

SegMed

Implementation of MedSAM for medical
Imaging and enhancement with FeatUp
and GAFL



<p> Hasaan Maqsood, Iana Kulichenko,
Daniil Volkov, Sergey Egorov</p>

TABLE OF CONTENTS.



01

Problem



02

Dataset



03

Baseline models



04

FeatUP



05

GAFL



06

Models
Comparison

Problem

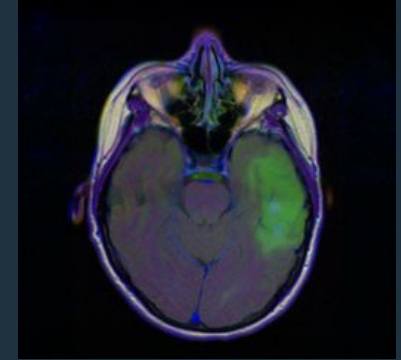
The problem we are addressing is the manual segmentation of brain tumors from MRI images. This process is critical for treatment planning but is often slow and prone to errors.

Our goal is to automate this process using the MedSAM model and enhance it with FeatUp and GAFL techniques. By doing so, we aim to improve the efficiency and accuracy of tumor segmentation.



DATASET

LGG Segmentation Dataset



The dataset consists of MRI brain images paired with FLAIR abnormality segmentation masks derived from The Cancer Imaging Archive (TCIA). It includes 110 cases from The Cancer Genome Atlas (TCGA), featuring lowergrade glioma patients. Each case includes pre-contrast, FLAIR, and post-contrast sequences, with segmentation masks provided as binary, single-channel images. The dataset is partitioned into training, testing, and validation sets, with 2828, 393, and 708 samples, respectively.

Reference: <https://registry.opendata.aws/tcga/>

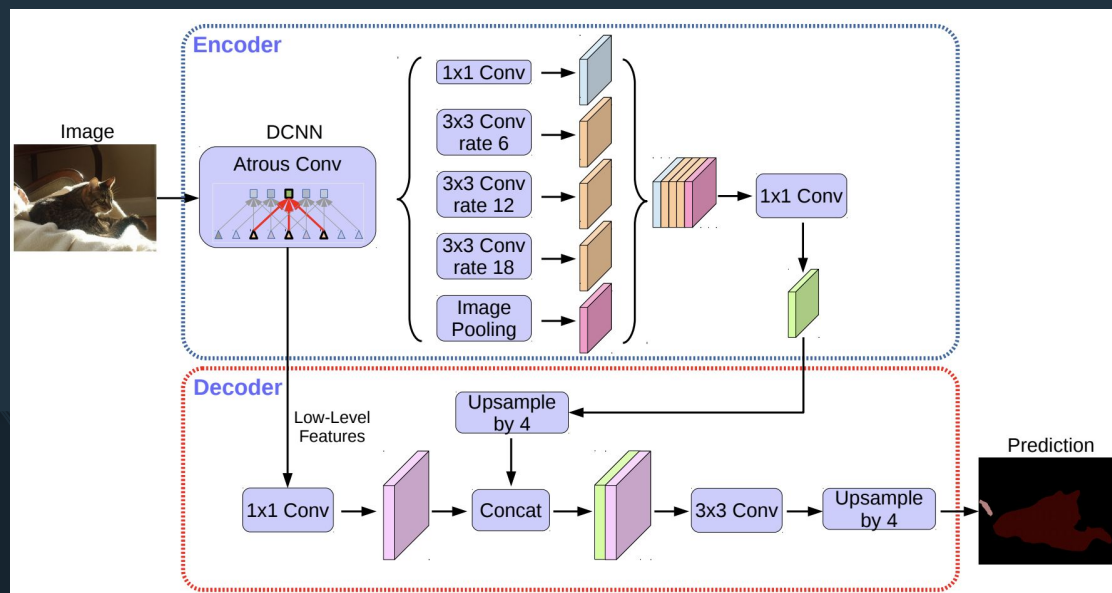


Milestones: 1st part

Baseline Models

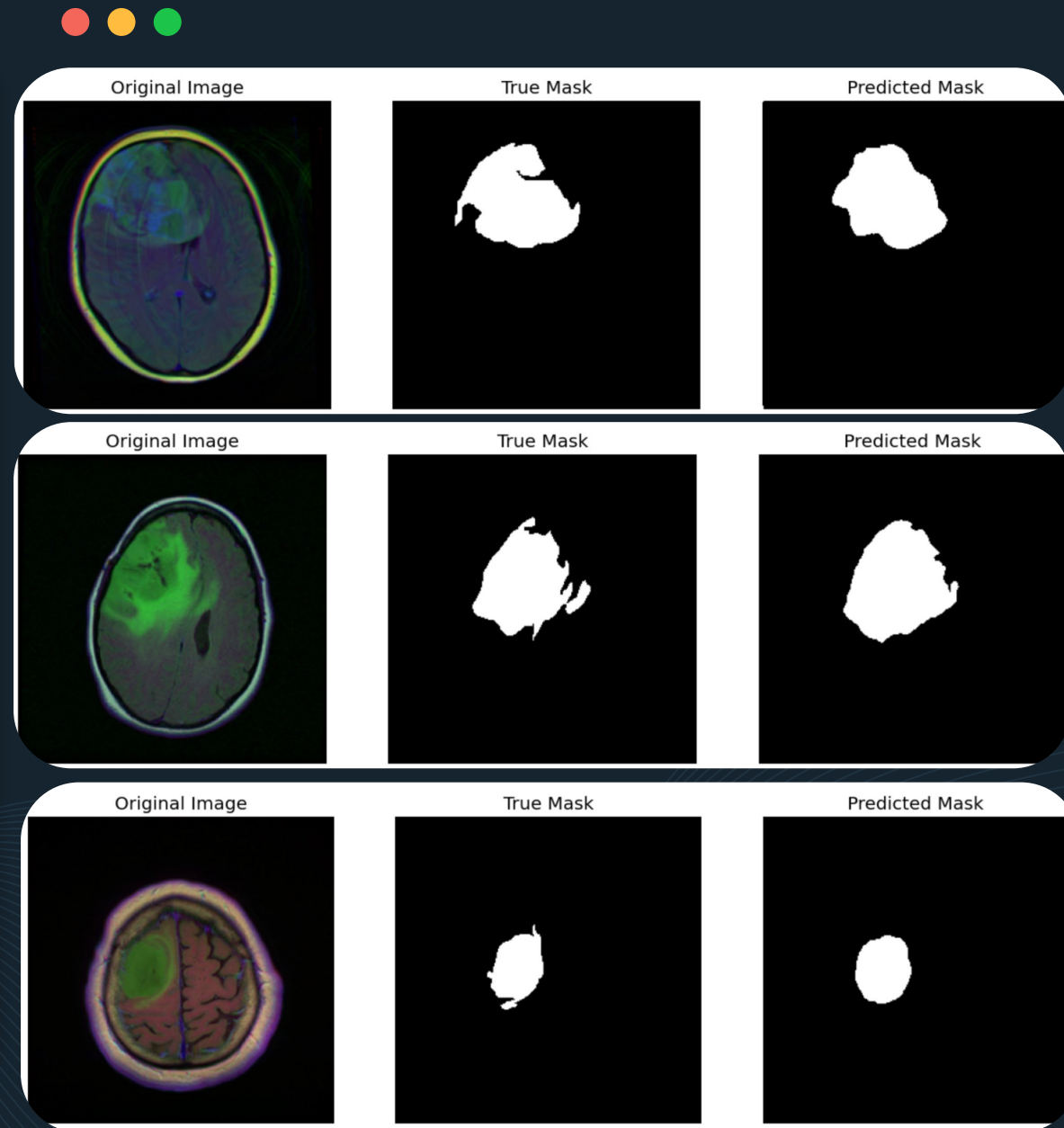
- Implemented baseline models: DeepLabV3+, U-Net, U-Net++, DeepLabV3, and Pyramid Attention Network (PAN).
- Best performing model: DeepLabV3+ with an IoU of 0.87

DeepLabV3+



Results Comparison of Baseline models

Model	Train Loss	Valid Loss	IOU
DeepLabV3+	1.6634	1.6662	0.8741
UNET	1.6691	1.6712	0.8517
UNET++	1.6730	1.67750	0.8268
DeepLabV3	1.6646	1.6700	0.7986
PAN	1.6656	1.6683	0.7970





Milestones: 2nd part

- SAM
- SAM + FeatUp
- SAM + GAFL

Advance Models

- After establishing a strong foundation with baseline models, our focus shifted towards advanced model enhancements.
- We aimed to further improve segmentation accuracy and efficiency by integrating cutting-edge techniques into the SAM model.
- These enhancements include the incorporation of FeatUp for improved feature resolution and GAFL for optimized frequency content.

MedSAM

Segment Anything Model

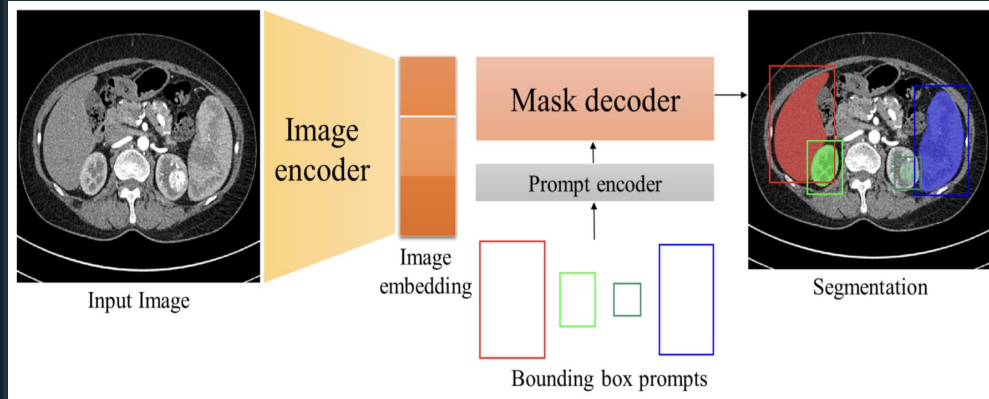


Image encoder is pre-trained Vision Transformer VIT, Prompt Encoder represented with an off-the-shelf text encoder from CLIP. Mask Decoder is modified Transformer decoder blocks.

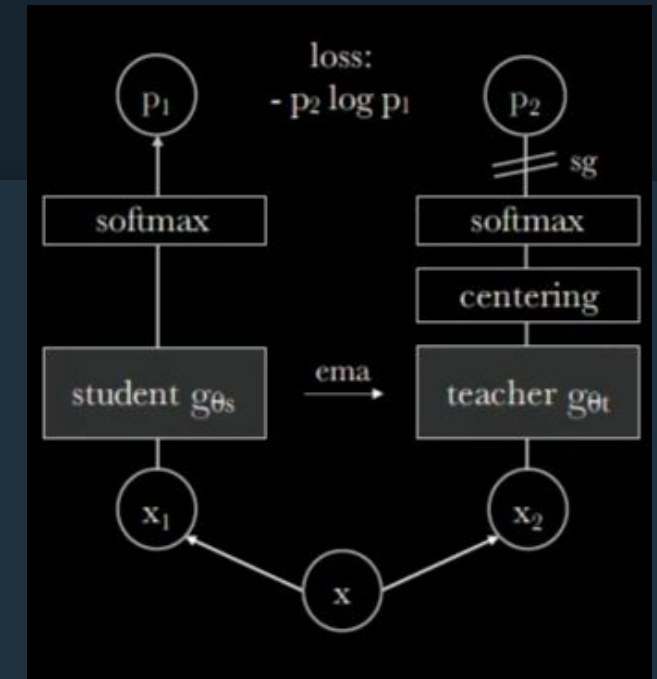
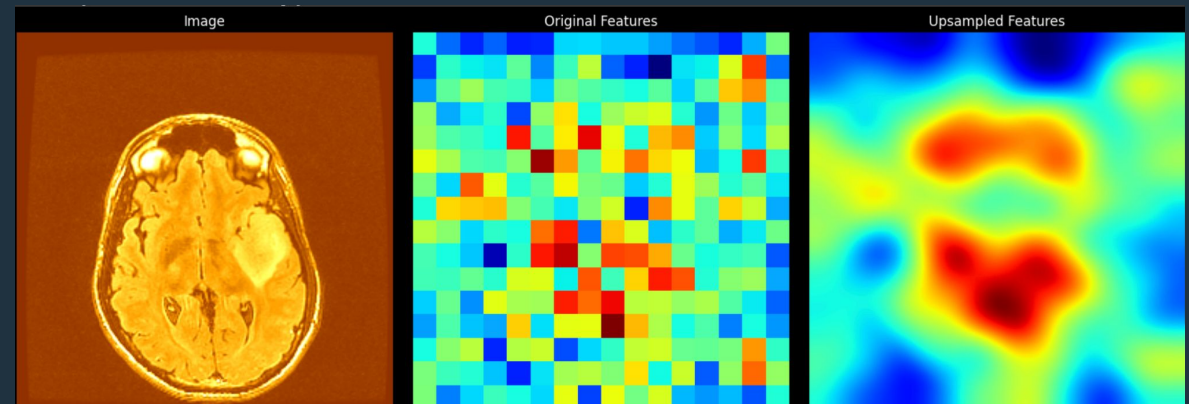
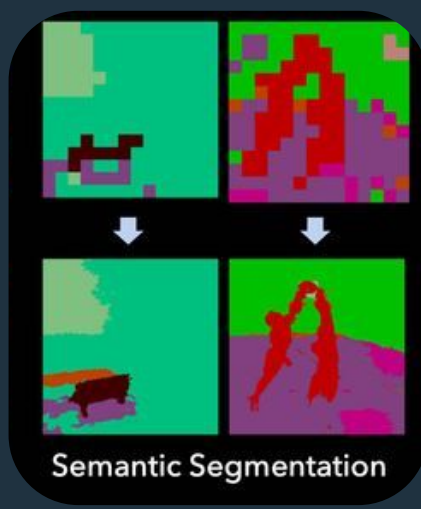
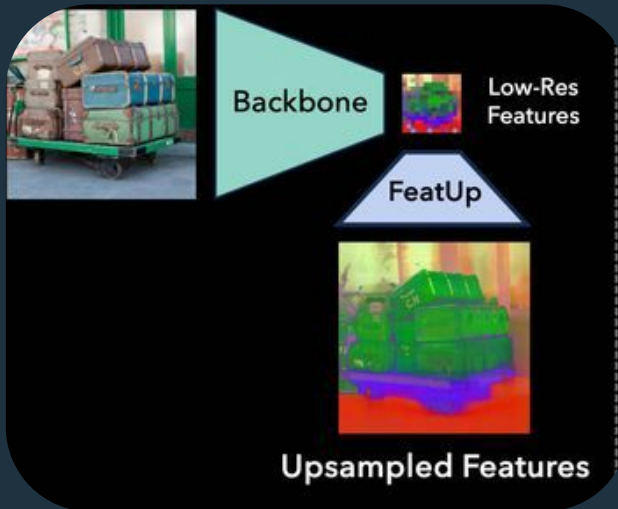


Reference:
<https://www.nature.com/articles/s41467-024-44824-z>

FeatUp: A Model-Agnostic Framework for Features at Any Resolution



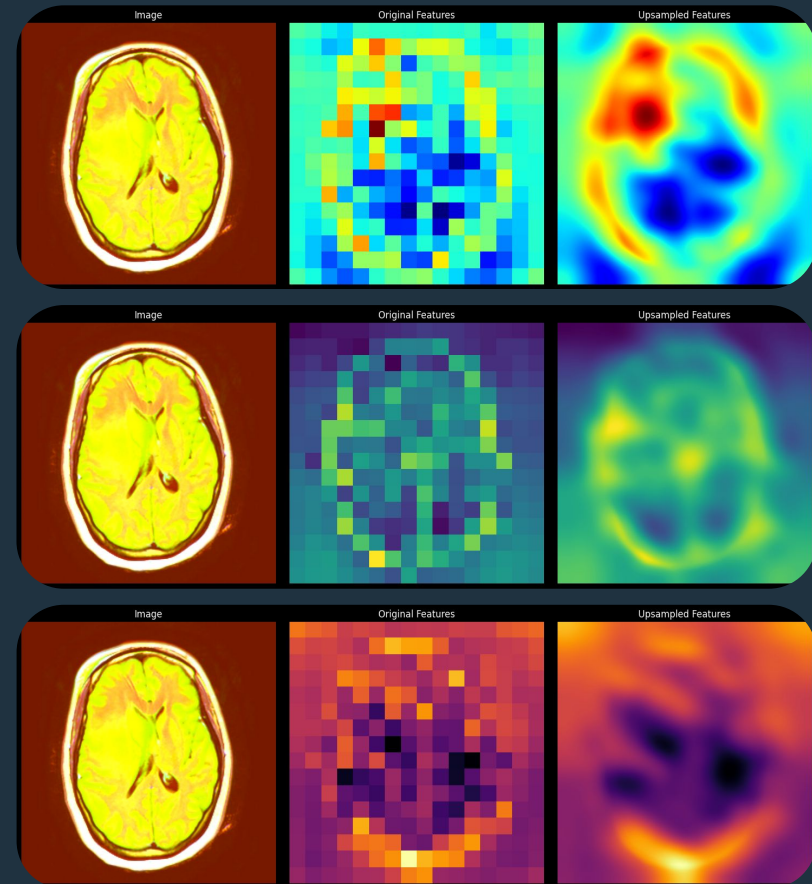
FeatUp upsamples image features from any model backbone, adding spatial resolution to existing semantics.



Reference:
<https://github.com/mhamilton723/FeatUp>

SAM/SAM+FeatUp Comparison

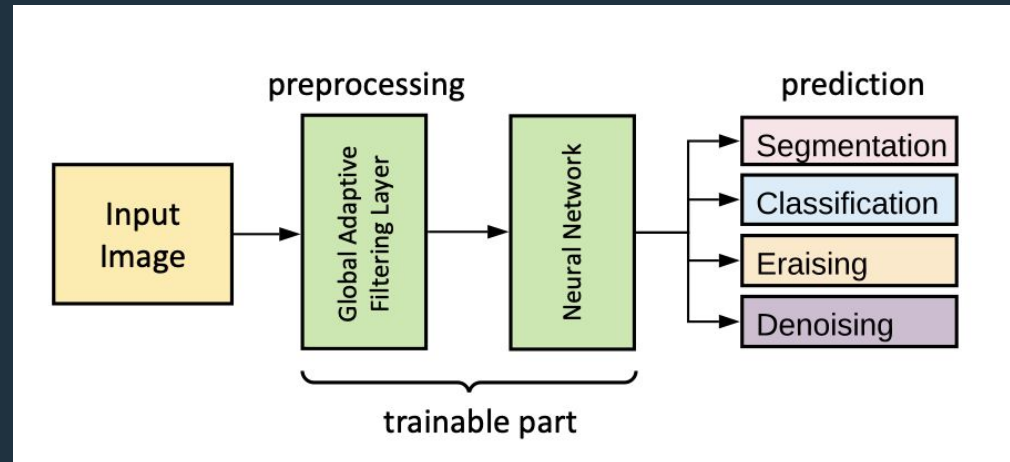
	SAM	SAM+FeatUP
Mean Loss	0.0442	0.0159
Mean IoU	0.8095	0.8536
Mean Accuracy	0.9721	0.9763
Mean Precision	0.5089	0.5729
Mean Recall	0.2857	0.4258
Mean F1 Score	0.2947	0.4351
Mean Dice Score	0.2870	0.3395



Reference:
<https://github.com/mhamilton723/FeatUp>

GAFL:

Global Adaptive Filtering Layer for Computer Vision



Algorithm 1: Global Adaptive Filtering Layer

Input: I – Initial image.

\mathcal{F} – Fast Fourier Transform operator.

- 1: $W_1, W_2, B_1, B_2 = \text{ReLU}(W_1, W_2, B_1, B_2);$
- 2: $F = \mathcal{F}I;$
- 3: $S = W_2 * \sigma(W_1 * |F| + B_1) + B_2;$
- 4: $S = S * F / |F|;$
- 5: $I' = \mathcal{F}^{-1}S;$

Output: I' – Image after global frequency filtering.

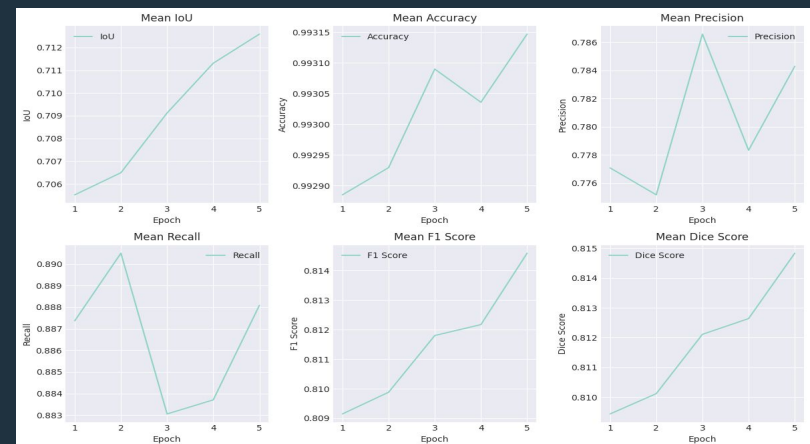
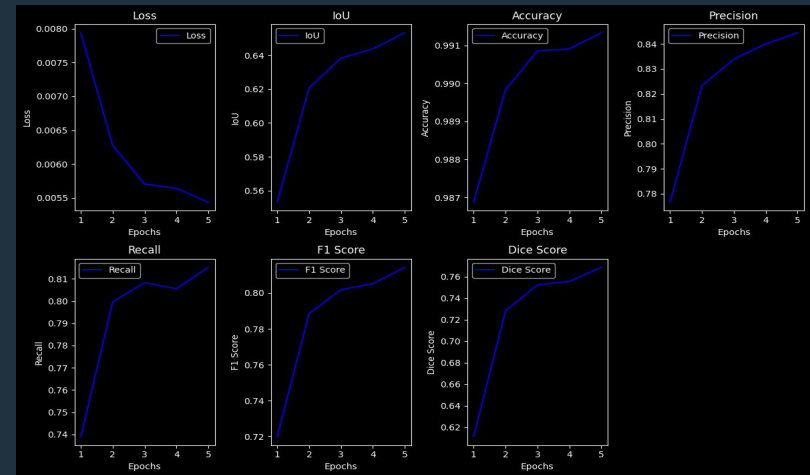
Reference:

<https://github.com/cviaai/GAFL>

SAM/SAM+GAFL Comparison



	SAM	SAM+GAFL
Mean Loss	0.0052	0.0062
Mean IoU	0.6757	0.9245
Mean Accuracy	0.9919	0.9698
Mean Precision	0.8609	0.9468
Mean Recall	0.8344	0.9289
Mean F1 Score	0.8348	0.9176
Mean Dice Score	0.7786	0.8124



Reference:

<https://github.com/mhamilton723/FeatUp>

Paired T-Test

statistical analysis

Null Hypothesis (H_0): There is no significant difference in the performance metrics between the models.

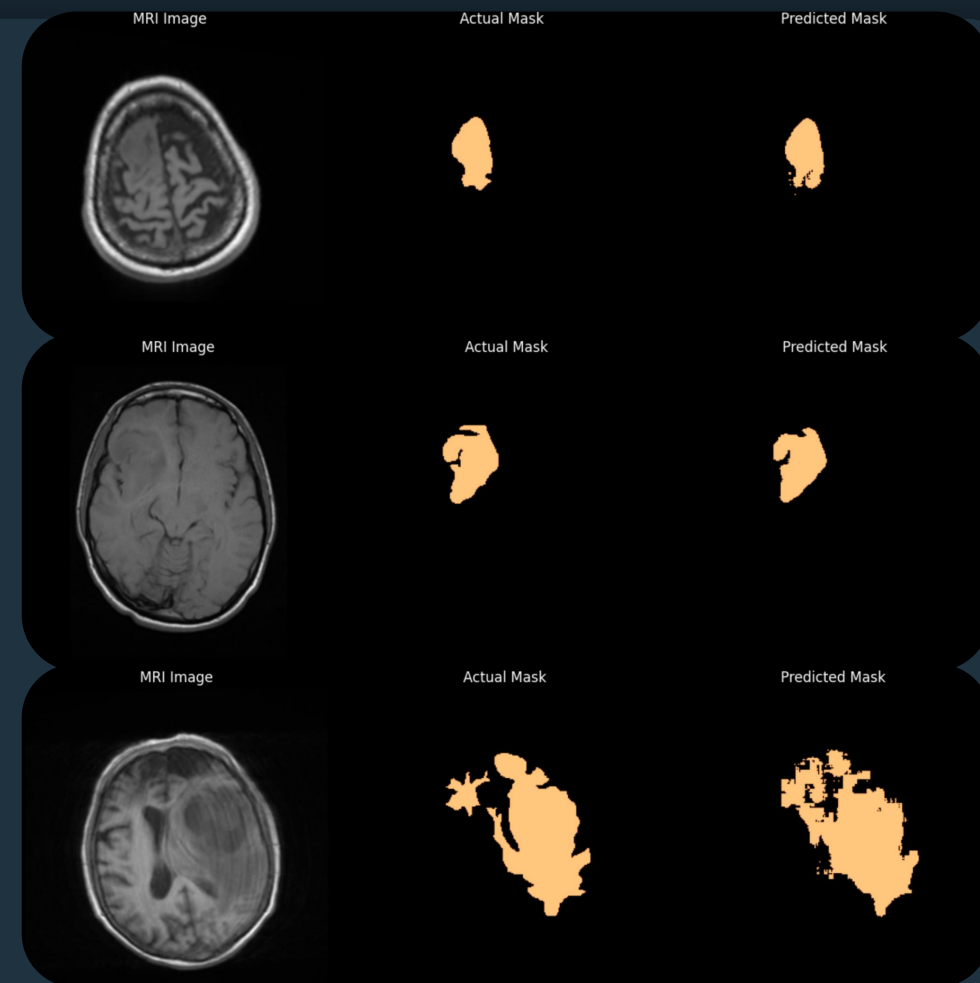
Alternative Hypothesis (H_1): There is a significant difference in the performance metrics between the models.

1. T-test between SAM and SAM+FeatUP:
 - t-statistic: 2.0755
 - p-value: 0.0832
2. T-test between SAM and SAM+GAFL:
 - t-statistic: -0.2600
 - p-value: 0.8035

- SAM vs. SAM+FeatUP: The p-value is 0.0832, which is greater than the common significance level of 0.05 but less than 0.10. This suggests that there is a marginally significant difference between the performance metrics of the SAM and SAM+FeatUP models. However, it is not strong enough to conclusively reject the null hypothesis at the 0.05 level.
- SAM vs. SAM+GAFL: The p-value is 0.8035, which is greater than the common significance level of 0.05. Therefore, we fail to reject the null hypothesis. This means that there is no statistically significant difference between the performance metrics of the SAM and SAM+GAFL models.

Results of Comparison models

	SAM	SAM+FeatUP	SAM+GAFL
Mean Loss	0.0052	0.0159	0.7004
Mean IoU	0.6757	0.8537	0.7004
Mean Accuracy	0.9919	0.9762	0.9926
Mean Precision	0.8609	0.5728	0.7704
Mean Recall	0.8344	0.4258	0.8901
Mean F1 Score	0.8348	0.4351	0.8043
Mean Dice Score	0.7786	0.3395	0.8045



Reference:
<https://github.com/mhamilton723/FeatUp>

Team



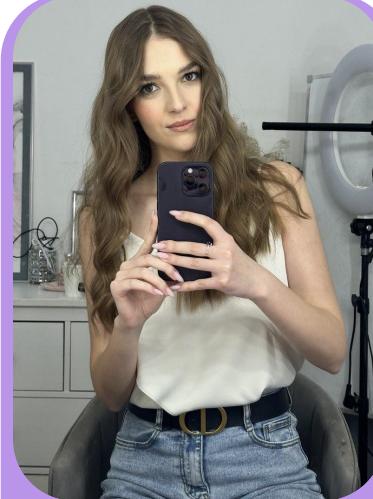
Hasaan Maqsood @Hasaan4

SAM, GAFL, FeatUp
integration,
presentation, reports



Sergey Egorov
@WastingT_me

Sam+FeatUp code
Presentation



Yana Kulichenko
@iani_kuli

Baseline models,
SAM, presentation,
reports



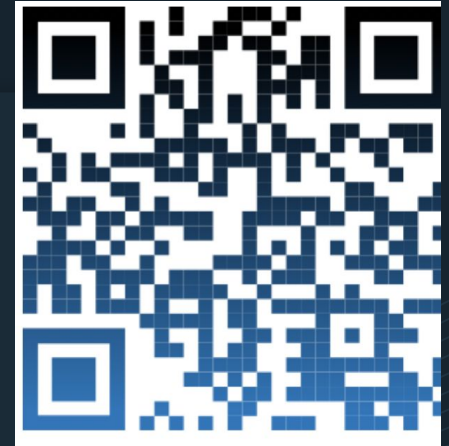
Daniil Volkov
@Gamma_function

Testing of FeatUp,
GAFL

THANK YOU!

Do you have any questions?

OUR GITHUB



@Hasaan45
@iani_kuli
@WastingT_me
@Gamma_function

