ICLR
International Conference On
Learning Representations

SCAD: Super-Class-Aware Debiasing for Long-Tailed Semi-Supervised Learning

Sunguk Jang*, Jinwoo Jeon*, Byung-Jun Lee

Korea University, AITRICS

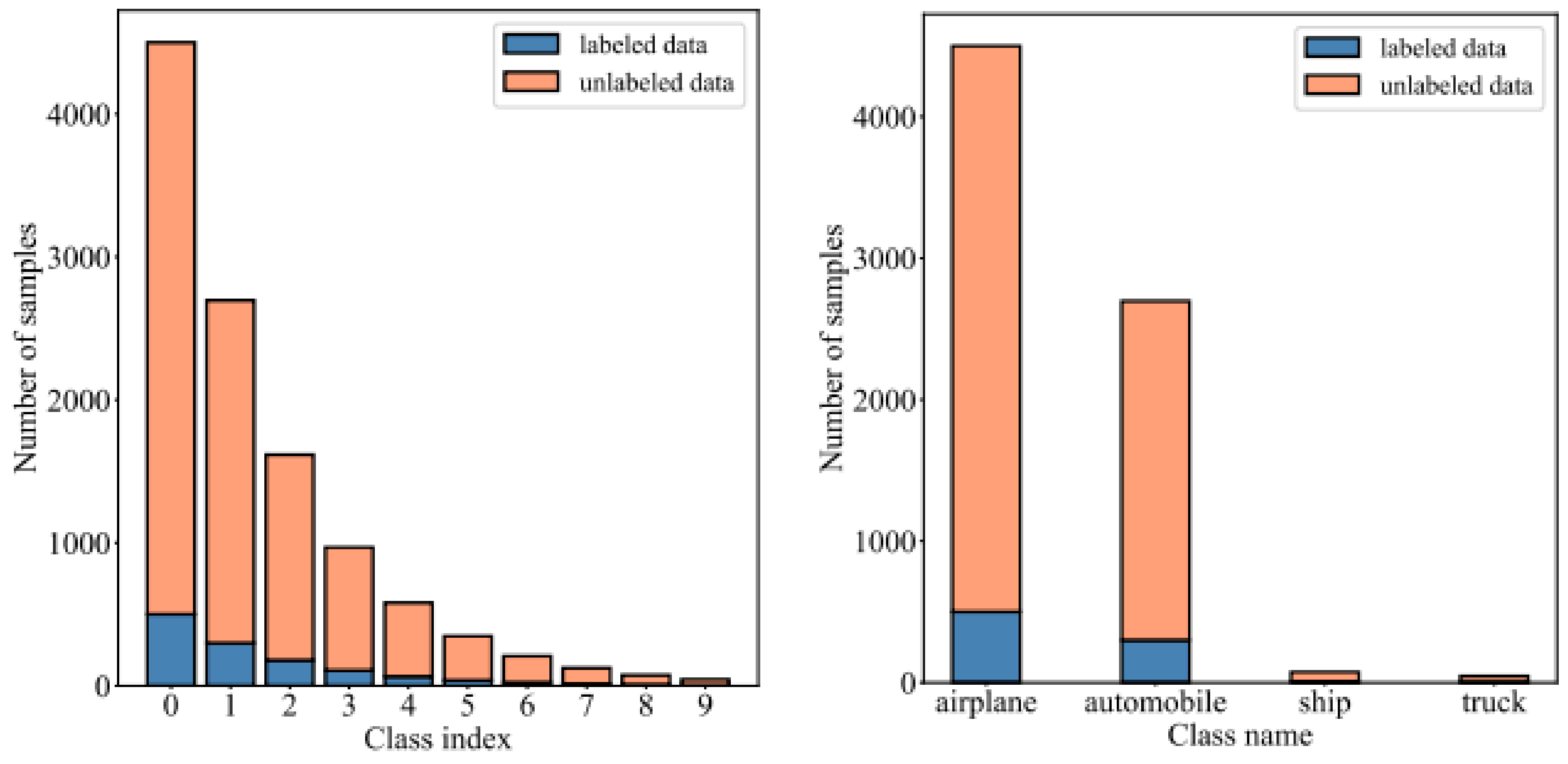
KOREA UNIVERSITY 1905

aitrics

Motivation

➤ Existing long-tailed learning methods focus on global class imbalance, but overlook semantic relationships among classes.

➤ This leads to biased predictions within semantically related groups, even when global prior is addressed.

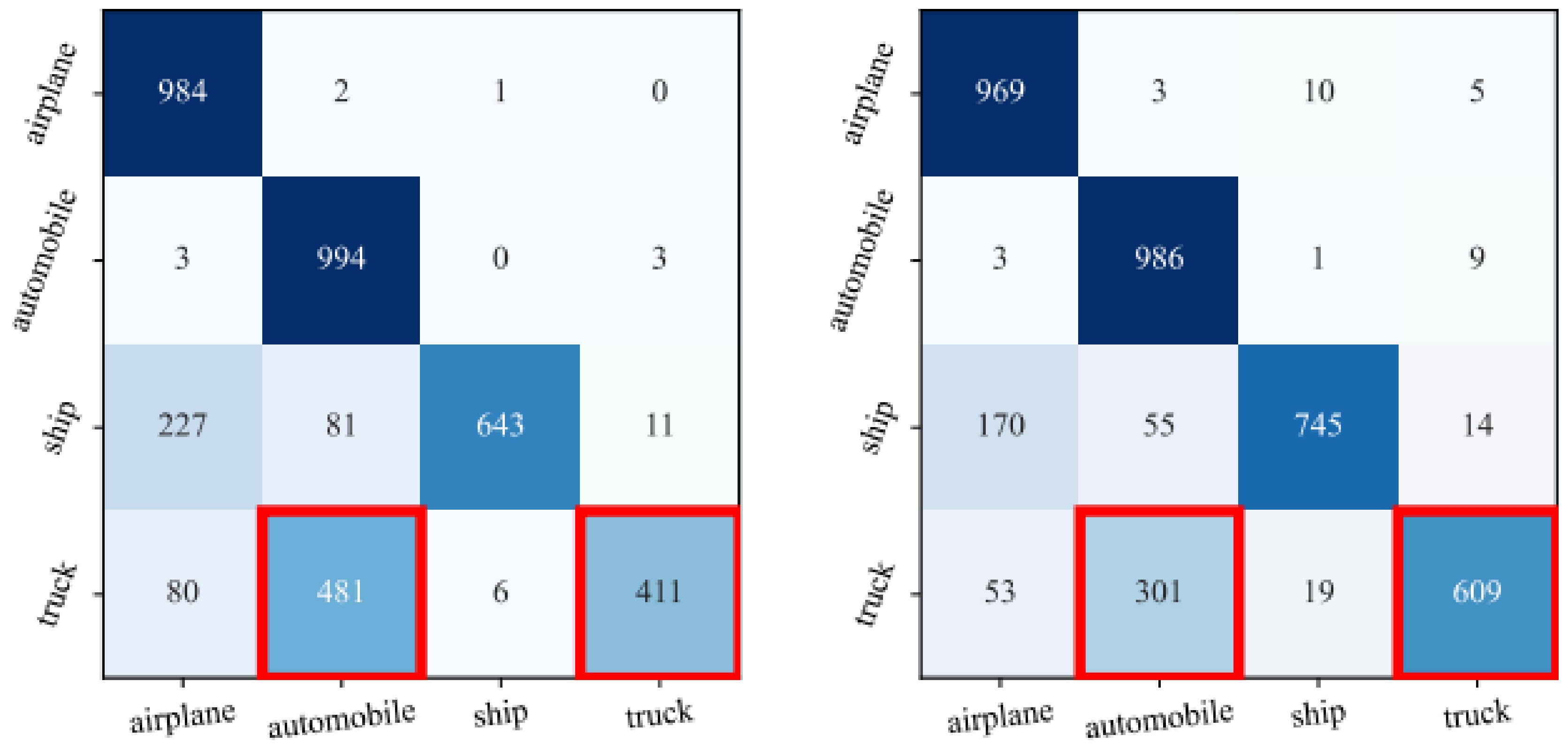


(a) Global Distribution (b) Local Distribution

Intra-Super-Class Imbalance Problem

- We identify a new challenge called intra-super-class imbalance, where imbalance persists within semantically similar classes.

- To address this, we propose SCAD, which dynamically adjusts logits based on each sample's predicted super-class.

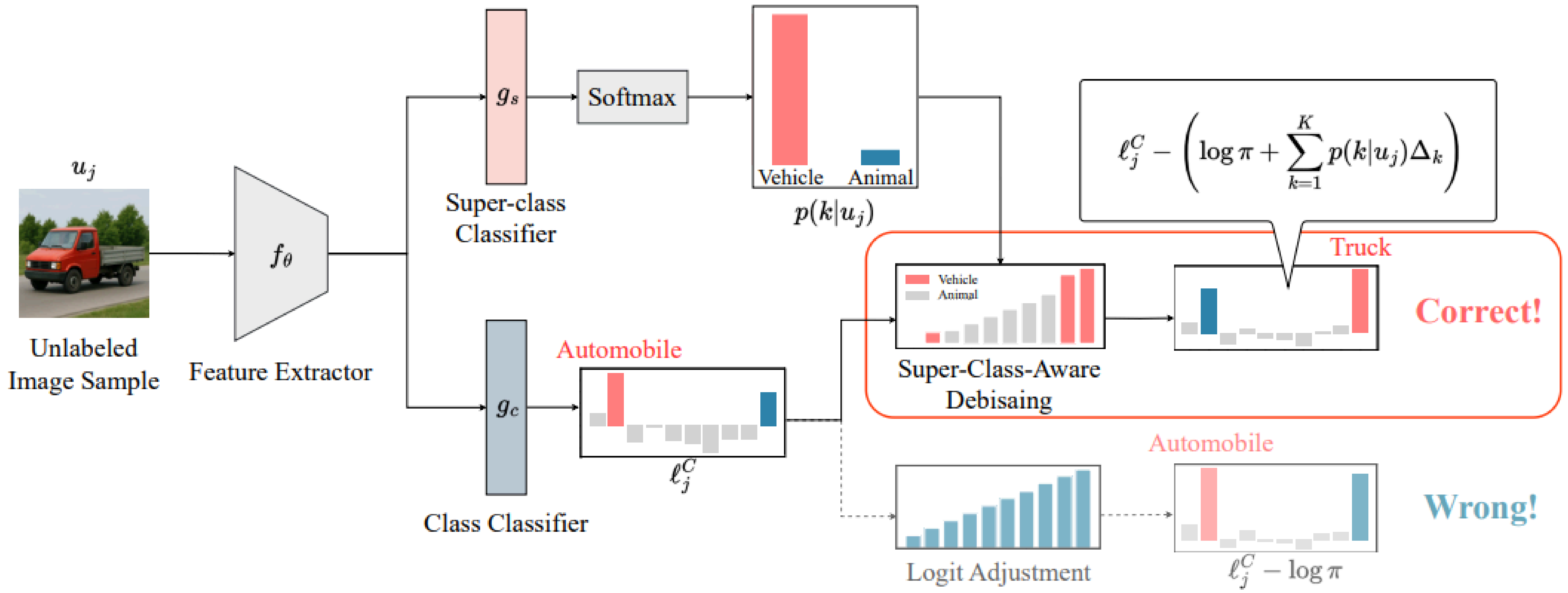


(c) FixMatch + LA (d) FixMatch + SCAD

Method

➤ We construct superclass groups using a pretrained text model (e.g., CLIP) to capture semantic relationships between classes, and then introduce an additional superclass classifier

➤ SCAD decomposes the adjustment into a global class prior and a local intra–super-class prior. While the global prior corrects overall class imbalance, the intra–super-class prior captures frequency within each semantic group, enabling debiasing among similar classes.



Experiment

Algorithm	CIFAR10-LT				CIFAR100-LT			
	$\gamma = \gamma_l = \gamma_u = 100$		$\gamma = \gamma_l = \gamma_u = 150$		$\gamma = \gamma_l = \gamma_u = 10$		$\gamma = \gamma_l = \gamma_u = 20$	
	$N_1 = 500$ $M_1 = 4000$	$N_1 = 1500$ $M_1 = 3000$	$N_1 = 500$ $M_1 = 4000$	$N_1 = 1500$ $M_1 = 3000$	$N_1 = 50$ $M_1 = 400$	$N_1 = 150$ $M_1 = 300$	$N_1 = 50$ $M_1 = 400$	$N_1 = 150$ $M_1 = 300$
Supervised	47.3 $\pm$0.95	61.9 $\pm$0.41	44.2 $\pm$0.33	58.2 $\pm$0.29	29.6 $\pm$0.57	46.9 $\pm$0.22	25.1 $\pm$1.14	41.2 $\pm$0.15
w/ LA	53.3 $\pm$0.44	70.6 $\pm$0.21	49.5 $\pm$0.40	67.1 $\pm$0.78	30.2 $\pm$0.44	48.7 $\pm$0.89	26.5 $\pm$1.31	44.1 $\pm$0.42
FixMatch	67.8 $\pm$1.13	77.5 $\pm$1.32	62.9 $\pm$0.36	72.4 $\pm$1.03	45.2 $\pm$0.55	56.5 $\pm$0.06	40.0 $\pm$0.96	50.7 $\pm$0.25
w/ DARP	74.5 $\pm$0.78	77.8 $\pm$0.63	67.2 $\pm$0.32	73.6 $\pm$0.73	49.4 $\pm$0.20	58.1 $\pm$0.44	43.4 $\pm$0.87	52.2 $\pm$0.66
w/ CReST+	76.3 $\pm$0.86	78.1 $\pm$0.42	67.5 $\pm$0.45	73.7 $\pm$0.34	44.5 $\pm$0.94	57.4 $\pm$0.18	40.1 $\pm$1.28	52.1 $\pm$0.21
w/ DASO	76.0 $\pm$0.37	79.1 $\pm$0.75	70.1 $\pm$1.81	75.1 $\pm$0.77	49.8 $\pm$0.24	59.2 $\pm$0.35	43.6 $\pm$0.09	52.9 $\pm$0.42
w/ DASO + Ours	75.1 $\pm$0.10	79.2 $\pm$1.60	69.0 $\pm$0.61	75.8 $\pm$0.33	50.0 $\pm$0.22	58.7 $\pm$0.28	44.5 $\pm$0.35	53.2 $\pm$0.50
FixMatch + LA	75.3 $\pm$2.45	82.0 $\pm$0.36	67.0 $\pm$2.49	78.0 $\pm$0.91	47.3 $\pm$0.42	58.6 $\pm$0.36	41.4 $\pm$0.93	53.4 $\pm$0.32
w/ DARP	76.6 $\pm$0.92	80.8 $\pm$0.62	68.2 $\pm$0.94	76.7 $\pm$1.13	50.5 $\pm$0.78	59.9 $\pm$0.32	44.4 $\pm$0.65	53.8 $\pm$0.43
w/ CReST+	76.7 $\pm$1.13	81.1 $\pm$0.57	70.9 $\pm$1.18	77.9 $\pm$0.71	44.0 $\pm$0.21	57.1 $\pm$0.55	40.6 $\pm$0.55	52.3 $\pm$0.20
w/ DASO	77.9 $\pm$0.88	82.5 $\pm$0.08	70.1 $\pm$1.68	79.0 $\pm$2.23	50.7 $\pm$0.51	60.6 $\pm$0.71	44.1 $\pm$0.61	55.1 $\pm$0.72
w/ DASO + Ours	81.6 $\pm$0.22	84.0 $\pm$0.99	74.5 $\pm$1.21	82.2 $\pm$0.14	51.8 $\pm$0.28	60.5 $\pm$0.21	45.7 $\pm$0.56	55.7 $\pm$0.57
FixMatch + ACR	81.6 $\pm$0.19	84.1 $\pm$0.39	77.0 $\pm$1.19	80.9 $\pm$0.22	51.3 $\pm$0.48	61.1 $\pm$0.11	44.8 $\pm$0.21	55.9 $\pm$0.31
w/ Ours	**83.5** $\pm$0.16	**85.5** $\pm$0.03	**78.6** $\pm$0.56	**83.3** $\pm$0.20	**52.7** $\pm$0.11	**61.8** $\pm$0.21	**45.8** $\pm$0.20	**56.4** $\pm$0.10

CIFAR10-LT, CIFAR100-LT

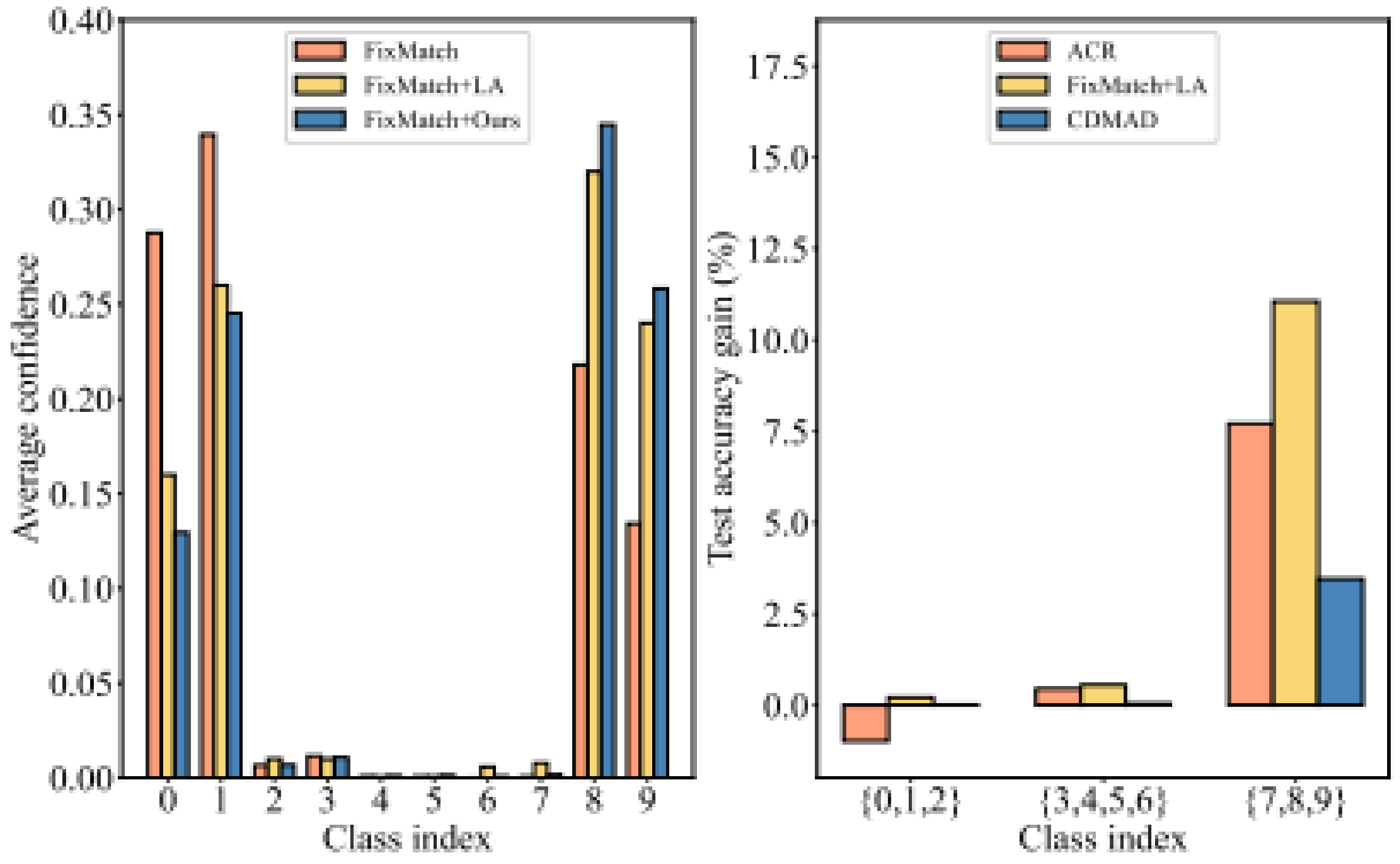
Algorithm / ImageNet-127	32 × 32	64 × 64
FixMatch	29.7	42.3
w/ DARP	30.5	42.5
w/ DARP+cRT	39.7	51.0
w/ CReST+	32.5	44.7
w/ CReST++LA	40.9	55.9
w/ CoSSL	43.7	53.9
w/ ACR	57.2	63.6
w/ Ours	60.1	66.7
w/ ACR + Ours	**60.5**	**67.0**

Imagenet-127

Experiment

Ablations	CIFAR10-LT	STL10-LT
FixMatch	67.8	56.1
+ Super-class learning	69.2	69.0
+ Logit-Adjustment (LA)	76.9	70.4
+ Super-class-Aware Debiasing (SCAD)	**78.7**	**71.3**

Algorithms	CIFAR100-LT
FixMatch + LA	47.3
w/ Ours with ground truth	50.3
w/ Ours with GloVe	49.7
w/ Ours with SBERT	50.1
w/ Ours with CLIP text encoder	50.4
w/ Ours with text-embedding-ada-002	**50.5**



Conclusion

- We introduce Super-Class-Aware Debiasing (SCAD), a novel framework designed to tackle the fundamental challenge of intra-super-class imbalance in long-tailed semi-supervised learning (LTSSL).

- SCAD outperforms strong baselines and debiasing methods across diverse datasets, demonstrating that leveraging semantic structure and local class relationships enables robust and generalizable learning in imbalanced settings.