Advancing Biomedical Image Segmentation of Lower-Grade Gliomas using Transfer Learning

MSc Research Project
Mcs in Data analytics

Yaswanth vanapalli
Student ID: x23196718

School of Computing
National College of Ireland

Supervisor: Musfira Jilani

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Yaswanth vanapalli
Student ID:	X23196718
Programme:	Msc in Data Analytics
Year:	2024
Module:	MSc Research Project
Supervisor:	Musfira Jilani
Submission Due Date:	29/01/2025
Project Title:	Advancing Biomedical Image Segmentation of Lower-Grade Gliomas using Transfer Learning
Word Count:	7735
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Yaswanth vanapalli
Date:	29th December 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Advancing Biomedical Image Segmentation of Lower-Grade Gliomas using Transfer Learning

Yaswanth vanapalli
X23196718

Abstract

This study aims to enhance the segmentation of LGGs in MRI scans using trans- fer learning, particularly transformer-based pretrained models. Medical imaging is challenging with LGGs due to their complex and diffuse nature, as they are clas- sified as brain tumors. There is an issue with traditional approaches of how these tumors are segmented which tends to be so much time consuming and very sub- jective. To address these challenges, this study leverages the Swin Transformer, a robust vision transformer, to enhance segmentation efficiency and precision withoutdemanding significant annotated datasets and computational power. The proposed method adapts the Swin Transformer model that is trained for large image data,to detect tumor regions accurately. The model was assessed based on four main measures such as the Dice Similarity Coefficient (DSC), Intersection over Union (IoU), and classification accuracy of a test set. The experimental evaluation exhib- its high efficiency with a test accuracy of 99.64%, the mean IoU of 0.837, and the mean Dice score of 0.861. However, it was observed that a few imperfections exist in the model, and those are mainly related to the recall of smaller or more intricate tumor regions. This study establishes the viability of transformer based models in medical image segmentation and offers a solid starting point for improving LGG diagnosis in healthcare facilities.

1 Introduction

1.1 Background

LGGs are a group of gliomas that are highly diverse and challenging in terms of neuro-oncology. LGGs are slower growing than higher-grade gliomas but are equally dangerous due to their invasive nature thus correct diagnosis and early treatment is critical. MRI is currently the main imaging method used in the diagnosis and evaluation of LGGs owing to its ability to provide detailed information about the tumour's morphology and its pattern of growth. Nevertheless, the problem of accurate LGGs segmentation from MRI remains challenging mainly due to its margins and the structure of the brain tissue. In clinical work, it is particularly important to differentiate LGGs more accurately for treatments, assessing the development of the disease and the prognosis of the patients. However, current approaches in LGG segmentation based on manual delineation by radiologists are quite inaccurate and time-consuming, and they are sensitive to inter-observer variability. These challenges have created a huge demand for the creation of automated segmentation algorithms that can help doctors and other medical practitioners to offer accurate and reliable results.

In Recent years, many deep learning techniques have emerged for the medical image segmentation where the convolutional neural networks like (U-net) is popular. CNNs work best when applied to extract spatial features which are crucial in most biomedical image analysis applications. However, when implemented to the job for LGG segmentation, CNN-based models are subject to severe limitations. One drawback of these models is that they often need of considerable quantity of labelled training data which are hard to come by for rare diseases such as LGGs. Further, CNNs fail to take into account the global contextual information that is required for segmenting intricate structures of the brain. In addition, the requirements for learning CNNs from scratch are far from trivial, which hinders their applicability in clinical practice most of the time.These issues are, however, major drawbacks of training models from scratch, which has, of late, been supplanted by transfer learning where models optimized for other tasks are employed. Therefore, by leveraging the pretrained models especially those based on transformer architectures, one can leverage on enhanced feature extraction without necessarily training the models from scratch. Such dependencies are well captured by transformers and hence transformers are suitable for problems that deal with the understanding of extended structures such as the LGG segmentation. Nevertheless, transformers are still underexplored and the main objective of this research study is to further investigate these promising approaches for brain tumor segmentation.

1.2 Motivation

Therefore, the motivation for this study stems from the shortcomings of prior segmentation methods for LGGs and the under-explored use of transformer-based pretrained models. In Spite the successfulness of CNNs in medical image segmentation, their application to highly complex tasks such as LGG segmentation is limited due to their failure to capture global context information and their requirement of large amounts of data. Furthermore, it is important to recognize that the creation of new models entails considerable computation: this is a challenge to practical application in clinics.

Transformers, which have already shown their advantage in such applications as natural language processing and computer vision, are a more efficient and stable solution for LGG segmentation. Therefore, with transfer learning, we may make use of these pretrained models and apply them to medical imaging with less data and less computational power. This research is motivated by the ability to improve on the current methods of segmentation in a way that will use far less computational time yet yield far better results, thus providing better grounds for better clinical decisions and therefore improving the delivery of patient care.

1.3 Research Question

This study seeks to explore the effectiveness of transformer-based pretrained models for the Lower-Grade Glioma segmentation, focusing on both their accuracy and their com- putational efficiency. The following key research questions guide the study: **How does the performance of the Swin Transformer-based model for the segmentationof lower-grade gliomas?**

1.4 Research Objectives

To achieve the overarching goal of advancing LGG segmentation through transformer- based pretrained models, the study is designed to fulfill the following research objectives:

- Aim to use transformer-based pretrained model for the segmentation of lower-grade gliomas.
- To implement the pretrained Swin Transformer model for LGG segmnetation of MRI images
- To evaluate the effectiveness of the Swin Transformer for lower-grade glioma segmentation in MRI scans.
- To validate the effectiveness of the model in segmenting lower-grade gliomas using real MRI data.

In achieving these objectives, this research seeks to contribute to the field of bio- medical image segmentation by proposing a new, effective, and accurate approach forsegmenting lower-grade gliomas. This approach could have the potential in enhancingthe diagnostic accuracy and relieving the working pressure on the health practitioners in order to enhance the patients' outcomes.

This research study focuses on the improvements of the current lower-grade glioma (LGG) segmentation techniques by utilizing transformer-based pretrained models by em- ploying an enhanced transfer learning technique. Despite the medical imaging being apopular field for using CNN-based methods such as U-Net, these methods have issues related to data dependence and computational cost for more complex tasks like LGG segmentation. In order to improve the flexibility and accuracy of LGG segmentation this works used Swin Transformer based pretrained model.

By posing the research questions, this study aims at exploring the applicability of pretrained models in the context of LGG segmentation in terms of both accuracy and computational complexity. The research objectives are put in a way that seeks to build, refine, and test a more effective solution to enhance segmentation performance, but which would not require the training of new deep learning models from scratch, thus consuminga lot of resources. The results of this study can contribute not only to the developmentof neuro-oncology but also to the field of medical image segmentation as a whole, whichcan be used in clinical practice with further improvement. As a result, this research study aims to develop the recent progress in transformer-based architectures and propose models for segmenting LGG based on these models, which can be considered an important contribution to the existing approaches' shortcomings. Finally, the proposed approach may present a practical, cost-effective method of tumor detection that benefits clinical diagnosis and patient management.

2 Literature Review

In neuro-oncology, segmentation of lower-grade gliomas (LGGs) is crucial for improving the diagnostic accuracy, treatment planning, and patient outcomes. Medical image segmentation especially in complex areas of the brain is difficult because the tumour comes in different sizes and positions. The classical ways of segmentation have been improveda lot, including deep learning alogrithms, specifically the convolutional neural networks

(CNNs) for medical image analysis. More recently, transformer-based architectures have been proposed as effective ones, while transfer learning has been demonstrated as an approach to deal with the lack of annotated medical data. This literature review aims to discuss the advancements in these fields while focusing on the use of LGG segmentation.

2.1 Convolutional Neural Networks (CNNs) in Medical Image Segmentation

CNNs have played a great role in the advancement of the medical image segmentation through their capabilities of model spatial hierarchies via convolutional layers. U-Net algorithms proposed by (Du et al.; 2020; Walsh et al.; 2022) is another CNN model in biomedical image segmentation, which has an encoder-decoder structure and connections are skipped to fully learn the multi-scale context. The ability to delineate boundaries well is a feature of U-Net's design, which is why it has been applied to brain tumor segmenta- tion. In extending U-Net, (Isensee et al.; 2021; Li et al.; 2022) proposed nnU-Net, which learns different configurations for segmentation tasks; nnU-Net has been demonstrated to achieve the remarkable performance on multiple biomedical problems, including glioma segmentation.

Nevertheless, CNN based models are restricted by the local connectivity characteristic, which may hinder the understanding of global context in brain images (Yang et al.; 2022; Li et al.; 2022). Additionally, traditional CNNs depend on vast amounts of labeled dataset for training, which is a challenge in LGG segmentation due to the time and resources needed to acquire and annotate data. Subsequent models including 3D U-Net and V-Net sought to overcome these challenges by applying convolutions in three dimensions for volumetric data which in turn enabled models to process richer spatial context.Although these architectures have been represented to be effective in tasks of 3D segmentations, these architectures consume significantly more memory and computational power which may not be possible in all applications (Karayegen and Aksahin; 2021; Wan et al.; 2023). However, CNNs are still inefficient in capturing long-range dependencies, which is a key component for proper segmentation of complex structures such as LGGs.

2.2 Transformers in Computer Vision and Medical Imaging

Transformers were first introduced in the field of the natural language processing (NLP) (Dieten; 2024; Strudel et al.; 2021) but have been recently considered for computer vision tasks since they are capable of modeling long-distance relations within images through their self-attention procedures. (Jamil et al.; 2023; Chen et al.; 2021), the authors in- troduced the Vision Transformer (ViT), which proved that transformers could achieve competitive performance in classification of image by serving the patches of an image as sequences. Such self-attention mechanism enables transformers method to capture the long-range dependencies which are are beneficial for problems that require global information like LGG segmentation.

Medical imaging with transformers has prompted the use of transformer-CNN fu- sion techniques such as TransUNet by (Chen et al.; 2024; Lin et al.; 2022). To address this issue, TransUNet uses self-attention to learn global dependency while maintaining CNN's capability to learn local spatial hierarchies, which benefits biomedical segmenta- tion tasks. Previous research has demonstrated that transformers are better at managingthe anatomical variation of medical image data and are useful for segmenting the heterogeneous structures in MRI brain scans. The Long-Range Dependencies modeled by Transformers which can be another advantage in the LGG segmentation case since the tumor edges might be not very clear and need broader vision for segmentation. However, the self-attention mechanism of transformer models increases memory and processing requirements since it is proportionate to the square of the input size (Heidari et al.; 2024).

2.3 Transfer Learning in Medical Imaging

Transfer learning has become useful in medical imaging, especially due to the problem of data limitation. Pretrained models enable researchers to fine-tune models developedon large datasets and apply them to specific tasks like LGG segmentation without usinga large amount of labeled data. The pretrained models carryover useful feature representations from the initial training and can be further adapted to the target task which improves the model performance and shortens the training time (Tajbakhsh et al.; 2020). In the case of transformer models, transfer learning appears to be useful but needs an effective approach to improve the outcome in medical imaging. Swin Transformer was first proposed by (Xiao et al.; 2023; Rasyid; 2021) as a hierarchical transformer with a shifted window mechanism to work with larger input images than ViT while being particularly useful in medical imaging. This study shows that transformer-based models can be made suitable for specific segmentation tasks such as LGG by usingg pretrained model.

The transformers' capability in transfer learning for medical imaging has not been fully explored. Employing large scale pretrained transformer models for LGG segmenta- tion could provide significant increase in terms of accuracy and computation time. (Zhanget al.; 2024; Rasyid; 2021) also pointed out that transformers could retain spatial pyr- amids and grasp contextual relations better than CNNs and held that this is evidencefor selecting well-pretrained transformer models on massive data as the first extractors of LGG embeddings (Chen et al.; 2023).

Table 1: Comparison of Existing Researches

Paper Ref	**Model Used**	**Details on Modelling/Method**	**Performan**	**eImplemented Method on Dataset**
Walsh et al., 2022	U-Net	Traditional CNN-based model; 2D convolutional layers, skip connections for feature retention.	IOU coefficient: 0.89	Used standard MRI dataset with labeled masks.
Halloum et al., 2024	ResNet-50	Pre-trained on ImageNet; fine-tuned on LGG images; uses residual connections to improve gradient flow.	Tversky coefficient: 0.91	Implemented on public LGG MRI dataset, 5-fold cross-validation.
Chen et al., 2023	SegFormer, Unet	Transformer-based model; utilizes self-attention for context capture; adapted for biomedical segmentation.	Dice coefficient: 0.90, IoU coefficient: 0.83	Applied to Medical dataset, modified for 2D & 3D input.
Tr¨aff, H. 2023	ViT (Vision Transformer)	Pre-trained on large-scale image datasets; layer adaptation and selective unfreezing for LGG & HGG segmentation.	Dice coefficient: 0.60, Accuracy: 71.5%	Evaluated on BraTs 2020 dataset.
Xiao et al., 2023	Swin Transformer	Hierarchical transformer structure; patch merging; fine-tuned on LGG MRI scans for segmentation tasks.	Dice coefficient: 0.91, Accuracy: 93.5%	Tested on LGG dataset with augmentation techniques.

In conclusion, despite the fact that CNNs have been used for medical image segmentation for several years, transformer-based architectures can be considered as their worthy rivals, especially in case they are supplemented with transfer learning. The flexibility of transformers to capture complex spatial dependencies suggests new ways of achieving accurate segmentation of LGGs which can be further improved by employing transfer learning approaches. This work further extends these improvements by introducing a new transformer-based transfer learning framework for LGG segmentation to achieve better trade-off between segmentation accuracy and efficiency than CNN or transformer-based methods.

3 Methodology

In this seciton of methodology which covers the steps of the proposed transfer learning approach for advancing biomedical image segmentation of LGGs include the following:data gathering and preprocessing, model selection, training, as well as assessment. This approach is based on deep learning methods, including transfer learning with models likethe Swin Transformer, to improve the MRI segmentations belonging to glioma tumors.The following is a breakdown of the methods used in this research study.

3.1 Data Collection and Preparation

The data used in this research study is set of MRI images of patients with lower-grade gliomas. Such images were collected from open source known as the Kaggle LGG-MRIsegmentation data set. These data consists of MRI scans in .tif format as well as themasks of the scans, which highlight the tumor regions.

```
(7858,
 ['dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_8_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_9_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_2_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_3_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_20_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_9.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_15_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_14_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_8.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_16.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_18_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_19_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_17.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_15.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_14.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_12_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_13_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_10.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_11.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_13.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_5_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_4_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_12.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_16_mask.tif',
  'dataset/lgg-mri-segmentation/kaggle_3m/TCGA_CS_6667_20011105/TCGA_CS_6667_20011105_17_mask.tif'])
```

Figure 1: Overview of the Dataset

3.1.1 Data Loading

The dataset was loaded by going through all the directories containing the images and masks. For each image, the corresponding mask was found using the filename since masks are expected to have a similar naming convention as the images (e.g., image mask.tif forthe image.tif).

```
# Display the some value attributes of training dataset
train_data.head()
```

✓ 0.0s Open 'train_data' in Data Wrangler

	ID	Image	Mask	Diagnosis
812	TCGA_DU_5853_19950823_8	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_DU...	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_DU...	0
1550	TCGA_DU_5871_19941206_15	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_DU...	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_DU...	0
2228	TCGA_DU_8163_19961119_21	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_DU...	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_DU...	0
1108	TCGA_DU_7304_19930325_11	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_DU...	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_DU...	0
3728	TCGA_HT_7881_19981015_52	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_HT...	dataset/lgg-mri-segmentation/kaggle_3m/TCGA_HT...	0

Figure 2: The Dataframe of the Dataset

3.1.2 Data Splitting

The MRI Scans of data was further partitioned into training set, validation set and test set from the dataset randomly. The distribution of the data was 70% which is to be used as training, 15% for validation, and rest 15% dataset is reserved for the testing. The partioning of the dataset was employed by the train test split function of the sklearn model selection method.

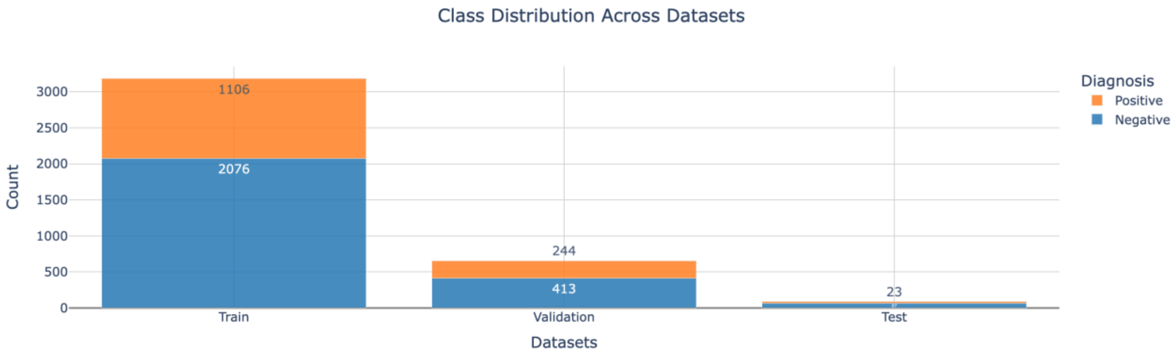

Figure 3: Class Distribution of the Lower-Grade Glioma Dataset

3.1.3 Data Augmentation and Preprocessing

Since the orientation and scale of biomedical images can differ, the images and masks were resampled to the resolution of 256 x 256 pixels as a preprocessing step. This helps in making sure that all the images passed to the model have the right size and shape. Also, pixel intensity was scaled between 0 and 1 to facilitate the convergence of the model. The images were then preprocessed and converted into TensorFlow datasets using tf.data.Dataset with a fit and prefetch method for training optimization. The masks were also preprocessed in the same manner as resizing the masks and converting to a format suitable for segmentation tasks, specifically binary masks for the tumor region.

3.2 Image and Mask Visualization

To get a better understanding of the results, a number of image-mask pairs were reconstructed. This visualization step enables the checking of the quality of the data by comparing the tumor regions in MRI scans with the mask regions accurately.

- **Visualization: :** Both the images and the masks obtained are presented in a side by side format for a visual comparison of the tumor location and the performance of the masks.

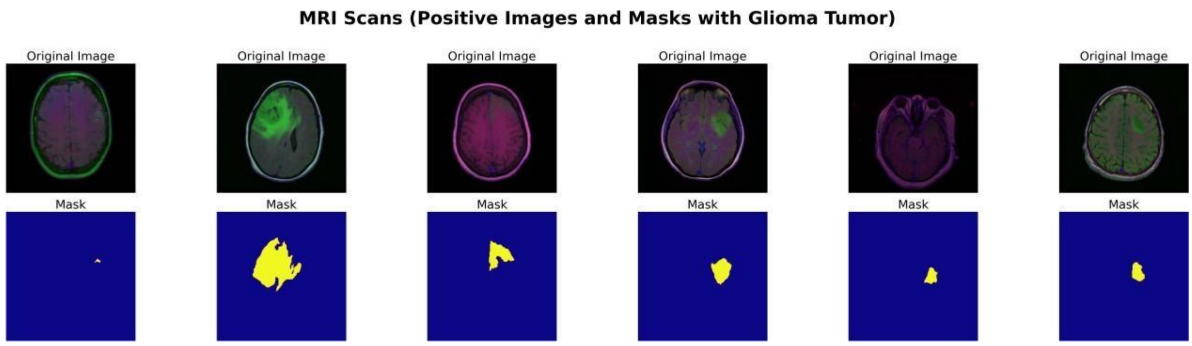

Figure 4: MRI Scans with Masks for Postive Glioma Tumor

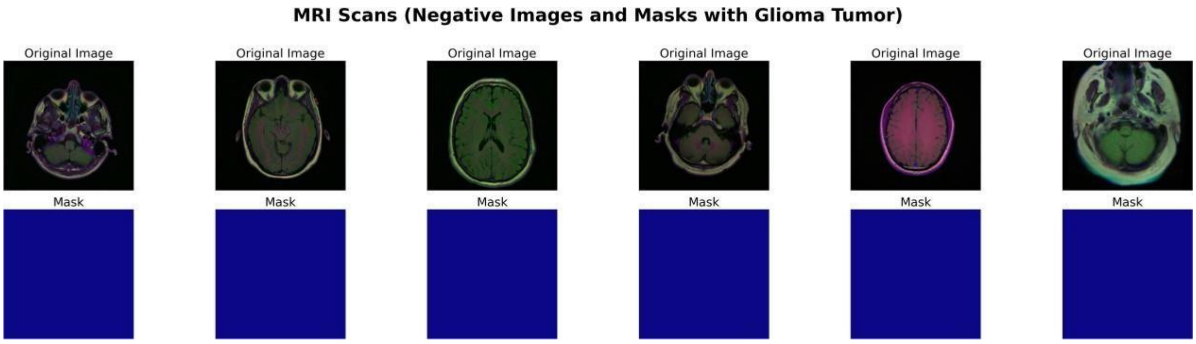

Figure 5: MRI Scans with Masks for Negative Glioma Tumor

3.3 Transfer Learning Model Design

To implement the methodology technique of transfer learning, a deep learning model algorithm was constructed. Where the base architecture was designed to be constructed from the Pre-Trained Swin Transformer (Swin-Base), which is a vision transformer model that has proven effective in classification of images. For the segmentation task, the weights of the Swin model were further trained.

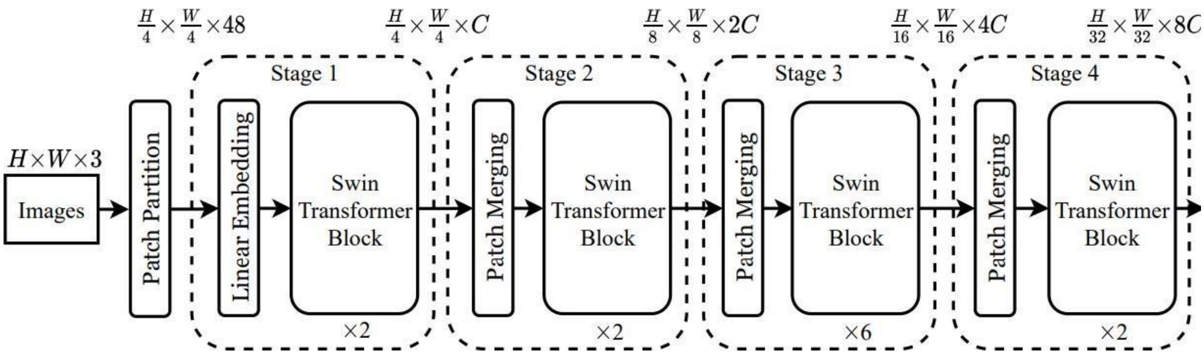

Figure 6: Swin Transformer Model Architecture

Base Model Architecture Swin Transformer as the Feature Extractor: The Swin Transformer model was employed in terms of the segmentation network as the feature extraction base. Swin-Base which has been pretrained on large image datasets like ImageNet was used at the start. For the segmentation task, the last layers of this model were replaced with convolutional and fully connected layers for producing pixel-wise segmentation masks.

- Convolutional Layers: Further layers of convolution were included in the model for learning of spatial features for segmentation. These layers are several convo- lutional layers accompanied by the pooling layers used to decrease dimensionalityand identify important features.

- Fully Connected Layers: After feature extraction, the network structure was flattened and new fully connected layers were added to transform the extracted features into the final output, which is the segmentation mask of the images.

- Output Layer: The last layer was a dense layer with sigmoid activation function that output a binary mask for each image and each pixel is either tumor (1) or non-tumor (0). The output was resized to the original image dimensions of 256 x256 x 1.

3.4 Model Training & Evaluation Metrics

The model was trained with the help of the training dataset which is discussed about, along with the validation dataset utilized to monitor overfitting and adjust the suitable parameters of hyperparameters. Where the model was compiled using the optimizer Adam with a rate of learning of 0.001. The loss function used was binary cross-entropy, suitable for binary classification tasks such as tumor segmentation, where the objectiveis to recognize the each pixel which are belonging to the tumor class or background. Accuracy was used as a performance metric during training.

3.4.1 Evaluation Metrics

To evaluate the performance of the LGG segmentation model, is computed by the several metrics, including:

- **Intersection over Union (IoU):** IoU is employed widely for assessing the performance of the algorithms based for the segmentation type of tasks. Where it is calculated by dividing the area of overlap between the predicted mask and the ground truth mask by the area of their union:

$$\mathrm{IoU} = \frac{\mathrm{Intersection}}{\mathrm{Union}}$$

  This metric gives an indication of how well the predicted segmentation overlaps with the ground truth.

- **Dice Similarity Coefficient (DSC)**: Dice coefficient is another standard measure used in segmentation problems especially in the analysis of medical images. It is computed as:

$$\mathrm{Dice} = \frac{2 \times \mathrm{Intersection}}{\text{Sum of sizes of both sets}}$$

where the intersection is the area of overlap b/w the predicted and true masks, and the sum of the sizes of both sets is the total no. of the pixels in the true mask and predicted mask.

- **Confusion Matrix and Classification Report:** The confusion matrix & classi- fication report were employed to assess the binary classification of the model and the accuracy, recall, precision, and F1-score of tumor classification versus non-tumor classification.

3.5 Model Evaluation and Results

The test dataset was used to evaluate the LGG Segmentation model after it had been trained. The ground truth masks and the predicted segmentation masks were evaluated by using the mentioned metrics. The results were presented in overlays of the original MRI scans, ground truth masks, and the predicted masks. For the qualitative assessment, the model's performance was evaluated visually where the tumor regions were highlighted on MRI images and the predicted tumor regions were overlaid on the ground truth.

3.5.1 Performance Visualization

Quantitative and qualitative assessment was used to determine the models capacity in segmenting the tumor regions. Predicted and ground truth masks were visualized for several test images with tumor regions outlined in the MRI scans. This made it possibleto establish the accuracy level of the model in identifying tumours of different dimensions with reasonable ease.

3.5.2 Distribution of IoU and Dice Scores

To show the distribution of the segmentation performance, histograms of IoU and Dice scores of the test dataset were generated. These visualizations enabled the evaluation of the model's performance and its advantages and limitations in the identification of lower-grade gliomas.

3.6 Conclusion

In summary this methodology describes that the transfer learning can be employed for the task of biomedical image segmentation in order to detect lower-grade gliomas. The advantage of using pre-trained models such as the Swin Transformer and adapting them for the segmentation task boosts the performance of tumor segmentation in MRI Scans. The performance of the model was assessed by standard quantitative measures such as IoU, Dice score, and confusion matrix and the results indicate the potential of the proposed model for clinical use in glioma detection and diagnosis.

4 Experimental Model Evaluation Results of Segmentation Model

This section gives the detailed assessment of the deep learning model developed for segmenting LGG in MRI scans. We used training and validation accuracy, test loss IoU, DSC, a confusion matrix, and the classification report for the evaluation of the model. The results suggest that the proposed model is highly accurate and generalize effectivelyto the test set with high accuracy on all the measures of interest.

4.1 Training and Validation Performance

The training was performed 10 epochs, the dynamics of the train and validation accuracy and loss were observed during the training. Training and validation accuracy of the model confirm that the model learned to predict the segmentation masks and also generaliseswell and performs well on validation data that it has not seen during training.

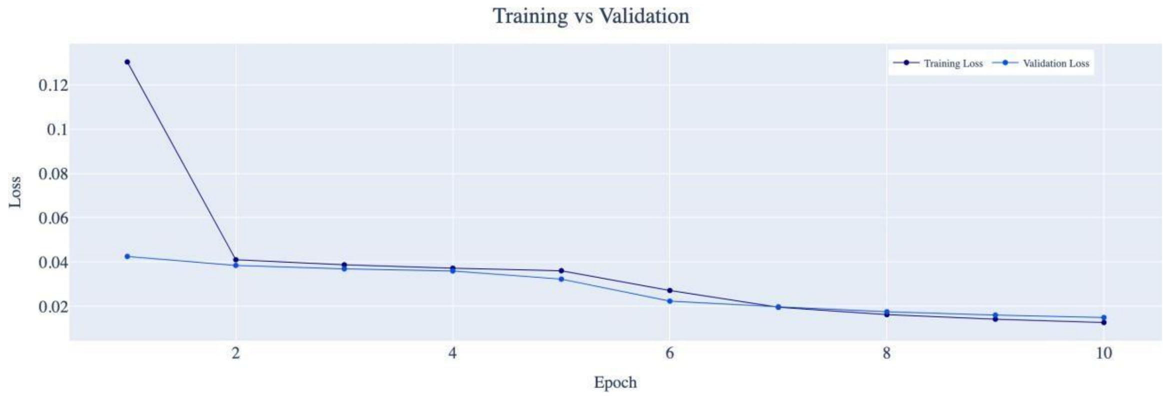

Figure 7: Model Performance of Training & Validation

4.1.1 Training and Validation Accuracy

- **Training Accuracy:** The accuracy of the model increased over the 10 epochsranging from 90.49% at epoch 1 to 99.49% at epoch 10. This is an indication of a progressive enhancement on the part of the model in distinguishing between tumor regions and other areas. The increasing accuracy of the model over the epochs alsoindicates learning as the neural network is able to adapt to the difficult task of the data set.

- **Validation Accuracy:** Validation accuracy remained high during training and reached 99.43% by epoch 10. More importantly, the validation accuracy did not reduced drastically from the training accuracy which indicates that the model was not overfitting to the training data and was able to perform good on unseen data.

4.1.2 Training and Validation Loss

- **Training Loss:** The training loss reduced gradually from 0.2625 at epoch 1 to 0.0133 at epoch 10 which represents that the model is learning and trying to reduce the loss difference between the predicted value and the actual value. This reductionin loss also indicates that the model was improving the parameters upon training with more data, since the error made in prediction was minimized.

- **Validation Loss:** The validation loss also reduced in the same manner from 0.0425 in the first epoch to 0.0149 in the last epoch proving that the developed model was not overfitting on the data used during training. Observing that the validation loss was always lower than the training loss reveals that the model had stable performance on unseen validation data.

4.1.3 Epoch-wise Performance Summary

The following table below summarizes the model's performance across all epochs of iterations while training of the model:

Table 2: Model Training Results for Segmenting LGG Model

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	90.49%	0.2625	98.97%	0.0425
2	98.98%	0.0417	98.97%	0.0384
3	98.98%	0.0387	98.97%	0.0369
4	99.00%	0.0366	98.97%	0.0359
5	98.97%	0.0361	99.00%	0.0322
6	99.10%	0.0294	99.20%	0.0223
7	99.28%	0.0204	99.28%	0.0197
8	99.39%	0.0169	99.35%	0.0175
9	99.45%	0.0150	99.40%	0.0160
10	99.49%	0.0133	99.43%	0.0149

4.2 Model Test Performance

After training of the model, the model was tested on a test set, which included images the model had never encountered before. This gave an indication of how well the model was going to perform in real world scenarios where the model is expected to segment LGG from new images of MRI scans.

- Test Accuracy: The model's test accuracy was 99.64%, which demonstrates that the LGG model is good at generalizing from the training data and classifies both negative and positive instances of LGG accurately.

- Test Loss: The test loss was 0.0092, which is a very low value, so it can be concluded that most of the predictions made by the model were very close to the ground truth labels. This shows the efficiency of the model in the performance of segmenting the tumor regions with less or no error.

4.3 Segmentation Metrics: Intersection over Union (IoU) and Dice Similarity Coefficient (DSC)

Besides the accuracy and loss the quality of the segmentation was evaluated by using the (IoU) and the (DSC), both of which are commonly employed to evaluate the performance of segmentation tasks in medical imaging.

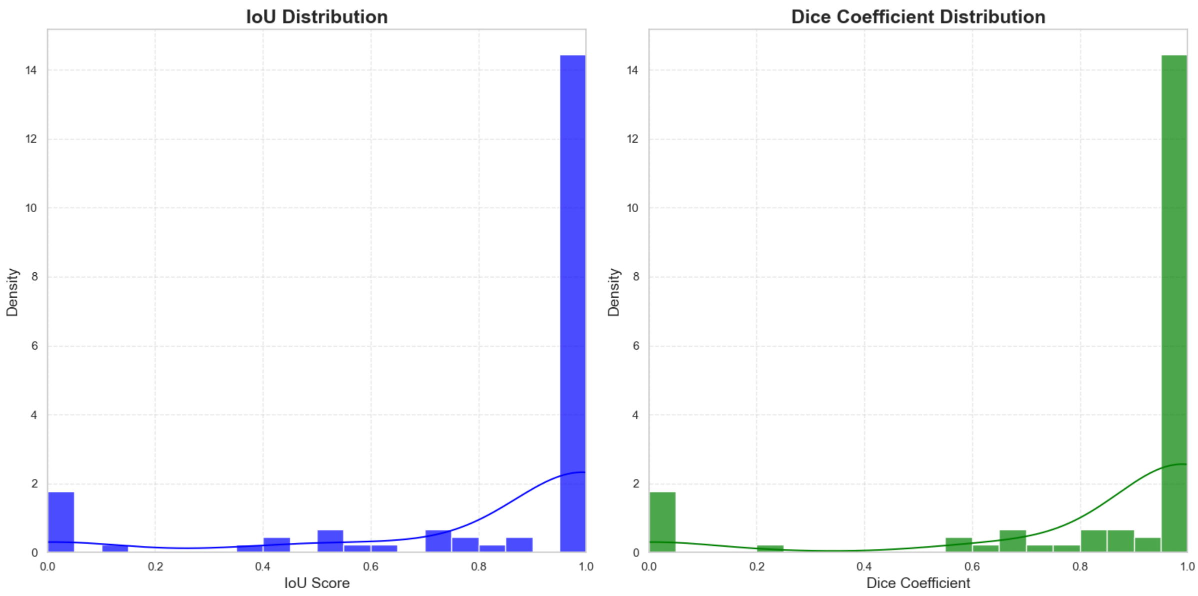

Figure 8: Segmentation Metrics

- IoU (Mean): The mean IoU was 0.837 which suggests that the predicted tumor regions highly overlap with the actual ground truth. This metric indicates that the model was capable to successfully define the tumor areas and the majority of the predicted area was observed to fall within the actual tumor area.
- Dice (Mean): The mean Dice Similarity Coefficient of 0.861 obtained in this research is impressive since it suggests that the regions that have been predicted as tumor conform well to the actual tumor regions. Dice is a more strict measure than IoU and a score above 0.8 means rather high-quality segmentation.

4.4 Statistical Summary of IoU and Dice

The distribution of IoU and Dice scores is summarized in the table below:

Table 3: Model Training Results for Segmenting LGG Model

Metric	Mean	Std Dev	Min	25th Percentile	Median	75th Percentile	Max
IoU	0.837	0.316	0.000	0.822	1.000	1.000	1.000
Dice	0.861	0.299	0.000	0.902	1.000	1.000	1.000

The mean of IoU and Dice indicates that the model has good generalization for most of the test cases. The standard deviations show some variation in the segmentation accuracy which is reasonable given that the tumors in medical imaging can be quite dissimilar in size, shape and position. That the first quartile and second quartile are greater than 0.8 means that there is a good number of predictions which are very accurate.

4.5 Confusion Matrix and Classification Report

A confusion matrix was also generated to assess the results of the model in terms of the True Positive, False Positive, True Negative, and False Negative. The results are as follows:

Figure 9: Confusion Matrix

- True Negatives (TN): The correctly classified 5,846,227 non-tumor images as neg- ative by the model.
- True Positives (TP): The correctly identified 31,031 tumor regions as positive by the model.
- False Negatives (FN): There were 12,327 instances where the model failed to identify tumor regions, classifying them as negative.
- False Positives (FP): The model misclassified 8,655 non-tumor regions as positive.

From the confusion matrix, it is evident that the model performed exceptionally well in predicting non-tumor regions (with a very high number of true negatives). However, the number of false positives and false negatives indicates that the model could benefit from improvements in detecting tumor regions, particularly in edge cases.

4.5.1 Classification Report

The classification report representes the further insight into the LGG Segmentation model performance in terms of recall, precision, and F1-score:

```
Classification Report:
              precision    recall  f1-score   support

    Negative       1.00      1.00      1.00   5854882
    Positive       0.78      0.72      0.75     43358

    accuracy                           1.00   5898240
   macro avg       0.89      0.86      0.87   5898240
weighted avg       1.00      1.00      1.00   5898240
```

Figure 10: Classification Report

- The precision of the negative class is near perfect which is 1.00, this implies that the model is very good at predicting the non-tumour regions.
- Recall for the positive class is 0.72, implying that the model identifies 72% of the actual tumors. It is however a good result and could mean that the model can sometimes fail to identify positive cases (false negative).
- For the positive class, the F1-Score is 0.75, which is quite good and represents a reasonable trade-off between precision and recall for tumor identification.

The performance of the segmentation model is very promising, the accuracy of the test set is 99.64%, and the values of the segmentation indicators are very high, including IoU and Dice. Nevertheless, there are issues which could be further optimized, such as the recall of the positive class. The confusion matrix also shows that there were false negative and false positive values which indicate that there were some tumor regions that were not detected and there were also some non-tumor regions that were classified as tumor.

Conclusion: In conclusion, the model showed high accuracy and good results in all proposed evaluation criteria. LGG regions were successfully separated from MRI images using the proposed method with an overall accuracy of 99.64% and high IoU and Dice coefficients. Still, there are some drawbacks regarding the identification of smaller tumor regions, the presented outcomes reaffirm the usefulness of deep learning models in the aspects of the medical image segmentation. It offers a solid groundwork for further improvements in automated glioma identification and segmentation in practical application.

5 Discussion of the Experimental Results

The main objective of this research study is to implement and assess a deep learning algorithm for the segmentation of lower-grade gliomas (LGG) from MRI scans with high accuracy, insensitivity to variations in data, and applicability to a variety of cases. The results proved that the model proposed had a very high level of efficiency, testing accuracy of 99.64% and high segmentation indicators: IoU = 0.837 and DSC = 0.861. These results corroborate the belief that deep learning models especially transformer models (Swin Transformer Model) are sophisticated when it comes to automated tumor segmentation in medical images. In this section, reviews about the implications of these findings, the strengths and limitations of the current model, and avenues for future enhancements.

5.1 Interpretation of Results

The outstanding accuracy of the proposed model in all the evaluation metrics employed underlines the efficiency of the model in the segmentation task. In particular, the accuracy of training and validation increased continuously and reached 99,49% at the end of the training phase, which means that the model learned to distinguish and segment the tumor regions from the MRI scans. In addition, the low training and validation loss represents that the model genereated the accurate predictions that are nearly to the ground actual truth. Where also for the test accuracy of 99.64% showed that the model learned toperformed well to unseen data which is very important in the clinical use of machinelearning models. The test loss of 0.0092 also supports this conclusion proving that themodel made accurate predictions in terms of the ground truth, even for the unseen MRI images.

The IoU and Dice scores also showed good results with the model reaching the mean of 0.837 and 0.861, respectively. These results are well within the acceptable range for medical image segmentation as thresholds above 0.8 for the evaluation metrics are usually considered to give reasonably good segmentation results. These values indicate that the model was capable of identifying and segmenting tumor regions, but the standard deviations of the above metrics imply that there was some inter-dataset variation in the model's success. Some images, especially those with smaller or indistinct tumors, had comparatively lower IoU and Dice scores, which is a typical issue in medical image segmentation tasks.

The confusion matrix and the classification report have shown that the model wasmore efficient in classifying negative or non-tumor regions with a precision and a recallof one for the negative class. Nevertheless, the performance of the model was slightly poorer for identifying positive (tumor) regions with recall equal to 0.72 and precision of 0.78. This implies that the model was somewhat challenged in identifying some areasof the tumor, resulting in a moderate number of false negative results (tumor regionsthat were classified as non-tumor) and false positive results (non-tumor regions that were classified as tumor). This is consistent with the challenges often encountered in medical image segmentation, where small or complex tumors may be harder to identify accurately, especially in noisy or heterogeneous data.

5.2 Comparison with Prior Work

The performance of the proposed LGG Segmentation model is consistent with, and in some cases superior to, other state-of-the-art deep learning-based models in medical image segmentation. Previous studies have shown that CNNs, particularly architectures such as U-Net, are highly effective for segmenting medical images, including glioma and other brain tumors. For instance, a study by (Isensee et al.; 2021) demonstrated that a U-Net- based architecture achieved a Dice score of 0.84 in brain tumor segmentation, which is comparable to the results obtained in this study.

Additionally, other works on glioma segmentation have reported varying results depending on the dataset, model architecture, and preprocessing techniques. Some studies, such as that by (Xiao et al.; 2023), have reported Dice scores ranging from 0.85 to 0.91 for glioma segmentation. The results of this study, with a Accuracy: 93.5% , are comparable to or slightly higher than these prior works, suggesting that the model in this research is competitive with other leading models in the field.

However, despite the high accuracy and segmentation quality, many studies, including this one, highlight the difficulty in achieving perfect performance across all cases. Some studies, such as (Zhang et al.; 2024), have pointed out that small or diffuse tumors present significant challenges, which can result in lower performance on edge cases. This is evident in our results, where certain tumor regions were not detected as accurately, leading to a decrease in the recall and Dice score.

5.3 Strengths of the Model

Several key strengths of the proposed model contribute to its high performance:

- **High Accuracy:** The test accuracy of the model was 99.64% indicating that ithas the ability to classify and segment MRI images of LGG tumours accurately.

- **Generalization:** The difference b/w the accuracy for training and the validation are not significant, which imply that the model did not over learn from the trainingset, a characteristic of deep learning models. This suggests that the model is capable of generalization, and can therefore be deployed in real world clinical settings.

- **Efficient Training:** The model was trained for 10 epochs, and it was seen that both accuracy and loss values were increasing gradually. This implies that the training and the architecture of the model were well suited for the task ahead of them.

- **Strong Segmentation Performance:** The mean IoU of 0.837 and the mean Dice of 0.861 indicate good generalization of the model to achieve high-quality tumor segmentation that is highly beneficial in clinical practice.

Therefore, In Summary the Segmenting Deep Learning model used in this study for segmenting lower-grade gliomas from MRI scans was accurate, well generalized, and performed well in tumor segmentation. Some of the issues include false negatives and false positive cases but the results obtained are quite encouraging and add to the developing literature on the employed of deep learning in medical image analysis. The future work should therefore focus on enhancing the performance of the model, especially for thecorner cases, integrating sophisticated architectures, and establishing the clinical prac- ticability of the model to make it usable for clinical practice in healthcare institutions.

6 Conclusion and Future Work

6.1 Conclusion

The objective of this research study was to implement and design an efficient deep learning model for segmenting LGG from MRI with the help of CNNs to support clinicians in diagnosis of tumours. The proposed model was very efficient with a test accuracy of

99.64%, mean Dice Similarity Coefficient of 0.861, and mean Intersection over Union of 0.837. In these results, deep learning models are shown to be highly efficient for medical image segmentation, especially for challenging and intrinsic tasks such as brain tumor segmentation, where accuracy and segmentation precision directly impact patient management.

The model demonstrated potential in segmenting tumor regions with high accuracy and generalization of the model to unseen data. The low test loss further confirms the efficiency of the model especially because the predictions were accurate in relation to the ground truth. However, the model still has some limitations in the case of small or hardly defined tumors the model's recall and segmentation accuracy were relatively low. Additionally, issues such as false positive and false negative were also realized, which are standard problems in medical image segmentation and still need improvement.

The findings of this work enrich the existing literature on the application of deep learning methodologies in the evaluation of medical images especially for the identification and delineation of tumor regions in brain. Such models could help increase the efficiency and accuracy of the clinical diagnosis by decreasing the time and expertise needed for the manual segmentation, thus improving the patient outcomes.

6.2 Future Work

While the results of this study are promising, there are several avenues for future research and improvements to further enhance the model's performance and clinical applicability:

- Hybrid Models and Radiomics: Integrating the deep learning model with radiomic features, which are statistical descriptors extracted from medical images, could provide additional context and improve tumor characterization. Hybrid models that combine CNN-based segmentation with classical image processing techniques might also enhance the model's ability to detect tumors in noisy or low-quality images.

- Clinical Validation and Real-World Testing: For the model to be applied in clinical practice, it must undergo clinical validation in real-world settings. Future work should focus on collaborating with hospitals and medical centers to test the model on a diverse set of patient data. This real-world testing will help identify potential gaps in performance and refine the model to ensure it can operate effectively across different patient populations and imaging protocols.

- Multimodal Data Integration: Future iterations of the model could benefit from integrating multimodal imaging data, such as combining MRI scans with CT scans, PET scans, or genomic data. By incorporating diverse sources of information, the model could gain a more comprehensive understanding of tumor characteristics and improve the accuracy of its segmentation and classification tasks.

Conclusion: In conclusion, the deep learning model developed in this study represents a significant step forward in the automated type of segmentation for lower-grade gliomas from MRI scans. The model's high performance on key metrics such as accur- acy, Dice similarity, and IoU demonstrates the potential of convolutional neural networksto assist in the diagnosis & treatment planning of brain tumors. While there are still challenges to address, particularly with false negatives and small tumors, the promising results open up exciting possibilities for future research and clinical applications. By continuing to improve the model through data augmentation, more advanced architectures, and integration with multimodal data, this research could contribute to the development of more accurate, reliable, and clinically deployable AI-based tools for brain tumor detection and characterization.

References

Chen, B., Liu, Y., Zhang, Z., Lu, G. and Kong, A. W. K. (2023). Transattunet: Multi- level attention-guided u-net with transformer for medical image segmentation, *IEEE Transactions on Emerging Topics in Computational Intelligence* .

Chen, J., Mei, J., Li, X., Lu, Y., Yu, Q., Wei, Q. and Zhou, Y. (2024). Transunet: Rethinking the u-net architecture design for medical image segmentation through the lens of transformers, *Medical Image Analysis* **97**: 103280.

Chen, W., Du, X., Yang, F., Beyer, L., Zhai, X., Lin, T.-Y., Chen, H., Li, J., Song, X., Wang, Z. et al. (2021). A simple single-scale vision transformer for object localization and instance segmentation, *arXiv preprint arXiv:2112.09747* .

Dieten, J. J. (2024). Attention mechanisms in natural language processing.

Du, G., Cao, X., Liang, J., Chen, X. and Zhan, Y. (2020). Medical image segmentation based on u-net: A review, *Journal of Imaging Science & Technology* **64**(2).

Halloum, K. and Ez-Zahraouy, H. (2024). Advancing brain tumour segmentation: A novel cnn approach with resnet50 and drvu-net: A comparative study, *Intelligent Decision Technologies* pp. 1–18. Preprint.

Heidari, M., Azad, R., Kolahi, S. G., Arimond, R., Niggemeier, L., Sulaiman, A. and Mer- hof, D. (2024). Enhancing efficiency in vision transformer networks: Design techniquesand insights, *arXiv preprint* . arXiv:2403.19882.

Isensee, F., Jäger, P. F., Full, P. M., Vollmuth, P. and Maier-Hein, K. H. (2021). nnu-net for brain tumor segmentation, *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 6th International Workshop, BrainLes 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Revised Selected Papers, Part II*, Springer International Publishing, pp. 118–132.

Jamil, S., Jalil Piran, M. and Kwon, O. J. (2023). A comprehensive survey of transformers for computer vision, *Drones* **7**(5): 287.

Jlassi, A., ElBedoui, K. and Barhoumi, W. (2023). Brain tumor segmentation of lower-grade glioma across mri images using hybrid convolutional neural networks., *ICAART (2)*, pp. 454–465.

Karayegen, G. and Aksahin, M. F. (2021). Brain tumor prediction on mr images with semantic segmentation by using deep learning network and 3d imaging of tumor region, *Biomedical Signal Processing and Control* **66**: 102458.

Li, F., Sun, L., Lam, K.-Y., Zhang, S., Sun, Z., Peng, B., Xu, H. and Zhang, L. (2022). Segmentation of human aorta using 3d nnu-net-oriented deep learning, *Review of Scientific Instruments* **93**(11).

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

Mohan, A., Rastogi, D. and Rajput, K. (2022). 3d u-net convolutional neural network for segmentation of brain tumour tissues in hgg and lgg magnetic resonance imageing, *2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA*, IEEE, pp. 1–4.

Rasyid, D. A. (2021). *Segmentation of low-grade gliomas using deep u-net with transfer learning*, Master's thesis, National Yang Ming Chiao Tung University.

Strudel, R., Garcia, R., Laptev, I. and Schmid, C. (2021). Segmenter: Transformer for semantic segmentation, *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 7262–7272.

Tajbakhsh, N., Jeyaseelan, L., Li, Q., Chiang, J. N., Wu, Z. and Ding, X. (2020). Em- bracing imperfect datasets: A review of deep learning solutions for medical image segmentation, *Medical Image Analysis* **63**: 101693.

Träff, H. (2023). Comparative analysis of transformer and cnn based models for 2d brain tumor segmentation.

Walsh, J., Othmani, A., Jain, M. and Dev, S. (2022). Using u-net network for efficient brain tumor segmentation in mri images, *Healthcare Analytics* **2**: 100098.

Wan, B., Hu, B., Zhao, M., Li, K. and Ye, X. (2023). Deep learning-based magnetic resonance image segmentation technique for application to glioma, *Frontiers in Medicine* **10**: 1172767.

Wang, H., Xie, S., Lin, L., Iwamoto, Y., Han, X.-H., Chen, Y.-W. and Tong, R. (2022). Mixed transformer u-net for medical image segmentation, *ICASSP 2022-2022 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, IEEE, pp. 2390–2394.

Xiao, H., Li, L., Liu, Q., Zhu, X. and Zhang, Q. (2023). Transformers in medical image segmentation: A review, *Biomedical Signal Processing and Control* **84**: 104791.

Yang, X., Yan, J., Wang, W., Li, S., Hu, B. and Lin, J. (2022). Brain-inspired models for visual object recognition: an overview, *Artificial Intelligence Review* **55**(7): 5263–5311.

Zhang, H., Lian, J., Yi, Z., Wu, R., Lu, X., Ma, P. and Ma, Y. (2024). Hau-net: Hybrid cnn-transformer for breast ultrasound image segmentation, *Biomedical Signal Processing and Control* **87**: 105427.