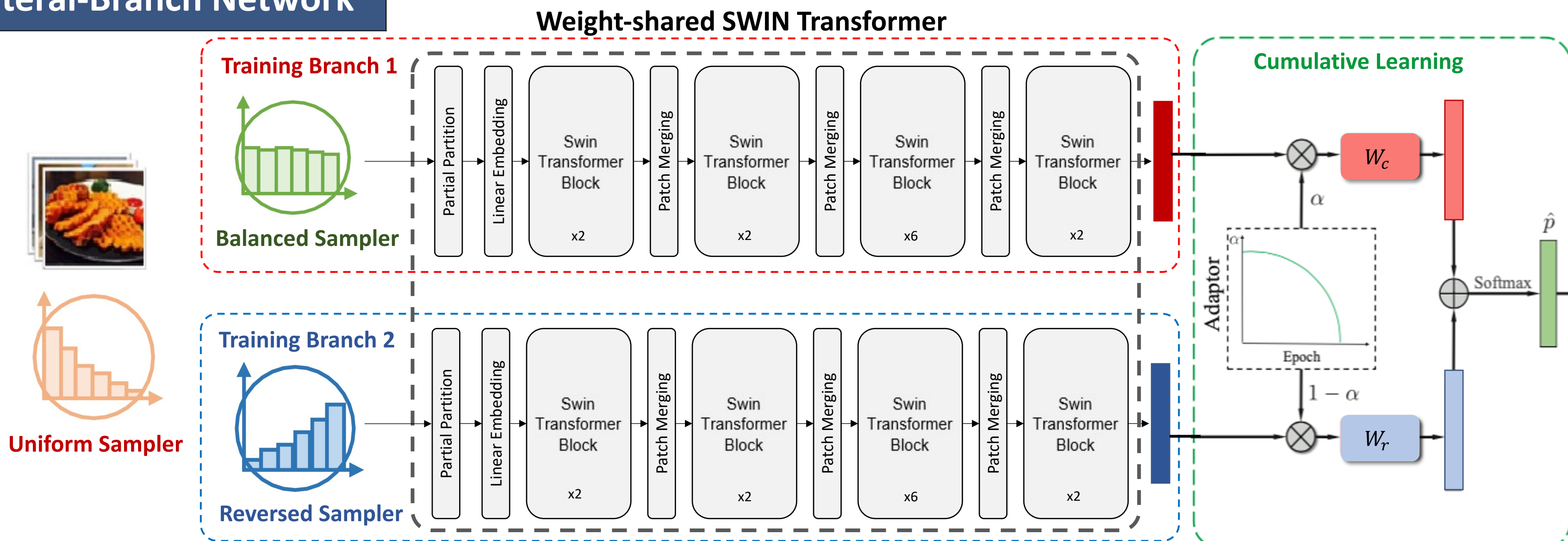
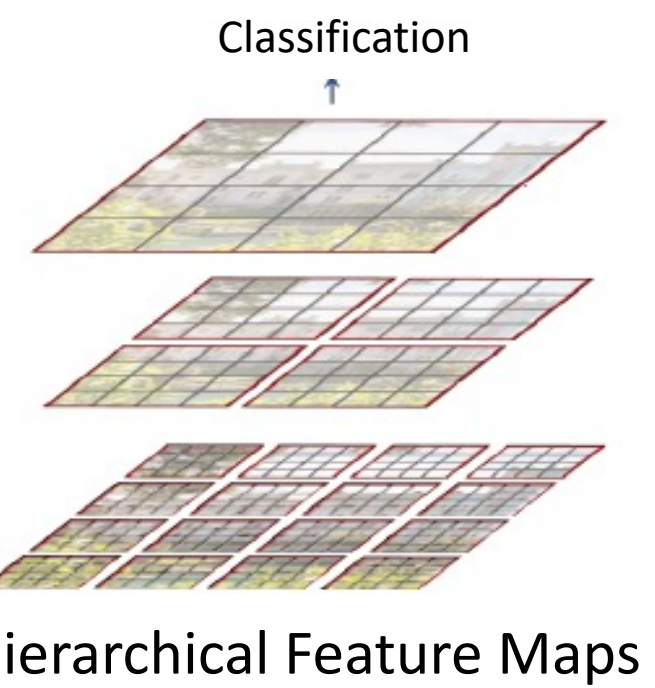
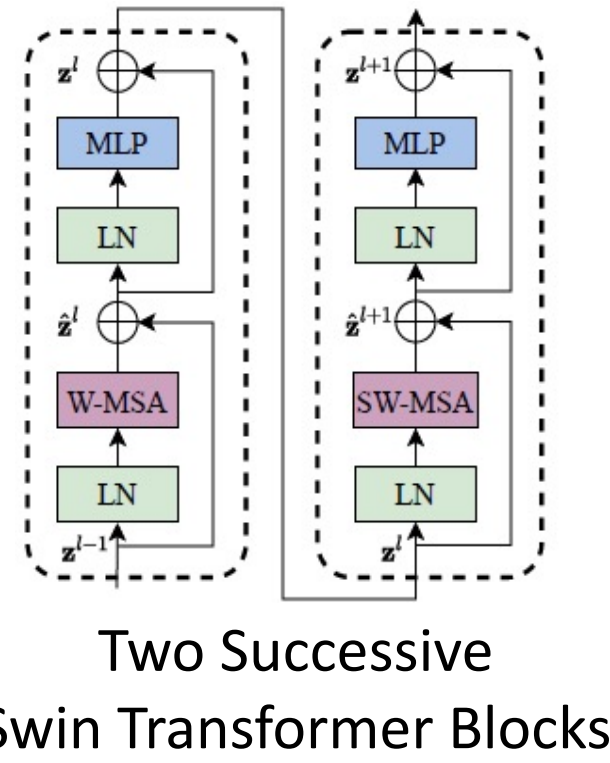


Model Architecture

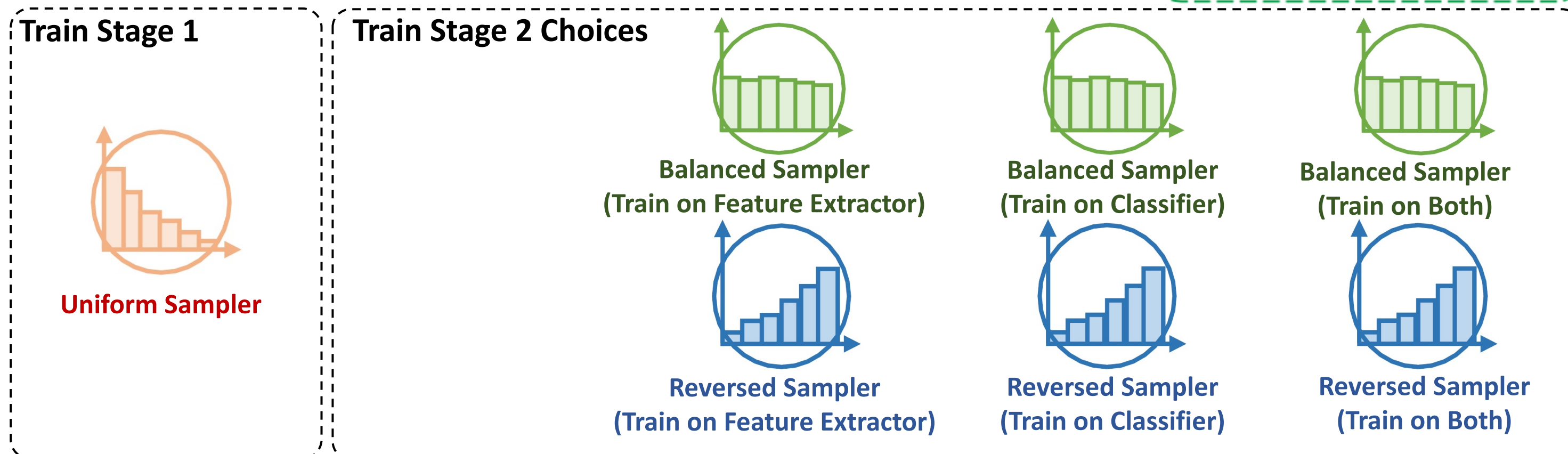
Bilateral-Branch Network



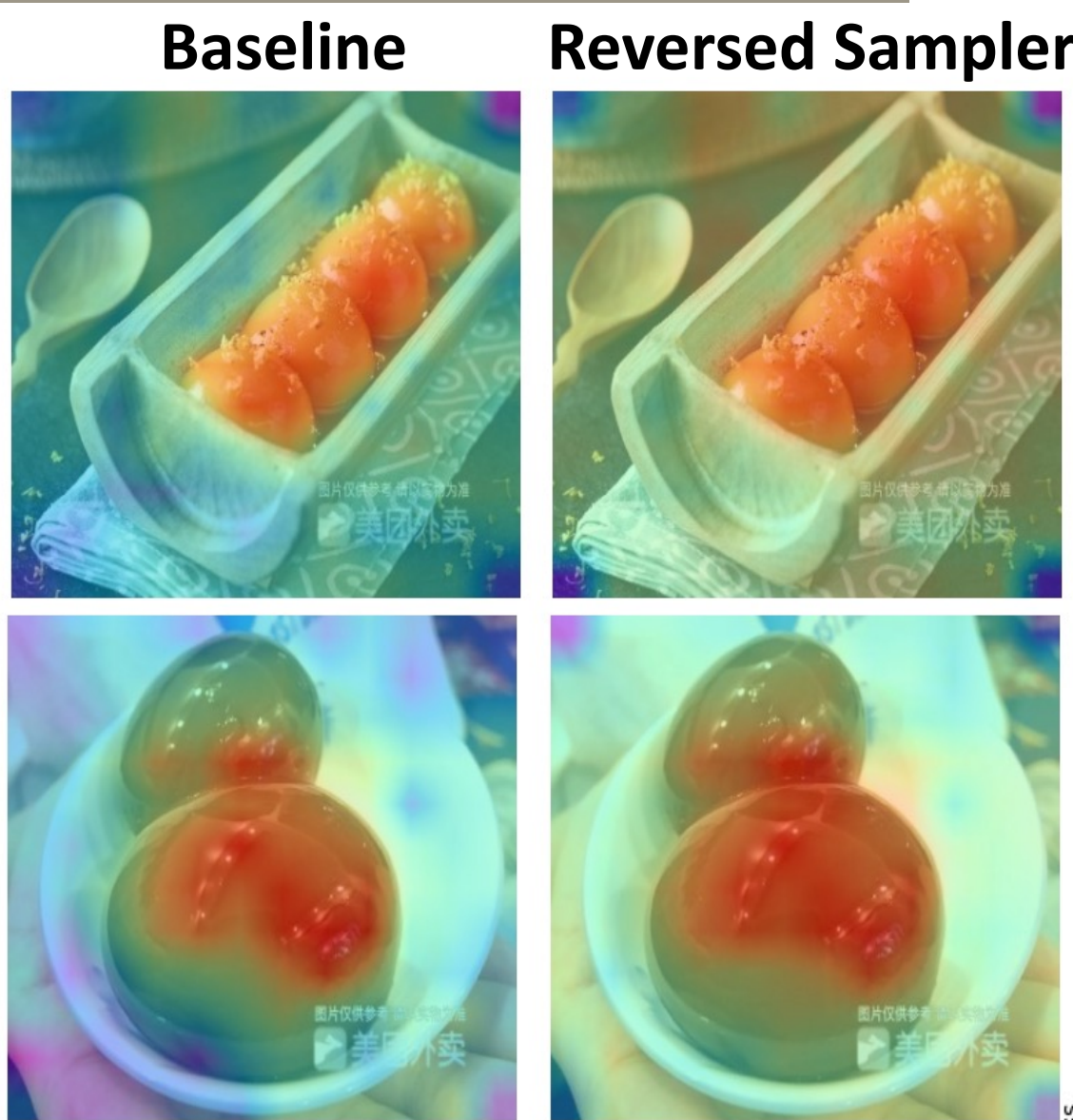
Swin Transformer Concept



Decoupling

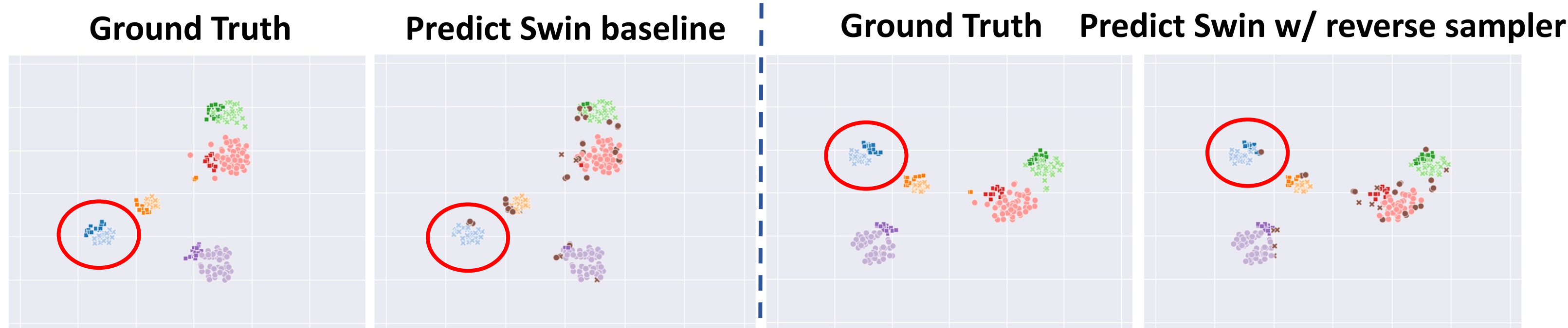


Visualization

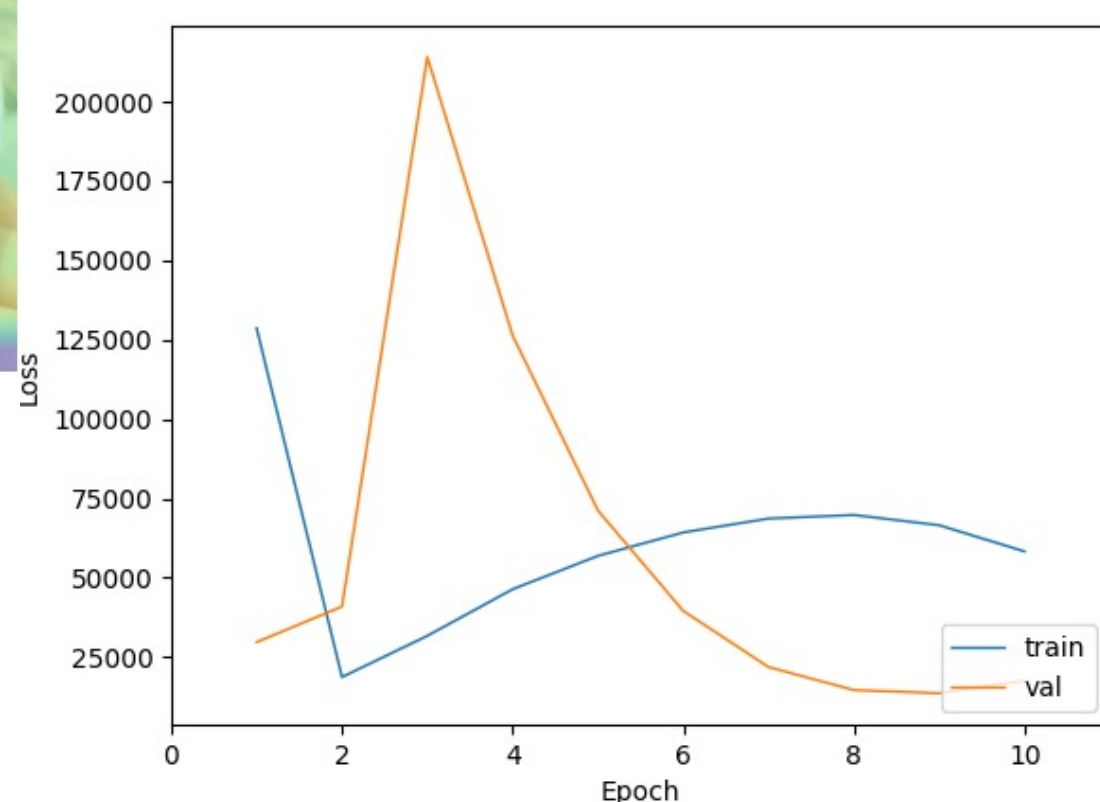


Attention Map (Decouple)

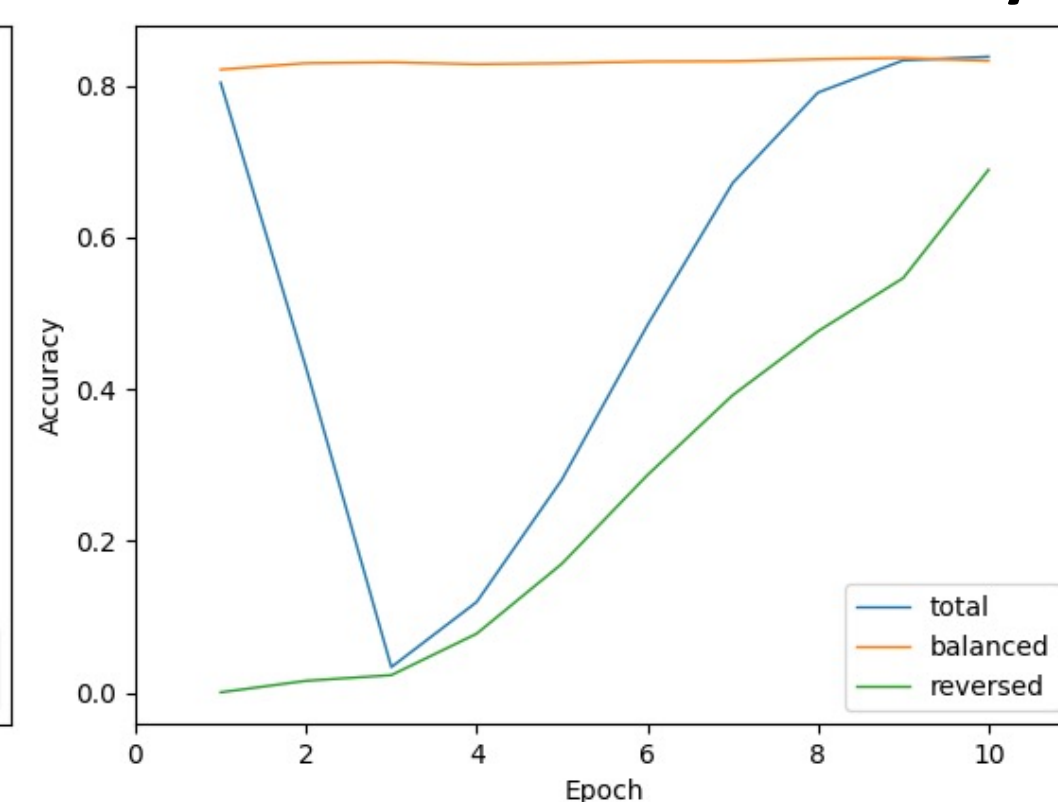
TSNE (Decouple)



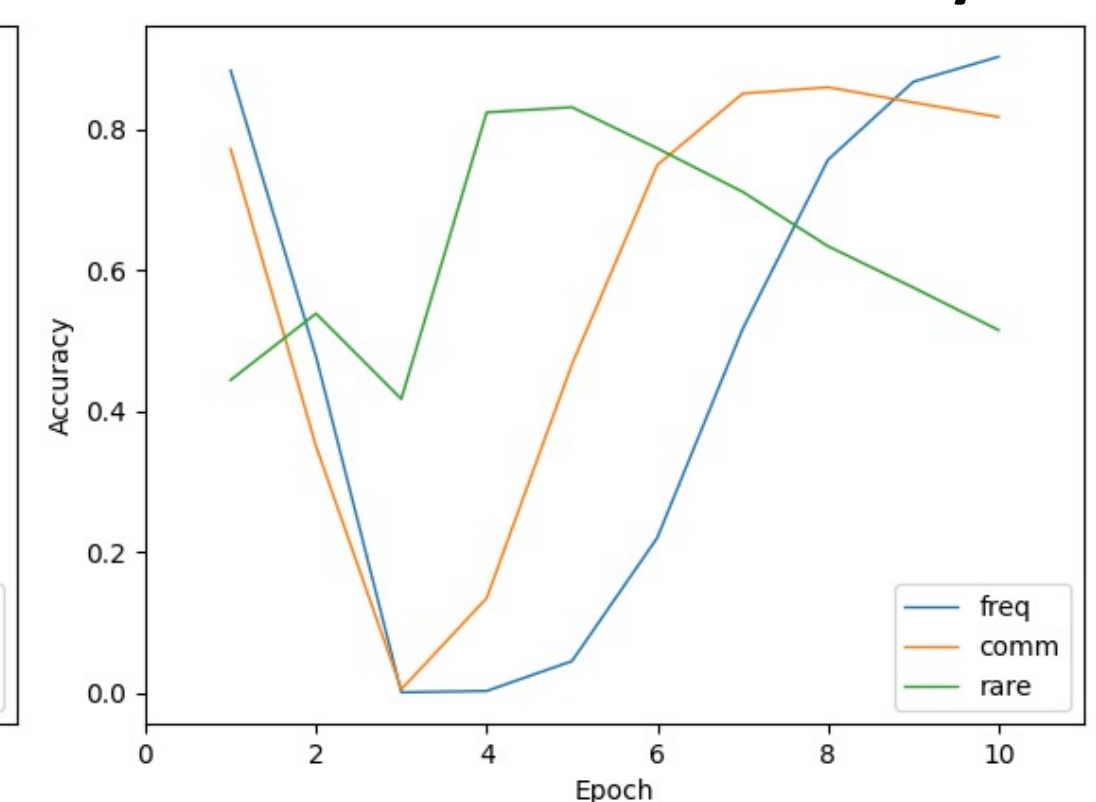
Loss



Branch Validation Accuracy



Track Validation Accuracy



Loss and Accuracy (Bilateral-Branch Network)

Experiments

Classifier

Backbone	Main	Frequent	Common	Rare
ResNest269	67.53%	86.68%	64.52%	20.46%
EfficientNet-B7	64.99%	82.65%	63.03%	21.79%
ViT	60.56%	83.61%	56.38%	8.34%
Swin Transformer	73.76%	90.68%	72.93%	31.20%

Bilateral-Branch Network

Backbone	Sampler	Main	Frequent	Common	Rare
Uniform	U				
	B				
Balanced	U				

Decoupling

Method	Sampler	Main	Frequent	Common	Rare
Stage 1 Training	B				
	U				
Stage 2 Trained on Feature	B				
	R				
Stage 2 Trained on MLP	B				
	R				
Stage 2 Trained on Both	B				
	R				

B: Balanced sampler / U: Uniform sampler / R: Reversed sampler