SAMMed-VR: Integrated Segment Anything Model in Virtual Reality for Supervised Brain Tumour Segmentation

Vahid Pooryousef* Monash University

Himashi Peiris† Monash University

Huiyi Liang‡ Monash University

Ang Li§ University of Queensland

Zhaolin Chen¶ Monash University

Tim Dwyer‖ Monash University

Jiazhou Liu** Monash University

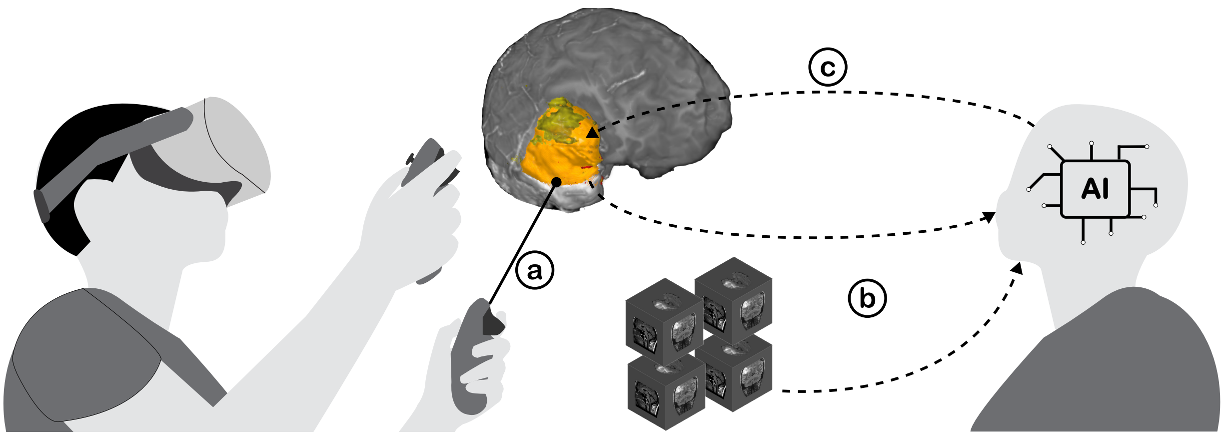

Figure 1: An overview of the workflow in our system: a) An expert makes a selection in VR, b) the AI segmentation model takes the selection with the brain tumour imaging data, and c) highlights predicted tumour regions in VR.

Abstract

Brain tumours differ significantly in shape, size, location, and contrast imperfections. Reliable segmentation is crucial for accurate identification and effective treatment. Recent advances in AI-based 3D medical image segmentation, such as SAM-Med3D, have improved automation; however, expert supervision remains vital for brain imaging. We present a virtual reality probe that integrates an expert-in-the-loop approach with SAM-Med3D, allowing users to iteratively enhance brain tumour segmentation by selecting points for AI refinement.

Index Terms: Brain Tumour Segmentation, AI model, VR

1 Introduction and Background

Due to their complex and heterogeneous nature, brain tumours pose challenges for diagnosis and treatment. Traditional tumour segmentation in MRI relies on manual annotation of three-dimensional scans, with experts marking hundreds of slices. The diverse appearance and unclear boundaries of brain tumours further complicate this process, making it time-consuming, subjective, and prone to error, which can delay diagnosis and limit treatment accessibility.

To overcome the limitations of manual annotation, recent studies have explored Machine Learning to automate segmentation. Deep Learning, especially Convolutional Neural Networks (CNNs), has become central to medical image segmentation, producing models that adapt to MRI data [2, 7]. However, these models depend heavily on training data and lack support for expert interaction during segmentation, reducing their reliability in practice. Inspired by Vision Foundation Models (VFMs) [3], newer models like SAM-Med3D [6] aim to address data limitations and enable user input. Since their initial predictions are often imprecise, expert involvement remains essential, especially for sensitive regions like brain tumours. Therefore, designing an interface for efficient interaction with model outputs is crucial.

On the other hand, immersive technologies, such as virtual reality (VR), offer an interaction paradigm for medical image visualisation by presenting data in 3D space, reducing occlusion, distortion, and information loss [4]. Beyond the benefits of 3D display, immersive technologies enable natural interactions (*e.g.*, [5]), enabling users to engage with data through gestures or controllers (*e.g.*, tapping or grabbing). In this research, we explore integrating AI models with VR to provide medical experts an interactive 3D interface for refining AI-predicted segmentation results.

2 Architecture

Figure 2 illustrates the architecture of our proposed system. Here, the imaging data is visualised in 3D for user analysis, with users able to iteratively select points to improve segmentation. These points are processed by the AI model, which refines the segmentation based on user input, and updated results are returned for further review. For the first iteration, the AI model segments the dataset independently, providing an initial context for further user refinement. We adopt an AI model extended from SAM-Med3D, as prior studies [6] have shown such models achieve acceptable accuracy (*i.e.*, dice score above 88% [2]) without human intervention. While the extent of data change per iteration is not immediately meaningful, we enhance the visualisation tool with techniques that help users

*e-mail: vahid.pooryousef@monash.edu
†e-mail: himashi.peiris@monash.edu
‡e-mail: mia.liang@monash.edu
§e-mail: ang.li@uq.edu.au
¶e-mail: zhaolin.chen@monash.edu
‖e-mail: tim.dwyer@monash.edu
**e-mail: joe.liu@monash.edu

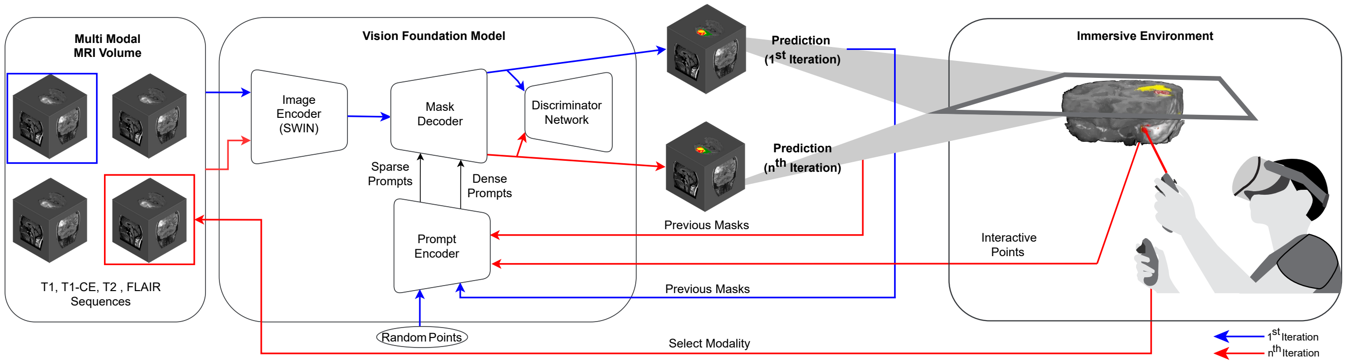

Figure 2: SAMMed-VR architecture.

conveniently review and select points.

3 System Features

AI Model. Inspired by SAM-Med3D [6], we implemented a VFM specifically for brain tumour segmentation task. In contrast to previous VFMs where multiple tasks are trained together on a large dataset, here we only focus on tumour subregion segmentation using four MRI sequences (*i.e*., T1, T2, T1CE and FLAIR). To train the model, we use 484 patient cases from MSD BraTS task [1]. With the co-registered 4 MRI sequences, altogether we have 1936 MRI scans for training, validation and testing.

We incorporate a discriminator network—a fully convolutional neural network—to improve the segmentation quality of the predictions. Additionally, the model employs a SWIN Transformer based image encoder to generate voxel embeddings instead of using a Vision Transformer. The model takes MRI data and user points as sparse prompts, with prior masks (initially empty) as dense prompts, to predict brain tumour regions, as illustrated in Fig. 2.

3D Visualisation Area. Figure 3.left, as well as visualised brain in Fig. 1 and Fig. 2 show the visualisation area. The tumour regions are highlighted with a primary colour (Red, orange and yellow represent Non-enhancing tumour (NCR/NET), Enhancing tumour (ET), Edema (ED) regions, respectively). Users can resize the brain by dragging the bounding box corners and rotate it along three axes via white handles at the box's edge centres. The other techniques have been utilised, include:

Cross-section Plane slices through the brain or segments, and can be freely moved or rotated (the green rectangle in Fig. 3.left).

Cutting Box cuts the regions on three axes (see Fig. 3.right) and follows the controller to help selecting points.

Point Selector is placed in a corner of the Cutting Box (red dot in Fig. 3.right) for clear visibility on three axes. To enhance precision despite VR hand-tracking inaccuracies, we use a two-stage process: users first move the target point freely, then make fine adjustments with controller sensitivity reduced tenfold.

Imaging viewers on Cross-section Plane and Cutting Box have been integrated, allowing users to view MRI images on the 3D model and thereby enhance decision-making when selecting points.

4 Discussion and Future Work

Our proposed system is designed to provide experts with as much relevant and effective information as possible, enabling accurate and rapid decision-making. By incorporating VR, we strive to achieve superior performance through quantitative comparison to traditional 2D desktop interfaces. We believe that incorporating a human-in-the-loop approach is essential in the era of AI integration, particularly for sensitive decision-making tasks such as brain tumour diagnosis. Immersive technologies offer a valuable platform

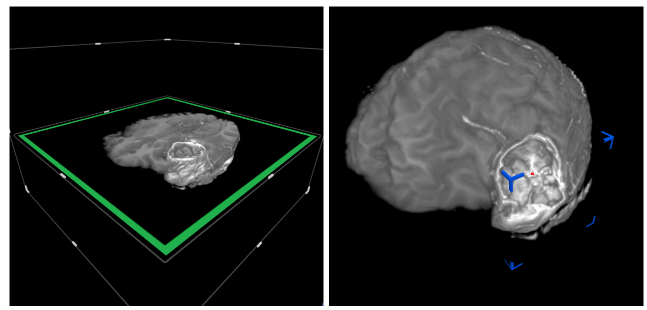

Figure 3: Cross-section Plane (left) and Cutting Box (right) for analysing internal regions of the brain.

for these tasks, but effective interaction and visualisation techniques are crucial to ensure informed decision-making while not hindering AI performance benefits.

In the future, we will test the effectiveness of this tool against automated and semi-automated desktop slicers. Additionally, future research may investigate advanced VR interaction techniques to enhance user performance and accuracy.

References

[1] M. Antonelli, A. Reinke, S. Bakas, K. Farahani, A. Kopp-Schneider, B. A. Landman, G. Litjens, B. Menze, O. Ronneberger, R. M. Summers, et al. The medical segmentation decathlon. *Nature communications*, 13(1):1–13, 2022. 2

[2] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2):203–211, 2021. 1

[3] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, et al. Segment anything. *arXiv preprint arXiv:2304.02643*, 2023. 1

[4] T. Munzner. *Visualization analysis and design*. CRC press, New York, NY, USA, 2014. 1

[5] M. Sousa, D. Mendes, S. Paulo, N. Matela, J. Jorge, and D. S. o. Lopes. Vrrrroom: Virtual reality for radiologists in the reading room. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 4057–4062, New York, NY, USA, 2017. Association for Computing Machinery. 1

[6] H. Wang, S. Guo, J. Ye, Z. Deng, J. Cheng, T. Li, J. Chen, Y. Su, Z. Huang, Y. Shen, et al. Sam-med3d: towards general-purpose segmentation models for volumetric medical images. *arXiv preprint*, 2023. 1, 2

[7] R. Wang, T. Lei, R. Cui, B. Zhang, H. Meng, and A. K. Nandi. Medical image segmentation using deep learning: A survey. *IET image processing*, 16(5):1243–1267, 2022. 1