Presentation on

In Search of State-of-the-art Implementation for Segmentation of Metallographic Images: New Dataset, Challenges, Task-Specific, Universal and Fusion Model approach.

(working title)

Bishal

04-30-2024

Computer Vision & Pattern Recognition Lab
kit 금오공과대학교 Kumoh National Institute of Technology

In Search of State-of-the-art Implementation for Segmentation of Metallographic Images: New Dataset, Challenges, Task-Specific, Universal and Fusion Model approach.

(working title)

Information Fusion

Supports open access

38.6
CiteScore

18.6
Impact Factor

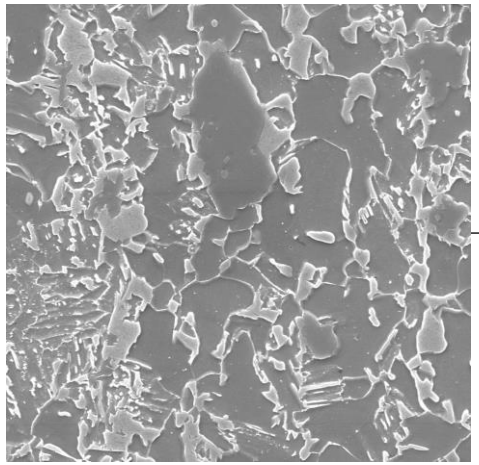
Paper Contributions

1. Introduction of new dataset for alloy steel image segmentation.
   - Current publicly available dataset - MetalDam, UCHS
2. Proposal of State-of-the-art task-specific model
   - Based on previously submitted paper to CAI 24 and ICASI 24
3. Comparison with vision foundational model – SAM and its adaption using LoRA
   - SAM
   - LoRA-SAM
     - No Guidance
     - Bounding Box Guidance
     - Ratio Guidance

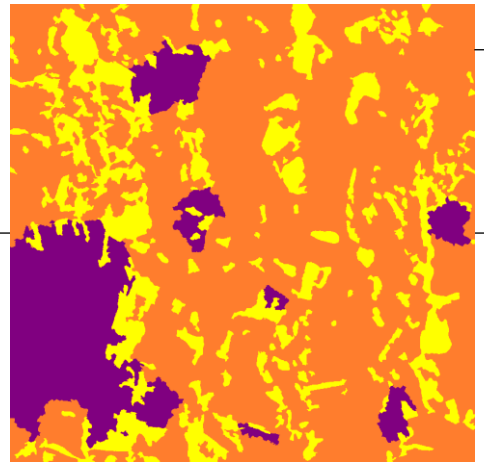
   These works itself can produce additional conference papers
4. Proposal of a fusion model → task-specific and vision foundational model
   - Based on previously submitted paper to ECCV 24

Today's presentation provides further brief into point 3

Preliminary Results

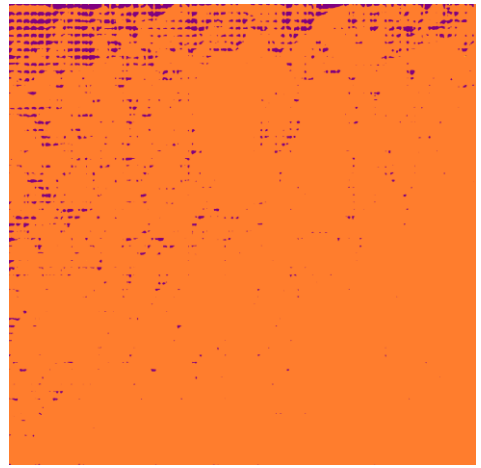


input

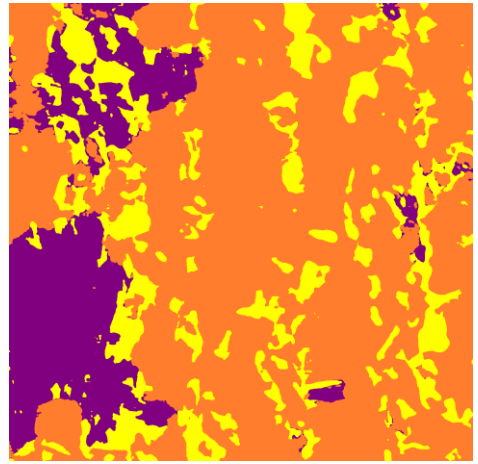


Bounding Boxes

SAM

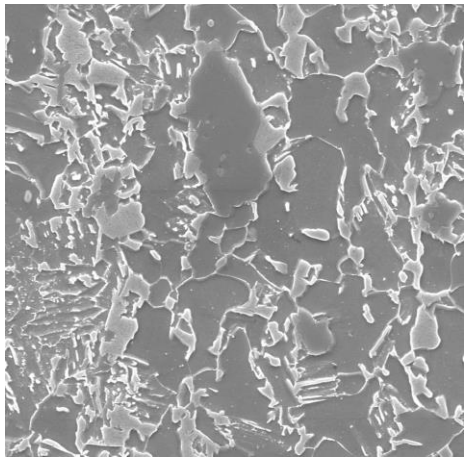


LoRA-SAM
R-512



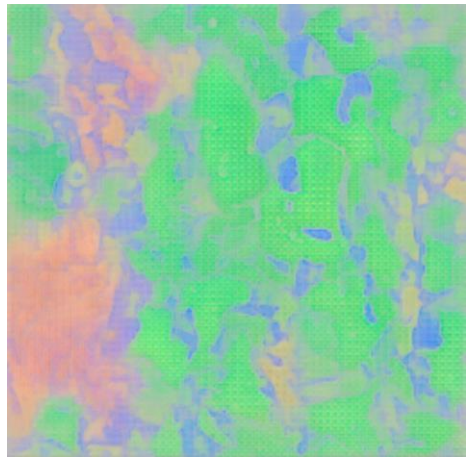
Preliminary Results

Under the hood -

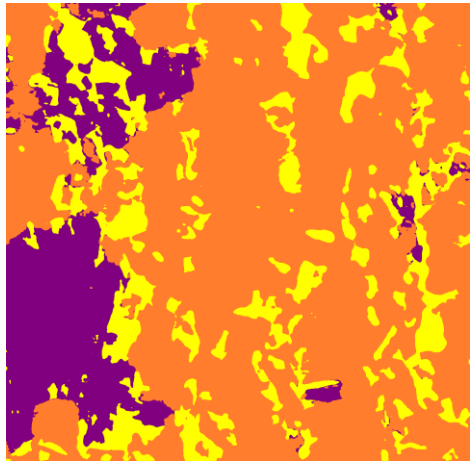


× 2700
SEM

LoRA-SAM
R-512



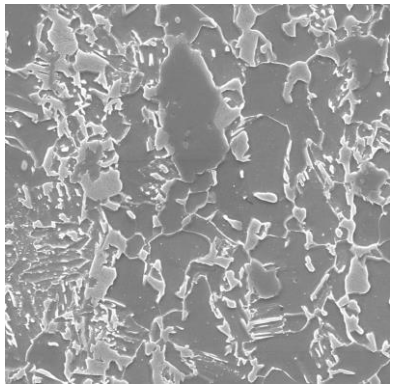
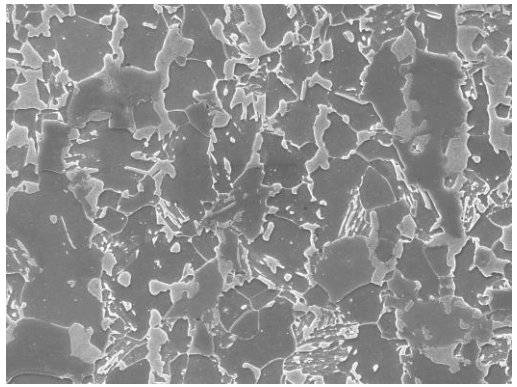
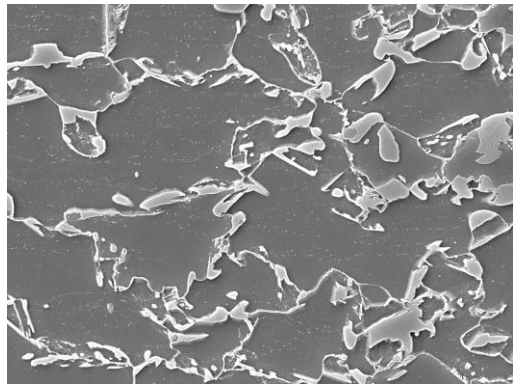
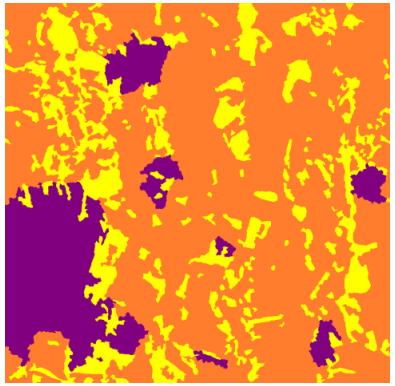
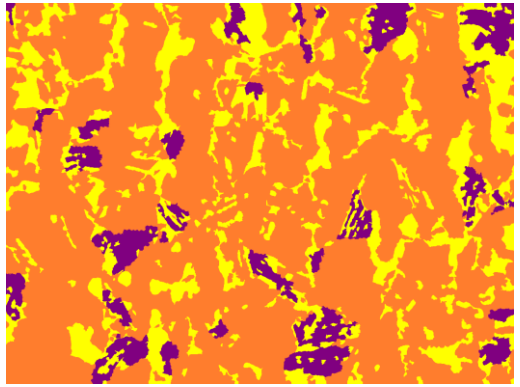
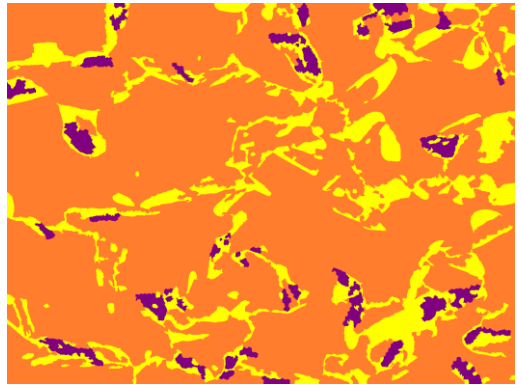
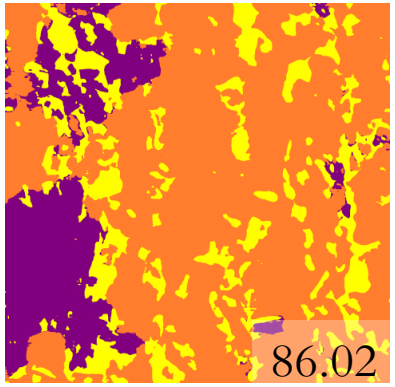
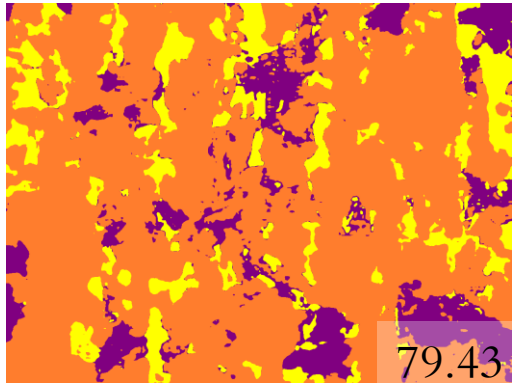
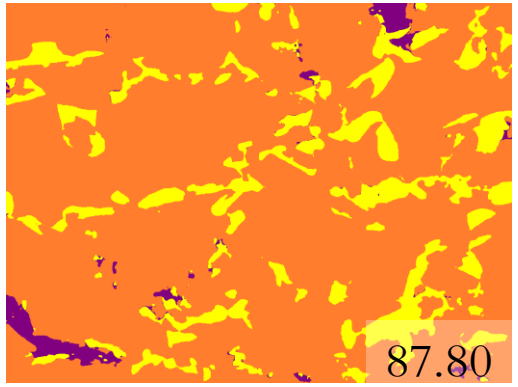
Soft Labels
per class



Softmax and
Recolor

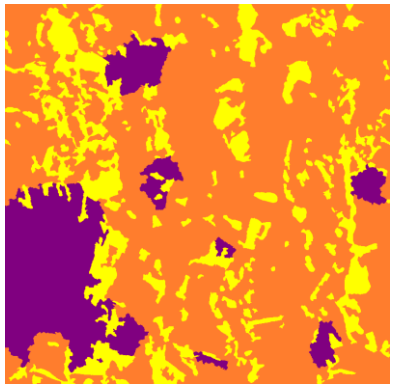
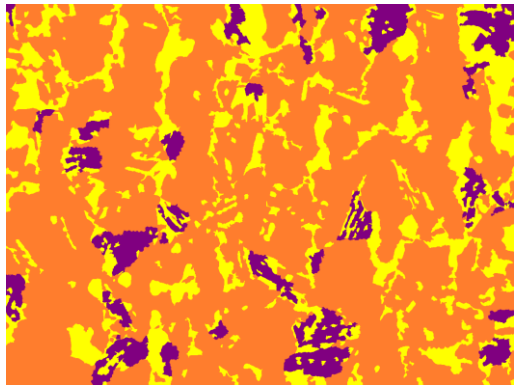
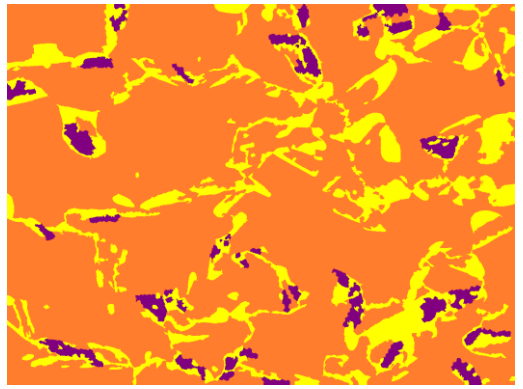
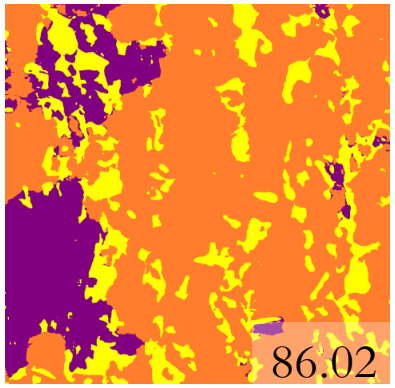
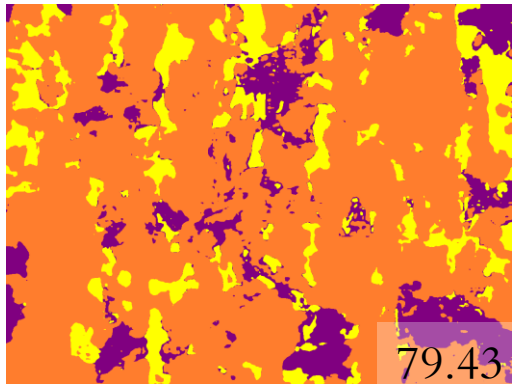
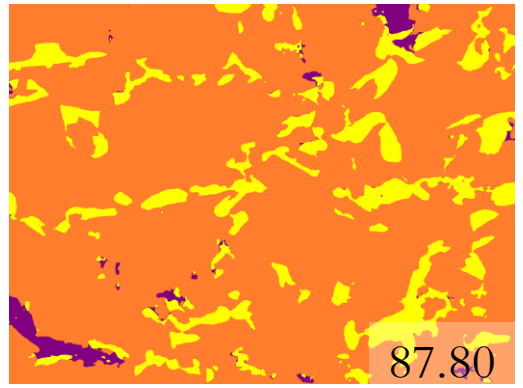
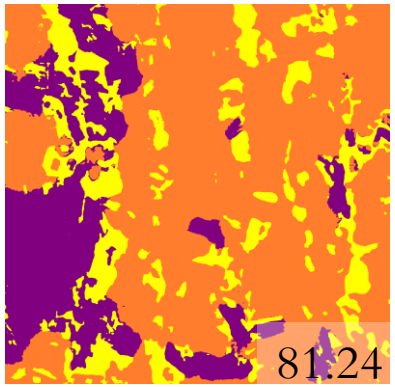
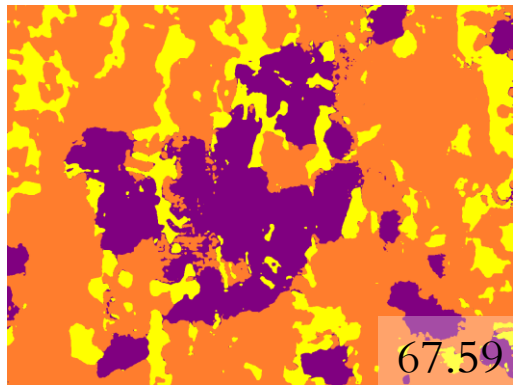
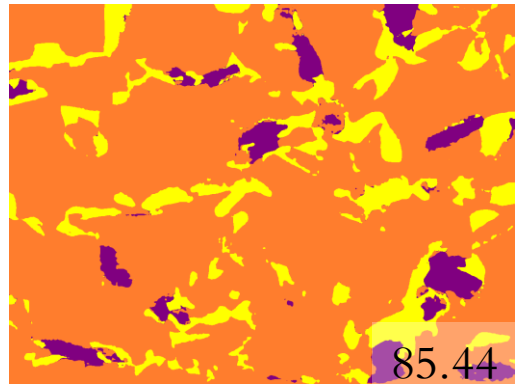
Preliminary Results

E-Type Test images
LoRA-SAM R-512
Dice

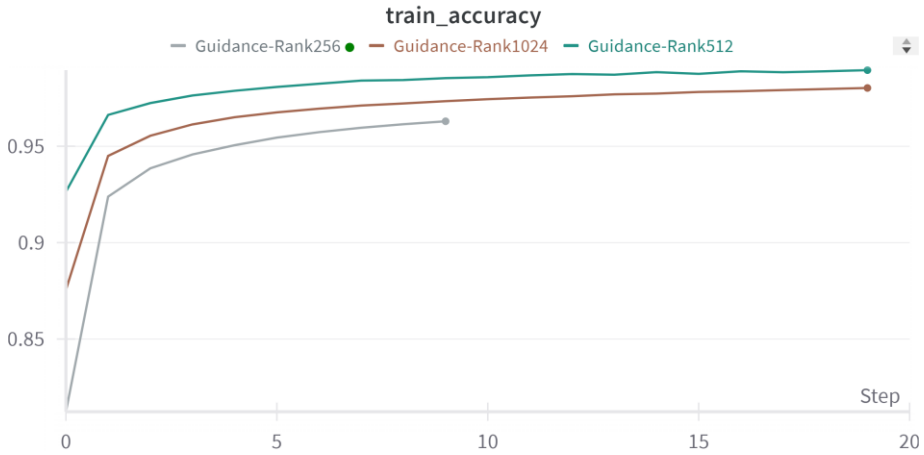
	× 2700	× 3000	× 5000
SEM			
GT			
Predicted	 86.02	 79.43	 87.80

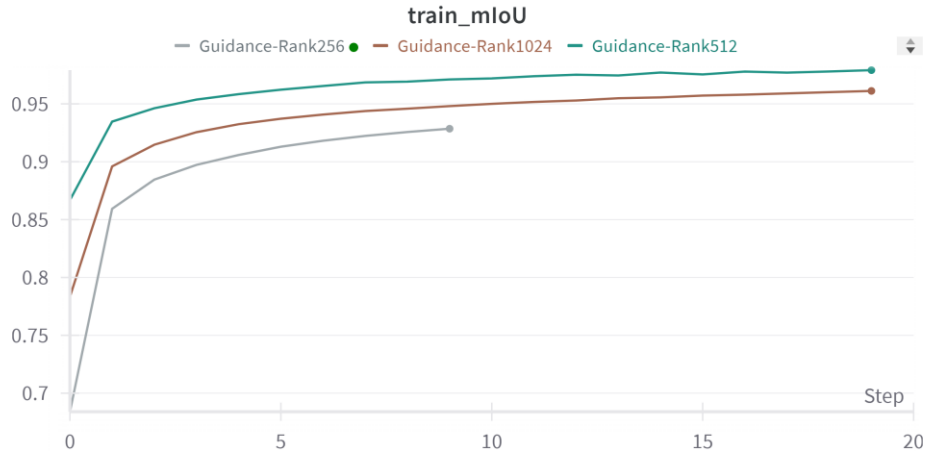
Preliminary Results

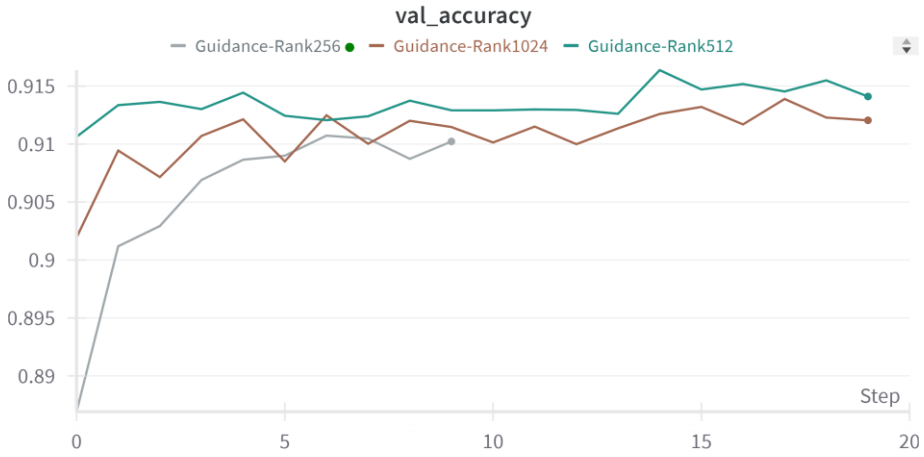
E-Type Test images
Dice

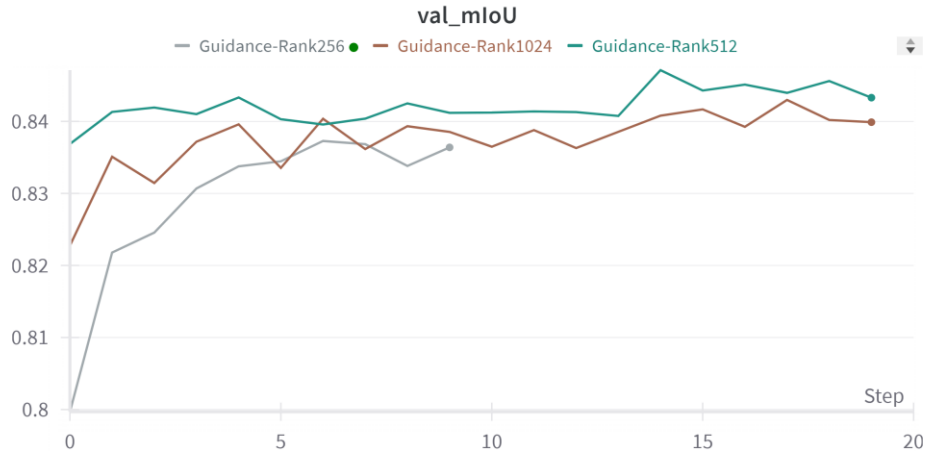
	× 2700	× 3000	× 5000	
GT				
Rank 512	 86.02	 79.43	 87.80	Mean – 84.42
Rank 1024	 81.24	 67.59	 85.44	Mean – 78.09

Preliminary Results



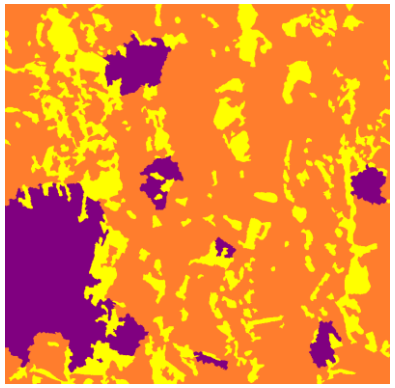
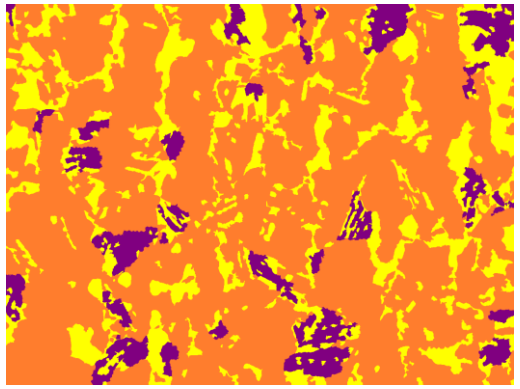
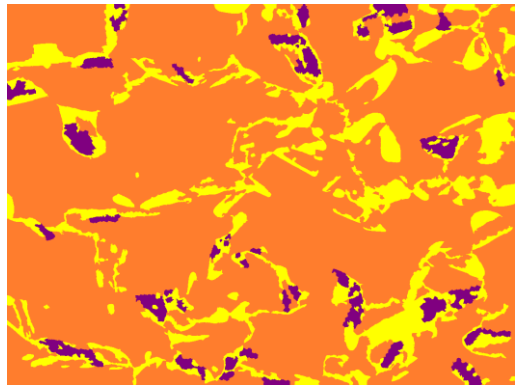
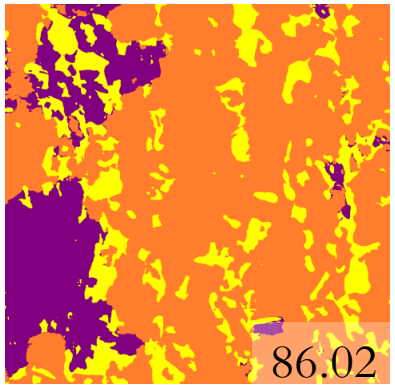
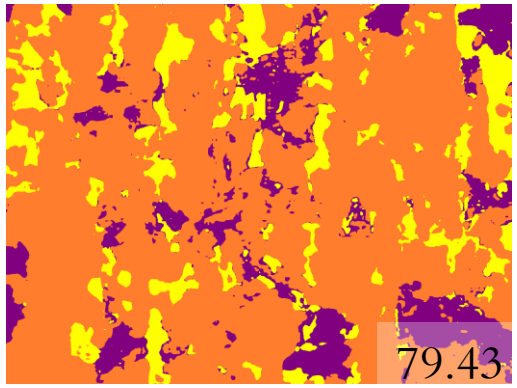
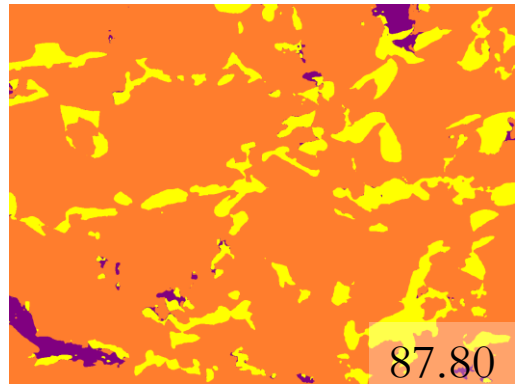
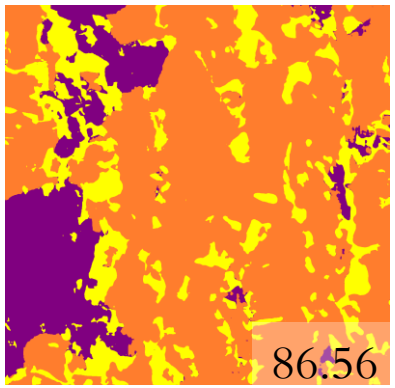
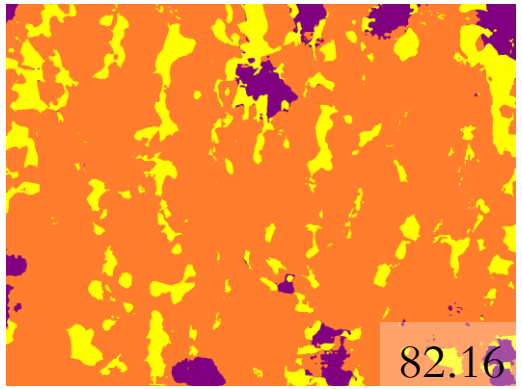
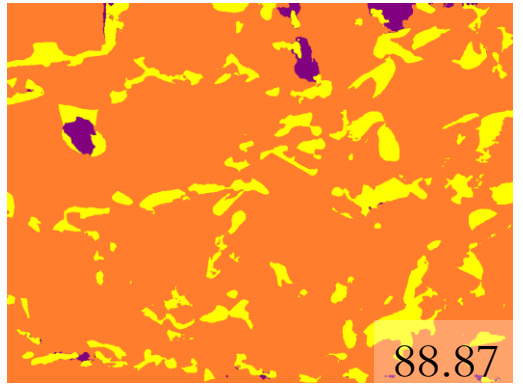




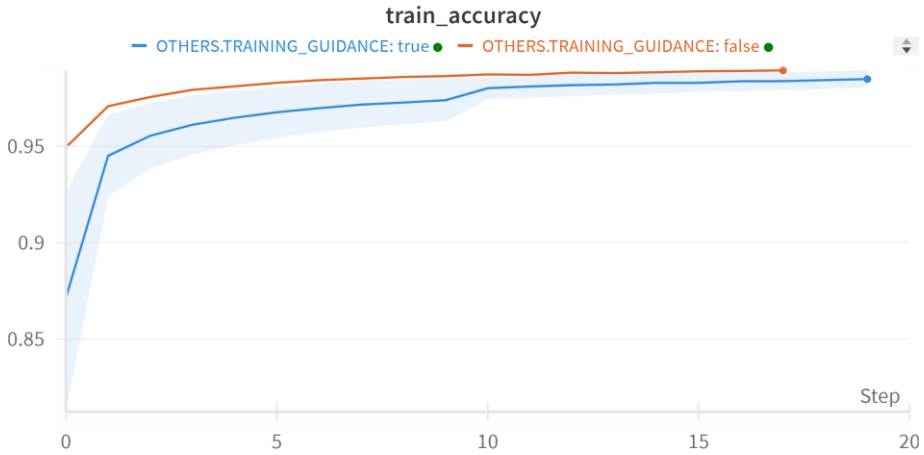


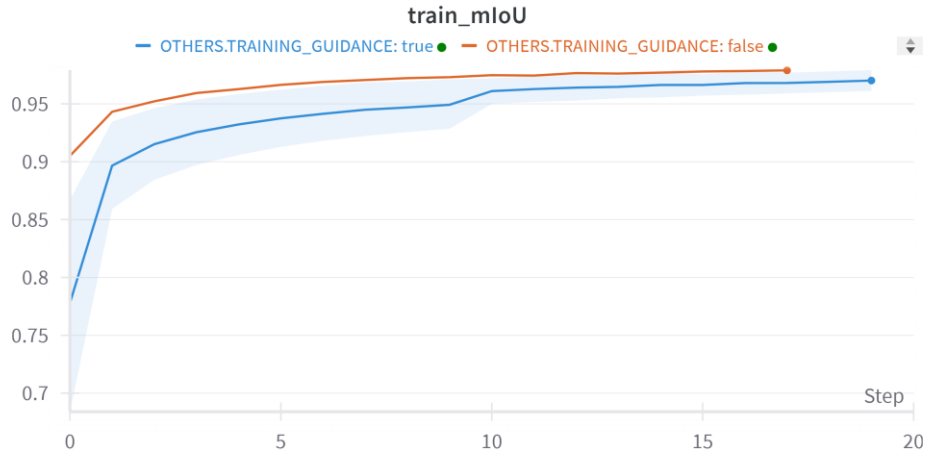
Preliminary Results

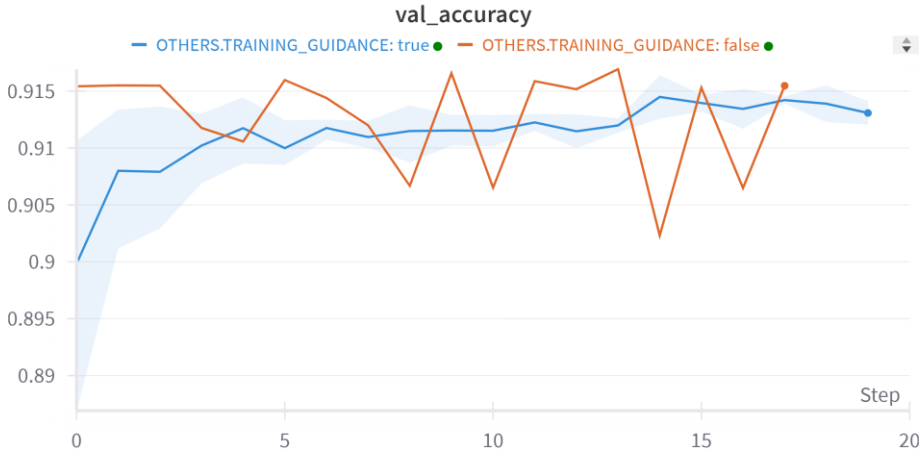
E-Type Test images
Dice

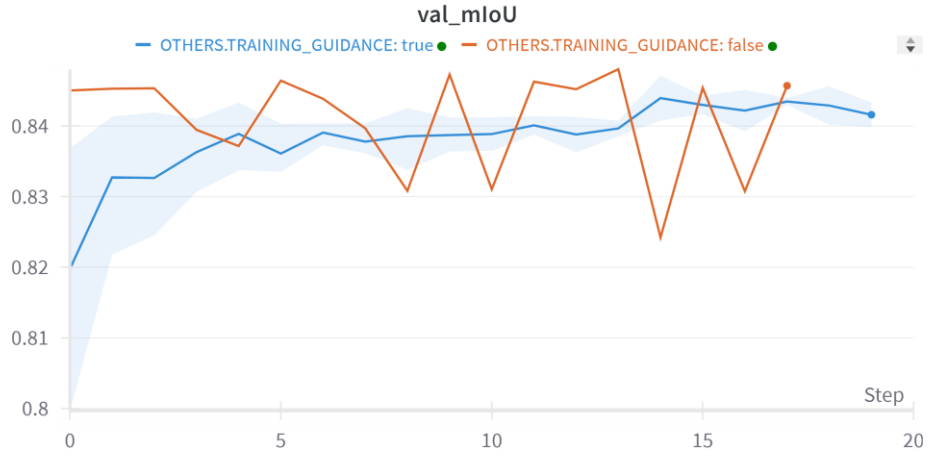
	× 2700	× 3000	× 5000	
GT				
BB Guidance Rank 512	 86.02	 79.43	 87.80	Mean – 84.42
No Guidance Rank 512	 86.56	 82.16	 88.87	Mean – 85.87

Preliminary Results





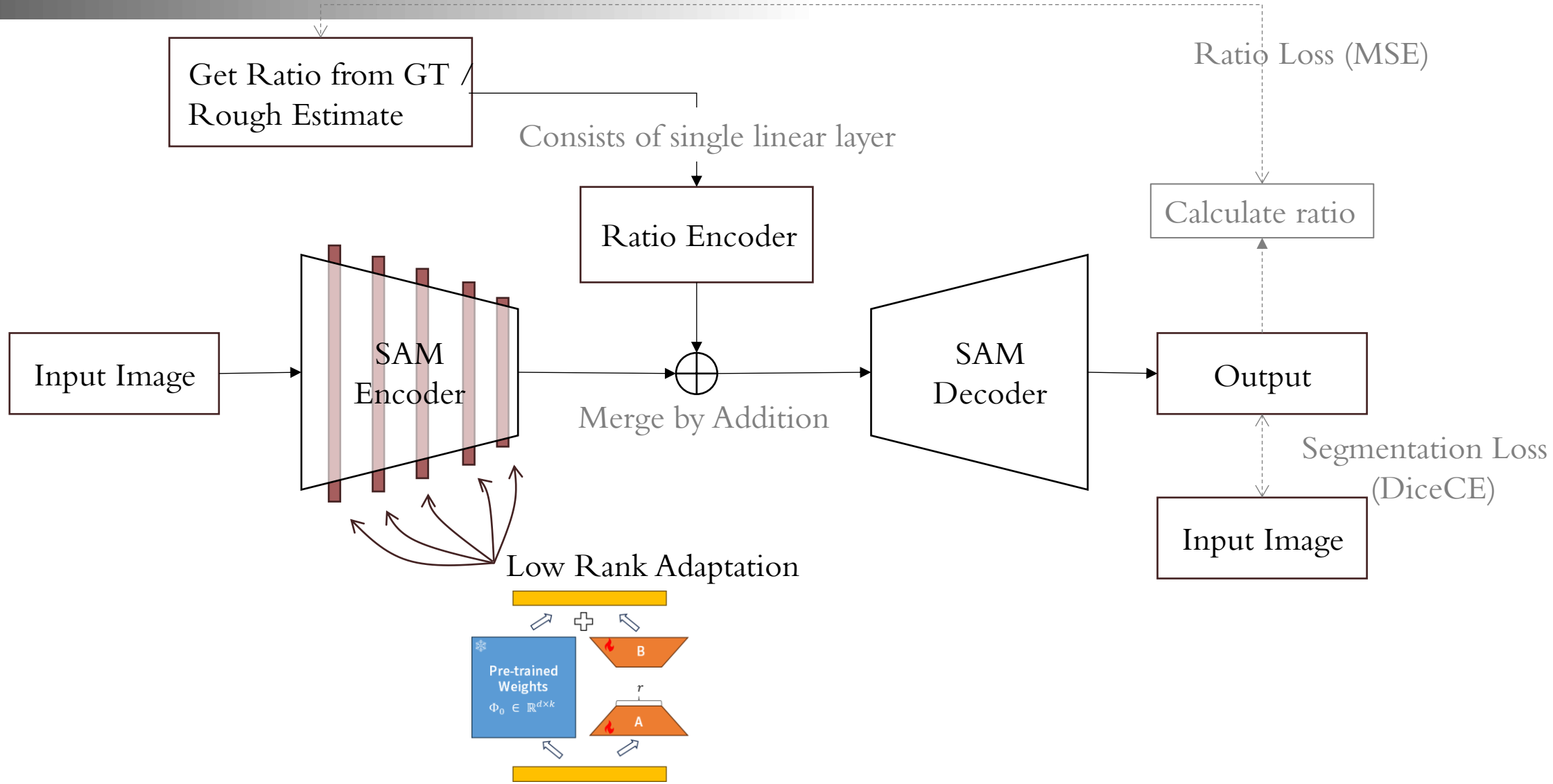




Preliminary Results

		× 2700	× 3000	× 5000	Avg.
Task-Specific	Enhanced U-Net3+	82.52	78.09	86.95	82.52
	Enhanced ELU-Net	**87.45**	**84.09**	**87.42**	**86.32**
VFM (0 Shot)	SAM	13.33	12.38	15.14	13.62
Adapted VFM	LoRA-SAM (256)	81.22	63.91	82.58	75.91
	LoRA-SAM (512)	<u>86.02</u>	<u>79.43</u>	<u>87.80</u>	<u>84.42</u>
	LoRA-SAM (1024)	81.24	67.59	85.44	78.09
	LoRA-SAM(512) [No Guidance]	<u>86.56</u>	<u>82.16</u>	<u>88.87</u>	<u>85.87</u>

Ratio Guidance



Attention Layers

Total Loss = 0.7 × segmentation loss + 0.3 × ratio loss

Experiments To-Do List

TO Do

- [ ] SAM-LoRA with BB Guidance - Rank 256 On-going
- [x] ~~SAM-LoRA with BB Guidance - Rank 512~~
- [x] ~~SAM-LoRA with BB Guidance - Rank 1024~~
- [ ] SAM-LoRA with BB Guidance - Rank 2046

- [ ] SAM-LoRA without Guidance - Rank 256
- [x] ~~SAM-LoRA without Guidance - Rank 512~~
- [ ] SAM-LoRA without Guidance - Rank 1024
- [ ] SAM-LoRA without Guidance - Rank 2046

- [ ] SAM-LoRA with ratio Guidance - Rank 256
- [ ] SAM-LoRA with ratio Guidance - Rank 512 On-going
- [ ] SAM-LoRA with ratio Guidance - Rank 1024
- [ ] SAM-LoRA with ratio Guidance - Rank 2046

- [ ] SAM-LoRA with BB & ratio Guidance - Rank 256
- [ ] SAM-LoRA with BB & ratio Guidance - Rank 512
- [ ] SAM-LoRA with BB & ratio Guidance - Rank 1024
- [ ] SAM-LoRA with BB & ratio Guidance - Rank 2046

- [ ] UNet
- [ ] UNet++
- [x] ~~UNet3+~~
- [x] ~~ELU-Net~~
- [ ] nnUNet

- [ ] UNet + SAM Fusion
- [ ] UNet++ + SAM Fusion
- [ ] UNet3+ + SAM Fusion
- [ ] ELU-Net + SAM Fusion

What other experiments would be beneficial?