SAMSeg: Modifying SAM's Framework for Intracranial Meningioma Segmentation from Brain MRIs

Tapera Chikumbu
CHKTAP011@myuct.ac.za
University of Cape Town
Cape Town, South Africa

Cassandra Wallace
WLLCAS004@myuct.ac.za
University of Cape Town
Cape Town, South Africa

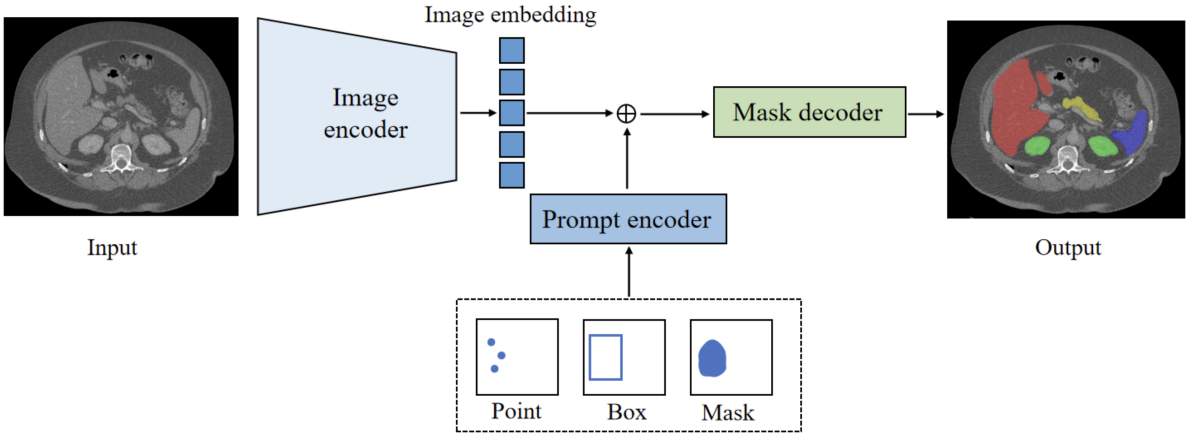

Figure 1: Architecture of SAM

ABSTRACT

The Segment Anything Model (SAM) has been shown to excel at segmenting natural images, but the direct application of SAM to medical images presents challenges due to inherent feature differences in medical data. This project investigates the potential for adapting SAM for automated segmentation of intracranial meningiomas on brain Magnetic Resonance Imaging (MRI) scans. Two approaches are proposed to adapt SAM: (1) integrating complete architectures like U-Net onto SAM, and (2) exploring 3D adaptor layers with or without an additional 3D encoder. This research contributes to the advancement of automated methods for objective and quantitative analysis of meningiomas, potentially contributing to advancements in clinical care for patients with this condition.

KEYWORDS

Computer Vision, Deep Learning, Medical Image Segmentation, Tumour Segmentation, Segment Anything Model

1 INTRODUCTION & PROJECT DESCRIPTION

Image Segmentation, the process of partitioning an image into meaningful regions-of-interest, is crucial in various computer vision applications [24]. Medical Image Segmentation (MIS) allows for accurate diagnosis, treatment planning, and disease monitoring [18]. However, traditional methods used by medical professionals are time-consuming and prone to errors due to subjectivity [9].

Deep Learning offers a promising alternative. The Segment Anything Model (SAM) excels in natural image segmentation (coloured images of everyday objects). Its vast training dataset and novel architecture allow it to achieve high accuracy without requiring specific object-based training, known as zero-shot learning [12]. However, applying SAM directly to medical images presents challenges due to their inherent differences to natural images: 3D structures, lower colour contrast, and higher noise levels [3, 7, 18].

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique that generates detailed images using magnetic fields and radio waves. With its ability to visualise soft tissues [11, 13], MRI is crucial for diagnosing and managing various conditions like tumour growth. A specific type of MRI, multi-parametric MRI (mpMRI), goes beyond basic anatomy by combining images from various MRI sequences to capture different physiological properties of tissues, helping in tumour characterisation [14]. For brain tumours, mpMRI typically includes a series of T1-weighted before contrast, T1-weighted after contrast, T2-weighted and T2-weighted Fluid Attenuated Inversion Recovery (FLAIR), providing a comprehensive picture of the tumour for tasks such as surgical planning and treatment response monitoring [13, 14].

Meningiomas are the most common intracranial tumour in adults. Although most are benign (non-cancerous), aggressive tumours require intensive treatment [19, 20]. Segmentation helps determine the size, location, and potential aggressiveness of meningiomas. Its segmentation is unique

due to often attaching or infiltrating the skull base, requiring the segmentation model to account for complex skull-tumour interfaces [5, 10].

While some research has explored adapting SAM to MIS, as further outlined in Section 2, it often focused on other medical modalities or creating a universal model. In addition, there is a research gap on the specific improvements these adaptations achieve compared to the baseline SAM performance on the same task. Our project aims to address this gap by investigating the potential of adapting SAM for automated intracranial meningioma segmentation from brain MRIs.

We propose two modifications to SAM's framework: (1) adding an additional architecture, such as a Convolutional Neural Network (CNN) [1] like U-Net [23], on top of SAM, and (2) adding 3D adaptor layers to SAM, with an optional 3D encoder.

Both models will be evaluated on the BraTS 2023 Meningioma Challenge dataset, a standardised benchmark with the largest collection of multi-label expert-annotated meningioma mpMRIs to date [13], comparing their performance to that of the original SAM model. This project contributes to advancements in both medical adaptation of SAM and automated meningioma segmentation, which could lead to improved clinical care. We aim to achieve improved segmentation performance compared to the baseline SAM model by leveraging its strengths while addressing its limitations for medical images.

2 RELATED WORK

Kirillov et al. **[12]** proposed the Segment Anything Model (SAM) as a promptable foundation model for natural image segmentation. Experiments show SAM outperforming prominent segmentation models such as U-Net and UC-TransNet on various tasks [26]. This is largely attributed to its use of a novel vision transformer (ViT) architecture. Notwithstanding, several underperformed scenes where the default SAM struggles to accurately generate semantic masks have been identified. These include shadow detection, salient object detection and medical image segmentation [2]. Only after optimising SAM for such scenes can comparable performance be achieved. Some adaptations may not match our use case but still present useful insights and roadblocks.

SAM can serve as a component in a larger segmentation pipeline. Leveraging the strengths of several models increases overall complexity but allows for more accurate segmenting. This is apparent in work by Li et al. [16] and Lin et al. [17]. The combinations of SAM and different neural networks achieve better performance than their constituent parts.

Other SAM adaptations are tailored for 3D. Gu et al. [6] slice 3D MRIs into 2D chunks that SAM can process. In addition, adapter blocks are incorporated into the image encoder and mask decoder. This allows for Parameter Efficient Fine-Tuning (PEFT). Most parameters are kept constant during training with the weights of these adapters being the focus of updates. This approach offers faster convergence times and lower resource use than the alternative of updating all parameters. Other models like Med-SA [25] and AdaptiveSAM [21] have found success using similar PEFT methods.

Standard slicing of 3D volumes sacrifices accuracy for simplicity as information between slices is lost. Quan et al. [22] combat this using a window to store a maximum of three adjacent slices. This window offers more contextual information than possible from segmenting each slice in isolation.

3 PROBLEM STATEMENT

Automated segmentation of brain tumours from MRIs plays a crucial role in clinical tasks, such as treatment planning and surgical intervention [24]. While deep learning approaches like the Segment Anything Model (SAM) have shown promise in image segmentation, directly applying SAM to medical images presents challenges due to inherent feature differences in medical data. SAM is primarily trained on 2D natural images and may not capture the complexities of 3D medical data like brain MRIs, which often have lower colour contrast and higher noise levels. This limitation can lead to suboptimal segmentation performance [3, 7, 18].

This project will investigate the potential of adapting SAM for automated segmentation of intracranial meningiomas in brain MRIs. Specifically, we address the research gap in which the improvement achieved by modifications to SAM compared to its baseline performance is not well explored. We propose two approaches to improve SAM's segmentation accuracy for brain tumour tasks:

(1) Integrating a separate, task-specific network, like a Convolutional Neural Network (CNN) architecture such as U-Net, on top of the pre-trained SAM.
(2) Incorporating 3D adaptor layers within SAM, with the possibility of including a dedicated 3D encoder specifically designed for processing 3D data.

Our problem statement for each approach is as follows:

(1) The baseline SAM model, trained on natural images, might not be optimal for Brain MRI Tumour Segmentation. This project investigates if adding a task-specific network on top of SAM can improve its segmentation performance compared to the baseline.
(2) SAM's architecture is not applicable for processing the 3D nature of Brain MRI data. This project investigates if incorporating 3D adaptor layers and/or a 3D encoder can improve its tumour segmentation performance compared to the baseline.

3.1 Research Questions

3.1.1 ***Framework Modification 1: Adding Another Network on SAM****. Can adding a task-specific network on top of the pre-trained SAM improve its intracranial meningioma segmentation performance on brain MRIs, compared to the baseline SAM performance on the same task?*

We hypothesise that adding a U-Net or similar CNN architecture on top of the pre-trained SAM will significantly improve its intracranial meningioma segmentation accuracy for brain MRI tasks compared to the baseline performance

on the same task, even with the added complexity. This additional network can learn task-specific features from brain MRIs, leading to more accurate segmentation.

3.1.2 ***Framework Modification 2: Adding 3D Adaptor Layers within SAM****. Can adding 3D adaptor layers and/or a 3D encoder within pre-trained SAM improve its intracranial meningioma segmentation performance on brain MRIs, compared to the baseline SAM performance on the same task?*

We hypothesise that a combination of 3D adaptor layers and a 3D encoder specifically designed for 3D data will significantly improve SAM's performance for brain MRI intracranial meningioma segmentation tasks. The 3D encoder can learn more robust feature representations from the MRI volumes, while the adaptor layers can further process these features for accurate segmentation.

3.2 Aims

By answering our hypotheses for our research questions, our project aims to achieve two primary goals:

- Evaluate the effectiveness of our two proposed modifications to the SAM framework for automated segmentation of intracranial meningiomas in brain MRIs.
- Achieve improvement in segmentation performance compared to the baseline SAM model on the BraTS 2023 Meningioma Challenge dataset.

By achieving these objectives, we hope to demonstrate the potential for adapting SAM for improved automated intracranial meningioma segmentation methods, potentially contributing to advancements in clinical care for patients with this condition.

4 PROCEDURES AND METHODS

This section outlines the procedures and methods we will employ to achieve our research objective. Our overall workflow can be seen in Figure 2.

4.1 Data Collection

We will utilise the BraTS 2023 Meningioma Challenge dataset for this research project. This dataset consists of pre-operative and pre-treatment MRIs of patients diagnosed with intracranial meningioma. Ground truth segmentation labels for three tumour compartments are included. These are visualised in Figure 3.

4.2 Data Pre-Processing

The pre-processing operations performed by challenge organisers before making the dataset available are summarised in Figure 4. For our experiments, a significant challenge lies in transforming these scans for suitable input into SAM. MRIs are inherently 3D volumetric data, while SAM is designed for 2D images. There are several approaches to address this challenge, each with its advantages and limitations:

- **Central Slice:** We could extract only the central slice from each MRI scan. However, this approach loses crucial information about the 3D structure of the tumour, potentially impacting segmentation accuracy.
- **Stacking Slices:** We could provide SAM with every slice of the scan as separate inputs. This approach preserves the 3D information but significantly increases training time and computational resources.
- **3D to 2D Conversion Techniques:** We could explore techniques like max pooling or average pooling to convert 3D volumes into 2D representations suitable for SAM. However, these techniques might lead to information loss.

We will determine the most suitable pre-processing steps for SAM input after evaluating the effectiveness of our fine-tuning approaches. Other data augmentation techniques will also be tested during training. These include random flipping, rotation, and scaling. Such steps are meant to improve the generality of the model.

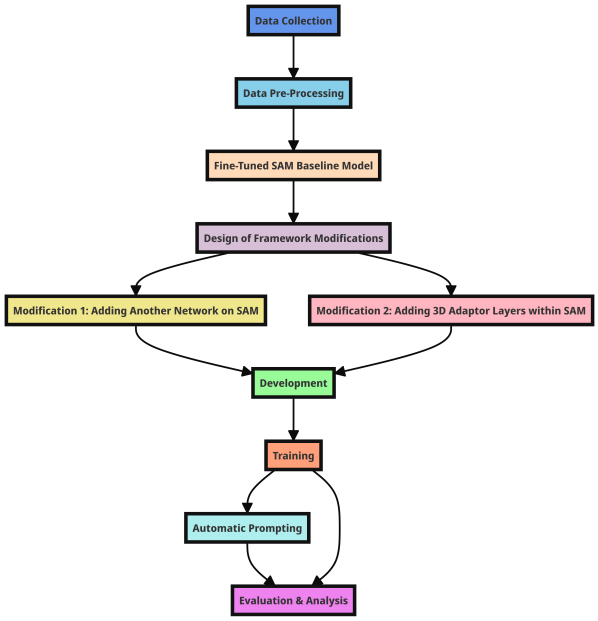

Figure 2: A flowchart that illustrates the overall workflow and key stages of our research project.

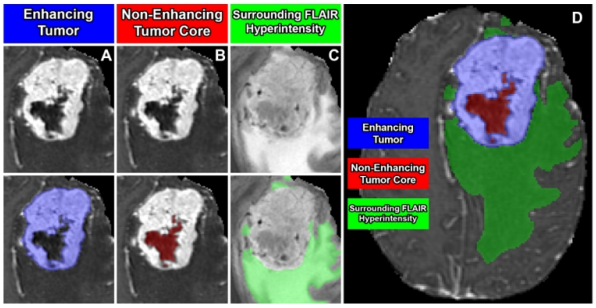

Figure 3: A visual sample from the BraTS 2023 Meningioma Challenge dataset, showing the three types of tumour segmentations [13].

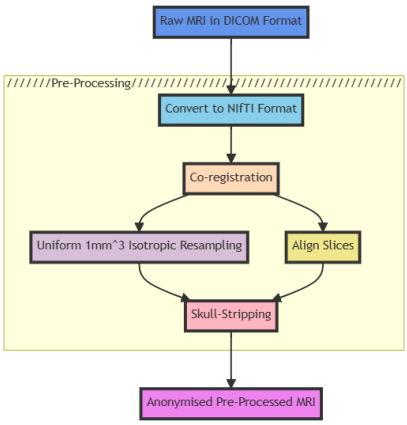

Figure 4: A flowchart of how the BraTS organisers pre-processed the Intracranial Meningioma data for the challenge [13].

4.3 Fine-Tuned SAM Baseline Model

We will establish a baseline model by fine-tuning the pre-trained SAM architecture on the BraTS Meningioma dataset. This involves utilising the pre-trained weights of the SAM model and adapting the final layers to the three-class intracranial meningioma segmentation task. We will focus on employing Parameter-Efficient Fine-Tuning (PEFT) techniques, and appropriate optimisation techniques and loss functions to train the model. This will allow us to evaluate the performance of both the vanilla/raw SAM and the fine-tuned SAM with our modified models.

4.4 Design of our Framework Modifications

This section outlines the two proposed framework modifications to the baseline SAM model.

4.4.1 ***Framework Modification 1: Adding Another Network On SAM.*** This framework modification explores integrating an additional entire network on top of the SAM for intracranial meningioma segmentation in brain MRIs.

The overall justification is that it would leverage the strengths of both models, to combine SAM's powerful feature extraction with the established MIS performance of a specialised architecture, for tailored use for intracranial meningioma segmentation on brain MRIs. In addition, this extra network can potentially learn more complex spatial relationships within the MRIs, leading to improved segmentation accuracy [1, 12, 16, 17].

However, adding a network introduces additional trainable parameters, leading to increased training time and computational demands. Efforts will be made to minimise model size to ensure efficient training while maintaining accuracy. In addition, carefully combining SAM's capabilities with the added network requires a well-defined, optimised strategy. Aspects like information flow between the networks will need to be addressed.

Many different types of networks and architectures could be integrated with SAM. For our project, we propose Convolutional Neural Networks (CNNs). They are a type of Deep

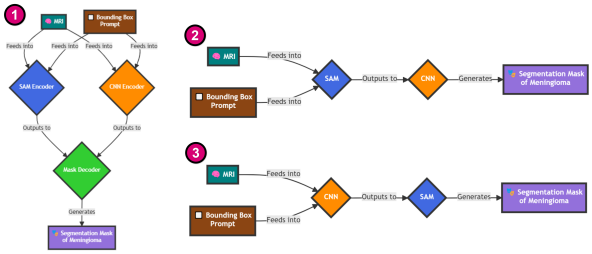

Figure 5: A visual of Framework Modification #1: (1) Sub-approach 1, where a CNN Encoder is used in parallel with SAM's Encoder. (2) Sub-approach 2, with SAM First, CNN Second. (3) Sub-approach 2, with CNN First, SAM Second.

Learning neural architecture specifically designed for image recognition and analysis tasks. They excel in capturing spatial relationships and extracting local and global features from images, making them highly effective in MIS [1]. Two sub-approaches will be considered.

Firstly, a parallel CNN encoder branch will be introduced alongside SAM encoders, as seen in Part 1 of Figure 5. Both encoders would independently process the input MRI. Their extracted features will then be combined, using techniques such as summation or channel-wise concatenation, which will be fed into a decoder network for the final segmentation mask prediction. This approach takes advantage of the strengths of both models while promoting information exchange.

Secondly, a pre-trained CNN architecture [1], such as U-Net [23], will be placed either before or after the SAM encoder. There are two potential workflows for this sub-approach:

- **SAM First, CNN Second:** The SAM performs the initial segmentation, leveraging its powerful feature extraction. The CNN then further refines these segmentations for improved detail and precision (Part 2 in Figure 5).
- **CNN First, SAM Second:** The CNN performs the initial segmentation, providing coarse localisation. The SAM then refines them to improve segmentation accuracy (Part 3 in Figure 5).

The chosen workflow (SAM First or CNN First) will be determined through experimentation.

Note: While some research has integrated additional networks with SAM, it has not been extensively explored in the medical field, particularly for brain MRIs and meningioma segmentation. We will draw inspiration from existing methods (e.g., ClipSAM [15], SAMUS [17]) while acknowledging their applications and challenges.

4.4.2 ***Framework Modification 2: Adding 3D Adaptor Layers Within SAM.*** ViT-B is a pre-trained ViT dominant in lightweight applications of SAM. Tuning all 91 million parameters of the transformer to make a task-specific ViT would be resource-intensive. Adaptation is popular in natural language processing to reduce the number of trainable parameters in a model [8]. Adaptor blocks work by project-

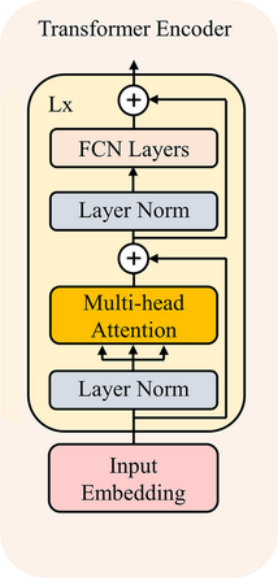

Figure 6: Architecture of a Transformer Encoder [25]

ing model parameters onto a smaller subspace. Most of the parameters are kept constant during training. Instead, those in this smaller subspace are updated in each training cycle. This reduces the computational cost of training very large models.

The encoders within ViT-B are composed of multi-head attention (MHA), fully connected network (FCN) and normalisation layers as seen in Figure 6. Adaptor blocks have been successfully added to the encoders in SAM [6, 25]. In particular, Wu et al. incorporate adaptors to down-sample images while introducing a Space-Depth Transpose (SD-Trans) block. Its architecture is seen in Figure 7. This improves the model performance on 3D data by incorporating volume depth into each input embedding. It is hypothesised

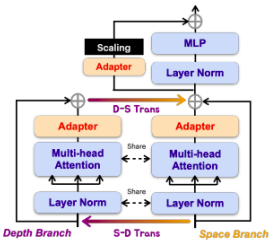

Figure 7: Adaptor placement in SD-Trans [25]

that such an adaptor structure can improve the performance of SAM on the meningioma dataset. To test this, adaptor blocks will be added to SAM in an attempt to achieve the depth retention of Med-SA. The Med-SA source code is available online and will be used to better understand how to place adaptor blocks within the transformer. Using this knowledge, the encoder architecture of Med-SA will be replicated on our SAM baseline through appropriate adaptor insertion.

Assuming successful replication of the SD-Trans, the model's performance on the meningioma dataset will show how adding adaptor blocks affects performance. Alternatively, other insertion points will be investigated for creation of a different model that still utilises adaptor blocks to incorporate 3D information. One such possibility is the use of pre-trained encoders designed for 3D data.

4.5 Development

4.5.1 Implementation Framework. PyTorch, a user-friendly deep learning framework, will be used for the efficient development of our modified SAM architectures. Meticulous code documentation and Git version control ensure project maintainability and collaboration.

4.5.2 Addressing Computational Demands. Training Deep Learning models, especially for MIS tasks, can be very computationally intensive. We anticipate this to be a significant challenge in our project due to the complexity of our models. We will address this with:

- **Code Optimisation & Resource Management:** Techniques like batching and gradient checkpointing will reduce memory usage and training time without sacrificing accuracy. We will closely monitor resource consumption to identify and address bottlenecks.
- **Cloud Computing Resources:** We will explore leveraging free cloud platforms with powerful GPUs to overcome the limitations of local hardware.
- **Staged Training:** A staged training approach will involve initially training a smaller model, gradually increasing complexity as it converges. This validates core functionalities early while minimising computational demands.

If computational limitations become a major obstacle, we might explore more significant changes, such as model compression techniques, which will reduce the size of our models and memory footprint, without sacrificing strategy.

4.5.3 Training Methodology. Both models will be trained on the BraTS 2023 Meningioma Challenge dataset [13]. Our training process will follow a well-defined pipeline with the following key steps:

(1) **Data Splitting:** Rigorous splitting ensures a balanced representation of tumour classes across training, validation, and testing sets.
(2) **Model Initialisation & Transfer Learning:** We will leverage our fine-tuned baseline SAM model as a starting point for both modified architectures (adding another network on top of SAM & adding 3D Adaptor Layers), allowing them to benefit from pre-trained features while learning meningioma-specific details.
(3) **Hyper-Parameter Tuning:** We will employ techniques like grid search or randomised search to optimise hyper-parameters such as learning rate, optimiser settings, and batch size for each model. This tuning will be conducted using the validation set to ensure generalisation performance.

(4) **Training Loop & Monitoring:** A comprehensive training loop will be implemented, monitoring key metrics within each epoch. Early stopping mechanisms will be implemented to stop training once validation performance plateaus to prevent overfitting.

Using these methods strategically, we aim to mitigate computational challenges. We acknowledge that training times may still be significant, and we will closely monitor progress to adjust hyper-parameters or explore alternative strategies if necessary.

4.6 Automatic Prompting Implementation (Optional)

As an extension, we plan to investigate incorporating automatic prompting techniques to further improve segmentation performance. This could involve dynamically generating task-specific prompts that guide the model during the training process. However, this will depend on available time and resources.

4.7 Evaluation & Analysis

Our focus is on segmentation accuracy. Dice loss and Intersection Over Union (IOU) scores will be used to express this accuracy. The Sørensen-Dice coefficient [4] compares the intersection of samples to the sum of the constituent samples and expresses the result as a percentage. In our case, we will be comparing ground-truth masks (A) and the predicted masks (B) for each model. Dice loss is the corresponding difference function, given by 1 - Dice coefficient.

$$\begin{aligned} Dice &= \frac{2|A \cap B|}{|A| + |B|} \\ L_{\text{Dice}} &= 1 - \frac{2|A \cap B|}{|A| + |B|} \end{aligned} \tag{1}$$

These calculations will be performed by a test script once a training cycle is complete. Data for this stage is taken from the testing split of images and masks. As they are excluded from training, this allows us to evaluate model performance on unseen data. Higher Dice scores correspond with better models. We will also express Dice scores on a three-point scale: Bad, Okay, Good. A mask is considered "Bad" if it has a Dice score below 50%. "Okay" scores are between 50% and 80%, while "Good" scores are above 80%. These values were chosen as the majority of challenge submissions have dice in the stated "Okay" range while the highest recorded score on the challenge is 86.6%. Notably, no submissions utilising SAM were identified.

5 ETHICAL, PROFESSIONAL, AND LEGAL CONSIDERATIONS

Our project utilises anonymised Brain MRI data from the BraTS 2023 challenge [13]. All patients provided informed consent for their data usage through the challenge's established collaborations with medical institutions.

Since the data is pre-anonymised using standard methods such as skull stripping [13], no additional ethics clearance is needed. We commit to upholding data privacy and security regulations and best practices set by the BraTS challenge organizers.

Academic integrity is paramount. We will attribute and cite the SAM framework and any modifications appropriately, ensuring copyright compliance and avoiding plagiarism.

Our modified SAMSeg models' code and relevant scripts will be released under the MIT license, fostering collaboration and advancing SAM in the medical field. Intellectual property from the project will be owned by the University of Cape Town.

We remain vigilant for unforeseen ethical considerations, adapting our approach as needed and consulting our supervisor or university ethics committees if new concerns arise.

6 ANTICIPATED OUTCOMES

6.1 Research

We will create fine-tuned SAM models focused on meningioma segmentation. These are expected to offer performance gains over the SAM baseline when tested on the same inputs.

Our major challenge will be fitting the model designs within time and budget constraints. If too complex, we may be unable to train and evaluate our models using the available resources within the time limit. The project cannot be extended beyond the due date.

6.2 Impact

The results of our experiments will give insight into the effectiveness of hybrid fine-tuning for SAM in this task. This work is expected to motivate research in other domains such as pediatric tumors and brain metastases.

Tumour segmentation tools are currently used in the diagnosis and treatment of patients. Our work serves as a basis for possible advancements that manufacturers can make to their products. This contributes to advancements in clinical care for patients with meningiomas and possibly other neurological conditions.

6.3 Success Factors

- Ability to transform 3D images into 2D with minimal data loss.
- Final model designs can be trained and evaluated using available computing power.
- Improved accuracy over the baseline version of SAM.

7 PROJECT PLAN

7.1 Risk Assessment

Our comprehensive Risk Assessment Matrix, with corresponding Mitigation, Monitoring, and Management strategies, can be found in Appendix A.

7.2 Resources Required

7.2.1 Equipment: Deep learning models for medical image segmentation require substantial computational power, particularly GPUs for accelerated computations. Access to

departmental computing clusters or high-performance facilities with powerful GPUs will be necessary.

The BraTS 2023 Intracranial Meningioma Challenge dataset is approximately 20GB. While our workstations may initially suffice for storage, additional needs may arise with data augmentation and model size. Alternative solutions like scalable cloud storage services (e.g., Google Cloud Storage) will be explored if necessary.

7.2.2 Personnel: **Tapera & Cassandra:** Responsible for project oversight, research, development, evaluation, and documentation. **Supervisors (Patrick Marais and Fred Nicolls):** Regular meetings for guidance, feedback, and troubleshooting.

7.2.3 Software: **Deep Learning Framework:** Utilising PyTorch or TensorFlow for building, training, and deploying SAM models.
Version Control: Employing Git for collaborative development and version control.
Libraries and Tools:

- Data Manipulation & Processing: NumPy, Pandas
- Image Processing: OpenCV, scikit-image, imgaug
- Medical Image Analysis: SimpleITK, MONAI
- Model Evaluation Metrics: scikit-learn for metrics like Dice coefficient and Jaccard Index.

Visualisation: matplotlib for monitoring training progress.
Medical Image Segmentation GUI [16]: Open-source GUI for potential qualitative evaluation of model performance and visualisation of segmentation results.

7.3 Deliverables

Table 1 shows our project deliverables and corresponding deadlines.

Deliverable	Deadline
Literature Review	25/03/2024
Project Proposal Presentation	23/04/2024
Project Proposal	30/04/2024
Ethics Application (if required)	06/05/2024
Project Progress Demonstration	22/07/2024
Draft Project Paper	23/08/2024
Final Project Paper	30/08/2024
Final Project Code & Documentation *- Original/Vanilla SAM Model* *- Fine-Tuned SAM Baseline Model* *- Framework Modified #1 SAM Model* *- Framework Modified #2 SAM Model*	09/09/2024
Final Project Demonstration	16/09/2024
Project Poster	27/09/2024
Project Website	04/10/2024
School of IT Showcase (Presentation & Demonstration)	22/10/2024

Table 1: Our Research Project Deliverables.

7.4 Timeline & Milestones

A Gantt chart with our full project timeline is visualised in Appendix B. Table 2 presents a summary of the key milestones.

No.	Milestone	Expected Date
1	Completion of Research & Discovery	25/03/2024
2	Completion of Selection of Framework Modification Approaches	12/04/2024
3	Completion of Dataset Selection	16/04/2024
4	Completion of Data Collection & Pre-Processing	11/05/2024
5	Completion of Fine-Tuned SAM Baseline Model	25/05/2024
6	Completion of Individual Framework Modification Models	05/07/2024
7	Completion of Midpoint Demonstration Progress Review	26/07/2024
8	Completion of Final Evaluation & Refinement	03/08/2024
9	Completion of First Draft of Project Papers	23/08/2024
10	Completion of Project	22/10/2024

Table 2: Key Project Milestones.

8 TASK ALLOCATION

Our task allocations can be seen below in Table 3.

Cassandra	Tapera
Data Loading & Transforming	
Baseline Fine-Tuned SAM	
Modification #1	Modification #2
Evaluation	
Project Paper Write-Up	Project Paper Write-Up
Final Deliverables (website, poster)	

Table 3: Our Research Project Task Allocation.

REFERENCES

[1] Rahul Chauhan, Kamal Kumar Ghanshala, and RC Joshi. 2018. Convolutional neural network (CNN) for image detection and recognition. In *2018 first international conference on secure cyber computing and communication (ICSCCC)*. IEEE, 278–282.

[2] Tianrun Chen, Lanyun Zhu, Chaotao Ding, Runlong Cao, Yan Wang, Zejian Li, Lingyun Sun, Papa Mao, and Ying Zang. 2023. SAM Fails to Segment Anything? – SAM-Adapter: Adapting SAM in Underperformed Scenes: Camouflage, Shadow, Medical Image Segmentation, and More. arXiv:2304.09148 [cs.CV]

[3] Ruining Deng, Can Cui, Quan Liu, Tianyuan Yao, Lucas W. Remedios, Shunxing Bao, Bennett A. Landman, Lee E. Wheless, Lori A. Coburn, Keith T. Wilson, Yaohong Wang, Shilin Zhao, Agnes B. Fogo, Haichun Yang, Yucheng Tang, and Yuankai Huo. 2023. Segment Anything Model (SAM) for Digital Pathology: Assess Zero-shot Segmentation on Whole Slide Imaging. arXiv:2304.04155 [eess.IV]

[4] Lee R. Dice. 1945. Measures of the Amount of Ecologic Association Between Species. *Ecology* 26, 3 (1945), 297–302. http://www.jstor.org/stable/1932409

[5] Alper Dincer, Saul F. Morales-Valero, Stephanie M. Robert, Joanna K. Tabor, Joseph O'Brien, Kanat Yalcin, Robert K. Fulbright, Zeynep Ersen-Omay, Ian F. Dunn, and Jennifer Moliterno. 2023. Surgical strategies for intracranial meningioma in the molecular era. *Journal of Neuro-Oncology* 162, 2 (April 2023), 253–265. https://doi.org/10.1007/s11060-023-04272-z

[6] Hanxue Gu, Roy Colglazier, Haoyu Dong, Jikai Zhang, Yaqian Chen, Zafer Yildiz, Yuwen Chen, Lin Li, Jichen Yang, Jay Willhite, Alex M. Meyer, Brian Guo, Yashvi Atul Shah, Emily Luo, Shipra Rajput, Sally Kuehn, Clark Bulleit, Kevin A. Wu, Jisoo Lee, Brandon Ramirez, Darui Lu, Jay M. Levin, and Maciej A. Mazurowski. 2024. SegmentAnyBone: A Universal Model that Segments Any Bone at Any Location on MRI. arXiv:2401.12974 [eess.IV]

[7] Chuanfei Hu, Tianyi Xia, Shenghong Ju, and Xinde Li. 2023. When SAM Meets Medical Images: An Investigation of Segment Anything Model (SAM) on Multi-phase Liver Tumor Segmentation. arXiv:2304.08506 [eess.IV]

[8] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. LoRA: Low-Rank Adaptation of Large Language Models. arXiv:2106.09685 [cs.CL]

[9] Fabian Isensee, Jens Petersen, Andre Klein, David Zimmerer, Paul F. Jaeger, Simon Kohl, Jakob Wasserthal, Gregor Koehler, Tobias Norajitra, Sebastian Wirkert, and Klaus H. Maier-Hein. 2018. nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation. arXiv:1809.10486 [cs.CV]

[10] Ho Kang, Joseph Nathanael Witanto, Kevin Pratama, Doohee Lee, Kyu Sung Choi, Seung Hong Choi, Kyung-Min Kim, Min-Sung Kim, Jin Wook Kim, Yong Hwy Kim, Sang Joon Park, and Chul-Kee Park. 2023. Fully Automated MRI Segmentation and Volumetric Measurement of Intracranial Meningioma Using Deep Learning. *Journal of Magnetic Resonance Imaging* 57, 3 (2023), 871–881. https://doi.org/10.1002/jmri.28332 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jmri.28332

[11] Girish Katti, Syeda Arshiya Ara, and Ayesha Shireen. 2011. Magnetic resonance imaging (MRI)–A review. *International journal of dental clinics* 3, 1 (2011), 65–70.

[12] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. 2023. Segment Anything. arXiv:2304.02643 [cs.CV]

[13] Dominic LaBella, Maruf Adewole, Michelle Alonso-Basanta, Talissa Altes, Syed Muhammad Anwar, Ujjwal Baid, Timothy Bergquist, Radhika Bhalerao, Sully Chen, Verena Chung, Gian-Marco Conte, Farouk Dako, James Eddy, Ivan Ezhov, Devon Godfrey, Fathi Hilal, Ariana Familiar, Keyvan Farahani, Juan Eugenio Iglesias, Zhifan Jiang, Elaine Johanson, Anahita Fathi Kazerooni, Collin Kent, John Kirkpatrick, Florian Kofler, Koen Van Leemput, Hongwei Bran Li, Xinyang Liu, Aria Mahtabfar, Shan McBurney-Lin, Ryan McLean, Zeke Meier, Ahmed W Moawad, John Mongan, Pierre Nedelec, Maxence Pajot, Marie Piraud, Arif Rashid, Zachary Reitman, Russell Takeshi Shinohara, Yury Velichko, Chunhao Wang, Pranav Warman, Walter Wiggins, Mariam Aboian, Jake Albrecht, Udunna Anazodo, Spyridon Bakas, Adam Flanders, Anastasia Janas, Goldey Khanna, Marius George Linguraru, Bjoern Menze, Ayman Nada, Andreas M Rauschecker, Jeff Rudie, Nourel Hoda Tahon, Javier Villanueva-Meyer, Benedikt Wiestler, and Evan Calabrese. 2023. The ASNR-MICCAI Brain Tumor Segmentation (BraTS) Challenge 2023: Intracranial Meningioma. arXiv:2305.07642 [cs.CV]

[14] Cheng Li, Wen Li, Chenyang Liu, Hairong Zheng, Jing Cai, and Shanshan Wang. 2022. Artificial intelligence in multiparametric magnetic resonance imaging: A review. *Medical Physics* 49, 10 (2022), e1024–e1054. https://doi.org/10.1002/mp.15936 arXiv:https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1002/mp.15936

[15] Shengze Li, Jianjian Cao, Peng Ye, Yuhan Ding, Chongjun Tu, and Tao Chen. 2024. ClipSAM: CLIP and SAM Collaboration for Zero-Shot Anomaly Segmentation. arXiv:2401.12665 [cs.CV]

[16] Yunxiang Li, Bowen Jing, Zihan Li, Jing Wang, and You Zhang. 2023. nnSAM: Plug-and-play Segment Anything Model Improves nnUNet Performance. arXiv:2309.16967 [cs.CV]

[17] Xian Lin, Yangyang Xiang, Li Zhang, Xin Yang, Zengqiang Yan, and Li Yu. 2023. SAMUS: Adapting Segment Anything Model for Clinically-Friendly and Generalizable Ultrasound Image Segmentation. arXiv:2309.06824 [cs.CV]

[18] Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. 2024. Segment anything in medical images. *Nature Communications* 15, 1 (Jan. 2024). https://doi.org/10.1038/s41467-024-44824-z

[19] Ilaria Maggio, Enrico Franceschi, Alicia Tosoni, Vincenzo Di Nunno, Lidia Gatto, Raffaele Lodi, and Alba A Brandes. 2021. Meningioma: not always a benign tumor. A review of advances in the treatment of meningiomas. *CNS Oncology* 10, 2 (2021), CNS72. https://doi.org/10.2217/cns-2021-0003 arXiv:https://doi.org/10.2217/cns-2021-0003 PMID: 34015955.

[20] Christian Ogasawara, Brandon D. Philbrick, and D. Cory Adamson. 2021. Meningioma: A Review of Epidemiology, Pathology, Diagnosis, Treatment, and Future Directions. *Biomedicines* 9, 3 (2021). https://doi.org/10.3390/biomedicines9030319

[21] Jay N. Paranjape, Nithin Gopalakrishnan Nair, Shameema Sikder, S. Swaroop Vedula, and Vishal M. Patel. 2023. AdaptiveSAM: Towards Efficient Tuning of SAM for Surgical Scene Segmentation. arXiv:2308.03726 [cs.CV]

[22] Quan Quan, Fenghe Tang, Zikang Xu, Heqin Zhu, and S. Kevin Zhou. 2023. Slide-SAM: Medical SAM Meets Sliding Window. arXiv:2311.10121 [cs.CV]

[23] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597 [cs.CV]

[24] Jamie Shotton, John Winn, Carsten Rother, and Antonio Criminisi. 2006. TextonBoost: Joint Appearance, Shape and Context Modeling for Multi-class Object Recognition and Segmentation. In *Computer Vision – ECCV 2006*, Aleš Leonardis, Horst Bischof, and Axel Pinz (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1–15.

[25] Junde Wu, Wei Ji, Yuanpei Liu, Huazhu Fu, Min Xu, Yanwu Xu, and Yueming Jin. 2023. Medical SAM Adapter: Adapting Segment Anything Model for Medical Image Segmentation. arXiv:2304.12620 [cs.CV]

[26] Yichi Zhang and Rushi Jiao. 2023. Towards Segment Anything Model (SAM) for Medical Image Segmentation: A Survey. arXiv:2305.03678 [eess.IV]

A APPENDIX A: RISK ASSESSMENT

Risk	Likelihood	Severity	Mitigation	Monitoring	Management
Limited Availability of Computational Resources (High Computational Demands of Deep Learning Training)	High	High	Leverage cloud computing resources, such as Google Colab, for access to powerful GPUs, when the department cluster is not available. Consider using acceleration libraries, such as CUDA, if applicable. Ensure code is optimised for efficiency.	Track memory usage and GPU/CPU utilisation during training of our model. Monitor training time per epoch to assess progress and identify bottlenecks.	If this risk becomes a reality, explore alternative approaches to reduce network complexity to scale our computational cost with available resources. Such as: Model compression techniques, such as pruning, to decrease computational load, without sacrificing accuracy. Reducing the size of the model or the resolution of the input.
Data Loss due to Software Corruption or Hardware Failure	Low	High	Utilise version control systems, like Git, for regular, online code backups, allowing reversion to previous versions if necessary. Regularly commit code to the repository and maintain a practice of autosaving.	Monitor the integrity of backups through periodic verification, ensuring there are no inconsistencies in our repository.	If this risk becomes a reality, restore data from backups and revert to a previous working version of the code from the repository. Since backups will be performed regularly, the amount of data loss will be minimal.
Project Delays or Timing Issues	Medium	High	Develop a realistic and achievable project plan, avoiding underestimation of task durations. Activate development as soon as possible. Allocate sufficient time for each task with enough buffer room for potential setbacks. Maintain open communication with each other and our supervisor about progress and potential delays.	Regularly track project progress through our Gantt chart to identify deviations from the schedule and make adjustments. Monitor milestones and deliverables closely.	If this risk becomes a reality, and delay(s) become unavoidable, inform our supervisor immediately and discuss potential solutions, such as extensions or reducing the scope if necessary. Prioritise critical tasks to ensure core deliverables are still met. If necessary, explore alternative approaches that require less time.
Limited Availability of High-Quality Medical Data	Low	High	Utilise data augmentation techniques to artificially expand the dataset (e.g. flipping, rotating). Explore transfer learning from a related pre-trained Brain MRI MIS model.	Track data distribution and address class imbalances during augmentation. Regularly monitor training & validation loss and performance metrics to detect potential overfitting due to limited data.	If this risk becomes a reality, collaborate with medical institutions to access larger labelled datasets, and possibly consolidate several datasets into one. Evaluate the effectiveness of data augmentation techniques and refine if necessary. Consider incorporating semi-supervised learning techniques if unlabelled data is available.
Scope Creep	Medium	Medium	Clearly define the project scope in writing, leaving no ambiguity. Confirm this with our supervisor ahead of time. Implement advanced features, such as automatic prompting, only if time allows.	Regularly review project goals and deliverables with the supervisor. Maintain open communication with our supervisor to discuss any potential scope changes.	If this risk becomes a reality, refocus efforts on core project objectives and ensure their successful implementation, before considering any extended features. Discuss potential delays with the supervisor if scope expansion is necessary.

Risk	Likelihood	Severity	Mitigation	Monitoring	Management
Unexpected Convergence Issues during Training	Medium	Medium	Experiment with different hyper-parameter settings (learning rates, batch sizes, optimisers) to find the optimal convergence configuration. Implement early stopping mechanisms to prevent overfitting if training plateaus. Allocate buffer time in our project plan for troubleshooting and refining our training approach based on unforeseen obstacles.	Track performance metrics (loss, accuracy) of the model to assess model improvement. Use visualisation tools to visualise the training loss curves and validation accuracy, to appropriately identify any signs of model stagnation or divergence. Monitor for common signs of overfitting, such as stagnation or divergence.	If this risk becomes a reality, pivot to a different training approach to vastly improve our model's stability. Such as: Utilise learning rate schedulers to adjust the learning rate during training. Implement weight decay to prevent exploding gradients. Employ dropout layers to reduce overfitting. Explore curriculum learning, where the model is trained progressively on more challenging tasks.
Emergencies in Personal Life (health, family, etc.)	High	Medium	Prioritise our health, and maintain a healthy work-life balance. Maintain clear communication channels between each other and our supervisor to discuss any personal challenges that arise.	Monitor personal well-being and workload distribution.	If this risk becomes a reality, inform our supervisor immediately, and discuss the possibility of flexible deadlines, acknowledging the unforeseenness of the personal circumstances. Collaborate to temporarily adjust workloads to minimise project impact during emergencies.

B APPENDIX B: GANTT CHART (TIMELINE AND MILESTONES)