國立臺灣大學 National Taiwan University
NTU DLCV Fall 2021

Long Tail Annihilator

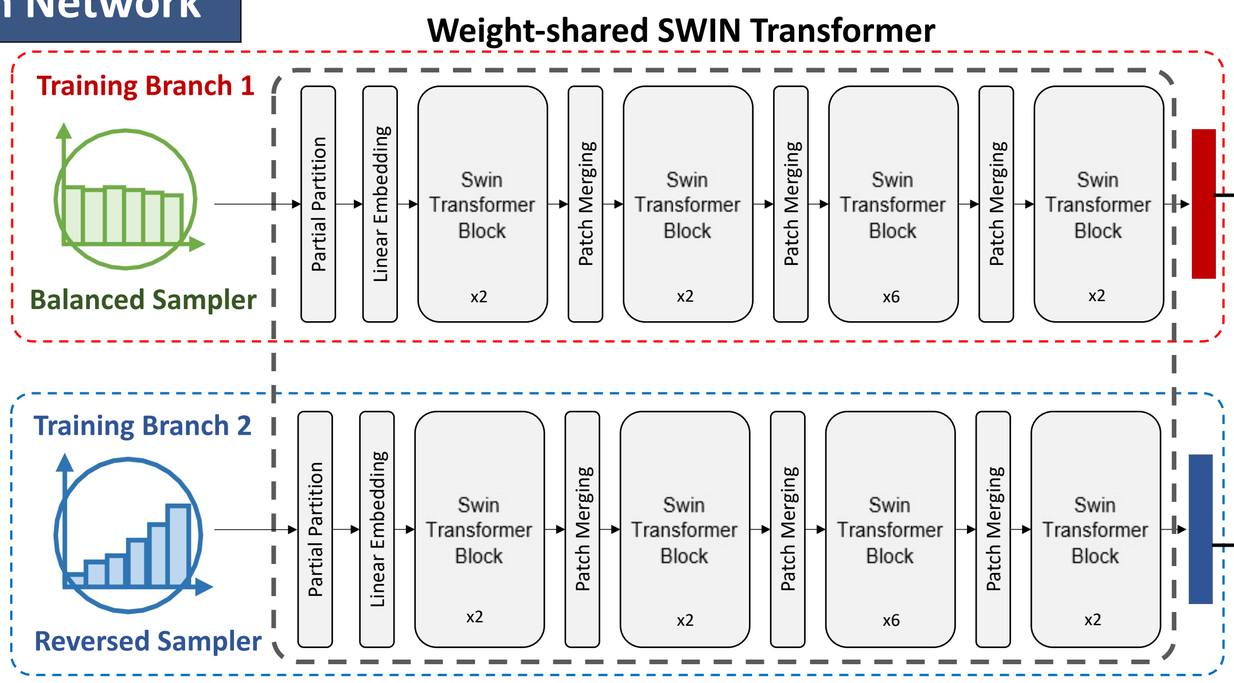
Yi-Hsien Lin D06943006
Hung-Yu Shu R09943021
Cheng-You Tsai R10943014
Chih-Ying Liu B07901039
Pin-Hua Lee B07303024

Model Architecture

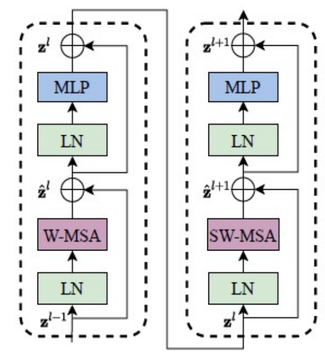
Bilateral-Branch Network

Uniform Sampler

Reversed Sampler

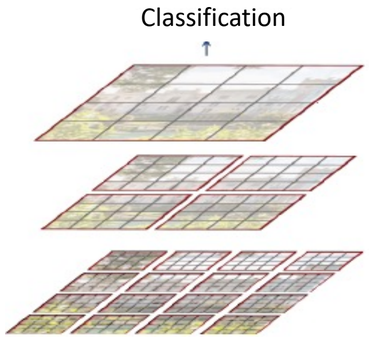


Swin Transformer Concept



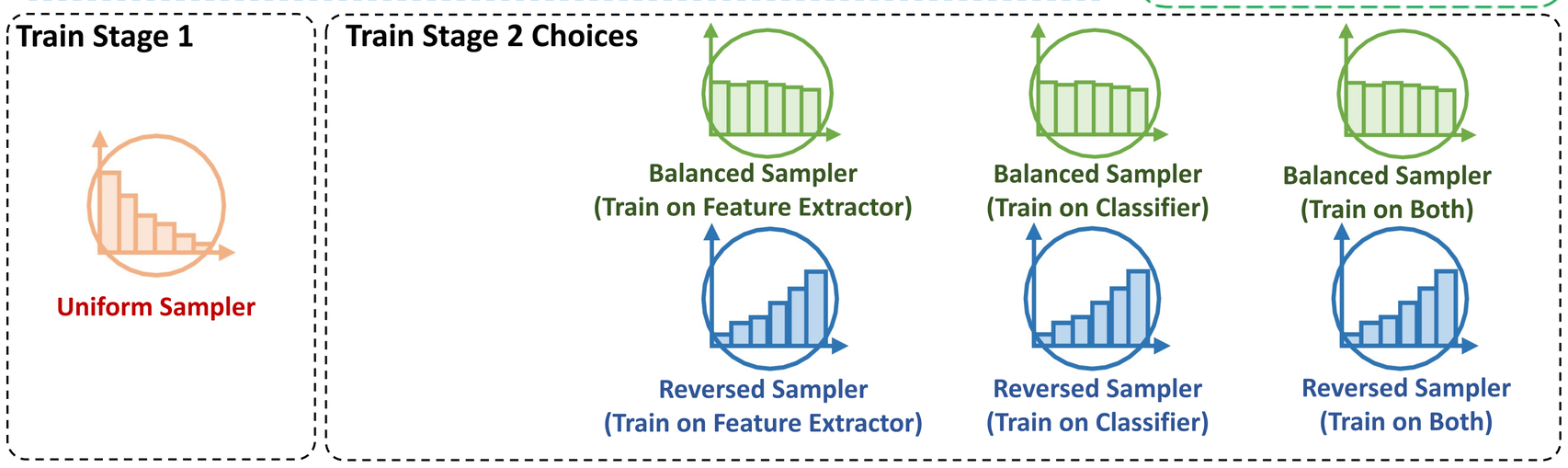
Two Successive Swin Transformer Blocks

Classification



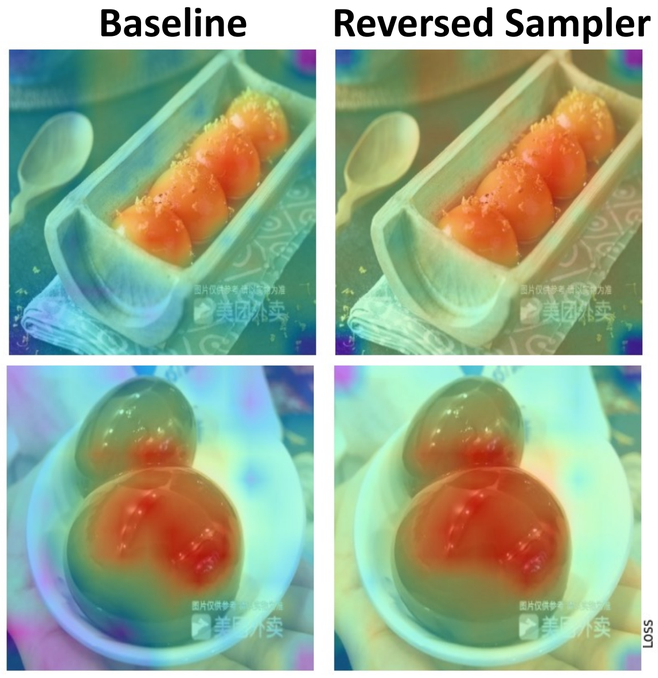
Hierarchical Feature Maps

Decoupling

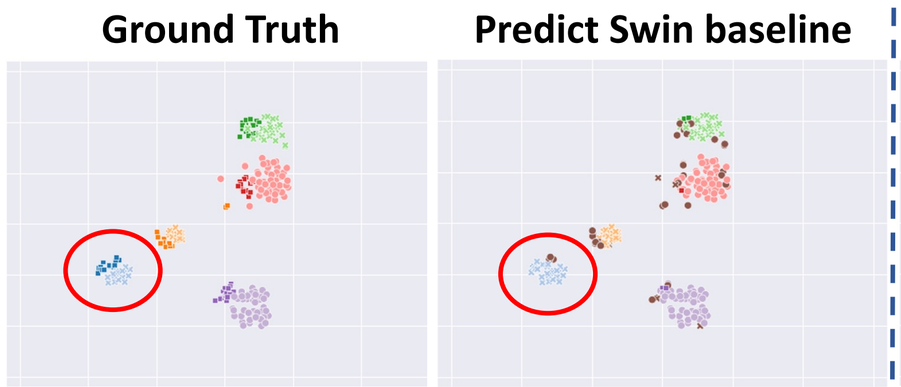


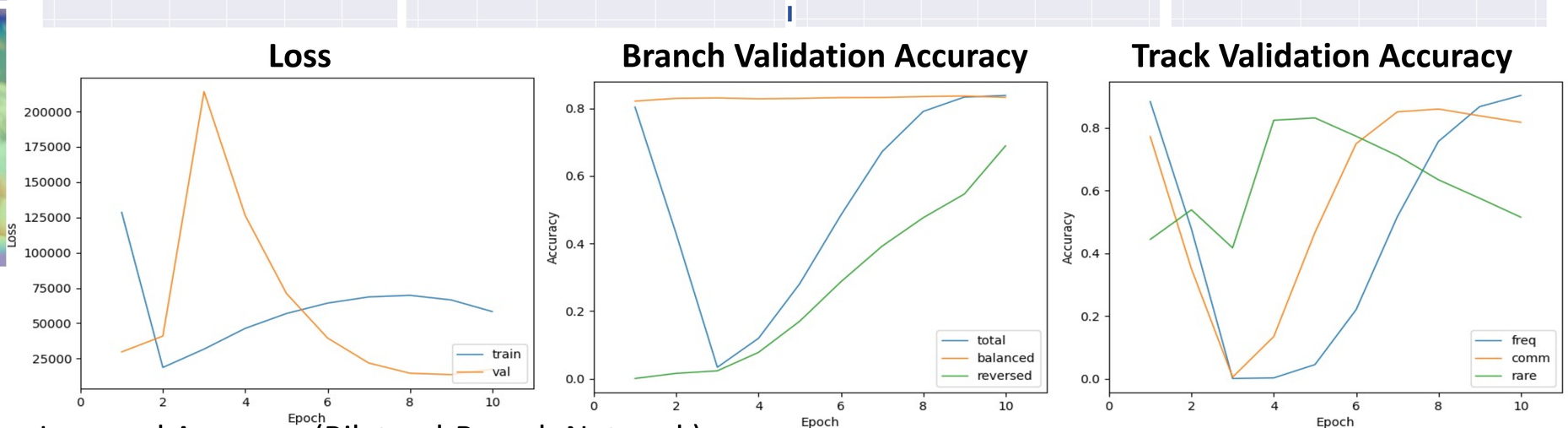
Visualization

TSNE (Decouple)



Attention Map (Decouple)





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/R	78.08%	89.67%	80.70%	26.40%
	B/R	81.24%	89.73%	81.78%	50.69%
Balanced	U/R	72.40%	91.20%	62.08%	3.30%
Balanced	B/R	78.48%	77.25%	84.30%	53.90%
Uniform1	▲ B/R	78.63%	75.81%	84.95%	56.28%

Decoupling

Method	Sampler	Main	Frequent	Common	Rare
Stage 1 Training	U1	65.40%	86.64%	51.10%	2.22%
	U10	73.76%	90.68%	72.93%	31.20%
Stage 2 Trained on Feature	B	79.20%	90.53%	78.50%	45.04%
	R	80.16%	88.53%	79.98%	53.16%
Stage 2 Trained on MLP	B	79.49%	90.27%	78.78%	48.43%
	▲ R	79.42%	88.92%	78.82%	58.93%
Stage 2 Trained on Both	B	77.12%	90.02%	77.00%	38.22%
	▲ R	79.30%	86.39%	79.34%	55.35%
Decoupling Resnet269 ▲		70.80%	70.09%	72.73%	59.37%
Test Time Augmenation + Ensemble 4 models ▲		85.53%	88.07%	85.04%	61.73%

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