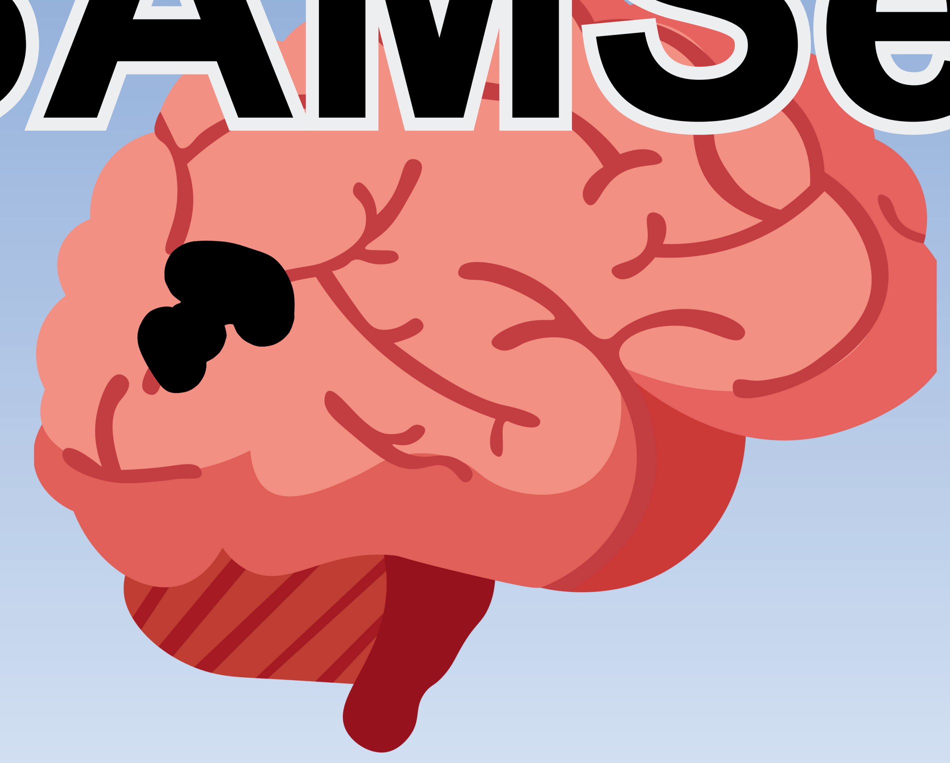
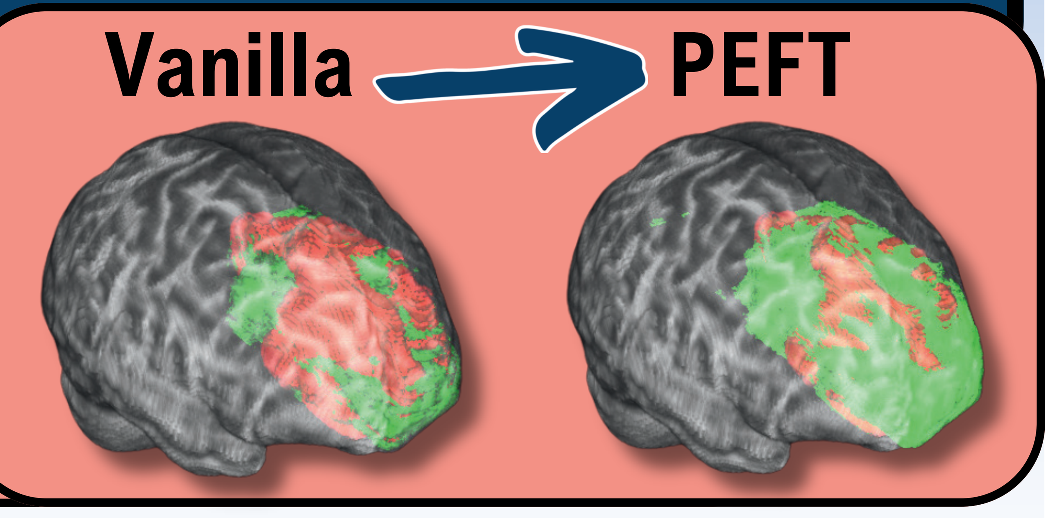
SAMSeg



Improving SAM for Brain Tumour Segmentation

The *Segment Anything Model (SAM)* is a state-of-the-art, promptable deep learning model that excels in 2D natural image segmentation. However, medical imaging, especially brain tumour segmentation, presents unique challenges due to the complexity of tumour boundaries and the multi-dimensionality of data. This project aims to enhance SAM's capability for brain tumour segmentation through *Parameter-Efficient Fine-Tuning (PEFT)* on the BraTS Intracranial Meningioma 2023 dataset. In addition, we introduce two framework modifications to augment SAM's performance: **U-SAM** and **LoRA.**



Vanilla & PEFT-SAM:

The baseline - **Vanilla SAM** - demonstrated strong performance for tumour segmentation - **84.1%**. Using bounding box prompts, we fine-tuned SAM on the BraTS dataset, achieving **87.7%** *(+3.6%)* PEFT-SAM sets the stage for our architectural modifications.

U-SAM:

U-SAM is a complementary, hybrid model that integrates SAM with U-Net's superior 3D spatial capabilities for better tumour boundary precision. We trained our 3D U-Net on the same dataset and achieved **83.5%**. Two variations were explored:

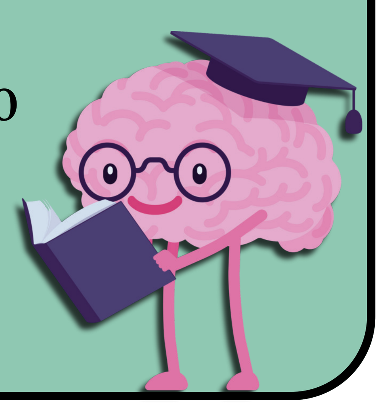
U-SAM-Combo:

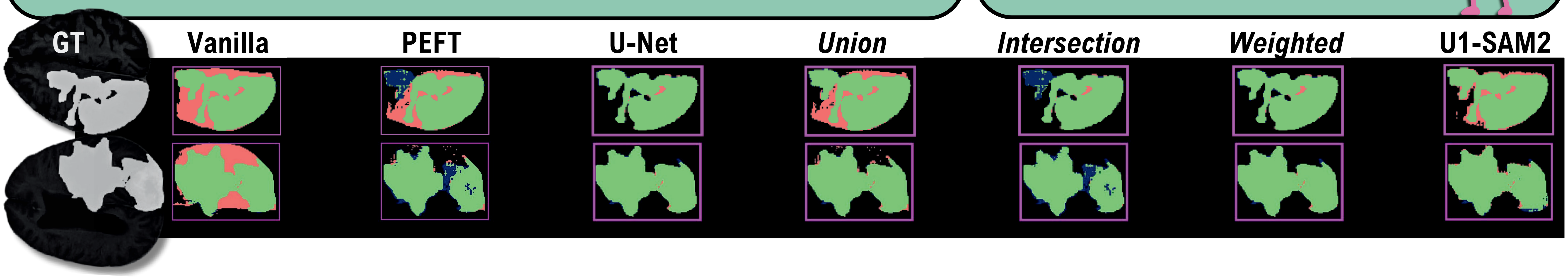
In **U-SAM-Combo**, SAM and U-Net branches were combined using different methods:

- *Union*: Combines predictions from SAM and U-Net.
- *Intersection*: Focuses on the overlap between SAM and U-Net.
- *Weighted*: Through intermediate evaluation, assigns weights by model performance.

U1-SAM2:

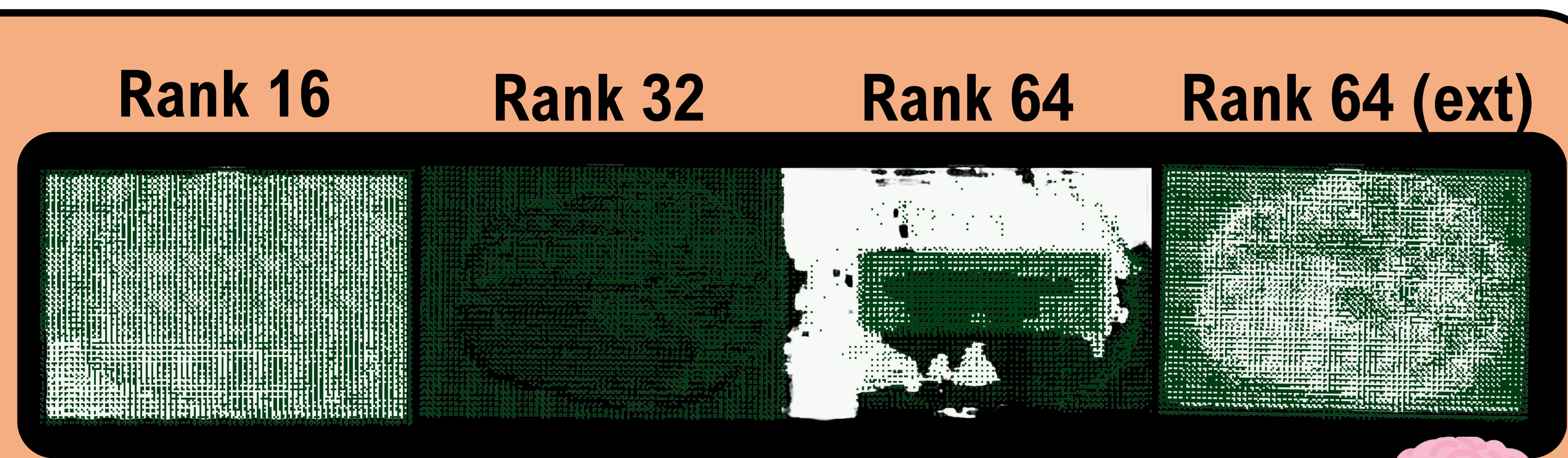
In **U1-SAM2**, U-Net's output is fed into SAM as a guidance mask, allowing the model to fine-tune its segmentation using U-Net's predictions.

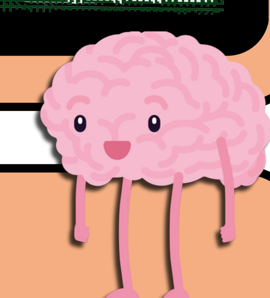




LoRA:

LoRA is a fine-tuning technique where pre-trained weights are decomposed into a pair of smaller matrices. The size of the smaller matrices is controlled by a rank parameter. We investigated how applying LoRA to SAM's image encoder at different ranks affects segmentation accuracy.





U-SAM Results

Model	Accuracy	Change from PEFT-SAM	Change from Vanilla SAM
U-SAM-Combo (*Union*)	88.6%	0.9%	4.5%
U-SAM-Combo (*Intersection*)	82.6%	-5.1%	-1.5%
U-SAM-Combo (*Weighted*)	91.8%	4.1%	7.7%
U1-SAM2	85.1%	-2.6%	1.0%

While **U1-SAM2** struggled with spurious predictions, **U-SAM-Combo** excelled, especially with *Union* and *Weighted*. The results clearly demonstrate U-SAM's potential, meeting our goal to enhance SAM's brain tumour segmentation accuracy.

LoRA Results

Model	Accuracy	Change from PEFT-SAM	Change from Vanilla SAM
LoRA-64	2.1%	-85.6%	-82.0%
LoRA-32	3.9%	-83.8%	-80.1%
LoRA-16	4.9%	-82.8%	-79.1%

Reduction in training loss was not reflected in validation performance. Smaller ranks offered better performance but still fell below baselines. When visualised, all LoRA models appear to learn entire image rather than tumour masks. Further experimentation could not be performed in time for more conclusive results that confirm and correct these suspicions.

Authors
Tapera Chikumbu | Cassandra Wallace
chktap011@myuct.ac.za | wllcas004@myuct.ac.za

Supervisors
Patrick Marais | Fred Nicolls
patrick@cs.uct.ac.za | fred.nicolls@uct.ac.za

UNIVERSITY OF CAPE TOWN
Department of Computer Science
dept@cs.uct.ac.za cs.uct.ac.za