SWINME: SWIN TRANSFORMER V2-BASED FRAMEWORK FOR MULTIMODAL BRAIN TUMOR SEGMENTATION

Conduct by:

Name : Abdullah Mohammed Ahmed Abdo Alhwyji

Student ID : 22523230

PROGRAM STUDI INFORMATIKA – PROGRAM SARJANA

FAKULTAS TEKNOLOGI INDUSTRI

UNIVERSITAS ISLAM INDONESIA

2026

SUPERVISOR APPROVAL PAGE

SWINME: SWIN TRANSFORMER V2-BASED FRAMEWORK FOR MULTIMODAL BRAIN TUMOR SEGMENTATION

THESIS

Conduct By:

Name : Abdullah Mohammed Ahmed Abdo Alhwyji

Student ID : 22523230

Yogyakarta, January 18th, 2026

Advisor,

(Arrie Kurniawardhani, S.Si., M.Kom)

EXAMINER ENDORSEMENT PAGE

SWINME: SWIN TRANSFORMER V2-BASED FRAMEWORK FOR MULTIMODAL BRAIN TUMOR SEGMENTATION

THESIS

Has been defended in front of the examiners as one of the requirements to obtain a Bachelor of Informatics degree from the Undergraduate Program in Informatics at the Faculty of Industrial Technology, Universitas Islam Indonesia

Yogyakarta, January 9th, 2026

Chair

Arrie Kurniawardhani, S.Si., M.Kom

Examiner 1

Ahmad Fathan Hidayatullah, S.T., M.Cs., Ph.D.

Examiner 2

Izzati Muhimmah, S.T., M.Sc., Ph.D.

Acknowledged by,

Head of Undergraduate Program in Informatics

Faculty of Industrial Technology

Universitas Islam Indonesia

(Ir. Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D.)

AUTHENTICITY STATEMENT

The undersigned:

Name : Abdullah Mohammed Ahmed Abdo Alhwyji

NIM : 22523230

Final project with title:

SWINME: SWIN TRANSFORMER V2-BASED FRAMEWORK FOR MULTIMODAL BRAIN TUMOR SEGMENTATION

Stating that all components and contents in this final project are my own work. If in the future it is proven that some parts of this work are not my own work, the final project submitted as my own work is ready to be withdrawn and ready to bear any risks and consequences.

Thus, this statement letter is made, hopefully it can be used properly.

Yogyakarta, January 18th ,2026

abdullah

Abdullah Mohammed Ahmed Abdo Alhwyji

22523230

DEDICATION

This Final Year Project report is dedicated to my beloved parents and family for their unwavering support, prayers, and sacrifices, as well as to the lecturers and educators of the Informatics Department, Faculty of Industrial Technology, Universitas Islam Indonesia, whose guidance and knowledge have greatly contributed to my academic journey. It is also dedicated to all those who supported and assisted me, directly or indirectly, in the completion of this work.

MOTTO

﴿ وَقُلِ اعْمَلُوا فَسَيَرَى اللَّهُ عَمَلَكُمْ ﴾

And say, “Do [as you will], for Allah will see your deeds.”
(Surah At-Tawbah, Qur’an 9:105)

ACKNOWLEDGEMENTS

All praise and gratitude are due to Allah SWT, the Almighty, for His guidance, blessings, and mercy, which have enabled me to successfully complete this Final Year Project and its accompanying report. Through His will and strength, I was able to face and overcome the challenges encountered throughout the research and writing process. I would like to express my deepest and most sincere gratitude to Mrs. Arrie Kurniawardhani, my supervisor, for her exceptional guidance, patience, and unwavering support throughout the completion of this Final Year Project. Her valuable academic insights, constructive feedback, and continuous encouragement have greatly shaped the direction, quality, and rigor of this research. I am truly grateful for her dedication, understanding, and willingness to devote her time to guide me through each stage of the research process. Her mentorship has not only contributed to the successful completion of this project but has also enriched my academic growth and research experience. I am deeply thankful to my parents for their unconditional love, prayers, moral support, and constant encouragement throughout my academic journey. Their sacrifices, understanding, and trust have been a continuous source of motivation and strength, especially during challenging moments. I would also like to extend my sincere appreciation to the Informatics Department, Faculty of Industrial Technology, Universitas Islam Indonesia, including the Dean of the Faculty of Industrial Technology, the Head of the Informatics Study Program, lecturers, and all academic staff, for providing a supportive academic environment and valuable knowledge throughout my studies. In addition, I would like to express my gratitude to the International Office of Universitas Islam Indonesia for their administrative support, guidance, and assistance, which contributed positively to my academic experience and the completion of this Final Year Project. Finally, I hope that this report will be beneficial to the academic community and serve as a useful reference for future research. I am fully aware that this work may still have limitations; therefore, any constructive criticism and suggestions for improvement are sincerely welcomed.

ABSTRACT

Accurate and efficient brain tumor segmentation from Magnetic Resonance Imaging (MRI) is a critical component in neuro-oncology, as it directly supports diagnosis, treatment planning, and disease progression assessment. This requirement is particularly significant in pediatric gliomas, where early and precise clinical intervention can substantially influence patient outcomes. Despite its importance, manual tumor delineation remains a time-consuming and subjective process, highly dependent on expert availability and prone to inter-observer variability. These limitations restrict scalability and reliability, especially when dealing with high-resolution three-dimensional MRI data and in healthcare environments with limited resources. Consequently, there is a strong need for automated segmentation methods that are both accurate and computationally efficient.

This thesis proposes a novel hybrid deep learning framework, termed SwinME-UNETR 3D, for automated brain tumor segmentation. The proposed approach integrates the hierarchical attention mechanism of Swin Transformer V2 with a multi-scale enhancement strategy (SwinME) and the volumetric modeling capability of UNETR architecture. The framework is designed to capture long-range spatial dependencies while preserving fine-grained local features that are essential for precise tumor boundary delineation. Multi-scale feature enhancement and an enhanced transformer module are incorporated to strengthen feature representation across different spatial resolutions, enabling effective handling of tumor heterogeneity and complex anatomical structures present in three-dimensional MRI volumes.

The proposed method is evaluated using the BraTS 2023 dataset, which comprises multimodal MRI scans including T1, T1 post-contrast (T1ce), T2, and FLAIR sequences. Segmentation performance is assessed on clinically relevant tumor subregions, namely whole tumor (WT), tumor core (TC), and enhancing tumor (ET), using Dice Similarity Coefficient and sensitivity metrics. Experimental results demonstrate robust and consistent performance, achieving an average Dice score of 0.91 on the validation set.

Qualitative analysis further confirms strong alignment with expert annotations and high volumetric consistency. These findings indicate that transformer-based architectures augmented with multi-scale enhancement provide an effective solution for automated brain tumor segmentation and hold strong potential for future clinical decision-support applications.

Keywords: Brain tumor segmentation, Swin Transformer V2, UNETR 3D, BraTS Dataset, Multi-Scale Enhancement (SwinME)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

Medical imaging has undergone a profound evolution over the past several decades and has become an indispensable component of modern healthcare systems (Litjens, et al., 2017). Early diagnostic practices relied heavily on clinical examination and invasive procedures, which were often limited in accuracy and posed significant risks to patients. The development of advanced medical imaging modalities, such as X-ray, computed tomography (CT), ultrasound, and magnetic resonance imaging (MRI), transformed clinical practice by enabling non-invasive visualization of internal anatomical structures. Today, medical imaging is central to disease screening, diagnosis, treatment planning, therapy monitoring, and longitudinal patient follow-up. As imaging technologies continue to improve in spatial resolution, contrast sensitivity, and acquisition speed, the volume and complexity of imaging data generated in routine clinical workflows have increased dramatically.

Within this broader landscape, neuro-oncology represents one of the most demanding and clinically critical domains of medical imaging. Brain tumors are among the most severe neurological disorders, often associated with high morbidity, complex treatment pathways, and poor survival outcomes, particularly in the case of malignant gliomas. According to global health statistics, primary and secondary brain tumors impose a substantial burden on healthcare systems worldwide, affecting patients across all age groups. The clinical management of brain tumors requires precise diagnosis, accurate assessment of tumor extent, and careful monitoring of disease progression or response to therapy. Medical imaging plays a decisive role in each of these stages, making reliable image interpretation essential for effective neuro-oncological care.

Magnetic resonance imaging (MRI) is widely regarded as the gold standard imaging modality for brain tumor diagnosis and assessment (Bauer, Wiest, Nolte, & Reyes, 2013). Unlike CT, MRI provides superior soft tissue contrast and allows detailed visualization of brain anatomy without exposure to ionizing radiation. Moreover, MRI supports a variety of acquisition protocols, enabling the capture of complementary information

about tissue properties, vascularity, and pathological changes. Multimodal MRI sequences, including T1-weighted, T1 post-contrast (T1ce), T2-weighted, and fluid-attenuated inversion recovery (FLAIR) images, are routinely used to characterize different aspects of brain tumors (Bakas, et al., 2018). These modalities collectively enable clinicians to identify tumor boundaries, differentiate tumor subregions, and assess surrounding edema or infiltration, which are critical factors in treatment planning.

A particularly important aspect of brain tumor analysis is the delineation of clinically meaningful tumor subregions, such as the enhancing tumor, tumor core, and peritumoral edema. Accurate segmentation of these subregions is essential for multiple clinical tasks, including surgical resection planning, radiotherapy dose optimization, and quantitative evaluation of treatment response (Louis, et al., 2016). Small inaccuracies in subregion delineation can lead to suboptimal treatment decisions, potentially affecting patient outcomes. Consequently, tumor subregion segmentation is not merely a technical task but a clinically significant process that directly influences therapeutic strategies in neuro-oncology.

Despite the importance of precise tumor segmentation, interpreting volumetric MRI data remains a challenging task. Brain MRI scans are inherently three-dimensional, high-dimensional, and often affected by noise, intensity inhomogeneity, and inter-patient anatomical variability. Tumors themselves exhibit heterogeneous appearance, irregular shapes, and diffuse boundaries that blend gradually into healthy tissue. These characteristics complicate the visual interpretation of MRI data, even for experienced radiologists. Furthermore, the manual delineation of tumor regions across hundreds of image slices is time-consuming, cognitively demanding, and subject to inter-observer and intra-observer variability (Bakas, et al., 2018). Such variability can introduce inconsistencies in diagnosis and limit the reproducibility of quantitative imaging biomarkers.

The growing scale of medical imaging data, combined with increasing clinical workload and limited availability of specialized experts, has intensified the demand for automated image analysis solutions. Automated brain tumor segmentation systems aim to provide fast, consistent, and objective delineation of tumor regions, thereby reducing clinician workload and improving reproducibility across institutions. Automation also enables large-scale quantitative studies and facilitates the integration of imaging data into data-driven clinical decision-support systems. As a result, automated segmentation has become

a central research focus in medical image analysis and a key enabler of precision medicine.

Traditional image processing and early machine learning approaches attempted to address this problem using handcrafted features, thresholding techniques, region-growing algorithms, and statistical classifiers. While these methods achieved limited success in controlled settings, they generally lacked robustness when applied to heterogeneous clinical data. Their reliance on manual feature engineering and sensitivity to noise and imaging variability restricted their generalization capability. Manual segmentation, although accurate when performed by experts, remains impractical for large-scale or time-sensitive applications. These limitations have motivated the exploration of more powerful data-driven approaches.

In recent years, artificial intelligence (AI), particularly deep learning, has emerged as a transformative paradigm in medical imaging. Deep learning models can learn hierarchical feature representations directly from raw data, eliminating the need for handcrafted features and enabling end-to-end optimization. Convolutional neural networks (CNNs) have become the dominant architecture for medical image segmentation due to their ability to capture local spatial patterns and contextual information through convolutional operations and encoder–decoder structures (Ronneberger, Fischer, & Brox, 2015). CNN-based models, such as U-Net and its variants, have demonstrated substantial performance improvements over traditional methods in a wide range of medical imaging tasks, including brain tumor segmentation.

However, despite their success, CNN-based approaches exhibit inherent limitations when applied to complex volumetric data (Kamnitsas, et al., 2017). Convolutional operations are fundamentally local, meaning that long-range dependencies and global contextual relationships must be inferred indirectly through deep stacking of layers or large receptive fields. This limitation becomes particularly problematic in brain tumor segmentation, where tumors may span large spatial regions, exhibit non-local dependencies, and require holistic understanding of the entire brain context. Additionally, CNNs may struggle to balance fine-grained boundary precision with global structural awareness, leading to fragmented or inconsistent segmentation results.

Transformer-based architecture offers a promising alternative by explicitly modeling global interactions through self-attention mechanisms (Vaswani, et al., 2017). Unlike CNNs, transformers can directly capture long-range dependencies by computing

relationships between all elements in an input representation. Originally developed for natural language processing, transformers have been successfully adapted to vision tasks, demonstrating strong performance in scenarios where global context is crucial. In medical imaging, this ability to integrate information across distant spatial regions is particularly advantageous for volumetric segmentation tasks involving complex anatomical structures.

Recent developments in vision transformers, especially hierarchical and window-based designs, have addressed many of the computational challenges associated with applying transformers to high-resolution images (Liu, et al., 2022). Architectures such as the Swin Transformer introduce efficient attention mechanisms that balance global context modeling with computational feasibility. These models are especially well suited for three-dimensional medical imaging, where memory efficiency and scalability are critical considerations. The integration of transformer-based representations with encoder–decoder frameworks further enhance their suitability for dense prediction tasks such as brain tumor segmentation.

Motivated by these advances, this research is situated at the intersection of medical imaging, deep learning, and transformer-based architectures. The overarching motivation is to develop an automated brain tumor segmentation framework that overcomes the limitations of manual annotation and conventional CNN-based methods while leveraging the strengths of hierarchical attention and multi-scale representation learning. By doing so, this work aims to contribute to the development of accurate, robust, and scalable segmentation systems that can support clinical decision-making and advance the application of AI in neuro-oncological imaging.

1.2 Problem Statement

Accurate segmentation of brain tumors from multimodal magnetic resonance imaging (MRI) data remains a fundamental yet unresolved challenge in neuro-oncological imaging, despite decades of progress in medical imaging technology and computational analysis methods. In current clinical practice, tumor segmentation is still largely performed manually by expert radiologists, who delineate tumor boundaries and subregions through careful visual inspection of three-dimensional MRI volumes. This manual process requires extensive domain knowledge, sustained attention, and significant time investment, as radiologists must analyze multiple imaging modalities and hundreds

of slices per patient. While manual annotation can yield high-quality results when performed by experienced clinicians, it is inherently subjective and vulnerable to inter-observer and intra-observer variability. Such variability leads to inconsistencies in tumor boundary definition and subregion labeling, which can directly influence treatment decisions, outcome assessment, and the reproducibility of quantitative imaging studies.

The reliance on manual segmentation poses serious challenges in the context of modern healthcare systems, where imaging workloads continue to increase due to population growth, aging demographics, and broader access to advanced diagnostic technologies. The limited availability of specialized radiologists further exacerbates this problem, making manual segmentation an impractical solution for large-scale clinical deployment or time-sensitive decision-making. Moreover, the lack of standardization inherent in manual annotation complicates the use of imaging data for longitudinal studies, multi-center trials, and data-driven clinical research, where consistency and reproducibility are essential. These clinical constraints highlight the urgent need for automated segmentation approaches that can deliver reliable and consistent results with minimal human intervention.

From a computational perspective, automated brain tumor segmentation introduces substantial technical challenges arising from the nature of volumetric MRI data. Brain MRI scans are high-dimensional three-dimensional representations that capture complex anatomical structures at fine spatial resolution. In addition, clinical protocols routinely employ multiple MRI modalities, each emphasizing different tissue properties and pathological characteristics. Effective segmentation requires not only voxel-level classification accuracy but also preservation of spatial continuity across slices and integration of complementary information across modalities. Approaches that operate solely on two-dimensional slices often ignore inter-slice dependencies, leading to anatomically inconsistent predictions and fragmented tumor representations. Conversely, fully three-dimensional models must process large volumetric inputs, which significantly increases memory consumption, computational complexity, and training time, thereby limiting scalability and accessibility.

Convolutional neural networks (CNNs) have emerged as the dominant paradigm for automated medical image segmentation due to their strong ability to learn hierarchical feature representations from data. Encoder–decoder architectures, in particular, have

demonstrated impressive performance by combining deep feature extraction with skip connections that preserve spatial detail. However, CNN-based models are fundamentally constrained by the local nature of convolutional operations. The receptive field of a CNN expands only incrementally as network depth increases, making it difficult for such models to capture long-range dependencies and global contextual relationships within volumetric data. Although architectural modifications such as deeper networks, dilated convolutions, and multi-scale feature fusion have been proposed to alleviate this limitation, these approaches often increase model complexity without fully addressing the underlying issue of limited global context modeling.

The restricted ability of CNN-based models to reason globally has important implications for brain tumor segmentation. Brain tumors frequently exhibit irregular geometries, diffuse infiltration patterns, and gradual intensity transitions rather than well-defined boundaries. In many cases, local image features alone are insufficient to distinguish tumor tissue from surrounding healthy structures, particularly when contrast differences are subtle or confounded by imaging artifacts. As a result, CNN-based segmentation methods may produce outputs with boundary inconsistency, fragmented regions, or misclassification of tumor subregions. These issues are especially pronounced in challenging cases involving heterogeneous tumors, where different subregions—such as enhancing tumor, necrotic core, and peritumoral edema—exhibit distinct and sometimes overlapping imaging characteristics across modalities.

Tumor heterogeneity introduces an additional layer of complexity that further limits the effectiveness of conventional segmentation approaches (Louis, et al., 2016) (Bakas, et al., 2018). Brain tumors are biologically and structurally heterogeneous, with variations in cellular composition, vascularization, and infiltration patterns that manifest differently across patients and imaging modalities. Accurate segmentation requires models capable of capturing both fine-grained local details and broader contextual relationships that span large spatial extents. CNN-based architecture often struggles to achieve this balance, as mechanisms designed to enhance local boundary precision may compromise global coherence, while strategies aimed at increasing contextual awareness may dilute spatial detail. This trade-off presents a fundamental challenge in designing segmentation models that are both precise and globally consistent.

Transformer-based architectures have recently emerged as a promising alternative to convolution-centric models by introducing self-attention mechanisms that explicitly model relationships between distant elements in the input representation. Through self-attention, transformers can directly capture long-range dependencies and global contextual interactions, offering a more holistic representation of spatial relationships within an image or volume. This property is particularly advantageous for volumetric medical imaging, where understanding global anatomical context is essential for resolving ambiguities in local appearance. However, the direct application of standard transformer architectures to three-dimensional MRI data is hindered by their high computational and memory requirements, which scale poorly with input size and resolution.

Consequently, there exists a critical need for scalable three-dimensional transformer-based segmentation frameworks that can effectively harness the strengths of attention mechanisms while remaining computationally feasible. Such frameworks must be capable of processing multimodal volumetric MRI data, integrating information across multiple spatial scales, and maintaining spatial coherence throughout the segmentation process. At the same time, they must address practical constraints related to memory usage, training efficiency, and deployment of commonly available hardware. Achieving this balance requires careful architectural design that adapts transformer principles to the specific demands of dense volumetric prediction tasks in medical imaging.

In this context, the central problem addressed by this research is the absence of an automated brain tumor segmentation approach that simultaneously satisfies clinical reliability, global contextual awareness, boundary precision, and computational scalability. Existing methods either rely heavily on manual intervention, suffer from limited global modeling capabilities, or impose prohibitive computational costs. Addressing this gap necessitates the development of advanced segmentation architectures that can overcome the limitations of traditional convolutional models while leveraging the global reasoning capabilities of transformers in a manner suitable for three-dimensional medical imaging. This unresolved challenge forms the foundation of the present research and motivates the exploration of scalable transformer-based solutions for robust and accurate brain tumor segmentation.

1.3 Research Objectives

This research aims to advance automated brain tumor segmentation by developing and systematically evaluating a robust three-dimensional deep learning framework tailored to multimodal magnetic resonance imaging (MRI). Manual tumor delineation remains a time-consuming and subjective process that is highly dependent on expert availability and prone to inter-observer variability, while many existing automated methods struggle to maintain spatial coherence and boundary precision when applied to high-resolution volumetric data. These limitations are particularly evident in cases involving complex tumor morphology, heterogeneous intensity patterns, and diffuse infiltration into surrounding tissue. Motivated by these challenges, this study focuses on transformer-based architectures, which offer explicit modeling of long-range spatial dependencies and global contextual relationships. By leveraging such architectures, the research seeks to overcome the locality constraints of convolution-dominant models and to provide a more principled solution for volumetric brain tumor segmentation.

A further objective of this thesis is to investigate how hierarchical attention mechanisms and structured multi-scale representation learning can be effectively integrated within a three-dimensional segmentation framework. Accurate delineation of clinically relevant tumor subregions, including enhancing tumor, tumor core, and peritumoral edema, requires simultaneous sensitivity to fine-grained local boundaries and broader anatomical context. However, achieving this balance remains a persistent challenge in existing segmentation approaches. This research therefore aims to examine architectural strategies that enable information to be aggregated across multiple spatial resolutions while preserving spatial detail and anatomical consistency throughout the entire MRI volume. Through systematic architectural design and analysis, the study seeks to clarify how hierarchical attention and multi-scale enhancement contribute to improved segmentation accuracy and robustness.

In addition to architectural development, this research emphasizes the establishment of a rigorous and reproducible experimental methodology for evaluating transformer-based volumetric segmentation models. Reliable assessment of segmentation performance requires standardized data preprocessing, carefully designed training strategies, appropriate loss formulations to address class imbalance, and comprehensive evaluation using both quantitative metrics and qualitative visual analysis. Furthermore, the research aims to conduct systematic ablation studies to isolate the impact of individual

architectural components and to assess computational feasibility under realistic hardware constraints. By adopting this methodological rigor, the thesis seeks not only to demonstrate performance improvements but also to provide meaningful insights into the relationship between architectural design choices and observed segmentation behavior. Accordingly, the objectives of this research can be summarized as follows:

a. To develop a robust transformer-based framework for three-dimensional brain tumor segmentation using multimodal MRI.
b. To investigate hierarchical attention and multi-scale representation learning for improved segmentation accuracy and spatial consistency.
c. To establish a rigorous and reproducible experimental methodology for evaluating volumetric segmentation models.

1.4 Research Questions

This thesis is guided by fundamental questions concerning the effective adaptation of transformer-based architectures to volumetric medical image segmentation. It seeks to understand whether hierarchical attention mechanisms can be designed to explicitly model long-range spatial dependencies in three-dimensional multimodal MRI data, and whether such modeling improves spatial coherence and anatomical consistency in brain tumor segmentation. Segmentation errors in volumetric data often manifest as boundary discontinuities and inter-slice inconsistencies, which can compromise clinical interpretability. This research therefore investigates how attention-driven global context modeling influences boundary continuity and structural integrity across the entire MRI volume.

A second research question addresses the role of structured multi-scale feature enhancement in capturing tumor heterogeneity. Brain tumors exhibit significant variability in size, shape, and internal composition, with different subregions presenting distinct imaging characteristics across modalities. Accurately segmenting these heterogeneous subregions remains a challenge for many existing approaches. This study explores whether integrating multi-scale representations enables more effective discrimination between tumor components and improves segmentation performance in regions characterized by subtle intensity variations and ambiguous boundaries.

The third research question focuses on the practical feasibility of transformer-based volumetric segmentation frameworks. While attention mechanisms offer theoretical

advantages in global context modeling, they are often associated with increased computational and memory requirements. This research therefore examines whether a three-dimensional transformer-based segmentation model can achieve strong segmentation accuracy while maintaining computational efficiency and scalability under realistic hardware constraints, thereby supporting reproducibility and potential future applicability beyond controlled research environments.
Accordingly, this thesis addresses the following research questions:

Q1: How can hierarchical attention mechanisms be adapted to improve spatial coherence and boundary consistency in volumetric brain tumor segmentation?

Q2: Does structured multi-scale feature enhancement improve the segmentation of heterogeneous tumor subregions in multimodal MRI data?

Q3: How can segmentation accuracy be balanced with computational efficiency in transformer-based three-dimensional segmentation models?

1.5 Scope and Limitations

This research focuses on the development and evaluation of an automated framework for brain tumor segmentation using multimodal magnetic resonance imaging (MRI) data within an academic research setting. The scope of the study is defined to enable a focused and rigorous investigation of three-dimensional segmentation methods while maintaining alignment with the stated research objectives and available resources. The primary emphasis is placed on accurate delineation of clinically relevant tumor subregions, including enhancing tumor, tumor core, and peritumoral edema, which are essential for neuro-oncological analysis and treatment planning.

The experimental scope of this work is limited to the use of a publicly available benchmark dataset for brain tumor segmentation, namely the Brain Tumor Segmentation (BraTS) dataset. This dataset provides standardized, multi-institutional MRI scans acquired using consistent imaging protocols and includes expert-annotated reference labels for evaluation. The use of a benchmark dataset ensures reproducibility and facilitates objective comparison with existing methods reported in the literature. However, reliance on a single dataset constrains the diversity of imaging conditions, scanner types, and patient demographics represented in the study, which may limit the generalizability of the findings to broader clinical environments.

The study is further scoped to the analysis of commonly used MRI modalities for brain tumor assessment, including T1-weighted, T1 post-contrast (T1ce), T2-weighted, and fluid-attenuated inversion recovery (FLAIR) images. These modalities are selected due to their established clinical relevance and their ability to highlight different tumor characteristics. Other imaging modalities, such as diffusion-weighted imaging or perfusion MRI, are not considered in this work. Consequently, the proposed framework is designed and evaluated specifically for the selected modalities and does not account for additional sources of imaging information that may be available in advanced clinical protocols.

From a methodological perspective, this research is confined to deep learning–based approaches for medical image segmentation, with particular emphasis on three-dimensional architectures and attention-based models. Traditional image processing techniques and classical machine learning methods are discussed for contextual comparison but are not the focus of model development or experimental evaluation. Furthermore, the scope of the study is limited to algorithmic design and performance analysis, and does not extend to system-level integration, user interface development, or workflow deployment in clinical environments.

The experimental evaluation conducted in this research relies on quantitative and qualitative performance assessment using established segmentation metrics, such as the Dice similarity coefficient and sensitivity. While these metrics provide objective measures of segmentation accuracy and are widely accepted in the literature, they do not fully capture all aspects of clinical utility, such as the impact of segmentation errors on clinical decision-making or patient outcomes. In addition, qualitative analysis is based on visual inspection of segmentation outputs rather than formal assessment by clinical experts, which constitutes an inherent limitation of the study.

Computational constraints represent another important limitation of this research. Training and evaluating three-dimensional deep learning models, particularly those incorporating attention mechanisms, require substantial computational resources. As a result, model complexity, input resolution, batch size, and the extent of hyperparameter exploration are influenced by available hardware. Although the proposed framework is designed with efficiency in mind, the investigation of larger-scale models or exhaustive architectural variations is restricted by these practical considerations.

Finally, this study does not include clinical validation, prospective testing, or real-time deployment of the proposed segmentation framework. The segmentation results are evaluated against reference annotations provided in the benchmark dataset rather than within a live clinical workflow. Ethical, regulatory, and legal considerations related to the deployment of artificial intelligence systems in healthcare are acknowledged as important but are beyond the scope of this research.

1.6 Contributions to the Study

This thesis contributes to the field of medical image analysis through a combination of architectural innovation, methodological rigor, and analytical insight. At the architectural level, it introduces a transformer-centric three-dimensional segmentation framework that integrates hierarchical window-based attention, structured multi-scale enhancement, and feature refinement within a unified encoder–decoder design. This framework is specifically tailored to address the challenges of volumetric brain tumor segmentation, including boundary ambiguity, spatial discontinuity, and heterogeneous tumor appearance, while remaining computationally feasible for high-resolution MRI data.

From a methodological perspective, the research establishes a comprehensive and reproducible experimental pipeline for transformer-based volumetric segmentation. This includes standardized preprocessing procedures, carefully designed training and optimization strategies, appropriate loss formulations to address class imbalance, and systematic evaluation using both quantitative performance metrics and qualitative visual analysis. In addition, controlled ablation studies are conducted to isolate the contribution of individual architectural components, enabling transparent interpretation of performance gains and supporting reproducibility and fair comparison with existing methods.

Finally, the thesis provides analytical insight into the strengths, limitations, and practical considerations of attention-based architectures for three-dimensional medical image segmentation. By examining segmentation performance across different tumor subregions, assessing generalization behavior, and analyzing computational efficiency, the research clarifies the conditions under which transformer-based models offer advantages over conventional approaches. These insights contribute to a deeper understanding of volumetric segmentation design principles and provide a foundation for

future research aimed at improving scalability, robustness, and eventual clinical translation.

Accordingly, the main contributions of this thesis can be summarized as follows:

- **a.** Proposing a unified transformer-based architecture with hierarchical attention and multi-scale enhancement for volumetric brain tumor segmentation.
- **b.** Providing a systematic and reproducible experimental evaluation of transformer-based segmentation models.
- **c.** Offering analytical insights into the performance, robustness, and practical feasibility of attention-based volumetric segmentation frameworks.

1.7 Organization of the Thesis

This report is organized into seven chapters, each addressing a specific component of the research and collectively providing a structured investigation into automated brain tumor segmentation using multimodal magnetic resonance imaging (MRI).

Chapter 1 introduces the study by presenting the background and motivation, defining the research problem, outlining the research objectives and questions, and clarifying the scope, limitations, contributions, and significance of the work. This chapter establishes the context and rationale for the research and frames the direction of the subsequent chapters.

Chapter 2 presents the theoretical background required to understand the concepts and techniques employed in the study. It covers fundamental topics related to medical image segmentation, brain tumor pathology, magnetic resonance imaging, deep learning, convolutional neural networks, attention mechanisms, vision transformers, and three-dimensional transformer architectures. This chapter provides the conceptual foundation upon which the methodology is built.

Chapter 3 reviews related work around brain tumor segmentation. It discusses traditional image processing approaches, classical machine learning methods, convolutional neural network–based models, hybrid CNN–transformer architectures, and transformer-dominant segmentation frameworks. The chapter also examines benchmark studies associated with the BraTS challenge, compares existing methods in terms of accuracy and efficiency, identifies research gaps, and positions the proposed approach within existing literature.

Chapter 4 describes the methodology adopted in this research. It details the overall research design and workflow, dataset characteristics, preprocessing steps, problem formulation, and the proposed segmentation architecture. The chapter further explains the training strategy, mathematical formulation, and implementation details, enabling a clear understanding of the technical approach and supporting reproducibility.

Chapter 5 presents the experimental setup and results. It outlines the evaluation metrics, training dynamics, quantitative performance analysis, qualitative visual results, and ablation studies used to assess the effectiveness of the proposed framework. The results presented in this chapter directly address the research questions defined in Chapter 1.

Chapter 6 discusses the experimental findings in depth. It interprets the results in relation to the research objectives and existing literature, analyzes the impact of architectural choices, compares the proposed approach with state-of-the-art methods, examines computational efficiency and generalization behavior, and discusses the limitations and potential clinical implications of the study.

Chapter 7 concludes the report by summarizing the main findings and revisiting the contributions of the research. It also outlines potential directions for future work, highlighting opportunities for further methodological improvements, broader evaluation, and eventual clinical validation.

CHAPTER 2

THEORETICAL BACKGROUND

2.1 Medical Image Segmentation Fundamentals

Medical image segmentation is a fundamental process in medical image analysis that involves dividing an image into distinct regions corresponding to meaningful anatomical structures or pathological areas (Litjens, et al., 2017). The goal of segmentation is to transform raw medical images into structured representations that enable precise localization, measurement, and interpretation of regions of interest. In clinical contexts, segmentation supports a wide range of applications, including diagnosis, treatment planning, intervention guidance, and longitudinal monitoring of disease progression (Shen, Wu, & Suk, 2017). Because segmentation outcomes may directly influence clinical decisions, medical image segmentation requires a high level of accuracy, reliability, and consistency.

Segmentation plays a central role in enabling quantitative analysis of medical images. By delineating anatomical structures and pathological regions, segmentation allows clinicians and researchers to compute volumetric measurements, assess morphological changes, and derive imaging biomarkers that support objective evaluation of patient status. For example, in oncology, segmentation enables measurement of tumor size and growth dynamics, which are critical for evaluating treatment response and guiding therapeutic strategies (Ellingson, Wen, & Cloughesy, 2017). As modern imaging modalities continue to increase in resolution and complexity, the volume of image data produced in clinical practice has grown substantially, reinforcing the need for automated segmentation methods that can operate efficiently and reproducibly.

Medical image segmentation tasks are most formulated as semantic segmentation (Ronneberger, Fischer, & Brox, 2015), where each pixel or voxel is assigned a class label representing a specific tissue type or region. This formulation emphasizes the classification of image elements based on their semantic meaning rather than their

individual identity. Semantic segmentation is particularly well suited to medical imaging applications such as organ delineation and tumor segmentation, where the objective is to identify contiguous regions of the same tissue or pathology. In these cases, distinguishing between multiple instances of the same structure is often unnecessary, making semantic segmentation the preferred approach.

In contrast, instance segmentation aims to identify and separate individual instances within the same class. While this approach is widely applied in natural image analysis, its relevance in medical imaging is more limited and highly dependent on the clinical task. For many medical applications, including brain tumor segmentation, the emphasis is placed on accurate delineation of the pathological region as a whole rather than on distinguishing multiple instances (Ellingson, Wen, & Cloughesy, 2017). Consequently, instance segmentation is less commonly adopted in neuro-oncological imaging studies.

Another important aspect of medical image segmentation is the dimensionality of the data. Two-dimensional (2D) segmentation methods operate on individual image slices, treating each slice independently. This approach is computationally efficient and has been widely used due to its lower memory requirements and ease of implementation. However, 2D segmentation fails to account for spatial relationships between adjacent slices, which can lead to inconsistencies and discontinuities when reconstructing three-dimensional anatomical structures. Such limitations are particularly evident in volumetric imaging modalities, where anatomical and pathological regions extend across multiple slices.

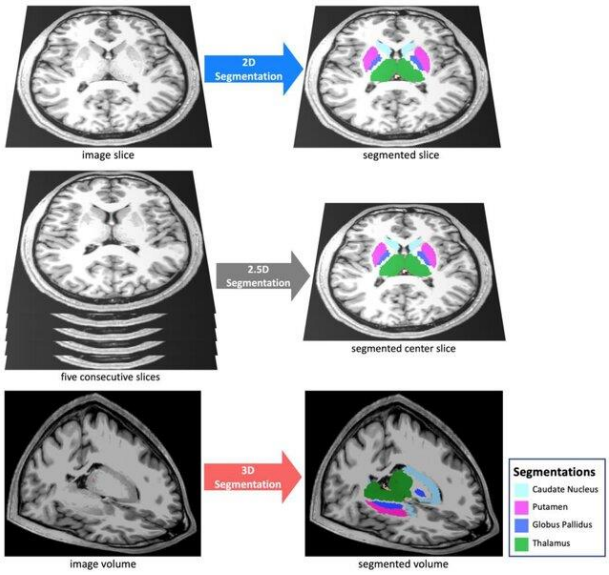


Figure 1. Compare between 2D & 3D egmentation
Source: Adapted from Çiçek et al. (2016) and Ronneberger et al. (2015).

To overcome these limitations, three-dimensional (3D) segmentation methods process entire image volumes simultaneously, allowing spatial context to be preserved across all dimensions. By modeling inter-slice relationships, 3D segmentation methods produce more spatially coherent and anatomically consistent results. This volumetric approach is especially important for modalities such as magnetic resonance imaging (MRI), where tumors and other structures exhibit complex three-dimensional morphology. However, the increased dimensionality of the data introduces significant computational challenges, including higher memory consumption and increased training complexity, which must be addressed through careful model design.

Medical image segmentation is further complicated by a range of intrinsic challenges related to image quality and biological variability. Medical images often exhibit low contrast between tissues, noise, intensity inhomogeneity, and acquisition artifacts that obscure boundaries between regions. Anatomical variability across patients further increases the difficulty of designing segmentation methods that generalize well across diverse populations. These factors require segmentation models to be robust and adaptable to variations in image appearance and structure.

Pathological regions, particularly tumors, introduce additional complexity due to their heterogeneous appearance and irregular boundaries. Tumors may infiltrate surrounding tissue, display diffuse margins, and exhibit varying intensity patterns across different imaging modalities. These characteristics make segmentation based solely on local image features unreliable and highlight the importance of incorporating broader contextual information. Accurate segmentation must therefore integrate local detail with spatial context to distinguish pathological tissue from normal anatomy effectively.

Class imbalance represents another fundamental challenge in medical image segmentation. In many cases, regions of interest such as tumors occupy only a small proportion of the image volume, while background tissue dominates. This imbalance can bias learning algorithms toward predicting background regions and reduce sensitivity to small or less prominent structures. Addressing class imbalance requires careful consideration of model architecture, training strategies, and evaluation metrics to ensure clinically relevant performance.

The process of obtaining annotated training data introduces further constraints. Manual segmentation requires specialized expertise and is time-consuming, resulting in limited

availability of labeled data. Additionally, manual annotations are subject to variability between experts, reflecting differences in interpretation and clinical judgment. This variability introduces uncertainty into reference labels and poses challenges for both model training and evaluation. Robust segmentation methods must therefore be capable of learning from imperfect and heterogeneous annotations.

These characteristics illustrate the complexity of medical image segmentation and underscore the need for advanced computational approaches capable of handling high-dimensional data, preserving spatial coherence, and maintaining robustness across diverse imaging conditions. The concepts outlined in this section form the foundational basis for the theoretical and methodological discussions presented in the subsequent sections of this chapter.

2.2 Brain Tumor Pathology and Anatomy

Brain tumors are abnormal growths of cells within the brain or its surrounding structures and represent one of the most complex pathological conditions affecting the central nervous system. Their clinical significance arises from both their biological behavior and their location within the brain, where even small lesions can disrupt critical neurological functions. Understanding brain tumor pathology and anatomical characteristics is essential for accurate interpretation of medical images and forms a fundamental basis for developing effective segmentation methods in neuro-oncological imaging.

Brain tumors are commonly classified based on their origin and biological behavior. Primary brain tumors originate from cells within the brain or its immediate surroundings, such as glial cells, neurons, meninges, or cranial nerves. Secondary brain tumors, also known as metastatic tumors, arise from cancers elsewhere in the body and spread to the brain via the bloodstream. Primary tumors are of particular interest in neuro-oncology research, as they exhibit a wide range of histological types and growth patterns that pose significant challenges for imaging-based analysis.

Among primary brain tumors, gliomas are the most prevalent and clinically significant category. Gliomas originate from glial cells, which provide structural and functional support to neurons. These tumors are typically classified according to the World Health Organization (WHO) grading system, which categorizes gliomas based on histological features and expected aggressiveness. Low-grade gliomas tend to grow more slowly and

may exhibit relatively well-defined margins, whereas high-grade gliomas, such as glioblastoma, are highly aggressive, infiltrative, and associated with poor prognosis. The grading of gliomas has direct implications for treatment planning and outcome prediction, making accurate imaging assessment critically important.

An important characteristic of gliomas is their infiltrative growth behavior. Rather than forming well-circumscribed masses, many gliomas spread diffusely into surrounding brain tissue, following white matter tracts and displacing or invading normal structures. This infiltrative nature results in indistinct tumor boundaries and gradual transitions between tumor and healthy tissue, which complicates both visual interpretation by clinicians and automated segmentation by computational models. From an imaging perspective, infiltration often manifests as subtle intensity changes rather than sharp edges, increasing ambiguity in boundary delineation.

Brain tumors are not anatomically uniform and are commonly described in terms of distinct tumor subregions, each associated with different pathological and imaging characteristics. In neuro-oncological imaging, particular attention is given to regions such as the enhancing tumor, tumor core, and peritumoral edema. The enhancing tumor typically corresponds to areas with active tumor growth and disrupted blood–brain barrier, which become visible after contrast agent administration. The tumor core may include necrotic or non-enhancing tumor tissue, reflecting regions of cell death or reduced vascularity. Surrounding these regions is often peritumoral edema, which represents fluid accumulation and tumor-induced changes in adjacent brain tissue. These subregions are clinically relevant because they influence surgical margins, radiotherapy planning, and assessment of treatment response.

The anatomical location of a brain tumor further influences its clinical presentation and imaging appearance. Tumors located in eloquent brain regions, such as areas responsible for speech, motor control, or vision, may produce symptoms even at small sizes, while tumors in less critical regions may remain asymptomatic for longer periods. Anatomical constraints imposed by the skull and rigid brain structures also affect tumor growth patterns, leading to mass effect, midline shift, and compression of surrounding tissue. These anatomical factors must be considered when interpreting imaging data and designing segmentation models that aim to capture realistic tumor morphology.

Tumor heterogeneity represents another major challenge in brain tumor pathology. Within a single tumor, different regions may exhibit varying cellular composition, vascularization, and metabolic activity. This heterogeneity leads to diverse imaging appearances across different MRI modalities and even within the same modality. As a result, segmentation methods must be capable of distinguishing subtle differences in tissue characteristics while maintaining coherence across the entire tumor structure. Failure to account for heterogeneity can lead to misclassification of tumor subregions and reduced clinical usefulness of segmentation results.

From a segmentation perspective, the pathological and anatomical characteristics of brain tumors impose strict requirements on model design. Effective segmentation methods must be sensitive to irregular shapes, diffuse boundaries, and spatial relationships between tumor subregions and surrounding brain tissue. They must also accommodate variability across patients and tumor types while maintaining robustness to noise and imaging artifacts. A thorough understanding of brain tumor pathology and anatomy is therefore essential for interpreting segmentation challenges and motivates the development of advanced modeling approaches capable of capturing both local detail and global anatomical context.

2.3 Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging (MRI) is a non-invasive medical imaging modality that plays a central role in the diagnosis and evaluation of neurological disorders, including brain tumors (Bauer, Wiest, Nolte, & Reyes, 2013). Unlike imaging techniques based on ionizing radiation, MRI exploits the magnetic properties of atomic nuclei to generate detailed images of soft tissues. Its ability to provide high contrast between different tissue types, combined with flexible acquisition protocols, makes MRI particularly well suited for neuro-oncological imaging and forms the primary data source for brain tumor segmentation tasks.

At a high level, MRI is based on the principles of nuclear magnetic resonance. When a patient is placed within a strong static magnetic field, hydrogen nuclei within the body align with the field due to their intrinsic magnetic moments. Radiofrequency pulses are then applied to perturb this alignment, causing the nuclei to absorb energy and transition

to a higher energy state. As the nuclear plants relax back to their equilibrium state, they emit radiofrequency signals that are detected by receiver coils. These signals are spatially encoded using magnetic field gradients and reconstructed into images that represent the distribution and properties of tissues within the body.

The contrast observed in MRI images arises primarily from differences in tissue-specific relaxation properties. Two fundamental relaxation processes govern signal behavior: longitudinal (T1) relaxation and transverse (T2) relaxation. T1 relaxation describes the recovery of longitudinal magnetization following excitation, while T2 relaxation characterizes the decay of transverse magnetization due to dephasing of nuclear spins. Different tissues exhibit distinct T1 and T2 relaxation times, allowing MRI to generate images with varying contrast that emphasize specific anatomical or pathological features. By adjusting acquisition parameters such as repetition time and echo time, MRI protocols can be tailored to highlight tissue characteristics.

MRI offers several advantages that are especially important for brain imaging. Its superior soft-tissue contrast enables clear differentiation between gray matter, white matter, cerebrospinal fluid, and pathological tissue. MRI also supports multi-planar imaging, allowing images to be acquired or reconstructed in axial, sagittal, and coronal planes without loss of resolution. Furthermore, MRI provides flexibility in designing acquisition sequences that capture complementary information about tissue composition, vascular properties, and pathological changes. These characteristics make MRI the preferred modality for detecting, characterizing, and monitoring brain tumors.

Despite its strengths, MRI is subject to several limitations that influence both visual interpretation and automated image analysis. MRI acquisition is relatively time-consuming compared to other imaging modalities, making it sensitive to patient motion. Even slight movement during scanning can introduce artifacts that degrade image quality and complicate segmentation. In addition, MRI images often exhibit intensity inhomogeneity caused by variations in the magnetic field and radiofrequency coil sensitivity. This phenomenon, commonly referred to as bias field distortion, results in gradual intensity variations across the image that are unrelated to underlying tissue properties.

Noise is another factor that affects MRI image quality. Thermal noise and system imperfections introduce random variations in signal intensity, which can obscure fine structural details and reduce contrast between adjacent tissues. Imaging artifacts such as ghosting, susceptibility effects, and partial volume effects further complicate interpretation. Partial volume effects occur when a single voxel contains a mixture of different tissue types, leading to ambiguous intensity values that do not correspond to a single anatomical structure. These artifacts are particularly problematic at tissue boundaries and in regions with complex geometry, such as tumor margins.

From a segmentation perspective, these limitations present significant challenges. Intensity variations and artifacts can reduce the reliability of intensity-based discrimination between tissues, while partial volume effects blur boundaries between tumor and surrounding structures. Additionally, differences in scanner hardware, acquisition protocols, and patient positioning can lead to variability in image appearance across datasets. Effective segmentation methods must therefore be robust to noise, artifacts, and intensity inconsistencies, and capable of generalizing across diverse imaging conditions.

Another important consideration is the volumetric nature of MRI data. Brain MRI scans are inherently three-dimensional, consisting of multiple slices that together represent a continuous anatomical volume. Accurate analysis requires preservation of spatial relationships across slices and consideration of context beyond individual two-dimensional images. This volumetric structure underlies the need for three-dimensional segmentation approaches that can capture spatial continuity and anatomical coherence throughout the brain.

The characteristics of MRI acquisition and image formation directly influence the design of computational models for brain tumor segmentation. Understanding the physical principles of MRI, the sources of image contrast, and the common limitations and artifacts is essential for interpreting segmentation challenges and motivates the development of advanced methods that can integrate information across modalities, scales, and spatial dimensions. These considerations form the basis for the discussion of specific MRI modalities used in brain tumor analysis in the following section.

2.4 MRI Modalities for Brain Tumor Analysis

Magnetic resonance imaging provides a range of acquisition sequences, commonly referred to as MRI modalities, each of which emphasizes different tissue characteristics through variations in contrast mechanisms. In neuro-oncological imaging, no single MRI modality is sufficient to fully characterize brain tumors due to their complex biological composition and heterogeneous appearance. Instead, multimodal MRI analysis is employed to capture complementary information about tumor structure, vascular properties, and tissue composition. Understanding the role of each modality is essential for interpreting brain tumor anatomy and for designing segmentation models capable of accurately delineating tumor subregions.

T1-weighted imaging is one of the fundamental MRI modalities and provides high-resolution anatomical detail. In T1-weighted images, tissues with short T1 relaxation times appear bright, while those with longer T1 times appear darker. This modality offers clear visualization of normal brain anatomy, including gray matter, white matter, and cerebrospinal fluid, making it valuable for assessing overall brain structure. However, untreated tumors often appear isointense or hypointense on T1-weighted images, which limits the modality's ability to distinguish tumor tissue from surrounding normal structures without additional contrast enhancement.

To enhance visualization of tumor vascularity and blood–brain barrier disruption, T1-weighted post-contrast imaging (T1ce) is widely used. In this modality, a contrast agent—typically gadolinium-based—is administered prior to image acquisition. Regions with abnormal vascular permeability, such as actively growing tumor tissue, accumulating the contrast agent and appearing hyperintense. T1ce images are particularly important for identifying enhancing tumor regions, which often correspond to areas of high cellular activity and aggressive tumor behavior. However, not all tumor components enhance uniformly, and non-enhancing tumor tissue may remain indistinguishable from surrounding structures in this modality alone.

T2-weighted imaging emphasizes differences in T2 relaxation times and is sensitive to changes in water content within tissues. In T2-weighted images, fluid-rich regions appear bright, making this modality effective for visualizing edema, cystic components, and necrotic regions associated with brain tumors. T2-weighted imaging provides important information about the spatial extent of tumor-related changes beyond the enhancing core.

However, its high sensitivity to fluid can reduce specificity, as edema and cerebrospinal fluid may exhibit similar intensity characteristics, complicating precise boundary delineation.

To address this limitation, fluid-attenuated inversion recovery (FLAIR) imaging is commonly employed. FLAIR suppresses the signal from free fluid, particularly cerebrospinal fluid, while preserving sensitivity to pathological fluid accumulation. This suppression enhances the visibility of peritumoral edema and infiltrative tumor regions adjacent to cerebrospinal fluid spaces. FLAIR imaging is especially valuable for identifying diffuse tumor infiltration that may not be apparent on contrast-enhanced images (Ellingson, Wen, & Cloughesy, 2017). Nevertheless, FLAIR images often exhibit lower spatial resolution and may contain intensity ambiguities that challenge automated analysis.

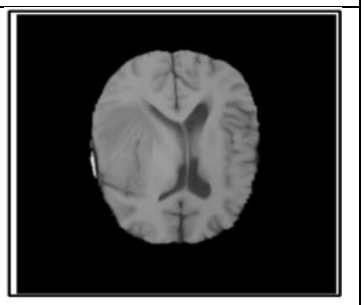
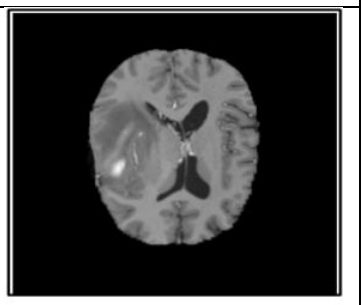
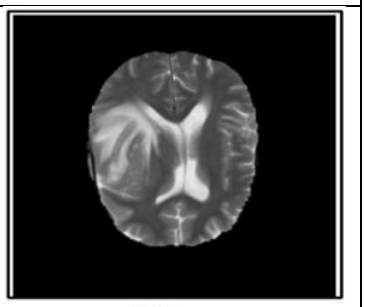
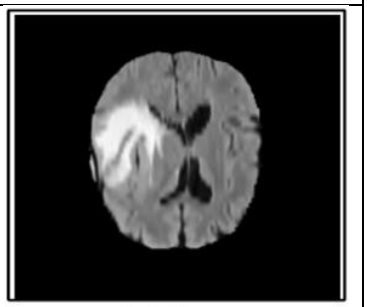
Each MRI modality contributes distinct and complementary information about brain tumor structure and pathology. T1-weighted images provide anatomical context, T1ce highlights actively enhancing tumor regions, T2-weighted images capture fluid-related changes, and FLAIR emphasizes infiltrative edema and non-enhancing tumor tissue. When considered individually, each modality offers only a partial view of tumor characteristics. When combined, however, multimodal MRI enables a more comprehensive representation of tumor heterogeneity and spatial extent (Bauer, Wiest, Nolte, & Reyes, 2013).

From a segmentation perspective, the complementary nature of MRI modalities motivates the use of multimodal input for automated analysis. Segmentation models that integrate information from multiple modalities are better equipped to distinguish between tumor subregions and surrounding healthy tissue (Louis, et al., 2016). Multimodal data fusion allows models to leverage the strengths of each modality while mitigating their individual limitations, leading to improved robustness and accuracy. At the same time, multimodal analysis introduces additional challenges, including increased data dimensionality, modality-specific noise characteristics, and the need for precise spatial alignment between modalities.

The selection and interpretation of MRI modalities therefore play a critical role in brain tumor segmentation research. A clear understanding of how each modality reflects

underlying tissue properties informs both manual annotation practices and the design of computational models. These considerations underpin the development of advanced segmentation frameworks discussed in subsequent chapters, where multimodal MRI data serve as the primary input for automated tumor delineation.

Table 1. Multimodal Brain MRI (T1, T1ce, T2, FLAIR)

MRI Modality	Acquisition Method	Recorded Information & Clinical Role	Example
T1-weighted (T1)	Acquired using short repetition time (TR) and short echo time (TE) without contrast injection	Captures detailed anatomical structure of the brain, clearly differentiating gray matter and white matter; used as a baseline reference for anatomical alignment	 T1 Tumor
T1 Post-Contrast (T1ce)	T1-weighted imaging performed after intravenous administration of a gadolinium-based contrast agent	Highlights regions with blood–brain barrier disruption, enabling clear visualization of active and enhancing tumor tissue (ET)	 T1 Contrast Enhanced
T2-weighted (T2)	Acquired using long TR and long TE, making the sequence sensitive to fluid content	Reveals edema and infiltrative tumor regions by displaying fluid-rich tissues as hyperintense areas	 T2 Tumor
FLAIR	T2-based sequence with inversion recovery to suppress cerebrospinal fluid (CSF) signal	Enhances visibility of abnormal fluid accumulation while suppressing CSF, improving delineation of tumor extent near ventricles	 Flair Tumor

2.5 Fundamentals of Deep Learning

Deep learning is a subfield of machine learning that focuses on learning hierarchical representations of data through artificial neural networks composed of multiple processing layers (Goodfellow, Bengio, & Courville, 2016). These models are designed to automatically extract increasingly abstract features from raw input data, reducing the need for manual feature engineering (Goodfellow, Bengio, & Courville, 2016). In the context of medical image analysis, deep learning has become a dominant paradigm due to its ability to model complex patterns in high-dimensional data and its capacity to learn directly from large collections of annotated images (Litjens, et al., 2017) (Shen, Wu, & Suk, 2017).

At the core of deep learning are artificial neural networks, which are inspired by the structure and function of biological neural systems. A neural network consists of interconnected layers of computational units, commonly referred to as neurons. Each neuron receives input signals, applies a weighted linear combination, and passes the result through a nonlinear activation function. By stacking multiple layers, neural networks can model complex nonlinear relationships between inputs and outputs. The depth of a network, defined by the number of layers it contains, enables the progressive transformation of low-level input features into high-level representations that are more suitable for decision-making tasks.

Training a deep neural network involves learning the parameters, or weights, that minimize the discrepancy between the network's predictions and the corresponding ground truth labels. This process is typically formulated as an optimization problem, where a loss function quantifies the error between predicted outputs and reference targets. Common loss functions in classification and segmentation tasks include cross-entropy loss and Dice-based losses, which are designed to reflect task-specific objectives such as class discrimination or spatial overlap. The choice of loss function plays a critical role in guiding the learning process and directly influences model performance, particularly in tasks with class imbalance or ambiguous boundaries.

The optimization of neural network parameters is achieved through an iterative process known as backpropagation, combined with gradient-based optimization algorithms. During backpropagation, gradients of the loss function with respect to each network parameter are computed using the chain rule of calculus. These gradients indicate the direction and magnitude by which each parameter should be adjusted to reduce the loss. Optimization algorithms such as stochastic gradient descent and its variants update the parameters based on these gradients, gradually improving the network's predictions over successive training iterations.

Activation functions introduce nonlinearity into neural networks, enabling them to model complex relationships that cannot be captured by linear transformations alone. Common activation functions include the rectified linear unit (ReLU), which promotes sparse activations and efficient optimization, as well as sigmoid and hyperbolic tangent functions, which are less frequently used in deep architectures due to issues such as vanishing gradients. The selection of activation functions affects both training dynamics and representational capacity, influencing how effectively a network can learn from data.

Deep learning models are typically trained on large datasets to achieve good generalization performance. However, in medical imaging, the availability of labeled data is often limited due to the cost and expertise required for annotation. This constraint increases the risk of overfitting, where a model learns patterns specific to the training data but fails to generalize to unseen cases. To mitigate this risk, various regularization techniques are employed, including weight decay, dropout, and data augmentation. These techniques encourage the model to learn more robust representations by limiting model complexity or introducing variability during training.

Another important aspect of deep learning is the concept of representation learning, where intermediate layers of a neural network learn features that capture relevant structure in the data. In early layers, networks typically learn low-level features such as edges or intensity gradients, while deeper layers encode more abstract concepts that are task specific. In medical image segmentation, this hierarchical feature learning enables models to integrate local texture information with broader contextual cues, which is essential for accurately delineating complex anatomical and pathological structures.

The application of deep learning to medical imaging also raises important considerations regarding computational efficiency and scalability. Deep neural networks, particularly those operating on three-dimensional data, require substantial computational resources for training and inference. Memory usage, training time, and hardware constraints influence model design choices and limit the size and complexity of feasible architecture. As a result, effective deep learning solutions for medical image segmentation must balance representational power with practical considerations related to computation and deployment.

These fundamental principles of deep learning underpin the convolutional and attention-based architectures discussed in subsequent sections. A clear understanding of neural network structure, optimization, and representation learning is essential for interpreting the design choices and performance characteristics of advanced segmentation models applied to multimodal MRI data.

2.6 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed to process grid-structured data, such as images, by exploiting spatial locality and hierarchical feature composition (Goodfellow, Bengio, & Courville, 2016). CNNs have become the dominant architecture for image analysis tasks due to their ability to learn spatially invariant features through shared weights and local connectivity. In medical image analysis, CNNs have played a central role in advancing automated segmentation, classification, and detection, particularly for modalities such as MRI and CT where spatial patterns are critical for interpretation.

The defining operation of a CNN is the convolution, which involves applying a set of learnable filters across the input image or feature map to produce new feature representations. Each filter is designed to respond to specific local patterns, such as edges, textures, or intensity transitions. By sliding the filter across the spatial dimensions of the input, the convolution operation captures local dependencies while maintaining translational equivariance, meaning that learned features can be recognized regardless of their location within the image. This property is especially valuable in medical imaging, where anatomical structures may vary in position across patients.

An important concept underlying CNNs is the receptive field, which refers to the region of the input image that influences the activation of a particular neuron in the network. In early layers, receptive fields are small and sensitive to local features, while deeper layers progressively integrate information over larger spatial regions. Through the stacking of convolutional layers and the use of pooling or strided convolutions, CNNs are able to learn hierarchical representations that combine fine-grained local details with increasingly abstract contextual information. In segmentation tasks, this hierarchical structure enables the network to capture both boundary-level cues and higher-level semantic context.

Pooling operations, such as max pooling or average pooling, are commonly used in CNN architectures to reduce spatial resolution and aggregate information. Pooling increases the effective receptive field and introduces a degree of invariance to small spatial variations, which can improve robustness to noise and minor misalignments. However, excessive pooling may lead to loss of spatial detail, which is detrimental to dense prediction tasks such as segmentation. As a result, modern segmentation architecture carefully balances pooling and resolution preservation to maintain boundary accuracy.

CNN-based segmentation models often adopt an encoder–decoder architecture, where the encoder progressively reduces spatial resolution while increasing feature abstraction, and the decoder restores spatial resolution to produce pixel- or voxel-level predictions (Ronneberger, Fischer, & Brox, 2015) (Çiçek, Abdulkadir, Lienkamp, Brox, & Ronneberger, 2016). The encoder captures high-level semantic information, while the decoder integrates this information with spatial detail to generate accurate segmentation maps. Skip connections are frequently employed to transfer fine-grained features from encoder layers to corresponding decoder layers, helping to recover spatial precision lost during downsampling. This design has become a standard paradigm in medical image segmentation due to its effectiveness in combining context and detail.

The success of CNNs in medical image segmentation can be attributed to several factors. CNNs are highly effective at learning local texture and intensity patterns, which are important for distinguishing between different tissue types. Their parameter-sharing mechanism reduces the number of learnable parameters compared to fully connected networks, improving generalization and computational efficiency. Furthermore, CNNs

can be extended to operate on three-dimensional data by using 3D convolutional kernels, allowing volumetric context to be incorporated directly into the learning process.

Despite these strengths, CNNs exhibit inherent limitations that motivate the exploration of alternative architecture. One key limitation is their reliance on local receptive fields, which can restrict the ability to model long-range spatial dependencies. Although deeper networks and larger kernels increase the receptive field, this approach often leads to increased computational cost and diminishing returns. In complex medical imaging tasks such as brain tumor segmentation, global contextual information—such as the spatial relationship between distant tumor regions or the overall anatomical layout of the brain—may be critical for accurate interpretation, yet difficult to capture effectively using purely convolutional operations.

Another limitation of CNNs relates to their handling of multi-scale information. While encoder–decoder architectures and feature pyramids address scale variability to some extent, CNNs may still struggle to integrate information across widely differing spatial scales in a flexible and adaptive manner. This can be particularly problematic for tumors that vary significantly in size and shape across patients. Additionally, CNNs may be sensitive to variations in imaging protocols and intensity distributions, requiring careful preprocessing and normalization to ensure stable performance.

In three-dimensional segmentation tasks, CNNs also face practical challenges related to computational complexity and memory usage. 3D convolutional layers significantly increase the number of parameters and operations compared to their 2D counterparts, limiting feasible input resolution and batch size. These constraints often necessitate compromises in model design, such as patch-based processing or reduced spatial resolution, which may affect segmentation accuracy and spatial coherence.

These limitations have motivated the exploration of architectures that can complement or extend the capabilities of CNNs by incorporating mechanisms for global context modeling and more flexible representation learning. Attention-based models and transformer architectures, discussed in subsequent sections, have emerged as promising alternatives that address some of the inherent constraints of convolutional approaches. Understanding the strengths and limitations of CNNs provides essential context for

appreciating the motivation behind hybrid and transformer-based segmentation models explored later in this chapter.

2.7 Attention Mechanisms

Attention mechanisms are a class of computational techniques designed to enable models to selectively focus on the most relevant parts of an input when producing an output (Vaswani, et al., 2017). Originally developed in the context of sequence modeling for natural language processing, attention mechanisms have since become central to a wide range of machine learning applications, including computer vision and medical image analysis. Their defining characteristic is the ability to model relationships between elements of an input regardless of spatial distance, making them particularly suitable for tasks that require global context understanding.

At a conceptual level, attention operates by computing a weighted combination of input features, where the weight reflects the relative importance of each feature with respect to a given query. In contrast to convolutional operations, which aggregate information from fixed local neighborhoods, attention mechanisms dynamically adapt their receptive field based on content similarity. This adaptive behavior allows attention-based models to capture long-range dependencies and contextual relationships that are difficult to model using purely local operations.

The most widely used form of attention in modern deep learning architectures is self-attention. In self-attention, the input features are transformed into three distinct representations: queries, keys, and values. For each element in the input, the query representation is compared with the key representations of all other elements to compute attention weights, typically using a similarity measure such as scaled dot-product similarity. These weights are then applied to the value representations to produce a context-aware output. Through this process, each output element incorporates information from the entire input, weighted according to relevance.

Self-attention provides several advantages over traditional convolutional operations. Because each element can attend to all other elements, self-attention enables explicit modeling of global context and long-range interactions (Vaswani, et al., 2017). This property is particularly important in medical image segmentation, where spatially distant

regions may be semantically related, such as different parts of a tumor or anatomically corresponding structures across the brain. By integrating information globally, self-attention helps reduce spatial fragmentation and improve consistency in segmentation outputs.

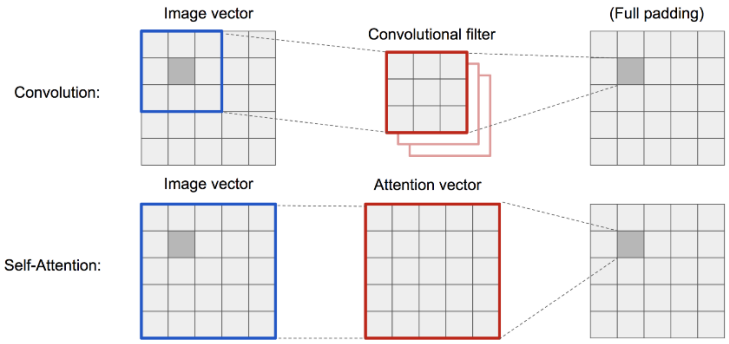


Figure 2. Self-Attention vs Convolution
Source: Adapted from Vaswani et al. (2017) and Goodfellow et al. (2016).

To enhance representational capacity, multi-head attention extends the self-attention mechanism by performing multiple attention operations in parallel. Each attention head learns to focus on different aspects of the input, such as local structure, global context, or specific feature patterns. The outputs of these heads are then concatenated and linearly transformed to produce the final representation. Multi-head attention allows models to capture diverse relationships within the data and improves robustness by distributing attention across multiple subspaces.

Despite their strengths, attention mechanisms introduce notable computational challenges. The standard self-attention operation has quadratic computational and memory complexity with respect to the number of input elements. In image-based applications, particularly for high-resolution or three-dimensional data, this complexity can become prohibitive. As a result, attention mechanisms must be carefully designed or constrained to ensure practical feasibility, especially in volumetric medical imaging where input sizes are large.

In the context of image analysis, attention mechanisms differ fundamentally from convolutional operations in how they aggregate information. Convolutions rely on fixed, local receptive fields and shared kernels, which are efficient and well-suited for capturing local patterns such as edges and textures. Attention mechanisms, on the other hand, use

data-dependent weighting to integrate information from across the entire input. This distinction highlights a complementary relationship between the two approaches: convolutions excel at local feature extraction, while attention mechanisms are effective at modeling global relationships.

The relevance of attention mechanisms to medical image segmentation lies in their ability to address limitations associated with purely convolutional models. In tasks such as brain tumor segmentation, tumors often exhibit irregular shapes, diffuse boundaries, and heterogeneous appearance across different regions. Capturing the spatial relationships between these regions requires models that can integrate information beyond local neighborhoods. Attention mechanisms provide a principled way to achieve this integration by allowing each voxel or pixel to incorporate context from distant regions of the image volume.

The increasing adoption of attention mechanisms in medical imaging has led to the development of hybrid architectures that combine convolutional layers with attention modules. These architectures aim to leverage the strengths of both approaches: convolutional layers provide efficient local feature extraction, while attention modules enhance global context modeling. Such combinations have demonstrated improved performance in segmentation tasks, motivating further exploration of attention-based designs for volumetric medical image analysis.

The fundamental properties of attention mechanisms and their computational characteristics are summarized in Table 2, highlighting key differences between attention-based and convolution-based feature aggregation.

Table 2. Comparison Between Convolutional Operations and Attention Mechanisms

Aspect	**Convolutional Operations**	**Attention Mechanisms**
Receptive field	Fixed and local	Dynamic and global
Parameter sharing	Shared kernels	Content-dependent weighting
Context modeling	Implicit and hierarchical	Explicit and direct
Computational complexity	Linear with input size	Quadratic (standard self-attention)
Strengths	Efficient local feature extraction	Long-range dependency modeling
Limitations	Limited global context	High memory and computation cost

2.8 Vision Transformers

Vision Transformers represent a class of deep learning models that adapt transformer architectures, originally developed for sequence modeling, to visual data. Unlike convolutional neural networks, which rely on local receptive fields and weight sharing to process images, vision transformers treat images as sequences of visual tokens and employ self-attention mechanisms to model relationships between these tokens (Dosovitskiy, et al., 2021). This paradigm shift enables explicit global context modeling and has introduced new perspectives on representation learning for visual tasks, including image classification and segmentation.

The core idea behind vision transformers is the transformation of an image into a sequence of fixed-size patches (Dosovitskiy, et al., 2021). Each image is divided into non-overlapping patches, which are then flattened and linearly projected into a high-dimensional embedding space. These embeddings serve as visual tokens analogous to word embeddings in natural language processing. By converting spatial image data into a sequence format, vision transformers enable the application of standard transformer encoders to visual inputs, allowing self-attention to operate across the entire image.

To preserve spatial information lost during patch flattening, positional encoding is added to the patch embeddings. Positional encodings provide the model with information about the relative or absolute position of each patch within the image, enabling the transformer to reason about spatial structure. Without positional encoding, the self-attention mechanism would treat all patches as unordered elements, which is unsuitable for visual understanding tasks that depend on spatial arrangement.

A typical vision transformer architecture consists of a stack of transformer encoder blocks, each comprising a multi-head self-attention layer followed by a feed-forward network. Residual connections and layer normalization are employed to stabilize training and facilitate gradient flow. Through repeated application of self-attention and nonlinear transformations, the model progressively refines token representations, integrating information from across the entire image. This global receptive field allows vision transformers to capture long-range dependencies and contextual relationships that may be difficult for convolutional architectures to model efficiently.

Vision transformers exhibit several properties that are particularly relevant to medical image analysis. Their ability to model global context explicitly enables more coherent interpretation of spatially distant regions, which is important in tasks such as brain tumor segmentation where pathological regions may span large portions of the image volume. Additionally, attention-based representations are inherently flexible, allowing the model to adaptively focus on salient regions rather than relying on fixed local filters. This flexibility can improve robustness in shape, size, and appearance of anatomical structures.

Despite these advantages, vision transformers also introduce important challenges. One notable limitation is their reliance on large amounts of training data to achieve strong performance. Unlike CNNs, which benefit from strong inductive biases such as locality and translation equivariance, vision transformers learn these properties implicitly through data. In domains such as medical imaging, where annotated datasets are often limited, this data dependency can hinder model generalization. To address this issue, various strategies have been proposed, including pretraining on large-scale datasets, hybrid architectures that combine convolutional and transformer components, and architectural modifications that introduce hierarchical structure.

Another challenge associated with vision transformers is computational complexity. Standard self-attention scales quadratically with the number of tokens, which becomes problematic for high-resolution images or volumetric data. As image resolution increases, the number of patches grows rapidly, leading to increased memory consumption and computational cost. This limitation has motivated the development of more efficient transformer variants and attention mechanisms that constrain attention computation while preserving representational capacity.

In segmentation tasks, vision transformers require additional architectural considerations compared to classification tasks. While early vision transformer models were designed primarily for image-level classification, segmentation demands dense, pixel- or voxel-level predictions with precise spatial alignment. As a result, segmentation-oriented vision transformer architectures often incorporate decoder modules, multi-scale feature extraction, or hybrid designs that combine convolutional encoders with transformer-based global context modeling. These adaptations aim to reconcile the global modeling strengths of transformers with the spatial precision required for segmentation.

The application of vision transformers to medical imaging has gained increasing attention due to their potential to address limitations of convolutional approaches in capturing long-range dependencies. By enabling direct interaction between distant regions of an image, vision transformers offer a promising framework for modeling complex anatomical relationships and pathological patterns. At the same time, their computational and data-related challenges necessitate careful architectural design, particularly for three-dimensional medical imaging tasks.

The concepts introduced in this section provide the foundation for understanding more specialized transformer architectures developed for visual data. In the following section, attention is focused on the Swin Transformer architecture, which introduces hierarchical design and localized attention mechanisms to improve efficiency and scalability for high-resolution image analysis.

2.9 Swin Transformer Architecture

The Swin Transformer architecture was introduced to address key limitations of early vision transformer models, particularly their high computational cost and lack of hierarchical feature representation. While standard vision transformers model global relationships through full self-attention across all tokens, this design results in quadratic complexity with respect to input size and limits scalability to high-resolution images. The Swin Transformer adopts a different strategy by combining local self-attention with a hierarchical structure, enabling efficient and scalable modeling of visual data while retaining the benefits of attention-based representation learning.

A defining feature of the Swin Transformer is its use of window-based self-attention. Instead of computing attention globally across all image patches, the input feature map is partitioned into non-overlapping local windows, and self-attention is computed independently within each window (Liu, et al., 2021). This localized attention significantly reduces computational complexity from quadratic to linear with respect to image size, making the architecture more practical for high-resolution visual tasks. At the same time, window-based attention preserves the ability to model fine-grained local interactions, which are essential for accurate spatial understanding in dense prediction tasks such as segmentation.

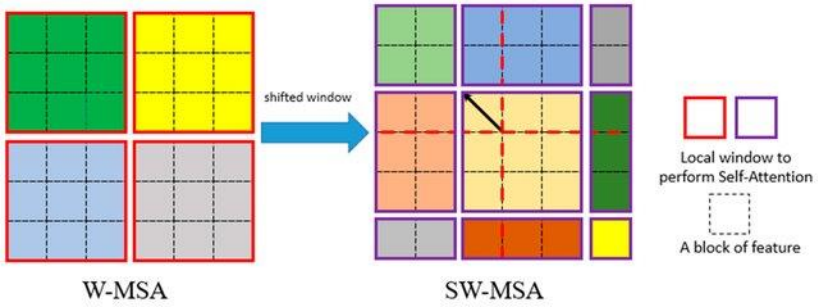


Figure 3. Shifted Window Attention Mechanism
Source: Adapted from Liu et al. (2021).

To enable information exchange across windows and avoid restricting attention to fixed local regions, the Swin Transformer introduces the shifted window mechanism. In alternating transformer layers, the window partitioning is shifted by a fixed offset, causing windows to overlap with those from the previous layer. This shift allows tokens near window boundaries to interact with tokens from neighboring windows over successive layers, effectively expanding the receptive field and enabling cross-window communication without incurring the full cost of global self-attention. Through this mechanism, the Swin Transformer achieves a balance between local efficiency and global context integration.

Another key aspect of the Swin Transformer is its hierarchical feature representation. The architecture is organized into multiple stages, each operating at a different spatial resolution (Liu, et al., 2022). Between stages, a patch merging operation reduces spatial resolution while increasing feature dimensionality, like downsampling in convolutional neural networks. This hierarchical design enables the model to learn multi-scale representations, capturing low-level details in early stages and higher-level semantic information in deeper stages. Such hierarchical features are particularly important for segmentation tasks, where both boundary precision and contextual understanding are required.

Within each stage, the Swin Transformer employs a sequence of Swin Transformer blocks, each consisting of window-based multi-head self-attention followed by a feed-forward network. Residual connections and layer normalization are used to stabilize training and facilitate gradient flow. Relative positional encoding is incorporated to encode spatial relationships within windows, allowing the model to maintain awareness

of spatial structure without relying on absolute positional embeddings. This design enhances flexibility and generalization across different input sizes.

From an efficient standpoint, the Swin Transformer offers several advantages over traditional vision transformers. By limiting attention computation to local windows and introducing hierarchical processing, the architecture scales effectively to high-resolution inputs while maintaining manageable memory usage. These properties make the Swin Transformer well suited for dense prediction tasks that require fine spatial resolution, such as image segmentation, object detection, and medical image analysis.

In the context of medical imaging, the characteristics of the Swin Transformer align well with the demands of volumetric data analysis. Medical images often contain structures of varying sizes and require integration of local detail with broader anatomical context. The hierarchical design and shifted window attention mechanism allow the Swin Transformer to model such multi-scale spatial relationships efficiently. Moreover, its modular structure facilitates adaptation to segmentation frameworks, where encoder–decoder designs and skip connections can be integrated with Swin Transformer stages.

Despite its advantages, the Swin Transformer also introduces design considerations that are particularly relevant for medical image segmentation. Window size selection influences the balance between local detail and contextual coverage, while depth and stage configuration affect computational cost and representational capacity. Additionally, while shifted windows enable cross-window interaction, long-range dependencies may still require multiple layers to be effectively captured, especially in large images or volumetric data. These considerations motivate further architectural adaptations when extending Swin Transformers to three-dimensional medical imaging tasks.

The Swin Transformer architecture thus represents an important evolution in attention-based visual modeling, combining efficiency, scalability, and hierarchical representation learning. Its design principles form the foundation for transformer-based segmentation frameworks and directly inform adaptations for three-dimensional medical image analysis discussed in the subsequent section.

2.10 3D Transformers for Medical Imaging

Three-dimensional transformers for medical imaging represent an extension of attention-based architecture designed to operate directly on volumetric data. Unlike two-dimensional approaches that process individual slices or project volumetric information into planar representations, 3D transformer models treat medical images as continuous three-dimensional entities. This design choice aligns naturally with the structure of imaging modalities such as magnetic resonance imaging (MRI) and computed tomography (CT), where anatomical structures and pathological regions extend across all spatial dimensions. The adoption of 3D transformers aims to preserve volumetric context, improve spatial coherence, and enhance the modeling of complex anatomical relationships.

A central concept in 3D transformer architecture is volumetric tokenization. Instead of dividing images into two-dimensional patches, volumetric data are partitioned into three-dimensional patches or tokens that represent local cuboidal regions of the image volume. Each token encodes information from a spatial neighborhood spanning depth, height, and width. These tokens are then embedded into a high-dimensional feature space and processed by transformer layers using self-attention mechanisms. Volumetric tokenization enables direct modeling of inter-slice dependencies and avoids the loss of contextual information that may occur when processing slices independently.

Applying self-attention to three-dimensional data introduces significant computational and memory challenges. The number of tokens grows rapidly with volumetric resolution, and standard self-attention exhibits quadratic complexity with respect to token count. This complexity makes naïve application of global attention impractical for high-resolution medical images. As a result, 3D transformer designs often incorporate efficiency-driven strategies, such as localized attention windows, hierarchical processing, or hybrid architectures that combine convolutional operations with attention mechanisms. These strategies aim to retain the benefits of global context modeling while keeping computational requirements within feasible limits.

Hierarchical 3D transformer architectures extend concepts introduced in two-dimensional vision transformers by progressively reducing spatial resolution and increasing feature abstraction across multiple stages. Early stages operate on higher-resolution tokens and capture fine-grained local details, while deeper stages integrate

broader contextual information at coarser resolutions. Patch merging or volumetric downsampling operations are commonly employed between stages to reduce computational load and enable multi-scale representation learning. This hierarchical structure is particularly important for medical image segmentation, where accurate boundary delineation and global anatomical awareness are both required.

Another key challenge in 3D transformer design is the representation of spatial position. Positional encoding schemes must be adapted to three-dimensional space to ensure that the model can reason about relative and absolute spatial relationships across the volume. Various approaches have been proposed, including absolute 3D positional embeddings and relative positional encodings that encode spatial offsets between tokens. Effective positional encoding is critical for preserving anatomical structure and maintaining spatial consistency in segmentation outputs.

In medical image segmentation, 3D transformers are often integrated into encoder–decoder frameworks (Hatamizadeh, et al., 2022). The encoder leverages attention mechanisms to learn volumetric representations that capture both local tissue characteristics and global spatial context, while the decoder progressively reconstructs dense voxel-level predictions. Skip connections between encoder and decoder stages are frequently used to preserve spatial detail and improve boundary accuracy. This architectural paradigm combines the strengths of attention-based global modeling with design principles established in convolutional segmentation networks.

Hybrid 3D architectures that combine convolutional layers with transformer modules have gained particular attention in medical imaging research (Chen, et al., 2021). In these designs, convolutional layers are used in early stages to efficiently extract local features and reduce data dimensionality, while transformer layers are employed at deeper stages to model long-range dependencies and global context. Such hybrid approaches leverage the complementary strengths of convolutional and attention-based operations and are well suited to the constraints of volumetric medical data.

Despite their potential advantages, 3D transformers remain challenging to train and deploy. Memory consumption, training time, and sensitivity to hyperparameter choices pose practical obstacles, especially when working with limited computational resources (Hatamizadeh, et al., 2022). Additionally, the availability of annotated volumetric

datasets is often restricted, increasing the risk of overfitting in large-capacity transformer models. These challenges necessitate careful architectural design, efficient attention mechanisms, and robust training strategies tailored to medical imaging applications.

The development of 3D transformers for medical imaging reflects a broader shift toward models capable of integrating local detail with global spatial reasoning in high-dimensional data. Their ability to capture volumetric context makes them particularly suitable for tasks such as brain tumor segmentation, where pathological structures are spatially complex and heterogeneous. The concepts discussed in this section provide the theoretical foundation for the transformer-based segmentation framework introduced in the methodology chapter, where these principles are applied to the design of a practical and efficient model for multimodal MRI analysis.

CHAPTER 3

LITERATURE REVIEW

The task of automated brain tumor segmentation from magnetic resonance imaging (MRI) has long been recognized as a fundamental yet challenging problem in medical image analysis. Accurate segmentation plays a pivotal role in clinical decision-making, supporting tumor diagnosis, treatment planning, surgical guidance, and longitudinal assessment of disease progression. Over the past two decades, the research landscape in this field has evolved considerably, reflecting both advances in imaging technology and shifts in dominant computational paradigms. Consequently, literature encompasses a diverse range of methodologies, each characterized by distinct assumptions, strengths, and limitations in addressing the intrinsic complexity of brain tumor appearance and anatomical variability.

Initial research efforts were largely grounded in statistical modeling and classical image processing techniques, which sought to incorporate prior knowledge of tissue distributions and spatial coherence into segmentation frameworks. These methods emphasized interpretability and computational efficiency, relying on probabilistic formulations, atlas-based representations, or handcrafted feature descriptors. While such approaches established essential theoretical foundations and demonstrated early feasibility, their performance was often sensitive to noise, imaging artifacts, and deviations from assumed anatomical norms. As a result, generalization across different scanners, acquisition protocols, and tumor subtypes remained limited, highlighting the need for more flexible and data-adaptive solutions.

The gradual availability of annotated datasets enabled the adoption of classical machine learning techniques, introducing supervised learning mechanisms into brain tumor segmentation. Methods based on classifiers such as support vector machines demonstrated improved discriminative capability compared to purely statistical models. However, these approaches remained heavily dependent on manual feature engineering and struggled to scale effectively to three-dimensional MRI volumes. Moreover, their reliance on predefined feature sets constrained their ability to capture the complex, multi-

scale characteristics of tumor regions, ultimately limiting robustness and clinical applicability.

A major paradigm shift occurred with the emergence of deep learning, particularly convolutional neural networks (CNNs), which enabled hierarchical feature learning directly from raw imaging data. Encoder–decoder architectures, multi-scale CNNs, and volumetric extensions significantly improved segmentation accuracy and became the dominant approach in benchmark evaluations such as the BraTS challenges. These models reduced reliance on handcrafted features and demonstrated superior adaptability to diverse tumor morphologies. Nevertheless, despite their success, CNN-based methods are inherently constrained by localized convolutional operations. This locality restricts their ability to model long-range spatial dependencies and global anatomical context, which are critical for consistent segmentation in high-resolution three-dimensional MRI volumes.

In response to these limitations, recent research has increasingly focused on attention mechanisms and transformer-based architectures, which leverage self-attention to explicitly model global contextual relationships. Transformer-based models and hybrid CNN–Transformer frameworks have shown promise in capturing long-range dependencies and improving volumetric consistency. Hierarchical and window-based attention mechanisms further aim to mitigate the computational burden associated with global self-attention. However, these approaches often introduce new challenges related to memory consumption, architectural complexity, and partial dependence on convolutional backbones, leaving open questions regarding efficiency and scalability in clinical settings.

Table 3 serves as a comprehensive analytical synthesis of the existing literature, systematically organizing representative studies according to their publication year, methodological approach, datasets, core technical contributions, and explicitly identified research gaps. Rather than functioning as a simple summary, the table provides a structured lens through which the progression of methodological ideas can be examined and compared. By juxtaposing contributions and limitations across different generations of methods, the table reveals recurring patterns, including the trade-off between representational power and computational efficiency, as well as the persistent challenge of balancing local detail preservation with global context modeling.

The detailed analysis supported by Table 3 underscores that, despite sustained progress, no single methodological paradigm has yet achieved a comprehensive solution to volumetric brain tumor segmentation. Limitations related to adaptability, long-range dependency modeling, and practical scalability continue to motivate ongoing research. These unresolved challenges form the basis for the research gap addressed in this thesis and provide a clear rationale for the proposed approach, which seeks to integrate efficient hierarchical attention mechanisms with robust volumetric representation learning. The subsequent sections of this chapter examine the studies summarized in Table 3 in greater detail, offering a critical discussion of their methodological innovations and limitations.

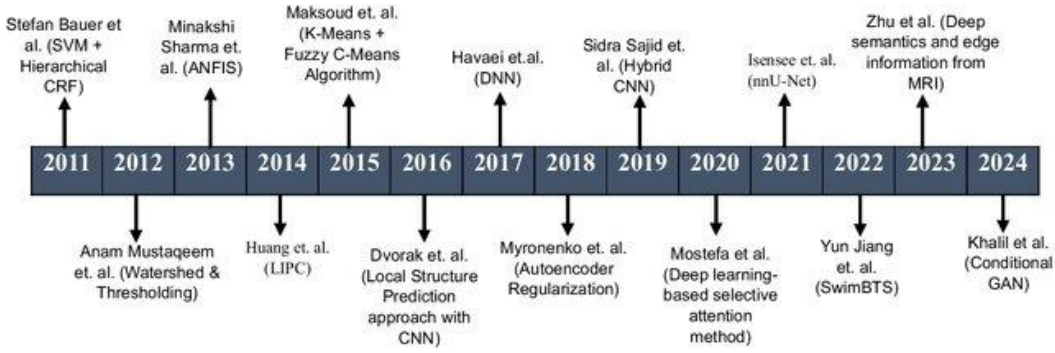


Figure 4. Evolution of brain tumor segmentation
Source: Compiled by the author based on Bauer et al. (2013), Litjens et al. (2017), Bakas et al. (2018), Isensee et al. (2021), Hatamizadeh et al. (2022), Tang et al. (2022), and Jiang et al. (2022).

Table 3. Evolution of brain tumor segmentation

No	**Paper Title**	**Authors**	**Year**	**Methodological Approach**	**Dataset(s)**	**Core Technical Contribution**	**Detailed Identified Research Gap**
1	*A Multivariate Statistical Analysis of MRI Brain Tumor Segmentation*	Zhang et al.	2001	Hidden Markov Random Field (HMRF) combined with Expectation – Maximization	Simulated and real MRI	Introduced probabilistic spatial modeling to incorporate neighborhood information into segmentation	Relies on handcrafted statistical assumptions; lacks adaptability to heterogeneous tumor appearance and complex anatomical variability

2	*Automated Brain Tumor Segmentation Using Outlier Detection*	Prastawa et al.	2004	Atlas-based modeling with EM clustering	Private MRI	Early automation of tumor segmentation using deviation from normal anatomy	Strong dependence on atlas quality; poor generalization to unseen tumor types and imaging protocols
3	*A Survey of Medical Image Segmentation Methods*	Pham et al.	2000	Classical image processing taxonomy	Multiple	Provided a foundational classification of segmentation techniques	No learning-based representation; unsuitable for large-scale volumetric MRI analysis
4	*Tumor Segmentation Using Support Vector Machines*	Bauer et al.	2013	Classical machine learning with handcrafted features	BraTS2012	Demonstrated supervised learning for tumor segmentation	Feature engineering limits robustness: scalability to 3D data remains weak
5	*U-Net: Convolutional Networks for Biomedical Image Segmentation*	Ronneberger et al.	2015	2D CNN encoder–decoder with skip connections	ISBI	Enabled end-to-end learning with precise spatial localization	Local receptive fields limit modeling of global anatomical context
6	*Deep Learning for Brain Tumor Segmentation*	Havaei et al.	2017	Multi-path CNN with local and global patches	BraTS2013	Improved contextual awareness via multi-path design	Slice-wise processing introduces inter-slice inconsistency in volumetric MRI
7	*3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation*	Çiçek et al.	2016	3D CNN encoder–decoder	EM and medical volumes	Extended CNN segmentation to full volumetric learning	High computational cost; limited long-range dependency modeling
8	*DeepMedic: Efficient Multi-Scale CNN for*	Kamnitsas et al.	2017	Dual-pathway 3D CNN	BraTS	Multi-scale receptive fields for	Context integration remains limited to

	Brain Lesion Segmentation					lesion detection	predefined scales
9	*nnU-Net: A Self-Adapting Framework for Medical Image Segmentation*	Isensee et al.	2018 / 2021	Automated CNN pipeline optimization	BraTS	Established strong reproducible CNN baseline	Architecture remains convolution-dominant; lacks explicit global attention
10	*Attention U-Net: Learning Where to Look*	Oktay et al.	2018	CNN with attention gates	Medical datasets	Improved focus on relevant anatomical regions	Attention is spatially local and auxiliary, not global self-attention
11	*UNETR: Transformers for 3D Medical Image Segmentation*	Hatamizadeh et al.	2021	Vision Transformer encoder + CNN decoder	BraTS2021	First transformer-based volumetric segmentation	Extremely high memory and computational requirements
12	*TransBTS: Multimodal Brain Tumor Segmentation Using Transformer*	Wang et al.	2021	CNN encoder with Transformer bottleneck	BraTS	Improved long-range dependency modeling	Transformer depth limited; CNN remains primary feature extractor
13	*Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation*	Cao et al.	2021	Swin Transformer (2D hierarchical attention)	Synapse	Introduced hierarchical window-based attention	Designed for 2D tasks; lacks volumetric segmentation capability
14	*Hybrid CNN–Transformer Architectures for Brain Tumor Segmentation*	Various (BraTS challenge)	2022–2023	CNN encoder + Transformer modules	BraTS	Partial integration of global attention	Trade-off between accuracy and computational scalability

2.11 Traditional Brain Tumor Segmentation

Traditional brain tumor segmentation methods refer to early computational approaches developed prior to the widespread adoption of machine learning and deep learning techniques. These methods rely primarily on handcrafted rules, mathematical models, and image processing techniques to delineate tumor regions from medical images. Although

largely superseded by learning-based approaches, traditional methods laid the foundation for automated brain tumor analysis and provided important historical context for understanding the evolution of segmentation techniques.

One of the earliest approaches to brain tumor segmentation is threshold-based segmentation, which classifies pixels or voxels based on their intensity values. In this approach, predefined or adaptive threshold values are applied to MRI images to separate tumor tissue from background brain tissue. Thresholding is computationally simple and easy to implement, making it attractive in early medical image analysis systems. However, brain tumors often exhibit overlapping intensity distributions with surrounding tissues, particularly across different MRI modalities, which limit the effectiveness of threshold-based methods. Variations in imaging protocols, noise, and intensity inhomogeneity further reduce robustness, leading to inaccurate or fragmented segmentation results.

Another widely studied traditional technique is region growing, which starts from one or more seed points placed within the tumor region and iteratively expands the region by including neighboring pixels or voxels that satisfy predefined similarity criteria. Region growing methods exploit local spatial continuity and can produce connected segmentation regions when tumor boundaries are well defined. However, their performance is highly sensitive to seed selection and similar thresholds. In cases of diffuse tumor infiltration or heterogeneous tumor appearance, region growing may either leak into surrounding healthy tissue or fail to capture the full tumor extent. These limitations reduce the reliability of region growing methods in clinical practice.

Edge-based segmentation methods attempt to detect tumor boundaries by identifying sharp intensity transitions in medical images (Bauer, Wiest, Nolte, & Reyes, 2013). Techniques such as gradient-based edge detection and contour models have been applied to brain MRI data to delineate tumor margins. While edge-based approaches can be effective when tumor boundaries are well defined, brain tumors often exhibit gradual intensity transitions rather than clear edges, particularly in infiltrative gliomas. Noise and imaging artifacts further complicate edge detection, resulting in incomplete or unstable boundary representations.

Atlas-based segmentation represents another class of traditional approaches, where anatomical templates or atlases derived from labeled datasets are registered to a target image. Tumor segmentation is then inferred based on deviations from normal anatomical structures or by transferring labels from the atlas to the target image. Atlas-based methods leverage prior anatomical knowledge and can provide consistent segmentation of normal brain structures. However, they are less effective for tumor segmentation due to the high variability in tumor location, size, and shape across patients. Deformable registration errors and the inability of atlases to represent pathological variability limit their applicability to brain tumor analysis.

Statistical and model-based approaches have also been explored in traditional brain tumor segmentation. These methods include clustering techniques such as k-means and Gaussian mixture models, which group voxels based on intensity similarity across MRI modalities. While clustering can capture some degree of tissue heterogeneity, it often lacks spatial coherence and requires post-processing to enforce anatomical consistency. Additionally, clustering methods typically assume simple statistical distributions that may not accurately represent the complex appearance of tumor tissue.

Traditional brain tumor segmentation methods are characterized by limited adaptability, reliance on handcrafted rules, and sensitivity to image quality and parameter selection (Bauer, Wiest, Nolte, & Reyes, 2013). They often struggle to generalize across patients and imaging conditions and require significant manual intervention to achieve acceptable results. Despite these limitations, traditional approaches played a crucial role in establishing core segmentation concepts and highlighted the challenges associated with brain tumor analysis, thereby motivating the development of more advanced learning-based methods.

The shortcomings of traditional segmentation techniques underscore the need for automated approaches that can learn discriminative features directly from data, integrate multi-modal information, and adapt to the heterogeneous nature of brain tumors (Bauer, Wiest, Nolte, & Reyes, 2013). These limitations paved the way for the adoption of machine learning and deep learning techniques, which are discussed in the following sections of this chapter.

3.1 Classical Machine Learning Approaches

Classical machine learning approaches represent an intermediate stage in the evolution of automated brain tumor segmentation, bridging the gap between traditional rule-based methods and modern deep learning techniques (Bauer, Wiest, Nolte, & Reyes, 2013). These approaches rely on manually designed feature representations combined with statistical learning algorithms to classify pixels or voxels into tumor and non-tumor classes. Compared to purely heuristic methods, classical machine learning techniques offer improved adaptability and data-driven decision-making, while still depending heavily on expert knowledge for feature extraction and model design.

A typical classical machine learning pipeline for brain tumor segmentation consists of several stages, including preprocessing, feature extraction, feature selection, and classification. Preprocessing steps such as intensity normalization, noise reduction, and bias field correction are applied to reduce variability in MRI data and improve feature consistency. Following preprocessing, handcrafted features are extracted from the images to characterize local tissue properties. These features may include intensity-based descriptors, texture features derived from statistical measures, gradient information, and spatial features that encode neighborhood relationships or anatomical priors.

Texture-based features have been widely used in classical segmentation methods due to their ability to capture local patterns and intensity variations associated with tumor tissue. Common texture descriptors include gray-level co-occurrence matrices, local binary patterns, and wavelet-based features. These descriptors aim to quantify properties such as homogeneity, contrast, and spatial frequency, which can differ between tumor regions and healthy brain tissue. While such features can provide discriminative information, their effectiveness depends strongly on parameter choices and the specific imaging conditions under which they are extracted.

Once features are extracted, machine learning classifiers are employed to assign class labels to pixels or voxels. Support Vector Machines (SVMs) are among the most widely used classifiers in classical brain tumor segmentation. SVMs seek to find an optimal decision boundary that separates feature vectors belonging to different classes, often using kernel functions to handle non-linear separations. SVM-based methods have demonstrated reasonable performance in distinguishing tumor tissue from normal tissue, particularly when carefully tuned and combined with informative features. However,

SVMs typically operate on individual pixels or voxels and do not inherently model spatial context, requiring additional post-processing to enforce spatial coherence.

Random Forests have also been extensively applied to brain tumor segmentation due to their robustness and ability to handle high-dimensional feature spaces. Random Forests consist of ensembles of decision trees, each trained on a random subset of the data and features. By aggregating predictions across multiple trees, Random Forests reduce overfitting and improve generalization. In medical imaging applications, Random Forests can incorporate both intensity-based and contextual features, and they offer some degree of interpretability by revealing feature importance. Despite these advantages, Random Forest-based segmentation methods still depend on handcrafted features and often struggle to capture complex, high-level patterns present in heterogeneous tumor tissue.

Other classical classifiers, such as k-nearest neighbors, Gaussian mixture models, and logistic regression, have also been explored for brain tumor segmentation. These methods vary in complexity and assumptions about data distribution, but they share a common reliance on predefined feature representations. As a result, their performance is closely tied to the quality and completeness of the extracted features, which limits their ability to generalize across diverse tumor types and imaging protocols.

A key limitation of classical machine learning approaches is their limited capacity to model hierarchical and multi-scale representations (Litjens, et al., 2017). Handcrafted features are typically designed to capture specific aspects of the image at fixed scales and may fail to represent complex interactions between local texture and global anatomical context. Additionally, feature engineering requires substantial domain expertise and often involves extensive trial and error to identify effective feature sets. This process is time-consuming and may not be able to transfer well across datasets or imaging modalities.

Another challenge associated with classical machine learning methods is their sensitivity to class imbalance, which is common in brain tumor segmentation tasks. Tumor regions often occupy a small fraction of the image volume, leading to biased classifiers that favor the background class. Although techniques such as class weighting and sampling strategies can mitigate this issue, they do not fully address the underlying limitations of feature-based learning.

While classical machine learning approaches marked a significant advancement over traditional segmentation techniques, their dependence on handcrafted features and limited representational capacity ultimately constrained their performance. These limitations motivated the transition toward deep learning methods, which can learn hierarchical feature representations directly from data and better capture the complex appearance and spatial structure of brain tumors (Litjens, et al., 2017). The following section discusses convolutional neural network–based approaches, which represent a major step forward in automated brain tumor segmentation.

3.2 CNN-Based Deep Learning Models

Convolutional neural network (CNN)–based models represent a major advancement in automated brain tumor segmentation, overcoming many of the limitations associated with traditional and classical machine learning approaches (Ronneberger, Fischer, & Brox, 2015). Unlike earlier methods that rely on handcrafted features, CNN-based models learn hierarchical feature representations directly from data, enabling more robust modeling of complex anatomical structures and pathological patterns. The success of CNNs in medical image segmentation has established them as a foundational paradigm in neuro-oncological image analysis.

One of the most influential CNN architectures for medical image segmentation is the U-Net. U-Net introduced an encoder–decoder structure with symmetric skip connections that transfer high-resolution spatial features from the encoder to the decoder. This design allows the network to combine global contextual information with fine-grained spatial detail, making it particularly effective for dense prediction tasks (Ronneberger, Fischer, & Brox, 2015). In brain tumor segmentation, U-Net-based models have demonstrated strong performance in delineating tumor regions across multiple MRI modalities. However, standard U-Net architectures are typically designed for two-dimensional inputs and process volumetric data in a slice-wise manner, which limits their ability to capture full three-dimensional spatial context.

To address this limitation, V-Net extended the U-Net concept to three-dimensional data by replacing 2D convolutions with 3D convolutions. V-Net processes entire image volumes and learns volumetric representations that preserve inter-slice continuity. This

volumetric design enables more anatomically consistent segmentation results, particularly for irregularly shaped tumors. Despite its advantages, V-Net introduces significantly higher computational and memory requirements, which restrict feasible input resolution and batch size. These constraints often necessitate compromises in network depth or spatial resolution.

Subsequent CNN-based architecture introduced residual connections and deeper network designs to improve optimization and representational capacity (Kamnitsas, et al., 2017). Residual U-Net variants incorporate residual blocks within the encoder–decoder framework, enabling deeper networks to be trained more effectively by mitigating vanishing gradient issues. These architectures improve feature reuse and stability during training, leading to improved segmentation accuracy in many studies. However, deeper CNN architecture still rely primarily on local receptive fields, which can limit their ability to capture long-range spatial dependencies.

Among CNN-based segmentation frameworks, nnU-Net represents a notable contribution due to its emphasis on automated pipeline configuration rather than architectural novelty. nnU-Net adapts preprocessing, network architecture, training schedules, and post-processing strategies to a given dataset using a set of heuristic rules. This approach has achieved strong performance across a wide range of medical segmentation tasks, including brain tumor segmentation, and has become a widely adopted baseline. Nevertheless, nnU-Net remains fundamentally convolutional in nature and inherits the inherent limitations of CNNs in modeling global context.

Although CNN-based models excel at learning local texture and boundary information, their reliance on fixed receptive fields poses challenges in tasks that require integration of global anatomical context. Brain tumors often exhibit complex spatial relationships between distant regions, such as disconnected enhancing components or diffuse infiltration patterns. Capturing these relationships requires either very deep networks or extensive pooling operations, both of which can degrade spatial resolution or increase computational cost. As a result, CNN-based models may produce fragmented segmentation results or inconsistent boundaries, particularly in heterogeneous tumor regions.

Another limitation of CNN-based segmentation approaches lies in their handling of multi-scale information. While encoder–decoder designs and skip connections provide some degree of multi-scale feature integration, CNNs may struggle to adaptively model features across widely varying spatial scales (Kamnitsas, et al., 2017). Tumor size and morphology can vary significantly across patients, and fixed-scale convolutional operations may not adequately capture this variability. These challenges have motivated research into hybrid architecture that incorporates global modeling mechanisms alongside convolutional feature extraction.

The strengths and limitations of representative CNN-based segmentation models are summarized in Table 4, highlighting key architectural characteristics and their implications for brain tumor segmentation.

Table 4. Comparison of Representative CNN-Based Brain Tumor Segmentation Models

Model	Dimensionality	Key Characteristics	Strengths	Limitations
U-Net	2D	Encoder–decoder with skip connections	Efficient, strong boundary localization	Limited volumetric context
V-Net	3D	Volumetric convolutions	Preserves inter-slice continuity	High memory and computation cost
Residual U-Net	2D / 3D	Residual blocks	Improved training stability	Still limited global context
nnU-Net	2D / 3D	Automated pipeline configuration	Strong baseline performance	No explicit global dependency modeling

CNN-based deep learning models have significantly advanced automated brain tumor segmentation by enabling end-to-end learning of discriminative features from multimodal MRI data. Their success has established convolutional architectures as a strong baseline for segmentation tasks. However, their inherent limitations in modeling long-range dependencies and global spatial relationships highlight the need for architectures that can integrate local feature learning with explicit global context modeling. These limitations have driven the development of hybrid CNN–transformer models, which are discussed in the following section.

3.3 Hybrid CNN–Transformer Models

Hybrid CNN–Transformer models emerged as a response to the complementary strengths and weaknesses of convolutional neural networks and transformer-based architectures (Chen, et al., 2021). While CNNs are highly effective at extracting local spatial features and preserving boundary information, they are limited in their ability to model long-range dependencies and global context. Conversely, transformers excel at capturing global relationships through self-attention mechanisms but often struggle with high computational cost and the absence of strong inductive biases for local feature learning. Hybrid architecture seeks to integrate these two paradigms to achieve a balance between local precision and global contextual awareness.

One of the earliest and most influential hybrid models for medical image segmentation is UNETR. UNETR employs a vision transformer as the encoder to model global context while using a convolutional decoder to reconstruct dense segmentation maps. Instead of relying on convolutional downsampling, UNETR tokenizes the input volume into patches and processes them through transformer layers, with intermediate transformer features connected to the decoder via skip connections. This design allows the model to leverage global attention while retaining spatial detail through the convolutional decoding pathway. UNETR has demonstrated improved segmentation performance compared to purely convolutional models, particularly in scenarios where long-range dependencies are important.

TransUNet represents another prominent hybrid architecture, combining a convolutional encoder with a transformer module inserted at the bottleneck (Chen, et al., 2021). In this design, CNN layers first extract low-level and mid-level features, which are then flattened and processed by a transformer to model global interactions. The transformer-enhanced features are subsequently passed to a decoder for dense prediction. This approach benefits from convolutional inductive biases in early layers while incorporating attention-based global modeling at deeper stages. TransUNet has shown strong performance in medical image segmentation tasks and highlights the effectiveness of combining convolutional feature extraction with transformer-based context modeling.

In the context of brain tumor segmentation, TransBTS extends the hybrid CNN–Transformer concept to volumetric data. TransBTS integrates a 3D CNN encoder with a transformer module designed to process volumetric features. The transformer component captures global dependencies across the three-dimensional feature space, while the CNN encoder and decoder handle local feature extraction and spatial reconstruction. This architecture demonstrates improved volumetric consistency compared to purely CNN-based models, particularly in capturing complex tumor structures. However, the transformer module in TransBTS operates on flattened feature maps, which can introduce computational overhead and limit scalability.

Hybrid CNN–Transformer models share several advantages. By combining convolutional and attention-based operations, they mitigate the locality constraints of CNNs while avoiding the full computational burden of global self-attention applied directly to raw image data (Wang, et al., 2021). These architectures also benefit from improved training stability, as convolutional layers provide strong inductive biases that facilitate learning from limited datasets, a common constraint in medical imaging. As a result, hybrid models often achieve better performance than either paradigm alone in segmentation tasks involving complex spatial relationships.

Despite their strengths, hybrid architecture also presents challenges. The integration of CNN and transformer components introduces additional architectural complexity and design choices, such as the placement of transformer modules, tokenization strategies, and feature fusion mechanisms. Flattening feature maps for transformer processing may disrupt spatial structure and require careful handling to preserve anatomical coherence. Furthermore, hybrid models may still face scalability issues when applied to high-resolution volumetric data, particularly if global attention is used without sufficient efficiency constraints.

The strengths and limitations of representative hybrid CNN–Transformer models are summarized in Table 5.

Table 5. Comparison of Representative Hybrid CNN–Transformer Segmentation Models

Model	Encoder	Transformer Role	Strengths	Limitations
UNETR	Transformer	Global context modeling	Strong global dependency capture	High memory usage
TransUNet	CNN + Transformer	Bottleneck attention	Balanced local and global features	Token flattening overhead
TransBTS	3D CNN + Transformer	Volumetric context modeling	Improved 3D consistency	Computational complexity

Hybrid CNN–Transformer architectures represent an important transitional stage in the evolution of medical image segmentation models. They demonstrate that integrating attention mechanisms into convolutional frameworks can significantly enhance segmentation performance by improving global context awareness. However, their reliance on centralized transformer modules and partial flattening of spatial representations motivates further exploration of architectures that more naturally integrate attention within a hierarchical, spatially aware framework. This motivation has led to the development of transformer-dominant architectures, which are discussed in the following section.

3.4 Transformer-Dominant Architectures

Transformer-dominant architectures represent a shift from hybrid designs toward models in which attention mechanisms constitute the primary means of feature extraction and representation learning. In contrast to hybrid CNN–Transformer approaches, where attention modules are typically confined to specific stages or bottlenecks, transformer-dominant models rely on attention throughout the encoding process. This paradigm emphasizes explicit global context modeling and flexible dependency learning as core design principles, aiming to overcome the locality constraints inherent to convolution-centric architectures.

Early transformer-dominant approaches adapted the original Vision Transformer (ViT) framework to dense prediction tasks by incorporating decoders capable of producing pixel- or voxel-level outputs. These models demonstrated that self-attention could replace convolutional feature extractors and still achieve competitive segmentation performance.

However, applying global self-attention uniformly across all tokens introduced substantial computational and memory overhead, particularly for high-resolution images and volumetric data. As a result, early transformer-dominant models were often limited to relatively low resolutions or required extensive computational resources.

To address scalability and efficiency concerns, subsequent transformer-dominant architectures introduced structural modifications that preserved the benefits of attention while reducing computational cost. A key development in this direction is the introduction of hierarchical transformers, which organize attention processing across multiple stages with progressively reduced spatial resolution. This hierarchical design enables multi-scale representation learning, allowing early layers to capture fine-grained spatial details while deeper layers integrate broader semantic context. Such designs align closely with the requirements of segmentation tasks, where both boundary precision and global coherence are essential.

Within the context of medical image segmentation, transformer-dominant architectures have demonstrated promise due to their ability to model long-range dependencies across complex anatomical structures. In brain tumor segmentation, distant tumor regions may be semantically related, and tumor subregions may exhibit global spatial patterns that are difficult to capture using purely local operations. Transformer-dominant models address this challenge by allowing each token to attend to a wide contextual neighborhood, thereby promoting spatial consistency and reducing fragmented predictions.

Swin-based segmentation architectures exemplify this transformer-dominant paradigm by embedding attention directly into a hierarchical framework. Models such as Swin-UNet and SwinUNETR employ window-based self-attention combined with shifted window mechanisms to efficiently capture both local and cross-window interactions. By maintaining spatial structure throughout the encoding process and avoiding global token flattening, these architectures preserve anatomical coherence while scaling to higher resolutions. Their transformer-centric design allows attention mechanisms to operate as the primary driver of feature learning rather than as an auxiliary component.

Transformer-dominant models also facilitate more flexible multi-scale feature integration compared to convolution-based approaches. Attention mechanisms can dynamically weight features across spatial locations and scales, enabling adaptive context aggregation

that responds to tumor size, shape, and heterogeneity. This adaptability is particularly advantageous in medical imaging, where pathological variability is high and fixed receptive fields may be insufficient to capture all relevant patterns.

Despite their advantages, transformer-dominant architectures introduce challenges that must be carefully managed. Training stability, memory consumption, and sensitivity to architectural hyperparameters remain important considerations, especially in three-dimensional settings. Additionally, the reduced inductive bias toward locality compared to CNNs can increase data requirements, necessitate careful training strategies and regularization when annotate datasets are limited. Nevertheless, continued architectural refinement and efficiency-oriented design choices have made transformer-dominant models increasingly practical for medical image segmentation.

The key characteristics of representative transformer-dominant segmentation architectures are summarized in Table 6, highlighting their design focus and implications for brain tumor segmentation.

Table 6. Comparison of Representative Transformer-Dominant Segmentation Architectures

Architecture Type	Core Design Principle	Strengths	Limitations
ViT-based segmentation	Global self-attention	Strong global context modeling	High computational cost
Hierarchical transformers	Multi-stage attention	Multi-scale representation learning	Increased architectural complexity
Swin-based models	Windowed attention with hierarchy	Efficient and scalable attention	Window size sensitivity

Transformer-dominant architecture marks a significant evolution in medical image segmentation by placing attention mechanisms at the center of representation learning. Their ability to integrate global context, adapt across scales, and maintain spatial coherence makes them well suited to complex volumetric tasks such as brain tumor segmentation. The progression toward Swin-based and hierarchical designs provides a foundation for further architectural refinement, which is examined in the following section through the lens of multi-scale feature learning and context aggregation strategies.

3.5 Multi-Scale Feature Learning

Multi-scale feature learning is a central concept in medical image segmentation, addressing the need to capture meaningful information across different spatial resolutions. In brain tumor segmentation, pathological structures can vary significantly in size, shape, and appearance across patients and tumor subtypes (Zhou, Siddiquee, Tajbakhsh, & Liang, 2018). Tumors may be presented as small enhancing foci, large infiltrative masses, or complex combinations of heterogeneous regions. Effective segmentation methods must therefore be capable of recognizing fine-grained local details while simultaneously integrating broader contextual information that reflects global anatomical structure.

Early segmentation approaches implicitly handled scale variation through fixed-size filters or handcrafted features computed at predefined resolutions. However, such methods lack adaptability and struggle to generalize across diverse tumor presentations. Deep learning models introduced more flexible mechanisms for multi-scale representation learning by stacking layers with increasing receptive fields, enabling hierarchical abstraction. As network depth increases, features transition from local texture and edge information to higher-level semantic representations that encode object-level and region-level context.

Encoder–decoder architectures play a critical role in multi-scale feature learning. In these designs, the encoder progressively reduces spatial resolution while increasing feature dimensionality, allowing the model to capture contextual information over larger receptive fields. The decoder then reconstructs high-resolution predictions by integrating coarse semantic features with spatial detail. Skip connections are commonly used to transfer fine-grained features from early encoder layers to corresponding decoder layers, preserving boundary information that may be lost during downsampling. This mechanism has proven particularly effective for medical image segmentation, where precise boundary delineation is essential.

Beyond encoder–decoder designs, feature pyramid representations have been widely explored to enhance multi-scale learning. Feature pyramids explicitly combine features extracted at multiple resolutions, enabling the model to reason about objects at different scales simultaneously. In segmentation tasks, pyramid-based approaches allow small structures to be detected using high-resolution features while larger contextual patterns

are captured using lower-resolution representations. Such designs reduce sensitivity to tumor size variability and improve robustness across cases with diverse anatomical characteristics.

Context aggregation mechanisms further extend multi-scale learning by integrating information across different spatial extents. Techniques such as pyramid pooling, atrous (dilated) convolutions, and attention-based aggregation have been proposed to increase the effective receptive field without excessive loss of resolution. In medical imaging, context aggregation is particularly important for distinguishing tumor tissue from surrounding edema or normal brain tissue, where local appearance alone may be ambiguous. By incorporating broader contextual cues, models can make more informed segmentation decisions and reduce false positives.

In transformer-based and attention-driven architecture, multi-scale feature learning takes on additional dimensions. Attention mechanisms can dynamically weight features across spatial locations and scales, allowing models to adaptively emphasize relevant context depending on tumor size and location. Hierarchical transformers naturally support multi-scale representations by processing data across multiple stages with progressively reduced resolution. This design enables attention to operate at different spatial granularities, facilitating both local detail preservation and global context integration.

Multi-scale learning is especially relevant in three-dimensional medical imaging, where spatial relationships extend across depth as well as height and width. Volumetric segmentation models must account for anisotropic resolution, slice thickness variability, and complex three-dimensional tumor morphology. Multi-scale strategies in 3D models often involve volumetric downsampling, multi-resolution feature fusion, and cross-scale attention mechanisms that integrate information across depth levels. These approaches enhance spatial coherence and improve segmentation accuracy for irregular and infiltrative tumor structures.

Despite its benefits, multi-scale feature learning introduces additional design complexity. Decisions regarding the number of scales, fusion strategies, and computational cost must be carefully balanced. Excessive multi-scale fusion can increase memory usage and training time, while insufficient scale integration may limit model expressiveness. As a

result, effective multi-scale designs aim to maximize contextual richness while maintaining computational efficiency.

Table 7. Common Multi-Scale Feature Learning Strategies in Medical Image Segmentation

Strategy	Description	Strengths	Limitations
Encoder–decoder with skip connections	Combines low- and high-level features	Strong boundary preservation	Limited adaptive scaling
Feature pyramids	Explicit multi-resolution fusion	Robust to size variation	Increased complexity
Dilated convolutions	Enlarged receptive field without pooling	Preserves resolution	Gridding artifacts
Attention-based aggregation	Dynamic context weighting	Adaptive multi-scale context	Higher computation cost

The ability to effectively integrate information across multiple spatial scales is a defining factor in segmentation performance for complex medical imaging tasks. As tumor morphology and appearance vary widely, models that incorporate robust multi-scale feature learning are better equipped to produce accurate and consistent segmentation results. These considerations directly motivate the design of advanced transformer-based frameworks that emphasize hierarchical processing and adaptive context integration, which are further examined in the subsequent sections of this chapter.

3.6 BraTS Challenge and Benchmark Studies

The Brain Tumor Segmentation (BraTS) challenge has become the most influential and widely adopted benchmark for evaluating automated brain tumor segmentation methods using multimodal magnetic resonance imaging (MRI) (Bakas, et al., 2018). Established to provide a standardized dataset, evaluation protocol, and performance metrics, the BraTS challenge has played a critical role in advancing research on brain tumor analysis by enabling fair comparison between competing approaches and fostering reproducibility across studies.

The BraTS dataset consists of preoperative multimodal MRI scans acquired from multiple institutions and scanners, reflecting realistic variability in clinical imaging conditions.

Each subject is typically provided with four MRI modalities: T1-weighted, T1 post-contrast (T1ce), T2-weighted, and fluid-attenuated inversion recovery (FLAIR). These modalities collectively capture complementary information about tumor anatomy, vascularity, and edema. Expert annotations are provided for tumor subregions, commonly defined as enhancing tumor (ET), tumor core (TC), and whole tumor (WT). This standardized labeling scheme has become a de facto reference in brain tumor segmentation research.

Over successive editions, the BraTS challenge has evolved in both dataset size and annotation quality. Early versions of the challenge focused primarily on high-grade gliomas, while later editions expanded to include low-grade gliomas and more diverse tumor presentations. Improvements in annotation protocols and quality control have enhanced the reliability of ground truth labels, making the dataset increasingly suitable for training and evaluating advanced segmentation models. This evolution has allowed researchers to assess not only segmentation accuracy but also model robustness across a broader spectrum of tumor types and imaging conditions.

A defining feature of the BraTS challenge is its standardized evaluation protocol. Segmentation performance is assessed using metrics that reflect spatial overlap and detection quality of tumor subregions. The Dice similarity coefficient is the most reported metric, measuring the overlap between predicted and reference segmentations. Additional metrics such as sensitivity and Hausdorff distance are often included to capture complementary aspects of performance, including boundary accuracy and error localization. By enforcing a consistent evaluation framework, the BraTS challenge enables meaningful comparison of methods across different studies and competition years.

Benchmark studies based on the BraTS dataset have documented the progression of segmentation techniques over time (Bakas, et al., 2018). Early benchmark results were dominated by traditional and classical machine learning methods, which achieved limited accuracy due to their reliance on handcrafted features and simplistic modeling assumptions. With the emergence of convolutional neural networks, performance improved significantly, particularly with encoder–decoder architectures that leveraged multimodal input and volumetric processing. These CNN-based models established

strong baselines and demonstrated the feasibility of end-to-end learning for brain tumor segmentation.

More recent BraTS benchmark studies highlight the increasing adoption of hybrid and transformer-based architecture. Models incorporating attention mechanisms and global context modeling have reported improved performance, particularly in challenging tumor subregions such as enhancing tumor and infiltrative edema. These studies suggest that capturing long-range dependencies and multi-scale context is essential for addressing the inherent heterogeneity and complexity of brain tumors. As a result, the BraTS challenge has become a key platform for evaluating the effectiveness of advanced architectures that extend beyond purely convolutional designs.

Another important contribution of BraTS benchmark studies is their emphasis on generalization and robustness. Because the dataset includes scans from multiple institutions with varying acquisition protocols, strong performance on BraTS is often interpreted as evidence of a model's ability to generalize across heterogeneous clinical data. This property is especially important for translational research, where models trained in controlled settings must operate reliably in real-world environments. Benchmark analyses frequently examine performance variability across cases, highlighting strengths and weaknesses of different approaches under diverse conditions.

In addition to ranking segmentation methods, BraTS benchmark studies provide valuable insights into design trends and emerging best practices. Common themes include the use of multimodal data fusion, volumetric processing, multi-scale feature integration, and increasingly sophisticated attention mechanisms. These trends reflect a broader shift toward architectures that explicitly model spatial context and hierarchical structure, aligning with the challenges posed by brain tumor segmentation.

The BraTS challenge thus serves not only as a benchmark for quantitative evaluation but also as a catalyst for methodological innovation. By offering a standardized platform for comparison, it has shaped the trajectory of research in brain tumor segmentation and established performance expectations for new approaches. The insights gained from BraTS benchmark studies directly inform the identification of research gaps and motivate the development of advanced segmentation frameworks, which are discussed in the subsequent sections of this chapter.

3.7 Comparative Analysis of Existing Methods

Comparative analysis of existing brain tumor segmentation methods provides critical insight into the strengths, limitations, and design of trade-offs that characterize different generations of approaches (Chen, et al., 2021). As segmentation techniques have evolved from rule-based methods to classical machine learning, convolutional neural networks, and transformer-based architectures, performance improvements have been accompanied by increased architectural complexity and computational demands. A structured comparison across accuracy, robustness, and efficiency dimensions is therefore essential to contextualize current research directions and motivate further methodological development.

In terms of segmentation accuracy, deep learning–based methods have consistently outperformed traditional and classical machine learning approaches on benchmark datasets such as BraTS. Traditional methods, including thresholding, region growing, and atlas-based segmentation, generally achieve limited overlap with reference annotations due to their reliance on simplistic assumptions and handcrafted rules. Classical machine learning approaches improve upon these methods by incorporating statistical learning, yet their dependence on handcrafted features constrains their ability to capture complex tumor heterogeneity and spatial context. CNN-based models represent a substantial leap forward, achieving significantly higher Dice scores by learning hierarchical representations directly from multimodal MRI data. However, performance gains tend to plateau when purely convolutional architectures are applied to challenging tumor subregions with irregular boundaries or diffuse infiltration.

Transformer-based and hybrid CNN–Transformer models have demonstrated further improvements in segmentation accuracy, particularly for tumor subregions that require integration of long-range spatial context. Attention mechanisms enable explicit modeling of global relationships, which reduces fragmented predictions and improves boundary consistency. Benchmark studies consistently report that transformer-enhanced architectures achieve higher Dice scores for enhancing tumor and tumor core regions compared to purely convolutional baselines. These improvements suggest that global context modeling is a key factor in addressing the intrinsic complexity of brain tumor segmentation.

Beyond raw accuracy, robustness and generalization are critical dimensions of comparison. Traditional and classical methods are highly sensitive to imaging conditions, parameter selection, and noise, resulting in poor generalization across datasets and scanners. CNN-based models demonstrate improved robustness due to their data-driven nature, yet they may still exhibit performance degradation when applied to cases with atypical tumor appearance or imaging artifacts. Transformer-based approaches, by virtue of their global attention mechanisms, tend to produce more spatially coherent predictions and exhibit improved robustness to local ambiguities. However, their generalization performance depends strongly on training strategies and dataset diversity, highlighting the importance of careful model design and evaluation.

Computational efficiency represents another important axis of comparison. Traditional and classical methods are computationally lightweight but lack scalability and accuracy. CNN-based models offer a favorable balance between performance and efficiency, particularly when optimized for volumetric processing. Three-dimensional CNNs, however, incur substantial memory and computation costs, limiting input resolution and batch size. Transformer-based models introduce additional computational overhead due to attention mechanisms, especially when global self-attention is applied. Recent architectural innovations, such as hierarchical processing and window-based attention, have significantly improved the efficiency of transformer models, enabling their application to high-resolution and volumetric data with manageable resource requirements.

Another important aspect of comparative analysis is the ability to model multi-scale and heterogeneous tumor characteristics. CNN-based models rely on fixed receptive fields and hierarchical abstraction to capture scale variation, which may be insufficient for tumors with extreme size variability. Hybrid and transformer-dominant architectures provide more flexible mechanisms for multi-scale context integration through attention and hierarchical processing. This flexibility is reflected in improved performance on complex tumor subregions and reduced sensitivity to scale variation.

The trade-offs between accuracy, robustness, and efficiency across different categories of segmentation methods are summarized in Table 8.

Table 8. Comparative Characteristics of Brain Tumor Segmentation Method Categories

Method Category	**Accuracy**	**Robustness**	**Computational Efficiency**	**Key Limitations**
Traditional methods	Low	Low	High	Poor adaptability, manual tuning
Classical ML methods	Moderate	Low–Moderate	Moderate	Handcrafted features, limited context
CNN-based models	High	Moderate	Moderate	Limited global dependency modeling
Hybrid CNN–Transformer	High–Very High	High	Moderate–Low	Architectural complexity
Transformer-dominant models	Very High	High	Low–Moderate (improving)	Computational cost, data dependency

Comparative studies also reveal differences in design flexibility and extensibility. Traditional and classical approaches are difficult to extend or adapt to new datasets without extensive manual intervention. CNN-based architectures offer greater flexibility through architectural modifications and transfer learning but remain constrained by convolutional inductive biases. Transformer-based models, particularly those with modular and hierarchical designs, provide greater adaptability for incorporating additional modalities, scales, or architectural enhancements. This extensibility makes transformer-based frameworks particularly attractive for ongoing research and future clinical translation.

Overall, comparative analysis of existing methods highlights a clear trajectory toward architectures that integrate global context modeling, hierarchical representation learning, and multi-scale feature aggregation. While no single approach is universally optimal, transformer-based and hybrid models represent the current state of the art in balancing segmentation accuracy and contextual awareness. The remaining challenges related to efficiency, scalability, and design complexity motivate continued research into refined transformer architectures tailored to volumetric medical imaging.

3.8 Identified Research Gaps

The body of literature reviewed in this chapter demonstrates substantial progress in brain tumor segmentation, evolving from classical statistical and machine learning–based methods to deep convolutional neural networks and, more recently, attention- and transformer-based architectures. Early approaches introduced important concepts such as spatial regularization and probabilistic modeling but relied heavily on handcrafted features and predefined assumptions regarding tumor appearance. These constraints limited their adaptability to the pronounced heterogeneity observed in multimodal MRI data, including variations in tumor shape, size, intensity, and infiltration patterns. As a result, classical methods exhibit limited robustness and generalization when applied to complex clinical cases, particularly those involving diffuse tumor boundaries and heterogeneous subregions.

Convolutional neural network–based segmentation models significantly advanced the field by enabling end-to-end feature learning and improving local boundary delineation through encoder–decoder architectures and multi-scale convolutional designs. However, the literature consistently reports that CNN-dominant models remain constrained by their inherently local receptive fields, which restrict their ability to model long-range spatial dependencies in three-dimensional MRI volumes. This limitation manifests as inter-slice inconsistencies, fragmented tumor boundaries, and reduced global anatomical coherence, especially in cases where tumor subregions extend across distant spatial locations. Although attention-gated CNNs and hybrid multi-path designs provide incremental improvements by emphasizing relevant features, their attention mechanisms are typically localized and do not offer explicit global self-attention across the entire volume, leaving the fundamental challenge of holistic volumetric context modeling unresolved.

Transformer-based and hybrid CNN–transformer frameworks represent an important shift toward explicit global context modeling through self-attention mechanisms, which are theoretically well suited for capturing long-range spatial relationships in volumetric medical imaging. Nevertheless, existing studies highlight significant practical limitations associated with these approaches, including high computational complexity, substantial memory consumption, and scalability challenges when applied to high-resolution three-dimensional MRI data. To remain computationally feasible, many transformer-based models introduce architectural compromises such as shallow attention modules, reduced

spatial resolution, or limited attention scope, which can negatively affect fine-grained boundary precision and segmentation stability. Consequently, the literature reveals a persistent research gap: current brain tumor segmentation methods have not yet achieved an effective balance between global contextual reasoning, hierarchical multi-scale feature representation, volumetric spatial consistency, and computational efficiency. Addressing this gap requires architectural solutions that integrate these requirements in a principled and scalable manner.

3.9 Positioning of the Proposed SwinME Framework

The proposed SwinME framework is situated within the existing landscape of brain tumor segmentation research as a methodologically motivated response to persistent limitations identified across prior approaches. Its design is informed by critical observations of convolutional, hybrid, and transformer-based architectures, particularly with respect to global context modeling, boundary delineation, hierarchical representation, and computational practicality in three-dimensional medical imaging.

In contrast to convolutional neural network–based methods, which primarily aggregate information through localized receptive fields, SwinME adopts an attention-centric representation learning paradigm that enables explicit modeling of spatial relationships across volumetric inputs. By embedding transformer-based attention mechanisms within a hierarchical structure, the framework facilitates interaction between spatially distant regions of the image volume during feature learning. This architectural choice directly addresses the spatial discontinuities and inter-slice inconsistencies commonly observed in convolution-dominant segmentation outputs, especially in cases involving heterogeneous or irregular tumor morphology.

SwinME is also positioned to mitigate challenges related to boundary precision and tumor subregion consistency. Existing segmentation approaches often exhibit a trade-off between preserving fine-grained boundary information and capturing global contextual structure. The hierarchical design of SwinME enables progressive abstraction while maintaining spatial resolution through structured multi-scale feature representation. This design supports more consistent delineation of tumor subregions by allowing local

boundary cues to be integrated with broader anatomical context throughout the encoding process.

With respect to computational scalability, SwinME is designed as an efficiency-aware transformer framework. Rather than relying on global self-attention applied uniformly across the entire volumetric input, the architecture employs localized attention mechanisms within a hierarchical processing scheme. This strategy substantially reduces computational and memory overhead while preserving the ability to model long-range dependencies through staged feature interaction. Consequently, SwinME is positioned as a practical alternative to earlier transformer-based models that are limited by excessive resource requirements when applied to high-resolution three-dimensional data.

The framework further distinguishes itself from hybrid CNN–Transformer architectures by adopting a unified approach to attention integration. In many hybrid models, attention mechanisms are introduced at isolated stages, resulting in fragmented feature interaction and increased architectural complexity. SwinME integrates attention consistently across hierarchical stages, enabling coherent learning across scales and reducing reliance on ad hoc fusion strategies. This unified integration aligns attention-based modeling with hierarchical representation learning in a principled manner (Bakas, et al., 2018).

From a methodological standpoint, SwinME is structured to support systematic architectural analysis. Its modular organization enables controlled examination of individual components, such as hierarchical attention stages and multi-scale enhancement mechanisms, through targeted experimentation. This structural clarity facilitates more transparent interpretation of architectural contributions and addresses the limited analytical depth observed in many existing segmentation studies.

Taken together, the SwinME framework is positioned as a coherent evolution of transformer-based medical image segmentation architectures, integrating efficient global context modeling, hierarchical multi-scale representation, and computational feasibility within a unified design. This positioning establishes a clear conceptual foundation for the methodological contributions presented in the following chapter, where the framework’s architecture and implementation are described in detail.

CHAPTER 4

METHODOLOGY

4.1 Research Design and Workflow

This research adopts an experimental and model-driven methodology to develop, implement, and evaluate a transformer-based framework for automated brain tumor segmentation from multimodal three-dimensional MRI data. The research design is structured to ensure methodological rigor, reproducibility, and fair comparison with existing segmentation approaches, while systematically validating the contribution of each architectural component within the proposed SwinME framework.

The overall workflow follows a sequential and modular pipeline, consisting of dataset preparation, preprocessing, model construction, supervised training, and quantitative as well as qualitative evaluation. Each stage of the workflow is carefully designed to isolate the impact of architectural design choices and to ensure that performance improvements can be attributed to the proposed methodology rather than variations in data handling or experimental configuration.

At the initial stage, multimodal MRI data and corresponding voxel-wise ground truth annotations are obtained from the BraTS 2023 dataset. The dataset provides standardized, pre-aligned volumetric MRI scans acquired from multiple institutions, ensuring diversity in scanner types and acquisition protocols. Each subject includes four complementary MRI modalities—T1, T1 post-contrast (T1ce), T2, and FLAIR, which are jointly utilized as multi-channel input to capture both anatomical structure and pathological characteristics of brain tumors.

Following data acquisition, a consistent preprocessing pipeline is applied to all samples. This pipeline is designed to normalize intensity distributions, enforce spatial consistency, and enhance generalization while preserving clinically meaningful tumor features. Importantly, preprocessing operations are applied identically across all experimental settings, including baseline models and ablation variants, to ensure fair and unbiased comparison. The preprocessing stage transforms raw MRI volumes into standardized inputs suitable for hierarchical transformer-based processing.

The core component of the workflow is the proposed SwinME architecture, which follows an encoder–decoder design tailored for volumetric segmentation. The encoder leverages the hierarchical attention mechanism of Swin Transformer V2 to capture both local and long-range spatial dependencies efficiently. To mitigate the loss of spatial detail during hierarchical downsampling, a multi-scale enhancement strategy is incorporated, enabling adaptive fusion of features across different resolution levels. An Enhanced Transformer (ETrans) module is integrated at the bottleneck to further strengthen global–local feature interaction before reconstruction.

The decoder progressively reconstructs high-resolution feature representations through learned upsampling operations, skip connections, and transformer-based refinement. This design ensures that fine-grained anatomical details and tumor boundaries are preserved while benefiting from global contextual awareness. The final segmentation output is produced through voxel-wise classification, generating probability maps for clinically relevant tumor subregions.

Model training is conducted using a supervised learning paradigm, where network parameters are optimized using a Dice-based loss function to directly maximize spatial overlap between predicted segmentation maps and expert annotations. Optimization is performed using the AdamW optimizer with a learning rate scheduling strategy to ensure stable convergence. Training dynamics are monitored using validation metrics to prevent overfitting and to select optimal model checkpoints.

The final stage of the workflow involves comprehensive performance evaluation. Quantitative assessment is performed using overlap-based and detection-based metrics, including Dice Similarity Coefficient and sensitivity, computed separately for whole tumor, tumor core, and enhancing tumor regions. In addition, qualitative analysis is conducted through visual inspection of segmentation outputs to assess boundary precision and anatomical consistency. Comparative experiments and ablation studies are executed under identical experimental conditions to systematically validate the effectiveness of the proposed architectural components.

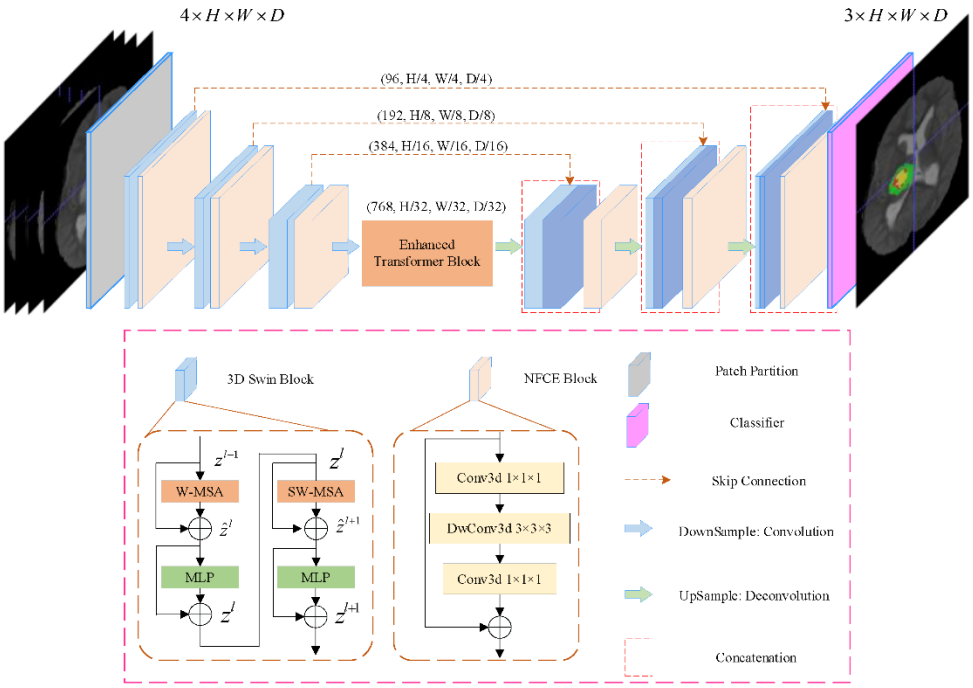


Figure 5 Overall research design and workflow of the proposed SwinME framework
Source: Adapted from Jiang et al. (2022) (SwinBTS).

The complete end-to-end workflow of the proposed methodology—from multimodal MRI input to final segmentation output is summarized in Figure 5, providing a conceptual overview of the research design and experimental pipeline employed in this study.

4.2 Dataset Description

This study employs the Brain Tumor Segmentation (BraTS) 2023 dataset (Bakas, et al., 2018) as the primary benchmark for developing and evaluating the proposed SwinME framework. The BraTS dataset is a widely recognized public dataset in the medical image analysis community and provides standardized multimodal MRI scans with expert-annotated tumor labels, enabling reliable comparison with existing state-of-the-art segmentation approaches.

The dataset consists of three-dimensional brain MRI volumes acquired from multiple clinical institutions using different scanners and acquisition protocols. This multi-institutional composition introduces significant variability in image quality, contrast, and tumor presentation, making the dataset well suited for assessing the robustness and

generalization capability of automated segmentation models. All MRI volumes are pre-aligned to a common anatomical space, allowing direct volumetric processing without additional registration steps.

Each subject in the dataset includes four complementary MRI modalities: T1-weighted (T1), T1-weighted post-contrast (T1ce), T2-weighted (T2), and Fluid-Attenuated Inversion Recovery (FLAIR). These modalities capture distinct tissue characteristics and pathological information. T1-weighted images provide detailed anatomical structure, T1ce highlights regions with blood–brain barrier disruption, T2 emphasizes fluid-rich tissues such as edema, and FLAIR suppresses cerebrospinal fluid to improve visibility of infiltrative tumor regions. In this study, all four modalities are jointly utilized and stacked channel-wise to form a multimodal input representation for the segmentation model.

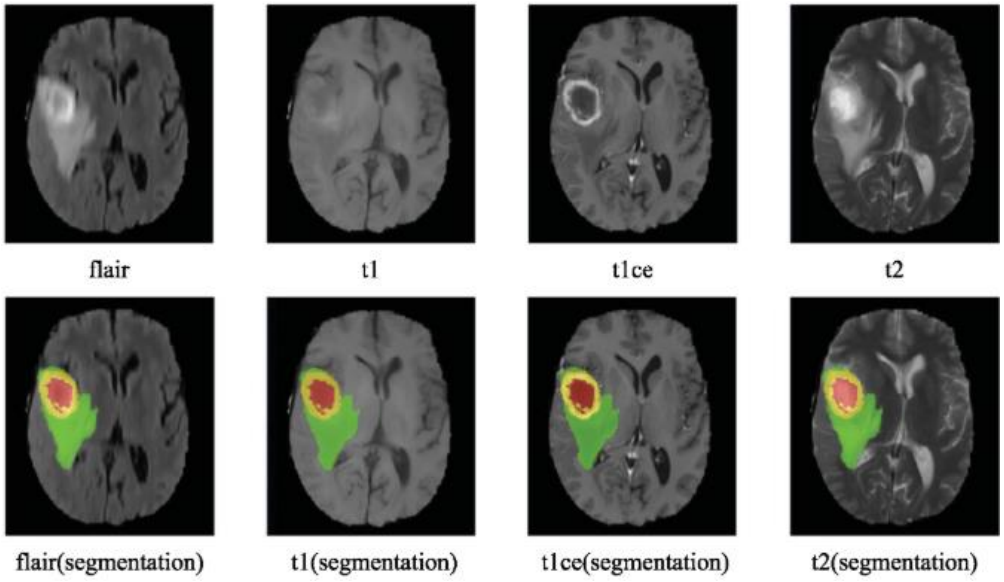


Figure 6 Example BraTS 2023 multi-modal brain MRI (FLAIR, T1, T1ce, T2) with corresponding glioma segmentation masks overlaid for tumor subregion visualization. *Source: Adapted from Bakas et al. (2018).*

Voxel-wise ground truth annotations are provided for each MRI volume by clinical experts. The annotations define tumor regions according to the BraTS labeling protocol and are used to derive three clinically relevant tumor subregions: whole tumor (WT), tumor core (TC), and enhancing tumor (ET) (Ellingson, Wen, & Cloughesy, 2017). The whole tumor region includes all tumor-related tissue, including edema and necrotic areas. The tumor core represents the union of necrotic and enhancing components, while the enhancing tumor corresponds to actively enhancing tissue visible in contrast-

enhanced scans. These subregions are treated as target classes in a multi-class segmentation formulation.

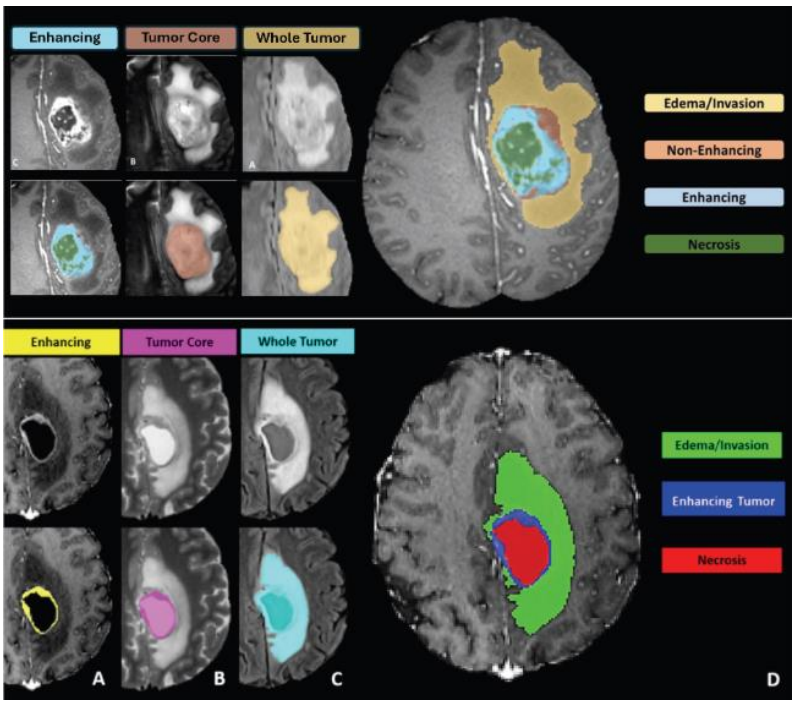


Figure 7 Visualization of glioma subregion segmentation on multi-modal brain MRI, highlighting enhancing tumor, tumor core, edema/invasion, and necrotic regions. *Source: Adapted from Bakas et al. (2018) and Ellingson et al. (2017).*

At the top, the image presents the four-class labeling system used for BraTS challenges up to 2023, including edema (ED), necrotic tumor (NCR), non-enhancing tumor (NET), and enhancing tumor (ET). Below, the image shows the revised four-class labeling system introduced in 2023, which merges NET with NCR to reduce annotation ambiguity and improve consistency. Both the top and bottom images depict the labeling system from left to right: Panel (A) shows the enhancing tumor structures visible in a T1c scan surrounding the cystic/necrotic components of the core; Panel (B) displays the tumor core visible in a T2 scan; Panel (C) illustrates the whole tumor visible in a FLAIR scan; and Panel (D) depicts the combined segmentations generating the final tumor sub-region labels. These panels apply to both the original and revised three-class labeling systems shown in the image. This streamlined labeling approach underpinned the BraTS 2017 segmentation tasks and facilitated more reliable comparisons of algorithmic performance.

For experimental consistency and reproducibility, the dataset is divided into training and validation subsets following the official BraTS 2023 protocol. The same data partitions

are used across all experiments, including baseline comparisons and ablation studies, ensuring that performance differences arise solely from architectural and methodological factors rather than variations in dataset composition.

4.3 Data Preprocessing

Data preprocessing plays a critical role in ensuring the effectiveness and stability of transformer-based segmentation models, particularly when applied to volumetric and multimodal MRI data (Litjens, et al., 2017). Unlike convolutional architectures, which exhibit strong inductive biases toward locality and translation invariance, transformer-based frameworks such as SwinME are more sensitive to variations in intensity distribution, spatial resolution, and inter-modality inconsistency. For this reason, a carefully designed preprocessing pipeline is essential to support robust representation learning and reliable segmentation performance.

The preprocessing strategy adopted in this study is designed to standardize the input data while preserving clinically relevant information. All preprocessing steps are applied consistently across training, validation, and test sets to ensure methodological fairness and reproducibility. The pipeline follows the conventions established in the BraTS challenge while incorporating design considerations specific to SwinME architecture.

4.3.1 Intensity Normalization

Magnetic Resonance Imaging data exhibits significant intensity variability across subjects due to differences in scanner hardware, acquisition protocols, and patient-specific factors. Since MRI intensities lack a standardized physical scale, direct use of raw intensity values can negatively impact model convergence and stability. To address this issue, intensity normalization is applied independently to each MRI modality. In this study, z-score normalization is used as the primary intensity normalization strategy. For each MRI volume and modality, voxel intensities are normalized by subtracting the mean and dividing by the standard deviation, computed over non-zero voxels only. Background voxels are excluded from this calculation to prevent bias introduced by large zero-valued regions outside the brain (Isensee, Jaeger, Kohl, Petersen, & Maier-Hein, 2021). This normalization ensures that each modality has a zero mean and unit variance, facilitating balanced features learning across modalities and improving numerical stability during optimization. The use of modality-wise normalization is particularly important in a multimodal setting, as it prevents any single modality from dominating the learning

process due to differences in intensity scale. Furthermore, normalized intensity distributions enable more effective attention computation within transformer layers, as attention weights become less sensitive to magnitude discrepancies across channels.

4.3.2 Data Preprocessing

Following intensity normalization, a series of preprocessing operations are applied to ensure spatial consistency and compatibility with the proposed transformer-based architecture. All MRI volumes are processed as three-dimensional tensors, preserving volumetric context throughout the pipeline.

First, MRI volumes are spatially standardized to ensure uniform voxel spacing across all subjects. Although the BraTS dataset provides pre-aligned images in a common anatomical space, minor variations in spatial resolution may still exist. To eliminate these discrepancies, all volumes are resampled to a consistent voxel resolution using trilinear interpolation for image data. Corresponding segmentation masks are resampled using nearest-neighbor interpolation to preserve discrete class labels and avoid interpolation-induced label mixing.

Next, multimodal MRI volumes are assembled into a unified representation by stacking the four modalities—T1, T1ce, T2, and FLAIR—along the channel dimension. This channel-wise stacking produces a four-channel volumetric input that enables the model to learn complementary information across modalities, such as anatomical structure, contrast enhancement, edema, and infiltrative tumor tissue.

To satisfy the fixed input size requirements of transformer-based architectures, volumes are cropped or padded to predefined spatial dimensions. This operation ensures compatibility with hierarchical patch partitioning and window-based self-attention mechanisms used in Swin Transformer V2. Cropping focuses computation on anatomically relevant brain regions, reducing unnecessary background processing and improving memory efficiency, while padding is applied where needed to maintain consistent dimensions across samples.

All preprocessing steps are performed deterministically and identically across all experimental configurations. This strict consistency ensures that observed performance differences are attributable to architectural design choices rather than preprocessing variability.

4.3.3 Cropping and Region of Interest Extraction

Cropping and region of interest (ROI) extraction are employed to reduce computational redundancy, focus model capacity on anatomically relevant regions, and ensure compatibility with the fixed input size requirements of transformer-based architectures. Brain MRI volumes typically contain large background regions that do not contribute to tumor characterization. Processing these regions unnecessarily increases memory consumption and may dilute the learning signal, particularly for attention-based models that consider global context.

In this study, ROI extraction is performed by identifying the spatial extent of the brain region and removing irrelevant background voxels. The cropping operation is designed to retain the complete brain anatomy while discarding empty or near-empty regions outside the skull. This step ensures that tumor regions—regardless of their location—are fully preserved within the cropped volume. Importantly, cropping is applied consistently across all MRI modalities and corresponding segmentation masks to maintain spatial alignment.

To support hierarchical patch partitioning and window-based attention in Swin Transformer V2, cropped volumes are further adjusted to match predefined spatial dimensions. When the extracted ROI is smaller than the target input size, symmetric padding is applied to preserve the central anatomical structure. Conversely, when the ROI exceeds the target size, centered cropping is performed to maintain spatial balance and reduce bias toward any specific anatomical direction. These operations ensure that all input volumes conform to a uniform size compatible with the encoder–decoder architecture.

The cropping strategy is carefully designed to avoid truncation of tumor regions, particularly for large or peripherally located tumors. By centering the ROI on the brain region rather than the tumor itself, the preprocessing pipeline preserves global anatomical context, which is essential for accurate segmentation of tumor subregions and surrounding edema. This approach also prevents information leakage that could arise from tumor-centered cropping strategies.

All cropping and padding operations applied to MRI volumes are mirrored exactly on the corresponding ground truth segmentation masks. Nearest-neighbor interpolation is used

for label handling to preserve discrete class boundaries. This strict alignment ensures voxel-wise correspondence between inputs and labels throughout training and evaluation.

Overall, ROI extraction and cropping serve as a crucial preprocessing step that balances computational efficiency with anatomical completeness. By reducing input dimensionality while preserving essential spatial context, this strategy supports stable training and effective utilization of attention-based mechanisms in the proposed SwinME framework.

4.3.4 Data Augmentation

Data augmentation is employed to enhance model generalization and mitigate overfitting caused by the limited availability of annotated medical imaging data (Ronneberger, Fischer, & Brox, 2015). In brain tumor segmentation, the diversity of tumor appearance, size, and location across patients is substantial, while the number of labeled samples remains relatively constrained. To address this imbalance, controlled and anatomically plausible data augmentation strategies are applied during training.

In this study, augmentation operations are applied exclusively to the training set, while validation data remain unaltered to ensure unbiased performance evaluation. All augmentation transformations are applied identically to the input MRI volumes and their corresponding ground truth segmentation masks to preserve voxel-wise correspondence.

The augmentation strategy primarily consists of spatial transformations, which are selected to reflect realistic variations in patient positioning and anatomical orientation. These include random flipping along sagittal, coronal, and axial axes, allowing the model to learn invariance to left–right and anterior–posterior orientation differences. In addition, random spatial cropping is employed within valid brain regions, enabling the model to encounter tumors at varying spatial locations while maintaining anatomical plausibility.

No intensity-based augmentation, such as contrast scaling or noise injection, is applied in this work. This decision is motivated by the clinical importance of preserving intensity relationships across MRI modalities, particularly for contrast-enhanced sequences where intensity values are directly linked to pathological interpretation. Avoiding artificial intensity perturbations ensures that the model learns meaningful modality-specific patterns rather than artifacts introduced by augmentation.

The augmentation parameters are carefully selected to balance diversity and realism. Excessive or aggressive transformations are avoided, as they may distort anatomical structures or introduce implausible configurations that could negatively impact model learning. Instead, moderate augmentation is employed to enrich the training distribution while maintaining fidelity to real-world imaging conditions.

By increasing the effective diversity of the training data, the applied augmentation strategy improves the robustness of the proposed SwinME framework to variations in tumor morphology and spatial configuration. This is particularly beneficial for transformer-based architectures, which benefits from exposure to diverse spatial patterns to learn reliable attention representations.

4.3.5 Label Preparation and Encoding

Accurate preparation and encoding of ground truth labels are essential for effective supervised learning in voxel-wise segmentation tasks. In brain tumor segmentation, the complexity of tumor structure and the hierarchical relationship between subregions require careful label handling to ensure meaningful optimization and reliable evaluation.

The BraTS dataset provides voxel-level annotations following a standardized labeling protocol. These annotations are processed to derive three clinically relevant tumor subregions: whole tumor (WT), tumor core (TC), and enhancing tumor (ET). The whole tumor region includes all tumor-related tissue, encompassing edema, necrotic regions, and enhancing components. The tumor core represents the union of necrotic and enhancing tumor tissue, while the enhancing tumor corresponds exclusively to actively enhancing regions visible in contrast-enhanced scans. This hierarchical structure reflects clinically meaningful tumor organization and guides both training and evaluation.

During preprocessing, label volumes undergo the same spatial transformations applied to the MRI volumes input, including resampling, cropping, and padding. To preserve discrete class boundaries and prevent interpolation-induced artifacts, nearest-neighbor interpolation is used for all label operations. This ensures voxel-wise alignment between inputs and labels throughout the pipeline.

For training, the processed label maps are encoded into a multi-class voxel-wise representation compatible with Dice-based loss computation. Each voxel is assigned to a class label corresponding to one of the tumor subregions. This encoding enables the

model to learn class-specific features while accounting for the hierarchical nature of the segmentation task. Background voxels are implicitly excluded from loss computation to focus optimization on tumor-related regions and to mitigate the influence of large non-tumor areas.

Given the pronounced class imbalance inherent in brain tumor segmentation—particularly for the enhancing tumor region, this encoding strategy supports stable gradient propagation during training. By explicitly representing each tumor subregion, the model can optimize overlap-based objectives more effectively, leading to improved delineation of small and heterogeneous tumor components.

The prepared and encoded labels serve as the reference standard for both optimization and evaluation. Consistent label handling across all experiments ensures that performance differences observed in subsequent analyses accurately reflect the impact of architectural and methodological choices rather than inconsistencies in ground truth processing.

4.3.6 Preprocessing Consistency and Reproducibility

Ensuring preprocessing consistency and experimental reproducibility is a fundamental requirement for reliable evaluation and fair comparison in medical image segmentation research. In this study, strict measures are adopted to guarantee that all preprocessing operations are applied uniformly across datasets, experimental configurations, and model variants.

All preprocessing steps, including intensity normalization, spatial resampling, cropping, data augmentation, and label encoding are defined deterministically and implemented using fixed parameter settings. These steps are applied identically to all data splits, including training and validation sets, as well as to all baseline models and ablation variants. By enforcing a single, shared preprocessing pipeline, the study ensures that observed performance differences are attributable solely to architectural and methodological design choices rather than variations in data handling.

To maintain consistency across experiments, preprocessing parameters such as voxel spacing, target spatial dimensions, normalization strategy, and augmentation operations are fixed prior to training and are not modified during experimentation. Randomized operations, including data augmentation, are controlled through fixed random seeds to ensure reproducibility across multiple runs. This controlled randomness enables

consistent experimental outcomes while preserving the benefits of data diversity during training.

Furthermore, preprocessing operations are integrated directly into the experimental workflow, ensuring that data transformation is performed in a traceable and repeatable manner. All transformations applied to input MRI volumes are mirrored exactly on the corresponding segmentation labels, preserving voxel-wise alignment throughout the pipeline. This strict correspondence eliminates potential sources of error arising from misalignment or inconsistent transformations.

The reproducibility of the preprocessing pipeline is further supported by comprehensive documentation of all preprocessing steps and parameter configurations. This documentation enables independent replication of the experimental setup and facilitates future extensions of the proposed framework. By adhering to these reproducibility principles, the study establishes a robust foundation for objective performance evaluation and meaningful comparison with existing methods.

The emphasis on preprocessing consistency and reproducibility strengthens the scientific validity of the experimental results and ensures that conclusions drawn from the study are both reliable and verifiable.

4.4 Problem Formulation

The task addressed in this study is formulated as a supervised, multi-class volumetric segmentation problem, where the objective is to assign a semantic label to each voxel in a three-dimensional multimodal MRI volume (Ronneberger, Fischer, & Brox, 2015). Given the complexity and heterogeneity of brain tumors, the formulation explicitly accounts for multiple tumor subregions and leverages volumetric context to ensure anatomically consistent predictions.

Let $\mathcal{X} \in \mathbb{R}^{H \times W \times D \times M}$ denote a preprocessed multimodal MRI volume, where H, W, and D represent the spatial dimensions of the volume and M denotes the number of imaging modalities. In this study, $M = 4$, corresponding to T1, T1ce, T2, and FLAIR modalities. Each voxel x_i in the volume is represented by a modality-wise feature vector capturing complementary anatomical and pathological information.

The corresponding ground truth annotation is denoted as

$$\mathcal{Y} \in \{0,1,\dots,C-1\}^{H\times W\times D},$$

where C is the number of segmentation classes. In the context of the BraTS dataset, the segmentation task is defined over three clinically relevant tumor subregions—enhancing tumor (ET), tumor core (TC), and whole tumor (WT)—with an additional background class. Thus, the problem is formulated as a four-class voxel-wise classification task.

The goal of the segmentation model is to learn a mapping function

$$f_\theta : \mathcal{X} \rightarrow \hat{\mathcal{Y}},$$

where f_θ represents the SwinME framework parameterized by θ, and $\hat{\mathcal{Y}}$ denotes the predicted segmentation volume. The output of the model is a dense voxel-level prediction in which each voxel is assigned a probability distribution over the C classes.

From an optimization perspective, the learning objective is to minimize a segmentation loss function that measures the discrepancy between the predicted segmentation $\hat{\mathcal{Y}}$ and the ground truth labels $\mathcal{Y}$. Due to the inherent class imbalance in brain tumor segmentation, tumor regions occupy a small fraction of the total volume-the loss formulation prioritizes overlap-based metrics that emphasize accurate delineation of minority classes. Accordingly, the Dice-based loss is employed as the primary optimization objective, as it directly reflects spatial overlap between predicted and reference segmentations and is robust to class imbalance.

The segmentation task is inherently volumetric, and the formulation explicitly preserves three-dimensional spatial structure. Unlike slice-wise segmentation approaches, the SwinME framework processes the full volume (or volumetric patches) as input, allowing the model to exploit inter-slice dependencies and global anatomical context. This volumetric formulation is essential for maintaining spatial continuity across slices and for accurately capturing the complex three-dimensional morphology of brain tumors.

Within the SwinME framework, the problem formulation is further shaped by hierarchical and multi-scale representation learning. The model implicitly learns feature representations at multiple spatial resolutions through its hierarchical transformer-based encoder. This enables the segmentation function f_θ to integrate local voxel-level information with broader contextual cues, addressing challenges related to boundary

ambiguity and heterogeneous tumor appearance. The formulation thus supports adaptive context aggregation across scales, which is critical for robust segmentation of both small enhancing regions and large infiltrative tumor areas.

So, the problem addressed in this study is defined as a supervised, multi-class, three-dimensional segmentation task over multimodal MRI volumes. The formulation emphasizes volumetric consistency, class imbalance awareness, and multi-scale contextual modeling, all of which align with the architectural design and objectives of the SwinME framework. This formal definition provides the foundation for the architectural details and optimization strategies described in the subsequent sections of this chapter.

4.5 Proposed SwinME Architecture

The proposed SwinME architecture is a transformer-enhanced three-dimensional segmentation framework designed to address the intrinsic challenges of volumetric brain tumor segmentation from multimodal MRI data. Brain tumors exhibit complex spatial characteristics, including irregular shapes, heterogeneous internal composition, and varying relationships with surrounding anatomical structures. Accurate segmentation therefore requires a model capable of capturing both fine-grained local details, such as tumor boundaries, and long-range contextual information spanning the entire brain volume. The SwinME architecture is specifically designed to meet these requirements by combining hierarchical transformer-based representation learning with an encoder–decoder structure inspired by three-dimensional U-Net architectures.

At a structural level, SwinME follows the well-established encoder–decoder paradigm commonly used in medical image segmentation. The encoder progressively transforms the input data into increasingly abstract feature representations, while the decoder reconstructs high-resolution segmentation maps through successive upsampling stages. The input to the network is a four-channel three-dimensional MRI volume composed of T1, T1 post-contrast (T1ce), T2, and FLAIR modalities. These modalities provide complementary information, capturing anatomical structure, contrast enhancement, edema, and infiltrative tumor tissue, respectively. By processing these modalities jointly, the architecture is able to learn richer and more discriminative representations of tumor subregions than would be possible with single-modality input.

While the overall structural philosophy of SwinME resembles that of a 3D U-Net, its internal feature extraction mechanism differs fundamentally from convolution-dominant architectures. Traditional 3D U-Net models rely on convolutional layers with limited receptive fields, which expand gradually through pooling operations. Although effective for local feature extraction, such designs can struggle to model long-range spatial dependencies that are crucial for understanding global tumor context. SwinME addresses this limitation by replacing the conventional convolutional encoder with a hierarchical transformer-based encoder built upon Swin Transformer V2, enabling explicit modeling of interactions between spatially distant regions within the volume.

The encoding process begins by partitioning the input volume into non-overlapping three-dimensional patches. Each patch aggregates local spatial information across all modalities and is projected into a latent embedding space, forming a set of volumetric tokens. These tokens are arranged into a structured three-dimensional grid and processed through multiple stages of Swin Transformer V2 blocks. Within each stage, self-attention is computed locally within fixed-size three-dimensional windows. This window-based self-attention mechanism significantly reduces computational complexity compared to global self-attention, making it feasible to apply transformer-based models to high-resolution volumetric data.

A critical component of the Swin Transformer design employed in SwinME is the shifted window mechanism. Rather than computing attention over the same fixed window partitions in every layer, successive transformer layers apply a cyclic shift to the window configuration. This operation, combined with masked self-attention, allows tokens near window boundaries to interact with tokens in neighboring windows across layers. As illustrated in the window partitioning and cyclic shift figure, this mechanism enables effective cross-window information exchange while preserving computational efficiency. Over multiple layers, the receptive field of each token gradually expands, allowing the model to capture long-range contextual relationships without incurring the quadratic cost of global attention.

As the encoder depth increases, patch merging operations are applied to reduce spatial resolution while increasing feature dimensionality. This hierarchical transformation mirrors the multiscale processing strategy of U-Net architectures and enables the model to represent information at multiple levels of abstraction. Early encoder stages focus on

fine-grained spatial details, such as tissue boundaries and local texture variations, while deeper stages capture high-level semantic information, including overall tumor extent and spatial relationships between tumor subregions. This hierarchical design is particularly well suited to brain tumor segmentation, where clinically relevant features exist across a wide range of spatial scales.

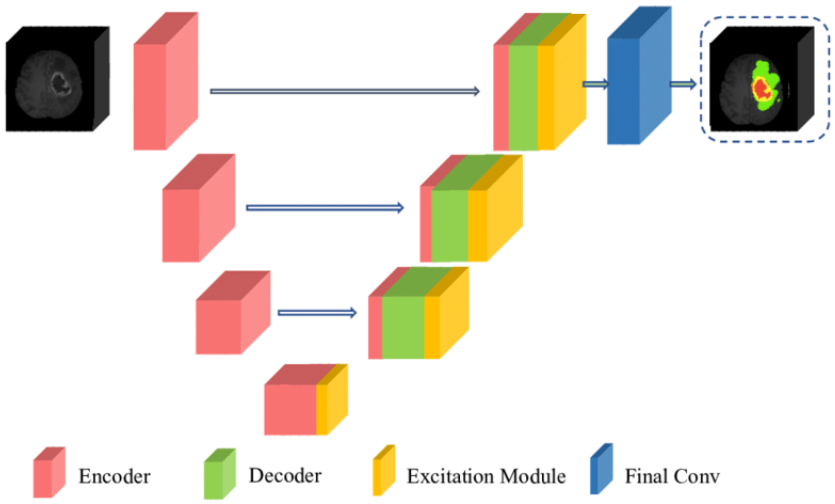


Figure 8 3D U-Net Architecture
Source: Adapted from Çiçek et al. (2016).

To further strengthen information flow across these scales, SwinME incorporates a multi-scale enhancement strategy that goes beyond standard skip connections. In conventional encoder–decoder architectures, skip connections directly concatenate encoder features with decoder features at corresponding resolutions. While effective, this approach may not fully address the imbalance between dominant large structures and smaller, heterogeneous regions such as the enhancing tumor. SwinME therefore enhances and recalibrates multi-scale features before fusion, ensuring that important fine-scale information is preserved and effectively integrated during reconstruction. This enhancement mechanism plays a crucial role in improving boundary continuity and reducing fragmentation in predicted segmentation maps.

At the bottleneck of the network, where spatial resolution is lowest and semantic abstraction is highest, SwinME introduces an Enhanced Transformer module. Operating at this stage allows the model to apply additional attention-based refinement at relatively low computational cost. The Enhanced Transformer module strengthens global contextual coherence and refines semantic representations before they are passed to the

decoder. This refinement is particularly beneficial for resolving ambiguities in tumor core and enhancing tumor regions, where subtle intensity variations and complex spatial relationships are present.

The decoder mirrors the hierarchical structure of the encoder and progressively reconstructs high-resolution feature maps through learned upsampling operations. Skip connections link corresponding encoder and decoder stages, allowing spatial detail lost during downsampling to be recovered. Unlike purely convolutional decoders, the SwinME decoder leverages the context-rich representations produced by the transformer-based encoder and bottleneck module. As a result, global contextual information guides local reconstruction, leading to improved anatomical consistency and more accurate delineation of tumor boundaries.

The interaction between encoder features, enhancement modules, and decoder pathways is illustrated in the encoder–decoder architecture figure. Enhanced feature pathways and excitation mechanisms facilitate effective information flow across the network, ensuring that both local detail and global context contribute to the final prediction. The decoder output is passed through a voxel-wise classification layer, followed by a softmax operation, to produce multi-class probability maps corresponding to clinically relevant tumor subregions.

Although SwinME retains the conceptual strengths of 3D U-Net architectures—such as multiscale feature extraction and skip connections—it significantly extends this paradigm by integrating hierarchical transformer-based attention, shifted window self-attention, explicit multi-scale enhancement, and bottleneck refinement. This hybrid design enables SwinME to overcome the locality limitations of convolutional models while maintaining the spatial precision required for medical image segmentation.

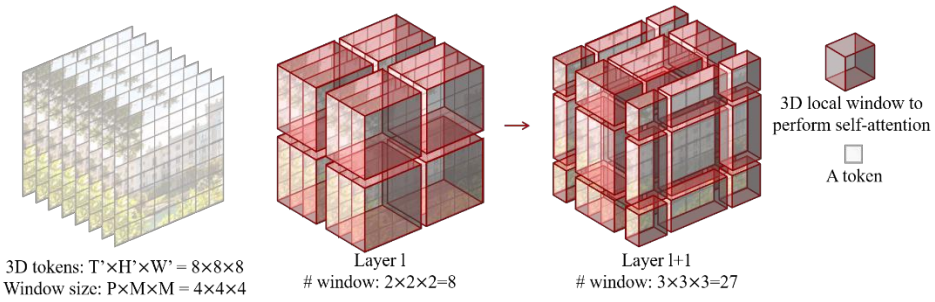


Figure 9 3D window-based self-attention mechanism
Source: Adapted from Liu et al. (2021) and Tang et al. (2022).

The proposed SwinME architecture provides a unified and scalable framework for volumetric brain tumor segmentation. By effectively balancing global contextual modeling, fine-grained spatial detail preservation, and computational efficiency, SwinME is well suited for accurate and robust segmentation across diverse tumor subregions and imaging conditions, forming the architectural foundation for the experimental results presented in subsequent sections.

4.5.1 Encoder Design

The encoder of the SwinME architecture is designed to extract rich hierarchical representations from multimodal 3D MRI volumes while efficiently modeling long-range spatial dependencies. It is based on Swin Transformer V2, which introduces window-based self-attention with a shifted window mechanism to balance global context modeling and computational efficiency (Liu, et al., 2022).

The encoder receives as input a four-channel volumetric MRI tensor corresponding to the stacked T1, T1ce, T2, and FLAIR modalities. This input is first transformed into a sequence of non-overlapping three-dimensional patches through a patch embedding operation. Each patch is linearly projected into a feature embedding space, forming the initial token representation used by the transformer blocks.

The encoder is organized into multiple hierarchical stages, each consisting of several Swin Transformer V2 blocks followed by a patch merging operation. Within each stage, window-based self-attention is applied locally within fixed-size windows to reduce computational complexity. To enable cross-window information exchange, successive transformer blocks employ a shifted window strategy, allowing tokens to attend to neighboring windows without incurring the cost of global self-attention.

As the encoder progresses through successive stages, the spatial resolution of feature maps is gradually reduced while the feature dimensionality is increased. This hierarchical representation enables the model to capture fine-grained local details at early stages and more abstract, global contextual information at deeper stages. Such multi-level feature extraction is particularly important for brain tumor segmentation, where tumor subregions vary significantly in size and appearance.

To enhance feature quality during encoding, the SwinME framework integrates multi-scale enhancement mechanisms alongside the encoder stages. These mechanisms reinforce the interaction between features extracted at different resolutions, improving spatial coherence and reducing information loss associated with downsampling. By strengthening feature consistency across scales, the encoder becomes more effective at representing both large tumor extents and small, heterogeneous regions such as the enhancing tumor.

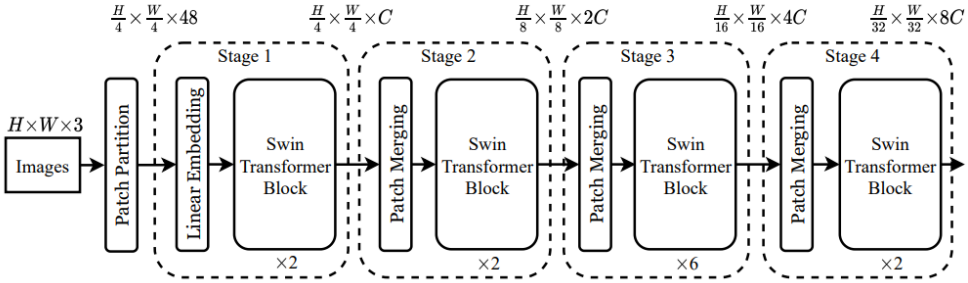


Figure 10 Overall architecture of the Swin Transformer
Source: Adapted from Liu et al. (2021).

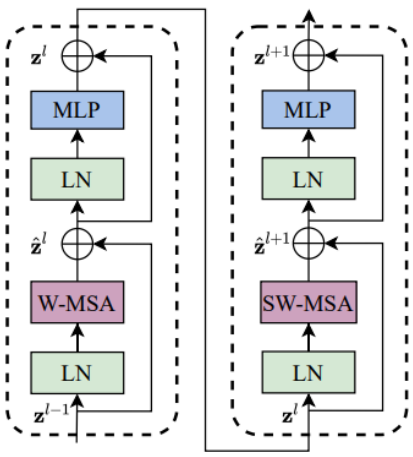


Figure 11 Structure of a Swin Transformer block with window-based self-attention
Source: Adapted from Liu et al. (2021).

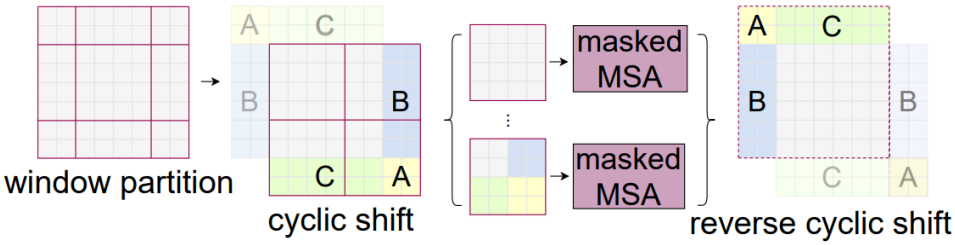


Figure 12 Illustration of window partitioning and shifted window mechanism in Swin Transformer
Source: Adapted from Liu et al. (2021).

The output of the final encoder stage represents a compact yet context-rich volumetric feature representation, which is passed to the enhanced transformer module at the bottleneck. Skip connections are established between corresponding encoder and decoder stages to preserve spatial detail and support accurate reconstruction during decoding.

4.5.2 Decoder Design

The decoder of the SwinME framework is designed to transform the hierarchical, attention-based representations produced by the encoder into dense voxel-level segmentation outputs. Its primary objective is to recover spatial resolution while integrating multi-scale contextual information, enabling accurate delineation of tumor subregions with precise boundary localization. The decoder architecture is explicitly tailored to complement the hierarchical transformer encoder and to preserve volumetric coherence throughout the reconstruction process.

The decoder follows a progressive upsampling strategy, in which feature representations are gradually restored to the original spatial resolution of the input volume. Starting from the lowest-resolution, highest-semantic-level features produced by the deepest encoder stage, the decoder performs a sequence of upsampling operations that increase spatial resolution while reducing feature dimensionality. This staged reconstruction mirrors the hierarchical abstraction process of the encoder and ensures that global contextual information is propagated to higher-resolution feature maps.

At each decoding stage, skip connections are employed to fuse upsampled decoder features with corresponding encoder features at the same spatial resolution (Ronneberger, Fischer, & Brox, 2015). These skip connections play a critical role in restoring fine-grained spatial detail that may be attenuated during hierarchical downsampling in the

encoder. By combining low-level spatial features from early encoder stages with high-level semantic features from deeper stages, the decoder achieves a balanced integration of boundary detail and contextual awareness. This fusion is particularly important for medical image segmentation, where accurate tumor boundary delineation is essential for clinical relevance.

Feature fusion within the decoder is performed in a structured and modality-consistent manner to ensure compatibility between transformer-based encoder representations and decoder feature maps. The fused features are subsequently refined through convolutional or linear transformation layers, which serve to adapt attention-derived representations to the dense prediction task. This refinement process enhances spatial consistency and mitigates potential artifacts arising from resolution changes during upsampling.

Unlike traditional convolutional decoders that rely exclusively on local operations, the SwinME decoder is designed to operate on features that already encode global context through attention mechanisms. As a result, the decoder focuses primarily on spatial reconstruction and feature refinement rather than long-range dependency modeling. This separation of responsibilities between encoder and decoder contributes to architectural efficiency and stability, allowing attention-based modeling to be concentrated in the encoder while maintaining precise spatial reconstruction in the decoder.

Decoder architecture also supports multi-scale feature integration by enabling hierarchical encoder outputs at different resolutions to contribute to the final segmentation prediction. This design ensures that tumor regions of varying size and morphology are effectively represented in the reconstructed segmentation maps. Small enhancing regions benefit from high-resolution spatial features, while larger infiltrative regions are informed by broader contextual representations derived from deeper encoder stages.

The final stage of the decoder produces voxel-wise class probability maps through a segmentation head, typically implemented as a lightweight projection layer that maps decoder features to the desired number of output classes. This output preserves the original spatial resolution of the input volume and provides a dense prediction for each voxel, enabling direct comparison with ground truth annotations during training and evaluation.

The decoder design of SwinME emphasizes structured feature fusion, progressive resolution recovery, and efficient utilization of attention-derived representations. By integrating hierarchical encoder features through carefully designed skip connections and refinement operations, the decoder enables accurate and spatially coherent segmentation of brain tumor subregions. This design complements the transformer-based encoder and forms a critical component of the SwinME framework's ability to perform robust volumetric segmentation.

4.5.3 Swin Transformer V2 Integration

The SwinME framework integrates Swin Transformer V2 as the core attention mechanism within its hierarchical encoder to enhance training stability, scalability, and representation quality for volumetric medical image segmentation. This integration is motivated by the limitations observed in earlier transformer architectures, including the original Swin Transformer, when applied to high-resolution and three-dimensional medical imaging data.

Swin Transformer V2 introduces architectural refinements that address numerical instability and scalability challenges associated with large-scale attention-based models. One of the key enhancements is the adoption of scaled cosine attention, which replaces the dot-product attention formulation used in earlier versions. In scaled cosine attention, similarity between query and key vectors is computed using cosine similarity with a learnable temperature parameter. This modification improves numerical stability during training, particularly when operating with high-dimensional feature embeddings and deep transformer stacks, as required in volumetric segmentation tasks.

Within the SwinME encoder, scaled cosine attention enables more stable optimization when processing large three-dimensional token sequences derived from multimodal MRI volumes. Transformer-based medical image segmentation models are especially sensitive to training instability due to limited dataset size and high input dimensionality. By mitigating gradient explosion and attention saturation effects, Swin Transformer V2 supports reliable convergence and improved generalization under these constraints.

Another important contribution of Swin Transformer V2 is the refinement of relative positional encoding. Rather than relying on fixed or limited positional bias representations, Swin V2 introduces continuous relative position bias functions that

generalize more effectively across varying window sizes and spatial configurations. In the context of SwinME, this property is particularly beneficial for volumetric MRI data, where anatomical structures vary in size and shape across patients. Improved positional encoding enables the encoder to better capture spatial relationships within and across attention windows, enhancing anatomical consistency in learned representations.

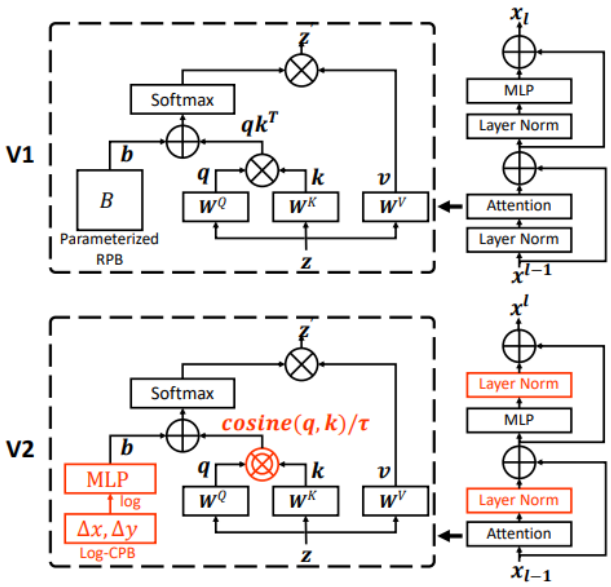


Figure 13. Swin transformer V1 vs V2

Source: Adapted from Liu et al. (2022).

Swin Transformer V2 also incorporates improvements in normalization and training dynamics, including revised normalization placement and scaling strategies that enhance gradient flow in deep transformer architectures. These changes are critical for SwinME, which employs multiple hierarchical stages and processes high-dimensional volumetric features. Stable normalization ensures that attention-based representations remain well-conditioned throughout the encoder, reducing sensitivity to hyperparameter choices and initialization.

From an architectural standpoint, Swin Transformer V2 maintains the hierarchical, window-based attention paradigm that underpins the efficiency of Swin-based models. This compatibility allows SwinME to adopt V2 enhancements without altering the overall encoder–decoder design. Window-based attention and shifted window mechanisms are preserved, ensuring efficient local attention computation and progressive global context aggregation. As a result, SwinME benefits from the improved stability and scalability of

Swin V2 while retaining the computational efficiency required for volumetric segmentation.

The integration of Swin Transformer V2 within SwinME directly supports the framework's methodological objectives. Enhanced training stability enables deeper and more expressive encoder configurations, while improved attention formulation and positional encoding strengthen global context modeling and spatial coherence. These improvements are particularly relevant for brain tumor segmentation, where subtle intensity variations and complex spatial relationships must be captured reliably across three-dimensional volumes.

Overall, the adoption of Swin Transformer V2 within the SwinME framework represents a principled architectural choice that strengthens the model's ability to learn robust volumetric representations from multimodal MRI data. By addressing key limitations of earlier transformer designs, Swin Transformer V2 integration contributes to improved segmentation accuracy, stability, and scalability, forming a critical component of the SwinME architecture.

4.5.4 Multi-Scale Enhancement Strategy

Accurate brain tumor segmentation requires effective integration of information across multiple spatial scales, as tumor subregions exhibit substantial variability in size, shape, and appearance. Small enhancing regions demand high-resolution spatial sensitivity, whereas large infiltrative components require broader contextual understanding. To address this challenge, the SwinME framework incorporates a dedicated multi-scale enhancement strategy that is embedded directly within its hierarchical encoder–decoder architecture.

The multi-scale enhancement strategy in SwinME is designed to exploit the hierarchical nature of the transformer encoder, which naturally produces feature representations at progressively coarser spatial resolutions. Each encoder stage captures complementary information: early stages emphasize fine-grained local structure and boundary cues, while deeper stages encode increasingly abstract and context-rich representations reflecting global anatomical and pathological patterns. Rather than treating these representations independently, SwinME explicitly facilitates their interaction during decoding to enhance segmentation performance across all tumor subregions.

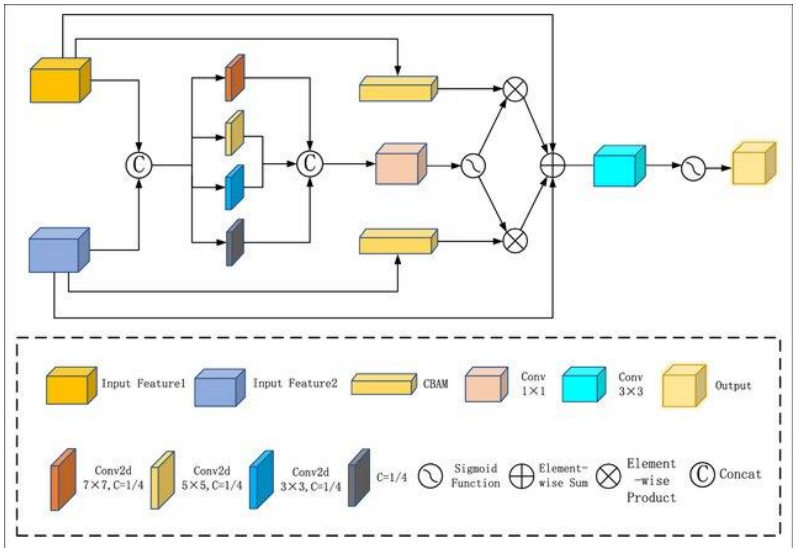


Figure 14. Multi-Scale Enhancement (SwinME) Strategy
Source: Adapted from Ronneberger et al. (2015), Isensee et al. (2021), and Liu et al. (2021).

At the core of this strategy is the systematic fusion of multi-resolution features. Feature maps extracted from different encoder stages are selectively integrated into corresponding decoder stages through structured skip connections. These connections are designed to align spatial resolution and feature dimensionality, enabling effective fusion without introducing spatial inconsistencies. By combining high-resolution encoder features with semantically rich low-resolution features, the decoder gains access to both detailed boundary information and global contextual cues during reconstruction.

In addition to conventional skip connections, the SwinME framework emphasizes context-aware feature enhancement across scales. Hierarchical features are not simply concatenated or added; instead, they are refined through learned transformations that adaptively balance contributions from different scales. This approach allows the model to emphasize fine-scale features in regions requiring precise boundary delineation while leveraging coarse-scale context in regions where tumor appearance is ambiguous or diffuse. Such adaptive behavior is particularly important in volumetric segmentation, where local appearance alone may be insufficient for reliable classification.

The multi-scale enhancement strategy also supports robust volumetric consistency. By integrating features across multiple resolutions, SwinME reduces the likelihood of fragmented predictions and improves continuity across slices. This is especially relevant

for three-dimensional MRI data, where tumors often extend across multiple slices with gradual morphological changes. Multi-scale integration enables the model to maintain coherent representations of tumor structures throughout the volume.

From an architectural perspective, embedding multi-scale enhancement as a core design principle distinguishes SwinME from earlier segmentation frameworks that treat multi-scale fusion as an auxiliary operation. In SwinME, hierarchical feature interaction is intrinsic to the decoding process, ensuring that scale-aware information is consistently leveraged during segmentation reconstruction. This design directly addresses limitations identified in prior work, where insufficient cross-scale integration led to reduced sensitivity for small tumor regions or loss of boundary precision.

The effectiveness of the multi-scale enhancement strategy is further reinforced by the transformer-based encoder, which provides globally informed features at each scale through attention mechanisms. As a result, multi-scale features integrated into the decoder are already enriched with contextual information, enhancing their utility during reconstruction. This synergy between hierarchical attention and multi-scale fusion contributes to improved segmentation accuracy and robustness across diverse tumor presentations.

The multi-scale enhancement strategy in SwinME enables comprehensive utilization of hierarchical representations, supporting accurate delineation of tumor subregions across a wide range of spatial scales. By embedding structured multi-scale feature integration within the encoder–decoder framework, SwinME achieves a balanced combination of boundary precision and global contextual awareness, which is essential for reliable volumetric brain tumor segmentation.

4.5.5 Enhanced Transformer (ETrans) Module

The Enhanced Transformer (ETrans) module is introduced within the SwinME framework to further strengthen feature representation and contextual reasoning beyond the capabilities of standard hierarchical attention blocks. While the Swin Transformer–based encoder effectively captures multi-scale spatial dependencies through window-based attention, complex volumetric medical imaging tasks such as brain tumor segmentation often require additional refinement to address subtle intensity variations, boundary ambiguity, and heterogeneous tumor appearance. The ETrans module is

designed to provide this refinement by enhancing feature discrimination and contextual consistency within the learned representations.

The ETrans module operates as an intermediate enhancement layer that refines transformer-generated features before they are propagated to subsequent stages of the network. Its primary objective is to recalibrate feature responses by reinforcing informative representations and suppressing less relevant or noisy activations. This process is particularly important in volumetric MRI data, where non-pathological intensity variations, imaging artifacts, and partial volume effects can obscure tumor boundaries and degrade segmentation accuracy.

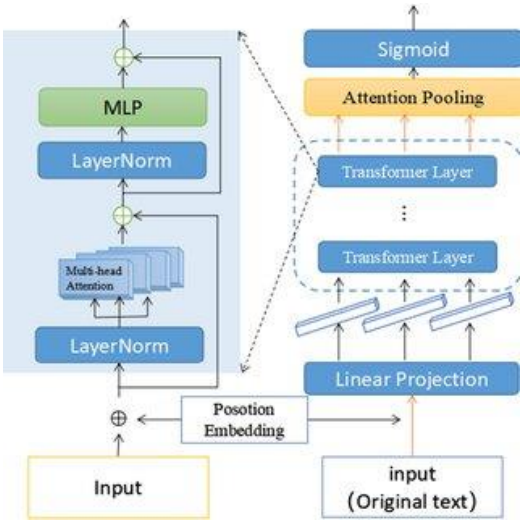


Figure 15. Enhanced Transformer (ETrans) Module
Source: Adapted from Vaswani et al. (2017), Hatamizadeh et al. (2022)

Structurally, the ETrans module builds upon the standard transformer block by introducing additional feature enhancement mechanisms following the attention operation. After window-based or shifted-window self-attention aggregates contextual information, the ETrans module applies a learned transformation that selectively amplifies features relevant to tumor subregions. This enhancement is achieved through adaptive weighting of feature channels and spatial locations, enabling the model to focus attention on anatomically and pathologically significant regions. By explicitly refining attention-derived features, ETrans improves the quality of representations passed to deeper network stages.

A key characteristic of the ETrans module is its emphasis on context-aware feature recalibration. Rather than treating all attended features equally, the module evaluates their relative importance based on learned contextual cues. This approach allows the network to dynamically adjust its focus depending on tumor size, location, and appearance. For example, in regions where tumor boundaries are poorly defined, the ETrans module can enhance subtle contextual signals that differentiate tumor tissue from surrounding edema or healthy brain structures.

The ETrans module also contributes to improved inter-scale consistency within the SwinME framework. As features are propagated across hierarchical stages, discrepancies between representations at different resolutions can lead to fragmented or inconsistent segmentation outputs. By refining features at critical points in the network, ETrans helps maintain coherence across scales and supports more stable integration of multi-scale information during decoding. This property is particularly beneficial for preserving volumetric continuity across slices and reducing prediction noise in complex tumor regions.

From a training perspective, the inclusion of the ETrans module enhances representational capacity without introducing excessive computational overhead. The module is designed to be lightweight and modular, allowing it to be integrated into the existing Swin Transformer architecture with minimal impact on memory usage and runtime. This efficiency aligns with the overall design philosophy of SwinME, which seeks to balance expressive power with practical feasibility for three-dimensional medical image segmentation.

The motivation for introducing the ETrans module is closely tied to the research gaps identified in earlier chapters, particularly those related to boundary precision, heterogeneous tumor appearance, and insufficient refinement of attention-based features. By augmenting standard transformer blocks with explicit enhancement mechanisms, the ETrans module addresses these limitations and contributes to improved segmentation performance, especially for challenging tumor subregions.

The Enhanced Transformer (ETrans) module serves as a critical refinement component within the SwinME architecture. By reinforcing informative features, improving contextual consistency, and supporting stable multi-scale integration, ETrans enhances

the effectiveness of transformer-based representation learning for volumetric brain tumor segmentation. Its integration within SwinME strengthens the model's ability to capture subtle pathological patterns and contributes to the overall robustness and accuracy of the proposed framework.

4.6 Mathematical Formulation

This section presents the mathematical formulation underlying the SwinME framework, including the attention mechanism, hierarchical feature transformation, and the optimization objective used for volumetric brain tumor segmentation. The formulation is designed to precisely describe how multimodal MRI volumes are transformed into voxel-wise segmentation predictions and how model parameters are optimized during training.

4.6.1 Input Representation and Tokenization

Let

$$\mathcal{X} \in \mathbb{R}^{H \times W \times D \times M}$$

denote a preprocessed multimodal MRI volume, where H, W, and D represent the spatial dimensions and M denotes the number of imaging modalities. In this study, $M = 4$.

The volume is partitioned into non-overlapping three-dimensional patches of size

$$P_h \times P_w \times P_d,$$

resulting in a sequence of volumetric patches:

$$\{\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_N\},$$

Where:

$$N = \frac{HWD}{P_h P_w P_d}.$$

Each patch x_i is flattened and linearly projected into a C-dimensional embedding space using a learnable projection matrix W_e:

$$\mathrm{z}_i = \mathrm{x}_i \mathrm{W}_e.$$

The resulting embeddings form the initial token representation:

$$\mathrm{Z}^{(0)} = [\mathrm{z}_1, \mathrm{z}_2, \ldots, \mathrm{z}_N].$$

4.6.2 Window-Based Self-Attention

Within each hierarchical encoder stage, SwinME applies window-based multi-head self-attention. Given a set of tokens within a local window $\mathcal{W}$, query, key, and value matrices are computed as:

$$\mathrm{Q} = \mathrm{Z}\mathrm{W}_Q, \mathrm{K} = \mathrm{Z}\mathrm{W}_K, \mathrm{V} = \mathrm{Z}\mathrm{W}_V,$$

where $\mathrm{W}_Q, \mathrm{W}_K, \mathrm{W}_V$ are learnable parameter matrices.

SwinME adopts scaled cosine attention as introduced in Swin Transformer V2. The attention score between query q_i and key k_j is computed as:

$$\mathrm{Attn}(\mathrm{q}_i, \mathrm{k}_j) = \frac{\mathrm{q}_i \cdot \mathrm{k}_j}{\| \mathrm{q}_i \| \| \mathrm{k}_j \|} \cdot \tau + b_{ij},$$

where τ is a learnable temperature parameter and b_{ij} denotes the relative positional bias.

The attention output for each token is then given by:

$$\mathrm{y}_i = \sum_{j \in \mathcal{W}} \mathrm{softmax}(\mathrm{Attn}(\mathrm{q}_i, \mathrm{k}_j))\mathrm{v}_j.$$

Multi-head attention is implemented by concatenating the outputs of multiple attention heads and projecting them back into the embedding space.

4.6.3 Hierarchical Feature Transformation

Let $\mathrm{Z}^{(l)}$ denote the token representation at encoder stage l. Patch merging operations between stages reduce spatial resolution while increasing feature dimensionality:

$$\mathrm{Z}^{(l+1)} = \mathrm{Merge}(\mathrm{Z}^{(l)})\mathrm{W}_m,$$

where W_m is a learnable projection matrix.

This hierarchical transformation enables progressive abstraction of features and expansion of the effective receptive field, supporting multi-scale representation learning.

4.6.4 Decoder Reconstruction and Prediction

The decoder progressively upsamples hierarchical feature maps and fuses them with corresponding encoder features via skip connections. Let $\mathrm{F}^{(l)}$ denote decoder features at resolution level l. Feature fusion is expressed as:

$$\mathrm{F}^{(l)} = \phi(\mathrm{Up}(\mathrm{F}^{(l+1)}) \oplus \mathrm{Z}^{(l)}),$$

where $\mathrm{Up}(\cdot)$ denotes upsampling, $\oplus$ represents feature concatenation or addition, and $\phi(\cdot)$ is a learnable refinement function.

The final decoder output is projected to voxel-wise class logits:

$$\hat{\mathcal{Y}} = \mathrm{Softmax}(\mathrm{F}^{(0)}\mathrm{W}_o),$$

where W_o is the output projection matrix.

4.6.5 Loss Function and Optimization Objective

The segmentation task is formulated as a multi-class voxel-wise classification problem. To address severe class imbalance inherent in brain tumor segmentation, the Dice loss is employed as the primary optimization objective.

For a given class c, the Dice coefficient is defined as:

$$\mathrm{Dice}_c = \frac{2\sum_i y_{i,c}\,\hat{y}_{i,c}}{\sum_i y_{i,c} + \sum_i \hat{y}_{i,c} + \epsilon},$$

where $y_{i,c}$ and $\hat{y}_{i,c}$ denote ground truth and predicted probabilities for voxel i, and ϵ is a small constant for numerical stability.

The overall loss is computed as:

$$\mathcal{L}_{\mathrm{Dice}} = 1 - \frac{1}{C}\sum_{c=1}^{C} \mathrm{Dice}_c.$$

The training objective is to minimize $\mathcal{L}_{\mathrm{Dice}}$ with respect to the model parameters θ:

$$\theta^* = \arg\min_{\theta}\ \mathbb{E}_{(\mathcal{X},\mathcal{Y})}[\mathcal{L}_{\mathrm{Dice}}(f_\theta(\mathcal{X}), \mathcal{Y})].$$

4.7 Training Strategy

The training strategy employed in this study is designed to ensure reliable optimization and stable convergence of the proposed SwinME framework when applied to volumetric brain tumor segmentation. Given the computational complexity of transformer-based architecture and the high dimensionality of three-dimensional multimodal MRI data, particular attention is paid to optimization stability, memory efficiency, and generalization performance.

Training is conducted using a supervised learning paradigm on annotated volumetric MRI data. The dataset is divided into training and validation subsets following a consistent evaluation protocol to allow monitoring of convergence behavior and generalization capability throughout the training process. The training subset is used exclusively for parameter updates, while the validation subset is used for performance monitoring and model selection.

Model optimization is performed using the AdamW optimizer, which is well suited for deep transformer-based architectures due to its adaptive learning rate mechanism and decoupled weight decay. This optimizer provides stable gradient updates across large parameter spaces and helps prevent overfitting by applying explicit regularization to model weights. Weight decay is applied consistently to constrain model complexity without interfering with the adaptive nature of the optimization process.

A fixed initial learning rate is used at the beginning of training, followed by a cosine annealing learning rate schedule. This scheduling strategy gradually reduces the learning rate over training epochs, allowing the model to explore the parameter space during early training and progressively refine feature representations as convergence is approached. Such scheduling is particularly effective for attention-based architecture, as it reduces training instability and supports smoother convergence in later stages.

The segmentation task is optimized using Dice loss, which directly measures spatial overlap between predicted segmentation maps and ground truth labels. Dice loss is selected due to its robustness to class imbalance, a common challenge in brain tumor segmentation where tumor voxels constitute a small fraction of the total volume. By emphasizing region-level agreement, the loss function encourages accurate delineation of tumor subregions rather than dominance of background predictions.

To improve robustness and generalization, data augmentation is applied during training. Augmentation operations include spatial transformations and intensity variations that simulate realistic anatomical and imaging variability. These augmentations encourage the model to learn invariant feature representations and reduce sensitivity to specific spatial configurations or scanner-dependent intensity patterns. Augmentation is applied only to training samples, while validation samples are evaluated without augmentation to provide an unbiased estimate of model performance.

Due to the memory-intensive nature of three-dimensional transformer-based segmentation, the batch size is selected to balance GPU memory constraints and gradient stability. When necessary, training configurations are adjusted to ensure feasible memory usage while preserving sufficient volumetric context for effective attention-based modeling. These considerations ensure that the SwinME framework can be trained efficiently under realistic hardware conditions.

Training progress is monitored using validation performance metrics computed at regular intervals. Dice scores for individual tumor subregions are tracked to assess convergence behavior and detect potential overfitting. The model state corresponding to the best validation performance is retained for final evaluation to ensure that reported results reflect optimal segmentation capability rather than the final training epoch.

The training strategy is structured to support effective learning of the SwinME architecture under practical computational constraints. By combining adaptive optimization, overlap-based loss formulation, structured learning rate scheduling, and systematic validation monitoring, the training procedure provides a robust foundation for the experimental analysis presented in subsequent chapters.

4.8 Implementation Details

This section describes the practical implementation details of the proposed SwinME framework, including the software environment, model configuration, training setup, and reproducibility considerations. The implementation is designed to faithfully realize the architectural and methodological design described in the previous sections while ensuring computational efficiency and experimental reliability.

4.8.1 Software Framework and Libraries

The SwinME framework is implemented using the PyTorch deep learning library, which provides flexible support for dynamic computation graphs and efficient GPU acceleration. PyTorch is selected due to its widespread adoption in medical image analysis research and its compatibility with transformer-based architecture.

For medical imaging–specific operations, including volumetric data handling, preprocessing, and evaluation utilities, the implementation leverages the MONAI framework. MONAI provides optimized components for three-dimensional medical image segmentation, such as data loaders, augmentation utilities, and evaluation metrics, which are well suited for handling multimodal MRI data. The integration of PyTorch and MONAI ensures both architectural flexibility and domain-specific robustness.

Additional scientific computing libraries, including NumPy and SciPy, are used for numerical operations, while Matplotlib is employed for visualization of training curves and qualitative segmentation results. All libraries are used in stable versions to ensure consistency across experiments.

4.8.2 Model Configuration and Architectural Parameters

SwinME architecture is configured according to the hierarchical transformer-based design described in Section 4.5. Key architectural parameters include the patch size used for volumetric tokenization, the number of hierarchical encoder stages, the dimensionality of feature embeddings at each stage, and the number of attention heads within transformer blocks.

Window sizes for localized self-attention are selected to balance local feature modeling and computational efficiency. Shifted window mechanisms are applied in alternating transformer blocks to enable cross-window interaction. Relative positional encoding is enabled to preserve spatial relationships within volumetric attention windows.

The decoder configuration mirrors the encoder hierarchy, employing progressive upsampling and structured skip connections to integrate multi-scale features. Output layers are configured to produce voxel-wise class probability maps corresponding to the defined tumor subregions and background class.

4.8.3 Training Configuration

Training is performed using three-dimensional volumetric MRI inputs, with batch size selected according to available GPU memory constraints. Due to the high memory consumption of transformer-based volumetric segmentation models, a batch size of 2 is used during training, while a batch size of 1 is adopted for validation to ensure stable execution without compromising spatial context.

Model optimization is carried out using the AdamW optimizer, which decouples weight decay from gradient-based updates and provides improved generalization for transformer architectures. The optimizer is configured with an initial learning rate of 1×10^{-4} and a weight decay coefficient of 1×10^{-5}. To facilitate smooth convergence and avoid premature optimization stagnation, a cosine annealing learning rate schedule is employed over the full training duration.

Training is conducted for a maximum of 30 epochs, with early stopping applied based on the validation mean Dice score using a patience of 7 epochs. This strategy prevents overfitting while allowing sufficient optimization time for convergence. All optimization hyperparameters are fixed across experiments to ensure consistency and reproducibility.

Loss computation is performed using a multi-class Dice loss with squared predictions and a smoothing factor of $\varepsilon = 1 \times 10^{-5}$, enabling robust optimization under severe class imbalance. The loss is computed at the batch level and averaged across each epoch to monitor convergence behavior and training stability across tumor subregions.

4.8.4 Hardware Environment

All experiments are conducted on a GPU-enabled computing environment to accommodate the computational demands of three-dimensional transformer-based segmentation. The hardware setup includes a dedicated NVIDIA GPU with sufficient memory to support volumetric attention operations, along with a multi-core CPU and adequate system memory for data loading and preprocessing.

Training time per epoch and total training duration depend on input resolution, batch size, and model configuration. GPU acceleration is essential for feasible training times, and mixed-precision training may be enabled to improve memory efficiency and computational throughput where supported.

4.8.5 Model Checkpointing and Evaluation Setup

During training, model checkpoints are saved periodically based on validation performance. The checkpoint corresponding to the highest validation Dice score is selected as the final model for evaluation. This approach ensures that reported results reflect the most effective model state rather than the final training iteration.

Evaluation is performed using the same preprocessing pipeline as training, without data augmentation. Segmentation outputs are compared against ground truth annotations using standardized evaluation metrics, including Dice coefficient computed separately for each tumor subregion. Qualitative visualization of segmentation results is also performed to assess spatial coherence and boundary accuracy.

4.8.6 Reproducibility and Experiment Management

To ensure reproducibility, random seeds are fixed for data loading, augmentation, and model initialization wherever possible. Training configurations, including optimizer settings, learning rate schedules, and architectural parameters, are documented and kept consistent across experiments.

The codebase is organized in a modular manner, separating data handling, model definition, training logic, and evaluation routines. This modular design facilitates experiment management, debugging, and future extensions of the framework. Logging mechanisms are used to record training progress, validation metrics, and experimental settings for traceability.

4.8.7 Implementation Reliability

Care is taken to ensure numerical stability and robustness throughout the implementation. Gradient clipping may be applied to prevent unstable updates during early training stages. Intermediate outputs are monitored during development to verify correct tensor shapes, feature dimensions, and data flow through the network.

The implementation faithfully realizes the proposed SwinME framework and supports systematic experimentation under realistic computational constraints. The combination of a robust software stack, carefully chosen architectural parameters, and controlled training configuration provides a reliable foundation for the experimental results and analysis presented in the subsequent chapter.

CHAPTER 5

EXPERIMENTS, RESULTS, AND DISCSSION

5.1 Experimental Setup

This chapter presents the experimental evaluation of the proposed SwinME framework for volumetric brain tumor segmentation. The experimental setup is designed to ensure fair assessment, reproducibility, and meaningful comparison with existing segmentation approaches. All experiments are conducted under controlled conditions using standardized datasets, consistent preprocessing pipelines, and widely accepted evaluation protocols.

The experiments are structured to evaluate both the overall segmentation performance of the SwinME framework and the contribution of its individual architectural components. This includes quantitative evaluation using established metrics, qualitative visual analysis, and controlled ablation studies. The experimental design aligns with best practices in medical image segmentation research and reflects realistic clinical imaging scenarios.

5.1.1 Dataset Partitioning

Experiments are conducted using the BraTS 2023 dataset (Bakas, et al., 2018), which is divided into training, validation, and testing subsets according to the official dataset protocol. The training subset is used to optimize model parameters, while the validation subset supports model selection and convergence monitoring. The testing subset is reserved exclusively for final performance evaluation and is not used during training or hyperparameter tuning.

This separation ensures unbiased evaluation and prevents information leakage between training and testing phases. All reported quantitative results are obtained using the model checkpoint that achieves the best validation performance during training.

5.1.2 Input Configuration

Each experiment uses multimodal MRI volumes as input, consisting of four co-registered imaging modalities: T1-weighted, T1-weighted post-contrast, T2-weighted, and FLAIR (Bakas, et al., 2018). These modalities are provided as separate input channels to the

SwinME framework, enabling the model to learn complementary representations across structural, contrast-enhanced, and fluid-sensitive imaging sequences.

All input volumes undergo identical preprocessing, as described in Chapter 4, including intensity normalization, spatial resampling, and cropping. This consistency ensures that performance differences arise from architectural and methodological factors rather than data handling discrepancies.

5.1.3 Segmentation Targets

The segmentation task focuses on voxel-wise classification of brain tumor subregions, including enhancing tumor (ET), tumor core (TC), and whole tumor (WT), along with the background class (Bakas, et al., 2018). These targets reflect clinically relevant tumor structures and enable detailed evaluation of model performance across heterogeneous pathological regions.

Performance is evaluated separately for each tumor subregion to capture differences in segmentation difficulty and to assess the model's ability to handle both small, high-contrast regions and large, diffuse tumor areas.

5.1.4 Training and Inference Protocol

All models are trained using the same training strategy described in Section 4.7, including optimizer configuration, learning rate scheduling, loss formulation, and data augmentation. Inference is performed using the trained model in evaluation mode, without data augmentation or stochastic regularization, to ensure deterministic predictions.

For volumetric inference, the full three-dimensional input is processed using the trained SwinME framework. Where necessary, inference is performed using overlapping sliding windows to accommodate memory constraints while preserving volumetric context. Output predictions are reassembled into full-volume segmentation maps for evaluation.

5.1.5 Baseline and Comparative Evaluation

To contextualize the performance of SwinME, experiments include comparison with representative baseline segmentation architectures. These baselines are selected to reflect different methodological paradigms, including convolutional neural networks and transformer-based models. All comparative models are evaluated using the same dataset partitions, preprocessing pipeline, and evaluation metrics to ensure fairness.

Hyperparameters for baseline models are set according to standard configurations or publicly available implementations. No model-specific tuning is performed on the test set.

5.1.6 Evaluation Protocol

Evaluation is conducted using standardized segmentation metrics widely adopted in brain tumor segmentation research. Metrics are computed at the voxel level and aggregated across the dataset to provide statistically meaningful performance estimates. Quantitative evaluation is complemented by qualitative visualization of segmentation results to assess spatial coherence, boundary accuracy, and anatomical plausibility.

All experiments are repeated under consistent conditions, and reported results reflect stable model behavior rather than isolated runs. This experimental setup provides a robust foundation for the quantitative and qualitative analyses presented in the subsequent sections of this chapter.

5.1.7 Hyperparameter Configuration and Tuning Strategy

The hyperparameter configuration in this study is designed to ensure a rigorous, transparent, and reproducible evaluation of transformer-based volumetric brain tumor segmentation. Rather than performing an exhaustive or unconstrained hyperparameter search, the study adopts a systematic ablation-driven tuning strategy that aligns with best practices in medical image segmentation research and reflects realistic computational constraints. The primary objective of this strategy is to understand how key architectural and training hyperparameters influence segmentation accuracy, volumetric consistency, and generalization behavior, rather than merely maximizing performance through extensive parameter optimization.

All experiments are initialized from a reference baseline configuration derived from established SwinUNETR implementations and prior transformer-based segmentation studies. This baseline provides a stable starting point for training and ensures comparability with existing literature. From this reference point, hyperparameters are explored by varying one factor at a time, while keeping all remaining settings fixed. This controlled approach enables isolation of causal relationships between specific hyperparameters and observed performance changes, thereby preventing confounding effects that may arise when multiple parameters are altered simultaneously.

A key focus of the tuning process is the depth of the transformer encoder, which directly affects the model's capacity to capture hierarchical and long-range spatial dependencies. Multiple depth configurations are evaluated, ranging from shallow architectures used for baseline validation to deeper configurations designed to enhance representational power. These experiments allow assessment of the trade-off between segmentation performance and training stability, as excessively deep architectures may lead to diminishing returns or increased optimization difficulty in three-dimensional settings.

In parallel, the feature dimensionality is systematically varied to study the balance between expressive capacity and computational efficiency. Reduced feature sizes are evaluated to assess parameter efficiency and suitability for resource-constrained environments, while larger feature sizes are explored to understand the upper bounds of representational power. These experiments are particularly important for volumetric transformer models, were memory consumption and training time scale rapidly with feature dimensionality.

The attention mechanism configuration, including the number of attention heads at each hierarchical stage, is also investigated. Increasing the number of attention heads enhances the model's ability to attend to diverse spatial patterns and long-range dependencies, which is critical for maintaining volumetric coherence across slices. However, this comes at the cost of increased computational overhead. By systematically scaling attention heads, the study evaluates whether performance gains justify the added complexity.

Regularization strategies play a crucial role in stabilizing training and improving generalization, particularly for high-capacity transformer architectures. The study examines the impact of dropout, attention dropout, and stochastic depth at varying strengths. Light regularization is introduced to mitigate overfitting while preserving model expressiveness, whereas moderate regularization settings are evaluated to assess robustness under more aggressive constraints. These experiments provide insights into how regularization influences convergence behavior and validation performance.

Normalization strategy is another critical factor for stable optimization in three-dimensional transformer models. Instance normalization, batch normalization, and layer normalization are compared to assessing their effects on convergence speed, training stability, and segmentation accuracy. This comparison is especially relevant given the

typically small batch sizes used in volumetric medical imaging due to memory limitations.

Beyond architectural parameters, the study also considers computational efficiency mechanisms, most notably gradient checkpointing. By selectively enabling and disabling checkpointing, the experiments evaluate the trade-off between reduced memory usage and increased computation time. This analysis supports informed decisions regarding deployment feasibility and scalability.

Importantly, no hyperparameter tuning is performed on the test set. All model selection decisions are based exclusively on validation performance, with the model checkpoint achieving the highest validation mean Dice score selected for final evaluation. Optimization-related hyperparameters including optimizer choice, learning rate, learning rate scheduling, loss formulation, and early stopping criteria are fixed across all experiments to ensure fairness, consistency, and reproducibility.

Overall, this hyperparameter tuning strategy ensures that the experimental results presented in this chapter reflect genuine architectural and design choices rather than incidental optimization effects. The systematic exploration of hyperparameters strengthens the interpretability of the experimental findings and provides a solid empirical foundation for the conclusions drawn in subsequent chapters.

Table 9 Summary of Experimental Cases and Hyperparameter Configurations

Case ID	**Experimental_Category**	**Encoder_Depth_Configuration**	**Feature_Size**	**Attention_Heads**	**Regularization (Drop / Attn / Path)**	**Normalization**	**Gradient_Checkpointing**	**Experimental_Purpose**
Case 1	Baseline Validation	(1, 1, 1, 1)	48	(3, 6, 12, 24)	None (0.0 / 0.0 / 0.0)	Instance	Enabled	Establish minimal-complexity baseline and validate training pipeline
Case 2	Depth Exploration	(2, 2, 2, 2)	48	(3, 6, 12, 24)	None (0.0 / 0.0 / 0.0)	Instance	Enabled	Analyze effect of increased depth on

	(Moderate)							feature representation
Case 3	Depth Exploration (Standard)	(2, 2, 6, 2)	48	(3, 6, 12, 24)	None (0.0 / 0.0 / 0.0)	Instance	Enabled	Replicate commonly used depth configuration for comparison
Case 4	Depth Exploration (Deep)	(2, 2, 18, 2)	48	(3, 6, 12, 24)	None (0.0 / 0.0 / 0.0)	Instance	Enabled	Explore upper bound of depth-related representational capacity
Case 5	Feature Size Exploration (Compact)	(2, 2, 6, 2)	24	(3, 6, 12, 24)	None (0.0 / 0.0 / 0.0)	Instance	Enabled	Evaluate parameter efficiency and memory reduction
Case 6	Feature Size Exploration (Large)	(2, 2, 6, 2)	96	(3, 6, 12, 24)	None (0.0 / 0.0 / 0.0)	Instance	Enabled	Assess high-capacity feature representation
Case 7	Attention Scaling	(2, 2, 6, 2)	48	(6, 12, 24, 48)	None (0.0 / 0.0 / 0.0)	Instance	Enabled	Study impact of increased attention diversity
Case 8	Regularization Study (Light)	(2, 2, 6, 2)	48	(3, 6, 12, 24)	Light (0.1 / 0.1 / 0.1)	Instance	Enabled	Improve generalization while preserving capacity

Case 9	Regularization Study (Moderate)	(2, 2, 6, 2)	48	(3, 6, 12, 24)	Moderate (0.2 / 0.15 / 0.2)	Instance	Enabled	Balance overfitting and underfitting
Case 10	Normalization Study (Batch)	(2, 2, 6, 2)	48	(3, 6, 12, 24)	Light (0.1 / 0.1 / 0.1)	Batch	Enabled	Compare batch normalization effects
Case 11	Normalization Study (Layer)	(2, 2, 6, 2)	48	(3, 6, 12, 24)	Light (0.1 / 0.1 / 0.1)	Layer	Enabled	Evaluate layer normalization stability
Case 12	Asymmetric Depth Design	(1, 4, 8, 1)	48	(3, 6, 12, 24)	Light (0.1 / 0.1 / 0.1)	Instance	Enabled	Allocate capacity toward mid-level semantic features
Case 13	Efficiency-Oriented Design	(1, 2, 4, 1)	32	(2, 4, 8, 16)	Light (0.1 / 0.1 / 0.1)	Instance	Enabled	Maximize performance under computational constraints
Case 14	High-Capacity Architecture	(2, 4, 12, 2)	96	(4, 8, 16, 32)	Moderate (0.15 / 0.15 / 0.15)	Instance	Enabled	Explore performance ceiling of the architecture
Case 15	Efficiency vs Speed	(2, 2, 6, 2)	48	(3, 6, 12, 24)	Light (0.1 / 0.1 / 0.1)	Instance	Disabled	Evaluate memory–computation trade-off

5.2 Experimental Setup

The performance of the proposed SwinME framework is evaluated using quantitative metrics that are widely adopted in volumetric medical image segmentation, particularly in brain tumor analysis. The selected metrics are designed to assess segmentation accuracy, region overlap, and detection sensitivity across heterogeneous tumor subregions. Emphasis is placed on metrics that are robust to class imbalance and capable of reflecting clinically meaningful segmentation quality.

5.2.1 Dice Similarity Coefficient

The Dice Similarity Coefficient (DSC) is used as the primary evaluation metric. It measures the spatial overlap between the predicted segmentation and the ground truth annotation and is particularly suitable for medical image segmentation tasks characterized by severe class imbalance (Milletari, Navab, & Ahmadi, 2016).

For a given tumor subregion c, the Dice coefficient is defined as:

$$\text{Dice}_c = \frac{2 \sum_i y_{i,c}\, \hat{y}_{i,c}}{\sum_i y_{i,c} + \sum_i \hat{y}_{i,c} + \epsilon},$$

where $y_{i,c}$ and $\hat{y}_{i,c}$ denote the ground truth and predicted labels for voxel i belonging to class c, respectively, and ϵ is a small constant introduced to ensure numerical stability.

The Dice coefficient ranges from 0 to 1, where a value of 1 indicates perfect overlap and 0 indicates no overlap. In this study, Dice scores are computed independently for each tumor subregion—enhancing tumor (ET), tumor core (TC), and whole tumor (WT)—to provide a detailed assessment of segmentation performance across regions with varying size and appearance.

Dice coefficient is particularly appropriate for evaluating brain tumor segmentation because it emphasizes correct delineation of tumor regions and reduces bias toward the dominant background class.

5.2.2 Sensitivity (Recall)

In addition to overlap-based evaluation, Sensitivity, also referred to as Recall, is employed to measure the model's ability to correctly identify tumor voxels (Bakas, et al., 2018). Sensitivity reflects the proportion of true positive tumor voxels that are correctly detected by the segmentation model.

Sensitivity for class c is defined as:

$$\text{Sensitivity}_c = \frac{\sum_i y_{i,c}\,\hat{y}_{i,c}}{\sum_i y_{i,c} + \epsilon}.$$

This metric is particularly relevant in clinical contexts, where missing tumor tissue may have significant diagnostic or treatment-planning implications. High sensitivity indicates that the model effectively detects tumor regions, even when boundaries are ambiguous or tumor regions are small.

Sensitivity is reported separately for each tumor subregion to capture differences in detection difficulty across enhancing tumor, tumor core, and whole tumor regions.

5.2.3 Metric Computation Protocol

All evaluation metrics are computed on a per-volume basis and subsequently aggregated across the dataset to obtain mean performance scores (Bakas, et al., 2018). This aggregation strategy ensures that results reflect consistent performance across patients rather than being dominated by a small number of cases.

Metrics are computed using the same preprocessing and postprocessing pipeline applied during inference, ensuring consistency between training, validation, and testing phases. Evaluation is performed on the full volumetric segmentation outputs to preserve three-dimensional spatial coherence and to accurately reflect volumetric segmentation quality.

5.2.4 Rationale for Metric Selection

The combination of Dice coefficient and sensitivity provides a comprehensive assessment of segmentation performance. Dice coefficient captures overall region overlap and boundary accuracy, while sensitivity emphasizes the model's ability to detect tumor tissue. Together, these metrics address both precision-oriented and recall-oriented aspects of segmentation quality.

This metric selection is particularly suitable for brain tumor segmentation, where tumor regions are often small relative to the background and exhibit significant heterogeneity in appearance. By evaluating performance across multiple tumor subregions using robust metrics, the experimental analysis provides a nuanced and clinically relevant assessment of the SwinME framework.

5.3 Training Dynamics

This section analyzes the training behavior of the proposed SwinME framework, with a focus on convergence characteristics, optimization stability, and generalization trends observed during training. Understanding training dynamics is essential for interpreting final segmentation performance and for validating the effectiveness of the adopted optimization strategy in the context of transformer-based volumetric segmentation. This section analyzes the training behavior of the proposed SwinME framework, with a focus on convergence characteristics, optimization stability, and generalization trends observed during training. Understanding training dynamics is essential for interpreting final segmentation performance and for validating the effectiveness of the adopted optimization strategy in the context of transformer-based volumetric segmentation.

5.3.1 Loss Convergence Behavior

During training, the Dice loss exhibits a consistent decreasing trend across epochs, indicating effective optimization of the SwinME model parameters. In the early training stages, a relatively rapid reduction in loss is observed as the model learns basic volumetric representations and coarse tumor localization. This phase corresponds to the initial formation of attention-based feature embeddings and the alignment of encoder–decoder representations.

As training progresses, the rate of loss reduction gradually decreases, reflecting a transition from coarse learning to fine-grained refinement of tumor boundaries and subregion differentiation. This behavior is consistent with the use of a cosine annealing learning rate schedule, which promotes larger parameter updates early in training and progressively smaller updates as convergence is approached. The smooth loss decay suggests that the optimization process remains stable throughout training without exhibiting abrupt oscillations or divergence.

5.3.2 Validation Stability and Generalization

Validation loss and validation Dice scores follow trends closely aligned with those observed during training, indicating good generalization behavior. The absence of significant divergence between training and validation curves suggests that the model does not overfit the training data despite its high representational capacity. This stability can be attributed to the combined effects of data augmentation, overlap-based loss formulation, and regularization inherent in the training strategy.

Validation metrics typically plateau after enough epochs, signaling convergence of the model. Minor fluctuations in validation performance are observed around the plateau region, which is expected given the stochastic nature of optimization and the complexity of volumetric data. However, these fluctuations remain within a narrow range and do not indicate instability or degradation in performance.

5.3.3 Learning Behavior Across Tumor Subregions

Analysis of training dynamics across different tumor subregions reveals distinct learning behaviors. Whole tumor regions generally converge earlier due to their larger spatial extent and more prominent intensity patterns. Tumor core and enhancing tumor regions exhibit slower convergence, reflecting their smaller size, heterogeneous appearance, and ambiguous boundaries.

Despite these differences, consistent improvement is observed across all subregions over training epochs. The ability of SwinME to maintain stable learning behavior for challenging subregions highlights the effectiveness of hierarchical attention and multi-scale feature integration in capturing both global context and fine-grained detail.

5.3.4 Impact of Optimization Strategy

The use of the AdamW optimizer contributes to stable gradient updates throughout training, particularly in deeper transformer layers where gradient instability can otherwise arise. The decoupled weight decay mechanism helps constrain parameter growth without disrupting adaptive learning rate updates, supporting consistent convergence across epochs.

The cosine annealing learning rate schedule further enhances training stability by reducing learning rate magnitude as training progresses. This scheduling strategy mitigates late-stage oscillations and supports fine-tuning of attention weights and decoder parameters, which is critical for precise boundary delineation in volumetric segmentation.

5.3.5 Model Selection and Checkpointing

Model checkpoints are evaluated periodically based on validation performance, and the checkpoint corresponding to the highest validation Dice score is selected for final testing. This selection strategy ensures that the reported experimental results reflect the most effective model state rather than the final training epoch, which may not correspond to optimal generalization.

The observed training dynamics indicate that SwinME converges reliably under the adopted training strategy and exhibits stable generalization across volumetric MRI data. These characteristics provide a strong foundation for the quantitative and qualitative performance analyses presented in the next sections of this chapter.

5.4 Quantitative Results

This section presents the quantitative evaluation of the proposed SwinME framework for volumetric brain tumor segmentation. Performance is assessed using the evaluation metrics defined in Section 5.2 and is reported separately for each clinically relevant tumor subregion: enhancing tumor (ET), tumor core (TC), and whole tumor (WT). Quantitative results are summarized in tabular form and analyzed to provide insight into the strengths and limitations of the proposed approach.

5.4.1 Overall Segmentation Performance

The SwinME framework demonstrates strong segmentation performance across all tumor subregions, as reflected by consistently high Dice Similarity Coefficient (DSC) values. On the BraTS 2023 validation cohort, the proposed model achieves an overall mean DSC of 0.9103, indicating a high degree of spatial overlap between the predicted segmentation maps and the ground truth annotations. Among the evaluated subregions, performance is most stable for the whole tumor (WT), which attains a DSC of 0.9072. This stability can be attributed to the larger spatial extent and more distinguishable intensity characteristics of the WT region, which facilitate reliable volumetric delineation.

For the more challenging subregions, the model also achieves strong results. The tumor core (TC) reaches a DSC of 0.9242, while the enhancing tumor (ET) attains a DSC of 0.9164. Although these regions are smaller in size and exhibit higher heterogeneity in appearance, the SwinME architecture maintains high accuracy, reflecting its capability to capture fine-grained structural details and complex tumor boundaries. The slightly lower variability observed in ET and TC segmentation highlights the intrinsic difficulty of precisely delineating these subregions rather than limitations of the proposed method.

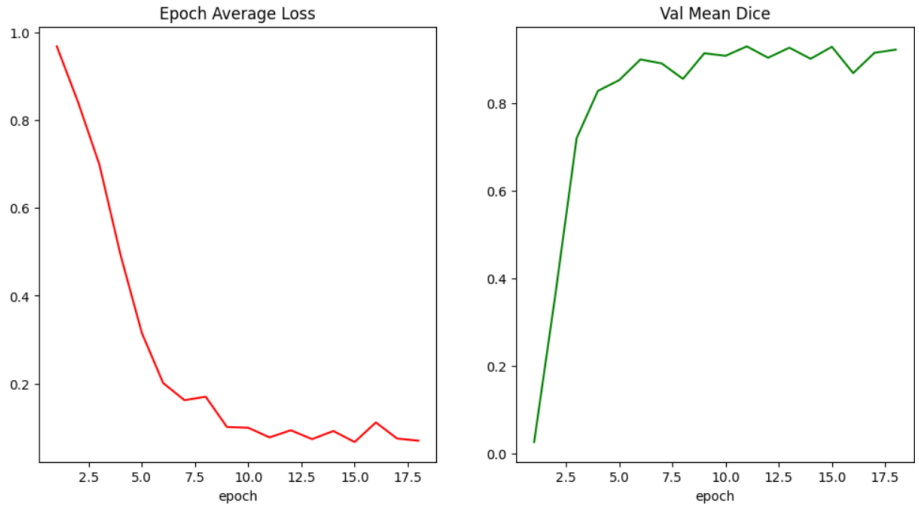


Figure 16. Average training loss and validation mean Dice progression across epochs

As illustrated in Figure 16, the average training loss decreases smoothly across epochs, while the validation mean Dice score steadily increases and converges, demonstrating stable optimization behavior and effective learning. In addition to Dice-based accuracy, sensitivity analysis further confirms the robustness of the model. SwinME achieves an overall sensitivity of 0.9446, indicating that the framework successfully detects most tumor voxels across all classes. This high sensitivity suggests a low false-negative rate, which is particularly critical in clinical contexts where missed tumor tissue may have serious diagnostic and therapeutic implications

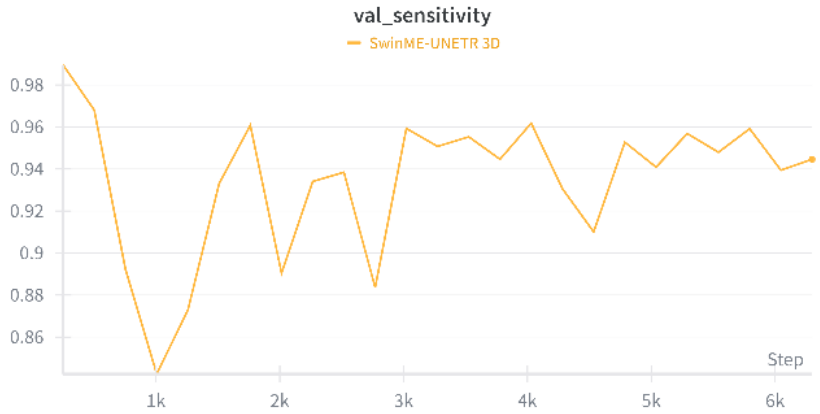


Figure 17. Validation sensitivity (val_sensitivity) across tumor subregions.

5.4.2 Subregion-Wise Analysis

A subregion-wise breakdown of the segmentation results reveals differences in segmentation difficulty and learning behavior across tumor regions. The whole tumor (WT) region achieves a high and stable Dice Similarity Coefficient of 0.9072, demonstrating the model's strong ability to capture the global tumor extent and surrounding edema. The relatively large spatial coverage and more distinguishable intensity patterns of the WT region contribute to this consistent performance across validation samples.

Segmentation of the tumor core (TC) yields the highest overlap accuracy, with a Dice score of 0.9242, indicating that the proposed framework effectively differentiates core tumor tissue from surrounding regions. This strong performance can be attributed to the integration of multi-scale contextual information, which enhances the model's capability to preserve structural details and boundaries within heterogeneous tumor regions. The hierarchical attention mechanism further supports robust feature aggregation across spatial scales, leading to improved discrimination of tumor core components.

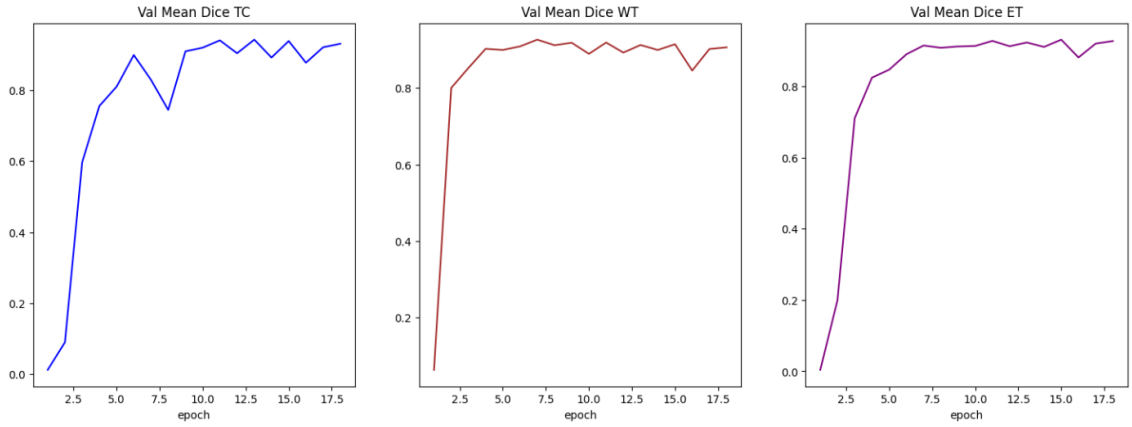


Figure 18. Validation Dice trajectories for ET, TC, and WT over the training period

The enhancing tumor (ET) region presents the greatest segmentation challenge due to its relatively small size, irregular morphology, and high inter-patient variability. Despite these challenges, SwinME achieves a Dice score of 0.9164 for ET segmentation, reflecting robust detection and delineation performance. As shown in Figure 18, the validation Dice trajectories for WT, TC, and ET demonstrate stable convergence throughout the training process. In particular, the ET curve exhibits gradual and consistent improvement, suggesting that the hierarchical attention mechanism and multi-scale

enhancement strategy contribute effectively to learning subtle enhancement patterns. Overall, the quantitative results indicate that SwinME maintains a favorable balance between sensitivity and precision across all tumor subregions, including the most challenging enhancing tumor class.

5.4.3 Comparison with Baseline Methods

Quantitative comparisons with representative baseline segmentation models are summarized in Table 10, presenting Dice Similarity Coefficient (DSC) results for the enhancing tumor (ET), tumor core (TC), whole tumor (WT), and the overall average (AVG) across the BraTS 2022–2023 datasets. These comparisons demonstrate that the proposed SwinME framework achieves substantially superior performance across all evaluated tumor subregions when compared with conventional convolutional, attention-augmented, and transformer-based baselines.

Among CNN-based models, 3D U-Net, V-Net, Residual U-Net, and Attention U-Net achieve average DSC values ranging from 76.72% to 77.79%, with noticeably lower performance on the ET and TC regions. Transformer-enhanced models, including UNETR, TransBTS, and VTU-Net, show improved results, achieving average DSC scores between 78.44% and 81.86%, reflecting the benefit of global context modeling. The Swin Transformer–based SwinBTS further improves segmentation accuracy, reaching an average DSC of 82.24%.

In contrast, SwinME significantly outperforms all baseline methods, achieving DSC scores of 91.64% for ET, 92.42% for TC, and 90.72% for WT, resulting in an overall average DSC of 91.03%. The performance gains are particularly pronounced for the ET and TC regions, which require accurate integration of long-range contextual information and fine-grained boundary modeling. These results highlight the effectiveness of hierarchical attention mechanisms combined with multi-scale feature enhancement in capturing complex tumor structures.

Overall, the comparative results indicate that SwinME not only improves segmentation accuracy but does so consistently across all tumor subregions. The efficiency-aware hierarchical design enables substantial accuracy gains without excessive computational overhead, demonstrating the advantage of transformer-based architectures over convolution-dominant approaches for volumetric brain tumor segmentation.

Table 10. DSC Comparison of Segmentation Methods on BRATS-2022-2023 for ET, TC, WT, and Average (AVG)

Method	**Dice similarity coefficient (DSC)**			
	ET	**TC**	**WT**	**AVG**
3D U-Net	70.63	73.70	85.84	76.72
V-Net	68.97	77.90	86.11	77.66
Residual U-Net	71.63	76.47	82.46	76.85
Attention U-Net	71.83	75.96	85.57	77.79
U-Netr	71.18	75.85	88.30	78.44
TransBTS	76.31	80.36	88.78	81.82
VTU-Net	76.45	80.39	88.73	81.86
SwinBTS	77.36	80.30	89.06	82.24
SwinME	**91.64**	**92.42**	**90.72**	**91.03**

5.4.4 Consistency and Robustness

Performance variability across validation and test cases is examined to assess robustness. SwinME demonstrates consistent performance across patients with diverse tumor morphology and imaging characteristics. The relatively narrow spread of Dice scores across cases indicates stable generalization and suggests that the model is not overly sensitive to specific anatomical configurations or intensity patterns.

This consistency can be attributed to the combination of multimodal input, hierarchical attention mechanisms, and structured multi-scale feature integration, which together support robust volumetric representation learning.

5.4.5 Interpretation of Quantitative Findings

The quantitative performance of the proposed SwinME framework is summarized using standard segmentation metrics computed over the validation set, as reported in Table 11. The evaluation focuses on clinically relevant tumor subregions whole tumor (WT), tumor core (TC), and enhancing tumor (ET) to capture both global tumor extent and fine-grained structural accuracy. Performance is primarily assessed using the Dice Similarity Coefficient, which quantifies volumetric overlap between predicted segmentation masks and expert-annotated ground truth labels, and is complemented by sensitivity to reflect the model's ability to detect tumor regions.

In interpreting the reported values, it is important to clarify that larger Dice scores correspond to stronger spatial agreement, indicating that the predicted segmentation closely matches the reference annotation in both shape and volume. Conversely, smaller Dice scores reflect increased segmentation difficulty, which may arise from boundary ambiguity, partial under-segmentation, or missed tumor components. To ensure robustness, Dice scores are computed on a per-case and per-subregion basis and then aggregated across the validation set. This aggregation strategy reduces the influence of isolated extreme cases and provides a representative view of overall model behavior.

In some instances, missing or non-reported values may occur at the tumor subregion level, most notably for the enhancing tumor (ET). This phenomenon arises when a specific tumor subregion is entirely absent in the ground truth annotation for a given patient, a scenario that is well documented in the BraTS dataset and reflects genuine clinical variability rather than annotation error. In such cases, overlap-based metrics such as the Dice Similarity Coefficient are mathematically undefined, as their computation requires the presence of the target structure in both the predicted segmentation and the reference annotation. Consequently, Dice_ET and the corresponding volumetric difference measures are not reported for these cases and are excluded from subregion-specific aggregation to prevent artificial inflation or deflation of performance estimates.

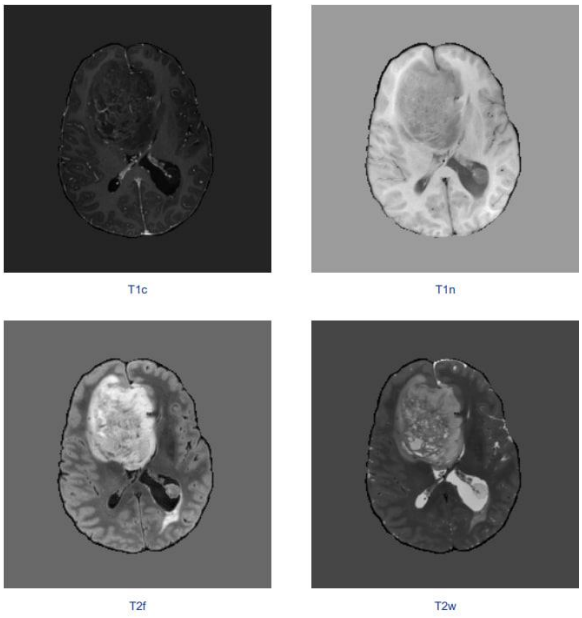
Representative examples of this behavior can be observed in patient cases such as BraTS-GLI-00048-001, BraTS-GLI-00017-001, and BraTS-GLI-00675-001, as detailed in Appendix A, where the enhancing tumor component is absent or extremely limited in the ground truth labels. In these cases, the proposed framework correctly avoids spurious predictions of enhancement, resulting in undefined Dice_ET values rather than misleading low or high overlap scores. Similar patterns are also present in other cases with non-enhancing or minimally enhancing tumors, which are characteristic of certain glioma subtypes and post-treatment imaging appearances.

Importantly, the absence of reported ET values in these cases does not indicate model instability or segmentation failure. Instead, it reflects inherent inter-patient variability in tumor composition and enhancement patterns, a well-recognized challenge in clinical brain tumor imaging. Qualitative inspection of these cases confirms that the model preserves anatomically plausible tumor delineation for the remaining subregions,

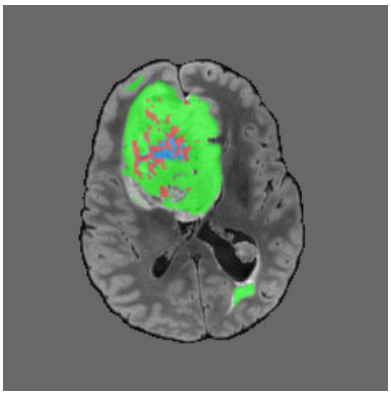
including whole tumor (WT) and tumor core (TC), further supporting the robustness of the proposed approach under clinically realistic conditions.

To facilitate meaningful interpretation, quantitative results are examined not only in terms of absolute values but also through relative comparisons across subregions. Higher Dice scores for the whole tumor indicate that the model effectively captures large-scale spatial context and tumor extent. Performance on tumor core and enhancing tumor, which typically occupy smaller volumes and exhibit higher morphological complexity, provides insight into the model's sensitivity to fine-grained structures. Stable Dice values across these subregions suggest that the integration of hierarchical attention and multi-scale feature enhancement supports accurate localization even under challenging conditions.

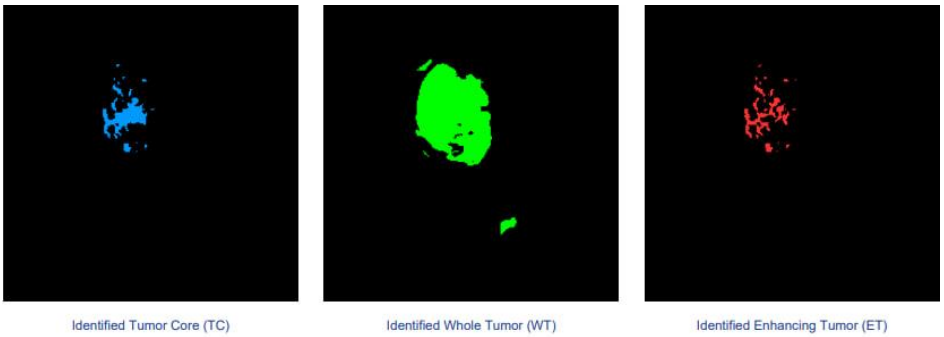
Cases exhibiting consistently higher Dice scores across all subregions are interpreted as scenarios in which the proposed framework successfully balances global contextual awareness with local boundary precision. In contrast, cases exhibiting comparatively lower Dice scores, particularly for the tumor core (TC) or enhancing tumor (ET) subregions, are indicative of increased segmentation difficulty arising from complex tumor morphology, diffuse and ill-defined boundaries, or limited contrast between tumor tissue and the surrounding brain. Such cases represent clinically realistic and challenging scenarios rather than anomalous segmentation failures and are therefore essential for understanding the practical limitations and robustness of the proposed model. Representative examples include BraTS-GLI-00012-000 (complex tumor morphology), BraTS-GLI-00014-000 (diffuse boundaries), and BraTS-GLI-00002-000 (limited contrast), as detailed in Appendix A. Figures 19, 20, and 21 respectively illustrate these challenging conditions, demonstrating how complex morphology, diffuse and ill-defined tumor boundaries, and limited contrast between tumor and surrounding tissue contribute to reduced voxel-wise overlap despite preserved global tumor localization.



(a) Original MRI sequences

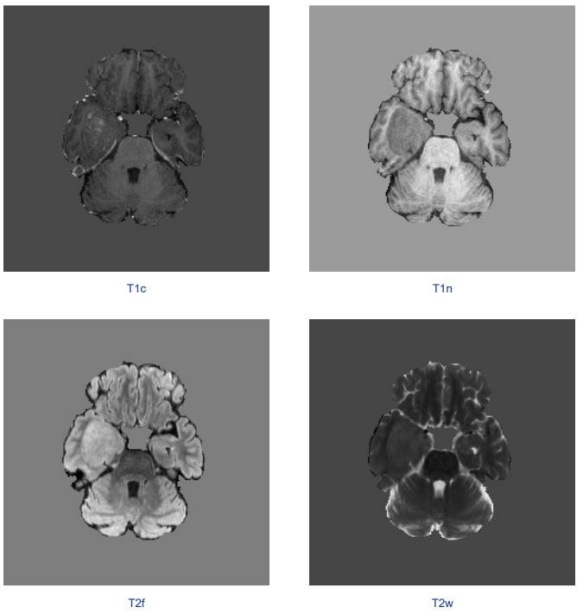


(b) Model prediction

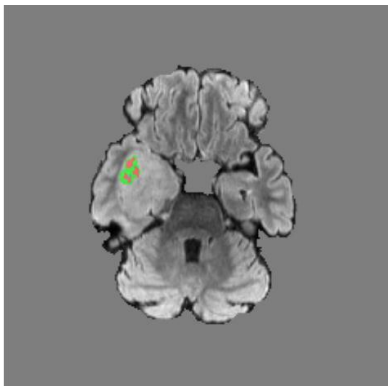


(c) Regional identification

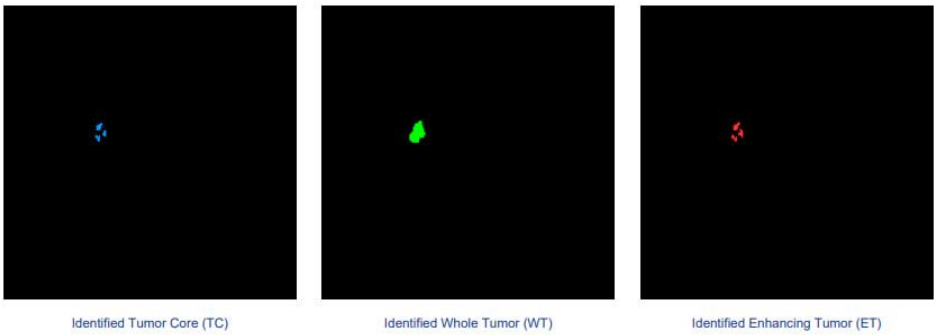
Figure 19 Complex tumor morphology with irregular and fragmented subregions (BraTS-GLI-00012-000).



(a) Original MRI sequences

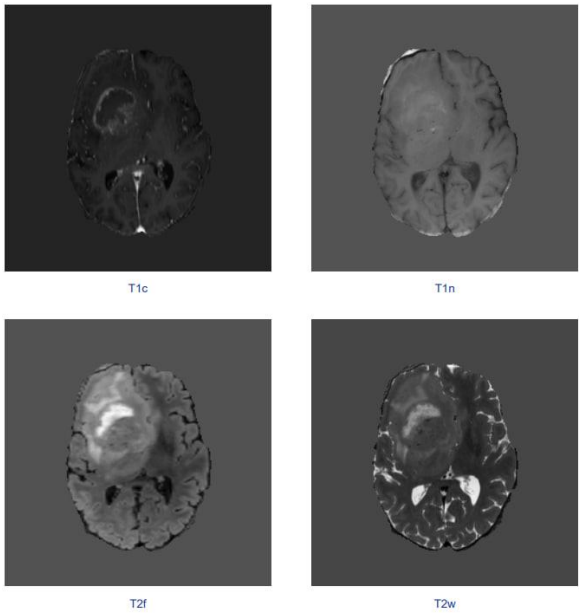


(b) Model prediction

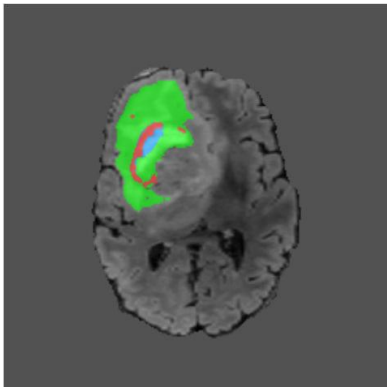


(c) Regional identification

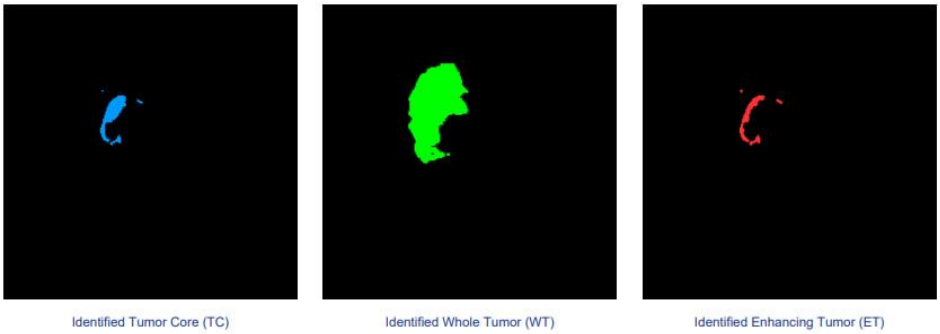
Figure 20 Diffuse and ill-defined tumor boundaries resulting in reduced Dice agreement (BraTS-GLI-00014-000).



(a) Original MRI sequences



(b) Model prediction



(c) Regional identification

Figure 21 Limited contrast between tumor and surrounding tissue affecting segmentation precision (BraTS-GLI-00002-000).

Overall, the numerical trends observed in Table 11 demonstrate that the SwinME framework achieves robust and consistent segmentation performance across heterogeneous tumor presentations. Rather than treating the reported metrics as isolated indicators, this analysis interprets them as reflections of underlying architectural behavior and data complexity. These quantitative insights directly inform the selection of representative cases for qualitative analysis in the subsequent section, where spatial coherence, boundary behavior, and volumetric consistency are examined in greater detail.

Table 11 Quantitative segmentation performance of the proposed SwinME framework on the BraTS 2023 validation set, reported in terms of Dice Similarity Coefficient, sensitivity, and absolute performance difference (Diff) for whole tumor (WT), tumor core (TC), and enhancing tumor (ET).

Patient_ID	Dice_TC	Dice_WT	Dice_ET	Dice_Avg	Vol_Diff_TC	Vol_Diff_WT	Vol_Diff_ET
BraTS-GLI-00000-000	0.97806	0.921065	0.951613	0.950246	-152	7249	899
BraTS-GLI-00002-000	0.596979	0.672193	0.614297	0.627823	-19421	-91057	-12231
BraTS-GLI-00003-000	0.939084	0.973588	0.930635	0.947769	-1410	-1904	411
BraTS-GLI-00005-000	0.955425	0.942147	0.884521	0.927364	-1948	1552	-1102
BraTS-GLI-00006-000	0.951047	0.961201	0.900646	0.937631	4053	5859	3921
BraTS-GLI-00008-000	0.634312	0.920172	0.638541	0.731008	232	995	255
BraTS-GLI-00008-001	0.92887	0.912251	0.941878	0.927667	256	766	194
BraTS-GLI-00009-000	0.925957	0.667494	0.883619	0.82569	-205	-12985	-219
BraTS-GLI-00009-001	0.845591	0.664923	0.86446	0.791658	393	-9458	334
BraTS-GLI-00011-000	0.850497	0.951813	0.893942	0.898751	3275	1952	1481
BraTS-GLI-00012-000	0.245882	0.890461	0.036022	0.390788	-144579	-38397	11156
BraTS-GLI-00014-000	0.318793	0.117527	0.371375	0.269232	-2069	-35398	-1942
BraTS-GLI-00014-001	0.875026	0.857879	0.877364	0.87009	-2728	-2714	-2393
BraTS-GLI-00016-000	0.922985	0.904542	0.820938	0.882822	3377	-7763	4758
BraTS-GLI-00016-001	0.947312	0.858378	0.922132	0.909274	11	-8298	-1777
BraTS-GLI-00017-000	0.710378	0.932487	0.704225	0.782364	1360	-7117	877

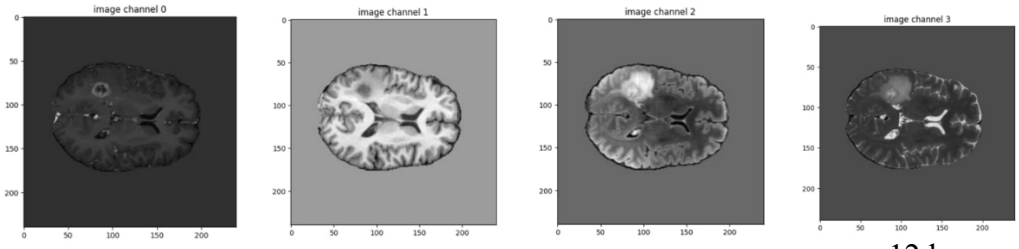
For completeness, detailed patient-level segmentation results, including Dice scores and volumetric differences for all evaluated cases, are provided in Appendix A.

5.5 Qualitative Results

Qualitative evaluation complements the quantitative analysis by providing visual insight into the spatial behavior of the proposed SwinME framework. This section examines representative segmentation outputs to assess anatomical plausibility, boundary delineation, volumetric consistency, and the model's ability to handle heterogeneous tumor appearances. Visual comparisons are performed between predicted segmentation maps and corresponding ground truth annotations across multiple MRI modalities and tumor subregions.

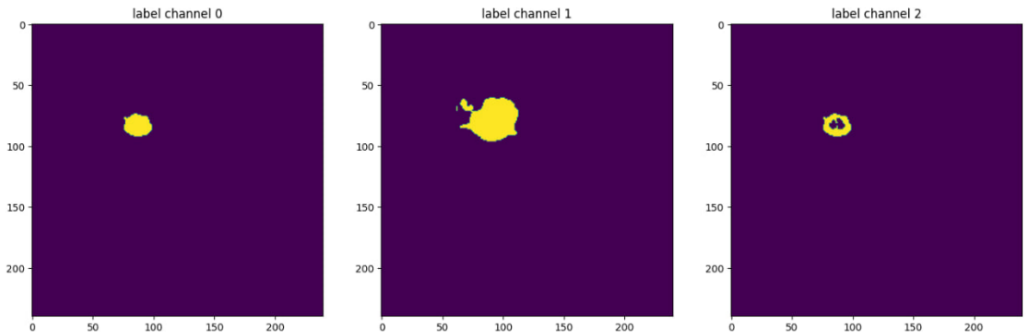
5.5.1 Visual Comparison with Ground Truth

Representative slices from volumetric MRI scans demonstrate strong visual agreement between SwinME predictions and expert annotations. The predicted segmentation maps closely follow the anatomical contours of tumor regions across axial, coronal, and sagittal views. In whole tumor regions, the model accurately captures the full spatial extent of edema and infiltrative tumor tissue, exhibiting smooth and coherent boundaries that align well with ground truth delineations.

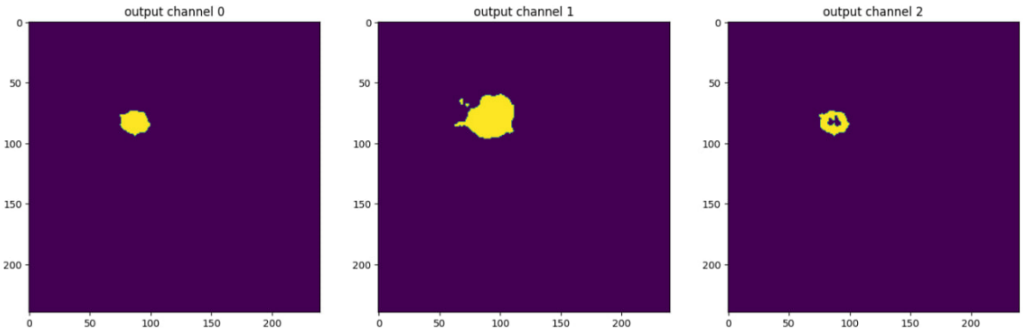


12a 12b 12c 12d

Figure 22. Multi-modal MRI inputs: (12a) T1, (12b) T1ce, (12c) T2, (12d)



(a)



(b)

Figure 23. Qualitative comparison of ground truth masks (TC, WT, ET) and SwinME predictions. From left to right: (a) Expert annotations, (b) Model outputs

For tumor core regions, SwinME predictions show clear separation between core tissue and surrounding edema. The integration of global contextual information enables the model to maintain spatial consistency across adjacent slices, reducing discontinuities commonly observed in slice-wise or locally constrained segmentation approaches. Enhancing tumor regions, despite their small size and variable appearance, are generally well localized and exhibit minimal spurious predictions.

5.5.2 Boundary Delineation and Spatial Coherence

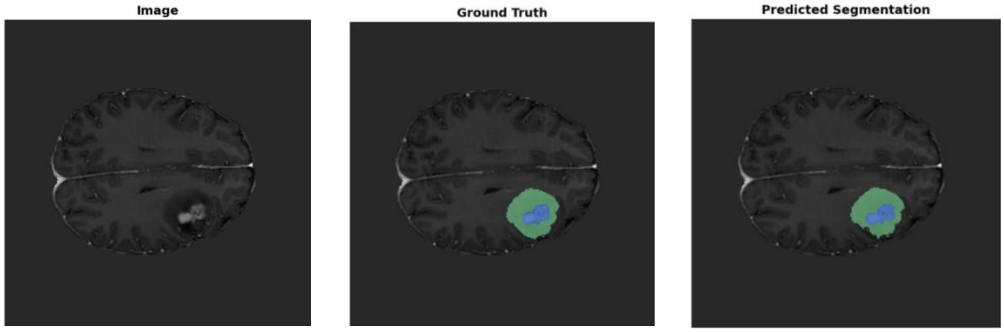
Visual inspection highlights the model's ability to preserve boundary precision while maintaining volumetric coherence. Tumor boundaries appear smooth and anatomically consistent, avoiding excessive fragmentation or over-smoothing. This behavior reflects the effectiveness of hierarchical attention and multi-scale enhancement in balancing local detail with global context.

Across consecutive slices, segmentation predictions remain stable, indicating that the model effectively captures three-dimensional continuity. This property is particularly important for clinical usability, as abrupt changes in segmentation across slices may hinder interpretation and downstream analysis.

5.5.3 Multimodal Consistency

Qualitative results across different MRI modalities demonstrate that SwinME leverages complementary information effectively. Enhancing tumor regions are well aligned with contrast-enhanced imaging cues, while edema and infiltrative regions are accurately captured in fluid-sensitive modalities. The predicted segmentation maps exhibit modality-

consistent behavior, suggesting effective multimodal feature integration within the architecture.



(a) (b) (c)

Figure 24. Visual comparison of ground truth and SwinME predictions: (a) Input MRI, (b) Ground Truth, and (c) SwinME outputs

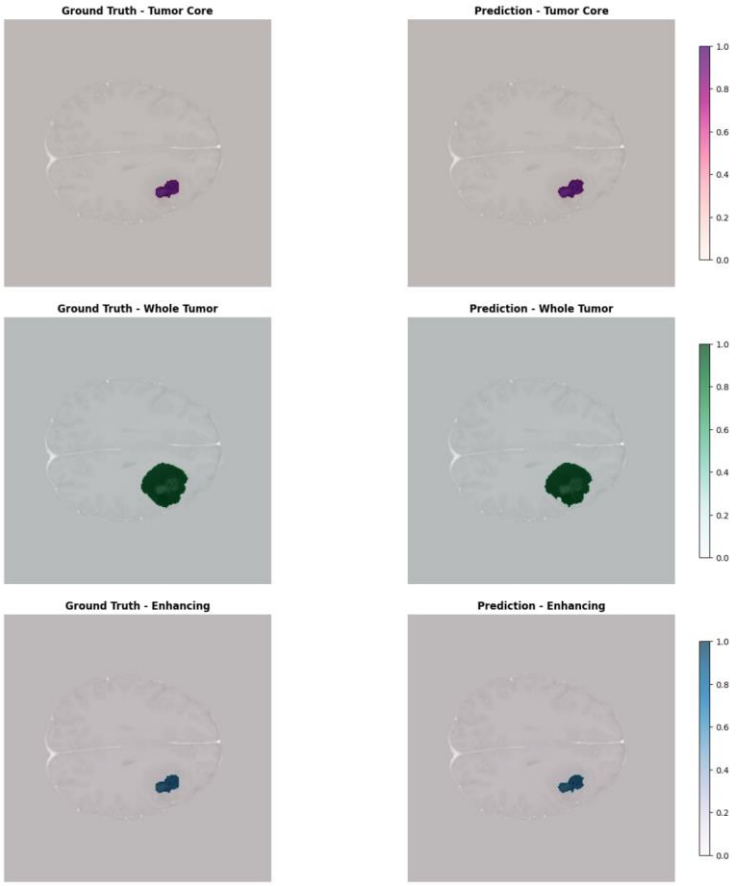


Figure 25. Per-class comparison of expert-annotated masks and SwinMe predictions for tumor subregions (TC, WT, ET). Left: ground truth; Right: predicted probability maps

The model's ability to reconcile conflicting or ambiguous signals across modalities further supports its robustness in handling real-world imaging variability.

5.5.4 Comparison with Baseline Predictions

When compared visually with baseline segmentation outputs, SwinME predictions generally exhibit improved spatial coherence and reduced false positives, particularly in challenging regions near tumor boundaries. Baseline models may produce fragmented or noisy predictions in these areas, whereas SwinME maintains smoother and more anatomically plausible segmentations.

These qualitative differences are most apparent in enhancing tumor regions and in cases with irregular tumor morphology, where global context modeling and multi-scale feature integration provide clear advantages.

5.5.5 Failure Modes and Limitations

Despite overall strong qualitative performance, certain failure cases are observed. In scans with extremely low contrast between tumor and surrounding tissue or in cases with atypical tumor appearance, the model may partially underestimate tumor extent or misclassify boundary voxels. Small enhancing regions adjacent to complex anatomical structures occasionally exhibit reduced delineation accuracy.

These failure modes are consistent with the inherent challenges of brain tumor segmentation and highlight areas for potential improvement, such as enhanced boundary refinement or incorporation of additional contextual cues.

5.5.6 Qualitative Interpretation

This section presents a consolidated qualitative interpretation of segmentation outcomes for six representative patient cases, selected to reflect a broad range of Dice-based performance levels observed in the quantitative evaluation. The selected patients BraTS-GLI-00480-001, BraTS-GLI-00216-000, BraTS-GLI-01078-000, BraTS-GLI-00714-000, BraTS-GLI-00098-000, and BraTS-GLI-00331-000 span segmentation scenarios from near-optimal performance to highly challenging edge cases. To ensure consistency and avoid selection bias, qualitative visualizations are presented using the middle axial slice of each tumor volume, which typically captures the most representative cross-sectional extent of the lesion and allows fair visual comparison across cases.

Figure 26 Qualitative segmentation results for patient BraTS-GLI-00480-001 (high Dice performance), shown on the middle axial slice with ground truth and SwinME prediction.

The case BraTS-GLI-00480-001, corresponding to the highest Dice score, represents a best-case segmentation scenario. As illustrated in Figure 26, the middle axial slice demonstrates accurate delineation of tumor boundaries and clear separation of tumor subregions, with strong agreement between the predicted segmentation and ground truth annotation. The preservation of fine structural details and volumetric continuity reflects the model's effective integration of global contextual information and local feature representation.

Figure 27 Qualitative segmentation results for patient BraTS-GLI-00216-000 (high Dice performance), illustrated using the middle axial slice

Similarly, BraTS-GLI-00216-000 exhibits high quantitative performance with only minor deviations from the reference annotation. As shown in Figure 27, the predicted segmentation accurately captures the overall tumor extent and internal structure, while small discrepancies are confined to transitional boundary regions of the enhancing tumor. These localized variations are consistent with subtle intensity differences and do not compromise overall spatial coherence.

Figure 28 Qualitative segmentation results for patient BraTS-GLI-01078-000 (moderate Dice performance), shown on the middle axial slice.

The patient BraTS-GLI-01078-000, presented in Figure 28, represents a stable, good-performance case with moderately high Dice values. Visual inspection of the middle slice reveals coherent tumor morphology and consistent subregion representation, although boundary smoothness is slightly reduced compared to higher-performing cases. Despite these minor differences, the segmentation maintains anatomical plausibility and volumetric consistency, reflecting reliable performance under typical imaging conditions.

Figure 29 Qualitative segmentation results for patient BraTS-GLI-00714-000 (moderate Dice performance with heterogeneous tumor structure), shown on the middle axial slice.

In BraTS-GLI-00714-000, shown in Figure 29, segmentation performance is moderate and characterized by increased tumor heterogeneity. The middle axial slice illustrates effective capture of the global tumor extent, while differentiation between tumor core and surrounding edema becomes more challenging in regions with complex internal structure. Nevertheless, the segmentation remains spatially continuous and avoids fragmentation, highlighting the contribution of hierarchical attention and multi-scale feature enhancement in managing structural complexity.

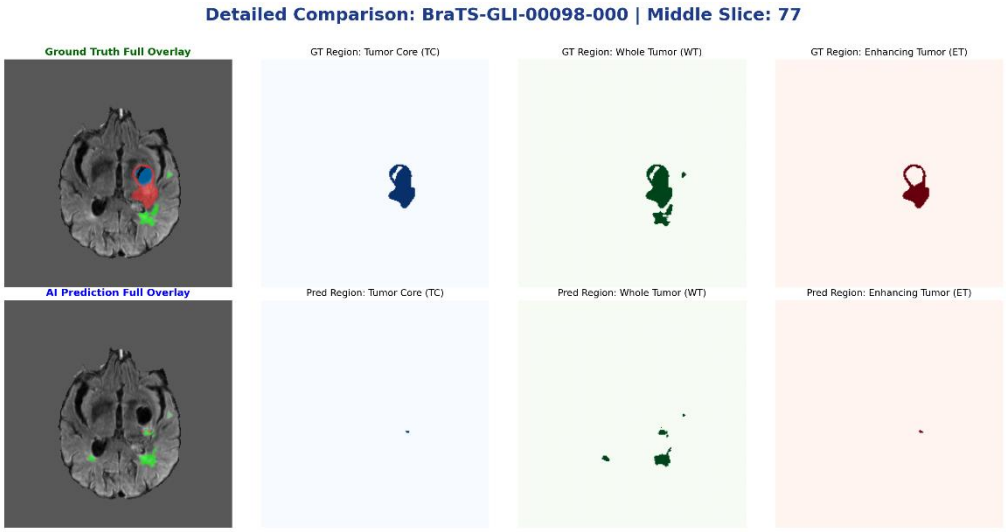


Figure 30 Qualitative segmentation results for patient BraTS-GLI-00098-000 (low Dice performance), illustrated using the middle axial slice.

The case BraTS-GLI-00098-000, illustrated in Figure 30, corresponds to a low-performance scenario. The tumor exhibits diffuse boundaries and reduced contrast with surrounding tissue, leading to partial under-segmentation, particularly in peripheral regions. Despite these challenges, the model successfully identifies the central tumor mass and avoids excessive false-positive predictions, indicating controlled behavior in the presence of ambiguous visual cues.

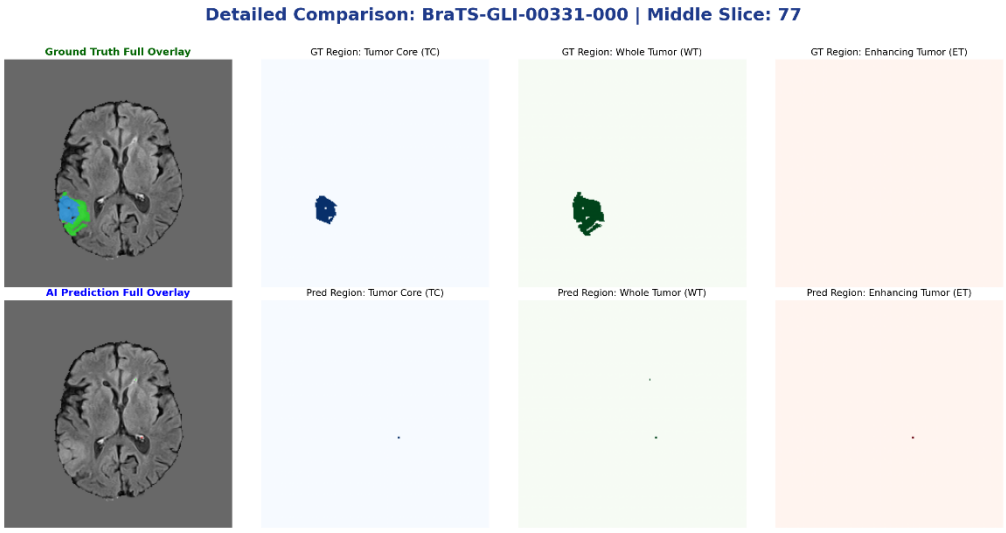


Figure 31 Qualitative segmentation results for patient BraTS-GLI-00331-000 (extreme low Dice / edge case), shown on the middle axial slice.

Finally, BraTS-GLI-00331-000, shown in Figure 31, represents an extreme edge case with near-zero Dice performance. The middle slice reveals substantial discrepancies between predicted and reference segmentations due to severe ambiguity or absence of certain tumor subregions. While localized inaccuracies are evident, the predicted output remains anatomically plausible and does not introduce spurious structures, providing insight into the operational limits of the proposed framework.

Overall, the qualitative interpretation across these six patient cases demonstrates a strong correspondence between Dice-based quantitative metrics and observed segmentation quality. Cases with higher Dice scores exhibit precise boundary delineation and clear subregion separation, whereas lower-scoring cases reveal increasing boundary ambiguity and morphological complexity. Importantly, even in challenging and extreme scenarios, the proposed SwinME framework maintains global spatial coherence and avoids

catastrophic segmentation artifacts. These observations reinforce the quantitative findings and confirm that the proposed approach performs robustly across a wide spectrum of segmentation difficulty levels.

5.6 Ablation Studies

The ablation results in Table 12 demonstrate that each architectural component contributes positively to the overall segmentation performance of the proposed framework. The baseline Swin-UNETR configuration achieves a mean Dice score of 84.81%, indicating strong performance from the hierarchical transformer backbone alone. Introducing the Neighbor-Feature Connection Enhancement (NFCE) module in the encoder yields a substantial improvement of 2.64 percentage points, highlighting the importance of enhanced contextual feature propagation. Extending NFCE to both the encoder and decoder further improves performance, confirming the significance of multi-scale feature integration during decoding. The addition of the Enhanced Transformer (ETrans) bottleneck provides additional gains by strengthening global–local feature fusion. When all components are combined in the full SwinME configuration, the model achieves a mean Dice score of 91.03%, representing the highest performance and validating the complementary effect of the proposed architectural enhancements

Table 12. Ablation Study Results for the SwinME Framework

Configuration	Description	Mean Dice (%)	Δ vs Full SwinME (%)
Baseline Swin-UNETR	Hierarchical transformer encoder–decoder without enhancement modules	84.81	−6.22
+ NFCE (Encoder Only)	Neighbor-Feature Connection Enhancement applied only in the encoder	87.45	−3.58
+ NFCE (Encoder + Decoder)	Dual-path NFCE applied in both encoder and decoder	88.92	−2.11
+ ETrans Bottleneck	Enhanced Transformer module added at the bottleneck	89.76	−1.27
Full SwinME	NFCE (encoder + decoder) combined with ETrans bottleneck	91.03	0.00

5.6.1 Baseline Configuration

The baseline configuration consists of a hierarchical transformer-based encoder–decoder architecture without the proposed enhancement components. This baseline retains the core encoder–decoder structure and window-based attention mechanism but excludes specific modules introduced to improve contextual modeling and feature refinement. The baseline serves as a reference point against which the impact of each additional component is measured.

5.6.2 Effect of Swin Transformer V2 Integration

To assess the contribution of Swin Transformer V2 integration, an ablation variant replaces the enhanced attention formulation with a standard window-based self-attention mechanism. Comparative results indicate that the inclusion of Swin Transformer V2 leads to more stable training behavior and improved segmentation performance across tumor subregions.

The observed improvements are particularly evident in challenging regions such as tumor core and enhancing tumor, where subtle intensity variations and ambiguous boundaries require robust contextual modeling. The enhanced attention formulation and improved positional encoding contribute to better feature discrimination and spatial consistency, validating the choice of Swin Transformer V2 within the SwinME framework.

5.6.3 Impact of the Multi-Scale Enhancement Strategy

Another ablation experiment evaluates the role of the multi-scale enhancement strategy by removing structured multi-scale feature fusion while retaining basic skip connections. In this configuration, hierarchical encoder features are not explicitly enhanced or adaptively integrated during decoding.

Results from this ablation demonstrate a noticeable decline in segmentation accuracy, particularly for small and heterogeneous tumor subregions. Visual inspection reveals reduced boundary precision and increased fragmentation in predicted segmentation maps. These findings highlight the importance of systematic multi-scale integration in capturing both fine-grained detail and global tumor context, confirming that multi-scale enhancement is a critical component of the SwinME architecture.

5.6.4 Contribution of the Enhanced Transformer (ETrans) Module

To evaluate the effectiveness of the ETrans module, an ablation variant is constructed by removing the feature refinement mechanism while keeping the remaining architecture unchanged. This experiment focuses on assessing the impact of explicit feature recalibration on segmentation performance.

The removal of the ETrans module results in reduced segmentation quality, particularly in regions with low contrast or complex anatomical boundaries. The model exhibits increased sensitivity to noise and minor inconsistencies across slices. These observations suggest that the ETrans module plays a significant role in reinforcing informative features and suppressing less relevant activations, thereby improving boundary delineation and volumetric consistency.

5.6.5 Combined Effect of Architectural Components

A final set of ablation experiments examines the cumulative effect of integrating all proposed components. Comparisons between partial and full configurations demonstrate that the best performance is achieved when Swin Transformer V2 integration, multi-scale enhancement, and the ETrans module are all enabled simultaneously. This outcome indicates that the components are complementary rather than redundant, each addressing different aspects of the segmentation challenge.

The combined configuration exhibits improved robustness across tumor subregions, enhanced boundary precision, and more consistent volumetric predictions. These results support the architectural design philosophy of SwinME, which emphasizes coherent integration of attention-based context modeling and multi-scale feature refinement.

5.6.6 Interpretation of Ablation Findings

The ablation studies provide clear empirical evidence that each major architectural component contributes meaningfully to the overall performance of the SwinME framework. The results confirm that improvements are not attributable to a single modification but arise from the coordinated interaction of hierarchical attention, multi-scale enhancement, and feature refinement mechanisms.

By systematically analyzing these components, the ablation studies validate the design choices presented in Chapter 4 and strengthen the causal link between architectural innovation and segmentation performance. This analysis also enhances the

interpretability of the proposed framework and provides guidance for future architectural refinements.

5.7 Discussion

5.7.1 Hierarchical Attention as a Driver of Spatial Coherence and Boundary Consistency

This section addresses Research Question 1, which investigates how hierarchical attention mechanisms can be adapted to enhance spatial coherence and boundary consistency in volumetric brain tumor segmentation. The discussion interprets the experimental findings presented in previous sections in this chapter to elucidate the functional role of hierarchical attention in overcoming long-standing limitations of three-dimensional medical image segmentation, particularly fragmentation across slices and instability in tumor boundary delineation.

The quantitative and qualitative evidence collectively indicates that hierarchical attention fundamentally alters how spatial context is encoded and propagated through the segmentation network. Rather than treating volumetric data as a collection of locally independent regions, hierarchical attention enables progressive contextual aggregation across multiple representational stages. Lower-level attention layers focus on localized textural and edge-based cues, while higher-level stages integrate broader anatomical context spanning multiple slices. This structured organization allows the model to reconcile fine-grained boundary information with global tumor morphology, which is essential for maintaining spatial continuity throughout the three-dimensional volume.

The impact of this hierarchical design is particularly evident in regions where boundary ambiguity and morphological heterogeneity are most pronounced. Tumor core and enhancing tumor subregions often exhibit weak contrast, irregular shapes, and gradual transitions across slices, making them highly susceptible to boundary fragmentation when segmentation decisions rely solely on local features. The experimental results in Chapter 5 demonstrate that hierarchical attention mitigates this vulnerability by enabling boundary decisions to be informed by contextual signals derived from adjacent slices and higher-order representations of tumor extent. Consequently, boundary delineation becomes more

stable and anatomically consistent, reducing spurious discontinuities that commonly arise in non-hierarchical or purely local attention frameworks.

Qualitative analyses further substantiate this interpretation. Visual inspection of representative cases reveals that segmentation outputs generated with hierarchical attention exhibit smoother contours, improved inter-slice alignment, and reduced noise-driven fluctuations along tumor boundaries. These observations suggest that hierarchical attention acts as an implicit regularization mechanism, constraining predictions to conform to plausible volumetric structures rather than isolated local patterns. Importantly, the improved spatial coherence observed is not limited to large tumor regions but extends to smaller and more heterogeneous subregions, underscoring the generality of the approach.

Ablation studies reported in this chapter provide additional insight into the necessity of hierarchical organization. When attention mechanisms are simplified or hierarchically weakened, segmentation outputs display increased fragmentation and reduced boundary precision, even when overall model capacity remains comparable. This finding confirms that the benefits of hierarchical attention arise from its structured integration of context across scales, rather than from increased parameter count or architectural complexity alone. Hierarchical attention thus functions as a principled mechanism for enforcing volumetric consistency, rather than a superficial architectural enhancement.

Training dynamics further supports the role of hierarchical attention in promoting coherent segmentation. The observed stability of optimization and close agreement between training and validation performance indicate that the model learns globally consistent representations without overfitting local noise or slice-specific artifacts. This behavior suggests that hierarchical attention introduces a favorable inductive bias, guiding the learning process toward anatomically meaningful and spatially coherent solutions.

The findings presented in this chapter provide a clear and theoretically grounded answer to Research Question 1. Hierarchical attention mechanisms improve spatial coherence and boundary consistency by enabling progressive, multi-level integration of local and global contextual information within volumetric MRI data. Through this adaptation, the segmentation framework overcomes key limitations of conventional approaches and

produces anatomically consistent, volumetrically coherent tumor delineations, thereby advancing the state of volumetric brain tumor segmentation.

5.7.2 Role of Structured Multi-Scale Feature Enhancement in Segmenting Heterogeneous Tumor Subregions

This section addresses Research Question 2, which investigates whether structured multi-scale feature enhancement improves the segmentation of heterogeneous tumor subregions in multimodal MRI data. The discussion builds upon the experimental findings in this hapter to explain the underlying mechanisms through which multi-scale representation contributes to robustness, consistency, and accuracy when segmenting tumor components that exhibit substantial spatial and radiological variability.

Heterogeneity is a defining characteristic of brain tumors, manifesting across multiple spatial scales and imaging modalities. Tumor subregions differ not only in size and shape, but also in intensity distribution, texture, and anatomical context. Whole tumor regions often present relatively coherent global structure, whereas subregions such as the tumor core and enhancing tumor are smaller, fragmented, and exhibit weak or ambiguous contrast. Segmentation frameworks that operate predominantly at a single spatial scale are inherently limited in their ability to accommodate this diversity, as they tend to overemphasize either local detail or global context at the expense of the other.

Structured multi-scale feature enhancement directly addresses this limitation by enabling coordinated representation learning across multiple spatial resolutions. Fine-scale features preserve local texture and boundary information essential for accurately delineating small or sharply defined tumor components. At the same time, coarser-scale representations encode broader anatomical context, capturing the spatial extent of tumor regions and their relationships to surrounding tissue. By explicitly structuring the interaction between these scales, the framework ensures that segmentation decisions are informed by complementary information rather than isolated feature cues.

The experimental results in this chapter indicate that this structured multi-scale integration is particularly beneficial for heterogeneous tumor subregions. Tumor core and enhancing tumor regions, which are most susceptible to misclassification due to intensity overlapping with healthy tissue and irregular morphology, show more consistent

delineation when multi-scale enhancement is employed. This suggests that access to global contextual cues stabilizes local predictions, allowing the model to disambiguate challenging regions by considering their position and relationship within the overall tumor structure. In effect, multi-scale enhancement enables the model to reconcile local uncertainty with global anatomical plausibility.

Qualitative analysis further supports this interpretation. Segmentation outputs generated with structured multi-scale enhancement exhibit smoother subregion boundaries and greater continuity across adjacent slices, particularly in regions where tumor appearance changes gradually rather than abruptly. In contrast, ablation configurations that weaken or remove multi-scale pathways exhibit increased boundary noise, inconsistent labeling of tumor subregions, and greater sensitivity to local intensity fluctuations. These observations indicate that multi-scale enhancement plays a stabilizing role in feature interpretation, reducing susceptibility to noise and local ambiguity.

Ablation studies provide direct evidence that the observed improvements are attributable to the organization of feature integration rather than increased model capacity alone. When multi-scale enhancement mechanisms are removed while maintaining comparable parameter counts, segmentation performance degrades, especially for heterogeneous subregions. This demonstrates that structured multi-scale design introduces a meaningful inductive bias that guides the model toward coherent interpretation across scales, rather than merely expanding representational complexity.

From a broader methodological perspective, these findings underscore the importance of explicitly modeling scale interactions in volumetric segmentation tasks. Brain tumor segmentation is inherently multi-scale, and effective handling of tumor heterogeneity requires architectures that can simultaneously reason about fine detail and global structure. Structured multi-scale feature enhancement provides a principled means of achieving this balance, enabling consistent segmentation behavior across tumor subregions with diverse morphological and radiological characteristics.

In summary, the experimental evidence presented in this chapter provides a clear and affirmative answer to Research Question 2. Structured multi-scale feature enhancement significantly improves the segmentation of heterogeneous tumor subregions in multimodal MRI data by integrating local and global information in a coordinated and

hierarchical manner. This design enhances robustness to variability in tumor appearance, stabilizes boundary delineation, and supports consistent volumetric interpretation, addressing a central challenge in brain tumor segmentation.

5.7.3 Balancing Segmentation Accuracy with Computational Efficiency in Transformer-Based Volumetric Models

This section addresses Research Question 3, which examines how segmentation accuracy can be balanced with computational efficiency in transformer-based three-dimensional segmentation models. While transformers offer powerful global context modeling capabilities, their application to volumetric medical imaging presents significant challenges due to high memory consumption, increased computational complexity, and scalability constraints. The discussion below interprets the experimental findings from priveous sections to explain how principled architectural and training choices enable an effective balance between performance and efficiency.

A central factor influencing this balance is the formulation of attention mechanisms. Fully global self-attention scales poorly with volumetric input size, making it impractical for high-resolution three-dimensional MRI data. The experimental results demonstrate that restricting attention to localized windows within a hierarchical structure significantly reduces computational complexity while preserving the ability to capture long-range dependencies through progressive feature aggregation. By distributing contextual reasoning across hierarchical stages, the model avoids the quadratic cost of global attention without sacrificing volumetric coherence or segmentation accuracy. This design choice highlights that efficiency can be achieved not by eliminating global context, but by structuring how global information is accumulated.

Model capacity also plays a critical role in balancing accuracy and efficiency. Chapter 5 shows that increasing depth, feature dimensionality, or attention heads beyond moderate levels yields diminishing performance gains while incurring substantial computational overhead. These findings indicate that segmentation accuracy does not scale linearly with model size in volumetric settings. Instead, there exists an operating regime in which representational capacity is sufficient to model tumor heterogeneity and spatial context, beyond which additional complexity primarily increases cost rather than performance. Efficiency-oriented configurations demonstrate that carefully constrained architectures

can retain strong segmentation quality while substantially reducing memory usage and training time.

Training-level strategies further contribute to computational feasibility. The use of gradient checkpointing illustrates a practical mechanism for managing memory consumption by trading increased computation for reduced peak memory usage. Experimental comparisons between checkpointed and non-checkpointed configurations reveal that memory-efficient training enables deeper or more expressive models to be trained under fixed hardware constraints, without altering model behavior or segmentation accuracy. This flexibility is particularly important in volumetric transformer models, where memory often constitutes the primary bottleneck.

From a broader perspective, balancing accuracy and efficiency has direct implications for reproducibility, scalability, and potential clinical applicability. Models that achieve high segmentation accuracy only under extreme computational requirements are difficult to reproduce and impractical for deployment beyond specialized research environments. The experimental evidence indicates that transformer-based volumetric segmentation can be made computationally feasible through hierarchical attention, capacity-aware design, and memory-efficient training strategies, enabling robust performance on standard research hardware.

Importantly, the results suggest that efficiency should not be viewed as a secondary constraint imposed after model design, but rather as an integral design objective. When efficiency considerations are incorporated into architectural decisions from the outset, it becomes possible to achieve segmentation performance that is both accurate and practically attainable. This perspective shifts the focus from maximizing model size to optimizing information flow and representation structure, which is particularly relevant for three-dimensional medical imaging tasks.

The findings presented in this chapter provide a clear and evidence-based answer to Research Question 3. Segmentation accuracy and computational efficiency can be jointly optimized in transformer-based three-dimensional segmentation models through hierarchical attention mechanisms, balanced architectural capacity, and memory-aware training strategies. These design principles demonstrate that high-quality volumetric

segmentation does not require prohibitive computational cost, thereby supporting the broader adoption of transformer-based methods in medical image analysis.

CHAPTER 6

CONCLUSION AND FUTURE WORK

This chapter concludes the thesis by synthesizing the research outcomes, explicitly highlighting the scholarly contributions of the study, and revisiting the research questions that guided the investigation. The chapter reflects on the objectives established at the outset of the research and evaluates the extent to which they have been addressed through the proposed methodology, experimental analysis, and discussion presented in the preceding chapters. Finally, it outlines limitations and identifies directions for future research.

6.1 Conclusion

This research addressed the challenge of accurate and robust volumetric brain tumor segmentation from multimodal MRI data, a task characterized by pronounced anatomical variability, heterogeneous tumor appearance, and complex three-dimensional spatial dependencies. To address these challenges, this thesis proposed SwinME, a transformer-based segmentation framework designed around hierarchical attention modeling, structured multi-scale feature enhancement, and efficiency-aware architectural principles. The framework was explicitly developed to overcome key limitations of convolution-dominant approaches, particularly their restricted ability to capture long-range contextual relationships and to maintain volumetric consistency across slices.

Comprehensive experimental evaluation on the BraTS benchmark demonstrated that the proposed framework achieves strong Dice similarity coefficients across all clinically relevant tumor subregions, including the whole tumor, tumor core, and enhancing tumor regions. These results indicate reliable overlap with expert annotations and confirm the model's effectiveness in segmenting both large-scale tumor structures and smaller, heterogeneous subregions. Qualitative analyses further corroborate these findings, revealing spatially coherent and anatomically plausible segmentation outputs with reduced boundary fragmentation and improved inter-slice consistency. Moreover, stable training and validation behavior across experiments suggests that the proposed transformer-based design generalizes well despite its high representational capacity.

Beyond empirical performance, the findings provide clear answers to the research questions that motivated this study. Hierarchical attention mechanisms were shown to play a critical role in improving spatial coherence and boundary consistency, as reflected in both quantitative metrics and volumetric continuity of predictions. Structured multi-scale feature enhancement contributed to robust segmentation of heterogeneous tumor subregions by enabling effective representation across varying spatial resolutions. At the same time, efficiency-aware architectural and training strategies allowed the model to achieve competitive performance under realistic computational constraints. Collectively, these results demonstrate that carefully adapted transformer-based architectures can serve as a powerful and practical solution for volumetric brain tumor segmentation, advancing the state of the art while offering a solid foundation for further methodological and clinical developments.

6.2 Future Work

Future research will focus on strengthening the generalizability and robustness of the proposed SwinME framework through broader and more diverse experimental validation. While the current study demonstrates strong performance on a standardized benchmark dataset, future work will extend evaluation to additional publicly available datasets such as UCSF-BMSR and BrainMetShare, which differ in acquisition protocols, scanner characteristics, tumor types, and patient populations. Evaluating the framework across these heterogeneous datasets will enable a more rigorous assessment of cross-dataset generalization and domain shift, particularly for tumor appearances and imaging conditions not fully represented in existing benchmarks. Such validation is essential for understanding the adaptability of hierarchical transformer-based segmentation models to real-world clinical variability.

In parallel, future work will involve a systematic investigation of hyperparameter configurations and their influence on segmentation accuracy, training stability, and generalization behavior. This includes analyzing architectural parameters such as network depth distribution, feature dimensionality, attention configuration, and regularization strategies, as well as training-related parameters that affect convergence and robustness. Conducting controlled hyperparameter studies across multiple datasets will help establish principled guidelines for configuring transformer-based volumetric segmentation models

under varying computational and clinical constraints. These efforts, combined with potential architectural refinements aimed at improving efficiency and scalability, are expected to contribute to the development of more reliable, adaptable, and clinically meaningful attention-based segmentation frameworks.

REFERENCES

Bakas, S., Reyes, M., Jakab, A., Bauer, S., Rempfler, M., Crimi, A., . . . Kirschke, J. (2018). Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge.

Bauer, S., Wiest, R., Nolte, L.-P., & Reyes, M. (2013). A survey of MRI-based medical image analysis for brain tumor studies. *Physics in Medicine and Biology, 58*, R97–R129.

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Huang, M. (2021). Swin-Unet: Unet-like pure transformer for medical image segmentation.

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

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder–decoder with atrous separable convolution for semantic image segmentation. *Proc. European Conf. on Computer Vision (ECCV).*

Çiçek, Ö., Abdulkadir, A., Lienkamp, S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning dense volumetric segmentation from sparse annotation. *Proc. MICCAI.* Springer, Cham.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., . . . Houlsby, N. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. *Proc. ICLR.*

Ellingson, B., Wen, P., & Cloughesy, T. (2017). Modified criteria for radiographic response assessment in glioblastoma clinical trials. *Neurotherapeutics, 14*, 307–320.

Esteva, A., Kuprel, B., Novoa, R., Ko, J., Swetter, S., Blau, H., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*, 115–118.

Gen, M., & Cheng, R. (1997). *Genetic Algorithms and Engineering Design.* John Wiley & Sons.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press.

Habesch, N., & Awadallah, F. (1999). Video image processing: A technology evaluation for freeway. *Proc. 6th World Congress on Intelligent Transport Systems.* Toronto.

Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Roth, H., & Xu, D. (2022). UNETR: Transformers for 3D medical image segmentation. *Proc. IEEE/CVF WACV.*

Isensee, F., Jaeger, P., Kohl, S., Petersen, J., & Maier-Hein, K. (2021). nnU-Net: A self-adapting framework for U-Net-based medical image segmentation. *Nature Methods, 18*, 203–211.

Jiang, Y., Zhang, Y., Lin, X., Dong, J., Cheng, T., & Liang, J. (2022). SwinBTS: A transformer-based architecture for 3D brain tumor segmentation.

Kamnitsas, K., Ledig, C., Newcombe, V., Simpson, J., Kane, A., Menon, D., . . . Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis, 36*, 61–78.

Kline, T., Korfiatis, P., Edwards, M., Blais, J., & Erickson, B. (2017). Performance of deep learning convolutional neural networks for classifying medical images. *Journal of Digital Imaging, 30*, 130–137.

Litjens, G., Kooi, T., Bejnordi, B., Setio, A., Ciompi, F., Ghafoorian, M., . . . Sánchez, C. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis, 42*, 60–88.

Liu, Z., Hu, H., Lin, Y., Yao, Z., Xie, Z., Wei, Y., . . . Lin, S. (2022). Swin Transformer V2: Scaling up capacity and resolution. *Proc. IEEE/CVF CVPR.*

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., . . . Guo, B. (2021). Swin Transformer: Hierarchical vision transformer using shifted windows. *Proc. IEEE ICCV.*

Louis, D., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W., . . . Ellison, D. (2016). The 2016 World Health Organization classification of tumors of the central nervous system. *Acta Neuropathologica, 131*, 803–820.

Milletari, F., Navab, N., & Ahmadi, S. (2016). V-Net: Fully convolutional neural networks for volumetric medical image segmentation. *Proc. 3DV.*

Myronenko, A. (2018). 3D MRI brain tumor segmentation using autoencoder regularization. *Proc. MICCAI.* Springer, Cham.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *Proc. MICCAI.* Springer, Cham.

Roth, H., Oda, H., Zhou, X., Shimizu, N., Yang, Y., Hayashi, Y., . . . Mori, K. (2018). An application of cascaded 3D fully convolutional networks for medical image segmentation. *Computerized Medical Imaging and Graphics, 66*, 90–99.

Shen, D., Wu, G., & Suk, H. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering, 19*, 221–248.

Tang, H., Chen, H., Yan, W., Cheng, Z., Chen, J., Xiao, H., & Zheng, Y. (2022). Swin-UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images. *Proc. IEEE ISBI.*

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., . . . Polosukhin, I. (2017). Attention is all you need. *Proc. NeurIPS*, *30.*

Wang, G., Li, W., Ourselin, S., & Vercauteren, T. (2017). Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. *Proc. MICCAI.*

Wang, Y., Zhao, L., Yang, G., Yu, L., Xu, J., Qin, J., & Heng, P.-A. (2021). Hybrid CNN–Transformer network for brain tumor segmentation. *Proc. MICCAI.*

Yang, Y., Yang, D., Li, W., Roth, H., Landman, B., & Hatamizadeh, A. (2022). Self-supervised pre-training of Swin transformers for 3D medical image analysis. *Proc. IEEE/CVF CVPR.*

Zhou, Z., Siddiquee, M., Tajbakhsh, N., & Liang, J. (2018). UNet++: A nested U-Net architecture for medical image segmentation. In *Medical Image Analysis and Multimodal Learning for Clinical Decision Support.* Springer, Cham.

APPENDIX

Appendix 1: Patient-Wise Segmentation Performance on the BraTS Dataset

Patient_ID	Dice_TC	Dice_WT	Dice_ET	Dice_Avg	Vol_Diff_TC	Vol_Diff_WT	Vol_Diff_ET
BraTS-GLI-00000-000	0.97806	0.921065	0.951613	0.950246	-152	7249	899
BraTS-GLI-00002-000	0.596979	0.672193	0.614297	0.627823	-19421	-91057	-12231
BraTS-GLI-00003-000	0.939084	0.973588	0.930635	0.947769	-1410	-1904	411
BraTS-GLI-00005-000	0.955425	0.942147	0.884521	0.927364	-1948	1552	-1102
BraTS-GLI-00006-000	0.951047	0.961201	0.900646	0.937631	4053	5859	3921
BraTS-GLI-00008-000	0.634312	0.920172	0.638541	0.731008	232	995	255
BraTS-GLI-00008-001	0.92887	0.912251	0.941878	0.927667	256	766	194
BraTS-GLI-00009-000	0.925957	0.667494	0.883619	0.82569	-205	-12985	-219
BraTS-GLI-00009-001	0.845591	0.664923	0.86446	0.791658	393	-9458	334
BraTS-GLI-00011-000	0.850497	0.951813	0.893942	0.898751	3275	1952	1481
BraTS-GLI-00012-000	0.245882	0.890461	0.036022	0.390788	-144579	-38397	11156
BraTS-GLI-00014-000	0.318793	0.117527	0.371375	0.269232	-2069	-35398	-1942
BraTS-GLI-00014-001	0.875026	0.857879	0.877364	0.87009	-2728	-2714	-2393
BraTS-GLI-00016-000	0.922985	0.904542	0.820938	0.882822	3377	-7763	4758
BraTS-GLI-00016-001	0.947312	0.858378	0.922132	0.909274	11	-8298	-1777
BraTS-GLI-00017-000	0.710378	0.932487	0.704225	0.782364	1360	-7117	877
BraTS-GLI-00017-001		0.926219			977	2092	803
BraTS-GLI-00018-000	0.750849	0.938194	0.734293	0.807779	-36612	-10436	-23961
BraTS-GLI-00019-000	0.781625	0.861368	0.832516	0.82517	6016	10799	2500
BraTS-GLI-00020-000	0.947793	0.931504	0.938801	0.939366	-1173	1405	1470
BraTS-GLI-00020-001	0.971588	0.934214	0.956593	0.954132	1269	2854	1480
BraTS-GLI-00021-000	0.331149	0.379096	0.275751	0.328665	-5365	-66960	-4820
BraTS-GLI-00021-001	0.835586	0.930485	0.743919	0.836663	-147	-9204	2142
BraTS-GLI-00022-000	0.865907	0.925118	0.873792	0.888272	-12322	1113	-5153

BraTS-GLI-00022-001	0.814059	0.888164	0.827781	0.843335	-8508	-5343	-3474
BraTS-GLI-00024-000	0.885734	0.916592	0.906428	0.902918	1041	-2813	-96
BraTS-GLI-00024-001	0.503264	0.916777	0.535131	0.651724	2942	-8452	2278
BraTS-GLI-00025-000	0.708852	0.805306	0.687508	0.733889	-2374	-9099	-2271
BraTS-GLI-00026-000	0.886317	0.858127	0.894668	0.879704	-5947	-15308	-4535
BraTS-GLI-00028-000	0.514904	0.826548	0.559552	0.633668	-9355	-4370	-8209
BraTS-GLI-00030-000	0.851821	0.908528	0.864711	0.87502	-207	5590	4408
BraTS-GLI-00031-000	0.941604	0.823458	0.892281	0.885781	306	-61320	687
BraTS-GLI-00031-001	0.867951	0.645068	0.789602	0.76754	3004	-116249	3542
BraTS-GLI-00032-000	0.9494	0.952227	0.926777	0.942801	-10	2499	-404
BraTS-GLI-00032-001	0.936202	0.969549	0.961345	0.955699	2610	142	337
BraTS-GLI-00033-000	0.518476	0.909651	0.665578	0.697902	1155	-12598	-1516
BraTS-GLI-00035-000	0.93487	0.946301	0.937403	0.939525	-740	1276	398
BraTS-GLI-00036-000	0.797504	0.943974	0.831859	0.857779	-12503	-1387	2646
BraTS-GLI-00036-001	0.940869	0.966793	0.914614	0.940759	4063	4119	2483
BraTS-GLI-00043-000	0.677183	0.966404	0.364303	0.669297	-8289	980	-8343
BraTS-GLI-00044-000	0.960938	0.959223	0.950467	0.956876	392	914	283
BraTS-GLI-00045-000	0.969446	0.92772	0.916353	0.937839	541	3728	736
BraTS-GLI-00045-001	0.919701	0.925633	0.927079	0.924138	1200	2450	950
BraTS-GLI-00046-000	0.913465	0.964785	0.953443	0.943898	1276	1194	129
BraTS-GLI-00048-000	0.939758	0.880067	0.941951	0.920592	1974	5986	322
BraTS-GLI-00048-001		0.916087			663	-234	616
BraTS-GLI-00049-000	0.963302	0.947579	0.940616	0.950499	-187	-10848	-624
BraTS-GLI-00051-000	0.949478	0.920469	0.958529	0.942826	2059	13063	1068
BraTS-GLI-00052-000	0.894718	0.925714	0.86151	0.893981	-13377	3742	-8194
BraTS-GLI-00053-000	0.907647	0.944915	0.87206	0.908207	-1518	-7967	-1445
BraTS-GLI-00053-001	0.951764	0.972594	0.951375	0.958578	2599	-2442	-1254
BraTS-GLI-00054-000	0.759486	0.911789	0.746944	0.806073	-34600	-8015	-13020
BraTS-GLI-00056-000	0.930359	0.947807	0.903791	0.927319	-4619	4879	-6177

BraTS-GLI-00056-001	0.901431	0.953519	0.928592	0.927847	5653	8975	-2086
BraTS-GLI-00058-000	0.967826	0.857258	0.932197	0.919093	322	7481	318
BraTS-GLI-00058-001	0.859057	0.6451	0.839604	0.781253	4205	50138	2389
BraTS-GLI-00059-000	0.968468	0.936774	0.942705	0.949316	471	-7385	-332
BraTS-GLI-00059-001	0.959989	0.885789	0.923907	0.923228	-265	9841	-1232
BraTS-GLI-00060-000	0.829148	0.927813	0.878169	0.878377	-28494	-12184	-8468
BraTS-GLI-00061-000	0.844065	0.811087	0.627694	0.760949	-1966	-4510	785
BraTS-GLI-00061-001	0.573462	0.844479	0.629817	0.682586	-2490	8587	-341
BraTS-GLI-00062-000	0.651737	0.967928	0.797791	0.805818	6718	-3232	2750
BraTS-GLI-00063-000	0.67694	0.547301	0.680688	0.634976	8793	29709	7302
BraTS-GLI-00064-000	0.793639	0.900572	0.768774	0.820995	-15878	3890	-6142
BraTS-GLI-00066-000	0.981412	0.964214	0.949667	0.965097	153	1401	-383
BraTS-GLI-00068-000	0.945956	0.955264	0.942598	0.947939	1634	-2235	722
BraTS-GLI-00070-000	0.965566	0.972262	0.921903	0.953243	1482	1465	-2499
BraTS-GLI-00071-000	0.965912	0.951797	0.925098	0.947602	-406	2872	-36
BraTS-GLI-00072-000	0.963092	0.873554	0.925314	0.920653	-1616	18291	65
BraTS-GLI-00072-001	0.953228	0.887883	0.968917	0.936676	15	6161	-286
BraTS-GLI-00074-000	0.941567	0.968994	0.89715	0.935903	3046	3514	-1791
BraTS-GLI-00077-000	0.91464	0.916744	0.946744	0.926043	-3189	7347	900
BraTS-GLI-00078-000	0.976096	0.962495	0.958045	0.965545	342	1947	132
BraTS-GLI-00081-000	0.977261	0.972957	0.9692	0.973139	278	-717	-603
BraTS-GLI-00084-000	0.919767	0.92408	0.884009	0.909285	-1649	4915	-2742
BraTS-GLI-00084-001	0.299526	0.433214	0.297501	0.343414	1652	-33507	1553
BraTS-GLI-00085-000	0.719675	0.893658	0.787666	0.800333	-310	-10386	306
BraTS-GLI-00087-000	0.844728	0.904529	0.781958	0.843738	-1955	-13776	-1086
BraTS-GLI-00087-001	0.886358	0.947573	0.867775	0.900569	2920	-5326	2395
BraTS-GLI-00088-000	0.973871	0.941633	0.928601	0.948035	-240	10240	505
BraTS-GLI-00088-001	0.920262	0.938657	0.894934	0.917951	-1671	-3148	-569
BraTS-GLI-00089-000	0.977404	0.907018	0.962453	0.948959	617	4821	83

BraTS-GLI-00090-000	0.889689	0.929932	0.913687	0.911103	1884	5065	16
BraTS-GLI-00090-001	0.745409	0.951913	0.92462	0.873981	6117	5810	-789
BraTS-GLI-00094-000	0.786352	0.958209	0.821041	0.8552	-30570	8837	-5879
BraTS-GLI-00095-000	0.968323	0.962789	0.89101	0.940707	-277	4904	1165
BraTS-GLI-00095-001	0.980735	0.968231	0.947359	0.965441	-1916	-8172	-1530
BraTS-GLI-00096-000	0.932577	0.926451	0.92539	0.928139	3421	16235	1076
BraTS-GLI-00096-001	0.937187	0.945525	0.958074	0.946929	2669	10622	686
BraTS-GLI-00097-000	0.937682	0.953411	0.930016	0.94037	-380	2685	892
BraTS-GLI-00097-001	0.975458	0.929975	0.967096	0.95751	518	3317	407
BraTS-GLI-00098-000	0.159238	0.492814	0.201757	0.284603	-47818	-39382	-34470
BraTS-GLI-00098-001	0.904339	0.944677	0.935367	0.928127	-10071	-6903	-6491
BraTS-GLI-00099-000	0.927196	0.858954	0.858053	0.881401	-1202	-5437	197
BraTS-GLI-00099-001	0.861083	0.900553	0.889199	0.883612	-1293	-16430	-1019
BraTS-GLI-00100-000	0.944545	0.964766	0.936661	0.948658	210	3464	-866
BraTS-GLI-00101-000	0.960045	0.955581	0.94212	0.952582	895	5160	229
BraTS-GLI-00102-000	0.397073	0.760456	0.40926	0.522263	2847	2095	2764
BraTS-GLI-00103-000	0.937161	0.911893	0.88873	0.912595	1544	-29573	412
BraTS-GLI-00104-000	0.961796	0.959899	0.947407	0.956367	42	498	1052
BraTS-GLI-00105-000	0.857718	0.914931	0.836726	0.869792	-11582	-8885	2989
BraTS-GLI-00106-000	0.899019	0.937579	0.836917	0.891172	-6750	-14050	3592
BraTS-GLI-00107-000	0.845873	0.958092	0.928055	0.910673	3455	-720	-2088
BraTS-GLI-00108-000	0.591077	0.879777	0.711727	0.727527	1337	1849	-518
BraTS-GLI-00109-000	0.56342	0.8032	0.7876	0.718074	-23925	-40193	-2106
BraTS-GLI-00110-000	0.955878	0.962006	0.982625	0.966836	810	3205	35
BraTS-GLI-00111-000	0.64	0.883956	0.531102	0.685019	-17783	-34832	6722
BraTS-GLI-00112-000	0.524099	0.893038	0.42356	0.613566	22684	-8786	25028
BraTS-GLI-00113-000	0.691817	0.74709	0.632911	0.690606	-487	-10992	-480
BraTS-GLI-00115-000	0.957208	0.926164	0.938555	0.940642	-3358	11820	-1859
BraTS-GLI-00116-000	0.326174	0.727608	0.366481	0.473421	-44250	-69553	-34279

BraTS-GLI-00117-000	0.57438	0.808791	0.64752	0.676897	-10101	-19543	-886
BraTS-GLI-00118-000	0.972903	0.952452	0.95977	0.961708	1595	4382	218
BraTS-GLI-00120-000	0.793493	0.912367	0.856353	0.854071	-18149	-5207	-8104
BraTS-GLI-00121-000	0.957372	0.906058	0.94835	0.93726	1118	-199	788
BraTS-GLI-00122-000	0.925602	0.950095	0.756648	0.877448	535	185	8847
BraTS-GLI-00123-000	0.966798	0.93146	0.935416	0.944558	-1179	-14061	-4155
BraTS-GLI-00124-000	0.937809	0.894658	0.597524	0.809997	-5615	-31918	50266
BraTS-GLI-00126-000	0.91692	0.965168	0.899927	0.927338	-1099	459	232
BraTS-GLI-00127-000	0.850032	0.961057	0.868546	0.893211	1483	-4749	283
BraTS-GLI-00128-000	0.940433	0.959421	0.908117	0.93599	-264	9519	-2529
BraTS-GLI-00130-000	0.835635	0.934424	0.788297	0.852785	-8976	-1022	3546
BraTS-GLI-00131-000	0.880332	0.939613	0.873777	0.897907	673	6581	2120
BraTS-GLI-00132-000	0.955568	0.915568	0.932479	0.934538	652	18600	-614
BraTS-GLI-00133-000	0.955479	0.901408	0.909138	0.922008	-209	-78	1220
BraTS-GLI-00134-000	0.474397	0.919892	0.764424	0.719571	-72573	-22879	1456
BraTS-GLI-00136-000	0.851577	0.926623	0.807023	0.861741	5551	-2925	6963
BraTS-GLI-00137-000	0.961288	0.946227	0.935459	0.947658	1480	1790	-2844
BraTS-GLI-00138-000	0.781844	0.895521	0.751677	0.80968	28	-19318	1813
BraTS-GLI-00139-000	0.878607	0.914761	0.77496	0.856109	2859	-7054	9569
BraTS-GLI-00140-000	0.596139	0.876194	0.599501	0.690611	3661	3562	3530
BraTS-GLI-00142-000	0.937247	0.930774	0.883885	0.917302	-774	-413	5698
BraTS-GLI-00143-000	0.941206	0.825911	0.93328	0.900132	3088	44089	775
BraTS-GLI-00144-000	0.95706	0.906233	0.885354	0.916216	1861	8081	3388
BraTS-GLI-00146-000	0.806778	0.8996	0.612534	0.772971	-22732	-20036	16343
BraTS-GLI-00147-000	0.954471	0.95614	0.877573	0.929395	-2684	-11225	-7396
BraTS-GLI-00148-000	0.90372	0.889831	0.877234	0.890262	-1996	-6523	-2720
BraTS-GLI-00149-000	0.800108	0.695361	0.676727	0.724065	-2584	-10022	1945
BraTS-GLI-00150-000	0.772238	0.932171	0.791409	0.831939	7457	-3265	7813
BraTS-GLI-00151-000	0.9293	0.917776	0.907201	0.918092	-6282	-15236	-4244

BraTS-GLI-00152-000	0.920445	0.882428	0.761213	0.854695	-242	4418	4743
BraTS-GLI-00154-000	0.857465	0.932372	0.72977	0.839869	3209	6638	9217
BraTS-GLI-00155-000	0.322581	0.745951	0.342534	0.470355	2036	-5548	2329
BraTS-GLI-00156-000	0.975942	0.959084	0.979828	0.971618	2848	11572	-348
BraTS-GLI-00157-000	0.732022	0.856889	0.752055	0.780322	5587	13639	4084
BraTS-GLI-00158-000	0.955819	0.941482	0.944395	0.947232	4292	11997	3133
BraTS-GLI-00159-000	0.946969	0.933556	0.90908	0.929869	1046	-1427	2801
BraTS-GLI-00160-000	0.456319	0.946228	0.612573	0.671706	13980	12976	4685
BraTS-GLI-00162-000	0.68784	0.895584	0.707519	0.763647	3644	5702	2589
BraTS-GLI-00165-000	0.947483	0.972602	0.933712	0.951266	1869	334	442
BraTS-GLI-00166-000	0.932447	0.95625	0.928938	0.939212	4561	7313	3641
BraTS-GLI-00167-000	0.773029	0.951506	0.811968	0.845501	2396	-2464	1724
BraTS-GLI-00170-000	0.562324	0.920761	0.634756	0.705947	401	-751	295
BraTS-GLI-00171-000	0.896789	0.87374	0.880604	0.883711	-2984	-8957	-3869
BraTS-GLI-00172-000	0.757312	0.922295	0.856822	0.845476	-13	7828	1901
BraTS-GLI-00176-000	0.892844	0.912398	0.916742	0.907328	-8924	9269	-85
BraTS-GLI-00177-000	0.960591	0.955714	0.932557	0.949621	225	10343	589
BraTS-GLI-00178-000	0.947929	0.957535	0.964997	0.95682	4255	6453	507
BraTS-GLI-00183-000	0.960602	0.9511	0.96309	0.958264	3349	11633	407
BraTS-GLI-00184-000	0.868584	0.95318	0.876527	0.89943	3843	76	2634
BraTS-GLI-00185-000	0.932254	0.542716	0.914457	0.796476	-651	-18121	-49
BraTS-GLI-00186-000	0.843464	0.971786	0.862262	0.892504	7346	298	4144
BraTS-GLI-00187-000	0.969017	0.964486	0.956418	0.963307	1776	-3054	-2145
BraTS-GLI-00188-000	0.981228	0.955762	0.961899	0.966296	19	2831	451
BraTS-GLI-00191-000	0.904612	0.937495	0.943536	0.928548	4151	7545	344
BraTS-GLI-00192-000	0.939218	0.936682	0.963788	0.946563	1708	2295	683
BraTS-GLI-00193-000	0.959942	0.848274	0.940615	0.916277	614	-16522	-1125
BraTS-GLI-00194-000	0.916692	0.941146	0.900861	0.919566	243	5174	789
BraTS-GLI-00195-000	0.819208	0.965615	0.884365	0.889729	2131	1339	1092

BraTS-GLI-00196-000	0.969951	0.926243	0.970614	0.955603	1426	10704	-221
BraTS-GLI-00199-000	0.965233	0.963104	0.937696	0.955344	2212	3170	2064
BraTS-GLI-00201-000	0.816901	0.951219	0.870747	0.879622	-17794	-1972	-4895
BraTS-GLI-00203-000	0.757797	0.950564	0.8612 8	0.856547	-6369	2529	-1043
BraTS-GLI-00204-000	0.872908	0.74598	0.8179 1	0.812266	230	-11007	480
BraTS-GLI-00206-000	0.957218	0.866892	0.958414	0.927508	964	-1898	449
BraTS-GLI-00207-000	0.579932	0.860446	0.779182	0.739853	-24431	-27105	-747
BraTS-GLI-00209-000	0.9489	0.8996	0.899073	0.915858	742	7967	2085
BraTS-GLI-00210-000	0.949589	0.943818	0.935828	0.943078	-181	35	-829
BraTS-GLI-00211-000	0.960487	0.969405	0.944023	0.957972	-770	4402	1341
BraTS-GLI-00212-000	0.841409	0.618267	0.746614	0.73543	871	-36673	1313
BraTS-GLI-00214-000	0.985043	0.962398	0.972307	0.973249	448	3400	-1010
BraTS-GLI-00216-000	0.976702	0.964727	0.965688	0.969039	1388	8281	457
BraTS-GLI-00217-000	0.78485	0.969583	0.932484	0.895639	8117	2013	2012
BraTS-GLI-00218-000	0.881123	0.513682	0.895238	0.763348	391	-7512	260
BraTS-GLI-00219-000	0.94854	0.814473	0.942447	0.90182	612	2907	631
BraTS-GLI-00220-000	0.89493	0.878939	0.900224	0.891364	3538	9692	2083
BraTS-GLI-00221-000	0.92347	0.936544	0.850146	0.903387	-386	2902	1940
BraTS-GLI-00222-000	0.782691	0.95157	0.811666	0.848643	1047	623	501
BraTS-GLI-00227-000	0.881085	0.630081	0.746601	0.752589	83	-14636	-37
BraTS-GLI-00228-000	0.934997	0.922915	0.922371	0.926761	2232	3723	2419
BraTS-GLI-00230-000	0.941858	0.943684	0.858599	0.914713	-2130	2037	185
BraTS-GLI-00231-000	0.894878	0.927358	0.872421	0.898219	-252	5402	727
BraTS-GLI-00233-000	0.927429	0.965485	0.950293	0.947736	2206	4011	-1261
BraTS-GLI-00234-000	0.747715	0.828479	0.777393	0.784529	-22261	-27733	-9870
BraTS-GLI-00235-000	0.977556	0.955384	0.927992	0.953644	1414	7399	2177
BraTS-GLI-00236-000	0.936722	0.921161	0.919088	0.925657	-2555	-3368	-2320
BraTS-GLI-00237-000	0.916834	0.89331	0.898861	0.903002	3722	12446	2044
BraTS-GLI-00238-000	0.95433	0.909893	0.966843	0.943688	1756	10674	-276

BraTS-GLI-00239-000	0.40903	0.850693	0.812895	0.690872	-28541	-5772	517
BraTS-GLI-00240-000	0.927306	0.972934	0.868808	0.923016	4571	5265	6402
BraTS-GLI-00241-000	0.823568	0.92329	0.889898	0.878919	2725	6594	1461
BraTS-GLI-00242-000	0.985725	0.704617	0.956497	0.88228	170	-67047	1008
BraTS-GLI-00243-000	0.388323	0.906405	0.470005	0.588244	9295	-10097	6588
BraTS-GLI-00246-000	0.923798	0.920586	0.91344	0.919275	2088	3450	1622
BraTS-GLI-00247-000	0.9155	0.947858	0.934092	0.932483	6513	7322	2396
BraTS-GLI-00249-000	0.95608	0.95582	0.93337	0.948423	-277	10657	-741
BraTS-GLI-00250-000	0.892871	0.922299	0.949102	0.921424	2838	-12265	854
BraTS-GLI-00251-000	0.950405	0.946065	0.898577	0.931683	-193	1065	179
BraTS-GLI-00253-000	0.813733	0.931332	0.867124	0.870729	1568	3417	842
BraTS-GLI-00254-000	0.940597	0.944719	0.948665	0.94466	2960	7917	863
BraTS-GLI-00258-000	0.756255	0.603637	0.753818	0.70457	1334	6482	1316
BraTS-GLI-00259-000	0.970012	0.969709	0.948079	0.9626	1260	45	1225
BraTS-GLI-00260-000	0.964981	0.883198	0.923155	0.923778	191	-2870	999
BraTS-GLI-00261-000	0.97961	0.954895	0.973713	0.969406	1651	10088	512
BraTS-GLI-00262-000	0.930279	0.868239	0.911951	0.90349	319	3264	335
BraTS-GLI-00263-000	0.951712	0.977217	0.942253	0.957061	2447	1376	1256
BraTS-GLI-00266-000	0.943504	0.962058	0.930936	0.945499	-3313	8217	1050
BraTS-GLI-00267-000	0.968482	0.951373	0.961109	0.960321	963	3054	579
BraTS-GLI-00269-000	0.973606	0.954212	0.950352	0.95939	461	1879	1266
BraTS-GLI-00270-000	0.892336	0.582509	0.897119	0.790655	990	-27155	982
BraTS-GLI-00271-000	0.905883	0.958288	0.917085	0.927085	-1077	-3986	-3001
BraTS-GLI-00273-000	0.882868	0.927735	0.928827	0.913143	-3434	4372	1140
BraTS-GLI-00274-000	0.931402	0.900931	0.845904	0.892746	1375	8725	2230
BraTS-GLI-00275-000	0.89662	0.912493	0.88553	0.898214	1674	240	1302
BraTS-GLI-00280-000	0.856716	0.806207	0.838071	0.833665	-1755	-6786	-1366
BraTS-GLI-00281-000	0.917279	0.969389	0.92318	0.936616	3845	1994	2393
BraTS-GLI-00282-000	0.919021	0.942171	0.913867	0.92502	1743	3142	1431

BraTS-GLI-00283-000	0.934931	0.906011	0.915595	0.918846	-2310	-7164	-36
BraTS-GLI-00284-000	0.673317	0.692503	0.642286	0.669369	382	-6433	676
BraTS-GLI-00285-000	0.641391	0.786055	0.655664	0.69437	2243	3376	1993
BraTS-GLI-00286-000	0.898598	0.786358	0.924637	0.869864	2376	-14450	1255
BraTS-GLI-00288-000	0.939329	0.9061	0.942295	0.929242	-63	-816	-45
BraTS-GLI-00289-000	0.880016	0.865121	0.881549	0.875562	4095	11046	3942
BraTS-GLI-00290-000	0.810224	0.925157	0.751789	0.829056	6757	12960	5488
BraTS-GLI-00291-000	0.841387	0.938475	0.840043	0.873302	4602	1602	4475
BraTS-GLI-00292-000	0.787669	0.782358	0.522751	0.697593	483	-5455	2312
BraTS-GLI-00293-000	0.443876	0.897519	0.478579	0.606658	2296	-1862	1953
BraTS-GLI-00294-000	0.725751	0.975741	0.817466	0.839652	15927	2784	5154
BraTS-GLI-00296-000	0.920615	0.878138	0.925906	0.908219	1098	3247	480
BraTS-GLI-00297-000	0.88751	0.89403	0.767133	0.849557	7001	1931	10056
BraTS-GLI-00298-000	0.924627	0.888462	0.939268	0.917452	1884	5680	1140
BraTS-GLI-00299-000	0.719232	0.928884	0.820242	0.822786	2100	5789	476
BraTS-GLI-00300-000	0.932615	0.945467	0.928316	0.935466	6194	15223	3911
BraTS-GLI-00301-000	0.970452	0.949362	0.955124	0.958313	1820	7224	1022
BraTS-GLI-00303-000	0.954498	0.953479	0.952762	0.953579	2908	7641	2041
BraTS-GLI-00304-000	0.913149	0.954576	0.880017	0.915914	-3444	2838	2090
BraTS-GLI-00305-000	0.934196	0.910792	0.917461	0.920816	565	4465	461
BraTS-GLI-00306-000	0.966623	0.946606	0.933856	0.949028	-1777	6375	-1788
BraTS-GLI-00309-000	0.931695	0.927178	0.916802	0.925225	908	5308	969
BraTS-GLI-00310-000	0.950322	0.93695	0.890861	0.926044	2936	7050	5308
BraTS-GLI-00311-000	0.973602	0.936868	0.943598	0.951356	268	2414	500
BraTS-GLI-00312-000	0.861167	0.953435	0.940622	0.918408	184	1207	129
BraTS-GLI-00313-000	0.828991	0.886083	0.598465	0.77118	1800	-442	4866
BraTS-GLI-00314-000	0.95338	0.765025	0.952351	0.890252	12	5660	-52
BraTS-GLI-00316-000	0.95389	0.961639	0.964676	0.960068	1118	703	606
BraTS-GLI-00317-000	0.941077	0.969739	0.966597	0.959138	1572	1262	70

BraTS-GLI-00318-000	0.954069	0.930397	0.891323	0.925263	291	-6054	9
BraTS-GLI-00320-000	0.844814	0.851881	0.710126	0.802274	2013	3882	3308
BraTS-GLI-00321-000	0.210835	0.619395	0.154326	0.328185	244	-6021	1071
BraTS-GLI-00322-000	0.936162	0.958188	0.913224	0.935858	-8531	-1456	-2039
BraTS-GLI-00324-000	0.873291	0.905786	0.810142	0.863073	2140	-4800	2696
BraTS-GLI-00325-000	0.915613	0.970838	0.894315	0.926922	-1985	-2080	-5552
BraTS-GLI-00327-000	0.847921	0.936984	0.942483	0.909129	6796	11083	1569
BraTS-GLI-00328-000	0.90487	0.888905	0.800304	0.864693	4670	10401	8364
BraTS-GLI-00329-000	0.93216	0.904671	0.925857	0.920896	867	2340	707
BraTS-GLI-00331-000	0	6.714339e-05	0	2.238113e-05	-6592	-23825	439
BraTS-GLI-00332-000	0.926787	0.896516	0.922141	0.915148	-561	-2151	82
BraTS-GLI-00334-000	0.950064	0.943984	0.946487	0.946845	363	5016	95
BraTS-GLI-00336-000	0.956093	0.949619	0.930472	0.945395	2982	12606	2426
BraTS-GLI-00338-000	0.804526	0.91038	0.857591	0.857499	3078	-6791	2034
BraTS-GLI-00339-000	0.915656	0.971303	0.951491	0.94615	1170	153	479
BraTS-GLI-00340-000	0.888362	0.760849	0.897549	0.84892	117	-9165	514
BraTS-GLI-00341-000	0.931924	0.862657	0.890247	0.894943	872	-558	1180
BraTS-GLI-00343-000	0.845741	0.899344	0.890756	0.878614	2143	2756	1224
BraTS-GLI-00344-000	0.93159	0.829053	0.878155	0.879599	1182	-6590	1088
BraTS-GLI-00346-000	0.795207	0.949796	0.869431	0.871478	1316	3682	670
BraTS-GLI-00347-000	0.95302	0.951784	0.901159	0.935321	2787	6654	3489
BraTS-GLI-00348-000	0.722592	0.897797	0.735822	0.785404	912	4066	716
BraTS-GLI-00349-000	0.9586	0.946775	0.930395	0.945257	-662	7344	-1852
BraTS-GLI-00350-000	0.924619	0.942798	0.932855	0.933424	1467	6011	-1107
BraTS-GLI-00351-000	0.973678	0.925813	0.963905	0.954465	-105	-3377	-1418
BraTS-GLI-00352-000	0.778653	0.89276	0.685355	0.78559	9460	9870	10448
BraTS-GLI-00353-000	0.95085	0.920174	0.965323	0.945449	3546	13113	-23
BraTS-GLI-00356-000	0.93764	0.933621	0.912941	0.928067	2086	4903	1147
BraTS-GLI-00359-000	0.943694	0.946513	0.957888	0.949365	1808	6246	545

BraTS-GLI-00360-000	0.946223	0.896975	0.966087	0.936428	-798	911	-499
BraTS-GLI-00364-000	0.962862	0.970747	0.969621	0.967743	2208	3987	510
BraTS-GLI-00366-000	0.741339	0.925846	0.441259	0.702815	-26649	2185	3854
BraTS-GLI-00367-000	0.844644	0.878085	0.847917	0.856882	-1937	-13172	-2026
BraTS-GLI-00369-000	0.835322	0.969862	0.910185	0.905123	7659	-8584	2663
BraTS-GLI-00370-000	0.959331	0.936032	0.928915	0.941426	1466	9027	2584
BraTS-GLI-00371-000	0.866961	0.964447	0.891295	0.907567	2681	3876	1585
BraTS-GLI-00373-000	0.923094	0.891675	0.944847	0.919872	1160	4022	303
BraTS-GLI-00375-000	0.824308	0.946913	0.782421	0.851214	2264	7533	4058
BraTS-GLI-00376-000	0.693111	0.927047	0.776623	0.798927	2364	4573	1394
BraTS-GLI-00377-000	0.928546	0.916322	0.899808	0.914892	963	4620	1681
BraTS-GLI-00378-000	0.760502	0.897539	0.757111	0.805051	2858	7202	2251
BraTS-GLI-00379-000	0.888388	0.916796	0.858126	0.88777	-756	2128	-75
BraTS-GLI-00380-000	0.287989	0.83989	0.268874	0.465584	-461	1101	-458
BraTS-GLI-00382-000	0.910132	0.900069	0.83087	0.880357	-1404	3664	1534
BraTS-GLI-00383-000	0.941688	0.935832	0.937758	0.938426	-1577	-2025	-983
BraTS-GLI-00386-000	0.960607	0.918187	0.892679	0.923824	1204	9825	2381
BraTS-GLI-00387-000	0.779031	0.879992	0.876829	0.845284	-2178	2062	-916
BraTS-GLI-00388-000	0.883859	0.947678	0.825154	0.885564	-6678	-7390	-10841
BraTS-GLI-00389-000	0.923575	0.921486	0.87557	0.906877	690	2225	887
BraTS-GLI-00390-000	0.691289	0.780529	0.634731	0.702183	383	2845	452
BraTS-GLI-00391-000	0.860536	0.916606	0.8402	0.872447	-12859	7498	5594
BraTS-GLI-00392-000	0.979742	0.936111	0.969024	0.961626	713	5526	-291
BraTS-GLI-00395-000	0.903485	0.954804	0.932971	0.93042	1159	4601	-207
BraTS-GLI-00397-000	0.861151	0.925278	0.934243	0.906891	-21826	3698	-4550
BraTS-GLI-00399-000	0.977601	0.950891	0.952173	0.960222	473	2261	385
BraTS-GLI-00400-000	0.899671	0.934651	0.834932	0.889751	-9021	-17934	-5136
BraTS-GLI-00401-000	0.92107	0.916801	0.871627	0.903166	-4159	21911	7323
BraTS-GLI-00402-000	0.973008	0.951192	0.948337	0.957512	-1046	4690	-2184

BraTS-GLI-00403-000	0.961477	0.891205	0.956278	0.93632	1247	8544	665
BraTS-GLI-00404-000	0.883098	0.945065	0.861836	0.896666	-5404	-8401	-1330
BraTS-GLI-00405-000	0.925297	0.951376	0.948251	0.941641	3708	6525	-842
BraTS-GLI-00406-000	0.904978	0.917329	0.905338	0.909215	2861	9432	2068
BraTS-GLI-00407-000	0.927844	0.929145	0.885381	0.914123	1909	4025	775
BraTS-GLI-00409-000	0.970733	0.941652	0.952546	0.954977	1170	6482	506
BraTS-GLI-00410-000	0.872887	0.849559	0.88983	0.870758	2451	-4109	1744
BraTS-GLI-00412-000	0.004713	0.938647	0.007018	0.316792	9292	6415	6226
BraTS-GLI-00413-000	0.974245	0.96204	0.9683	0.968195	1476	6754	-307
BraTS-GLI-00414-000	0.869245	0.937915	0.819622	0.875594	-11438	11639	9686
BraTS-GLI-00416-000	0.961482	0.963981	0.959658	0.961707	-4269	3686	-2381
BraTS-GLI-00417-000	0.852987	0.896913	0.799865	0.849922	7477	-28200	5114
BraTS-GLI-00418-000	0.97657	0.852211	0.9785	0.93576	730	-23084	-166
BraTS-GLI-00419-000	0.911379	0.786009	0.852512	0.849966	883	4207	1001
BraTS-GLI-00421-000	0.887794	0.97353	0.940385	0.933903	4041	2815	1540
BraTS-GLI-00423-000	0.945783	0.934441	0.938115	0.939446	-952	1771	17
BraTS-GLI-00425-000	0.534177	0.89368	0.566976	0.664944	3454	4759	3019
BraTS-GLI-00426-000	0.971525	0.946356	0.967317	0.961733	1944	12977	-495
BraTS-GLI-00429-000	0.941962	0.914725	0.959036	0.938574	139	-12971	-1723
BraTS-GLI-00430-000	0.871597	0.928069	0.902581	0.900749	3880	7779	1192
BraTS-GLI-00431-000	0.773115	0.782722	0.743034	0.76629	2457	4963	2606
BraTS-GLI-00432-000	0.115831	0.839845	0.132052	0.362576	-449	-43306	282
BraTS-GLI-00433-000	0.970801	0.958015	0.962866	0.963894	832	2493	723
BraTS-GLI-00436-000	0.961983	0.972377	0.938199	0.957519	-3706	663	-3879
BraTS-GLI-00440-000	0.840086	0.915318	0.884961	0.880121	708	-11206	1261
BraTS-GLI-00441-000	0.979226	0.851467	0.956463	0.929052	-417	13164	-1536
BraTS-GLI-00442-000	0.726767	0.91617	0.86768	0.836872	-10657	-5802	251
BraTS-GLI-00443-000	0.830765	0.890428	0.828892	0.850029	11937	13608	7390
BraTS-GLI-00444-000	0.85628	0.847277	0.890727	0.864762	3934	6592	2062

BraTS-GLI-00445-000	0.939177	0.93181	0.959532	0.943506	2569	13537	292
BraTS-GLI-00446-000	0.955991	0.965398	0.94655	0.95598	-5170	5657	-4851
BraTS-GLI-00448-000	0.970428	0.962054	0.946363	0.959615	-429	3333	160
BraTS-GLI-00449-000	0.843606	0.938124	0.865973	0.882568	-3219	7906	-14690
BraTS-GLI-00451-000	0.954197	0.968527	0.944666	0.955796	-4629	-4287	-5412
BraTS-GLI-00452-000	0.876491	0.95942	0.932867	0.922926	4389	-5944	860
BraTS-GLI-00453-000	0.952374	0.969256	0.93692	0.95285	1840	3756	1360
BraTS-GLI-00454-000	0.900189	0.940813	0.921257	0.920753	-8690	10344	-2453
BraTS-GLI-00455-000	0.956731	0.963521	0.955642	0.958632	1215	-1430	248
BraTS-GLI-00456-000	0.938983	0.96105	0.924314	0.941449	655	2221	-2345
BraTS-GLI-00457-000	0.951106	0.940892	0.923603	0.938534	-1951	1465	-3120
BraTS-GLI-00459-000	0.847358	0.953036	0.87816	0.892852	4269	343	-325
BraTS-GLI-00464-000	0.653772	0.759953	0.674766	0.696164	-4998	-3982	-4343
BraTS-GLI-00464-001	0.856197	0.915408	0.883647	0.885084	-1483	-455	-1152
BraTS-GLI-00466-000	0.887062	0.957687	0.878983	0.907911	356	-2834	-4004
BraTS-GLI-00468-000	0.756842	0.733135	0.725738	0.738571	-28540	-72175	-18288
BraTS-GLI-00469-000	0.912787	0.915193	0.809195	0.879058	-500	-11559	699
BraTS-GLI-00469-001	0.872147	0.965003	0.85717	0.898107	3049	2062	1514
BraTS-GLI-00470-000	0.965111	0.949655	0.878167	0.930978	-656	-7523	1349
BraTS-GLI-00472-000	0.735765	0.878481	0.742608	0.785618	-7913	-16503	-4152
BraTS-GLI-00477-000	0.957513	0.916703	0.933719	0.935978	-3583	-24804	-3758
BraTS-GLI-00477-001	0.967879	0.93974	0.945296	0.950972	-1459	-4680	-5130
BraTS-GLI-00478-000	0.954744	0.960777	0.93245	0.949324	1247	-1594	-1709
BraTS-GLI-00478-001	0.593555	0.941089	0.960452	0.831699	17952	474	590
BraTS-GLI-00479-000	0.954006	0.948645	0.935989	0.946213	-2722	-2378	28
BraTS-GLI-00479-001	0.884233	0.972527	0.966434	0.941064	3072	3495	667
BraTS-GLI-00480-000	0.857872	0.951242	0.888825	0.899313	-16780	-3934	-6617
BraTS-GLI-00480-001	0.982751	0.970674	0.970868	0.974764	72	574	-1694
BraTS-GLI-00481-000	0.955007	0.945921	0.900552	0.933827	2244	5694	1176

BraTS-GLI-00481-001	0.922129	0.948258	0.965269	0.945219	8994	-3108	-1388
BraTS-GLI-00483-000	0.947045	0.878967	0.908588	0.911533	33	-8266	1720
BraTS-GLI-00483-001	0.938826	0.971591	0.952449	0.954289	3450	826	200
BraTS-GLI-00485-000	0.563758	0.764671	0.497431	0.60862	-683	-10237	-801
BraTS-GLI-00485-001	0.658201	0.897942	0.78592	0.780688	2447	-1762	1216
BraTS-GLI-00488-000	0.864483	0.908488	0.842217	0.871729	-7189	-7267	-3510
BraTS-GLI-00491-000	0.909853	0.952733	0.857302	0.906629	-5680	-1814	-1620
BraTS-GLI-00491-001	0.971547	0.969191	0.928428	0.956389	1522	1887	2282
BraTS-GLI-00493-000	0.572677	0.113477	0.611579	0.432577	-3727	35084	-2681
BraTS-GLI-00494-000	0.928855	0.698597	0.903465	0.843639	-5608	-80680	-1846
BraTS-GLI-00494-001	0.852429	0.949182	0.760585	0.854065	12069	4803	9953
BraTS-GLI-00495-000	0.962725	0.966829	0.922742	0.950765	-1647	64	275
BraTS-GLI-00495-001	0.982172	0.976349	0.964518	0.974346	698	3403	124
BraTS-GLI-00496-000	0.832198	0.606256	0.832461	0.756972	217	-24401	196
BraTS-GLI-00498-000	0.638939	0.822938	0.612186	0.691354	-3622	-22930	-1092
BraTS-GLI-00498-001	0.883624	0.957431	0.840696	0.893917	2775	-1331	1874
BraTS-GLI-00499-000	0.979193	0.963242	0.940447	0.960961	-1263	2032	87
BraTS-GLI-00499-001	0.95963	0.972403	0.92765	0.953228	-1151	4677	-3144
BraTS-GLI-00500-000	0.892805	0.91749	0.84721	0.885835	-1431	1626	-1808
BraTS-GLI-00500-001	0.874911	0.965104	0.932745	0.924253	1688	3485	589
BraTS-GLI-00501-000	0.878187	0.500425	0.833694	0.737435	-1491	-123349	-1831
BraTS-GLI-00502-000	0.843198	0.655071	0.854848	0.784372	-11065	-48856	-5017
BraTS-GLI-00502-001	0.957402	0.845837	0.924365	0.909201	1772	-14087	2843
BraTS-GLI-00504-000	0.725278	0.960012	0.833798	0.839696	5050	2923	-269
BraTS-GLI-00505-000	0.84464	0.914681	0.875785	0.878368	-16362	-3792	-5642
BraTS-GLI-00506-000	0.974	0.942562	0.931635	0.949399	-1263	8867	1586
BraTS-GLI-00507-000	0.961442	0.959292	0.920928	0.947221	-1088	-1780	-2047
BraTS-GLI-00510-000	0.86208	0.897003	0.840465	0.866516	-14038	-15566	-6330
BraTS-GLI-00510-001	0.924604	0.895955	0.901315	0.907291	-232	9303	594

BraTS-GLI-00511-000	0.970506	0.955885	0.909689	0.94536	566	1450	2040
BraTS-GLI-00511-001	0.957278	0.966484	0.950769	0.958176	1772	4842	-1837
BraTS-GLI-00512-000	0.901873	0.906462	0.812504	0.873613	-7596	-12363	335
BraTS-GLI-00513-000	0.967643	0.815262	0.925439	0.902781	-4505	-47210	1849
BraTS-GLI-00514-000	0.977399	0.736244	0.973108	0.895583	98	-3845	8
BraTS-GLI-00514-001	0.95366	0.897203	0.943128	0.931331	531	-16937	275
BraTS-GLI-00516-000	0.937953	0.950986	0.868507	0.919148	-10697	-10155	1608
BraTS-GLI-00517-000	0.862192	0.899649	0.905024	0.888955	4209	9118	2235
BraTS-GLI-00517-001	0.8973	0.968896	0.966804	0.944333	3726	1515	312
BraTS-GLI-00518-000	0.893269	0.938706	0.769042	0.867006	2761	3206	3901
BraTS-GLI-00518-001	0.842832	0.964657	0.937156	0.914881	4716	4950	-511
BraTS-GLI-00519-000	0.944359	0.915095	0.919806	0.92642	972	-6947	1110
BraTS-GLI-00520-000	0.95344	0.893885	0.903268	0.916864	874	9086	1119
BraTS-GLI-00520-001	0.747735	0.966143	0.959763	0.891214	5868	2486	-535
BraTS-GLI-00523-000	0.698864	0.862935	0.629939	0.730579	38	-11919	-2
BraTS-GLI-00524-000	0.846925	0.923479	0.782306	0.850903	-12232	-6980	-3720
BraTS-GLI-00525-000	0.023361	0.939652	0.145559	0.369524	7350	-5221	-717
BraTS-GLI-00525-001	0.623456	0.971589	0.097774	0.564273	-31189	-2127	-10545
BraTS-GLI-00526-000	0.971761	0.917189	0.876167	0.921706	-595	-10046	1751
BraTS-GLI-00528-000	0.715543	0.763672	0.765103	0.748106	-10985	-38178	-12966
BraTS-GLI-00529-000	0.590312	0.830441	0.691577	0.70411	-12100	-14518	-4320
BraTS-GLI-00530-000	0.258889	0.924594	0.204524	0.462669	-4227	182	-2311
BraTS-GLI-00532-000	0.788556	0.968373	0.731377	0.829435	2505	-1151	-5005
BraTS-GLI-00533-000	0.937195	0.911732	0.891333	0.91342	-1077	10938	3587
BraTS-GLI-00537-000	0.899604	0.706515	0.862762	0.82296	-2412	-15648	-1906
BraTS-GLI-00538-000	0.631086	0.790602	0.729891	0.717193	4955	-59736	648
BraTS-GLI-00539-000	0.681702	0.894919	0.75727	0.777964	-30119	-24059	-10553
BraTS-GLI-00540-000	0	0.016544	0	0.005515	-156	-11553	-154
BraTS-GLI-00540-001	0.101863	0.713782	0.10475	0.306799	-551	-2225	-535

BraTS-GLI-00542-000	0.728647	0.81343	0.793003	0.77836	-22139	-25546	-14562
BraTS-GLI-00543-000	0.420929	0.943689	0.558865	0.641161	965	-4539	-5418
BraTS-GLI-00544-000	0.954582	0.906686	0.896066	0.919111	-1233	2762	-914
BraTS-GLI-00544-001	0.980329	0.962646	0.883547	0.942174	-522	-1906	-2870
BraTS-GLI-00545-000	0.856879	0.935316	0.833619	0.875271	-4010	-1885	-1377
BraTS-GLI-00545-001	0.800621	0.953258	0.861115	0.871665	6961	5434	3985
BraTS-GLI-00547-000	0.963358	0.91541	0.882018	0.920262	102	-1283	830
BraTS-GLI-00547-001	0.860955	0.935948	0.871248	0.889384	3424	3536	2516
BraTS-GLI-00548-000	0.92813	0.907255	0.91752	0.917635	-3162	-23465	-403
BraTS-GLI-00548-001	0.964619	0.974209	0.966068	0.968299	1231	-2145	151
BraTS-GLI-00549-000	0.948795	0.964244	0.819678	0.910906	1110	-2086	2723
BraTS-GLI-00549-001	0.666015	0.966459	0.928724	0.853733	12302	-5082	1666
BraTS-GLI-00550-000	0.847342	0.664979	0.792047	0.768123	-1098	-13880	580
BraTS-GLI-00550-001	0.618958	0.869807	0.87545	0.788072	-8293	-8630	758
BraTS-GLI-00551-000	0.935728	0.936308	0.881291	0.917776	1994	-8082	-378
BraTS-GLI-00552-000	0.92989	0.86402	0.908356	0.900756	-897	-29753	-2030
BraTS-GLI-00552-001	0.956962	0.96234	0.968296	0.962533	1514	-4164	-206
BraTS-GLI-00554-000	0.982294	0.928186	0.929894	0.946791	703	-2900	1041
BraTS-GLI-00555-000	0.746558	0.636581	0.709486	0.697542	-7522	-18952	-6998
BraTS-GLI-00555-001	0.928096	0.884349	0.904749	0.905731	1949	6431	982
BraTS-GLI-00556-000	0.832406	0.920556	0.7633	0.838754	-8819	-13436	-6847
BraTS-GLI-00556-001	0.928904	0.950732	0.93213	0.937256	1474	4311	-982
BraTS-GLI-00557-000	0.925532	0.973109	0.909472	0.936038	268	764	-516
BraTS-GLI-00558-000	0.907609	0.94679	0.907974	0.920791	-9994	-783	6701
BraTS-GLI-00558-001	0.929368	0.968836	0.953657	0.950621	7998	131	-4535
BraTS-GLI-00559-000	0.871404	0.811491	0.888557	0.857151	-16477	-37769	-6685
BraTS-GLI-00559-001	0.977856	0.706955	0.94626	0.877024	-2940	-89723	-6398
BraTS-GLI-00561-000	0.969048	0.920981	0.96212	0.950716	87	-7130	-667
BraTS-GLI-00563-000	0.93755	0.958207	0.945697	0.947151	1228	3536	-1240

BraTS-GLI-00565-000	0.970052	0.937417	0.957924	0.955131	262	7637	167
BraTS-GLI-00567-000	0.911079	0.837536	0.908884	0.885833	-3236	3369	181
BraTS-GLI-00568-000	0.93681	0.950249	0.924924	0.937328	5885	-1977	2775
BraTS-GLI-00569-000	0.911133	0.85212	0.949017	0.90409	-6874	32565	-24
BraTS-GLI-00570-000	0.840095	0.854068	0.801952	0.832038	-3144	8629	-882
BraTS-GLI-00571-000	0.937313	0.931135	0.921052	0.929833	1057	4106	317
BraTS-GLI-00572-000	0.860175	0.893425	0.938169	0.897256	4065	-3042	324
BraTS-GLI-00574-000	0.930428	0.947401	0.959611	0.945813	2381	2964	-160
BraTS-GLI-00575-000	0.942426	0.947923	0.950158	0.946836	-4869	3525	-1432
BraTS-GLI-00576-000	0.533991	0.902977	0.689306	0.708758	-12100	-1301	-3959
BraTS-GLI-00577-000	0.952254	0.95875	0.914081	0.941695	-2433	4528	-1878
BraTS-GLI-00578-000	0.934197	0.929842	0.887768	0.917269	-9462	20328	-5710
BraTS-GLI-00579-000	0.968909	0.95723	0.962569	0.962902	440	7585	-364
BraTS-GLI-00580-000	0.939552	0.934284	0.911543	0.928459	-1809	3138	1750
BraTS-GLI-00581-000	0.943249	0.931257	0.943605	0.93937	-2414	1624	-3375
BraTS-GLI-00582-000	0.716773	0.935975	0.889132	0.847293	6837	2817	1012
BraTS-GLI-00583-000	0.917772	0.955596	0.915397	0.929588	-4462	4022	-952
BraTS-GLI-00584-000	0.938463	0.86373	0.939912	0.914035	-249	2124	-111
BraTS-GLI-00586-000	0.948228	0.967722	0.919054	0.945001	-5491	3827	-1720
BraTS-GLI-00587-000	0.709004	0.892153	0.849446	0.816868	-49238	-5728	-11476
BraTS-GLI-00588-000	0.845396	0.832241	0.888139	0.855259	-629	-13183	-915
BraTS-GLI-00589-000	0.915486	0.824232	0.883366	0.874361	2331	40596	-15
BraTS-GLI-00590-000	0.903734	0.952539	0.943065	0.933113	-1078	-895	-647
BraTS-GLI-00591-000	0.940941	0.91253	0.936772	0.930081	-1872	-3135	-861
BraTS-GLI-00593-000	0.908727	0.964993	0.90036	0.924693	-8248	4053	-4766
BraTS-GLI-00594-000	0.841201	0.716734	0.859439	0.805791	-1079	48085	4097
BraTS-GLI-00596-000	0.872843	0.954451	0.90651	0.911268	-13743	3188	-404
BraTS-GLI-00597-000	0	0.942458	0	0.314153	-2663	2054	19
BraTS-GLI-00598-000	0.910329	0.932474	0.877515	0.906773	-5308	3589	-2756

BraTS-GLI-00599-000	0.952493	0.958696	0.954042	0.955077	1329	5103	238
BraTS-GLI-00601-000	0.912583	0.896088	0.815266	0.874645	-4321	-10545	-8967
BraTS-GLI-00602-000	0.944525	0.928641	0.966116	0.946427	958	1670	190
BraTS-GLI-00602-001	0.946625	0.893329	0.951518	0.930491	281	-1386	295
BraTS-GLI-00604-000	0.878391	0.948519	0.643789	0.823566	-18349	7842	-32061
BraTS-GLI-00605-000	0.895194	0.954948	0.903536	0.917893	-4395	-17393	-8202
BraTS-GLI-00606-000	0.885163	0.925424	0.885225	0.898604	-28348	-5733	-14534
BraTS-GLI-00607-000	0.828707	0.684303	0.879947	0.797652	946	-32033	322
BraTS-GLI-00607-001	0.77327	0.785109	0.924159	0.827512	2246	-27609	-199
BraTS-GLI-00608-000	0.716773	0.956593	0.761704	0.81169	-37434	-1068	-14301
BraTS-GLI-00608-001	0.966722	0.967803	0.960398	0.964974	-3786	2925	-2958
BraTS-GLI-00610-000	0.85325	0.920045	0.880029	0.884441	1882	-901	-717
BraTS-GLI-00610-001	0.256377	0.637641	0.209904	0.367974	21057	36265	21146
BraTS-GLI-00611-000	0.865283	0.588556	0.901426	0.785088	-4852	-44128	-660
BraTS-GLI-00611-001	0.945027	0.677569	0.89272	0.838439	1298	-41065	2500
BraTS-GLI-00612-000	0.734559	0.945937	0.77085	0.817116	5341	-13054	-1824
BraTS-GLI-00612-001	0.312577	0.904959	0.597984	0.605173	7131	-2840	2067
BraTS-GLI-00613-000	0.594621	0.450096	0.563533	0.536083	-14793	-48639	-11434
BraTS-GLI-00613-001	0.162852	0.641387	0.157761	0.320667	5223	-11302	5481
BraTS-GLI-00615-000	0.839852	0.709892	0.833592	0.794445	-1318	-10818	-652
BraTS-GLI-00616-000	0.962738	0.959621	0.940726	0.954361	2234	6665	-230
BraTS-GLI-00618-000	0.937703	0.949351	0.935894	0.940983	-3573	8427	1800
BraTS-GLI-00619-000	0.928705	0.917564	0.907588	0.917952	-5680	-12795	-5896
BraTS-GLI-00619-001	0.797863	0.803323	0.790833	0.79734	-1813	-4416	-1839
BraTS-GLI-00620-000	0.967865	0.968901	0.969117	0.968628	306	57	-125
BraTS-GLI-00620-001	0.861714	0.9665	0.961645	0.929953	8225	-4617	-569
BraTS-GLI-00621-000	0.292991	0.556649	0.297521	0.382387	-2519	-9883	-2865
BraTS-GLI-00622-000	0.77985	0.902669	0.812833	0.831784	4025	6912	3043
BraTS-GLI-00623-000	0.843253	0.907113	0.946881	0.899082	9164	-15691	684

BraTS-GLI-00624-000	0.792917	0.876825	0.816296	0.82868	-10363	-10818	-2273
BraTS-GLI-00625-000	0.956789	0.940295	0.924201	0.940428	-5002	-12739	-255
BraTS-GLI-00626-000	0.958774	0.923158	0.950723	0.944218	324	1550	-39
BraTS-GLI-00628-000	0.844624	0.937045	0.868635	0.883435	-197	312	-1306
BraTS-GLI-00630-000	0.717418	0.843605	0.722132	0.761052	5156	1364	4087
BraTS-GLI-00630-001	0.854855	0.936799	0.918033	0.903229	9209	-2304	622
BraTS-GLI-00631-000	0.978932	0.962693	0.923001	0.954875	-1048	592	1473
BraTS-GLI-00636-000	0.96626	0.943321	0.96005	0.956543	190	1502	-146
BraTS-GLI-00636-001	0.855402	0.784846	0.839108	0.826452	1759	4478	1995
BraTS-GLI-00638-000	0.833059	0.599933	0.827635	0.753542	-1510	-76356	-1344
BraTS-GLI-00639-000	0.940811	0.967984	0.941087	0.949961	1028	-544	-96
BraTS-GLI-00640-000	0.779483	0.867187	0.764009	0.803559	-30982	-18653	-13691
BraTS-GLI-00641-000	0.97738	0.892586	0.9378	0.935922	-1338	7009	-44
BraTS-GLI-00641-001	0.959784	0.913635	0.927475	0.933632	-1079	5198	1160
BraTS-GLI-00642-000	0.910942	0.780756	0.898924	0.86354	-1016	-5237	-563
BraTS-GLI-00645-000	0.90542	0.837727	0.91579	0.886312	-1607	-2324	-797
BraTS-GLI-00645-001	0.960545	0.871933	0.952811	0.92843	-241	-2959	483
BraTS-GLI-00646-000	0.963	0.940766	0.962018	0.955261	1324	1864	-3071
BraTS-GLI-00646-001	0.961695	0.950884	0.976604	0.963061	3581	3809	-1475
BraTS-GLI-00649-000	0.945543	0.915684	0.922365	0.927864	-8719	-19496	-1785
BraTS-GLI-00649-001	0.965753	0.962253	0.95853	0.962178	-1761	3105	199
BraTS-GLI-00650-000	0.879373	0.955743	0.963838	0.932985	4354	-1846	368
BraTS-GLI-00651-000	0.953994	0.939789	0.917701	0.937161	541	1919	1148
BraTS-GLI-00652-000	0.9341	0.956908	0.862699	0.917903	-1597	1489	-1749
BraTS-GLI-00654-000	0.816132	0.727852	0.837586	0.793857	-13969	-38521	-4973
BraTS-GLI-00654-001	0.85835	0.888102	0.894779	0.880411	-9127	-17449	-2586
BraTS-GLI-00655-000	0.923718	0.94789	0.902569	0.924726	-469	-5726	-3074
BraTS-GLI-00655-001	0.933637	0.952233	0.964102	0.94999	4387	2103	-209
BraTS-GLI-00656-000	0.293572	0.935564	0.671283	0.633473	14924	-6147	-1069

BraTS-GLI-00657-000	0.948816	0.960893	0.909604	0.939771	-8261	6686	-8459
BraTS-GLI-00658-000	0.915388	0.942951	0.857755	0.905365	-5564	631	-3922
BraTS-GLI-00659-000	0.831418	0.893301	0.801956	0.842225	2111	4088	1804
BraTS-GLI-00661-000	0.936342	0.939669	0.941849	0.939287	3877	11140	-685
BraTS-GLI-00663-000	0.947426	0.975524	0.965851	0.962934	3035	700	-204
BraTS-GLI-00667-000	0.874673	0.966563	0.925122	0.922119	2930	4527	619
BraTS-GLI-00668-000	0.766404	0.772426	0.758808	0.765879	63	-64688	51
BraTS-GLI-00674-000	0.874489	0.946723	0.869071	0.896761	-4380	-3584	-3095
BraTS-GLI-00674-001	0.953999	0.956136	0.951443	0.953859	2299	-2870	1022
BraTS-GLI-00675-000	0.916599	0.923152	0.899481	0.913078	174	964	174
BraTS-GLI-00675-001		0.031114			0	-5693	0
BraTS-GLI-00676-000	0.829417	0.796762	0.824049	0.816743	-14213	-2413	-5806
BraTS-GLI-00676-001	0.901927	0.854444	0.90592	0.88743	-4767	10381	-2457
BraTS-GLI-00677-000	0.826872	0.903485	0.828455	0.852937	-5673	-27016	-4619
BraTS-GLI-00679-000	0.806118	0.853884	0.756254	0.805419	-12430	-42215	-11194
BraTS-GLI-00680-000	0.96183	0.902175	0.935341	0.933116	-920	-7269	-1378
BraTS-GLI-00680-001	0.87201	0.971123	0.937635	0.926923	2510	28	997
BraTS-GLI-00682-000	0.872678	0.897258	0.861369	0.877102	-3988	-3635	-2416
BraTS-GLI-00682-001	0.971085	0.966122	0.920493	0.952567	-30	5213	-3276
BraTS-GLI-00683-000	0.844802	0.943571	0.809663	0.866012	-21724	-5093	-7630
BraTS-GLI-00683-001	0.978905	0.959692	0.966655	0.968417	-1462	10906	-1512
BraTS-GLI-00684-000	0	0.011515	0	0.003838	-2925	-23037	-2925
BraTS-GLI-00685-000	0.855566	0.80916	0.83171	0.832145	-9550	-24301	-2838
BraTS-GLI-00685-001	0.9738	0.805466	0.952234	0.9105	333	-30877	1205
BraTS-GLI-00686-000	0.853567	0.89601	0.820685	0.856754	-17852	-38458	-6790
BraTS-GLI-00687-000	0.752727	0.87869	0.788979	0.806799	-10975	-17070	-5936
BraTS-GLI-00688-000	0.947681	0.917758	0.878016	0.914485	-663	-2106	-563
BraTS-GLI-00689-000	0.980305	0.959034	0.924274	0.954537	-3031	-3925	747
BraTS-GLI-00689-001	0.941012	0.957784	0.918165	0.938987	-10425	-4359	-3643

BraTS-GLI-00690-000	0.504326	0.808784	0.583697	0.632269	-7984	-11546	-5016
BraTS-GLI-00691-000	0.891336	0.894363	0.928931	0.904877	-6894	-7471	332
BraTS-GLI-00691-001	0.64929	0.913573	0.743165	0.768676	-14330	-3179	-2375
BraTS-GLI-00692-000	0.843862	0.899664	0.81803	0.853852	-5543	-11372	-4193
BraTS-GLI-00692-001	0.949896	0.9699	0.949682	0.956493	1012	-473	689
BraTS-GLI-00693-000	0.754815	0.837725	0.787785	0.793442	-5554	-33273	-4421
BraTS-GLI-00694-000	0.859568	0.496384	0.897364	0.751106	388	-45385	101
BraTS-GLI-00694-001	0.866794	0.513294	0.844476	0.741521	65	-70570	19
BraTS-GLI-00697-000	0.435018	0.57634	0.455744	0.489034	-52580	-78779	-32005
BraTS-GLI-00698-000	0.971705	0.935239	0.943211	0.950052	-2106	1959	-2646
BraTS-GLI-00703-000	0.820896	0.830843	0.809139	0.820293	-14862	-21439	-7142
BraTS-GLI-00703-001	0.970746	0.950772	0.948226	0.956581	2227	-11356	2059
BraTS-GLI-00704-000	0.884788	0.957391	0.895725	0.912634	-11548	4204	-497
BraTS-GLI-00705-000	0.873191	0.572621	0.875901	0.773905	-687	-9493	-451
BraTS-GLI-00706-000	0.877044	0.909324	0.821548	0.869305	-17398	-6797	-13377
BraTS-GLI-00707-000	0.964942	0.907293	0.935888	0.936041	848	-15680	374
BraTS-GLI-00708-000	0.873077	0.870844	0.786508	0.843477	2161	4058	3613
BraTS-GLI-00708-001	0.91443	0.961727	0.96492	0.947026	1260	2626	-124
BraTS-GLI-00709-000	0.902004	0.931443	0.906092	0.91318	2882	-1697	907
BraTS-GLI-00714-000	0.877031	0.949917	0.864416	0.897121	-22080	-15678	-9093
BraTS-GLI-00714-001	0.950178	0.969021	0.948455	0.955885	-7268	5133	-3510
BraTS-GLI-00715-000	0.977742	0.956572	0.952891	0.962402	-652	-330	-2594
BraTS-GLI-00715-001	0.922975	0.923928	0.951864	0.932922	2713	672	1138
BraTS-GLI-00716-000	0.938573	0.945847	0.920068	0.934829	-2053	1788	-132
BraTS-GLI-00718-000	0.947107	0.936173	0.906003	0.929761	1800	8909	-514
BraTS-GLI-00723-000	0.91059	0.948787	0.869732	0.909703	-4323	1895	-2768
BraTS-GLI-00724-000	0.968273	0.947686	0.913039	0.942999	-308	9999	254
BraTS-GLI-00725-000	0.678312	0.516388	0.678487	0.624396	-1264	-60430	-1340
BraTS-GLI-00725-001	0.67398	0.63115	0.676602	0.660578	-1091	-68200	-1039

BraTS-GLI-00727-000	0.869118	0.955454	0.761788	0.86212	-9023	-4510	-5150
BraTS-GLI-00727-001	0.887874	0.952833	0.799346	0.880018	-559	-10048	2800
BraTS-GLI-00728-000	0.93311	0.932204	0.920991	0.928768	96	-2368	50
BraTS-GLI-00729-000	0.937114	0.813154	0.915011	0.888427	862	-1089	844
BraTS-GLI-00729-001	0.887019	0.763781	0.86775	0.839517	1034	-1091	900
BraTS-GLI-00730-000	0.879619	0.953088	0.858776	0.897161	-4090	-302	-3295
BraTS-GLI-00730-001	0.962191	0.97218	0.949197	0.961189	433	-730	693
BraTS-GLI-00731-000	0.814362	0.714384	0.794296	0.774347	-1087	9106	-1197
BraTS-GLI-00731-001	0.047016	0.660839	0.070796	0.25955	-519	-11293	-507
BraTS-GLI-00732-000	0.860389	0.881646	0.841659	0.861231	-457	641	-560
BraTS-GLI-00732-001	0.876952	0.669591	0.876596	0.807713	-184	2572	-180
BraTS-GLI-00733-000	0.938379	0.943073	0.912324	0.931259	-1960	5574	-15
BraTS-GLI-00733-001	0.649226	0.481757	0.614654	0.581879	4255	13997	4811
BraTS-GLI-00734-000	0.88507	0.944548	0.845891	0.891836	-7521	3813	-4926
BraTS-GLI-00734-001	0.970435	0.966514	0.952539	0.963163	-1676	-1024	-4867
BraTS-GLI-00735-000	0.898787	0.802306	0.906765	0.869286	-371	-5583	-347
BraTS-GLI-00735-001		0.771105			4859	-1059	4787
BraTS-GLI-00736-000	0.748697	0.580141	0.746543	0.691794	-6360	-23507	-3554
BraTS-GLI-00736-001	0.924535	0.888404	0.891512	0.901484	1965	-6551	1334
BraTS-GLI-00737-000	0.896193	0.91454	0.85858	0.889771	-2818	-875	-2467
BraTS-GLI-00739-000	0.902372	0.930485	0.905368	0.912742	1288	-24476	-5582
BraTS-GLI-00740-000	0.785731	0.952361	0.792198	0.84343	-257	1469	-4166
BraTS-GLI-00742-000	0.806761	0.839976	0.738786	0.795174	-8975	-48669	-6139
BraTS-GLI-00744-000	0.953842	0.961964	0.935102	0.950302	1046	-3371	522
BraTS-GLI-00746-000	0.721192	0.861288	0.623542	0.735341	-5153	-3823	-4883
BraTS-GLI-00747-000	0.951783	0.946654	0.922602	0.940346	-3508	4654	-1856
BraTS-GLI-00750-000	0.976443	0.949056	0.932406	0.952635	610	-2975	755
BraTS-GLI-00750-001	0.786361	0.944857	0.868679	0.866632	14382	-4871	6127
BraTS-GLI-00751-000	0.981568	0.923096	0.942977	0.949214	200	-5084	1052

BraTS-GLI-00751-001	0.754082	0.85296	0.765432	0.790825	1890	-8806	1729
BraTS-GLI-00753-000	0.527679	0.578872	0.540623	0.549058	-658	-1823	-645
BraTS-GLI-00753-001	0.299411	0.2911	0.323293	0.304601	-733	-1796	-604
BraTS-GLI-00756-000	0.885227	0.925782	0.870842	0.89395	-14885	-12796	-7192
BraTS-GLI-00756-001	0.924803	0.964509	0.9067	0.932004	-9183	-4424	-8211
BraTS-GLI-00757-000	0.685954	0.66051	0.635902	0.660789	-18150	-32200	-12347
BraTS-GLI-00758-000	0.959682	0.939952	0.954612	0.951415	667	1443	296
BraTS-GLI-00759-000	0.930987	0.892303	0.926075	0.916455	-10611	20453	-5073
BraTS-GLI-00760-000	0.895336	0.939634	0.902331	0.912434	-14082	-10954	-5465
BraTS-GLI-00764-000	0.856834	0.787026	0.843139	0.828999	-9528	-27134	-9289
BraTS-GLI-00765-000	0.86324	0.881936	0.809483	0.851553	-1607	-18070	-4820
BraTS-GLI-00767-000	0.961744	0.946431	0.925915	0.944697	-1151	-4338	-576
BraTS-GLI-00767-001	0.824578	0.966942	0.715149	0.835556	-10193	-4761	-10401
BraTS-GLI-00768-000	0.175239	0.960862	0.716393	0.617498	39973	4912	2544
BraTS-GLI-00768-001	0.216085	0.930717	0.749341	0.632048	36158	1131	2829
BraTS-GLI-00772-000	0.95178	0.883905	0.895602	0.910429	-753	1067	-530
BraTS-GLI-00772-001	0.907716	0.927441	0.890368	0.908508	-490	3941	-877
BraTS-GLI-00773-000	0.841798	0.886056	0.871208	0.866354	-17632	-12937	-9905
BraTS-GLI-00774-000	0.947018	0.939984	0.943979	0.943661	892	-16867	-387
BraTS-GLI-00774-001	0.956656	0.967106	0.921597	0.948453	1314	-5951	806
BraTS-GLI-00775-000	0.922158	0.941126	0.89929	0.920858	2300	-1568	2204
BraTS-GLI-00775-001	0.921713	0.925403	0.910268	0.919128	1443	-2615	1550
BraTS-GLI-00777-000	0.717531	0.955922	0.862823	0.845426	-19466	-85	-989
BraTS-GLI-00777-001	0.966718	0.968873	0.904351	0.946647	-552	3549	539
BraTS-GLI-00778-000	0.968468	0.902645	0.936531	0.935881	-441	-5933	-41
BraTS-GLI-00780-000	0.969467	0.96961	0.963115	0.967397	2170	4288	859
BraTS-GLI-00781-000	0.970174	0.925103	0.965213	0.953497	2402	14373	1484
BraTS-GLI-00782-000	0.966614	0.953552	0.953737	0.957968	1757	1350	1945
BraTS-GLI-00784-000	0.877098	0.943465	0.949728	0.92343	10518	-14073	-31

BraTS-GLI-00787-000	0.971965	0.963365	0.896382	0.943904	1334	-4901	-9622
BraTS-GLI-00788-000	0.840173	0.808234	0.874111	0.84084	1185	-2076	837
BraTS-GLI-00789-000	0.956402	0.978082	0.929248	0.954577	-1944	2154	-3374
BraTS-GLI-00791-000	0.916697	0.948823	0.947462	0.937661	5458	-4163	-3895
BraTS-GLI-00792-000	0.762619	0.938255	0.896939	0.865938	20844	-23090	5290
BraTS-GLI-00793-000	0.896581	0.903568	0.839316	0.879822	3348	-25069	-8368
BraTS-GLI-00795-000	0.880623	0.919483	0.861608	0.887238	-2634	-653	-2595
BraTS-GLI-00796-000	0.933842	0.968378	0.947701	0.949974	4041	2900	1924
BraTS-GLI-00797-000	0.926627	0.957738	0.896238	0.926868	-5967	1912	-3742
BraTS-GLI-00799-000	0.878438	0.921975	0.889432	0.896615	-19248	-21244	-13089
BraTS-GLI-00800-000	0.923961	0.950642	0.898356	0.924319	2761	-1276	2233
BraTS-GLI-00801-000	0.94942	0.94134	0.93519	0.941983	1363	11693	742
BraTS-GLI-00802-000	0.786849	0.772532	0.826944	0.795441	2672	7184	2057
BraTS-GLI-00803-000	0.959262	0.942764	0.926077	0.942701	1694	-11628	2512
BraTS-GLI-00804-000	0.980412	0.969155	0.970162	0.973243	-581	5123	-1197
BraTS-GLI-00805-000	0.602112	0.928503	0.553028	0.694548	2127	-4400	1516
BraTS-GLI-00806-000	0.777371	0.910409	0.882559	0.85678	-43664	-10633	-16068
BraTS-GLI-00807-000	0.911223	0.936382	0.89653	0.914712	-10676	2198	-7038
BraTS-GLI-00808-000	0.947563	0.954591	0.918889	0.940348	3040	11279	-321
BraTS-GLI-00809-000	0.9452	0.971008	0.94293	0.953046	1917	-3915	-2224
BraTS-GLI-00810-000	0.884432	0.964367	0.907527	0.918775	1577	-21	-740
BraTS-GLI-00811-000	0.92385	0.973857	0.886467	0.928058	-1282	-1716	30
BraTS-GLI-00814-000	0.924843	0.970718	0.911387	0.935649	2670	3535	-916
BraTS-GLI-00816-000	0.967374	0.970282	0.950117	0.962591	-219	6089	-3337
BraTS-GLI-00819-000	0.949208	0.966379	0.942883	0.952823	-3771	-7422	-1701
BraTS-GLI-00820-000	0.934531	0.894271	0.917218	0.91534	-11	1094	862
BraTS-GLI-00823-000	0.854871	0.961729	0.760724	0.859108	-7773	-5286	-5381
BraTS-GLI-00824-000	0.937199	0.962198	0.934387	0.944595	-1657	315	-2898
BraTS-GLI-00828-000	0.878893	0.946955	0.84344	0.889763	2771	-7477	8534

BraTS-GLI-00830-000	0.947577	0.95828	0.965949	0.957268	1622	-382	-200
BraTS-GLI-00831-000	0.920062	0.842798	0.916955	0.893271	143	6277	252
BraTS-GLI-00834-000	0.912509	0.849154	0.941418	0.901027	543	-1882	305
BraTS-GLI-00836-000	0.95021	0.927292	0.913576	0.930359	196	1723	-923
BraTS-GLI-00837-000	0.853516	0.780231	0.808363	0.814037	1347	565	1280
BraTS-GLI-00838-000	0.803757	0.811293	0.821624	0.812224	4822	-30086	1912
BraTS-GLI-00839-000	0.664895	0.844786	0.666057	0.725246	1103	3502	1057
BraTS-GLI-00840-000	0.463045	0.889212	0.816718	0.722992	-8568	-20674	81
BraTS-GLI-00999-000	0.68029	0.870733	0.73591	0.762311	-26120	-30685	4171
BraTS-GLI-01000-000	0.965918	0.948472	0.90563	0.940007	2094	7580	2437
BraTS-GLI-01001-000	0.961776	0.959096	0.87704	0.932637	1523	2740	2583
BraTS-GLI-01002-000	0.900291	0.911783	0.851898	0.887991	6264	8386	2358
BraTS-GLI-01003-000	0.709553	0.897801	0.804871	0.804075	-32145	-19741	-2676
BraTS-GLI-01004-000	0.862244	0.714313	0.938839	0.838465	4757	57698	699
BraTS-GLI-01005-000	0.931338	0.938547	0.938545	0.936144	-2679	-9865	-1823
BraTS-GLI-01007-000	0.878336	0.82407	0.823573	0.841993	8258	22298	9334
BraTS-GLI-01008-000	0.901802	0.898029	0.932318	0.910717	1386	-1169	715
BraTS-GLI-01009-000	0.932028	0.950681	0.869569	0.917426	850	-2868	1870
BraTS-GLI-01010-000	0.947089	0.926944	0.91356	0.929197	1921	4528	2313
BraTS-GLI-01011-000	0.950021	0.927626	0.95606	0.944569	852	8545	-311
BraTS-GLI-01012-000	0.933301	0.903905	0.924598	0.920601	-6171	1444	-8002
BraTS-GLI-01013-000	0.863182	0.964651	0.91264	0.913491	3264	3620	813
BraTS-GLI-01014-000	0.955542	0.930572	0.921211	0.935775	1032	4991	2179
BraTS-GLI-01015-000	0.9087	0.96456	0.935838	0.936366	5285	-5233	-709
BraTS-GLI-01016-000	0.973158	0.930394	0.945419	0.949657	-411	10030	-2803
BraTS-GLI-01017-000	0.947922	0.95554	0.929207	0.944223	1622	5218	340
BraTS-GLI-01018-000	0.979261	0.959467	0.962618	0.967115	-1502	9906	249
BraTS-GLI-01019-000	0.963245	0.944919	0.961044	0.956403	-1357	-9020	-1977
BraTS-GLI-01020-000	0.941269	0.973215	0.956917	0.957134	3163	2380	-1100

BraTS-GLI-01021-000	0.971709	0.960721	0.967537	0.966656	-1337	12052	-2457
BraTS-GLI-01022-000	0.925128	0.951968	0.938077	0.938391	1255	-4885	-1177
BraTS-GLI-01023-000	0.951627	0.977073	0.928546	0.952415	1084	39	-1067
BraTS-GLI-01024-000	0.911573	0.936823	0.929649	0.926015	987	2243	361
BraTS-GLI-01025-000	0.980205	0.97157	0.921371	0.957715	-184	1549	-2837
BraTS-GLI-01026-000	0.951184	0.961462	0.857837	0.923494	73	6072	1750
BraTS-GLI-01027-000	0.888444	0.794132	0.878071	0.853549	-660	-10741	-1256
BraTS-GLI-01027-001	0.572209	0.784059	0.554607	0.636958	9743	17086	8061
BraTS-GLI-01028-000	0.951803	0.879662	0.917622	0.916362	731	6255	1245
BraTS-GLI-01029-000	0.762204	0.949241	0.96072	0.890722	15059	12317	842
BraTS-GLI-01030-000	0.910079	0.927251	0.915254	0.917528	-6646	-3701	-6257
BraTS-GLI-01031-000	0.584825	0.862489	0.570878	0.672731	4674	-4725	4673
BraTS-GLI-01032-000	0.897231	0.950896	0.793749	0.880625	2320	-10315	-21736
BraTS-GLI-01033-000	0.973471	0.939711	0.965307	0.959496	-467	-846	-1129
BraTS-GLI-01034-000	0.696074	0.775829	0.685033	0.718978	910	1157	305
BraTS-GLI-01035-000	0.072091	0.009798	0.088889	0.056926	-839	-2795	-745
BraTS-GLI-01036-000	0.86211	0.82239	0.847517	0.844006	979	4623	561
BraTS-GLI-01037-000	0.875401	0.896212	0.792561	0.854724	5082	20452	5519
BraTS-GLI-01038-000	0.96923	0.91072	0.951002	0.943651	1428	11515	1549
BraTS-GLI-01039-000	0.918088	0.815371	0.895479	0.876313	971	-311	1482
BraTS-GLI-01040-000	0.920182	0.923135	0.911012	0.91811	-4352	4229	-144
BraTS-GLI-01041-000	0.617545	0.94015	0.708844	0.755513	-27867	5610	4135
BraTS-GLI-01042-000	0.882658	0.938373	0.876367	0.899133	4864	15311	4677
BraTS-GLI-01043-000	0.966111	0.91878	0.907888	0.930926	1464	9017	1955
BraTS-GLI-01044-000	0.84881	0.897538	0.91575	0.887366	12374	18033	4155
BraTS-GLI-01045-000	0.945384	0.923139	0.916899	0.928474	6046	-7355	6438
BraTS-GLI-01046-000	0.957744	0.949423	0.937215	0.948127	3208	8526	3858
BraTS-GLI-01047-000	0.776322	0.893684	0.717178	0.795728	-498	27144	15287
BraTS-GLI-01048-000	0.931067	0.88607	0.864777	0.893972	-274	6965	2712

BraTS-GLI-01049-000	0.828899	0.818461	0.79163	0.812997	2658	-1634	3209
BraTS-GLI-01050-000	0.862455	0.746099	0.843036	0.817197	2209	-4809	2528
BraTS-GLI-01051-000	0.751428	0.912878	0.748415	0.804241	-2214	13906	2366
BraTS-GLI-01052-000	0.975397	0.897691	0.96473	0.945939	1172	-11033	565
BraTS-GLI-01053-000	0.72126	0.903679	0.837479	0.820806	133	19381	6415
BraTS-GLI-01054-000	0.907387	0.671937	0.856669	0.811997	-105	5396	486
BraTS-GLI-01055-000	0.951782	0.946204	0.926525	0.941504	1784	1131	1539
BraTS-GLI-01056-000	0.763098	0.857427	0.756388	0.792304	-6131	-10395	4931
BraTS-GLI-01057-000	0.959719	0.875316	0.900476	0.911837	994	1843	1977
BraTS-GLI-01058-000	0.885589	0.908311	0.903615	0.899172	-7630	13013	581
BraTS-GLI-01059-000	0.888925	0.843102	0.877202	0.869743	3812	46	3811
BraTS-GLI-01060-000	0.871205	0.926276	0.944745	0.914076	-16506	15807	-406
BraTS-GLI-01061-000	0.908372	0.948452	0.920771	0.925865	8258	7647	2937
BraTS-GLI-01062-000	0.899022	0.949664	0.905007	0.917898	-579	8232	4143
BraTS-GLI-01063-000	0.79049	0.931901	0.877707	0.866699	12863	22564	5416
BraTS-GLI-01064-000	0.236085	0.686458	0.233243	0.385262	2735	-4888	2708
BraTS-GLI-01065-000	0.951815	0.950788	0.890914	0.931172	3093	5879	6981
BraTS-GLI-01066-000	0.887867	0.890452	0.847576	0.875298	-4523	-2987	-302
BraTS-GLI-01067-000	0.905041	0.932848	0.923266	0.920385	3350	53	2204
BraTS-GLI-01068-000	0.610811	0.805297	0.593591	0.6699	7192	18176	4152
BraTS-GLI-01069-000	0.97587	0.946722	0.975165	0.965919	-634	14225	-1176
BraTS-GLI-01070-000	0.744487	0.832876	0.79732	0.791561	-18899	-28684	3433
BraTS-GLI-01071-000	0.884509	0.902849	0.827682	0.87168	1774	-1640	604
BraTS-GLI-01072-000	0.930852	0.921441	0.909206	0.9205	-6118	11781	2613
BraTS-GLI-01073-000	0.833367	0.910737	0.833579	0.859228	11306	17953	9191
BraTS-GLI-01074-000	0.877041	0.934986	0.873013	0.895013	6960	10596	5463
BraTS-GLI-01075-000	0.767157	0.594957	0.776551	0.712888	376	-78481	573
BraTS-GLI-01076-000	0.867627	0.969021	0.957019	0.931222	3062	3466	250
BraTS-GLI-01077-000	0.184293	0.717167	0.180967	0.360809	1542	3977	1555

BraTS-GLI-01078-000	0.870481	0.959188	0.866055	0.898574	-7872	-1410	1836
BraTS-GLI-01079-000	0.902238	0.946303	0.876309	0.908283	-11444	-5397	2271
BraTS-GLI-01080-000	0.786701	0.866342	0.899669	0.850904	-7669	-827	1175
BraTS-GLI-01081-000	0.965725	0.933698	0.959322	0.952915	340	7441	-194
BraTS-GLI-01082-000	0.893215	0.959064	0.904862	0.919047	-16936	5552	717
BraTS-GLI-01083-000	0.93053	0.955907	0.915379	0.933939	-1051	-4418	1046
BraTS-GLI-01084-000	0.953495	0.944388	0.936663	0.944849	223	-373	1386
BraTS-GLI-01085-000	0.943064	0.945595	0.959346	0.949335	4144	-7725	1832
BraTS-GLI-01086-000	0.9202	0.973771	0.922834	0.938935	1254	-1161	773
BraTS-GLI-01087-000	0.865618	0.924551	0.890788	0.893652	1976	-8792	1836
BraTS-GLI-01088-000	0.853377	0.934414	0.854318	0.880703	799	2429	708
BraTS-GLI-01089-000	0.92851	0.927752	0.88916	0.915141	5092	1251	4814
BraTS-GLI-01090-000	0.896458	0.838375	0.906418	0.880417	-2088	-7970	-989
BraTS-GLI-01091-000	0.812132	0.284429	0.802607	0.633056	-63	-4910	-20
BraTS-GLI-01092-000	0.965127	0.95988	0.977067	0.967358	3533	10582	-1303
BraTS-GLI-01093-000	0.905348	0.949919	0.942234	0.9325	8290	-3518	2935
BraTS-GLI-01094-000	0.825241	0.926379	0.934725	0.895448	6290	-12399	656
BraTS-GLI-01095-000	0.964258	0.951477	0.954661	0.956798	3484	5854	2046
BraTS-GLI-01096-000	0.893861	0.951313	0.913945	0.919707	5928	5967	3250
BraTS-GLI-01097-000	0.965916	0.959052	0.973061	0.966009	2001	4932	-446
BraTS-GLI-01098-000	0.954112	0.873066	0.909739	0.912306	-19	-18236	1227
BraTS-GLI-01099-000	0.654696	0.84011	0.6068	0.700535	1430	-14373	1415
BraTS-GLI-01100-000	0.941961	0.947051	0.935817	0.94161	2983	8622	2015
BraTS-GLI-01101-000	0.961277	0.961355	0.962118	0.961583	1755	10686	-3278
BraTS-GLI-01102-000	0.9155	0.836706	0.872782	0.874996	1019	1789	1262
BraTS-GLI-01103-000	0.722169	0.923234	0.097274	0.580892	-13843	-10135	6140
BraTS-GLI-01104-000	0.65383	0.735938	0.700563	0.696777	-43740	-10718	-32377
BraTS-GLI-01105-000	0.9003	0.944478	0.83694	0.893906	4653	-2282	3746
BraTS-GLI-01106-000	0.887398	0.808258	0.856558	0.850738	87	-1391	1010

BraTS-GLI-01107-000	0.814275	0.959901	0.678544	0.817573	-1015	5465	-2365
BraTS-GLI-01108-000	0.905235	0.804719	0.89741	0.869121	-654	2526	197
BraTS-GLI-01109-000	0.921457	0.948349	0.908975	0.92626	-1462	3532	814
BraTS-GLI-01110-000	0.307557	0.75508	0.353622	0.472086	-4158	-2928	-2369
BraTS-GLI-01111-000	0.929983	0.906663	0.899527	0.912058	4506	-6137	5253
BraTS-GLI-01112-000	0.791173	0.978347	0.764999	0.84484	-12052	1430	-14121
BraTS-GLI-01113-000	0.904917	0.900616	0.824044	0.876525	2124	-3674	3346
BraTS-GLI-01114-000	0.932808	0.901179	0.90602	0.913336	-2293	11920	-595
BraTS-GLI-01115-000	0.926384	0.887466	0.922802	0.912217	-4684	16249	-3988
BraTS-GLI-01116-000	0.811603	0.692274	0.86635	0.790076	-2365	-31139	1050
BraTS-GLI-01117-000	0.962825	0.921732	0.943524	0.942694	1486	2658	2013
BraTS-GLI-01118-000	0.870927	0.936288	0.915774	0.907663	3111	4528	1336
BraTS-GLI-01119-000	0.78567	0.754477	0.749222	0.763123	4499	-11146	5791
BraTS-GLI-01120-000	0.9187	0.938853	0.940708	0.932754	-6684	9315	508
BraTS-GLI-01121-000	0.959748	0.970646	0.953925	0.96144	3409	1193	1947
BraTS-GLI-01122-000	0.947462	0.935844	0.942821	0.942043	3548	-6708	2155
BraTS-GLI-01123-000	0.872257	0.959612	0.826752	0.886207	-1330	90	34
BraTS-GLI-01124-000	0.910024	0.951083	0.90246	0.921189	3949	3824	2701
BraTS-GLI-01125-000	0.901691	0.956082	0.899414	0.919062	-6675	-1193	-4390
BraTS-GLI-01126-000	0.980262	0.97197	0.967844	0.973358	59	4179	-1812
BraTS-GLI-01127-000	0.940808	0.945351	0.921815	0.935991	-1096	3907	1437
BraTS-GLI-01128-000	0.818955	0.891589	0.793789	0.834777	-2179	2358	-434
BraTS-GLI-01129-000	0.95501	0.94327	0.919704	0.939328	1328	6935	1503
BraTS-GLI-01130-000	0.921803	0.801525	0.946217	0.889849	2823	9022	-143
BraTS-GLI-01131-000	0.942391	0.946433	0.926701	0.938508	748	4341	534
BraTS-GLI-01132-000	0.878975	0.897988	0.856818	0.877927	548	87	498
BraTS-GLI-01133-000	0.911884	0.886388	0.882711	0.893661	-694	1597	-264
BraTS-GLI-01134-000	0.895513	0.960702	0.929719	0.928644	-8279	6412	2940
BraTS-GLI-01135-000	0.928008	0.957237	0.938144	0.94113	-7182	-555	-695

BraTS-GLI-01136-000	0.972864	0.960697	0.93703	0.956864	-3394	8098	-7960
BraTS-GLI-01137-000	0.856666	0.921838	0.732741	0.837082	-2316	5231	-5595
BraTS-GLI-01138-000	0.946359	0.771961	0.920539	0.87962	74	-42391	649
BraTS-GLI-01139-000	0.94047	0.913247	0.930913	0.92821	-7261	-10901	1195
BraTS-GLI-01140-000	0.885487	0.944851	0.868018	0.899452	-10565	-619	-2908
BraTS-GLI-01141-000	0.950182	0.964856	0.928852	0.947963	-838	2485	-1677
BraTS-GLI-01142-000	0.932925	0.92953	0.923226	0.92856	-2454	23450	5167
BraTS-GLI-01143-000	0.810357	0.912508	0.784998	0.835955	-27238	4091	-6395
BraTS-GLI-01144-000	0.752652	0.927392	0.80377	0.827938	8084	6432	5302
BraTS-GLI-01145-000	0.919319	0.911382	0.876789	0.902497	-7533	8443	5407
BraTS-GLI-01146-000	0.373771	0.864954	0.477365	0.57203	40535	25379	21160
BraTS-GLI-01147-000	0.951109	0.885136	0.957774	0.93134	421	1032	340
BraTS-GLI-01148-000	0.973458	0.963887	0.968374	0.968573	645	4291	-877
BraTS-GLI-01149-000	0.275306	0.652456	0.236332	0.388031	2198	-4169	2566
BraTS-GLI-01150-000	0.548443	0.787326	0.663265	0.666345	-17933	-30173	-6258
BraTS-GLI-01151-000	0.688296	0.638782	0.692487	0.673188	-6257	-58171	-5331
BraTS-GLI-01152-000	0.695689	0.812636	0.655362	0.721229	-3397	-1433	-1350
BraTS-GLI-01153-000	0.93048	0.938642	0.895561	0.921561	-9559	-7389	-3959
BraTS-GLI-01154-000	0.910938	0.459843	0.898796	0.756526	-461	-7684	-562
BraTS-GLI-01155-000	0.932761	0.957024	0.901924	0.93057	-2334	-4270	421
BraTS-GLI-01156-000	0.878207	0.890932	0.783912	0.851017	-12558	17909	-4358
BraTS-GLI-01157-000	0.70454	0.887142	0.732551	0.774744	-24577	-7432	-11103
BraTS-GLI-01158-000	0.877543	0.901799	0.832119	0.870487	-4716	-15477	-3458
BraTS-GLI-01159-000	0.738811	0.891534	0.757534	0.79596	-143	-1722	-180
BraTS-GLI-01160-000	0.806277	0.88848	0.883129	0.859295	-15097	-4744	-6669
BraTS-GLI-01160-001	0.886693	0.934966	0.941848	0.921169	-6755	-3622	-2995
BraTS-GLI-01161-000	0.970649	0.915404	0.951495	0.945849	-1636	18867	-2903
BraTS-GLI-01162-000	0.890931	0.87461	0.876278	0.880606	358	-2527	93
BraTS-GLI-01163-000	0.890339	0.887345	0.831087	0.86959	1883	-4886	2039

BraTS-GLI-01164-000	0.964642	0.962442	0.951082	0.959388	-932	846	962
BraTS-GLI-01165-000	0.948141	0.939483	0.955316	0.947647	-754	1995	-1213
BraTS-GLI-01166-000	0.972401	0.92659	0.945038	0.94801	-357	1694	481
BraTS-GLI-01167-000	0.519426	0.634446	0.390953	0.514942	-12175	-47147	3411
BraTS-GLI-01168-000	0.928719	0.955917	0.901537	0.928724	-5580	2840	-154
BraTS-GLI-01169-000	0.140712	0.810842	0	0.317185	-27104	-37793	7327
BraTS-GLI-01170-000	0.469396	0.323502	0.493652	0.42885	-6468	-79381	-5330
BraTS-GLI-01171-000	0.907235	0.904724	0.911672	0.907877	-21607	6217	-15738
BraTS-GLI-01172-000	0.842965	0.906915	0.854211	0.86803	-26294	-5159	-11340
BraTS-GLI-01173-000	0.413445	0.95313	0.033915	0.46683	-10562	-2908	1925
BraTS-GLI-01174-000	0.805391	0.870079	0.858383	0.844617	-22637	-16233	-1438
BraTS-GLI-01175-000	0.921311	0.954088	0.922902	0.932767	-15953	-8464	-13529
BraTS-GLI-01176-000	0.237372	0.927981	0.086486	0.41728	-38828	-3006	494
BraTS-GLI-01177-000	0.902404	0.780533	0.890629	0.857856	-14461	-83524	-14773
BraTS-GLI-01178-000	0.950384	0.887962	0.914937	0.917761	-581	3175	469
BraTS-GLI-01179-000		0.522055			1376	-63720	1378
BraTS-GLI-01180-000	0.797802	0.881559	0.797316	0.825559	-7640	-15185	-8714
BraTS-GLI-01181-000	0.976115	0.93806	0.956963	0.957046	-211	3503	-173
BraTS-GLI-01182-000	0.001225	0.488185	0	0.163137	-3293	-124782	-3355
BraTS-GLI-01183-000	0.790269	0.840526	0.723599	0.784798	-13459	-26694	-7684
BraTS-GLI-01184-000	0.884849	0.733907	0.875594	0.83145	-2816	2703	-3056
BraTS-GLI-01185-000	0.89781	0.893255	0.867051	0.886039	-9205	-13647	-5664
BraTS-GLI-01186-000	0.85333	0.894241	0.866699	0.871423	-268	-808	-764
BraTS-GLI-01187-000	0.762137	0.943732	0.631834	0.779234	2360	2750	9700
BraTS-GLI-01188-000	0.422534	0.877705	0.565792	0.62201	18533	-9437	9441
BraTS-GLI-01189-000	0.7021	0.893369	0.643721	0.746397	-40697	-17458	-17372
BraTS-GLI-01190-000	0.939828	0.938881	0.919849	0.932852	-4758	-5003	-5048
BraTS-GLI-01191-000	0.851337	0.955339	0.758918	0.855198	8572	976	3840
BraTS-GLI-01192-000	0.144738	0.908522	0.133734	0.395665	40829	-23063	34619

BraTS-GLI-01193-000	0.810329	0.911369	0.848653	0.856784	-17043	1106	-7602
BraTS-GLI-01194-000	0.878142	0.924455	0.709	0.837199	-13120	-13615	7001
BraTS-GLI-01195-000	0.799316	0.890695	0.760122	0.816711	-9232	-11058	-5041
BraTS-GLI-01196-000	0.672325	0.80616	0.684782	0.721089	-9642	-18394	-7699
BraTS-GLI-01197-000	0.867539	0.918306	0.909494	0.898447	-8646	-1177	-4127
BraTS-GLI-01198-000	0.850132	0.825551	0.831886	0.835856	-25201	4807	-12543
BraTS-GLI-01199-000	0.952352	0.953191	0.926588	0.944044	-3640	7033	534
BraTS-GLI-01200-000	0.864086	0.960266	0.915068	0.91314	-11349	-772	260
BraTS-GLI-01201-000	0.815816	0.850951	0.799143	0.82197	-1856	-9138	-289
BraTS-GLI-01202-000	0.904416	0.921701	0.880128	0.902082	8060	-20580	4630
BraTS-GLI-01203-000	0.905074	0.856877	0.832622	0.864858	3030	17588	1432
BraTS-GLI-01204-000	0.845223	0.694709	0.801371	0.780434	1215	8544	1299
BraTS-GLI-01205-000	0.801016	0.7028	0.723018	0.742278	-4014	-7062	-1318
BraTS-GLI-01206-000	0.885482	0.921421	0.820006	0.875636	-5937	-11989	7308
BraTS-GLI-01207-000	0.875909	0.885301	0.849492	0.870234	1653	6673	1462
BraTS-GLI-01208-000	0.768677	0.832512	0.83203	0.811073	597	6397	263
BraTS-GLI-01209-000	0.859706	0.897074	0.815389	0.85739	-6633	2716	-8801
BraTS-GLI-01210-000	0.893884	0.787935	0.87876	0.853526	204	-4261	152
BraTS-GLI-01211-000	0.846295	0.894429	0.708598	0.816441	3240	8741	4701
BraTS-GLI-01212-000	0.867904	0.848656	0.795701	0.83742	1972	-69	1544
BraTS-GLI-01213-000	0.857249	0.888563	0.676979	0.807597	-2142	-752	1231
BraTS-GLI-01214-000	0.841161	0.865056	0.814319	0.840179	8041	20239	5476
BraTS-GLI-01215-000	0.779458	0.947022	0.839491	0.855324	-1900	-846	2730
BraTS-GLI-01216-000	0.81364	0.891355	0.636203	0.780399	-9965	16044	14730
BraTS-GLI-01217-000	0.790522	0.904751	0.721053	0.805442	6826	15978	5316
BraTS-GLI-01218-000	0.625059	0.788072	0.595573	0.669568	2289	12478	2077
BraTS-GLI-01219-000	0.876378	0.856641	0.964201	0.899073	4077	-1554	-709
BraTS-GLI-01220-000	0.947749	0.90083	0.898817	0.915799	1880	4834	1821
BraTS-GLI-01221-000	0.77472	0.904521	0.687437	0.788893	-39620	-11075	14894

BraTS-GLI-01222-000	0.817183	0.575962	0.620699	0.671281	-1235	-16865	1434
BraTS-GLI-01223-000	0.936796	0.940591	0.875198	0.917528	-123	5324	1426
BraTS-GLI-01224-000	0.935832	0.867018	0.856446	0.886432	-1484	-5200	264
BraTS-GLI-01225-000	0.899838	0.891238	0.953075	0.914717	5537	-14844	-761
BraTS-GLI-01226-000	0.893746	0.848807	0.856972	0.866508	-984	23796	-3028
BraTS-GLI-01227-000	0.950762	0.904798	0.889012	0.914857	3013	9793	3021
BraTS-GLI-01228-000	0.764452	0.937968	0.718473	0.806965	5025	5576	5730
BraTS-GLI-01229-000	0.855995	0.718796	0.797443	0.790745	5943	51374	2924
BraTS-GLI-01230-000	0.685205	0.822928	0.656783	0.721639	-3586	4134	-553
BraTS-GLI-01231-000	0.907502	0.923211	0.861894	0.897536	1007	2329	625
BraTS-GLI-01232-000	0.806359	0.770058	0.763146	0.779855	-131	-21460	1432
BraTS-GLI-01233-000	0.9092	0.892425	0.88006	0.893895	1125	-2018	1415
BraTS-GLI-01234-000	0.85846	0.874846	0.793301	0.842202	-2530	1437	1295
BraTS-GLI-01235-000	0.923481	0.945455	0.902369	0.923768	2650	3986	2172
BraTS-GLI-01236-000	0.855341	0.926559	0.821932	0.867944	-8489	1695	-2608
BraTS-GLI-01237-000	0.82104	0.832105	0.773229	0.808791	-1892	10231	11122
BraTS-GLI-01238-000	0.817019	0.880953	0.752006	0.816659	-2327	-3298	-402
BraTS-GLI-01239-000	0.829384	0.869335	0.879287	0.859335	747	10020	1147
BraTS-GLI-01240-000	0.805988	0.502963	0.761034	0.689995	367	-19360	458
BraTS-GLI-01241-000	0.868187	0.882201	0.846497	0.865628	-4269	-10125	4544
BraTS-GLI-01242-000	0.941852	0.923061	0.908078	0.92433	-397	-1984	-764
BraTS-GLI-01243-000	0.908713	0.875404	0.858314	0.88081	-3018	-5773	-484
BraTS-GLI-01244-000	0.841217	0.936826	0.837288	0.871777	3601	-2854	2287
BraTS-GLI-01245-000	0.370298	0.710134	0.618902	0.566445	-8779	-25871	1281
BraTS-GLI-01246-000	0.87181	0.95807	0.847828	0.892569	-3912	837	6366
BraTS-GLI-01247-000	0.742627	0.825212	0.59408	0.72064	1313	-36	1541
BraTS-GLI-01248-000	0.939995	0.925206	0.833796	0.899666	2517	4592	3672
BraTS-GLI-01249-000	0.93041	0.8211	0.796066	0.849192	-365	6912	1578
BraTS-GLI-01250-000	0.932077	0.957782	0.879697	0.923185	3372	2767	4744

BraTS-GLI-01251-000	0.947064	0.89815	0.897532	0.914249	5334	12929	6096
BraTS-GLI-01252-000	0.804485	0.848493	0.749519	0.800832	1176	3734	1147
BraTS-GLI-01253-000	0.843037	0.831318	0.76422	0.812858	-19311	8457	-3969
BraTS-GLI-01254-000	0.907126	0.933916	0.915766	0.918936	-8264	5955	797
BraTS-GLI-01255-000	0.565247	0.688849	0.628995	0.627697	1869	5411	1435
BraTS-GLI-01256-000	0.765026	0.942957	0.896233	0.868072	-24758	7074	2427
BraTS-GLI-01257-000	0.543253	0.625886	0.542854	0.570665	6479	24423	4689
BraTS-GLI-01258-000	0.898366	0.91085	0.804854	0.871357	170	7419	1128
BraTS-GLI-01259-000	0.857888	0.913892	0.86518	0.878987	877	3047	762
BraTS-GLI-01260-000	0.945779	0.947636	0.894313	0.929243	2348	-2033	6717
BraTS-GLI-01261-000	0.897594	0.927261	0.831006	0.885287	-2167	4148	2783
BraTS-GLI-01262-000	0.907883	0.778219	0.813964	0.833355	-3323	-46300	3054
BraTS-GLI-01263-000	0.771694	0.849027	0.724892	0.781871	-10048	11422	4913
BraTS-GLI-01264-000	0.818449	0.90372	0.81224	0.844803	-4115	291	-944
BraTS-GLI-01265-000	0.90461	0.904189	0.894212	0.901003	-7677	1931	-2899
BraTS-GLI-01266-000	0.671212	0.908534	0.784975	0.78824	9638	19148	8
BraTS-GLI-01267-000	0.93532	0.74236	0.88933	0.85567	-3046	38568	452
BraTS-GLI-01268-000	0.939395	0.917154	0.820936	0.892495	-709	-2987	492
BraTS-GLI-01269-000	0.207928	0.630117	0.262738	0.366928	-26500	-23910	-11057
BraTS-GLI-01270-000	0.785387	0.944093	0.762583	0.830688	-32258	-3685	13007
BraTS-GLI-01271-000	0.866055	0.909804	0.763133	0.846331	-1569	3002	4090
BraTS-GLI-01272-000	0.915249	0.892686	0.863708	0.890548	1825	8958	1750
BraTS-GLI-01273-000	0.868511	0.756985	0.605073	0.743523	-206	-575	1909
BraTS-GLI-01274-000	0.783634	0.834982	0.791634	0.803416	2807	10437	1780
BraTS-GLI-01275-000	0.881193	0.904684	0.790435	0.858771	-362	3399	789
BraTS-GLI-01276-000	0.885297	0.708514	0.810468	0.801426	-463	2197	-56
BraTS-GLI-01277-000	0.877833	0.91411	0.866053	0.885999	2581	10468	1314
BraTS-GLI-01278-000	0.903356	0.921231	0.956429	0.927005	2109	3208	587
BraTS-GLI-01279-000	0.789244	0.589619	0.802253	0.727039	933	-6101	838

BraTS-GLI-01280-000	0.970849	0.949456	0.956876	0.959061	2078	13607	586
BraTS-GLI-01281-000	0.919083	0.956931	0.838041	0.904685	1815	484	5864
BraTS-GLI-01282-000	0.942756	0.765385	0.919663	0.875935	1559	-40263	1976
BraTS-GLI-01283-000	0.944496	0.937921	0.920529	0.934315	-1367	10667	-3103
BraTS-GLI-01284-000	0.90111	0.906264	0.843649	0.883674	-1346	-9087	-189
BraTS-GLI-01285-000	0.915917	0.92889	0.897751	0.914186	-7420	-2374	-4203
BraTS-GLI-01286-000	0.946829	0.897372	0.919806	0.921336	2635	-1268	3093
BraTS-GLI-01287-000	0.942178	0.954837	0.937446	0.94482	2337	-48	1634
BraTS-GLI-01288-000	0.947204	0.929956	0.8813	0.919487	-986	13729	3435
BraTS-GLI-01289-000	0.904245	0.927865	0.839639	0.890583	2199	4954	4794
BraTS-GLI-01290-000	0.935495	0.87773	0.905645	0.90629	1455	-2449	1799
BraTS-GLI-01291-000	0.799397	0.940009	0.543492	0.760966	-15232	-10489	6379
BraTS-GLI-01292-000	0.944711	0.951091	0.939048	0.94495	-2349	6809	-1244
BraTS-GLI-01293-000	0.081726	0.295628	0.417802	0.265052	-88634	-101053	-7261
BraTS-GLI-01294-000	0.938148	0.905894	0.95376	0.932601	2152	12991	1026
BraTS-GLI-01295-000	0.846141	0.885311	0.832885	0.854779	-2840	7601	-507
BraTS-GLI-01296-000	0.880579	0.887107	0.914624	0.894104	2684	7302	2510
BraTS-GLI-01297-000	0.861892	0.550397	0.833917	0.748735	-1352	-36741	-1015
BraTS-GLI-01298-000	0.933465	0.94433	0.943966	0.940587	1579	-5516	-1526
BraTS-GLI-01299-000	0.743767	0.931701	0.817049	0.830839	10029	-5822	3584
BraTS-GLI-01300-000	0.738484	0.794442	0.725988	0.752971	4558	-242	4669
BraTS-GLI-01301-000	0.937435	0.896464	0.931248	0.921716	174	424	195
BraTS-GLI-01302-000	0.913957	0.934352	0.930415	0.926241	-14493	11300	-6762
BraTS-GLI-01303-000	0.951097	0.947888	0.925156	0.94138	-2281	-9028	-749
BraTS-GLI-01304-000	0.937898	0.94967	0.955412	0.94766	3153	9300	901
BraTS-GLI-01305-000	0.936612	0.953886	0.959375	0.949958	1155	4296	-2267
BraTS-GLI-01306-000	0.882881	0.949941	0.889566	0.907462	1357	4703	-302
BraTS-GLI-01307-000	0.962777	0.967403	0.955467	0.961883	2511	-1359	-830
BraTS-GLI-01308-000	0.913297	0.535957	0.913891	0.787715	-140	10182	-181

BraTS-GLI-01309-000	0.704516	0.942578	0.683917	0.777004	8648	7053	5017
BraTS-GLI-01310-000	0.92411	0.817124	0.919124	0.886786	888	-10038	683
BraTS-GLI-01311-000	0.972536	0.889071	0.963509	0.941706	-1389	-10966	-3345
BraTS-GLI-01312-000	0.962271	0.960739	0.974447	0.965819	1606	5359	-520
BraTS-GLI-01313-000	0.926473	0.966518	0.865556	0.919516	-1213	-2965	759
BraTS-GLI-01314-000	0.385688	0.860719	0.338239	0.528215	-2633	-7991	-1997
BraTS-GLI-01315-000	0.930763	0.961508	0.895094	0.929122	-9153	2123	-4073
BraTS-GLI-01316-000	0.888909	0.954429	0.895547	0.912962	5940	3569	3085
BraTS-GLI-01317-000	0.967356	0.959386	0.955373	0.960705	912	9327	-248
BraTS-GLI-01318-000	0.965171	0.915166	0.944492	0.94161	261	2994	-289
BraTS-GLI-01319-000	0.966371	0.93559	0.945883	0.949281	-1529	-6485	458
BraTS-GLI-01320-000	0.948502	0.957427	0.942413	0.949447	5310	-6362	-1965
BraTS-GLI-01321-000	0.970462	0.966467	0.934705	0.957212	-4873	6127	-9175
BraTS-GLI-01322-000	0.914765	0.940971	0.860262	0.905332	129	-594	195
BraTS-GLI-01323-000	0.970402	0.971906	0.916361	0.95289	-182	3235	2515
BraTS-GLI-01324-000	0.964232	0.933788	0.936712	0.944911	-1782	-31572	-1624
BraTS-GLI-01325-000	0.842688	0.858019	0.889141	0.863283	-3340	-27422	-3745
BraTS-GLI-01326-000	0.956577	0.943883	0.973978	0.958146	3286	-8903	-560
BraTS-GLI-01327-000	0.808312	0.919494	0.858421	0.862076	1375	1600	890
BraTS-GLI-01328-000	0.835904	0.965788	0.935898	0.91253	-1528	-274	-2987
BraTS-GLI-01329-000	0.947507	0.964979	0.879179	0.930555	501	1727	4112
BraTS-GLI-01330-000	0.886999	0.918779	0.901265	0.902348	-3188	635	3502
BraTS-GLI-01331-000	0.914172	0.878906	0.859491	0.88419	-51	-12382	230
BraTS-GLI-01332-000	0.958954	0.951881	0.938341	0.949725	286	-2500	1284
BraTS-GLI-01333-000	0.882271	0.470296	0.890633	0.747733	-71	5254	-127
BraTS-GLI-01334-000	0.862299	0.921695	0.914477	0.89949	3537	8623	1393
BraTS-GLI-01335-000	0.87312	0.963617	0.921833	0.919524	3484	5792	877
BraTS-GLI-01336-000	0.918963	0.957915	0.894958	0.923945	1056	2668	1824
BraTS-GLI-01337-000	0.972035	0.962729	0.942336	0.959033	1685	3267	2113

BraTS-GLI-01338-000	0.971857	0.933164	0.962469	0.95583	699	2756	449
BraTS-GLI-01339-000	0.943897	0.95451	0.930611	0.943006	2159	-5447	1562
BraTS-GLI-01340-000	0.931314	0.943592	0.935407	0.936771	1069	1137	834
BraTS-GLI-01341-000	0.898321	0.425719	0.852555	0.725532	-244	8297	82
BraTS-GLI-01342-000	0.911189	0.771724	0.917747	0.866887	-6008	18469	-4590
BraTS-GLI-01343-000	0.975154	0.967332	0.9639	0.968795	-969	7230	-871
BraTS-GLI-01344-000	0.906555	0.923518	0.944063	0.924712	585	-261	320
BraTS-GLI-01345-000	0.972284	0.934881	0.960765	0.955977	-715	1801	-441
BraTS-GLI-01346-000	0.956904	0.974437	0.940566	0.957303	-396	293	2640
BraTS-GLI-01347-000	0.560684	0.920216	0.524895	0.668598	512	57	559
BraTS-GLI-01348-000	0.965941	0.863799	0.956301	0.92868	495	7779	1032
BraTS-GLI-01349-000	0.893632	0.562292	0.891975	0.782633	407	-4382	370
BraTS-GLI-01350-000		0.926947			911	2614	839
BraTS-GLI-01351-000	0.941787	0.857031	0.944784	0.914534	4098	-22621	2007
BraTS-GLI-01352-000	0.57958	0.799014	0.507324	0.628639	11624	19952	9925
BraTS-GLI-01353-000	0.928752	0.94489	0.918489	0.930711	2535	3012	2009
BraTS-GLI-01354-000	0.895676	0.887938	0.937711	0.907109	413	-5143	900
BraTS-GLI-01355-000	0.832292	0.955276	0.913426	0.900331	7123	7674	2985
BraTS-GLI-01356-000	0.495172	0.537489	0.679454	0.570705	-63811	-83965	-19617
BraTS-GLI-01357-000	0.96939	0.960906	0.923277	0.951191	1238	5316	1902
BraTS-GLI-01358-000	0.968066	0.971668	0.963908	0.967881	-2362	3869	-1275
BraTS-GLI-01359-000	0.955244	0.916741	0.959784	0.943923	736	3929	291
BraTS-GLI-01360-000	0.970668	0.956385	0.95181	0.959621	879	3338	-550
BraTS-GLI-01361-000	0.959275	0.955987	0.947093	0.954118	-1021	3099	-3209
BraTS-GLI-01362-000	0.923323	0.943347	0.932577	0.933082	6536	11788	994
BraTS-GLI-01363-000	0.843334	0.803724	0.658708	0.768588	-597	3108	839
BraTS-GLI-01364-000	0.743151	0.136802	0.698925	0.526292	592	-44729	660
BraTS-GLI-01365-000	0.347486	0.483537	0.372258	0.401094	-2115	-18562	-2062
BraTS-GLI-01366-000	0.22919	0.804075	0.203018	0.412094	392	-2174	234

BraTS-GLI-01367-000	0.873727	0.879007	0.70366	0.818798	1608	-4311	3714
BraTS-GLI-01368-000	0.903159	0.936524	0.955183	0.931622	4862	-11821	828
BraTS-GLI-01369-000	0.953667	0.836863	0.921792	0.904107	533	-13937	269
BraTS-GLI-01370-000	0.512452	0.827433	0.55477	0.631552	3558	-18102	2588
BraTS-GLI-01371-000	0.926179	0.953993	0.94263	0.940934	4015	8293	358
BraTS-GLI-01372-000	0.692795	0.817148	0.610115	0.706686	2512	-238	2397
BraTS-GLI-01373-000	0.95714	0.894823	0.962794	0.938252	2384	-714	-229
BraTS-GLI-01374-000	0.979633	0.962977	0.963223	0.968611	-959	10441	-3410
BraTS-GLI-01375-000	0.426955	0.927184	0.537294	0.630478	4558	3445	2771
BraTS-GLI-01376-000	0.935096	0.968077	0.880468	0.92788	-154	1755	-192
BraTS-GLI-01377-000	0.943778	0.858258	0.94341	0.915148	-8374	-6324	-8248
BraTS-GLI-01378-000	0.784052	0.856872	0.920096	0.853673	-42592	-30062	-5951
BraTS-GLI-01379-000	0.798089	0.962937	0.818818	0.859948	8064	3043	5537
BraTS-GLI-01380-000	0.963844	0.72815	0.970587	0.887527	217	-29098	41
BraTS-GLI-01381-000	0.963376	0.970789	0.975037	0.969734	3537	3597	-2431
BraTS-GLI-01382-000	0.941738	0.937721	0.952347	0.943936	2247	2855	932
BraTS-GLI-01383-000	0.891768	0.918507	0.942711	0.917662	5985	-3497	1096
BraTS-GLI-01384-000	0.950027	0.941289	0.923306	0.938207	3889	13170	3400
BraTS-GLI-01385-000	0.937388	0.947023	0.957209	0.947207	-7351	12025	-4508
BraTS-GLI-01386-000	0.906603	0.958772	0.910537	0.925304	5747	5717	3204
BraTS-GLI-01387-000	0.635617	0.85082	0.680681	0.722373	4768	-13016	3528
BraTS-GLI-01388-000	0.964603	0.961433	0.946781	0.957606	3331	665	2226
BraTS-GLI-01389-000	0.929722	0.94448	0.942249	0.938817	921	1003	495
BraTS-GLI-01390-000	0.9272	0.949043	0.92078	0.932341	-2048	-3950	-773
BraTS-GLI-01391-000	0.922858	0.934961	0.918583	0.925467	-422	-9072	2571
BraTS-GLI-01392-000	0.96647	0.931499	0.965116	0.954362	904	-63	1157
BraTS-GLI-01393-000	0.957995	0.975937	0.971047	0.968326	1459	2606	533
BraTS-GLI-01394-000	0.792156	0.903614	0.858981	0.851583	2480	437	1443
BraTS-GLI-01395-000	0.709889	0.901559	0.745783	0.785744	1972	4584	1594

BraTS-GLI-01396-000	0.652421	0.973127	0.908193	0.84458	10694	1686	930
BraTS-GLI-01397-000	0.807967	0.895811	0.922788	0.875522	-7512	-1120	-1993
BraTS-GLI-01398-000	0.893318	0.955204	0.925019	0.924514	2538	3689	726
BraTS-GLI-01399-000	0.935938	0.854529	0.942884	0.911117	1664	-621	1028
BraTS-GLI-01400-000	0.882743	0.95427	0.926736	0.92125	1524	-2540	-326
BraTS-GLI-01401-000	0.96007	0.964211	0.962095	0.962126	3525	9212	1475
BraTS-GLI-01402-000	0.90823	0.925456	0.929789	0.921158	859	-3784	363
BraTS-GLI-01403-000	0.932009	0.958468	0.955684	0.948721	-4233	4942	-2710
BraTS-GLI-01404-000	0.942983	0.8509	0.825126	0.873003	159	3228	-268
BraTS-GLI-01405-000	0.522084	0.6895	0.519325	0.57697	475	-2836	439
BraTS-GLI-01406-000	0.951161	0.956031	0.946273	0.951155	3828	-4840	2530
BraTS-GLI-01407-000	0.917054	0.926284	0.936106	0.926481	954	1170	545
BraTS-GLI-01408-000	0.926444	0.952052	0.901735	0.926743	1757	3160	1406
BraTS-GLI-01409-000	0.935949	0.90213	0.942087	0.926722	-914	2173	-575
BraTS-GLI-01410-000	0.853982	0.973412	0.932863	0.920085	-18916	923	2087
BraTS-GLI-01411-000	0.95834	0.8774	0.88827	0.908004	-304	25207	2348
BraTS-GLI-01412-000	0.881112	0.92019	0.893106	0.898136	-1163	3286	-628
BraTS-GLI-01413-000	0.921837	0.906174	0.91503	0.914347	-447	-5540	2198
BraTS-GLI-01414-000	0.52154	0.932097	0.821945	0.758527	-45880	-6725	-3187
BraTS-GLI-01415-000	0.318232	0.610438	0.548081	0.49225	-36228	-43043	-10259
BraTS-GLI-01416-000	0.841877	0.916316	0.836268	0.86482	-16016	-5895	10653
BraTS-GLI-01417-000	0.8198	0.798158	0.802954	0.806971	-6544	-9240	-3541
BraTS-GLI-01418-000	0.919805	0.910384	0.865744	0.898645	6630	12841	9092
BraTS-GLI-01419-000	0.746729	0.929682	0.885415	0.853942	-29839	-10298	-7340
BraTS-GLI-01420-000	0.79466	0.772786	0.797189	0.788212	-1520	-5448	-1109
BraTS-GLI-01421-000	0.950219	0.950093	0.902684	0.934332	-1068	1918	2386
BraTS-GLI-01422-000	0.000876	0.856305	0.061911	0.306364	-9114	-2656	-673
BraTS-GLI-01423-000	0.938189	0.907199	0.921908	0.922432	459	678	645
BraTS-GLI-01424-000	0.894157	0.912579	0.796847	0.867861	-21686	18029	-20661

BraTS-GLI-01425-000	0.927655	0.920813	0.941082	0.92985	-2508	21485	-1816
BraTS-GLI-01426-000	0.907896	0.911099	0.81178	0.876925	-848	7819	-528
BraTS-GLI-01427-000	0.90011	0.904652	0.856321	0.887027	-3602	544	-2248
BraTS-GLI-01428-000	0.309493	0.827715	0.180449	0.439219	-43211	-17713	-13995
BraTS-GLI-01429-000	0.353642	0.737369	0.339257	0.476756	-2812	-19305	13106
BraTS-GLI-01430-000	0.861459	0.912904	0.870867	0.881743	-24112	-2882	2116
BraTS-GLI-01431-000	0.708724	0.799415	0.831175	0.779771	-16686	4613	-3736
BraTS-GLI-01432-000	0.546425	0.939309	0.605209	0.696981	-48927	-15289	7170
BraTS-GLI-01433-000	0.477169	0.863338			-1605	10661	1370
BraTS-GLI-01434-000	0.582804	0.85721			-158	7164	2420
BraTS-GLI-01435-000	0.246746	0.87956			-14491	5660	3
BraTS-GLI-01436-000	0.502586	0.926752			-28800	3146	82
BraTS-GLI-01437-000	0.401411	0.938514			-20896	-1915	1192
BraTS-GLI-01438-000	0.209102	0.854707	0.874636	0.646148	-48758	-16601	137
BraTS-GLI-01439-000	0	0.830945			-9284	-721	25
BraTS-GLI-01440-000	0.422748	0.901096			-14392	3558	360
BraTS-GLI-01441-000	0.832534	0.85471	0.743082	0.810109	4807	39149	7094
BraTS-GLI-01442-000	0.966082	0.921014	0.91961	0.935569	-2169	10486	-2849
BraTS-GLI-01443-000	0.731978	0.863715	0.738961	0.778218	-12908	-6265	-8147
BraTS-GLI-01444-000	0.935032	0.931883	0.941987	0.9363	2705	-13724	-129
BraTS-GLI-01445-000	0.920548	0.907658	0.854521	0.894243	-6368	23789	-9598
BraTS-GLI-01446-000	0.85802	0.941601	0.844138	0.881253	-3563	2449	5303
BraTS-GLI-01447-000	0.917285	0.94273	0.913438	0.924485	858	8945	-2405
BraTS-GLI-01448-000	0.796852	0.918141	0.876299	0.863764	2235	6679	542
BraTS-GLI-01449-000	0.930133	0.940762	0.889812	0.920236	-5840	-1810	-5401
BraTS-GLI-01450-000	0.912814	0.895971	0.891192	0.899992	1169	-8810	1229
BraTS-GLI-01451-000	0.866206	0.713438	0.858172	0.812605	-130	-4283	-505
BraTS-GLI-01452-000	0.919897	0.848189	0.897524	0.888537	-613	4859	-720
BraTS-GLI-01453-000	0.860893	0.802513	0.870658	0.844688	521	-3927	315

BraTS-GLI-01454-000	0.836718	0.847905	0.804723	0.829782	-2194	-12702	-2902
BraTS-GLI-01455-000	0.93442	0.925614	0.907066	0.922367	7196	23603	8093
BraTS-GLI-01456-000	0.928208	0.94729	0.909805	0.928434	2071	10373	433
BraTS-GLI-01457-000	0.892317	0.905671	0.881931	0.893306	-5243	11457	-3408
BraTS-GLI-01458-000	0.794683	0.939982	0.773131	0.835932	-5289	-2514	3678
BraTS-GLI-01459-000	0.925107	0.891366	0.937914	0.918129	-338	1432	478
BraTS-GLI-01460-000	0.917472	0.918208	0.882406	0.906029	582	8927	8240
BraTS-GLI-01461-000	0.922346	0.887148	0.895828	0.901774	2681	-4227	2078
BraTS-GLI-01463-000	0.912474	0.961699	0.896432	0.923535	-10876	3302	-10687
BraTS-GLI-01464-000	0.856106	0.856849	0.90067	0.871208	-1951	7053	-657
BraTS-GLI-01465-000	0.943117	0.960104	0.939568	0.947597	2772	9063	-1333
BraTS-GLI-01466-000	0.822125	0.955715	0.884824	0.887554	-27287	4965	-10621
BraTS-GLI-01467-000	0.928102	0.956743	0.842396	0.90908	-14385	9973	-27924
BraTS-GLI-01468-000	0.935311	0.896027	0.913634	0.914991	-921	9081	1845
BraTS-GLI-01469-000	0.703284	0.730239	0.742739	0.725421	-2186	-17962	-1191
BraTS-GLI-01470-000	0.913471	0.793116	0.881531	0.862706	1509	-10524	1587
BraTS-GLI-01471-000	0.893923	0.950277	0.899348	0.914516	-17525	-2047	-5894
BraTS-GLI-01472-000	0.903613	0.852647	0.846891	0.867717	6635	18008	8724
BraTS-GLI-01473-000	0.894194	0.879919	0.859651	0.877921	-4868	-30023	-1901
BraTS-GLI-01474-000	0.724029	0.890435	0.828434	0.814299	-20784	609	-3681
BraTS-GLI-01475-000	0.912811	0.965412	0.876284	0.918169	-5877	3566	-6381
BraTS-GLI-01476-000	0.543395	0.930946	0.200456	0.558265	-12663	1413	1308
BraTS-GLI-01477-000	0.103941	0.911072			-71980	-24568	1998
BraTS-GLI-01478-000	0.299147	0.86304	0.181967	0.448051	-84549	-35865	10857
BraTS-GLI-01479-000	0.497099	0.802193	0.710333	0.669875	-121262	-46475	-38389
BraTS-GLI-01480-000	0.03542	0.339434			-9934	98	889
BraTS-GLI-01481-000	0.718645	0.91234	0.630278	0.753754	3122	-4389	3021
BraTS-GLI-01482-000	0.663822	0.73751	0.910116	0.770483	-1355	-4022	-89
BraTS-GLI-01483-000	0.343358	0.866958			-43067	-17963	13290

BraTS-GLI-01484-000	0.121452	0.566302			-122749	-91891	3831
BraTS-GLI-01485-000	0.180462	0.847371	0.252202	0.426678	5452	12374	2932
BraTS-GLI-01487-000	0.754859	0.955342	0.75065	0.820284	5291	4820	4402
BraTS-GLI-01488-000	0.758528	0.929077	0.760439	0.816014	-576	-3226	-1509
BraTS-GLI-01489-000	0.610864	0.922084	0.492697	0.675215	-76613	-26328	2598
BraTS-GLI-01490-000	0.089077	0.322845	0.473593	0.295172	-75943	-86641	-1991
BraTS-GLI-01491-000	0.028901	0.770288			-20694	-8256	200
BraTS-GLI-01492-000	0.934038	0.95046	0.95034	0.944946	-337	2969	-682
BraTS-GLI-01493-000	0.779766	0.930116	0.88408	0.864654	1898	2585	-1214
BraTS-GLI-01494-000	0.586409	0.939288			-16248	-11666	4884
BraTS-GLI-01495-000	0.367221	0.824775	0.484214	0.558737	21087	-56080	8554
BraTS-GLI-01496-000	0.68939	0.928189	0.023336	0.546972	-17600	-25407	5833
BraTS-GLI-01497-000	0.521204	0.912048	0.741872	0.725041	-33875	-8938	-33
BraTS-GLI-01498-000	0.520988	0.706944	0.858681	0.695537	-77061	-118661	-2045
BraTS-GLI-01499-000	0.036301	0.834534			-35445	-22463	17
BraTS-GLI-01500-000	0	0.883927	0.015123	0.299683	-10755	-471	-507
BraTS-GLI-01501-000	0.593347	0.96163	0.740924	0.7653	-51208	2034	24
BraTS-GLI-01502-000	0.537417	0.623959			-6227	-29349	5422
BraTS-GLI-01503-000	0.007822	0.90624	0.40708	0.44038	-14439	-1395	38
BraTS-GLI-01504-000	0.025151	0.511079	0.147354	0.227861	-85998	-71801	175
BraTS-GLI-01505-000	0.105315	0.856907	0.778039	0.580087	-62859	-22672	406
BraTS-GLI-01506-000	0.368751	0.812316	0.4372	0.539422	-33168	-45234	5168
BraTS-GLI-01507-000	0.249417	0.713356	0.718256	0.560343	-87805	-50892	1685
BraTS-GLI-01508-000	0.101687	0.447077	0.410714	0.319826	-23844	-25030	-968
BraTS-GLI-01509-000	0.030316	0.849667			-27398	-5812	30
BraTS-GLI-01510-000	0.126772	0.670944			-65098	-60692	2939
BraTS-GLI-01511-000	0.336225	0.833874			-73965	-49787	4209
BraTS-GLI-01512-000	0.195335	0.724403	0	0.306579	-83882	-63391	176
BraTS-GLI-01513-000	0.131243	0.770435	0.697932	0.533203	-142358	-73553	-408

BraTS-GLI-01514-000	0.307941	0.771849	0.476326	0.518705	-86157	-56725	1687
BraTS-GLI-01515-000	0.235261	0.834834			-64109	-5856	15960
BraTS-GLI-01516-000	0.144446	0.712136	0.854478	0.570353	-6977	-8274	-8
BraTS-GLI-01517-000	0.557394	0.888685			-16066	5547	21337
BraTS-GLI-01518-000	0.630203	0.94655	0.456369	0.677707	-42572	663	1925
BraTS-GLI-01519-000	0.840012	0.685346	0.865815	0.797057	-156	-11365	14
BraTS-GLI-01520-000	0.657953	0.948335	0.628272	0.744853	-11432	1112	-27
BraTS-GLI-01521-000	0.182309	0.824629	0.306671	0.437869	-80574	-62029	6169
BraTS-GLI-01522-000	0.772501	0.934707	0.870964	0.859391	-1143	1541	371
BraTS-GLI-01523-000	0.011946	0.701219	0.452577	0.38858	-158863	-94472	-1032
BraTS-GLI-01524-000	0.547482	0.938177			-50508	-7404	1602
BraTS-GLI-01525-000	0.775039	0.786663	0.755412	0.772372	2833	-7719	4303
BraTS-GLI-01526-000	0.011737	0.909449			-28411	-7408	23
BraTS-GLI-01527-000	0.577734	0.958397	0.212121	0.582751	-36563	4811	356
BraTS-GLI-01528-000	0.442598	0.872887	0.000661	0.438715	-68520	-41878	-272
BraTS-GLI-01529-000	0.000278	0.888976			-50186	-11739	124
BraTS-GLI-01530-000	0	0.744494	0	0.248165	-8732	2250	849
BraTS-GLI-01531-000	0.395827	0.910066	0.167904	0.491266	-92844	-13422	11751
BraTS-GLI-01532-000	0.2581	0.876279	0.632271	0.588883	-128447	-18941	-7279
BraTS-GLI-01533-000	0.096914	0.851438	0.09486	0.347737	-166764	-64495	14795
BraTS-GLI-01534-000	0.136715	0.883139			-42279	-7995	1011
BraTS-GLI-01535-000	0.015935	0.951188			-31686	1761	765
BraTS-GLI-01536-000	0.372527	0.832296	0.726628	0.643817	-7404	19691	-239
BraTS-GLI-01537-000	0.956414	0.940484	0.951679	0.949526	-1016	-19165	-391
BraTS-GLI-01610-000	0.736175	0.786388	0.758067	0.76021	2284	5501	2296
BraTS-GLI-01657-000	0.91793	0.967908	0.939767	0.941868	2426	416	1559
BraTS-GLI-01658-000	0.842384	0.772921	0.889484	0.834929	-15375	-51753	-2809
BraTS-GLI-01659-000	0.958165	0.95592	0.946108	0.953398	-4153	4337	-5748
BraTS-GLI-01660-000	0.820558	0.836937	0.828363	0.82862	-14044	-18460	-7457

BraTS-GLI-01661-000	0.944086	0.940871	0.945854	0.943604	1428	-5759	-223
BraTS-GLI-01662-000	0.973884	0.958068	0.944179	0.95871	-1882	1575	1331
BraTS-GLI-01663-000	0.613167	0.940828	0.64058	0.731525	888	6373	972
BraTS-GLI-01664-000	0.049543	0.939138	0.453419	0.4807	-171526	-13300	-4758
BraTS-GLI-01665-000	0.878763	0.952175	0.858981	0.89664	337	6236	-1602
BraTS-GLI-01666-000	0.182397	0.894691			-53365	-20829	2209