Updates 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

05-23-2024

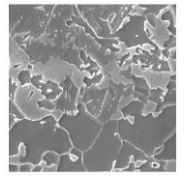
Computer Vision & Pattern Recognition Lab
kit 금오공과대학교 Kumoh National Institute of Technology

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	<u>86.32</u>	
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)	86.02	79.43	87.80	84.42	
	LoRA-SAM (1024)	81.24	67.59	85.44	78.09	
	LoRA-SAM (2048)	85.85	82.15	<u>88.37</u>	85.45	
	LoRA-SAM(512) [No Guidance]	<u>86.56</u>	<u>82.16</u>	<u>88.87</u>	85.87	
(a)	LoRA-SAM + Ratio + BB (512)	86.86	81.82	88.18	85.62	↑ 1.2%
(b)	LoRA-SAM + Ratio (512)	<u>86.65</u>	<u>83.67</u>	**89.48**	**86.60**	↑ 2.18%

(a) is the result that I presented on Tuesday, but I had misreported it. It was actually LoRA-SAM + BoundingBox + Ratio
(b) is the result with only LoRA-SAM and Ratio guidance.

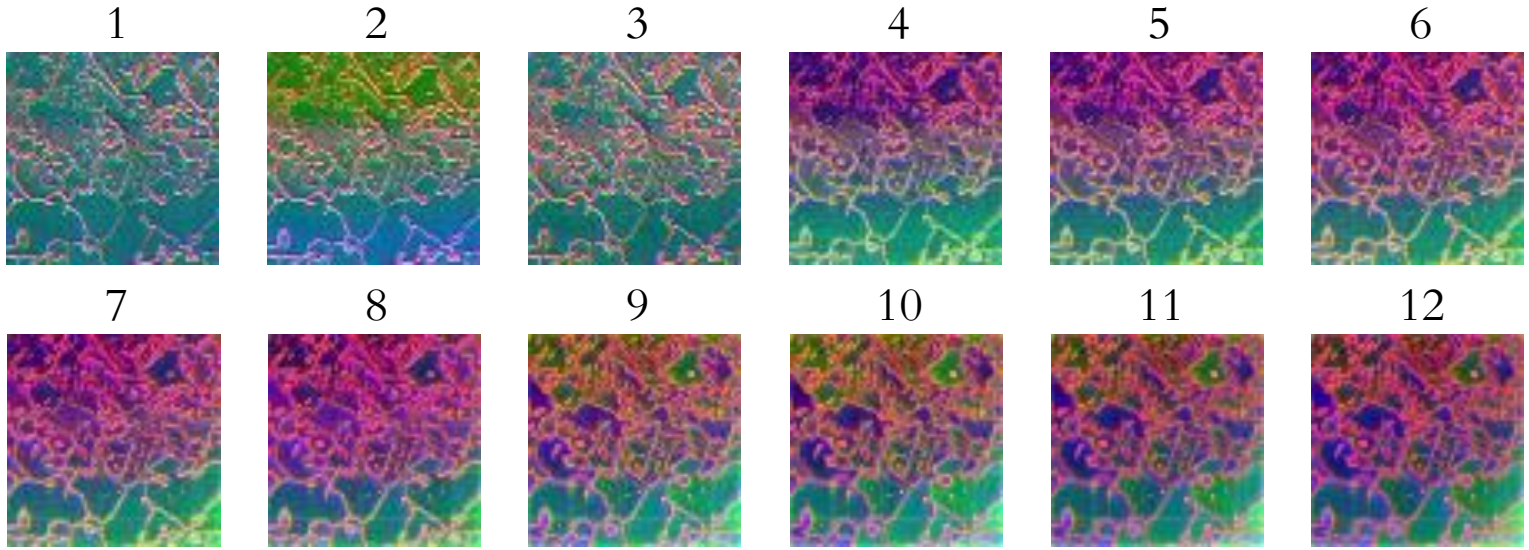
Visualization – Image Encoder



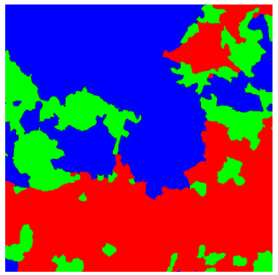
Input Image

SAM Layers

Baseline SAM



The model does not show any effect through the layers

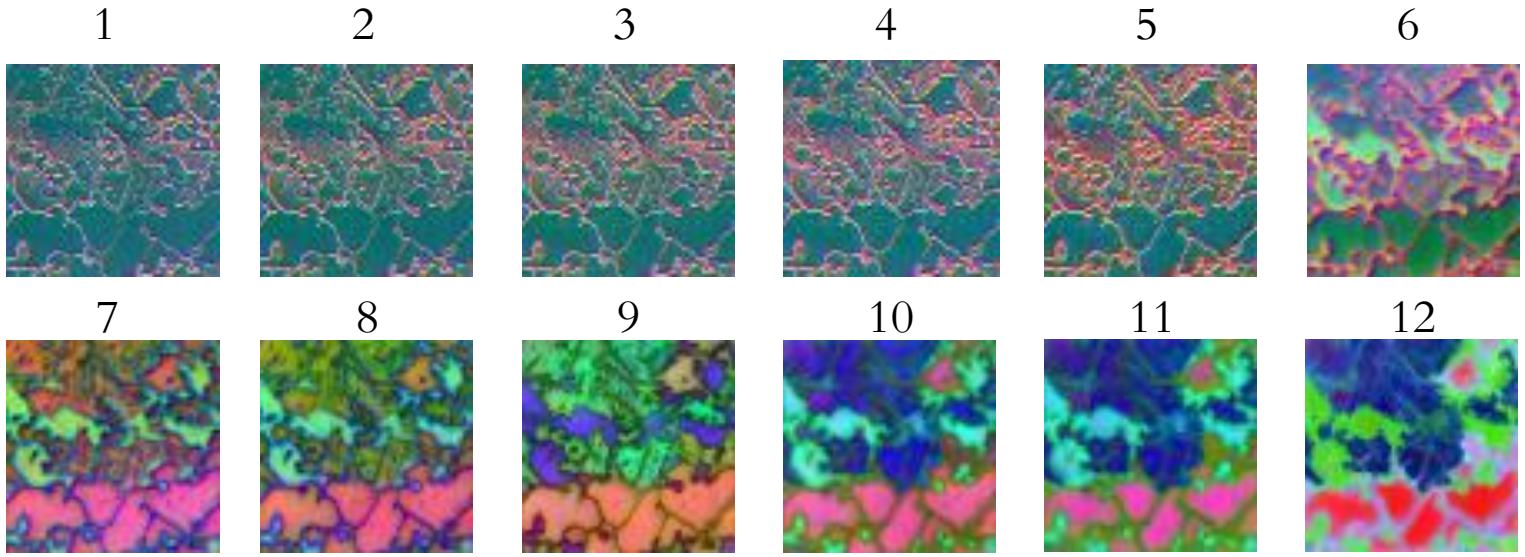


Target Image

$64, 64, 768 \rightarrow 64, 64, 3$

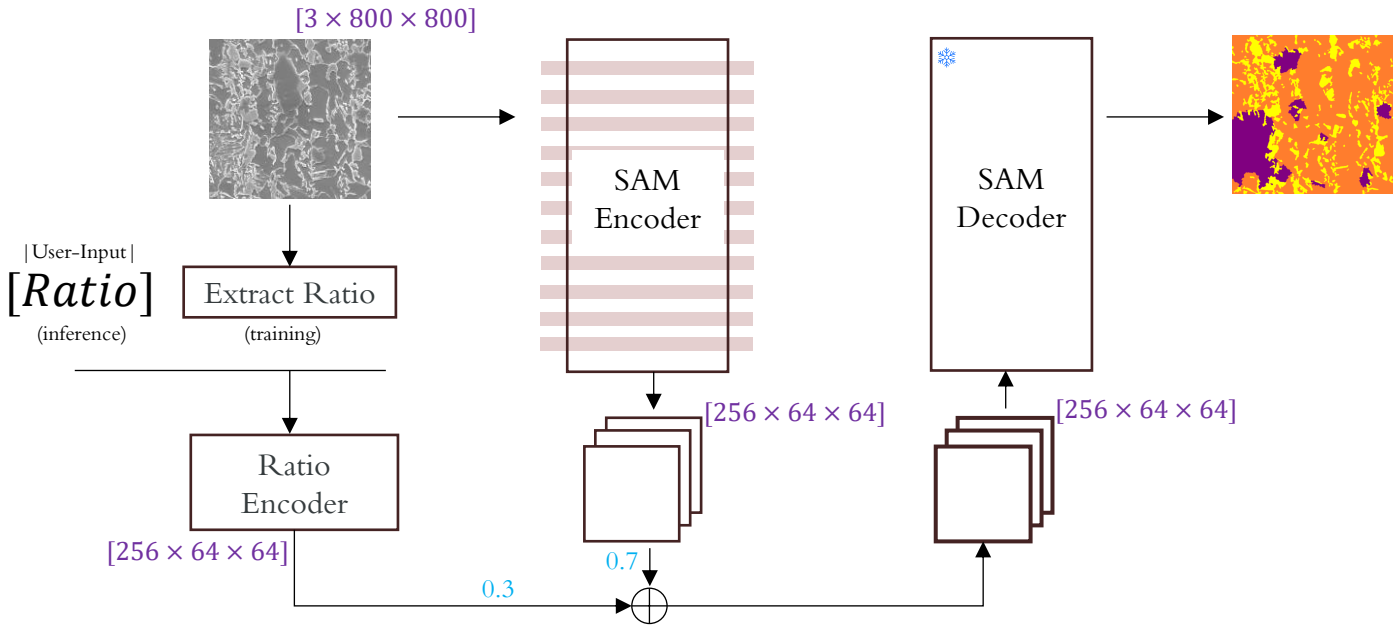
SAM-LoRA Layers

Our Adapted SAM



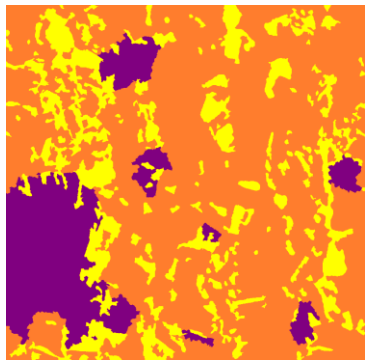
The model gradually shows distinction between phases

Ratio Guidance



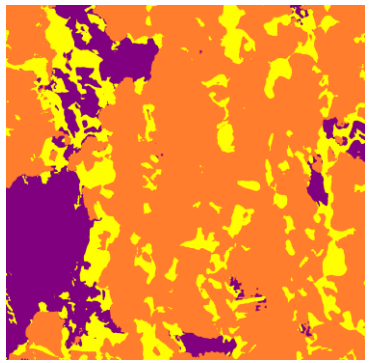
Effects of Ratio Encoding

Is the ratio encoder any effective?

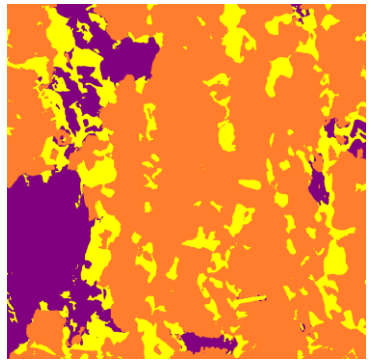


target

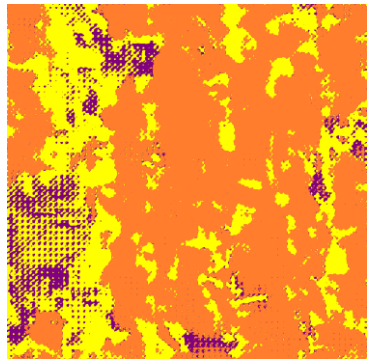
(a)



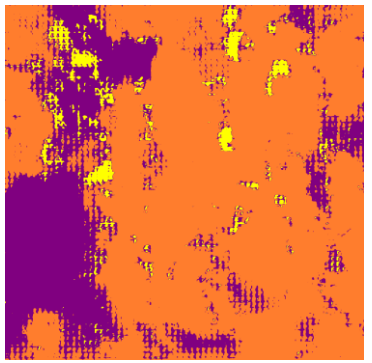
[0, 0, 0]
Dice → 84.89
(b)



[0.1, 0.7, 0.2]
Dice → 85.62
(c)



[0.1, 0.5, 0.4]
Dice → 56.41
(d)



[0.4, 0.5, 0.1]
Dice → 20.44
(e)

←Ratio provided during inference

Increasing the ratio of the martensite( ) and bainite ( ) also increases their respective amount in the prediction output.

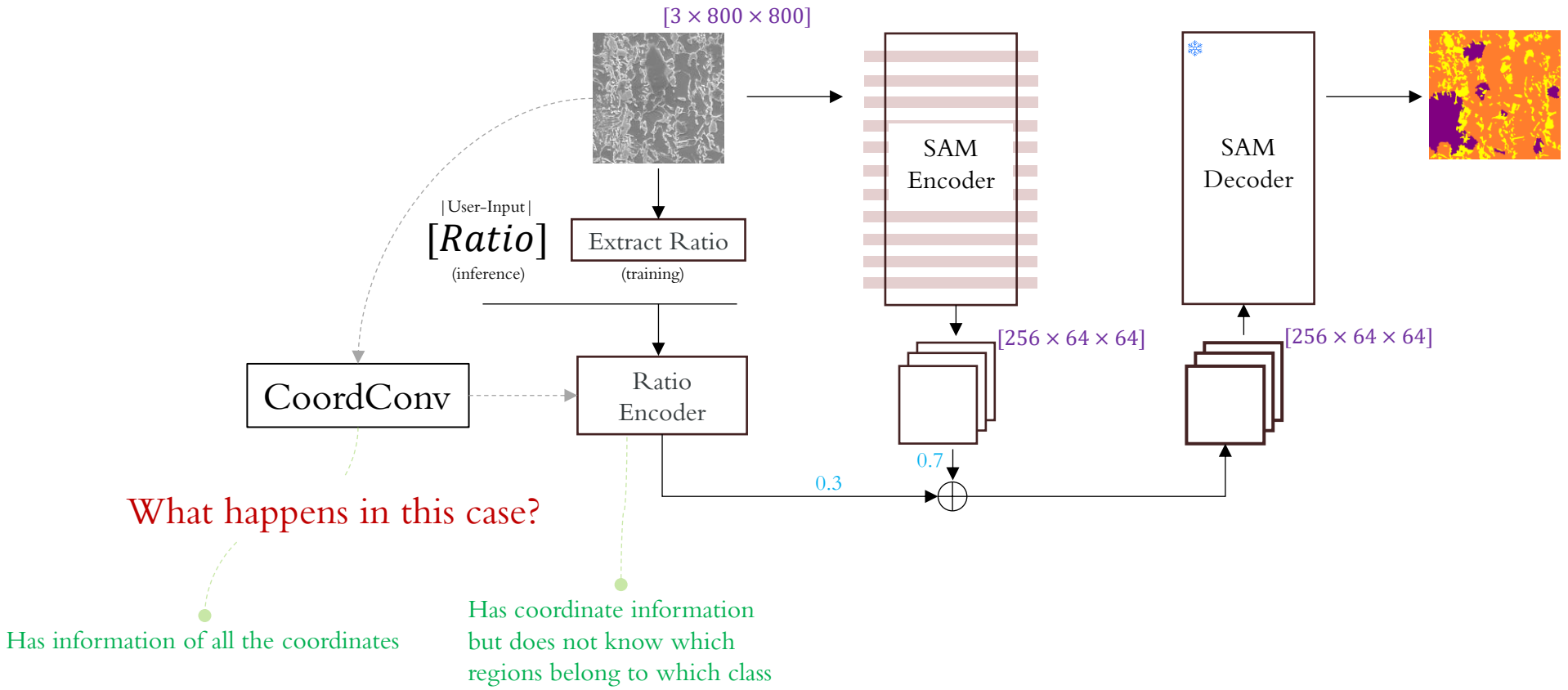
So, we can say that ratio encoder is effective in controlling the amount of phase present but fails to localize the changes – that is where the phase components should be present.

This is to be expected as the ratio encoder output is just added into the SAM's image encoder output without any co-relation

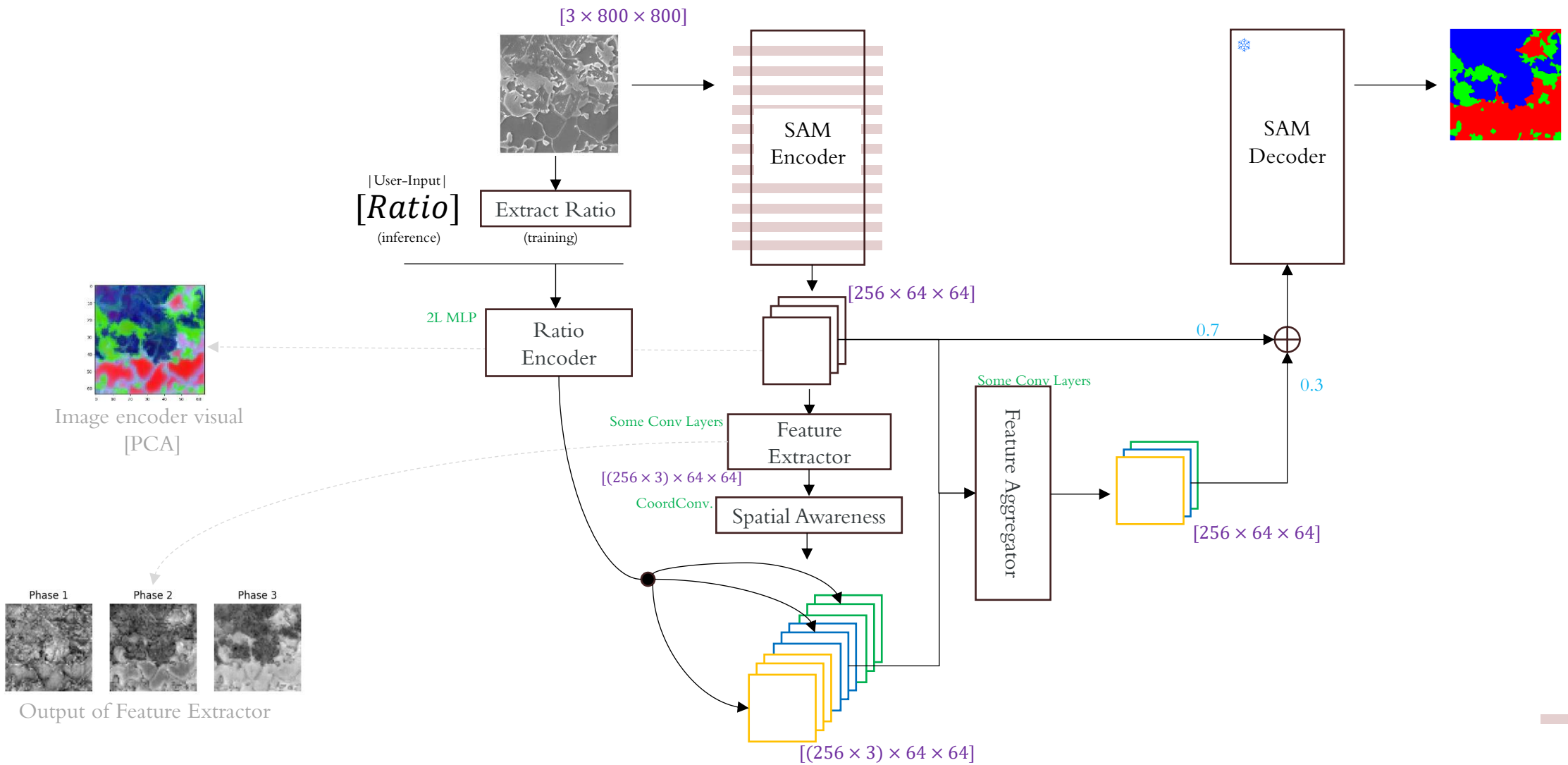
(Spatial) Ratio Encoding

Is the ratio encoder any effective? Yes, but no spatial awareness

How can we add spatial awareness? → CoordConv? HOW?



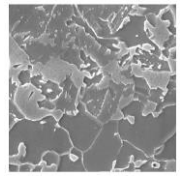
Ratio Guidance with Spatial Awareness



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	<u>86.32</u>
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)	86.02	79.43	87.80	84.42
	LoRA-SAM (1024)	81.24	67.59	85.44	78.09
	LoRA-SAM (2048)	85.85	82.15	<u>88.37</u>	85.45
	LoRA-SAM(512) [No Guidance]	<u>86.56</u>	<u>82.16</u>	<u>88.87</u>	85.87
	LoRA-SAM + Ratio + BB (512)	86.86 ↑ 0.84%	81.82 ↑ 2.39%	88.18 ↑ 0.38%	85.62 ↑ 1.2%
	LoRA-SAM + Ratio (512)	<u>86.65</u> ↑ 0.63%	<u>83.67</u> ↑ 4.24%	89.48 ↑ 1.68%	86.60 ↑ 2.18%
	LoRA-SAM + Ratio (Spatial-v1)(512)	<u>86.74</u> ↑ 0.72%	<u>83.25</u> ↑ 3.82%	**89.91** ↑ 2.11%	**86.63** ↑ 2.21%

Visualization – Image Encoder

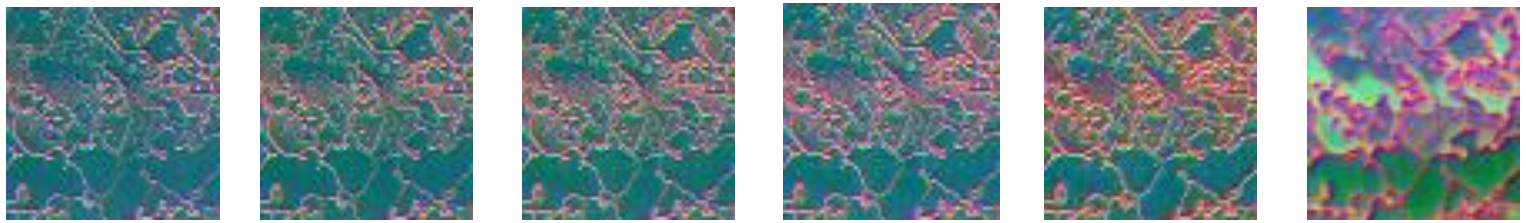


Input Image

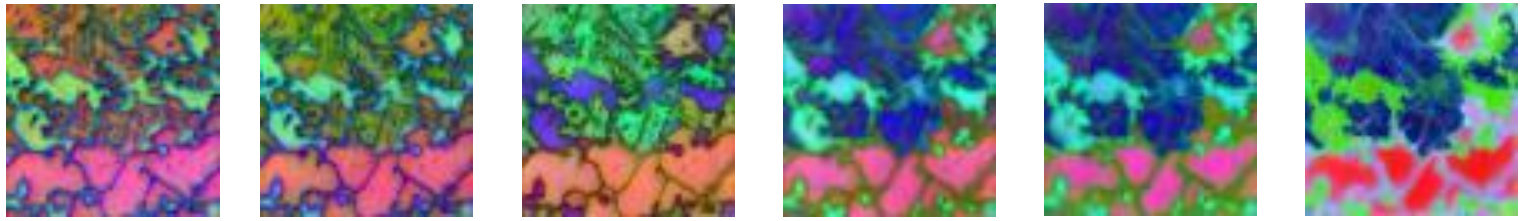
SAM-LoRA Layers

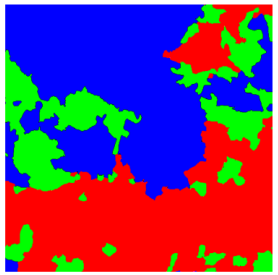
Our Adapted SAM

1 2 3 4 5 6



7 8 9 10 11 12

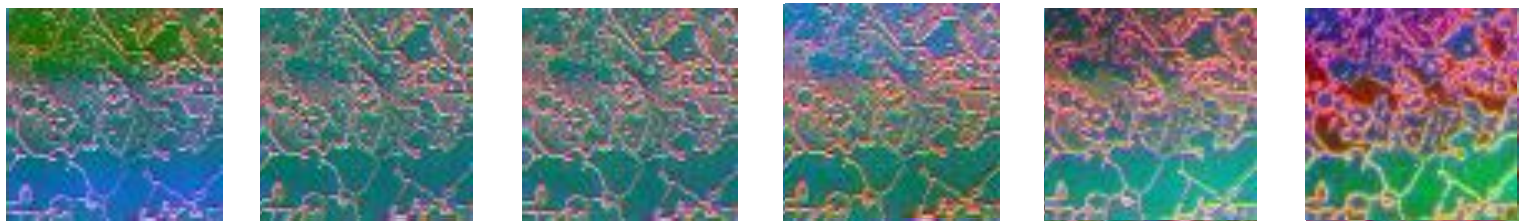




Target Image

$64, 64, 768 \rightarrow 64, 64, 3$

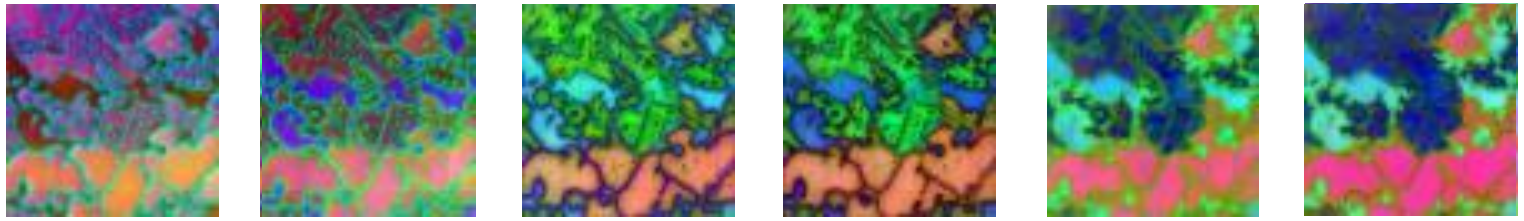
1 2 3 4 5 6



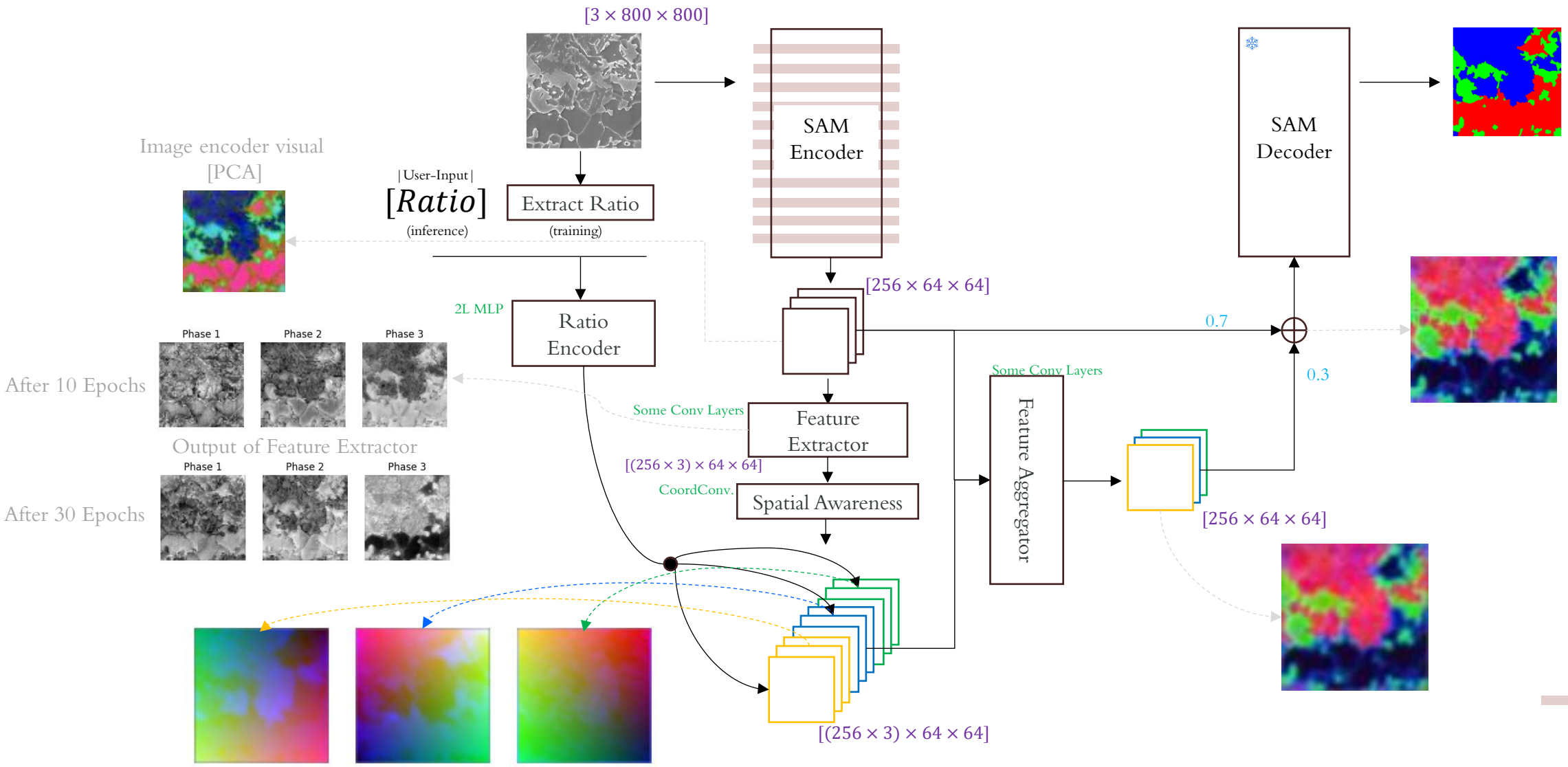
SAM-LoRA Layers

Spatial-Ratio v1

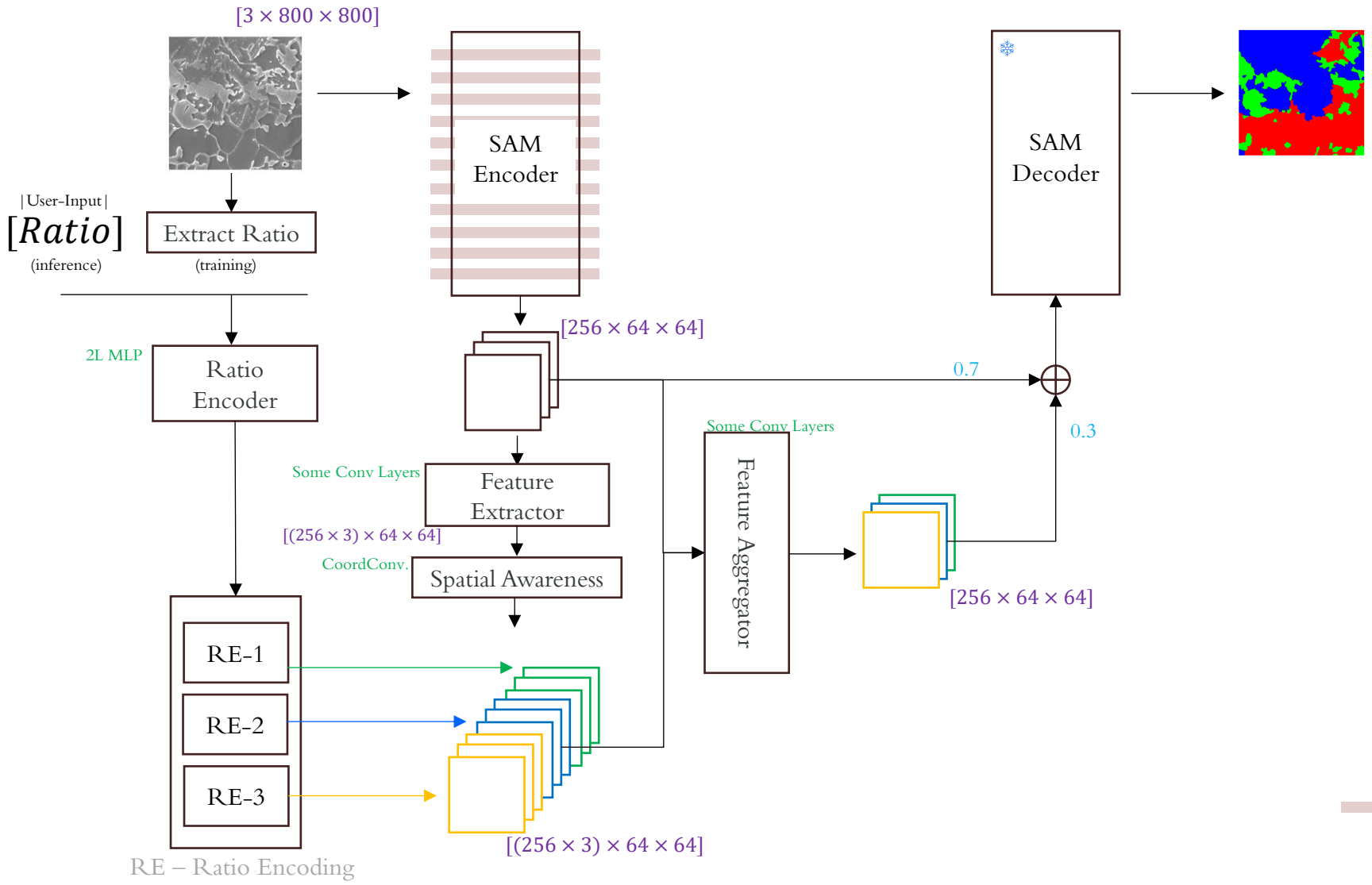
7 8 9 10 11 12



Ratio Guidance with Spatial Awareness



Ratio Guidance with Spatial Awareness



LoRA

Ratio Guidance with Spatial Awareness