors [12].

1.4 Symptoms and Diagnosis of Brain Tumors

Depending on the tumor's location and rate of growth, brain tumors can present with a variety of neurological symptoms. Since delayed detection can result in severe neurological impairments, early diagnosis is essential [16].

Symptom	Description
Headaches	A typical early symptom that is frequently worse in the morning and linked to nausea [14].
Seizures	May indicate a brain tumor for the first time, especially in low-grade gliomas [17].
Vision Impairment	Tumors near the optic nerve may cause blurry vision or double vision [12].
Cognitive Decline	Confusion and memory loss are frequently mistakenly attributed to aging [16].
Speech & Motor Difficulties	Language and mobility can be affected by frontal or parietal lobe tumors [18].

Table 1.2: Common Neurological Symptoms of Brain Tumors

1.4.1 Common Symptoms of Brain Tumors

1.4.2 Methods of Diagnosis

- **Clinical Examination:** A neurological assessment is performed to examine cognitive function, reflexes, and coordination [16].
- **Neuroimaging:** The gold standard for identifying brain cancers is MRI with contrast [19]. CT scans can be used to detect bleeding or big tumors [20].
- **Biopsy:** A surgical or needle biopsy is performed to determine tumor type and grade [14].

1.5 Medical Imaging in Healthcare

Medical imaging is a very important part of disease diagnosis at an early stage, treatment planning, and following up on patients' status. In neuro-oncology, MRI, CT, and PET scans are utilized to evaluate tumor size, position, and metabolism [21].

Computed tomography (CT) scanning and magnetic resonance imaging (MRI) are the pillars of tumor diagnosis in the brain. Imaging modalities offer differentiation among malignant and benign tumors, measure tumor growth, and monitor for response to therapy [22].

1.5.1 Common Imaging Techniques for Brain Tumors

1.5.1.1 X-ray Imaging

1. X-ray imaging was among the earliest of the medical imaging technologies to be used in neuro-oncology. Its use in diagnosing brain tumors is, nevertheless, limited since it is unable to image soft tissue features as indicated by Figure 1.5 [22].

2. X-ray imaging is now mainly utilized for the detection of fractures of the skull and intracranial abnormalities, rather than for brain malignancies [23].

Figure 1.5: Cranial X-ray image.
[24]

1.5.1.2 Computed Tomography (CT) Scan

1. CT scans provide quick and effective imaging, especially in acute settings where prompt decision-making is critical [22].

2. As seen in Figure 1.6 CT scans can pick up on hemorrhagic brain tumors and tumor calcifications that MRI could miss [21].

3. However, CT scans provide less soft tissue contrast than MRI, making them less helpful for precise tumor localization [25].

Figure 1.6: Computed Tomography (CT) scan of the brain.
[24]

1.5.1.3 Ultrasound Imaging

1. The ultrasound scanning is not routinely used for adult brain tumors because of the inadequate penetration of the skull. It is used primarily for the neonatal brain imaging when ventricular abnormalities and intracranial hemorrhage in newborns are to be detected [22][21].

2. Intraoperative ultrasound can help in real-time navigation of brain tumor resections [25].

1.5.1.4 Magnetic Resonance Imaging (MRI)

1. The gold standard in the diagnosis of brain tumors is MRI, which produces high-resolution soft tissue images, Figure 1.7 illustrates that [21].

2. Specialized MRI sequences, including diffusion-weighted imaging (DWI), perfusion-weighted imaging (PWI), and MR spectroscopy (MRS), give accurate details on tumor characteristics [25].

3. MRI can identify tumor vascularity, quantify blood-brain barrier integrity, and quantify tumor response to treatment [21].

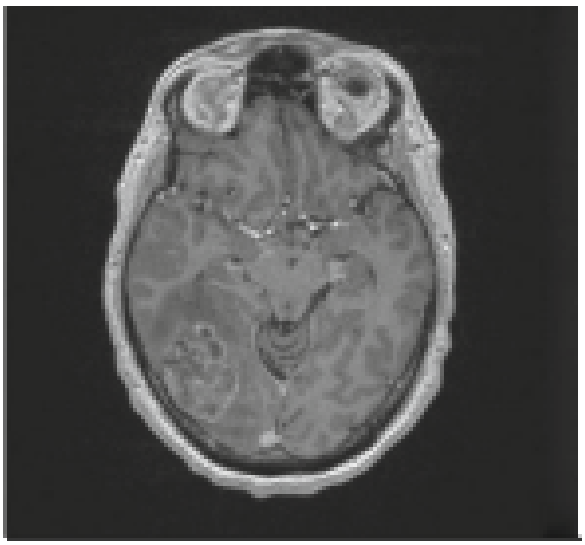

Figure 1.7: Magnetic Resonance Imaging (MRI) scan of the brain.
[24]

1.6 Magnetic Resonance Imaging (MRI) and Its Role in Neuroimaging

MRI aligns the hydrogen nuclei of the body using magnetic fields and radio waves, forming high-resolution images as they resume their normal state. MRI does not employ ionizing radiation as CT scans and X-rays do and is therefore safer when used repeatedly. Its ability to distinguish between soft tissues based on water and fat content makes it very suitable for brain imaging [26][27]. As evident from Figure 1.8, MRI scans can distinctively highlight various types of tumors in various planes, with the tumors outlined in red, facilitating accurate diagnosis and evaluation.

1.6.1 Types of MRI Sequences

Different MRI sequences are used to highlight various tissue properties, Figure 1.9 illustrate this:

- **T1-Weighted Imaging (T1WI):** Tumors appear darker than normal tissue, allowing for comprehensive anatomical viewing [26].
- **T2-Weighted Imaging (T2WI):** Highlights fluid-filled areas, making tumor-related edema more noticeable [29].

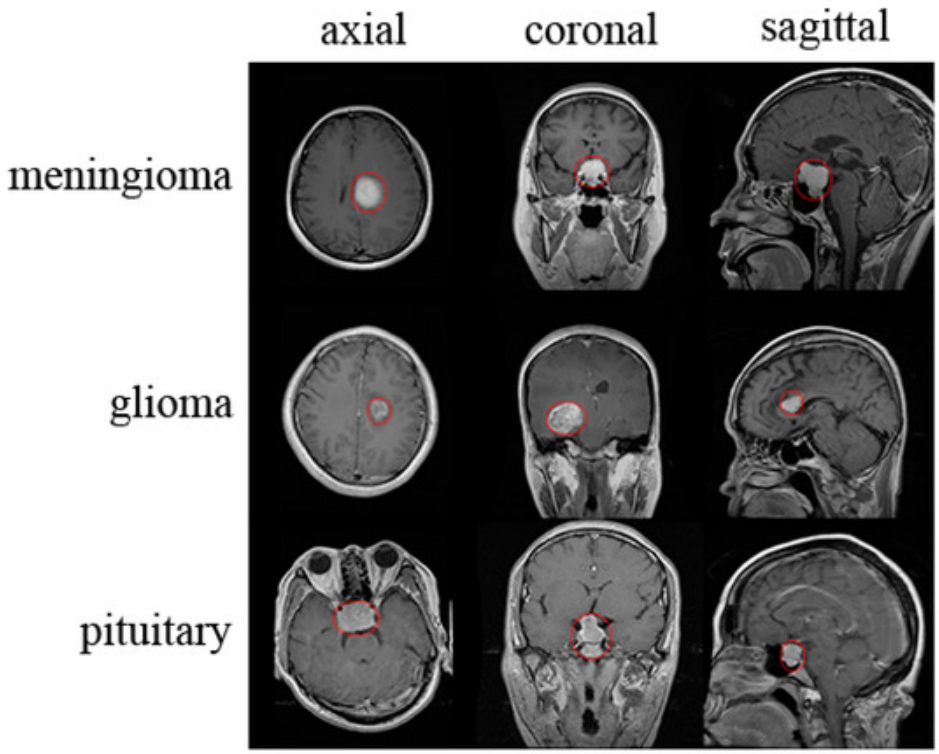

Figure 1.8: MRI scans showing different tumor types in various planes, with tumors outlined in red.

[28]

- **Fluid-Attenuated Inversion Recovery (FLAIR):** Suppresses fluid signals to improve tumor contrast with cerebrospinal fluid [30].

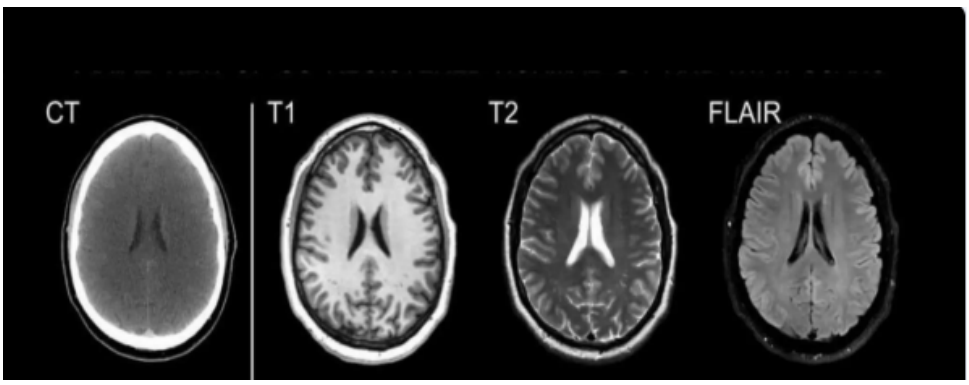

Figure 1.9: Comparison of Brain Imaging Modalities and MRI Sequences.
[31]

1.6.2 Advantages of MRI in Brain Tumor Diagnosis

MRI has various benefits in the diagnosis and monitoring of brain tumors:

- Superior soft tissue contrast allows for clear tumor visualization.

- Multi-sequence imaging enables differentiation of tumor components (solid vs. necrotic areas).
- No ionizing radiation exposure, making it safer for long-term monitoring [32].

1.6.3 Challenges of MRI in Brain Tumor Imaging

Aside from its advantages, MRI also has limitations:

- Expensive and difficult to access in environments with limited resources.
- Increased scan times, which might cause discomfort for patients.
- Quality of the image can be degraded by noise and artifacts, particularly in motion-sensitive patients [27].

1.7 Challenges in Brain Tumor Detection and Segmentation

Brain tumors are difficult to segment and detect since they are complicated and multiplex. There are certain significant challenges that affect accuracy and consistency of diagnosis.

1.7.1 Variability in Tumor Characteristics

Since brain tumors come in all shapes, sizes, and locations, they can be difficult to locate. Segmentation is made more difficult since they are very similar to normal tissue.

1.7.2 Challenges in Manual Tumor Segmentation

Manual segmentation, being time-consuming and subjective, may lead to differences among radiologists and potential inconsistency in diagnosis.

1.7.3 The Need for Automated Tumor Detection

The Need for Automated Tumor brain Detection Machine learning They enhance speed and accuracy and reduce human errors in tumor diagnosis. Machine learning algorithms are highly used in medical imaging.

1.8 Conclusion

In this chapter, we discussed the overall principles of brain tumors, their definition, classification, signs and symptoms, and methods of diagnosis. We saw how crucial medical imaging, specifically MRI, is in the detection of tumors and briefly discussed the issues of manual segmentation and its limited accuracy and efficiency. We also presented the automated detection of tumors as a solution to this problem. The next chapter will cover intelligent computing techniques, how smarter algorithms segment brain tumors more effectively, and how they improve diagnostic precision.

Chapter 2

Deep Learning-Based Techniques for Brain Tumor Segmentation

2.1 Introduction

Artificial Intelligence (AI) has become a disruptive force in healthcare due to the increasing need for accurate and quick diagnosis, particularly in the area of medical imaging. The development of AI techniques for brain tumor segmentation is examined in this chapter, starting with conventional machine learning methods and moving on to more sophisticated deep learning strategies. Convolutional neural networks (CNNs) and the development of the YOLO (You Only Look Once) architecture are given particular attention, leading to the real-time segmentation and detection capabilities of YOLOv8. The benefits of contemporary architectures in clinical applications are demonstrated by comparative studies of various segmentation techniques.

2.2 Artificial Intelligence in Healthcare

For a very long time, artificial intelligence has been influencing technical innovation. It is crucial that the digital revolution of modern life leads to the emergence of new technology tools through AI-based automated solutions. The impact of the shift has been felt in both technical and everyday domains as a result of the use of intense AI approaches. Despite its widespread use across numerous industries, artificial intelligence is closely related to healthcare and medi-

cal applications. AI-based solutions have been guaranteeing state-of-the-art results in medical issues, from drug discovery to genomics to medical diagnosis and treatment planning. The literature is particularly interested in image-based solutions among these problem orientations since AI-powered automated analyses provide accurate detections and quicker decision-making [33].

2.2.1 AI in Brain Tumor Analysis and it's Techniques

Brain tumour segmentation has been revolutionised by artificial intelligence, first through conventional machine learning (ML) approaches and more recently through deep learning (DL) techniques. The generalisation and accuracy of traditional machine learning techniques, including support vector machines (SVMs) and decision trees, are limited by their reliance on manually created features and their inability to handle the intricacy of MRI data [34]. Convolutional neural networks (CNNs), which automatically extract pertinent features and have shown higher performance in segmentation tasks, are among the DL designs that researchers have increasingly embraced to tackle these difficulties [33][35]. The advantages of the shift from classical machine learning to deep learning in medical image analysis are discussed in this section.

2.2.1.1 Traditional Machine Learning Methods

In the initial days of brain tumour analysis using automated methods, classical machine learning (ML) algorithms, i.e., Support Vector Machines (SVMs), Decision Trees (DTs), and k-Nearest Neighbours (k-NN), were extensively used. These algorithms rely on human-engineered features extracted from MRI scan parameters, i.e., features like shape, intensity, and texture by manual extraction. These features are subsequently employed for tumour type classification or the detection of abnormal regions in cerebral images [34].

Despite their utility, Machine Learning approaches come with several limitations [34]:

- **Inadequacy in managing intricate, high-dimensional picture data:** medical imaging can have complex, high-dimensional image data which are hard to handle.
- **Manual feature engineering:** conventional segmentation techniques need manual feature engineering by hand, for example, pixel intensity, texture, geometry, and morpho-

logical features. It is a slow process and demands a lot of expert experience.

- **Poor generalisation:** imaging modality changes, tumor heterogeneity, and dataset variability substantially impact the performance of traditional segmentation methods.

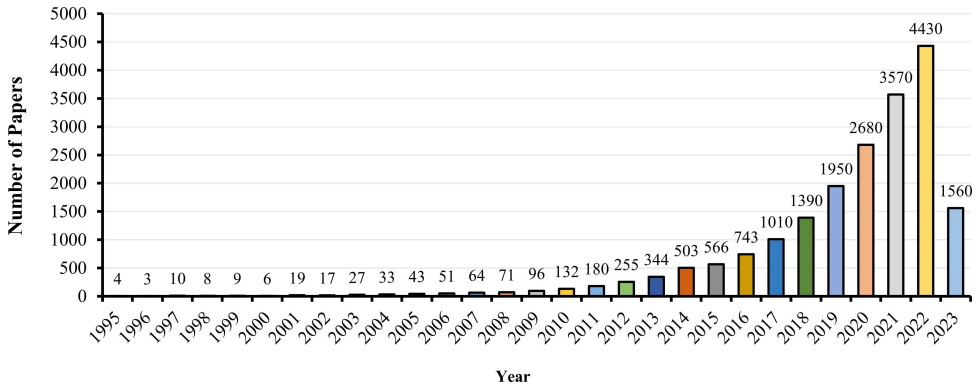

Figure 2.1: The number of published papers on brain tumor segmentation using artificial intelligence in the corresponding year.
[36]

2.2.1.2 Shift Towards Deep Learning

Deep learning algorithms can automatically extract features and patterns from unstructured data. Given these advantages, the discipline has switched heavily towards deep learning (DL) approaches, notably convolutional neural networks (CNNs). Unlike typical machine learning models, deep learning approaches employ raw picture data to learn meaningful features during training, eliminating the requirement for manual feature selection [34].

2.3 Image segmentation

Image segmentation is the technique of dividing or partitioning an image into parts, called segments. It is mostly useful for applications like image compression or object recognition, because for these types of applications, it is inefficient to process the whole image. So, image segmentation is used to segment the parts from image for further processing. There exist several image segmentation techniques, which partition the image into several parts based on certain image features like pixel intensity value, color, texture, etc [37].

2.4 Deep Learning for Medical Image Analysis

Deep Learning is crucial for enhanced medical image applications, given the effectiveness of AI in this field. Over the previous decade, Deep Learning-based research has significantly improved the accuracy of medical picture recognition. Using Deep Learning models on medical picture data enabled automated diagnosis of rare diseases, medical concerns, and early cancer detection. Deep Learning algorithms have proven effective in analysing medical images and improving diagnosis applications [33].

2.4.1 Advantages of Deep Learning in Brain Tumor Segmentation

- **Accuracy and Robustness:** Self-attention approaches have been proven to improve segmentation accuracy by focussing on specific cancer locations while simultaneously boosting overall accuracy [36].

  The multi-task learning model also shown significant improvements in brain tumour classification and segmentation, with a promising accuracy of 97% for each task [38].

  Deep learning algorithms were also successful in processing and objectively analysing massive amounts of MRI-based image data [34].

- **Automation of Complex Tasks:** Automated segmentation based on deep learning has demonstrated promising results, helping determine the tumor location, size, and shape using automated methods [36].

2.5 Deep Learning Architectures for Segmentation

Deep learning image segmentation architectures have experienced huge growth, and different well-known models and approaches have been proposed. The architectures embody the growth and variety of deep learning segmentation algorithms, with each having its own strengths and uses. We will review these architectures in this section.

2.5.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are the backbone of present-day image segmentation since they can learn spatial hierarchies and information at numerous scales. Various CNN-based architectures have been suggested over the years to enhance segmentation accuracy, computational time, and contextual awareness. As shown in Figure 2.2, the prevailing CNN architectures are as follows [39]:

- **Fully Convolutional Networks (FCNs):** The foundational model introduced dense, end-to-end pixel-wise prediction via convolutional layers.
- **HRNet:** Maintains high-resolution representations across the network, resulting in fine-grained segmentation.
- **DeepLabv3+:** Advanced architecture that combines atrous convolution and spatial pyramid pooling for effective multi-scale context acquisition.
- **PSPNet:** Pyramid pooling is used to incorporate global and local data, which works well for complex scene segmentation.
- **SegFormer:** A modern hybrid that combines transformer-based encoders and lightweight decoders to provide high performance without position encoding.

U-Net and YOLO have received special attention due to their versatility and domain-specific benefits. In the following sections, we will go over these two architectures in datail, emphasising their design ideas, strengths, and applications in medical segmentation tasks.

2.6 U-Net for Image Segmentation

U-Net is a model of convolutional neural network (CNN) for biomedical image segmentation.

Olaf Ronneberger et al. introduced the network for the first time in 2015, and it has been among the most well-liked deep learning architectures for medical image segmentation ever since. U-Net's main contribution is its ability to generate high-quality segmentations from very limited numbers of training images. This has particular application in medical contexts, as labeled data is often scarce and costly to acquire [41].

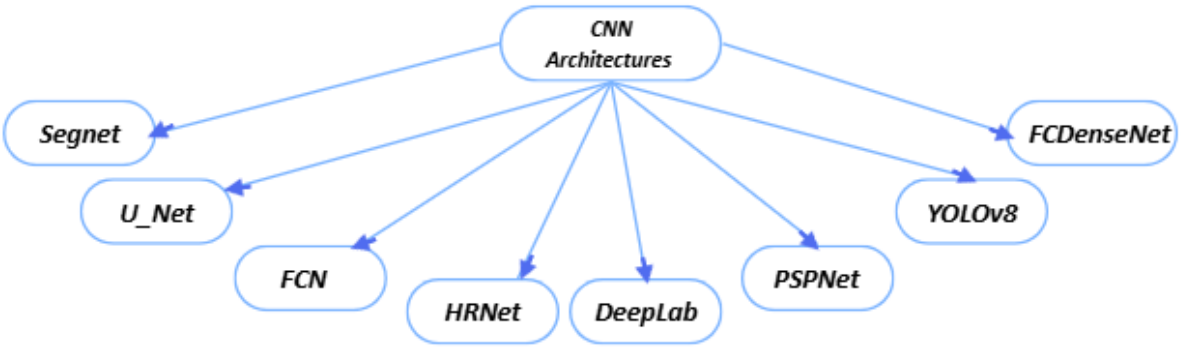

Figure 2.2: Deep Learning-based CNN Architecture.
[40]

2.6.1 U-net architecture

The popular design for challenges related to the segmentation of medical images, like brain tumor segmentation from MRI images, is the U-Net architecture. It's made to work with relatively few training samples and yields precise segmentations. Here is a detailed walkthrough of the U-Net model[41] , as shown in Figure 3.5.

1. **Input layer:** A single-channel MRI image is input into the network, typically preprocessed to the size of 572x572 pixels.

$$I_{\text{input}} \in R^{H \times W \times C} \tag{1}$$

   where H, W, and C stand for the input MRI images' height, width, and number of channels, respectively.

2. **Contracting path(Encoder):** This includes convolutions and maximum pooling can be represented as:

$$(I * K)_{i,j} = \sum_{m} \sum_{n} I_{m,n} K_{i-m,\, j-n} \tag{2}$$

   where:

   - I is the input image or feature map,

- K is the convolutional kernel (or filter),
- $*$ denotes the convolution operation.

Each convolution operation is typically followed by a nonlinear activation function, such as the Rectified Linear Unit (ReLU), and a downsampling step via a 2×2 max-pooling operation.

The encoder is composed of repeated applications of two unpadded 3×3 convolutional layers, each followed by ReLU activation. Downsampling is achieved by stride-2 convolutions or 2×2 max pooling.

At each downsampling step, the number of feature channels doubles, increasing from 64 in the initial layer to 1024 in the deepest layer.

3. **Bottleneck:** This part is directly after the last down sampling step and usually includes a ReLU function after two 3x3 convolutions.

4. **Expansive path (decoder):** The transposed convolution, or up-convolution operation, is defined as:

$$U(I,K)_{x,y} = \sum_i \sum_j I_{i,j} K_{x-i\cdot s, y-j\cdot s} \tag{3}$$

where:

- s is the stride of the transposed convolution,
- U denotes the up-convolution operation,
- K is the convolutional kernel.

In the expansive path, the upsampled feature map is concatenated with the corresponding feature map from the contracting path. Before concatenation, the encoder feature map is cropped as needed to match the spatial dimensions. The concatenation operation is denoted by:

$$C(F_{\text{contracting}}, F_{\text{up}}) \tag{4}$$

where:

- $F_{\text{contracting}}$ is the feature map from the encoder path,
- F_{up} is the upsampled feature map from the decoder path,
- $C(\cdot,\cdot)$ represents the concatenation operation.

After concatenation, the result is passed through two consecutive 3×3 unpadded convolutional layers, each followed by a Rectified Linear Unit (ReLU) activation function. This process refines the combined features and gradually reconstructs the spatial resolution. At each step in the decoder, the number of feature channels is halved, in contrast to the encoder where the number of channels is doubled.

5. **Final layer:** A 1x1 convolution is used to map each 64-component feature vector to the desired number of classes. In this case, it's mapping 2 channels for tumor segmentation Mathematically:

$$(I * K)_{i,j} = I_{i,j} \tag{5}$$

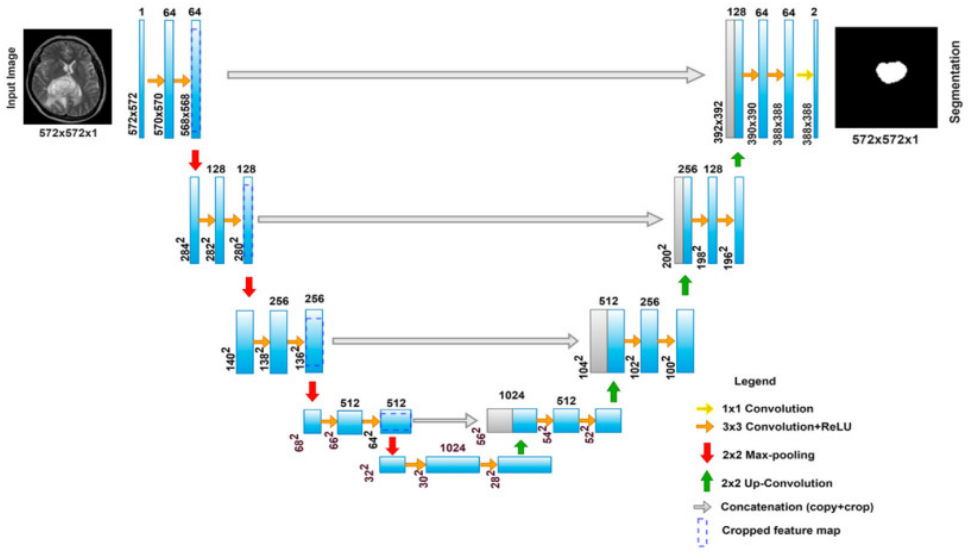

Figure 2.3: Architecture of a U-Net Model for brain tumor segmentation. [40]

2.6.2 Applications and Advantages

U-Net has emerged as the strongest framework for pixel-wise image segmentation tasks, particularly in applications requiring precise boundary delineation. The encoder-decoder architecture of the framework, with skip connections, enables concurrent extraction of contextual and spatial information and thus precise object shape and position reconstruction.

Key applications and advantages include [41]:

- **Medical Imaging:**
  - extensively employed for applications like MRI image segmentation of brain tumors.
  - Highly effective even with small training datasets, addressing the common challenge of limited annotated medical data.
  - Supports real-time segmentation in clinical contexts.
- **High Accuracy with Limited Data:**
  - Designed to perform well with relatively few training samples.
  - Perfect for fields where gathering data is costly or time-consuming.
- **Precise Boundary Preservation:**
  - The network's ability to preserve small spatial features through skip connections enhances segmentation accuracy, particularly in close proximity to object boundaries.
- **Panoptic Segmentation Capabilities:**
  - labels both the class and individual instances by combining semantic and instance segmentation (e.g., separating tumors from surrounding tissues).
- **Generalizability to Other Domains:**
  - Beyond medical imaging, U-Net has been successfully applied in agriculture (e.g., crop/plant segmentation), environmental monitoring, and satellite imagery analysis.

2.7 YOLO Architecture Evolution

Because of its single-pass processing and fast inference speeds, the You Only Look Once (YOLO) technique has become a cornerstone of real-time object detection. Since its inception in 2015, YOLO has gone through numerous major iterations, each with significant im-

provements in accuracy, speed, and functionality. More recent versions have added instance segmentation, broadening YOLO's application to complicated vision tasks like medical image analysis. Figure 2.4 visually summarizes the progression of YOLO versions from YOLOv1 to YOLOv8.

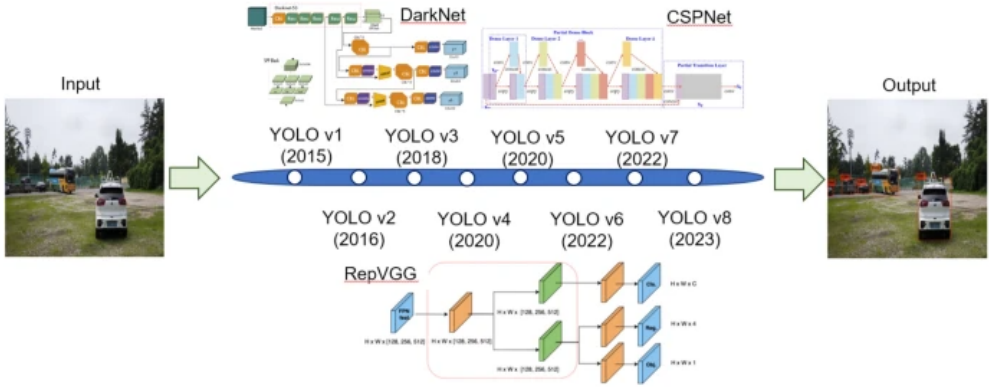

Figure 2.4: YOLO Evolution Timeline: Architectural Backbones and Version Progression (2015–2023)

[42]

2.7.1 Timeline of Key YOLO Versions and Innovations

Table 2.1 provides a summary of YOLO's architectural advancements and main functionality between versions [42].

2.7.2 YOLOv8 for Segmentation

YOLOv8 has been shown to be an effective object detection and instance segmentation tool with significant advancements in segmentation accuracy and speed. Perhaps the most striking aspect of this model is that it supports integrating segmentation heads, enabling it to output pixel-level masks along with bounding boxes. Due to its twin abilities, YOLOv8 is extremely beneficial in medical image processing applications where precise boundary delineation is vital, such as in brain tumor studies [43]. Its potential as an effective solution to real-time applications in clinical diagnosis, surgical assisting systems, and mobile health technologies is underscored by its ability to conduct detection and segmentation simultaneously and thereby enable quicker processing without any loss of accuracy [44].

Version	Year	Backbone	Key Features
YOLOv1	2015	Darknet	Grid-based detection, single-stage architecture, quick but less accurate. Introduced real-time detection using a single forward pass.
YOLOv2	2016	Darknet-19	Anchor boxes, batch normalization, residual learning, global average pooling, improved small object detection.
YOLOv3	2018	Darknet-53	Multi-scale prediction using Feature Pyramid Networks (FPN), ResNet-based skip connections, sigmoid output for multi-label classification.
YOLOv4	2020	CSPDarknet-53	PANet, Spatial Pyramid Pooling (SPP), large receptive field, quantization and advanced data augmentation.
YOLOv5	2021	CSPNet	Modular architecture (s, m, l, x), enhanced training pipeline, lightweight design, widely used in deployment.
YOLOv6	2022	RepVGG (EfficientRep)	EfficientRep backbone, Rep-PAN neck, quantization and distillation support, optimized for deployment.
YOLOv7	2022	E-ELAN	Bag-of-freebies training strategy, extended E-ELAN structure, real-time inference with improved accuracy.
YOLOv8	2023	Modified CSPDarknet	Anchor-free detection, segmentation support, unified head for detection/classification/segmentation.

Table 2.1: YOLO Version Evolution: Backbone Architectures and Key Features

2.7.3 Core Architecture of YOLOv8

YOLOv8 is a widely adopted model for real-time object detection and segmentation tasks. It introduces several improvements over its predecessors while maintaining high efficiency. The YOLOv8 model, as shown in Figure 2.5, is described in depth here [41]:

1. **Backbone (feature extraction):** Extracting features at various scales from the input image is the spine's responsibility. The basic structure of the backbone in YOLOv8s can be expressed as:

$$y = \text{ReLU}\left(\text{BatchNorm}\left(\text{Conv}(x)\right)\right) \tag{6}$$

where:

- Conv(x) is the convolution operation applied to the input x,
- BatchNorm denotes batch normalization,
- ReLU represents the rectified linear activation function.
  - It starts with the Stem Layer, typically the first few convolutional layers of the network that begin processing the raw input image.
  - A sequence of convolutional layers is used by subsequent layers, often called stages, to extract features at different degrees of abstraction.
  - The network can record more complex data as the depth (number of channels) increases as the layer becomes deeper, but the spatial resolution decreases.
  - The Spatial Pyramid Pooling-Fast (SPPF) layer aggregates features at multiple scales and maintains important spatial information, which is essential for detecting objects of various sizes.
  - Cross-Stage Partial (CSP) Layers aim to mitigate the vanishing gradient problem and reduce computational cost by splitting the feature map and merging it after processing.

2. **Neck (Feature aggregation and refinement):** The neck refines and aggregates features to prepare for the detection task. It utilizes a Top-Down and Bottom-Up approach, which involves combining employing conjunction and up-sampling, characterize maps of the spinal cord's various phases. This section ensures that the semantically more low-resolution features from the deeper layers are integrated with the higher-resolution features of the early layers. CSP Layers are used again to refine and mix these features further. The basic structure of the Neck in YOLOv8s can be expressed as:

$$y = \text{BiFPN}\left(\text{PANet}(x)\right) \tag{7}$$

where:

- x represents the input feature maps from the backbone,

- PANet(x) refers to path aggregation network, which optimizes feature pyramid networks for better feature fusion,
- BiFPN denotes a bidirectional feature pyramid network, which further refines and integrates features at multiple scales.

3. **Head(Detection):** The head module is responsible for predicting outcomes based on common traits. It contains Decoupled Heads, one for bounding box (BBox) prediction, another for class probability (Cls.), and a third for object-ness (Obj.), indicating whether the box contains an object. Each decoupled head uses a series of convolutional layers, ConvModules, that specialize in its respective task. Loss functions like IoU Loss for bounding box accuracy, BCE (Binary Cross Entropy) Loss for objectness, and nCE Loss for class predictions are applied to train the model effectively. The basic structure of the Head in YOLOv8s can be expressed as:

$$y = \text{Sigmoid}\left(\text{Convt}(f)\right) \tag{8}$$

where:

- f represents the input feature map from the neck,
- Conv indicates the convolutional layers applied to the feature map,
- The output is subjected to the sigmoid activation function to provide the final predictions.

2.7.4 Strengths and Practical Use Cases

The ability of YOLOv8 to address a number of crucial requirements in brain tumor segmentation justifies its use in this thesis. In particular, it is very appropriate for clinical applications due to its unified architecture, deployment effectiveness, and real-time processing capability. The reasoning is explained in detail in the following aspects:

- **Unified Detection and Segmentation Model:** Ultralytics released YOLOv8 in 2023, deploying it as a unified object detection, instance subdivision, and image classification model training framework through a recently released repository. This illustrates that

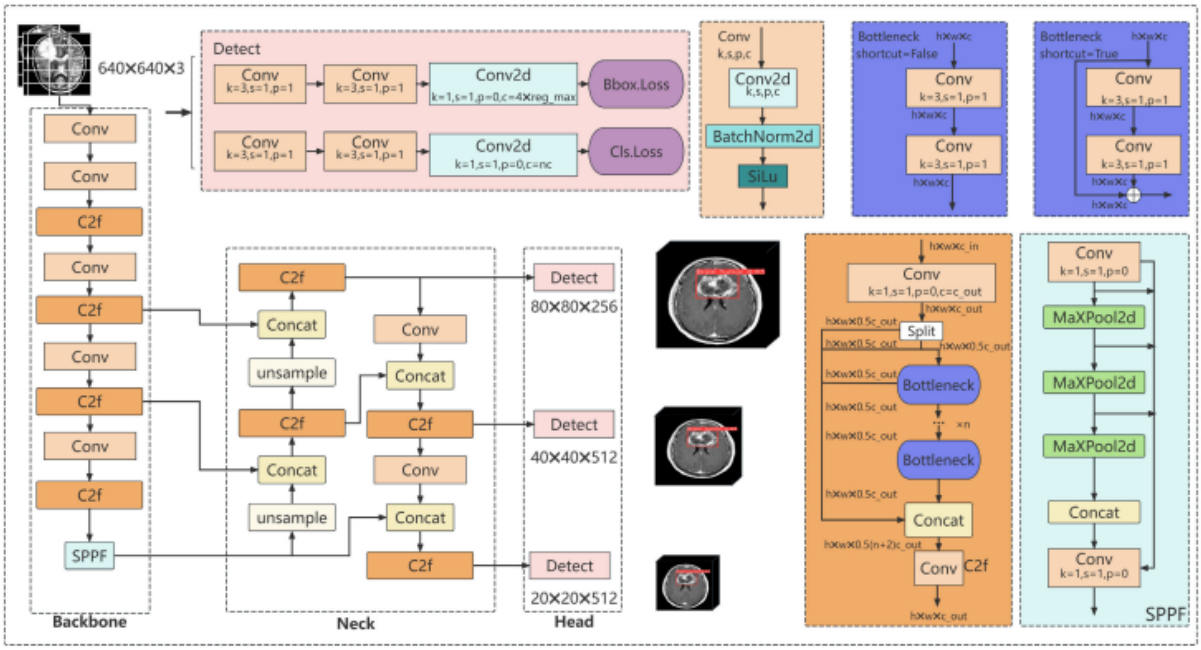

Figure 2.5: YOLOv8 network structure.
[43]

YOLOv8 is a multi-purpose architecture that can perform detection and segmentation [42].

- **Lightweight and Efficient:** YOLOv8 continues to be real-time efficient with a more streamlined network design and better detection performance. It is lower in parameters compared to the previous models and includes a modular design that consumes less computing power. It is therefore highly applicable for deployment in low-resource settings, including embedded systems, mobile health, and handheld diagnostic devices. Having the ability to leverage a strong segmentation model with low hardware requirements is significant in clinical settings where high-performance computing hardware might not be accessible, for instance, in rural clinics, point-of-care devices, or telemedicine systems [43].

- **Real-Time, High-Resolution Suitability:** High-resolution input real-time performance is one of the main requirements in clinical imaging settings, which YOLOv8 was engineered to achieve. YOLOv8's new segmentation head and anchor-free detection strategy allow for efficient and effective tumor detection without sacrificing performance or clinical usefulness. Recent research has confirmed its usefulness, showing that it is both faster and more accurate on a variety of medical datasets [44]

2.8 Comparative Analysis of Segmentation Techniques

Table 2.2 provides a comparative review of different segmentation approaches, emphasizing their key features, benefits, and drawbacks. With an emphasis on their applicability for brain tumor segmentation, this comparison highlights the transition from manual and traditional methods to contemporary deep learning techniques like U-Net and YOLOv8.

Segmentation Method	**Description**	**Strengths**	**Limitations**
Manual Segmentation	Carried out by hand by radiologists or medical professionals to draw the borders of tumors in MRI images.	High accuracy in skilled hands; No computational cost	Time-consuming; High variation between observers; Unscalable for large datasets
Classical Image Processing	Uses rule-based algorithms like region growing, edge detection, and thresholding.	Easy to implement; Low computational requirements	Noise-sensitive; Limited generalization; Poor performance on complex or overlapping structures
U-Net-Based Deep Learning	Uses encoder-decoder CNNs to learn hierarchical features for precise segmentation (e.g., U-Net, DeepLab).	High accuracy; Widely used in biomedical segmentation	Requires large annotated datasets; Slower inference than YOLO
YOLOv8-Based Segmentation	A unified, anchor-free model performing real-time detection and segmentation.	Real-time inference; Combines detection and segmentation; Suitable for deployment	Relatively new in biomedical imaging; Dependent on data quality and tuning

Table 2.2: Comparative analysis of segmentation techniques used in brain tumor imaging [2].

2.9 Performance Metrics-Based Comparison of Segmentation Methods

A comparison and summary of the significant measures, i.e., accuracy, speed, level of automation, training complexity, and real-time capability, are presented in Table 2.3 to provide a better insight into the relative performance of the segmentation methods.

2.10 Conclusion

This chapter offered an in-depth analysis of various deep learning algorithms used in medical imaging, particularly for brain tumor segmentation. We began by emphasising the critical importance of deep learning in improving the accuracy and automation of medical picture pro-

Segmentation Method	**Accuracy**	**Speed**	**Automation**	**Training Complexity**	**Real-Time Capable**
Manual Segmentation	Very High (expert-based)	Very Slow	None	N/A	No
Classical Image Processing	Low–Moderate	Fast	Partial	Low	Yes
U-Net-Based Deep Learning	High	Moderate	High	High	No
YOLOv8-Based Segmentation	High	Very Fast	High	Moderate	Yes

Table 2.3: Comparison of segmentation methods based on performance metrics.

cessing. Several benefits were highlighted, including increased robustness and performance in complicated segmentation tasks involving self-attention and multi-task learning methods. We also looked at major deep learning architectures including CNNs, U-Net, and attention-based models, each of which has its own set of advantages for tumor detection and classification in MRI scans.

In the following chapter, we will look at the proposed methodology, which incorporates various deep learning techniques into a comprehensive system for accurate and automated brain tumor segmentation. This will cover the dataset, preprocessing methods, model architecture, and performance evaluation strategies.

Chapter 3

Segmentation Models: Results and Comparison

3.1 Introduction

This chapter describes the pragmatic application of our brain tumor segmentation method employing deep learning tools. We concentrate on contrasting two models: YOLOv8, a real-time detection and segmentation framework, and U-Net, a semantic segmentation architecture designed for medical imaging. This chapter addresses the operational environment, dataset preparation, model development, training methodologies, evaluation measures, and comparative analysis.

3.2 Project Framework

The creation of an efficient system for brain tumor segmentation utilizing deep learning depends on a systematic framework designed to incorporate and evaluate two sophisticated neural network architectures. YOLOv8 and U-Net. This framework includes data collection, preprocessing, model design, training, evaluation, and comparison. Attempts to maintain scientific rigor and consistency throughout the experimentation phase. By applying both object detection and semantic segmentation algorithms to the same dataset, the system offers a full study of each model's performance in identifying tumor regions from brain MRI scans.

3.2.1 Objectives of the Study

The fundamental purpose of this study is to analyze and compare the effectiveness of two deep learning models, YOLOv8 and U-Net, in the task of segmenting brain tumors. YOLOv8 offers a rapid and efficient approach to object recognition and instance segmentation, which makes it suitable for real-time applications. In contrast, U-Net is specifically intended for biomedical image segmentation, giving fine-grained pixel-level accuracy. By assessing their performance on the same dataset, this work aims to discover which model gives the most accurate and clinically meaningful findings for the detection and delineation of brain tumors.

3.3 Dataset Preparation

This section covers the data preparation pipeline used prior to training the deep learning models. It describes the major phases of dataset gathering, cleaning, formatting, and transformation to provide consistent and trustworthy input for both YOLOv8 and U-Net. Special care was taken in the conversion of COCO-formatted annotations into binary masks for U-Net, as well as the arrangement of the data set into clearly structured subsets for training, validation, and testing. These preprocessing processes were needed to ensure an accurate and optimum segmentation performance across both systems.

3.3.1 Dataset Source

Roboflow, an online resource that offers annotated image datasets for computer vision tasks, provided the dataset utilized in this study. It can export in a variety of formats, including COCO JSON, and is especially made for object detection and segmentation issues.

The chosen dataset consists of MRI scans of the brain, each of which has been annotated to indicate the existence of brain tumors. Polygonal coordinates that define the borders of tumor regions are included in the annotations, which are supplied in the COCO (Common Objects in Context) format.

3.3.1.1 Dataset Characteristics

The main properties of the dataset are summarized in Table 3.1.

Property	Description
Source	Roboflow
Domain	Medical Imaging – Brain MRI
Annotation Format	COCO JSON (polygonal segmentation)
Classes	Tumor (binary segmentation: tumor vs. background)
Total Images	6,638
Image Format	JPEG (.jpg)
Color Mode	Grayscale (processed as single-channel input)

Table 3.1: Dataset Characteristics

3.3.2 Dataset Splitting

We manually divided the dataset into three separate subsets to make sure we were analyzing both models fairly and consistently:

- **Training set:** The model is trained to identify tumors from the available data using this subset. During this phase, the model acquires the ability to distinguish between healthy tissue and tumor regions by analyzing their respective characteristics.
- **Validation set:** After training, the validation set is used to fine-tune the model's hyper-parameters and measure its performance. It helps in choosing the most effective model configuration and optimizing its parameters without overfitting to the training data.
- **Test set:** The test set is used to objectively assess the final performance of the chosen model. This subset is maintained fully distinct during training and validation to ensure an unbiased evaluation of the model's capacity to detect tumors in previously unreported data.

Typically, the dataset is split as follows:

  - 70 % for the training set
  - 15 % for the validation set
  - 15 % for the test set

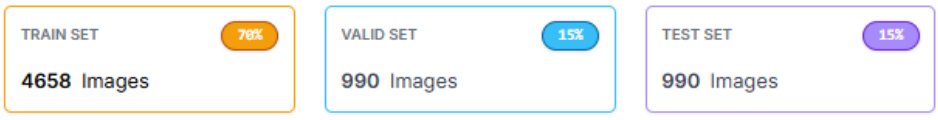

Figure 3.1: Training, Validation, and Test Data Distribution.

3.3.3 Mask Generation for U-Net

Unlike YOLOv8, which can use COCO-formatted annotations for object detection and segmentation, the U-Net model requires binary mask images for pixel-level training. To handle this, a special Python script was created to convert the polygon-based annotations in the COCO JSON file into equivalent binary masks. This script reads the segmentation coordinates and creates a binary mask with the tumor region filled in white on a black backdrop. The code for doing this conversion is provided below.

```
def convert_coco_to_masks(split):
    # Define the path to the COCO annotations JSON file for the given dataset split (train/valid/test)
    ann_path = f"/content/brain_tumor_data/{split}/_annotations.coco.json"
    # Define the directory containing the images for this split
    img_dir = f"/content/brain_tumor_data/{split}"
    # Define the output directory where the generated masks will be saved
    mask_dir = os.path.join(img_dir, "masks")
    # Create the masks directory if it doesn't already exist
    os.makedirs(mask_dir, exist_ok=True)
    # Load the COCO annotation data from the JSON file
    with open(ann_path) as f:
        coco_data = json.load(f)
    # Create a dictionary mapping image IDs to their metadata (file name, width, height, etc.)
    images_info = {img["id"]: img for img in coco_data["images"]}
    # Iterate through all annotations in the dataset
    for ann in tqdm(coco_data["annotations"], desc=f"Processing {split}"):
        image_id = ann["image_id"]                    # Get the image ID for this annotation
        segmentation = ann["segmentation"]            # Get the segmentation polygon coordinates
        image_info = images_info[image_id]            # Retrieve metadata for the corresponding image
        height = image_info["height"]
        width = image_info["width"]
        filename = image_info["file_name"]
```

```
        # Initialize a blank mask (same size as the image)
        mask = np.zeros((height, width), dtype=np.uint8)
        # Fill the mask using the polygon coordinates
        for seg in segmentation:
            pts = np.array(seg).reshape((-1, 2)).astype(np.int32)
            cv2.fillPoly(mask, [pts], color=255)        # Draw white
    (255) polygon on the black mask
        # Construct the output file name for the mask
        mask_name = os.path.splitext(filename)[0] + "_mask.png"
       # Save the mask image to the designated masks/ folder
        Image.fromarray(mask).save(os.path.join(mask_dir, mask_name))
```

3.4 Development Environment

The deep learning models were developed and implemented on a cloud-based architecture that provided great computational efficiency, collaboration, and easy access to tools and datasets. This environment consists of two parts: the hardware platform and the software tools utilised in the project.

3.4.1 Hardware Environment

Both models were trained and tested entirely in the cloud using Google Colab Pro, which provided dedicated GPU resources and facilitated the deployment of deep learning workflows.

- **Platform:** Google Colab Pro
- **Accelerator:** NVIDIA Tesla T4
- **RAM:** Between 13 GB and 25 GB, depending on the session configuration
- **Storage:** Dataset and model files were stored in Google Drive. The drive was mounted in the Colab environment to allow direct access for reading and writing data during training and testing.

3.4.2 Software Environment

To support the entire machine learning workflow, from data collection to model evaluation, a mix of cloud platforms, tools, and Python modules were used.

1. **Google Colab :** Google Colaboratory, often known as Colab, is a cloud-based platform built by Google to promote learning and research in artificial intelligence, particularly machine learning. Colab, built on the Jupyter Notebook technology, enables users to write, execute, and share Python code in an interactive environment similar to Google Docs. One of its main advantages is that it comes preconfigured with major AI libraries like TensorFlow, Keras, and Matplotlib, which reduces experiment setup time significantly. Furthermore, Colab provides free access to GPU accelerators via Google's cloud infrastructure, making it a great tool for students, academics, and developers to prototype and train deep learning models without requiring specific hardware [1].

2. **Google Drive :** Google Drive is a cloud storage service provided by Google that allows users to securely store and access files online. It works seamlessly with Google Colab and offers a dependable solution for storing datasets, preserving trained models, and loading resources during training and evaluation. In this research, Google Drive was used to store the raw MRI images, annotation files, U-Net generated binary masks, and model checkpoints. The integration with Colab enabled real-time read/write access without the need for manual uploads or downloads, which considerably simplified the data flow. [2].

3. **Roboflow :** Roboflow is a web-based platform that simplifies the process of creating, managing, and delivering datasets for computer vision applications. It has a user-friendly interface that allows users to input images, annotate them with bounding boxes or segmentation masks, and export the dataset in a variety of formats such as COCO, YOLO, and Pascal VOC. Roboflow also includes preprocessing techniques like image scaling, augmentation, and format conversion, which are critical for maintaining dataset quality and consistency. In this research, Roboflow was used to annotate brain MRI images and export them in COCO format, which is compatible with the YOLOv8 and U-Net pipelines. Its ability to handle versioning and dataset organization made it an indispens-

[1] Accessed on May 9, 2025, from the official website:https://colab.research.google.com
[2] Accessed on May 9, 2025, from the official website:https://drive.google.com

able tool during the dataset preparation phase. [3].

3.4.2.1 Programming Language (Python)

Python was chosen as the primary programming language for this project due to its open-source nature, user-friendly syntax, and wide use in scientific computing, artificial intelligence, and machine learning. Its high-level structure and diverse library ecosystem make it an excellent tool for academics with moderate programming experience.

Python offers a flexible and modular framework for data collection, processing, model construction, and visualisation. Coding in this study was done mostly in Jupyter Notebook environments using Google Colab, which allowed for interactive experimentation, real-time debugging, and result visualisation. Python's flexibility enabled smooth integration with TensorFlow and PyTorch, two of the most popular deep learning frameworks, allowing us to quickly construct and analyse YOLOv8 and U-Net models [45].

3.4.2.2 Python Libraries

A number of Python libraries were essential to the execution and implementation of our deep learning workflow.

- **TensorFlow/Keras:** This framework was used to create and train the U-Net model for semantic segmentation. TensorFlow delivers robust GPU acceleration, while Keras provides a high-level API for model creation, training, and evaluation.
- **scikit-learn:** A machine learning library for Python that offers easy-to-use and efficient tools for data analysis along with model evaluation, such as metrics including accuracy, precision, recall, and confusion matrices.
- **NumPy:** NumPy, an essential library for numerical computations, was employed for array manipulation, executing mathematical operations, and organizing pixel-wise image data utilized in training and evaluation.
- **Matplotlib:** A very popular Python data visualization library. It provides tools to create static, animated, and interactive plots, e.g., loss curves and segmentation results, that are often used to monitor and visualize model performance in deep learning workflows.

[3] Accessed on May 9, 2025, from the official website:https://roboflow.com

- **Ultralytics:** A Python library enabling simple access to the YOLOv5 and YOLOv8 models for object detection and segmentation with better performance and simplified training procedures.
- **Roboflow:** A data preparation, augmentation, and export platform and Python SDK widely utilized to automate the workflow of computer vision projects such as training YOLO models.
- **IPython.display:** A component of the IPython library that enables direct rendering of dynamic content like images, videos, and formatted outputs in Colab or Jupyter notebooks.

3.5 Implementation Process

This section details the implementation processes carried out for each of the two deep learning models: YOLOv8 and U-Net. For clarity and consistency, each model was constructed and tested using a standardized workflow consisting of four primary phases: preprocessing, model design, training, and evaluation.

3.5.1 YOLOv8 Implementation

The YOLOv8 model was used for instance segmentation, exploiting its capability to detect and delineate tumor regions using bounding boxes and masks. The implementation followed these phases:

3.5.1.1 Preprocessing

In this phase, the dataset was downloaded and prepared directly via the Roboflow Python API, which allowed for a seamless download and formatting of images and annotations suitable for YOLOv8 segmentation.

The following code snippet was used to install Roboflow and fetch the dataset:

```
!pip install roboflow
from roboflow import Roboflow
rf = Roboflow(api_key="Fmcw01cZSQfM0SpRXYHI")
```

```
project = rf.workspace("douaa-ywhtp").project("brain-tumor-segmentation
    -dwapu-qrwyh")
version = project.version(1)
dataset = version.download("yolov8")
```

This command automatically

- Downloads the dataset in **YOLOv8 format,**
- Unzips and organizes it into the correct folder structure.
- Includes image files and corresponding polygon annotations in .txt format.

3.5.1.2 Model Configuration

Once the dataset was downloaded and organized by Roboflow, the following step was to configure the YOLOv8 model for training. This was achieved by producing a YAML configuration file called data.yaml, which defines the dataset structure and important metadata.

Below is the actual **data.yaml** content used in this project:

```
names:
- tumor
nc: 1

roboflow:
  license: CC BY 4.0
  project: brain-tumor-segmentation-dwapu-qrwyh
  url: https://universe.roboflow.com/douaa-ywhtp/brain-tumor-
    segmentation-dwapu-qrwyh/dataset/1
  version: 1
  workspace: douaa-ywhtp

train: /content/brain-tumor-segmentation-1/train/images
val: /content/brain-tumor-segmentation-1/valid/images
test: /content/brain-tumor-segmentation-1/test/images
```

Explanation of Key Fields

- **names:** Specifies the name(s) of the target class. In this case, the only class is `"tumor"`.

- **nc:** Defines the number of classes. For this project, the value is set to 1, indicating binary segmentation.
- **train, val, test:** These fields specify the paths to the image directories used for training, validation, and testing, respectively.
- **roboflow:** Contains metadata about the dataset version, license type, project name, and Roboflow workspace. This block is automatically added when downloading the dataset using the Roboflow API.

3.5.1.3 Training

The YOLOv8 segmentation model was then trained using Ultralytics' CLI interface with the dataset and config file. In this paper, we implement the lightweight model yolov8n-seg.pt (nano) since it is efficient to train even on limited resources, yet still capable of supporting segmentation within datasets.

The training started with the following script:

```
!yolo task=segment mode=train data={dataset.location}/data.yaml model="
    yolov8n-seg.pt" epochs=50 imgsz=640
```

Breakdown of the script

- **task=segment:** Specifies that the model should perform instance segmentation.
- **mode=train:** Sets the operation mode to training.
- **data={dataset.location}/data.yaml:** Points to the dataset configuration file downloaded via Roboflow.
- **model="yolov8n-seg.pt":** Chooses the YOLOv8 "nano" segmentation variant as the base model.
- **epochs=50:** Defines the number of training epochs.
- **imgsz=640:** Sets the input image size for the model (images are automatically resized to 640×640).

During training, the Ultralytics YOLOv8 framework accomplished several critical activities automatically. Data augmentation techniques, including random flipping, scaling, and rotation, were utilized to increase the model's applicability and resilience by boosting dataset variability. The batch size was constantly adjusted based on the available GPU memory, ensuring optimum exploitation of resources. In addition, the platform enabled real-time tracking of critical training and validation parameters such as training loss, validation loss, segmentation mask loss, bounding box loss, and mean Average Precision (mAP) for both bounding boxes and masks.

3.5.1.4 Evaluation

The YOLOv8 model was evaluated after 50 training epochs. The purpose was to evaluate the model's capacity to detect and segment tumors in MRI data. To provide an objective evaluation, the previously set test dataset was used.

The model's effectiveness was assessed using widely accepted classification and segmentation performance indicators such as Precision, Accuracy, Recall, and F1 Score. These indicators give an entire overview of how well the model distinguishes between tumor and non-tumor regions in medical imaging.

To assess the performance of a segmentation model, we should use the following four categories:

- **True Positive (TP):** Tumor regions correctly predicted as tumors.
- **True Negative (TN):** Non-tumor areas correctly identified as non-tumor.
- **False Positive (FP):** Non-tumor areas incorrectly predicted as tumor.
- **False Negative (FN):** Tumor regions incorrectly labeled as non-tumor.

These values constitute the basis for computing the evaluation metrics:

- **Precision**

  Precision is the proportion of accurately predicted positive observations among all observations projected as positive. It demonstrates the model's capacity to reduce false positive classifications. A higher precision suggests that the model's prediction of a tumor is more

likely to be accurate.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{3.1}$$

- **Accuracy**

  Accuracy measures the categorization model's overall ability in properly predicting both positive and negative outcomes. It provides a global assessment of model performance by taking into consideration both true positives and true negatives.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{3.2}$$

- **Recall (Sensitivity)**

  Recall measures the model's ability to correctly identify all relevant positive cases. In the context of tumor segmentation, it determines how many real tumor areas are successfully identified. A high recall means that the model misses very few tumors, reducing false negatives.

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3.3}$$

- **F1 Score**

  The F1 Score represents the harmonic mean of precision and recall. It gives a balanced measure that takes into account both false positives and false negatives. Precision relies on the accuracy of positive predictions, while recall stresses completeness. The F1 score is very beneficial for achieving a balance between precision and recall.

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.4}$$

3.5.1.5 Results and Discussion

The YOLOv8 segmentation model was trained and evaluated over 50 epochs, with separate training and validation subsets. This split ensured an unbiased evaluation of the model's generalization ability. The training process was thoroughly monitored using accuracy and loss metrics, with evaluation carried out on a separate test set.

Training Curves Analysis

Figure 3.2 shows the progress of the YOLOv8 model's training and validation metrics, which include box loss, segmentation loss, classification loss, distributional focal loss (dfl), precision, recall, and mAP. These curves provide information about the model's learning behavior over 50 epochs. During the first 10 epochs, both training and validation metrics improved dramatically, showing that the model adapted quickly to dataset patterns and extracted features efficiently. This time is distinguished by a sharp decrease in loss and a significant rise in precision and recall, indicating the model's early learning phase.

From epoch 10 to 50, the curves gradually improve with modest movements, indicating that the model's parameters were fine-tuned on more complicated and confusing data. Loss values gradually decreased, but precision and recall metrics stabilized, indicating that the model had reached convergence.

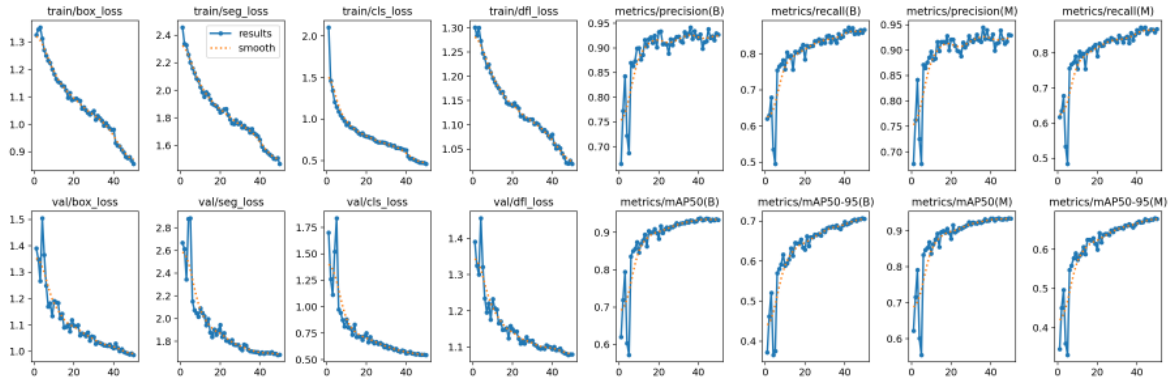

Figure 3.2: Training and validation loss and metric curves for YOLOv8 over 50 epochs.

Table 3.2 presents the evaluation metrics obtained from the YOLOv8 model for both bounding box detection and segmentation mask tasks.

Metric	Bounding Box	Segmentation Mask
Precision (P)	0.929	0.930
Recall (R)	0.857	0.859
mAP@0.50	0.935	0.935
mAP@0.50:0.95	0.708	0.679

Table 3.2: YOLOv8 Performance Summary

The YOLOv8 model consistently displayed a high degree of performance across object detection and segmentation tasks. With precision and recall values exceeding 85% in both tasks,

the model effectively detected and localized tumor regions with high consistency. Bounding box-based evaluation slightly outperformed mask segmentation, especially in the more stringent mAP@0.50:0.95 score, which indicates the model's capacity to properly fit bounding regions around tumors. In spite of this, the metrics associated with the segmentation mask were robust, verifying that YOLOv8 was capable of learning the spatial structure and boundary of tumor regions. Such a balance between the accuracy of detection and segmentation quality demonstrates the potential of YOLOv8 for real-time medical imaging tasks where both speed and accuracy are crucial.

3.5.2 U-Net Implementation

U-Net was utilized as a pixel-level segmentation model for the detection of brain tumors in MRI images. Its encoder-decoder architecture makes it highly proficient in the task of medical image segmentation. The preprocessing, model design, training, and evaluation procedures followed in this implementation are given in the following steps.

3.5.2.1 Preprocessing

This step utilized the dataset developed out of Roboflow and converted to binary masks for training the U-Net model. Contrary to YOLOv8, whose functionality is dependent on polygon-based COCO annotations, U-Net needs pixel-wise binary masks that distinctly define tumor areas from the remainder of the background.

Every split of the dataset (training, validation, and testing) contained :

- **Input images** in grayscale .jpg format.
- **Binary masks** in .png format located in a subfolder named masks/.

All images and masks have been resized to a uniform resolution of 128×128 pixels for compatibility with the model input layer and for memory efficiency during training. Pixel values have been normalized to [0, 1]. This function processes images and masks through resizing, normalizing, and pairing:

```
def load_dataset(split):
    image_dir = os.path.join(BASE_PATH, split)  # Path to images (e.g.,
    train/, valid/, test/)
```

```
    mask_dir = os.path.join(image_dir, 'masks')  # Path to
corresponding masks folder
    X, Y = [], []  # Lists to store images and masks
    for filename in sorted(os.listdir(image_dir)):
        # Select only image files (skip mask filenames found in same
dir)
        if filename.endswith('.jpg') or filename.endswith('.png'):
            if 'mask' in filename:
                continue
            # Build full paths to image and corresponding mask
            img_path = os.path.join(image_dir, filename)
            mask_path = os.path.join(mask_dir, os.path.splitext(
filename)[0] + '_mask.png')
            # Load grayscale image and mask
            image = cv2.imread(img_path, cv2.IMREAD_GRAYSCALE)
            mask = cv2.imread(mask_path, cv2.IMREAD_GRAYSCALE)
            # Resize and normalize
            image = cv2.resize(image, (128, 128)) / 255.0
            mask = cv2.resize(mask, (128, 128)) / 255.0
            # Append to the lists
            X.append(image)
            Y.append(mask)
    # Convert lists to NumPy arrays and reshape for model input
    X = np.array(X).reshape(-1, 128, 128, 1)
    Y = np.array(Y).reshape(-1, 128, 128, 1)
    return X, Y
```

3.5.2.2 Model Architecture

The U-Net architecture implemented in this project follows the classical encoder-decoder structure, most appropriate for biomedical image segmentation. It consists of a contracting path (encoder) for capturing context and a symmetric expanding path (decoder) enabling precise localization.

The model has been built using **TensorFlow/Keras** and follows the classical U-Net design. It includes:

- **Three downsampling blocks**, each composed of two convolutional layers with ReLU

activation, followed by max pooling to reduce spatial dimensions.

- **Bottleneck layer**, which serves as a bridge between the encoder and decoder paths and increases feature abstraction.
- **Three upsampling blocks**, where each block uses transposed convolution (or upsampling), followed by two convolutional layers. Each block also includes skip connections that concatenate features from the corresponding encoder block.
- **Final layer**: a single `Conv2D` layer with sigmoid activation to generate a binary segmentation mask matching the input dimensions.

The following code defines the U-Net architecture:

```
def unet_model(input_size=(128, 128, 1)):
    inputs = tf.keras.Input(input_size)
    # Encoder
    c1 = Conv2D(32, 3, activation='relu', padding='same')(inputs)
    c1 = Conv2D(32, 3, activation='relu', padding='same')(c1)
    p1 = MaxPooling2D()(c1)
    c2 = Conv2D(64, 3, activation='relu', padding='same')(p1)
    c2 = Conv2D(64, 3, activation='relu', padding='same')(c2)
    p2 = MaxPooling2D()(c2)
    # Bottleneck
    c3 = Conv2D(128, 3, activation='relu', padding='same')(p2)
    c3 = Conv2D(128, 3, activation='relu', padding='same')(c3)
    # Decoder
    u4 = UpSampling2D()(c3)
    u4 = concatenate([u4, c2])
    c4 = Conv2D(64, 3, activation='relu', padding='same')(u4)
    c4 = Conv2D(64, 3, activation='relu', padding='same')(c4)
    u5 = UpSampling2D()(c4)
    u5 = concatenate([u5, c1])
    c5 = Conv2D(32, 3, activation='relu', padding='same')(u5)
    c5 = Conv2D(32, 3, activation='relu', padding='same')(c5)
    outputs = Conv2D(1, 1, activation='sigmoid')(c5)
    model = tf.keras.Model(inputs=[inputs], outputs=[outputs])
    return model
```

3.5.2.3 Training Setup

Once the model was defined, it was then compiled and trained with the Adam optimizer, as well as a custom loss function combining Binary Cross-Entropy (BCE) and Dice loss. This combination enhances segmentation performance by taking into account both pixel-wise classification accuracy and spatial alignment between predicted and ground-truth masks.

The training was conducted using the following configuration:

- **Epochs:** 100
- **Batch size:** 8
- **Input size:** 128×128 grayscale images
- **Loss function:** Binary Cross-Entropy (BCE) + Dice loss
- **Optimizer:** Adam
- **Metrics:** Accuracy

The training code is implemented as follows:

```
# Define Dice Loss
def dice_loss(y_true, y_pred):
    smooth = 1e-6
    y_true_f = tf.keras.backend.flatten(y_true)
    y_pred_f = tf.keras.backend.flatten(y_pred)
    intersection = tf.reduce_sum(y_true_f * y_pred_f)
    return 1 - (2. * intersection + smooth) / (tf.reduce_sum(y_true_f)
   + tf.reduce_sum(y_pred_f) + smooth)
# Combine BCE + Dice
def bce_dice_loss(y_true, y_pred):
    bce = tf.keras.losses.binary_crossentropy(y_true, y_pred)
    return bce + dice_loss(y_true, y_pred)
# Compile and train
model = unet_model()
model.compile(optimizer='adam', loss=bce_dice_loss, metrics=['accuracy'
   ])

checkpoint = ModelCheckpoint(filepath='best_model.h5', monitor='
```

```
    val_loss',
                                save_best_only=True, mode='min', verbose
    =1)

history = model.fit(X_train, Y_train,
                    validation_data=(X_val, Y_val),
                    epochs=100,
                    batch_size=8,
                    callbacks=[checkpoint])
```

During training, both training and validation accuracy and loss were monitored to evaluate model convergence and generalization.

3.5.2.4 Evaluation and Discussion

In quantifying the performance of the U-Net architecture in segmenting brain tumors, two key metrics were prioritized: Intersection over Union (IoU) and Dice coefficient. These metrics are significant in medical image segmentation as they quantify spatial agreement between predicted tumor areas and their corresponding ground truth annotations.

- **IoU (Intersection over Union):** calculates the intersection between the predicted and actual mask, which measures how accurately the model has localized the tumor regions.
- **Dice Coefficient:** measures the similarity between true and predicted masks, placing greater emphasis on corresponding regions and providing a robust metric for medical segmentation quality.

After training for 100 epochs, the model performed as follows on the test set:

- **IoU:** 0.7019
- **Dice:** 0.8248
- **Precision:** 0.8712
- **Recall:** 0.7832
- **F1 Score:** 0.8248

The values exhibited demonstrate a high precision in segmentation, which means the model has

learned effectively to localize and segment tumor regions and achieve a good balance between recall and precision. The striking feature is that the Dice coefficient exhibits a high spatial agreement between predicted and ground-truth tumor borders, a feature that is clinically very important.

Training Performance

In order to better understand the model's learning behavior, the training and validation performance curves were plotted across 100 epochs. Figure 3.3 shows the progression of loss and accuracy during training, making it easier to determine convergence and generalization.

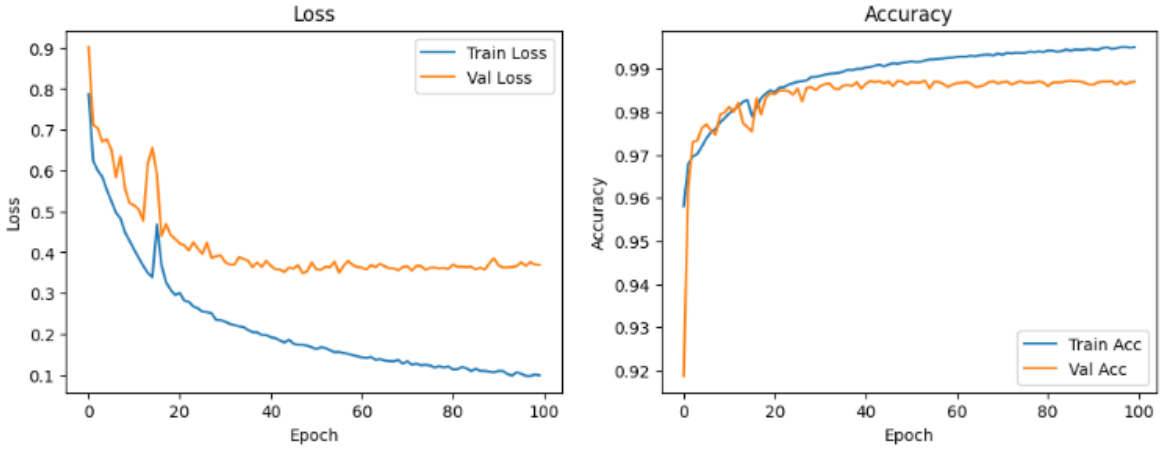

Figure 3.3: Training and validation loss and accuracy curves over 100 epochs.

- The loss curves follow a steep downward trajectory, particularly the training loss, which is a sign that the model kept enhancing its match with the data. Validation loss converges to a constant value after epoch 30, demonstrating convergence and an insignificant degree of overfitting.
- The accuracy plots demonstrate rapid learning during the first few epochs, with training accuracy above 99

3.6 Comparison and Evaluation

Having trained YOLOv8 and U-Net models, a comparative analysis was made on the identical dataset and metrics. This section gives a detailed comparison of their performance with a focus both on quantitative outcomes and qualitative results.

3.6.1 Quantitative Results

The table 3.3 above illustrates the best performance metrics for both models (YOLOv8 Mask and U-Net):

Metric	U-Net	YOLOv8 (Mask)	Analysis
Precision	0.8712	**0.930**	YOLOv8 is more confident (fewer false positives)
Recall	0.7832	**0.859**	YOLOv8 detects more tumor regions
IoU	**0.7019**	0.679 (approx.)	U-Net has better overlap with true mask
Dice	**0.8248**	0.820 (est.)	U-Net slightly better in boundary accuracy
F1 Score	0.8248	**0.893**	YOLOv8 has a better balance of precision and recall

Table 3.3: Quantitative performance comparison between YOLOv8 Mask and U-Net

The findings indicate that YOLOv8 Mask performs better than U-Net in terms of Precision, Recall, and F1 Score, therefore illustrating its enhanced capacity for detection and prediction of tumor areas. On the other hand, U-Net records better IoU and Dice values, indicating its performance in generating more accurate and precise mask boundaries, a key attribute for application in medical imaging.

3.6.2 Qualitative Comparison

Visual comparisons were performed using test images for each model, displayed below:

- **YOLOv8 Predictions:** Figure 3.4 illustrates the tumor regions detected and segmented by YOLOv8.
- **U-Net Predictions:** Figure 3.5 illustrates the tumor regions segmented by U-Net with precise boundaries.

3.6.3 Discussion

A comparison was made on the performance of the top-performing segmentation models, U-Net and YOLOv8 (Mask). The test results show that both models demonstrate very good performance under different conditions.

YOLOv8 (Mask) demonstrated better recall and precision scores, subsequently proving its effectiveness for immediate tumor detection with high accuracy. The high F1 scores confirm

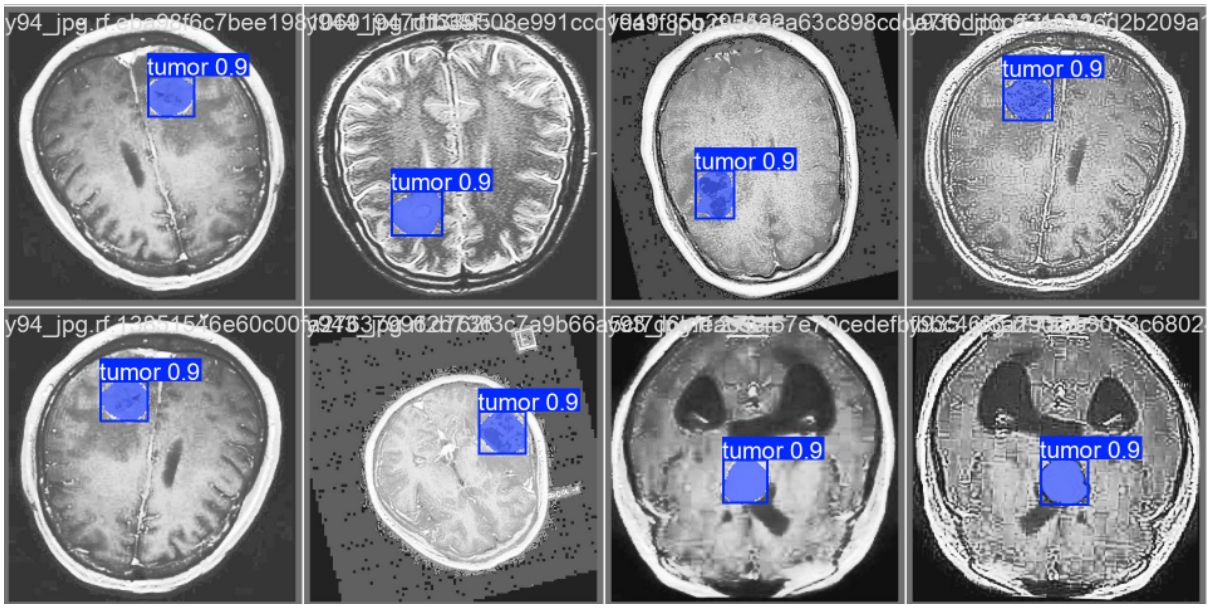

Figure 3.4: Tumor regions detected and segmented by YOLOv8.

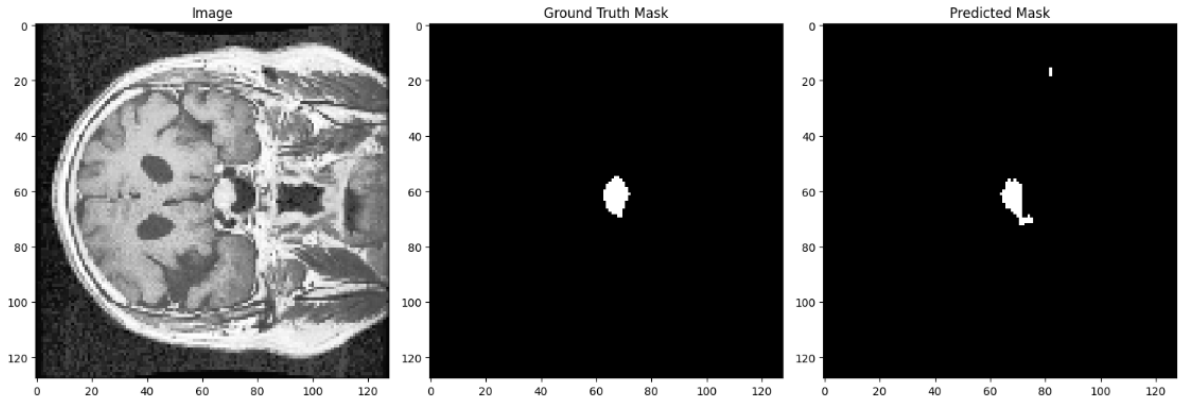

Figure 3.5: Tumor regions segmented by YOLOv8.

that the model is able to balance tumor region detection (recall) with minimized errors (precision). YOLOv8, therefore, is specifically suitable for applications that require increased speed and real-time processing, for instance, in automated screening systems.

U-Net, on the other hand, excelled in IoU and Dice coefficient metrics, highlighting its strength in producing accurate and smooth segmentation boundaries. This characteristic is particularly beneficial for medical applications, where precise delineation of tumor regions is essential. U-Net's architecture, which focuses on pixel-level accuracy, ensures high-quality segmentation that can support detailed diagnostic analysis.

3.7 Conclusion

In this chapter, we have expressed the development, implementation, and evaluation of two advanced deep learning models for the segmentation of brain tumors: U-Net and YOLOv8. The chapter began with a detailed study of the dataset preparation along with the methodical training processes adopted for the two models. Our comparative analysis, encompassing both quantitative measures and qualitative visual results, highlighted the strengths and weaknesses of each approach. Specifically, YOLOv8 was seen to perform better in terms of precision, recall, and F1-score, thereby making it extremely well-suited for rapid, real-time segmentation. On the other hand, U-Net was seen to provide better results in terms of IoU and Dice scores, highlighting its capability to provide highly accurate and smooth segmentation boundaries, a requirement of medical imaging applications. The choice between the two models would finally depend upon the particular requirements of the application—whether the application necessitates convenient segmentation or high spatial accuracy.

General Conclusion

This thesis investigated the use of deep learning techniques for the segmentation of brain tumors based on MRI imaging. Concretely, two state-of-the-art models were compared and implemented: U-Net, with its high accuracy in pixel-level segmentation, and YOLOv8, with its ability for rapid and precise object detection. The results of our investigation indicated that every model possesses unique advantages: U-Net attained enhanced efficacy concerning Intersection over Union (IoU) and the Dice coefficient, which are vital for accurate boundary delineation in medical imaging contexts. In contrast, YOLOv8 exhibited elevated precision and recall metrics, thereby rendering it more appropriate for swift segmentation activities.

Although both models exhibited high efficacy, the research pointed out the intrinsic trade-offs between processing speed and segmentation accuracy. YOLOv8's fast computing capacity combined with high detection accuracy makes it preferable for real-time applications, but high spatial fidelity needed for clinical diagnosis makes U-Net's pixel segmentation more appropriate for this purpose.

Also, this study sheds light on the possibility of hybrid models that leverage the best of YOLOv8 and U-Net. The Enhanced TumorNet model, as a hybrid method that takes advantage of YOLOv8 for fast detection and U-Net for precise segmentation, was identified as an engaging direction. This model optimizes the segmentation task by initially speeding up the detection of tumor regions with YOLOv8 and then improving these detections with U-Net, thereby achieving a trade-off between effectiveness and accuracy.

Future work can be guided towards further fine-tuning this hybrid model, e.g., by integrating transfer learning methods, augmenting training data with synthetic images, and testing the model in real clinical environments. Another possible direction of exploration is multi-modal data fusion, e.g., fusing MRI with other imaging modalities (e.g., CT) for improved tumor de-

tection and segmentation performance. Furthermore, real-time application of the hybrid model to a clinical environment would provide additional confirmation of its real-world relevance, that it is healthcare provider-friendly.

This thesis has made immense contributions to the area of medical image analysis through an extensive evaluation of two state-of-the-art deep learning models specifically designed for brain tumor segmentation, and by suggesting a hybrid method to enhance performance.

References

[1] M. C. Support, "Types of brain tumour - macmillan cancer support," 2024, accessed: March 20, 2025. [Online]. Available: https://www.macmillan.org.uk/cancer-information-and-support/brain-tumour/types-of-brain-tumour

[2] M. Y. Shams, W. M. Elmessery, A. A. T. Oraiath, A. Elbeltagi, A. Salem, P. Kumar, T. M. El-Messery, T. A. El-Hafeez, M. F. Abdelshafie, G. G. A. El-Wahhab, I. S. El-Soaly, and A. E. Elwakeel, "Automated on-site broiler live weight estimation through yolo-based segmentation," *Smart Agricultural Technology*, vol. 10, p. 100828, 2025. [Online]. Available: https://doi.org/10.1016/j.atech.2025.100828

[3] P. T. Schoenemann, "Evolution of the size and functional areas of the human brain," *Annual Review of Anthropology*, vol. 35, pp. 379–406, 2006.

[4] M. A. Hofman, "Evolution of the human brain: when bigger is better," *Frontiers in Neuroanatomy*, vol. 8, p. 15, 2014.

[5] J. Stiles and T. L. Jernigan, "The basics of brain development," *Neuropsychology Review*, vol. 20, pp. 327–348, 2010.

[6] D. S. Bassett and M. S. Gazzaniga, "Understanding complexity in the human brain," *Trends in Cognitive Sciences*, vol. 15, no. 5, pp. 200–209, 2011.

[7] J. H. Medicine, "Anatomy of the brain," 2024, accessed: March 18, 2025. [Online]. Available: https://www.hopkinsmedicine.org/health/conditions-and-diseases/anatomy-of-the-brain

[8] L. G. Ungerleider and J. V. Haxby, "'what' and 'where' in the human brain," *Current Opinion in Neurobiology*, vol. 4, no. 2, pp. 157–165, 1994.

[9] Q. B. Institute, "Brain anatomy and function," 2024, accessed: March 18, 2025. [Online]. Available: https://www.health.qld.gov.au/abios/asp/brain

[10] W. L. Nowinski, "Introduction to brain anatomy," in *Biomechanics of the Brain*, K. Miller, Ed. Springer, 2011, pp. 5–40. [Online]. Available: http://www.springer.com/978-1-4419-9996-2

[11] K. Herholz, K.-J. Langen, C. Schiepers, and J. M. Mountz, "Brain tumors," *Seminars in Nuclear Medicine*, vol. 42, no. 6, pp. 356–370, 2012.

[12] L. M. DeAngelis, "Brain tumors," *New England Journal of Medicine*, vol. 344, no. 2, pp. 114–123, 2001.

[13] M. L. Bondy, M. E. Scheurer, B. Malmer, J. S. Barnholtz-Sloan, F. G. Davis, D. Il'yasova, C. Kruchko, B. J. McCarthy, P. Rajaraman, J. A. Schwartzbaum, S. Sadetzki, B. Schlehofer, T. Tihan, J. L. Wiemels, M. Wrensch, and P. A. Buffler, "Brain tumor epidemiology: Consensus from the brain tumor epidemiology consortium," *Cancer*, vol. 113, no. 7 Suppl, pp. 1953–1968, 2008.

[14] V. Harris and H. Harris, "Types, symptoms, and current treatment protocols for brain tumors," *International Journal of Advanced Engineering Technologies and Innovations*, vol. 10, no. 2, pp. 352–364, 2024. [Online]. Available: https://www.researchgate.net/publication/385721793

[15] S. R. Gesaka, P. M. Okemwa, and P. M. Mwachaka, "Histological types of brain tumors diagnosed at the kenyatta national hospital between 2016 and 2019: a retrospective study," *Discover Oncology*, vol. 15, p. 39, 2024. [Online]. Available: https://doi.org/10.1007/s12672-024-00893-6

[16] B. Alther, V. Mylius, M. Weller, and A. Gantenbein, "From first symptoms to diagnosis: Initial clinical presentation of primary brain tumors," *Clinical & Translational Neuroscience*, vol. 4, no. 1, pp. 1–7, 2020. [Online]. Available: https://journals.sagepub.com/doi/10.1177/2514183X20968368

[17] J. Choi and H. J. Cho, "Comprehensive transcriptomic profiling of diverse brain tumor types uncovers complex structures of the brain tumor microenvironment," *Biomedicines*,

vol. 12, p. 506, 2024. [Online]. Available: https://www.mdpi.com/2227-9059/12/3/506

[18] F. Ghandour, A. Squassina, R. Karaky, M. Diab-Assaf, P. Fadda, and C. Pisanu, "Presenting psychiatric and neurological symptoms and signs of brain tumors before diagnosis: A systematic review," *Brain Sciences*, vol. 11, no. 3, p. 301, 2021. [Online]. Available: https://doi.org/10.3390/brainsci11030301

[19] A. Perkins and G. Liu, "Primary brain tumors in adults: Diagnosis and treatment," *American Family Physician*, vol. 93, no. 3, pp. 211–217, 2016. [Online]. Available: https://www.aafp.org/pubs/afp/issues/2016/0201/p211.html

[20] S. R. Chandana, S. Movva, M. Arora, and T. Singh, "Primary brain tumors in adults," *American Family Physician*, vol. 77, no. 10, pp. 1423–1430, 2008. [Online]. Available: https://www.aafp.org/afp/2008/0515/p1423.html

[21] M. Kim and H. S. Kim, "Emerging techniques in brain tumor imaging: What radiologists need to know," *Korean Journal of Radiology*, vol. 17, no. 5, pp. 598–619, 2016. [Online]. Available: https://doi.org/10.3348/kjr.2016.17.5.598

[22] D. Ganguly, S. Chakraborty, M. Balitanas, and T. hoon Kim, "Medical imaging: A review," in *SUComS 2010, Communications in Computer and Information Science (CCIS)*. Springer-Verlag Berlin Heidelberg, 2010, vol. 78, pp. 504–516.

[23] Z. H. Cho, D. Mccaughey, E. L. Hall, and R. P. Kruger, "Houndsfield nobel prize," *ResearchGate*, 2013. [Online]. Available: https://www.researchgate.net/publication/235968323

[24] W. Liu, S. Wang, and X. Zhang, "A study of classification and feature selection algorithms in high-dimensional biomedical data," *International Journal of Information Technology*, vol. 14, no. 1, pp. 37–45, 2008.

[25] W. B. Overcast, K. M. Davis, C. Y. Ho, G. D. Hutchins, M. A. Green, B. D. Graner, and M. C. Veronesi, "Advanced imaging techniques for neuro-oncologic tumor diagnosis, with an emphasis on pet-mri imaging of malignant brain tumors," *Current Oncology Reports*, vol. 23, p. 34, 2021. [Online]. Available: https://doi.org/10.1007/s11912-021-01020-2

[26] M. Symms, H. R. Jäger, K. Schmierer, and T. A. Yousry, "A review of structural magnetic

resonance neuroimaging," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 75, pp. 1235–1244, 2004. [Online]. Available: http://jnnp.bmj.com

[27] D. B. Plewes and W. Kucharczyk, "Physics of mri: A primer," *Journal of Magnetic Resonance Imaging*, vol. 35, pp. 1038–1054, 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/jmri.23642

[28] M. M. Badža and M. Barjaktarović, "Classification of brain tumors from mri images using a convolutional neural network," *Applied Sciences*, vol. 10, no. 6, p. 1999, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/6/1999

[29] M. Weininger, B. Lauterbach, S. Knop, T. Pabst, W. Kenn, D. Hahn, and M. Beissert, "Whole-body mri of multiple myeloma: Comparison of different mri sequences in assessment of different growth patterns," *European Journal of Radiology*, vol. 69, pp. 339–345, 2009. [Online]. Available: https://doi.org/10.1016/j.ejrad.2007.10.025

[30] G. Widmann, B. Henninger, C. Kremser, and W. Jaschke, "Mri sequences in head & neck radiology – state of the art," *Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, vol. 189, pp. 413–422, 2017. [Online]. Available: http://dx.doi.org/10.1055/s-0043-103280

[31] R. Mutiki, "Mri basics: How to read and understand mri sequences," Published on SlideShare, accessed: March 23, 2025. [Online]. Available: https://www.slideshare.net/rameshmutiki/mri-basics-how-to-read-and-understand-mri-sequences

[32] M. C. Mabray, R. F. B. Jr, and S. Cha, "Modern brain tumor imaging," *Brain Tumor Research and Treatment*, vol. 3, no. 1, pp. 8–23, 2015. [Online]. Available: http://dx.doi.org/10.14791/btrt.2015.3.1.8

[33] A. N. Onaizah, Y. Xia, and K. Hussain, "FL-SiCNN: An improved brain tumor diagnosis using siamese convolutional neural network in a peer-to-peer federated learning approach," *Alexandria Engineering Journal*, vol. 114, pp. 1–11, 2025, available online 26 November 2024. [Online]. Available: https://doi.org/10.1016/j.aej.2024.11.063

[34] M. Hassan, A. A. Fateh, J. Lin, Y. Zhuang, G. Lin, H. Xiong, Z. You, P. Qin, and H. Zeng, "Unfolding Explainable AI for Brain Tumor Segmentation," *Neurocomputing*,

vol. 599, p. 128058, 2024, available online 18 June 2024. [Online]. Available: https://doi.org/10.1016/j.neucom.2024.128058

[35] C. M. Umarani, S. Gollagi, S. Allagi, K. Sambrekar, and S. B. Ankali, "Advancements in deep learning techniques for brain tumor segmentation: A survey," *Informatics in Medicine Unlocked*, vol. 50, p. 101576, 2024, available online 22 August 2024. [Online]. Available: https://doi.org/10.1016/j.imu.2024.101576

[36] M. F. Ahamed, M. M. Hossain, M. Nahiduzzaman, M. R. Islam, M. R. Islam, M. Ahsan, and J. Haider, "A review on brain tumor segmentation based on deep learning methods with federated learning techniques," *Computerized Medical Imaging and Graphics*, vol. 110, p. 102313, 2023, available online 24 November 2023. [Online]. Available: https://doi.org/10.1016/j.compmedimag.2023.102313

[37] D. Kaur and Y. Kaur, "Various image segmentation techniques: A review," *International Journal of Computer Science and Mobile Computing*, vol. 3, no. 5, pp. 809–814, 2014.

[38] S. Kordnoori, M. Sabeti, M. H. Shakoor, and E. Moradi, "Deep multi-task learning structure for segmentation and classification of supratentorial brain tumors in mr images," *Interdisciplinary Neurosurgery: Advanced Techniques and Case Management*, vol. 36, p. 101931, 2024, available online 25 November 2023. [Online]. Available: https://doi.org/10.1016/j.inat.2023.101931

[39] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, vol. 86, no. 11. IEEE, 1998, pp. 2278–2324.

[40] M. I. Nazir, A. Akter, M. A. H. Wadud, and M. A. Uddin, "Utilizing customized cnn for brain tumor prediction with explainable ai," *Heliyon*, vol. 10, no. 10, p. e38997, 2024, available online 9 October 2024. [Online]. Available: https://doi.org/10.1016/j.heliyon.2024.e38997

[41] W. Zafar, G. Husnain, A. Iqbal, A. S. Alzahrani, M. A. Irfan, Y. Y. Ghadi, M. S. AL-Zahrani, and R. S. Naidu, "Enhanced tumornet: Leveraging yolov8s and u-net for superior brain tumor detection and segmentation utilizing mri scans," *Results in Engineering*, vol. 24, p. 102994, 2024, available online 29 September 2024. [Online].

Available: https://doi.org/10.1016/j.rineng.2024.102994

[42] C. H. Kang and S. Y. Kim, "Real-time object detection and segmentation technology: an analysis of the yolo algorithm," *JMST Advances*, vol. 5, pp. 69–76, 2023, published online 8 September 2023. [Online]. Available: https://doi.org/10.1007/s42791-023-00049-7

[43] Q. Yao, D. Zhuang, Y. Feng, Y. Wang, and J. Liu, "Accurate detection of brain tumor lesions from medical images based on improved yolov8 algorithm," *IEEE Access*, vol. 12, pp. 144 260–144 277, 2024. [Online]. Available: https://doi.org/10.1109/ACCESS.2024.3472039

[44] M. G. Ragab, S. J. Abdulkadir, A. Muneer, A. Alqushaibi, E. H. Sumiea, R. Qureshi, S. M. Al-Selwi, and H. Alhussian, "A comprehensive systematic review of yolo for medical object detection (2018 to 2023)," *IEEE Access*, vol. 12, pp. 57 815–57 832, 2024, open Access under CC BY-NC-ND 4.0 License.

[45] I. Kovács, T. Székely, P. Pogány, S. Takács, M. Erős, and B. Járay, "Utilizing the open-source programming language python to create interactive quality assurance dashboards for diagnostic and screening performance in cytology," *Journal of the American Society of Cytopathology*, vol. 13, pp. 309–318, 2024. [Online]. Available: https://doi.org/10.1016/j.jasc.2024.03.007