Mitigating Confidence Miscalibration in Open-World Semi-Supervised Learning

Wenqiang Wu, Feng Wang*, Jiye Liang, Liang Bai.

Key Laboratory of Computational Intelligence and Chinese Information Processing Ministry of Education,

School of Computer and Information Technology, Shanxi University, China

Type:

NeurIPS 2026 Oral

Presenter:

Feng Wang

Time:

2026-12-XX,

ICC Sydney,

Sydney, AUS

***Corresponding author.**

NEURAL INFORMATION PROCESSING SYSTEMS

山西大学 SHANXI UNIVERSITY

山西大学 计算机与信息技术学院（大数据学院）
School of Computer and Information Technology (School of Big Data), Shanxi University

CONTENTS

1 Background & Motivation

2 Research Method

3 Experiment Results

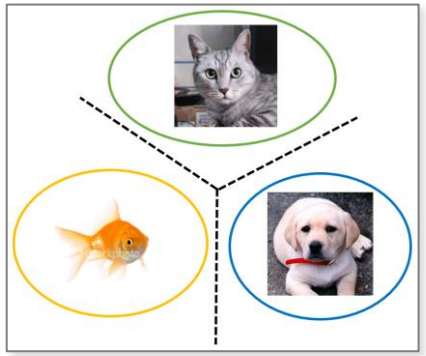
4 Conclusion & Discussion

1 Background & Motivation

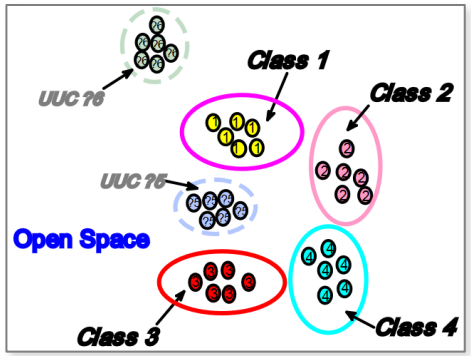
NEURAL INFORMATION PROCESSING SYSTEMS

SHANXI UNIVERSITY

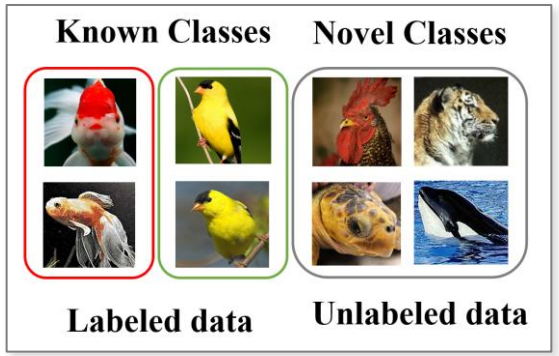
Background



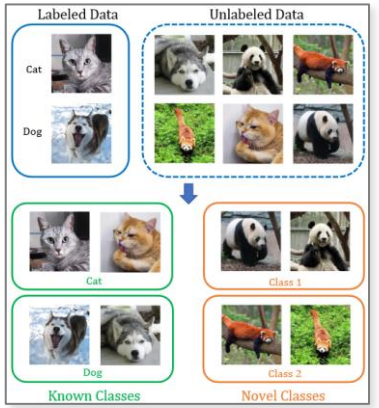
Closed-World Learning



Open Set Learning



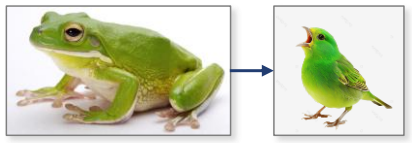
Novel Class Discovery



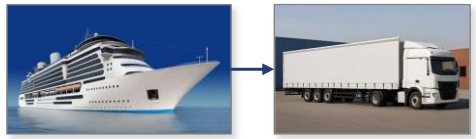
Open-World Learning

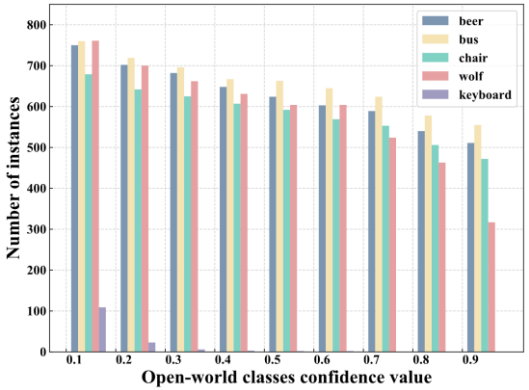
Motivation

All-class distribution error prediction

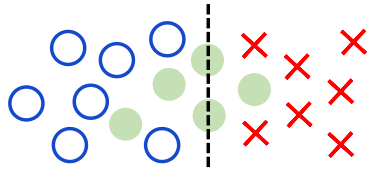


High-confidence misjudgment of known-class distribution

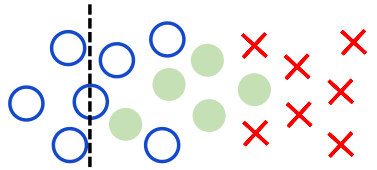




Fixed low threshold



Fixed high threshold



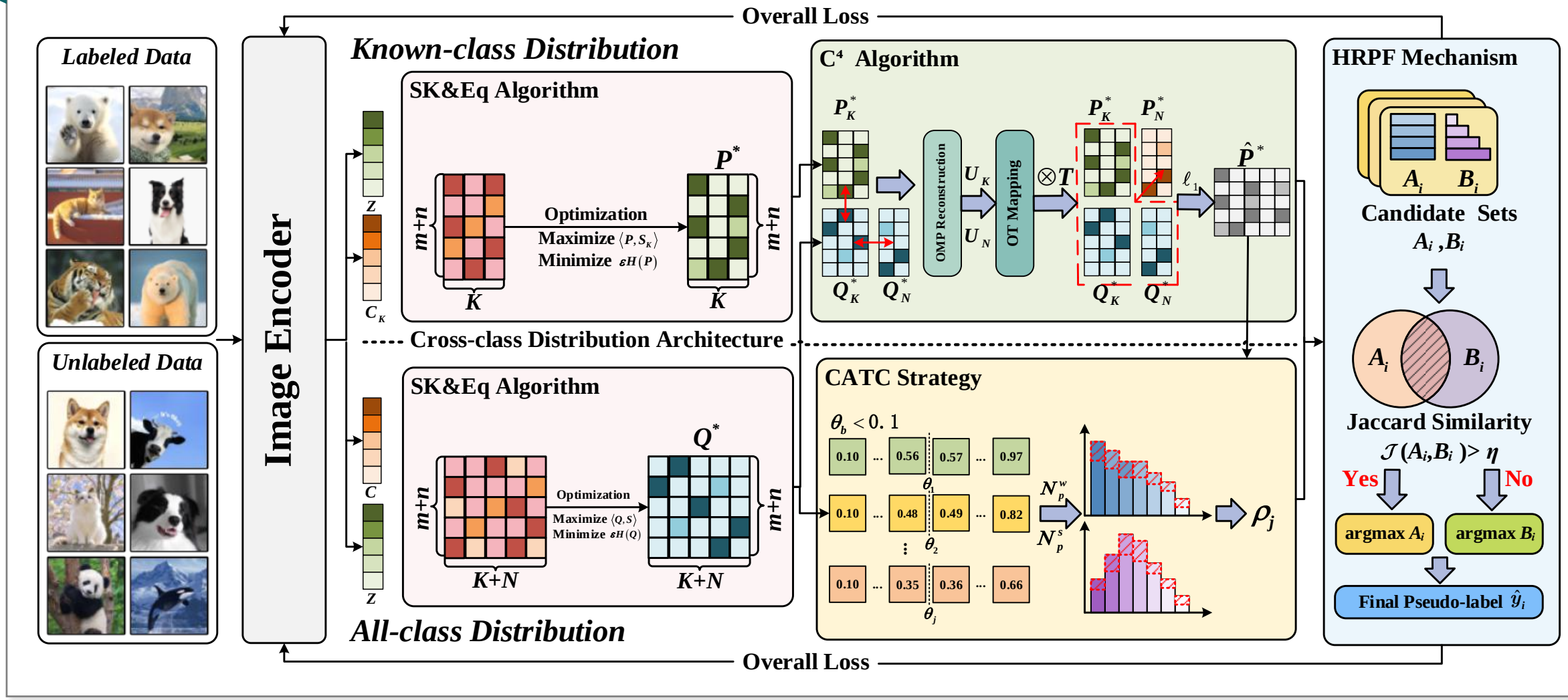
- Known-class bias.
- Single-distribution discriminant bias.
- Inaccurate confidence thresholds.

Known-class Set: Y_l; Unknown-class Set: Y_u. $Y_l \subset Y_u$; $Y_l \cap Y_u = \emptyset$; $Y_{\text{novel}} = Y_u \backslash Y_l$.

NEURAL INFORMATION PROCESSING SYSTEMS

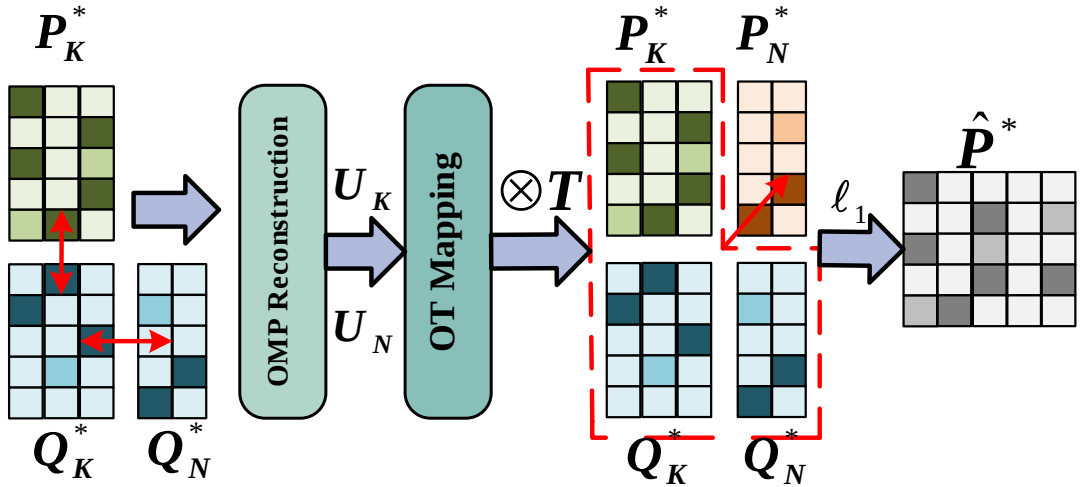
SHANXI UNIVERSITY

Overall Framework



2 OpenMCM

C⁴ Algorithm

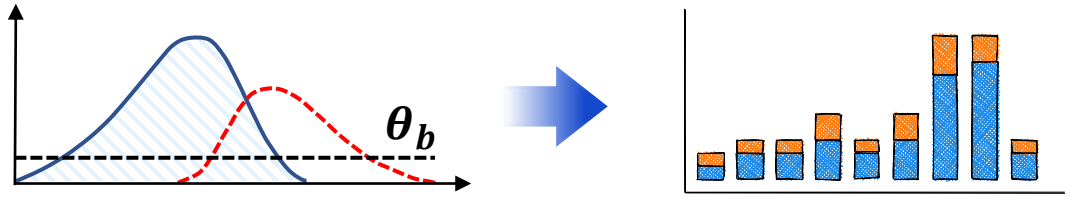


1. **Spare-reconstruction**
2. **OT Mapping**
3. **Normalization**

$$P_N^* = K \times P_K^* \times T \qquad \hat{P}^* = \ell_1 - ([P_K^*, P_N^*])$$

$\hat{P}^*$ & Q^* for subsequent filtering

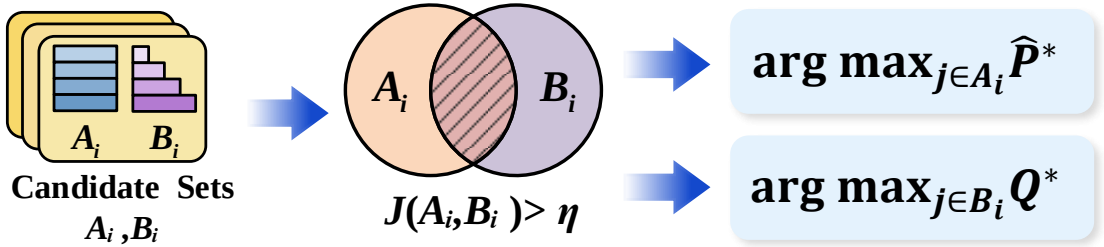
CATC Strategy



- Confidence-Based Candidate Filtering.
- Quantile-Based Class-Specific Thresholding.
- Dynamic Threshold Calibration.

$$\theta_j = \mathcal{Q}\left\{\mathcal{F}\left(p_i^j \middle| p_i^j > \theta_b\right), \frac{\rho}{100}\right\}$$

HRPF Mechanism



$$A_i \,\&\, B_i = \emptyset \Longrightarrow J(A_i, B_i) = 0 \Longrightarrow \arg\max_{j \in B_i} Q^*$$

From Sec 3.2, Sec 3.3, Sec 3.4 and Eq. (9).

3 Experiment

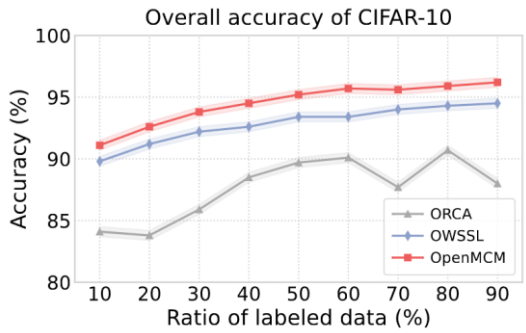
Results

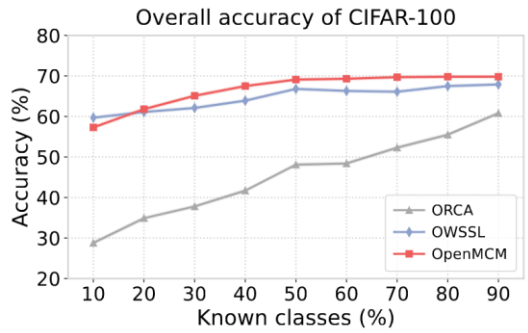
	Methods	Venue	CIFAR-10 Known	CIFAR-10 Novel	CIFAR-10 All	CIFAR-100 Known	CIFAR-100 Novel	CIFAR-100 All	ImageNet-100 Known	ImageNet-100 Novel	ImageNet-100 All
I	†FixMatch [11]	NeurIPS'20	71.5	50.4†	49.5	39.6	23.5†	20.3	65.8	36.7†	34.9
	†DS^{3}L [17]	ICML'20	77.6	45.3†	40.2	55.1	23.7†	24.0	71.2	32.5†	30.8
	†CGDL [23]	CVPR'20	72.3	44.6†	39.7	49.3	22.5†	23.5	67.3	33.8†	31.9
II	*DTC [26]	ICCV'19	53.9*	39.5	38.3	31.3*	22.9	18.3	25.6*	20.8	21.3
	*RankStats [27]	TPAMI'22	86.6*	81.0	82.9	36.4*	28.4	23.1	47.3*	28.7	40.3
	*UNO [33]	ICCV'21	91.6*	69.3	80.5	68.3*	36.5	51.5	–	–	–
	*SimCLR [72]	ICML'20	58.3*	63.4	51.7	28.6*	21.1	22.3	39.5*	35.7	36.9
III	ORCA [35]	ICLR'22	88.2	90.4	89.7	66.9	43.0	48.1	89.1	72.1	77.8
	GCD [38]	CVPR'22	78.4	79.7	79.1	68.5	33.5	45.2	82.3	58.3	68.2
	NACH [39]	NeurIPS'22	89.5	92.2	91.3	68.7	47.0	52.1	91.0	75.5	79.6
	OpenLDN [40]	ECCV'22	92.3	86.8	88.6	67.3	40.4	49.4	74.5	45.7	55.3
	TRSSL [41]	ECCV'22	94.9	89.6	92.2	68.5	52.1	60.3	82.6	67.8	75.4
	OpenNCD [42]	IJCAI'23	88.4	90.6	90.1	69.7	43.4	49.3	90.0	77.5	81.6
	TIDA [43]	NeurIPS'23	94.2	93.4	93.8	73.3	56.6	65.3	83.4	71.2	77.6
	SSOC [44]	ICDM'23	93.3	92.6	92.8	69.0	48.0	53.1	91.4	78.4	82.7
	OWSSL [45]	PR'24	94.9	91.8	93.4	77.9	55.7	66.8	90.0	68.2	79.1
	OwMatch [46]	NeurIPS'24	93.0	**95.9**	94.4	74.5	55.9	65.1	91.7	72.0	81.8
	TRAILER [47]	CVPR'24	93.4	95.0	94.4	69.7	48.7	55.6	91.4	**82.4**	85.3
	DSCL [48]	TNNLS'25	93.9	93.9	93.9	73.8	56.7	62.3	90.6	81.3	84.4
	OpenMCM	–	**95.3**	95.1	**95.2**	**78.4**	**59.8**	**69.1**	**92.0**	81.8	**86.9**

Key Results

- **SOTA overall ACC: 95.2% / 69.1% / 86.9%**
- **Gain over prior best: +0.8 / +2.3 / +1.6**
- **Balanced recognition: Strong performance on known & novel classes.**

Additional Experiment





Label Scarcity

- **Consistently high ACC across different labeled ratios.**
- **91.1% overall ACC with only 10% labeled data.**

Novel-Class Ratios

- **Strong ACC across different novel-class ratios.**
- **Demonstrates robustness in highly open-world settings**

☺ **Robust to label scarcity and increasing openness.**

From Tab 1, Fig 7-(I).A and Fig 8-(II).A.

4 Conclusion & Discussion

NEURAL INFORMATION PROCESSING SYSTEMS

SHANXI UNIVERSITY 山西大学

Conclusion

- ☐ Proposes a class-discriminative method.
- ☐ Maps confidence distributions across class different spaces.
- ☐ Dynamic per-sample class assignment.

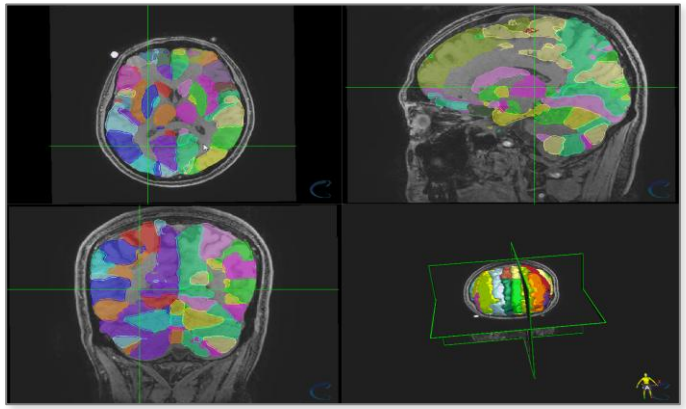
- ✓ Reliable model calibration.
- ✓ Balanced learning pace across classes.
- ✓ Mitigate the Matthew effect.

☺ This study reveals cross-semantic data correlations, providing new insights for Open-World Learning models.

Broader impacts









Limitations

- ☐ Class Number Prior.
- ☐ Extreme long-tailed distributions data.
- ☐ Fine-grained ambiguity.

- ✓ Robust class number estimation.
- ✓ Improve long-tail distribution calibration.
- ✓ Fine-grained class relationship modeling.

Thanks

Q&A