



Input

Patch Partition

Linear Embedding

Swin Transformer
Block $\times 2$

Patch Merging

Swin Transformer
Block $\times 2$

Patch Merging

Swin Transformer
Block $\times 18$

Patch Merging

Swin Transformer
Block $\times 2$

Avgpool

2.1 3.2 4.9 5.5 7.5 9.0

Original Features and Labels

Sampling

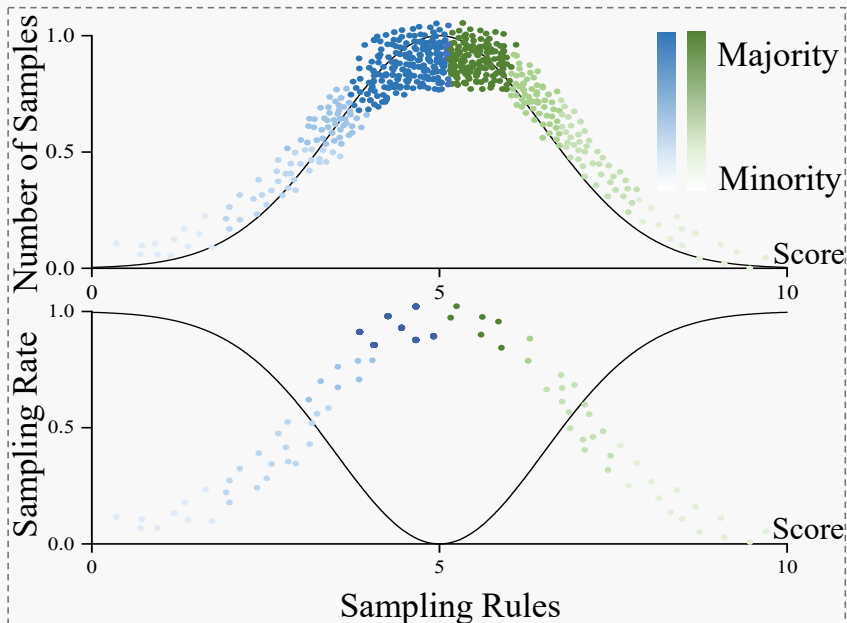
2.1 3.2 7.5 9.0

Sampled Features and Labels

Mixup

2.6 8.4

Mixed Features and Labels



(a) TFA (optional)

2.1 2.6 3.2 4.9 5.5 7.5 8.4 9.0

$M_{2,B}^f = \cos(\text{feature}, \text{feature})$

Feature Similarity

$M_{2,B}^l = \varphi(2.6, 3.2)$

Label Similarity

$M_{1,1}^f$	$M_{1,2}^f$...	$M_{1,B}^f$		$M_{1,1}^l$	$M_{1,2}^l$...	$M_{1,B}^l$
$M_{2,1}^f$	$M_{2,2}^f$		$M_{2,B}^f$?	$M_{2,1}^l$	$M_{2,2}^l$		$M_{2,B}^l$
...		
$M_{B,1}^f$	$M_{B,2}^f$...	$M_{B,B}^f$		$M_{B,1}^l$	$M_{B,2}^l$...	$M_{B,B}^l$

Feature Aligning by Label

2.1 2.6 3.2 4.9 5.5 7.5 8.4 9.0

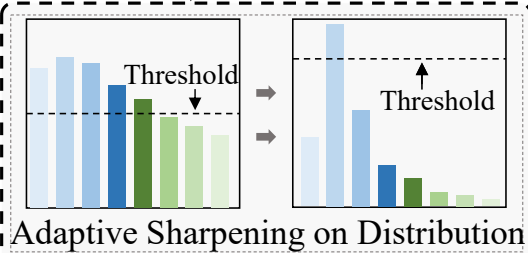
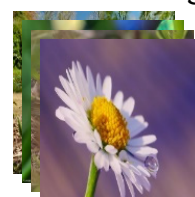
Aligned Features

(b) FLSA (optional)

MLP

Predicted Score

Unlabeled images



Pseudo-label

(c) APDS (optional)