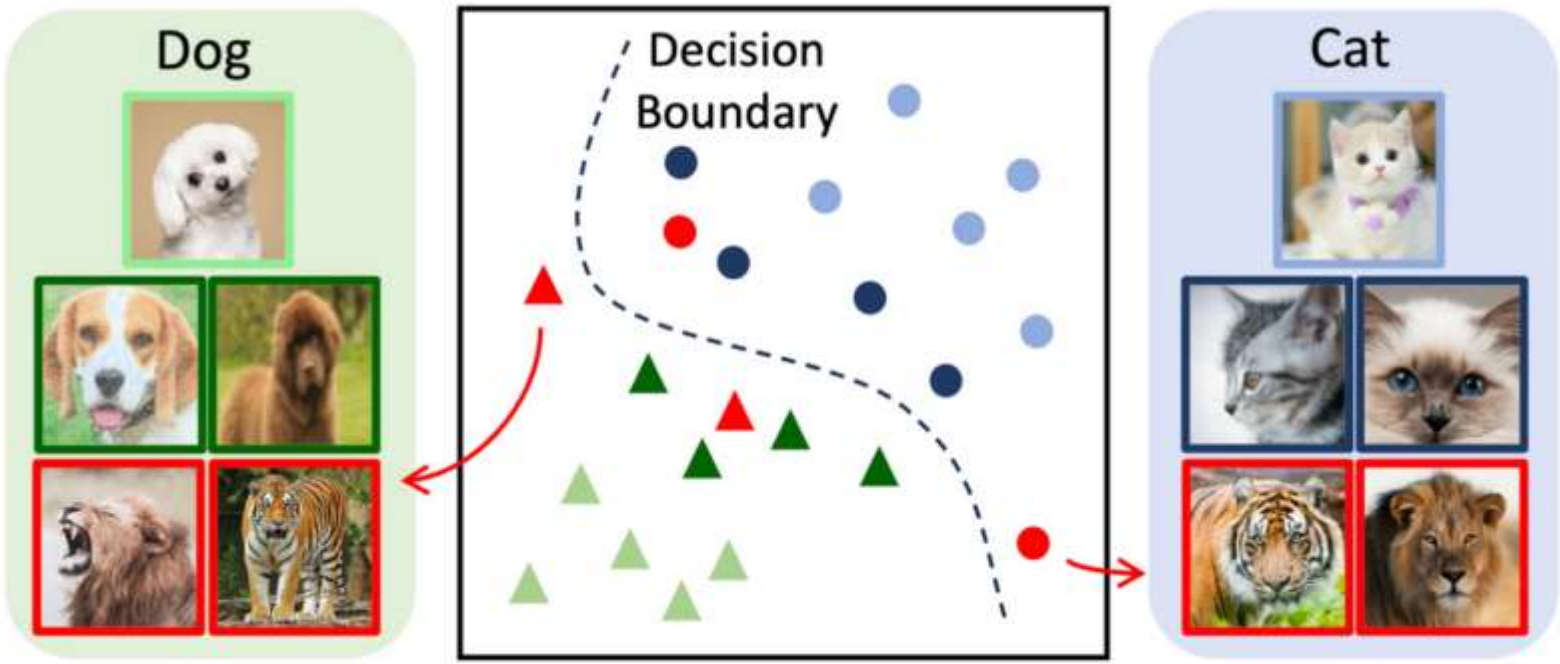
Open-set Active Learning

Motivation

- Tradition AL methods generally assumed that labeled and unlabeled data are drawn from the same class distribution.

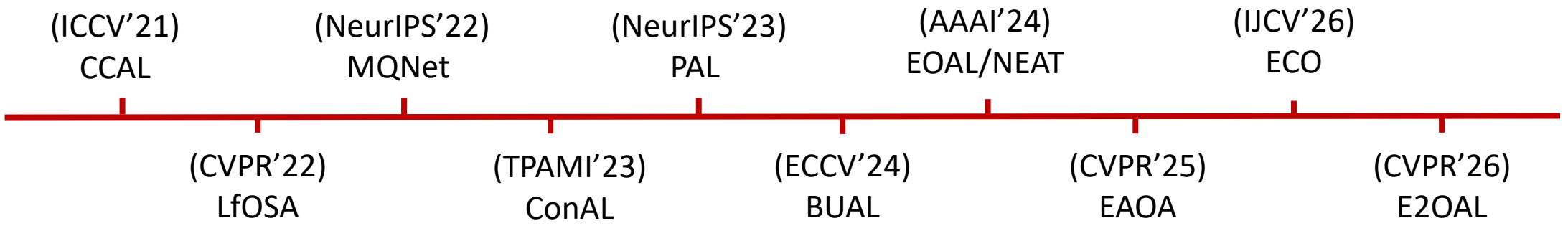


Open-set class / Unknown class / Out-of-distribution (OOD) examples

OOD examples exhibit high uncertainty and diversity because they share neither class-distinctive features nor other inductive biases with ID examples

⇨ Traditional AL fails.

Open-set Active Learning



ICCV'21: CCAL & TPAMI'23: ConAL

Contrastive Coding for Active Learning under Class Distribution Mismatch

Pan Du[1,2], Suyun Zhao[1,2,*], Hui Chen[2,†], Shuwen Chai[2,†], Hong Chen[1,2], Cuiping Li[1,2]
Key Lab of Data Engineering and Knowledge Engineering of MOE Renmin University of China[1]
Renmin University of China, Beijing, China[2]
{dupan, zhaosuyun, chenhui1025, chaishuwen, chong, licuiping}@ruc.edu.cn

ICCV'21: CCAL & TPAMI'23: ConAL

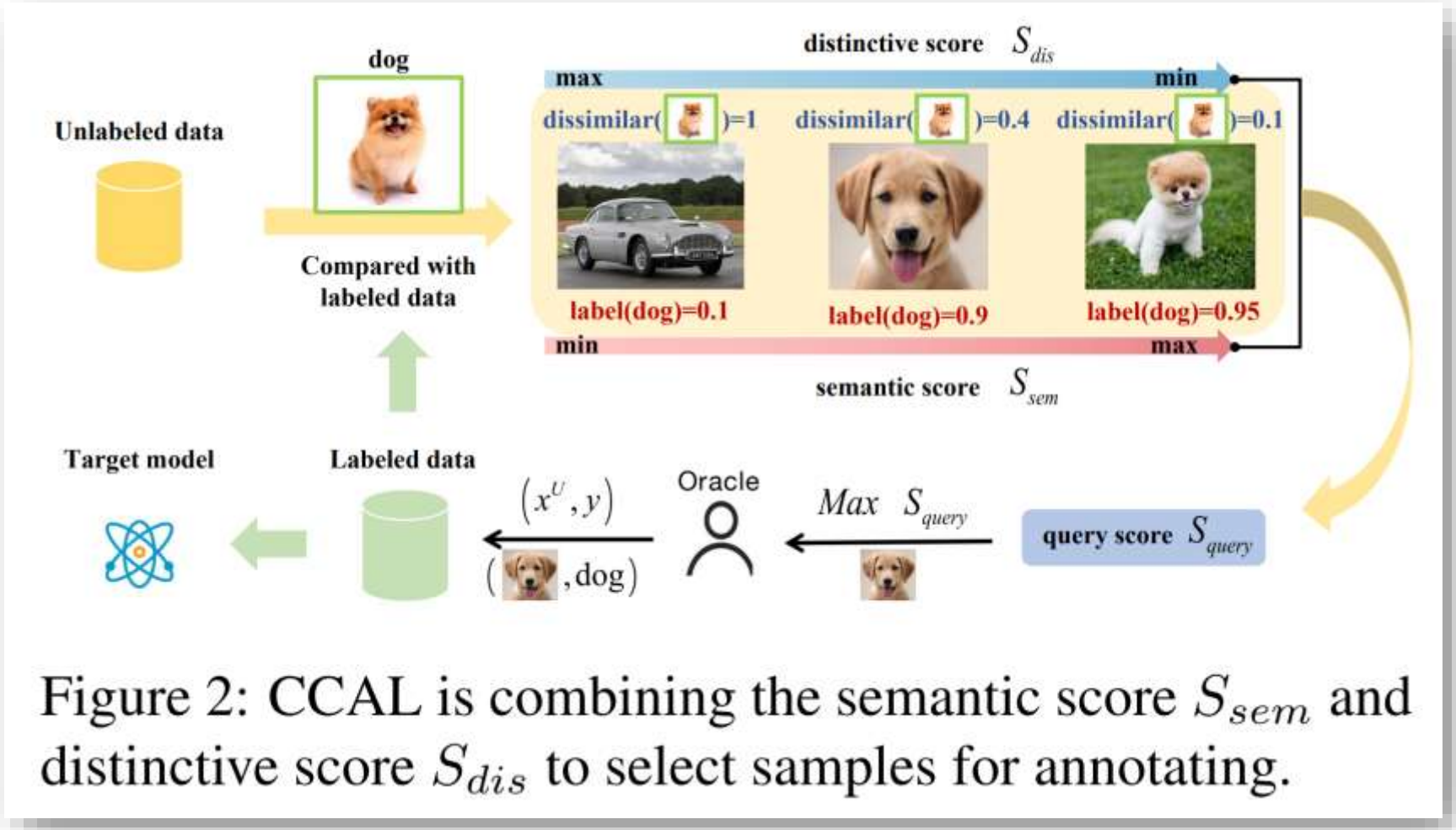


Figure 2: CCAL is combining the semantic score S_{sem} and distinctive score S_{dis} to select samples for annotating.

- Learn distinctive and semantic features.
- Query samples with high semantic consistency and strong distinguishability.

ICCV'21: CCAL & TPAMI'23: ConAL

- Learning semantic features

Loss

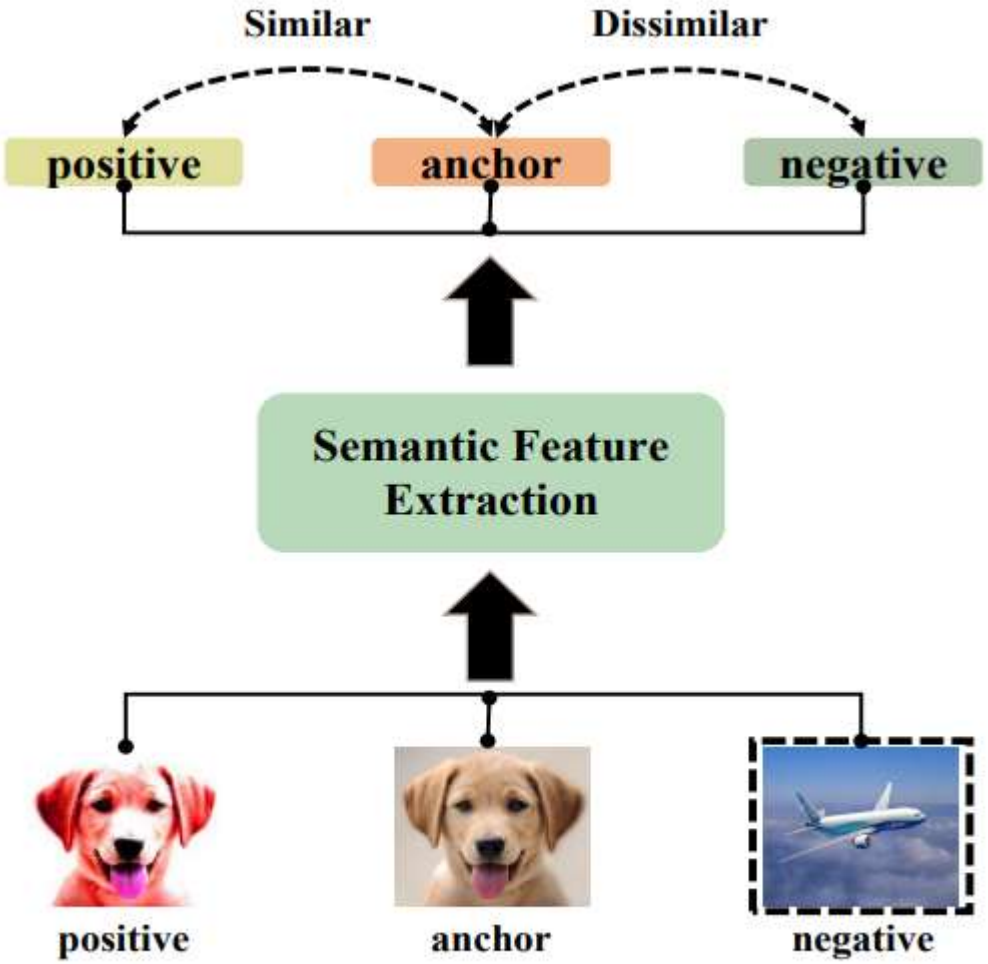
$$L_{sem}(B;\aleph) = \frac{1}{2\,|B|}\sum_{i=1}^{|B|} L_{con}(\tilde{x}_i^{(1)}, \tilde{x}_i^{(2)}, \tilde{B}_{-i}) + L_{con}(\tilde{x}_i^{(2)}, \tilde{x}_i^{(1)}, \tilde{B}_{-i}) \quad (2)$$

where $\tilde{B} = \{\tilde{x}_i^{(1)}\}_{i=1}^{|B|} \cup \{\tilde{x}_i^{(2)}\}_{i=1}^{|B|}$, $\tilde{B}_{-i} = \{\tilde{x}_j^{(1)}\}_{j\neq i} \cup \{\tilde{x}_j^{(2)}\}_{j\neq i}$, $x_i \in X^L \cup X^U$, and L_{con} is a contrastive learning loss [3].

Semantic score

$$S_{sem}(x_i^U) = \sigma(\max_j cos(z_s(x_j^L), z_s(x_i^U))), \quad (3)$$

where $\sigma(s) = (s - min(S_{sem}(X^U)))/(max(S_{sem}(X^U)) - min(S_{sem}(X^U)))$, $x_i^U \in X^U$.



(a) Contrast of semantics.

ICCV'21: CCAL & TPAMI'23: ConAL

- Learning distinctive features

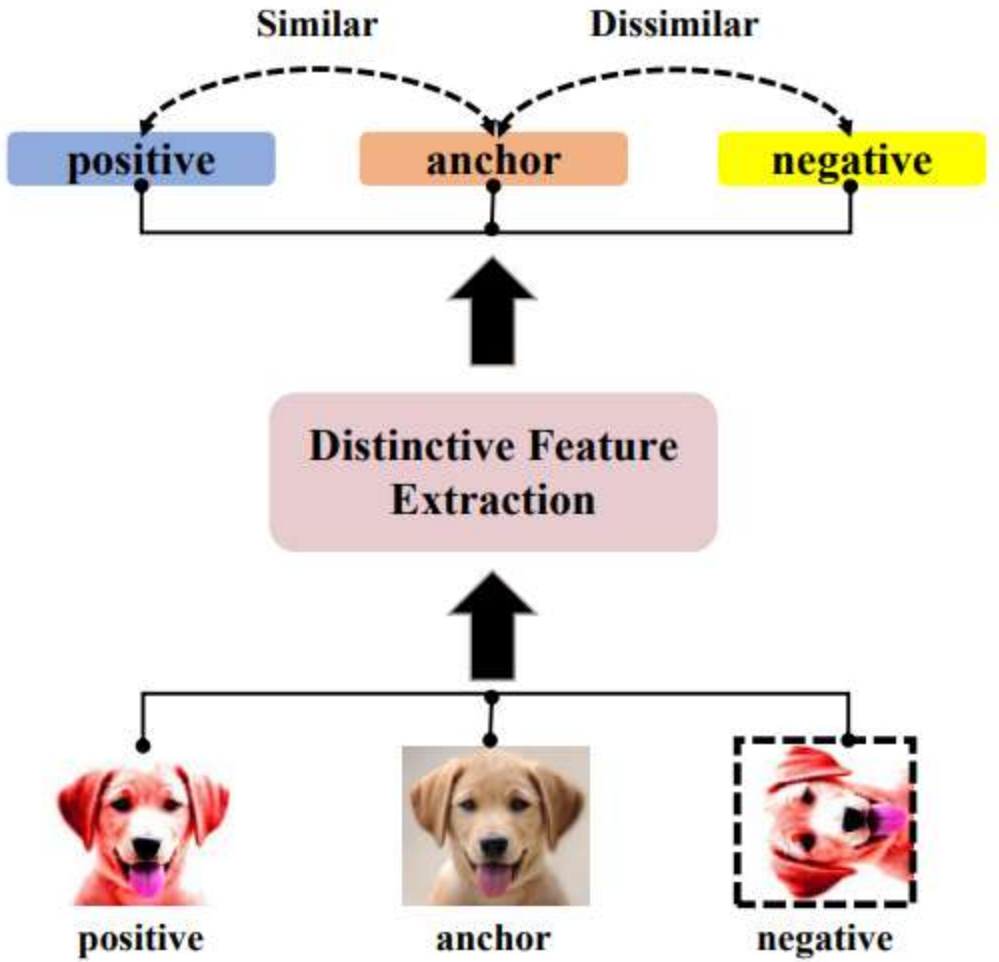
Loss

$$L_{dis}(B;\Re;\aleph) = \frac{1}{2|B|}\frac{1}{|r|}\sum_{i=1}^{|B|}\sum_{k\in r}[L_{con}(\tilde{x}^{(1)}_{\Re_k(i)}, \tilde{x}^{(2)}_{\Re_k(i)}, \tilde{B}^{\Re}_{-i}) + L_{con}(\tilde{x}^{(2)}_{\Re_k(i)}, \tilde{x}^{(1)}_{\Re_k(i)}, \tilde{B}^{\Re}_{-i})] \quad (4)$$

Distinctive score

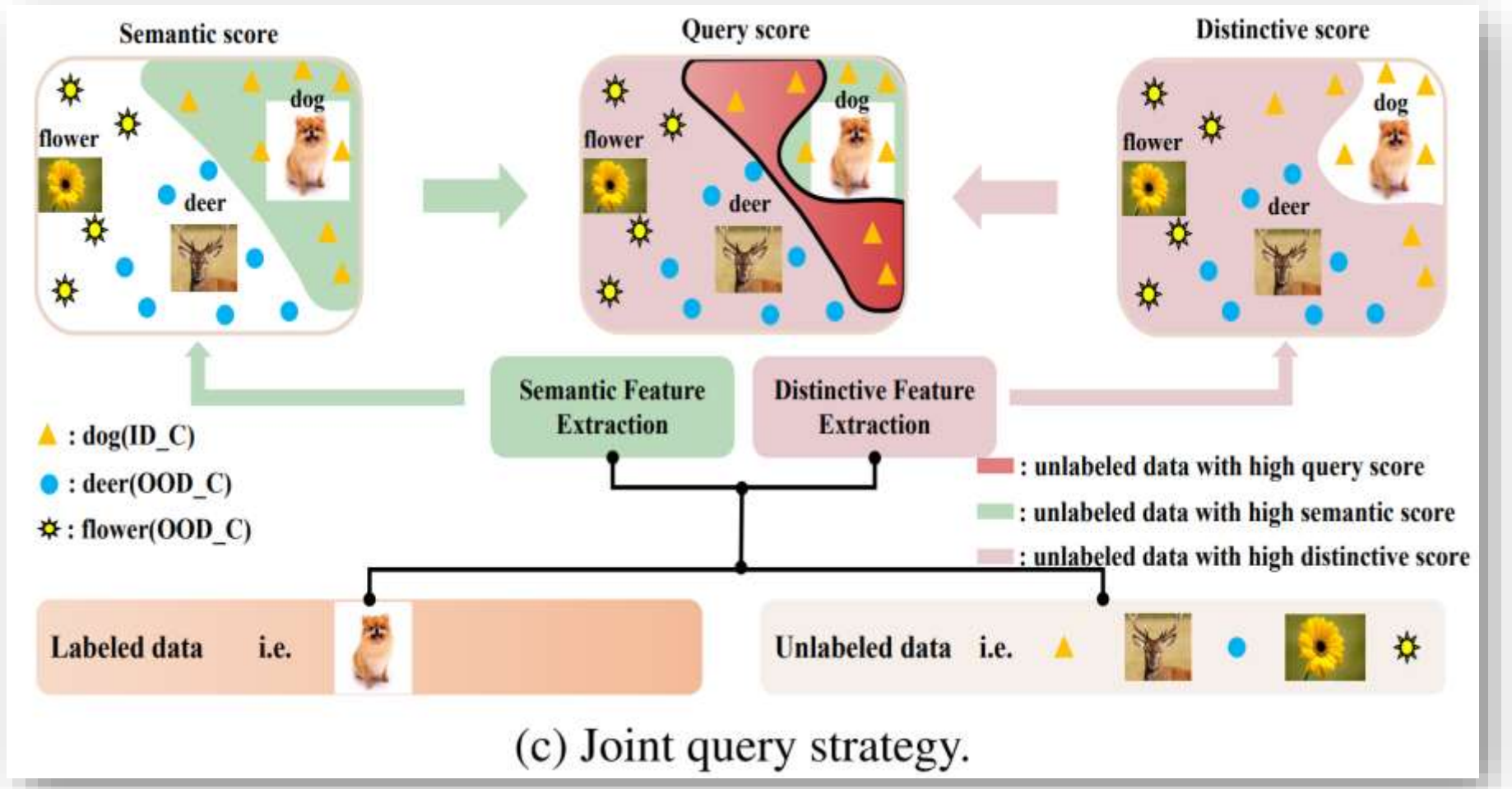
$$S_{dis}(x^U_i) = 1 - \sigma(cos(z_d(x^U_i), z_d(x^L_{i,st})) - cos(z_d(x^U_i), z_d(x^L_{i,nd})) + cos(z_d(x^L_{i,st}), z_d(x^L_{i,nd}))) \quad (5)$$

In Eq.5, $cos(z_d(x^U_i), z_d(x^L_{i,st})) - cos(z_d(x^U_i), z_d(x^L_{i,nd}))$ measures the difference of x^U_i to labeled samples.



(b) Contrast of distinctiveness.

ICCV'21: CCAL & TPAMI'23: ConAL



(c) Joint query strategy.

$$S_{query}(x_i^U) = tanh\left[\psi\left(S_{sem}(x_i^U)\right)\right] + S_{dis}(x_i^U) \quad (6)$$

where $\psi\left(S_{sem}(\cdot)\right) = k \times \left(S_{sem}(\cdot) - t\right)$

ICCV'21: CCAL & TPAMI'23: ConAL

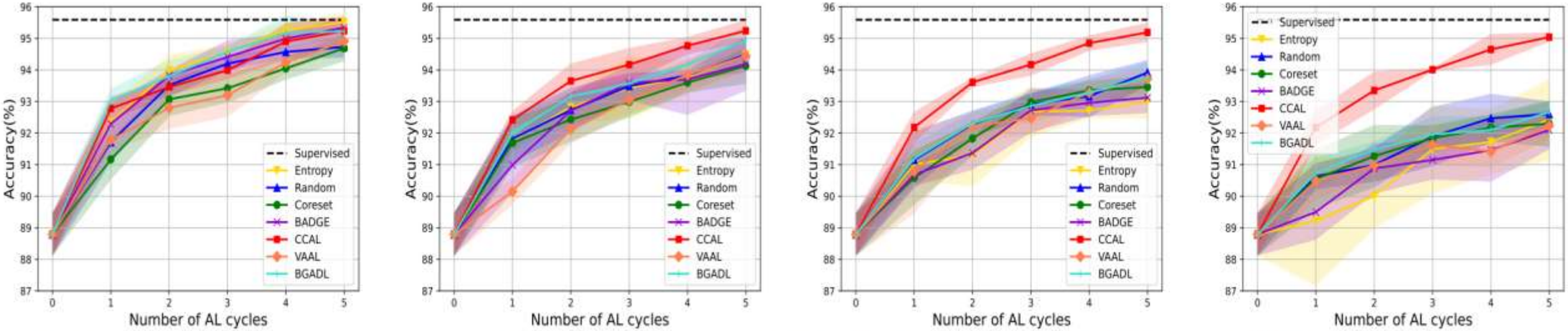


Figure 4: Classification accuracy of CCAL and compared AL algorithms on CIFAR10 under different mismatches.

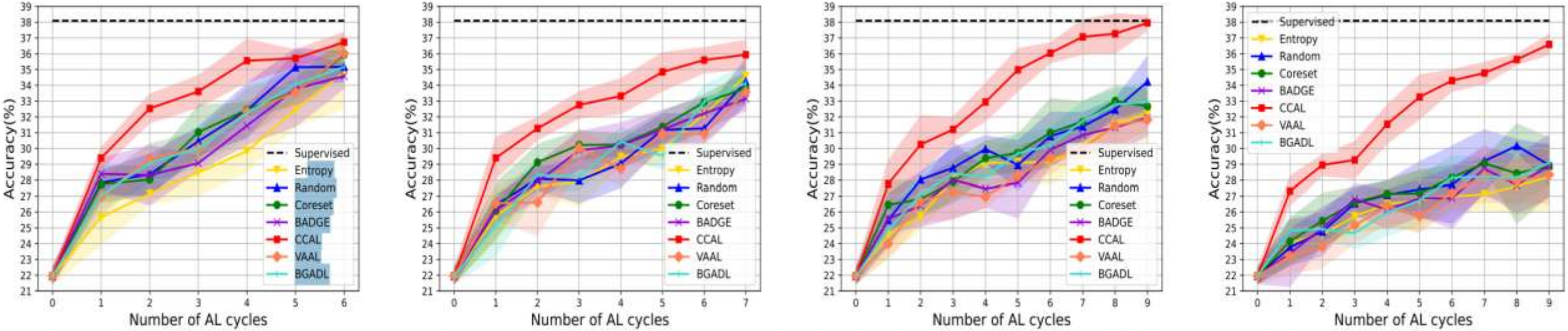


Figure 5: Classification accuracy of CCAL and compared AL algorithms on CIFAR100 under different mismatches.

CVPR'22: LfOSA

Active Learning for Open-set Annotation

Kun-Peng Ning[1*] Xun Zhao[2†] Yu Li[3*] Sheng-Jun Huang[1†]
[1]Nanjing University of Aeronautics and Astronautics
[2]Applied Research Center, Tencent PCG [3]International Digital Economy Academy (IDEA)
{ningkp,huangsj}@nuaa.edu.cn emmaxunzhao@tencent.com liyu@idea.edu.cn

CVPR'22: LfOSA

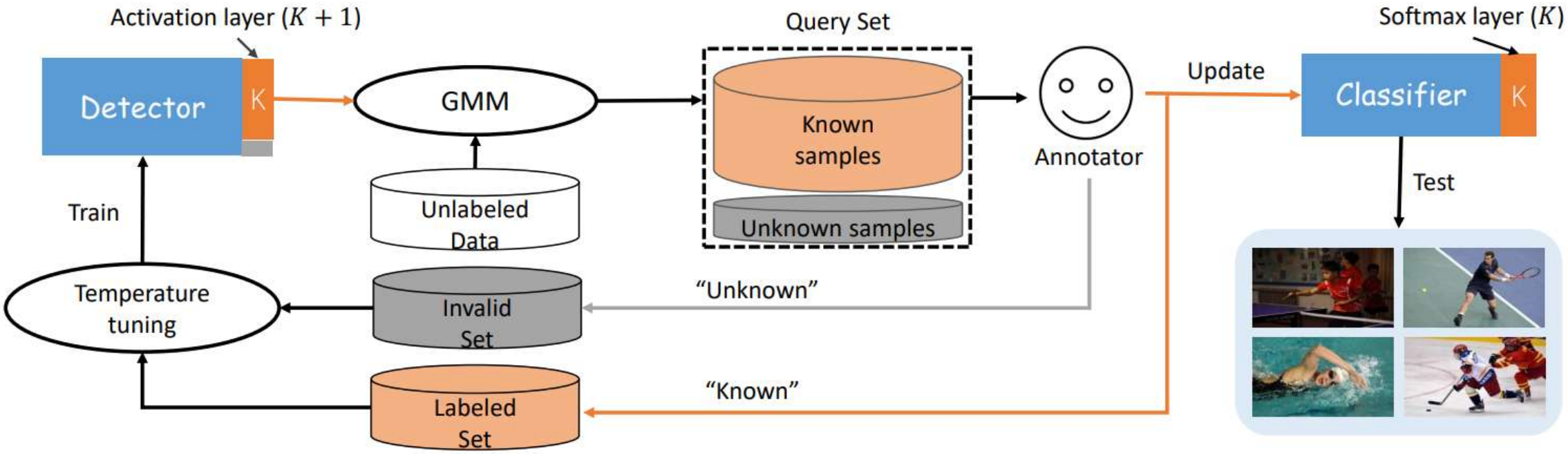


Figure 2. The framework of LfOSA. It includes two networks for detection and classification. The detector attempts to construct a query set for annotation by GMM modeling. After labeling, two networks will be updated for next iteration.

while $c = 1, 2, ..., K$ do

 # Collect the MAV set for each prediction class c

$$mav^c = \{mav_i^c | f_{\theta_D}(x_i) = c, \forall x_i \in D_U\}$$

 # Obtain known probability by GMM

$$\mathcal{W}^c = GMM(mav^c, \theta_D)$$

end

 # Merge and sort the probability sets of all classes

$$\mathcal{W} = sort(\mathcal{W}^1 \cup \mathcal{W}^2 \cup ... \cup \mathcal{W}^K)$$

 # Obtain the query set

$$X^{query} = \{(x_i, w_i) | w_i \geq \tau, \forall (x_i, w_i) \in (D_U, \mathcal{W})\}$$

CVPR'22: LfOSA

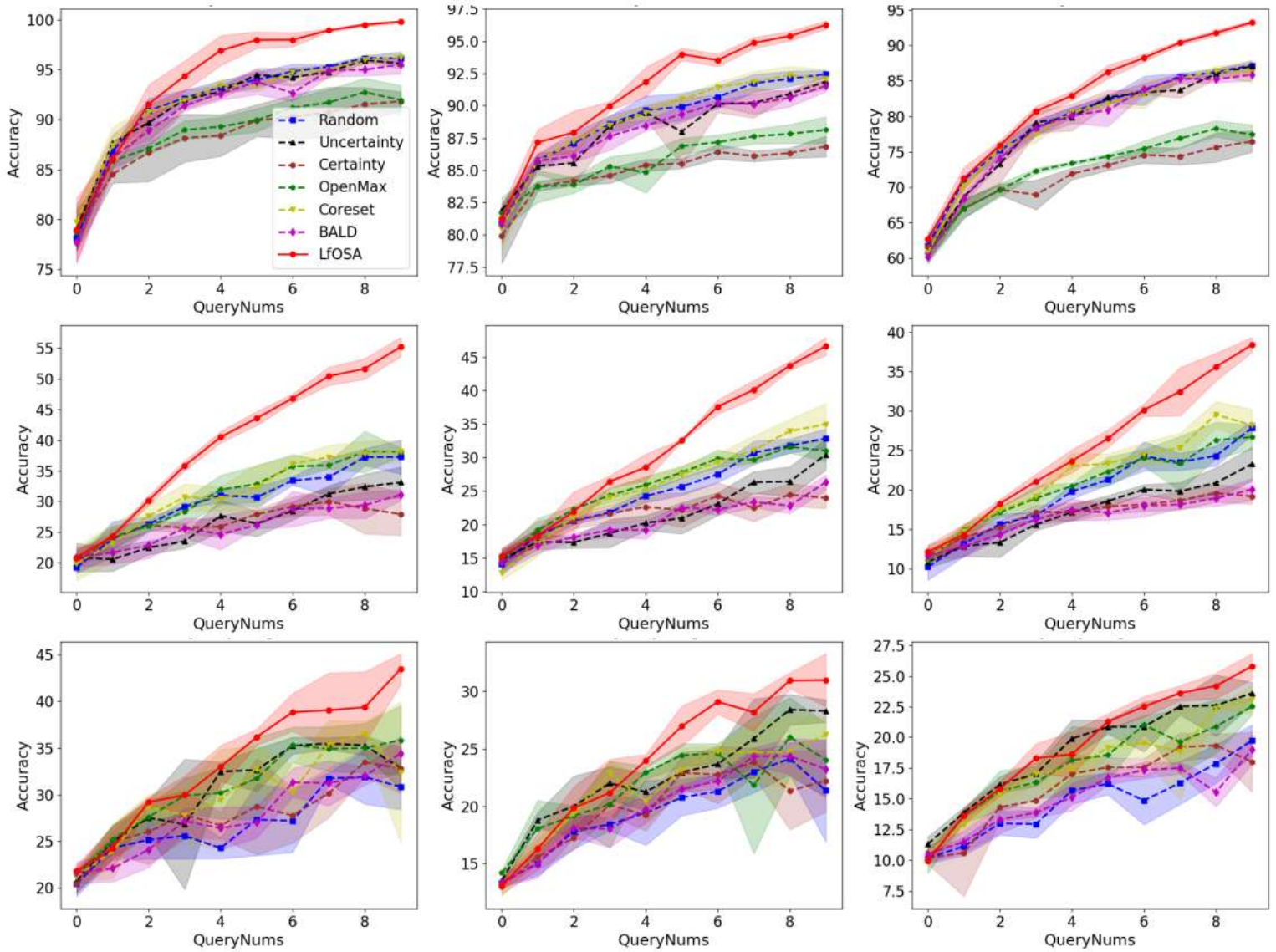
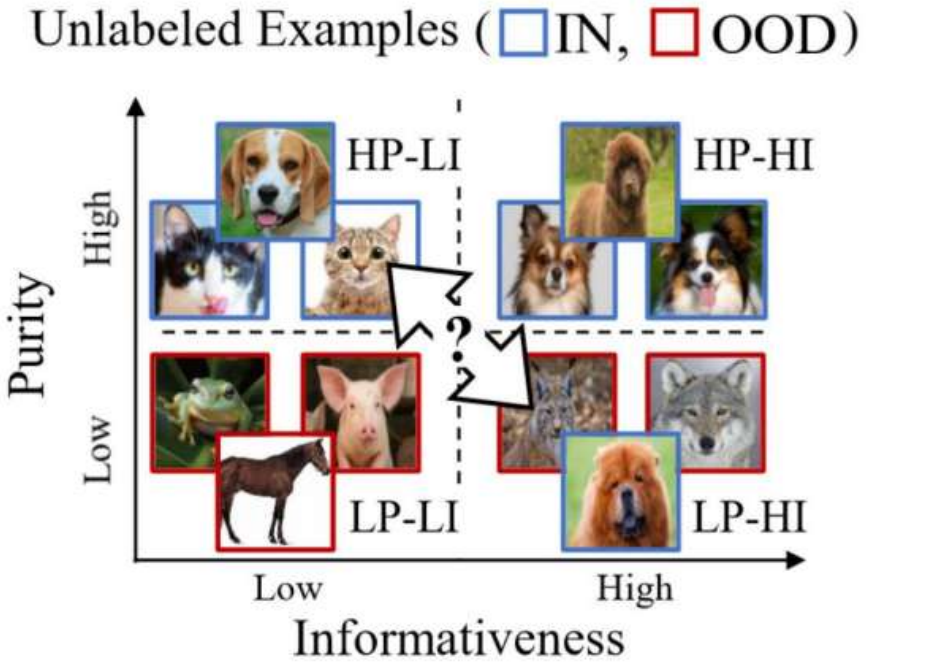


Figure 5. Classification performance comparison on CIFAR10 (first row), CIFAR100 (second row) and Tiny-Imagenet (third row) with 20% (first column), 30% (second column) and 40% (third column) mismatch ratio.

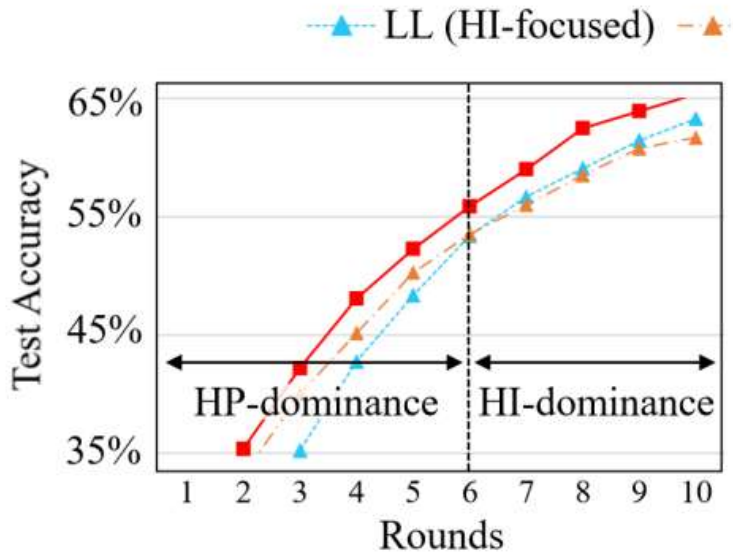
Meta-Query-Net: Resolving Purity-Informativeness Dilemma in Open-set Active Learning

Dongmin Park[1], **Yooju Shin**[1], **Jihwan Bang**[2,3], **Youngjun Lee**[1], **Hwanjun Song**[2*], **Jae-Gil Lee**[1*]
[1] KAIST, [2] NAVER AI Lab, [3] NAVER CLOVA
{dongminpark, yooju.shin, youngjun.lee, jaegil}@kaist.ac.kr
{jihwan.bang, hwanjun.song}@navercorp.com

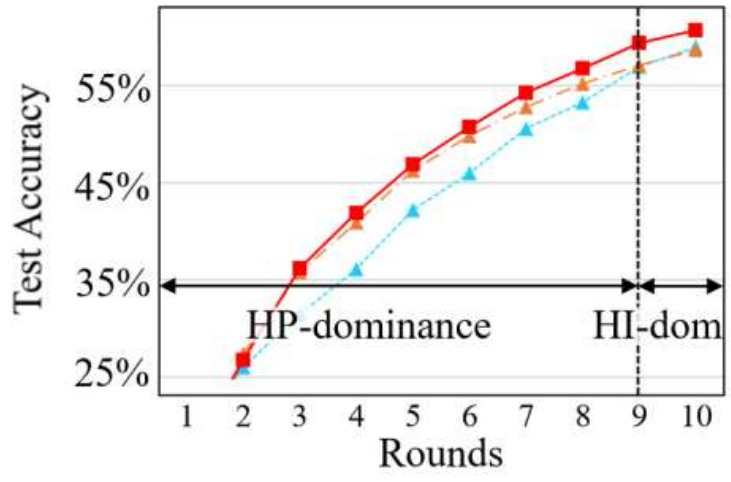
NeurIPS'22: MQNet



(a) Purity-informativeness Dilemma.



(b) 10% Open-set Noise.



(c) 30% Open-set Noise.

NeurIPS'22: MQNet

- $P(x)$: purity score of an example x
- $I(x)$: informativeness score of an example x
- $z_x = \langle P(x), I(x) \rangle$
- $\Phi(z_x)$: a score combination function

if $\mathcal{P}(x_i) > \mathcal{P}(x_j)$ and $\mathcal{I}(x_i) > \mathcal{I}(x_j)$,

⇩

$$\Phi(z_{x_i}) > \Phi(z_{x_j})$$

if $\mathcal{P}(x_i) > \mathcal{P}(x_j)$ and $\mathcal{I}(x_i) < \mathcal{I}(x_j)$

or $\mathcal{P}(x_i) < \mathcal{P}(x_j)$ and $\mathcal{I}(x_i) > \mathcal{I}(x_j)$,

⇩

?

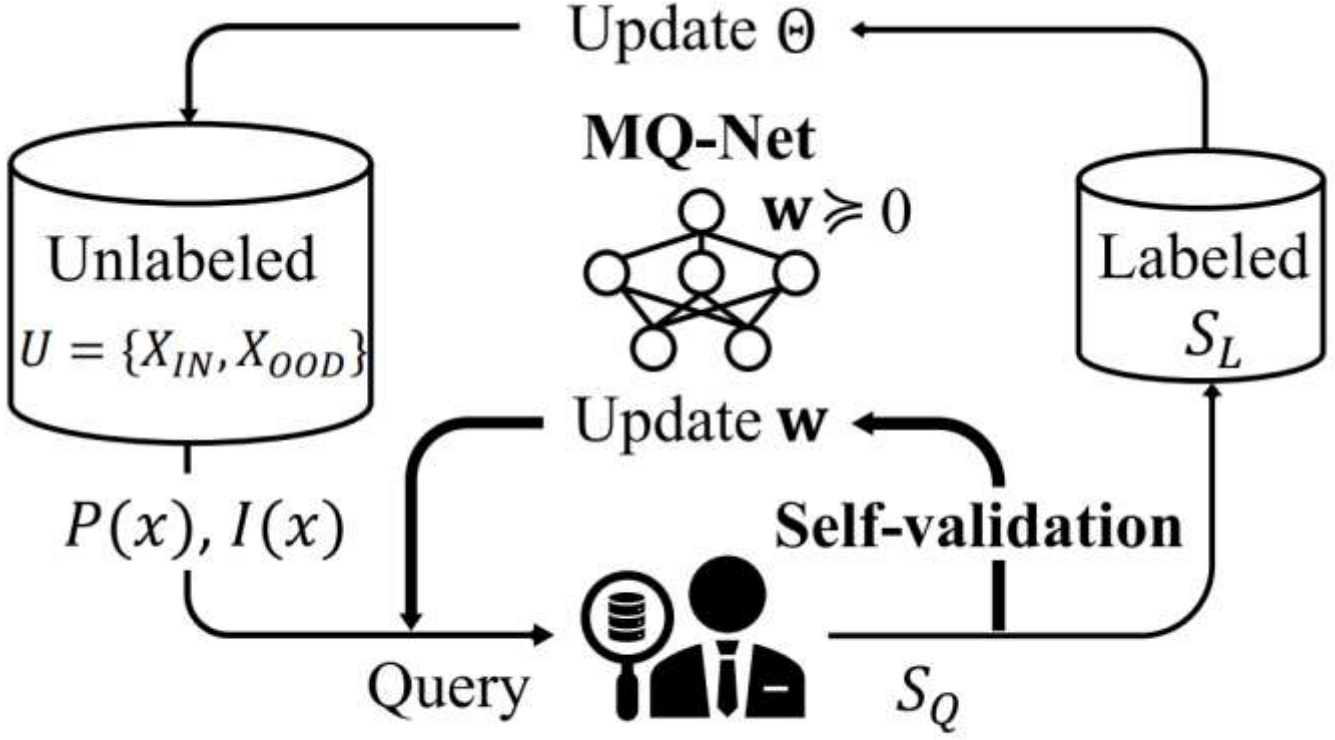


Figure 2: Overview of MQ-Net.

NeurIPS'22: MQNet

- Target model objective:

$$\ell_{mce}(x) = \mathbb{1}_{[l_x=1]}\ell_{ce}\big(f(x;\Theta), y\big)$$

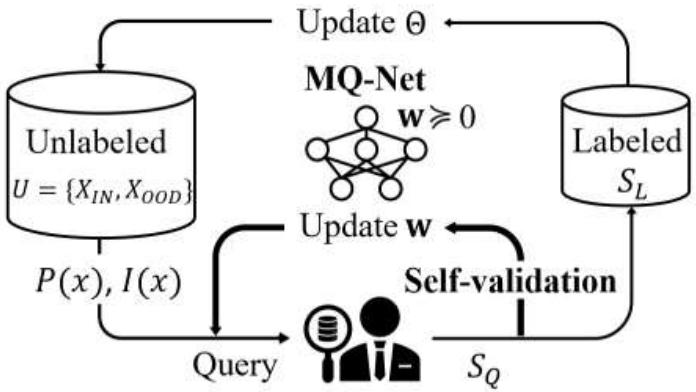


Figure 2: Overview of MQ-Net.

- Meta objective:

$$\mathcal{L}(S_Q) = \sum_{i\in S_Q}\sum_{j\in S_Q} \max\Big(0, -\text{Sign}\big(\ell_{mce}(x_i), \ell_{mce}(x_j)\big)\cdot\big(\Phi(z_{x_i};\mathbf{w}) - \Phi(z_{x_j};\mathbf{w}) + \eta\big)\Big)$$

$\Phi(z_{x_i};\, w)$ is forced to be higher than $\Phi(z_{x_j};\, w)$ if $l_{mce}(x_i) > l_{mce}(x_j)$

$\Phi(z_{x_i};\, w)$ is forced to be lower than $\Phi(z_{x_j};\, w)$ if $l_{mce}(x_i) < l_{mce}(x_j)$

$$s.t.\ \forall x_i, x_j,\ \text{if}\ \mathcal{P}(x_i) > \mathcal{P}(x_j)\ \text{and}\ \mathcal{I}(x_i) > \mathcal{I}(x_j),\ \text{then}\ \Phi(z_{x_i};\mathbf{w}) > \Phi(z_{x_j};\mathbf{w})$$

NeurIPS'22: MQNet

Table 1: Last test accuracy (%) at the final round for CIFAR10, CIFAR100, and ImageNet. The best results are in bold, and the second best results are underlined.

Datasets		CIFAR10 (4:6 split)				CIFAR100 (40:60 split)				ImageNet (50:950 split)			
Noise Ratio		10%	20%	40%	60%	10%	20%	40%	60%	10%	20%	40%	60%
Non-AL	RANDOM	89.83	89.06	87.73	85.64	60.88	59.69	55.52	47.37	62.72	60.12	54.04	48.24
Standard AL	CONF	<u>92.83</u>	91.72	88.69	85.43	62.84	60.20	53.74	45.38	63.56	<u>62.56</u>	51.08	45.04
	CORESET	91.76	91.06	89.12	86.50	63.79	62.02	56.21	48.33	63.64	62.24	55.32	49.04
	LL	92.09	91.21	<u>89.41</u>	86.95	<u>65.08</u>	<u>64.04</u>	56.27	48.49	63.28	61.56	55.68	47.30
	BADGE	92.80	<u>91.73</u>	89.27	86.83	62.54	61.28	55.07	47.60	<u>64.84</u>	61.48	54.04	47.80
Open-set AL	CCAL	90.55	89.99	88.87	<u>87.49</u>	61.20	61.16	<u>56.70</u>	50.20	61.68	60.70	<u>56.60</u>	51.16
	SIMILAR	89.92	89.19	88.53	87.38	60.07	59.89	56.13	<u>50.61</u>	63.92	61.40	56.48	<u>52.84</u>
Proposed	**MQ-Net**	**93.10**	**92.10**	**91.48**	**89.51**	**66.44**	**64.79**	**58.96**	**52.82**	**65.36**	**63.08**	**56.95**	**54.11**
% improve over 2nd best		0.32	0.40	2.32	2.32	2.09	1.17	3.99	4.37	0.80	1.35	0.62	2.40
% improve over the least		3.53	3.26	3.33	4.78	10.6	8.18	9.71	16.39	5.97	3.92	11.49	20.14

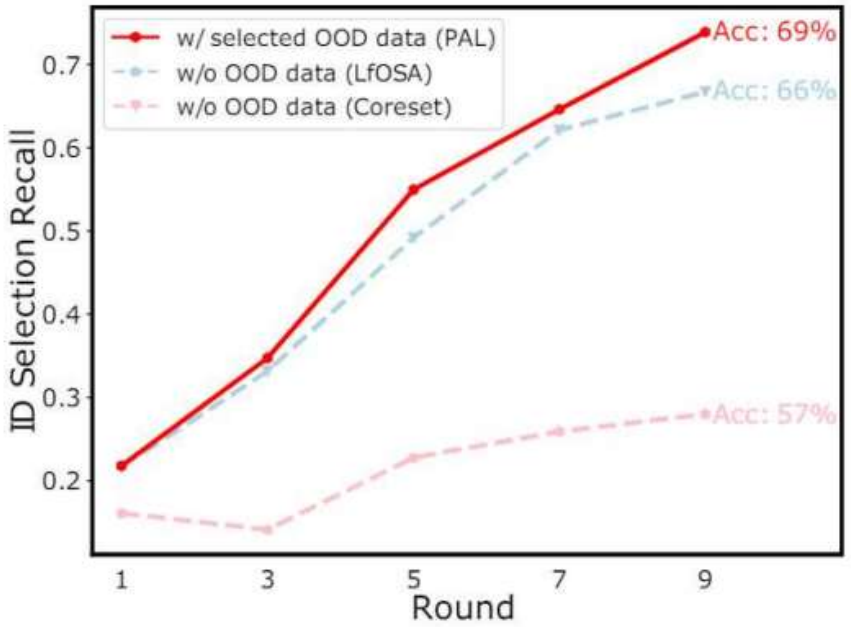
Not All Out-of-Distribution Data Are Harmful to Open-Set Active Learning

Yang Yang[1], **Yuxuan Zhang**[1], **Xin Song**[2], **Yi Xu**[3*]
[1]Nanjing University of Science and Technology
[2]Baidu Talent Intelligence Center, Baidu Inc
[3]Dalian University of Technology
{yyang, xuan_zhang}@njust.edu.cn, songxin06@baidu.com, yxu@dlut.edu.cn

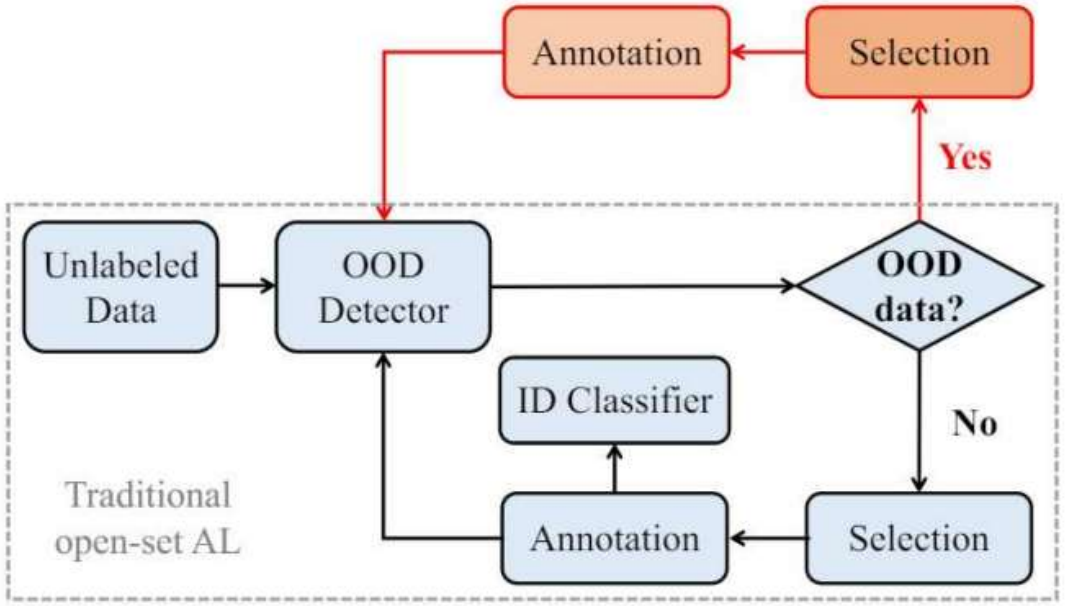
NeurIPS'23: PAL

Limitations:

- ✓ Existing methods mainly focus on learning the ID classifier, consequently restraining the resolution of the OOD detector.
- ✓ Valuable OOD instances can promptly strengthen the boundary between ID and OOD instances, and further improve the ID purity for classifier retraining.



(*a*) The power of OOD instances.



(*b*) The proposed PAL process

NeurIPS'23: PAL

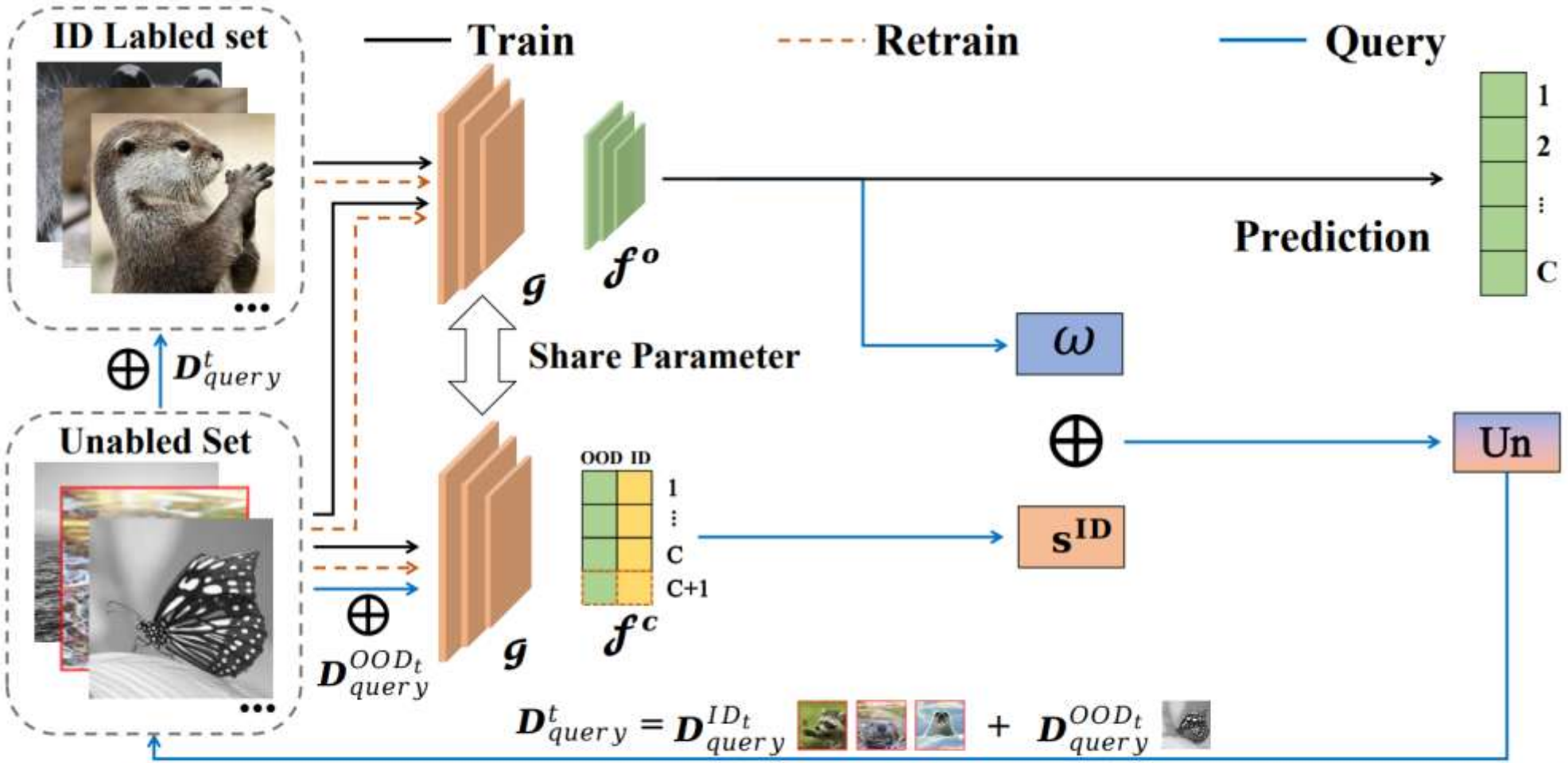


Figure 2: **Framework of the proposed Progressive Active Learning (PAL).** We initialized the ID classifier and OOD detector with the ID labeled data. During the query rounds, we explicitly query both pseudo-ID and pseudo-OOD instances with the designed sampling criterion to enhance the ID classifier and OOD detector.

NeurIPS'23: PAL

- Uncertainty weight

$$\ell_{OVA} = -\sum_{i=1}^{N_l}(\log \mathbf{p}^{y_{\mathbf{x}_i^l}}(p=1|\mathbf{x}_i^l) + \min_{c \neq y_{\mathbf{x}_i^l}} \log \mathbf{p}^c(p=0|\mathbf{x}_i^l))$$

$$s^{ID} = 1 + \mathbf{p}^{\hat{c}}(p=1|\mathbf{x}^u)\log \mathbf{p}^{\hat{c}}(p=1|\mathbf{x}^u)$$

$$\text{where } \hat{c} = \arg\max_c \mathbf{p}^c(p=1|\mathbf{x}^u).$$

- Meta-weight

$$\min_{\omega} \sum_{i=1}^{N_l} \ell_{sup}(\mathbf{x}_i^l, \mathbf{y}_i^l, \hat{\Theta}(\omega)),$$

$$s.t. \quad \hat{\Theta}(\omega) = \arg\min_{\Theta} \sum_{i=1}^{N_l} \ell_{sup}(\mathbf{x}_i^l, \mathbf{y}_i^l, \Theta) + \sum_{j=1}^{N_u} \omega_j \ell_{un}(\mathbf{x}_j^u, \Theta)$$

NeurIPS'23: PAL

- Progressive sampling

$$un(\mathbf{x}^u) = \omega + \mu s^{ID}$$

⇩

The selected b instances including N_{query}^{ID} pseudo-ID instances and N_{query}^{OOD} pseudo-OOD instances.

- At the first query, OOD instances with the lowest $un(x^u)$ are selected, $N_{query}^{OOD} \gg N_{query}^{ID}$.
- After that, the OVA classifiers are extended to $C+1$ classifiers, and the $un(x^u)$ is given by

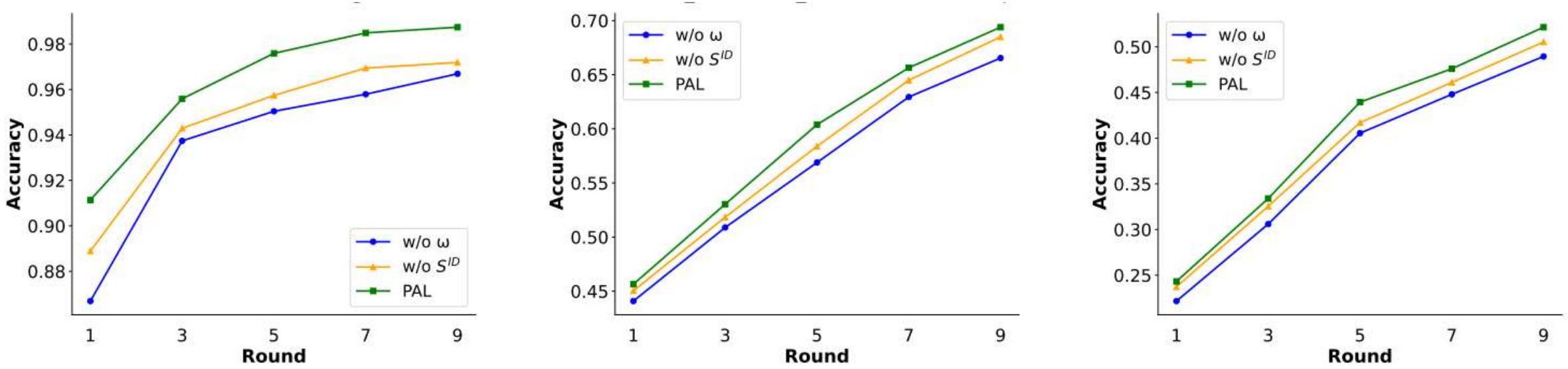
$$un(\mathbf{x}^u) = \begin{cases} \omega + \mu(1 - s^{ID}), & \text{if } \hat{c} = C+1 \\ \omega + \mu s^{ID}, & \text{otherwise} \end{cases}$$

NeurIPS'23: PAL

Table 1: Comparison of testing accuracy (%) on CIFAR-10, CIFAR-100, and Tiny-Imagenet datasets with an ID proportion of 20%. The best results are highlighted in bold, and the second-best results are underlined.

Datasets	CIFAR-10					CIFAR-100					Tiny-Imagenet				
Rounds	1	3	5	7	9	1	3	5	7	9	1	3	5	7	9
Label Only	78.6					42.5					21.4				
Random	88.1	93.1	95.9	96.9	97.4	44.6	49.2	52.2	54.5	56.7	22.9	26.8	32.7	37.8	39.2
Uncertainty	88.2	93.0	96.1	97.2	97.5	44.4	49.4	49.7	54.8	55.3	21.6	28.5	35.6	37.6	41.1
Certainty	88.1	90.8	91.6	92.4	93.0	45.0	50.7	53.1	54.4	54.9	21.7	27.9	35.4	39.8	44.7
Coreset	86.5	94.7	95.8	96.7	96.8	43.6	49.5	51.9	55.0	57.0	23.1	26.3	33.0	39.9	43.3
BALD	84.6	93.1	94.3	96.4	96.7	42.9	48.1	52.3	54.4	56.2	22.9	26.8	32.9	38.1	39.2
OpenMax	81.3	85.8	86.6	87.0	90.7	45.0	47.3	50.1	53.9	56.0	22.0	26.1	32.2	36.9	41.9
CCAL	88.9	94.3	96.2	97.4	97.7	45.1	50.9	53.4	57.2	60.4	23.2	28.5	35.6	40.6	44.9
MQnet	88.8	94.9	96.8	97.4	97.8	45.3	51.1	57.9	59.1	61.3	23.8	28.6	35.7	42.0	45.8
LfOSA	84.2	95.4	97.1	97.5	98.3	45.6	52.2	59.0	62.5	66.1	22.6	28.8	36.4	43.7	47.9
PAL	**91.1**	**95.6**	**97.6**	**98.5**	**98.7**	**45.6**	**53.0**	**60.0**	**65.6**	**69.4**	**24.3**	**33.4**	**43.9**	**47.6**	**52.1**

NeurIPS'23: PAL



(a) CIFAR-10 (b) CIFAR-100 (c) Tiny-Imagenet

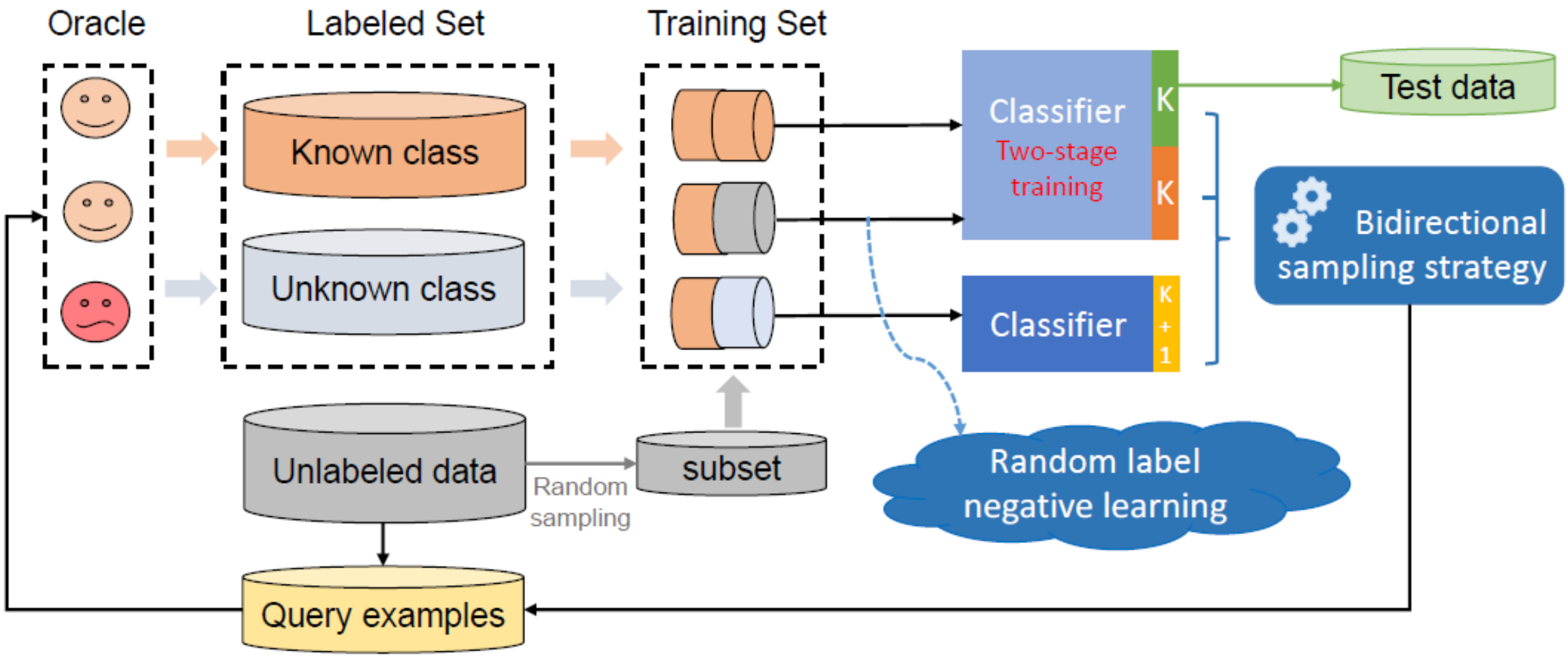
Figure 8: Ablation study conducted on CIFAR-10, CIFAR-100, and Tiny-Imagenet datasets with an ID proportion of 20%.

Bidirectional Uncertainty-Based Active Learning for Open-Set Annotation

Chen-Chen Zong[1], Ye-Wen Wang[1], Kun-Peng Ning[2], Hai-Bo Ye[1], and Sheng-Jun Huang[1]†

[1] Nanjing University of Aeronautics and Astronautics
[2] Peking University

ECCV'24: BUAL

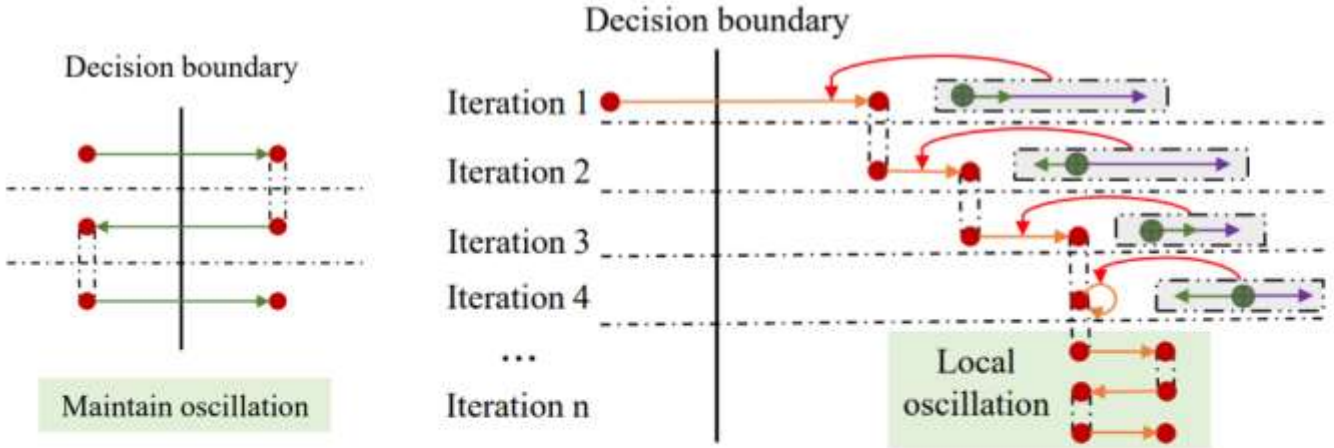


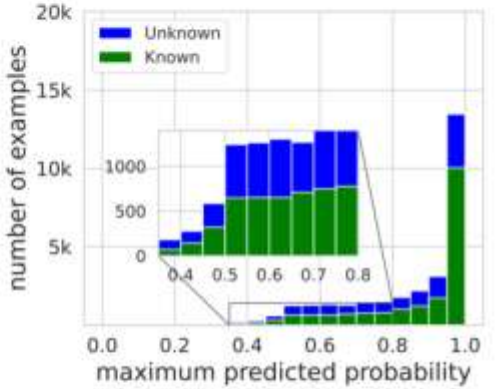
$$\ell_{NL}(f, \bar{y}) = -\sum_{k=1}^{K} \bar{y}_k \log(1 - p_k).$$

ECCV'24: BUAL

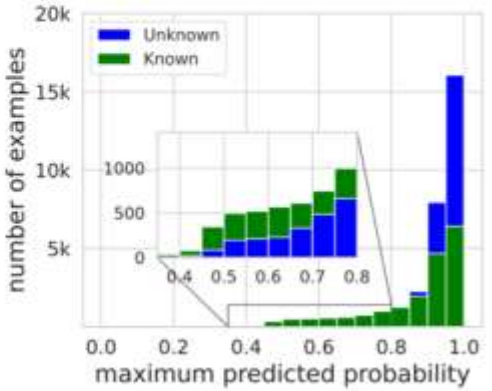
- 随机标签负学习微调 ⇒ 将未知类示例推向表征空间高置信度区域
  - 基于标记样本训练目标模型，随后重置分类器头参数
  - 对于标记样本，采样补标签进行负学习训练
  - 对于未标记样本，每个轮次随机赋予标签进行负学习训练

- 为什么随机标签负学习微调有效?





(a) w/o NL



(b) w/ NL

紫色箭头 →：标记样本的整体更新梯度

绿色箭头 →：样本自身提供的更新梯度

ECCV'24: BUAL

- 双向不确定性采样策略 ⇒ 将闭集策略扩展到开集设置
  - 基于随机标签进行微调随有效，但模型预测的稳定性不强
  - 在数据集开放性较低时，倾向于采用目标模型的预测不确定性
  - 在数据集开放性较高时，采用负学习微调模型的预测不确定性

$$\boldsymbol{x}^* = argmax_x p_{K+1}^{aux}(\boldsymbol{x}) unc_n + r\left[1 - p_{K+1}^{aux}(\boldsymbol{x})\right] unc_p$$

- 扩展闭集主动学习策略到开集设置
  - Bidirectional Least Confident Sampling (B-LC)

$$\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}} p_{K+1}^{aux}(\boldsymbol{x})\left[1 - \boldsymbol{\mathcal{P}}_{y^-}^{-}(\boldsymbol{x})\right] + r\left[1 - p_{K+1}^{aux}(\boldsymbol{x})\right]\left[1 - \boldsymbol{p}_{y^+}^{+}(\boldsymbol{x})\right]$$

  - Bidirectional Margin-Based Sampling (B-Margin)

$$\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}} p_{K+1}^{aux}(\boldsymbol{x})\left[\boldsymbol{\mathcal{P}}_{y_1^-}^{-}(\boldsymbol{x}) - \boldsymbol{\mathcal{P}}_{y_2^-}^{-}(\boldsymbol{x})\right] + r\left[1 - p_{K+1}^{aux}(\boldsymbol{x})\right]\left[\boldsymbol{p}_{y_1^+}^{+}(\boldsymbol{x}) - \boldsymbol{p}_{y_2^+}^{+}(\boldsymbol{x})\right]$$

  - Bidirectional Entropy-Based Sampling (B-Entropy)

$$\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}} p_{K+1}^{aux}(\boldsymbol{x})\left[-\boldsymbol{\mathcal{P}}_{y^-}^{-}(\boldsymbol{x}) \log \boldsymbol{\mathcal{P}}_{y^-}^{-}(\boldsymbol{x})\right] + r\left[1 - p_{K+1}^{aux}(\boldsymbol{x})\right]\left[-\boldsymbol{p}_{y^+}^{+}(\boldsymbol{x}) \log \boldsymbol{p}_{y^+}^{+}(\boldsymbol{x})\right]$$

ECCV'24: BUAL

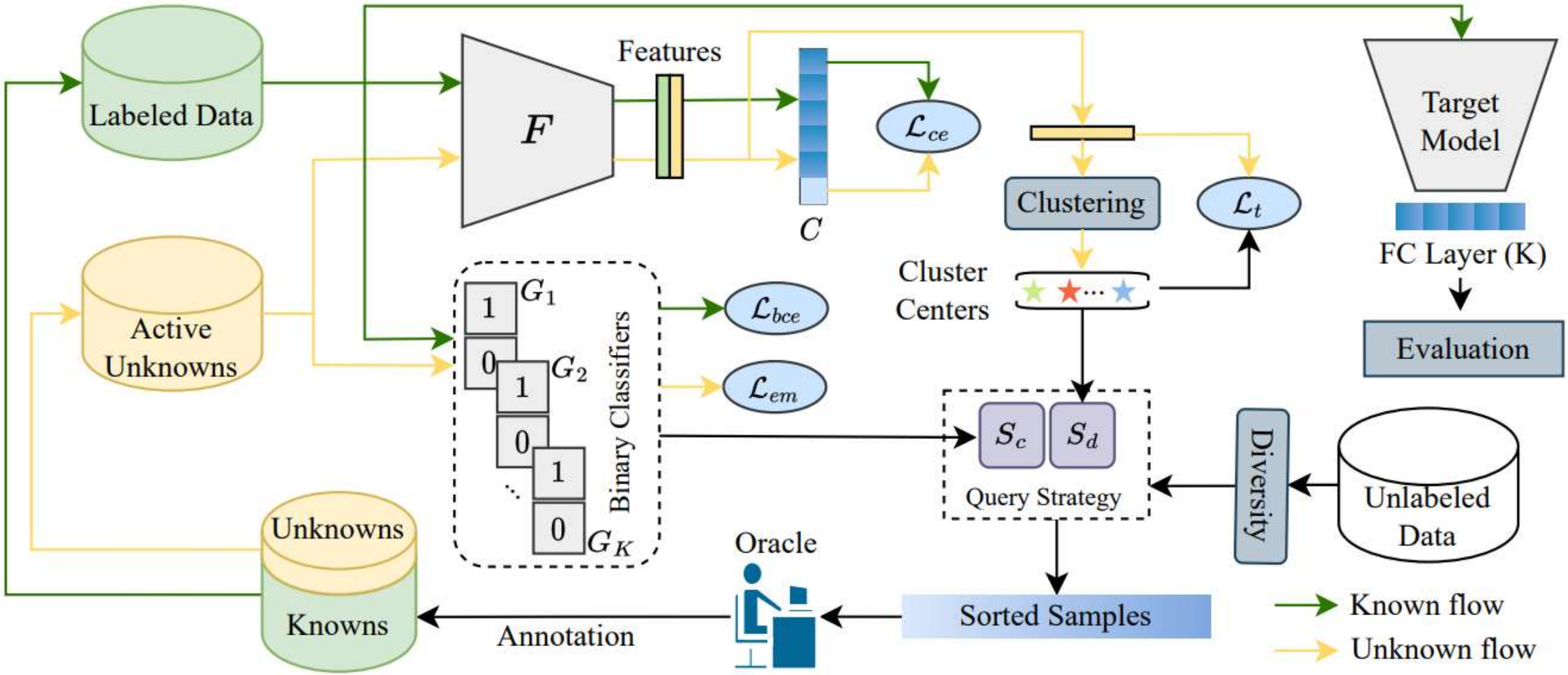
Datasets		CIFAR-10				CIFAR-100				Tiny-Imagenet			
Openness Ratio		0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
(1)	Random	83.3	82.5	87.2	96.9	57.6	58.3	58.7	61.2	45.7	47.2	50.9	55.0
(2)	LC	84.3	81.6	87.5	96.2	55.8	54.6	54.0	56.2	44.8	45.9	48.4	51.6
	Margin	86.0	84.1	89.0	97.0	59.3	59.6	58.8	58.9	46.4	47.1	50.8	54.0
	Entropy	85.4	83.4	88.0	96.8	57.1	56.8	55.7	56.4	44.6	44.5	46.9	50.7
(3)	Coreset	85.0	81.8	86.4	97.4	60.2	61.2	61.8	64.2	46.2	47.8	51.8	54.0
(4)	BADGE	86.8	84.2	89.2	96.4	60.2	60.8	60.4	62.0	46.3	47.8	51.8	53.3
(5)	LfOSA	73.7	78.7	87.0	98.6	52.3	56.6	62.4	68.2	42.5	46.6	52.4	59.9
	CCAL	80.8	81.5	88.0	98.1	55.9	60.0	64.7	67.7	44.4	46.3	50.3	57.0
(6)	DIAS	81.8	80.7	83.0	94.0	55.7	56.1	56.9	57.2	43.1	45.1	47.5	54.4
Ours	B-LC	**87.0**	87.2	92.5	**99.1**	59.3	62.8	67.5	**72.1**	45.7	48.7	54.7	60.6
	B-Margin	86.5	87.0	**92.6**	98.9	**60.9**	**63.1**	**68.3**	71.5	**46.5**	**49.5**	**55.7**	**61.2**
	B-Entropy	86.9	**87.4**	**92.6**	**99.1**	58.9	61.7	66.9	71.4	45.4	47.5	55.2	61.0

Entropic Open-Set Active Learning

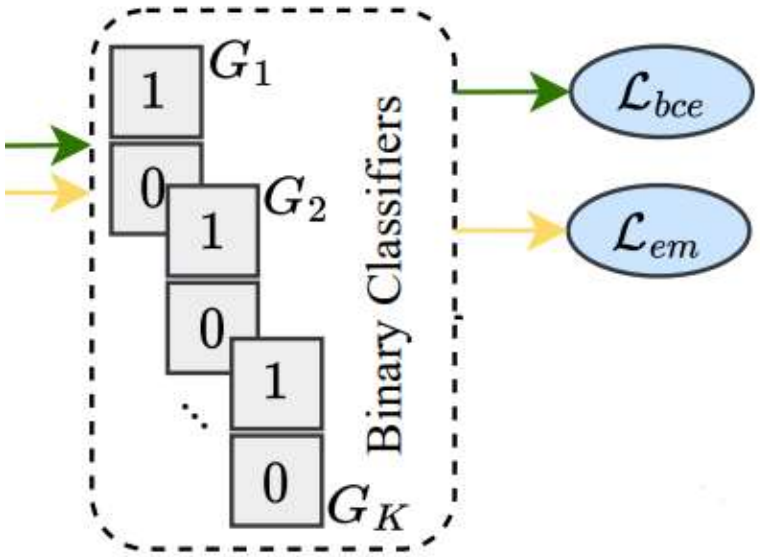
Bardia Safaei[1], Vibashan VS[1], Celso M. de Melo[2], Vishal M. Patel[1]

[1]Johns Hopkins University, Baltimore, MD, USA
[2]DEVCOM Army Research Laboratory, Adelphi, MD, USA
{bsafaei1, vvishnu2}@jhu.edu, celso.m.demelo.civ@army.mil, vpatel36@jhu.edu

AAAI'24: EOAL


Labeled Data
Active Unknowns
Unknowns
Knowns
F
Features
C
$\mathcal{L}_{ce}$
Clustering
$\mathcal{L}_t$
Cluster Centers
Target Model
FC Layer (K)
Evaluation
G_1
G_2
G_K
Binary Classifiers
$\mathcal{L}_{bce}$
$\mathcal{L}_{em}$
S_c
S_d
Query Strategy
Diversity
Unlabeled Data
Oracle
Annotation
Sorted Samples
Known flow
Unknown flow

AAAI'24: EOAL



$$\mathcal{L}_{bce} = \frac{1}{n_l} \sum_{(\mathbf{x}_i, y_i) \in \mathcal{D}_L} -\log(p^{y_i}) - \min_{j \neq y_i} \log(1 - p^j)$$

$$\mathcal{L}_{em} = \frac{1}{K \cdot n_{au}} \sum_{\mathbf{x} \in \mathcal{D}_{AU}} \sum_{i=1}^{K} -\frac{1}{2}\log(p^i) - \frac{1}{2}\log(1 - p^i)$$

- Closed-set Entropy

$$S_c(\mathbf{x}) = \frac{1}{K \cdot \log(2)} \sum_{i=1}^{K} H_i(\mathbf{x}),$$

where $H_i(\mathbf{x})$ denotes the entropy of G_i given as:

$$H_i(\mathbf{x}) = -p^i \cdot \log(p^i) - \left(1 - p^i\right) \cdot \log(1 - p^i).$$

AAAI'24: EOAL

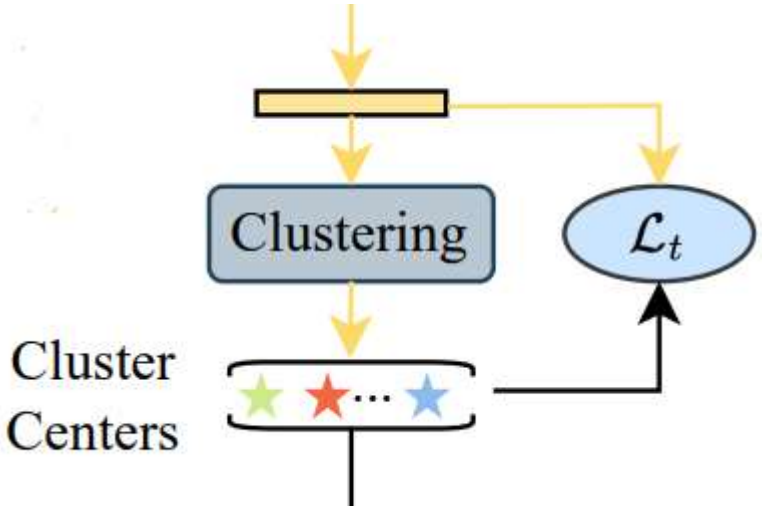
Distance-based Entropy

if $\mathcal{D}_{AU} \neq \emptyset$ **do**
  Cluster the features of $\mathcal{D}_{AU}$ into K clusters
  For cluster i, compute the center $\mathbf{c}_i$ using Eq. 5
  $\forall x \in D_U$, compute $S_d(x)$ via Eq. 6
end if

$$\mathbf{c}_i = \frac{\sum_{(\mathbf{x},\hat{y})\in\mathcal{D}_{AU}} \mathbb{I}\{\hat{y}=i\}\cdot F(\mathbf{x})}{\sum_{(\mathbf{x},\hat{y})\in\mathcal{D}_{AU}} \mathbb{I}\{\hat{y}=i\}}$$

$$S_d(\mathbf{x}) = \frac{-1}{\log(K)} \sum_{i=1}^{K} q_i(\mathbf{x}) \cdot \log(q_i(\mathbf{x})),$$

$$q_i(\mathbf{x}) = \frac{e^{-\|F(\mathbf{x})-\mathbf{c}_i\|/T}}{\sum_{j=1}^{K} e^{-\|F(\mathbf{x})-\mathbf{c}_j\|/T}},$$



$$\mathcal{L}_t = \frac{1}{n_{au}} \sum_{\mathbf{x}\in\mathcal{D}_{AU}} \log\left(1+\sum_{j\neq\hat{y}}^{K} e^{D_{\hat{y}}-D_j}\right) + \beta D_{\hat{y}},$$

where $D_i = \|F(\mathbf{x}) - \mathbf{c}_i\|$.

AAAI'24: EOAL

Query Strategy

$$S(\mathbf{x}) = S_c(\mathbf{x}) - S_d(\mathbf{x}).$$

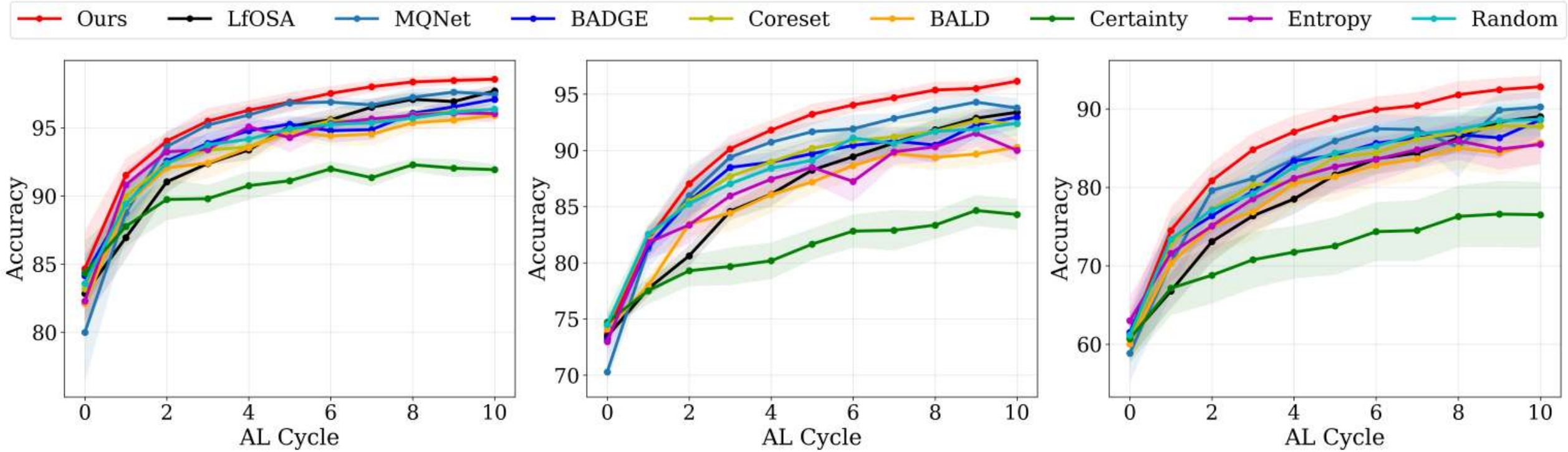


Figure 3: Classification accuracy comparison on CIFAR-10 (mismatch ratios from left to right: 20%, 30%, and 40%).

AAAI'24: EOAL

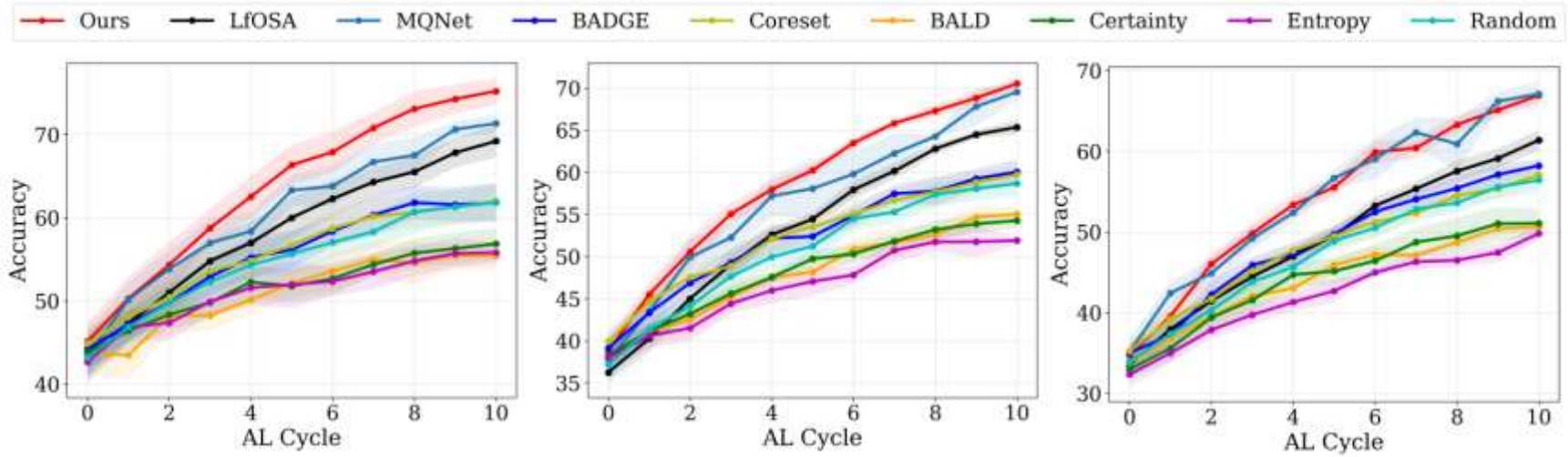


Figure 4: Classification accuracy comparison on CIFAR-100 (mismatch ratios from left to right: 20%, 30%, and 40%).

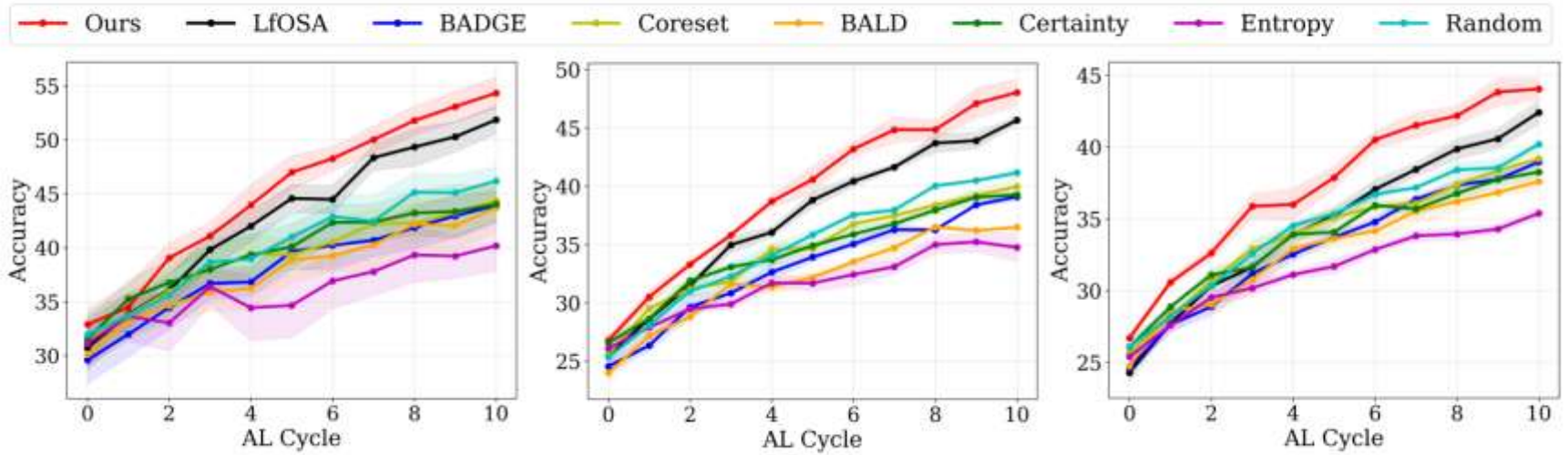


Figure 5: Classification accuracy comparison on TinyImageNet (mismatch ratios from left to right: 20%, 30%, and 40%).

Inconsistency-Based Data-Centric Active Open-Set Annotation

Ruiyu Mao, Ouyang Xu, Yunhui Guo

The University of Texas at Dallas
{ruiyu.mao, oxu, yunhui.guo}@utdallas.edu

AAAI'24: NEAT

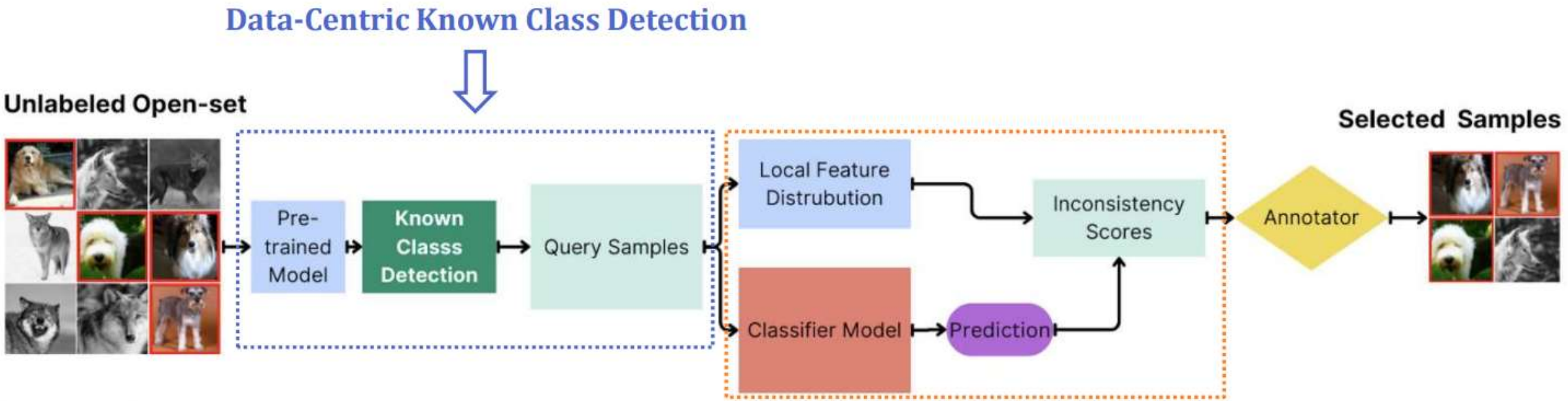


Figure 9: The architecture of NEAT includes a classifier model and a local feature distribution extractor, designed to identify both known and unknown samples.

AAAI'24: NEAT

Data-Centric Known Class Detection

- **Label clusterability.**
  - Samples with similar features should belong to the same class.

- **Feature extraction.**
  - CLIP provides high-quality features for calculating feature similarity.

- **Known class detection.**
  a. Each $N_k(x) \in L$ represents the k-th closest samples in L to the unlabeled example x in U.
  b. Compute the count of neighbors with known and unknown classes for each x.
  c. Classify an unlabeled sample x as belonging to a known class if all of its neighboring samples are from known classes.

AAAI'24: NEAT

Inconsistency-Based Active Learning

a. Given the K-nearest neighbors $\{N_i(x)\}_{i=1}^{K} \subset L$ of the example x.

b. Construct a vector $V_x \in \mathbb{R}^C$ with

$$V_x[c] = \sum_k 1(Y_k(x) = c).$$

c. Normalize V_x via the softmax function to be a probabilistic vector $\widetilde{V}_x$.

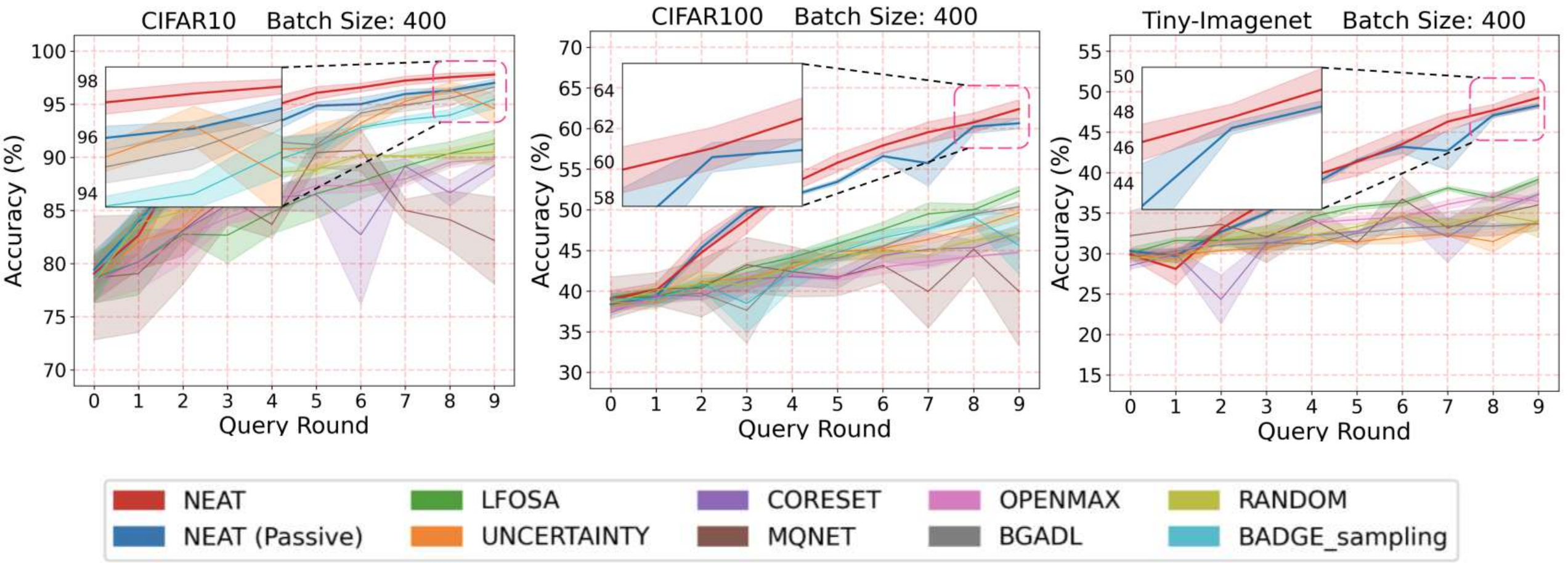
d. The inconsistency is computed using cross-entropy as

$$I(\mathbf{x}) = -\sum_{c=1}^{C} P_{\mathbf{x}}[c] \log \tilde{V}_{\mathbf{x}}[c].$$

e. Rank all the identified known classes examples using $I(x)$.

f. Select the top B examples for labeling.

AAAI'24: NEAT

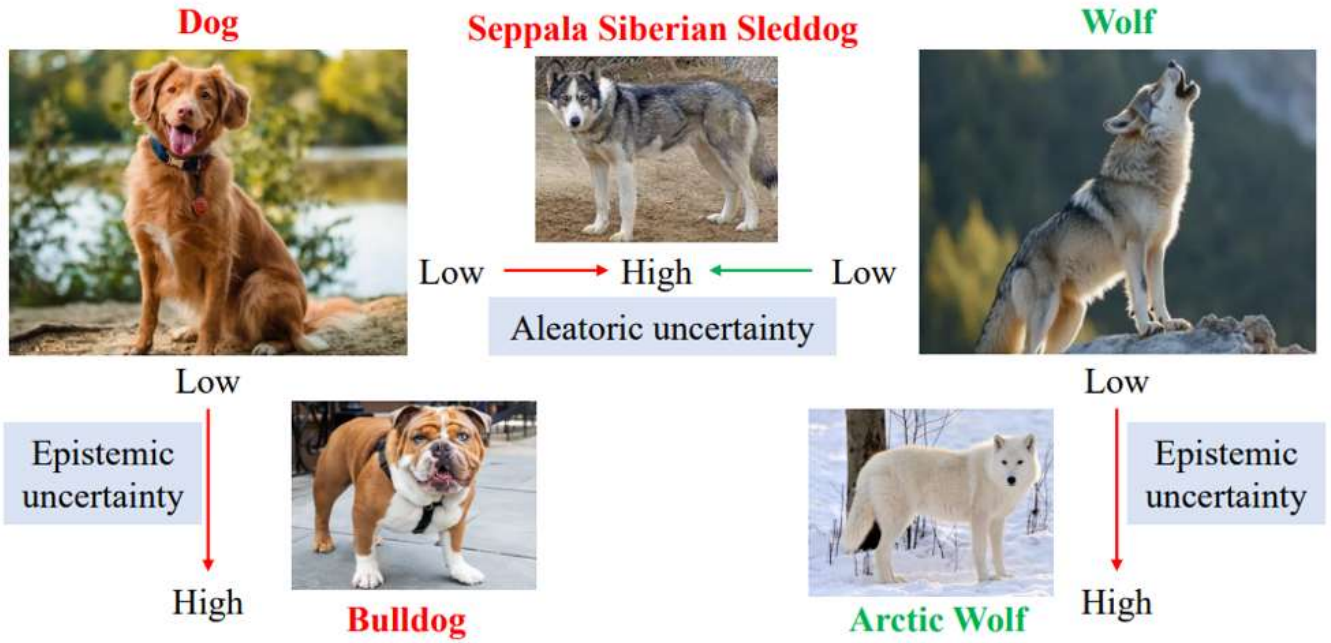

CIFAR10 Batch Size: 400
CIFAR100 Batch Size: 400
Tiny-Imagenet Batch Size: 400
Accuracy (%)
Query Round
NEAT
NEAT (Passive)
LFOSA
UNCERTAINTY
CORESET
MQNET
OPENMAX
BGADL
RANDOM
BADGE_sampling

Rethinking Epistemic and Aleatoric Uncertainty for Active Open-Set Annotation: An Energy-Based Approach

Chen-Chen Zong, Sheng-Jun Huang*
Nanjing University of Aeronautics and Astronautics
Nanjing, 211106, China
{chencz,huangsj}@nuaa.edu.cn

CVPR'25: EAOA

- 在开集环境下需同时考虑认知不确定性和数据不确定性

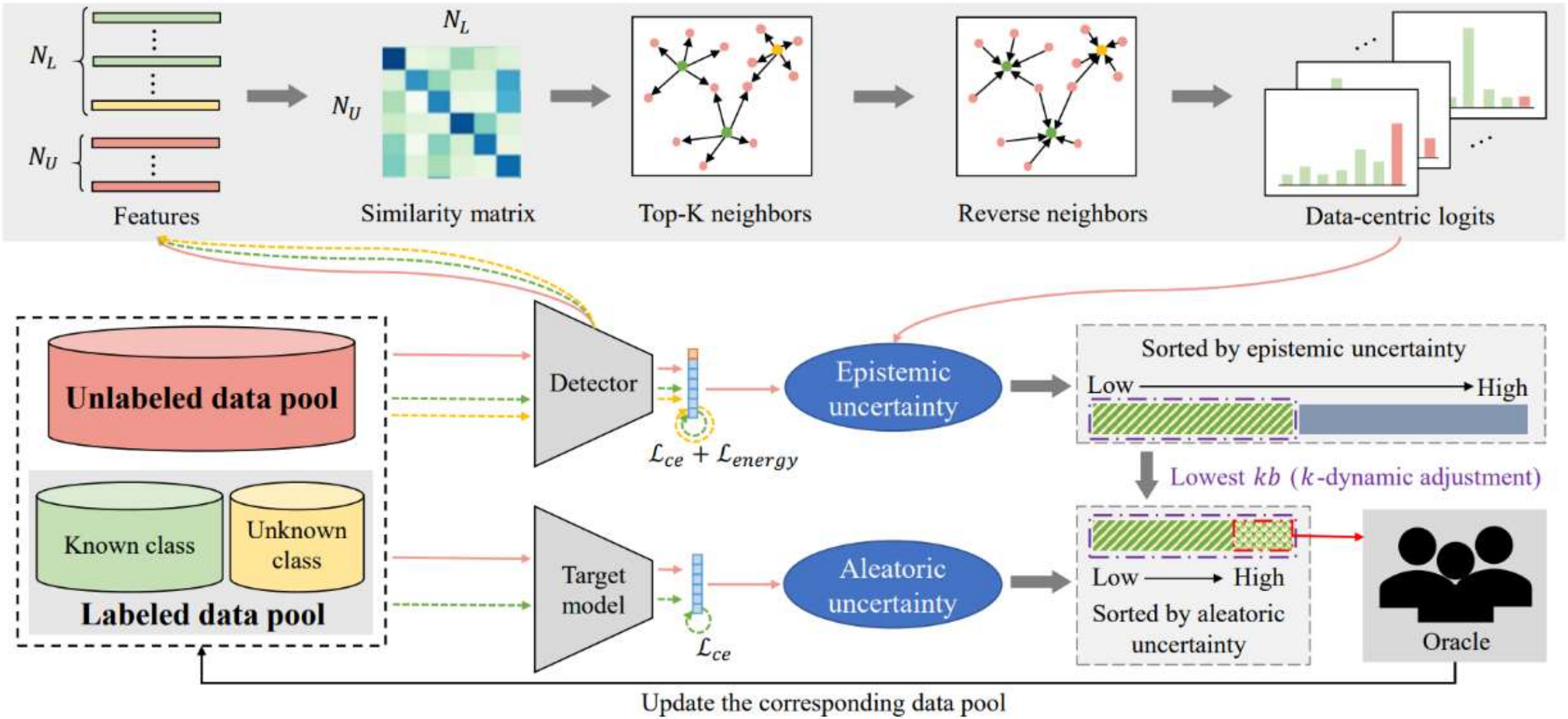


- 认知不确定性 ⇒ "模型不知道" ⇒ $p(x)$

数据不确定性仅在闭集设定下有意义

- 数据不确定性 ⇒ "数据本身就不确定" ⇒ $p(y|x) = \frac{p(x,y)}{p(x)}$

CVPR'25: EAOA


N_L
N_U
Features
N_L
N_U
Similarity matrix
Top-K neighbors
Reverse neighbors
Data-centric logits
Unlabeled data pool
Known class
Unknown class
Labeled data pool
Detector
$\mathcal{L}_{ce} + \mathcal{L}_{energy}$
Epistemic uncertainty
Sorted by epistemic uncertainty
Low
High
Lowest kb (k-dynamic adjustment)
Target model
$\mathcal{L}_{ce}$
Aleatoric uncertainty
Low
High
Sorted by aleatoric uncertainty
Oracle
Update the corresponding data pool

CVPR'25: EAOA

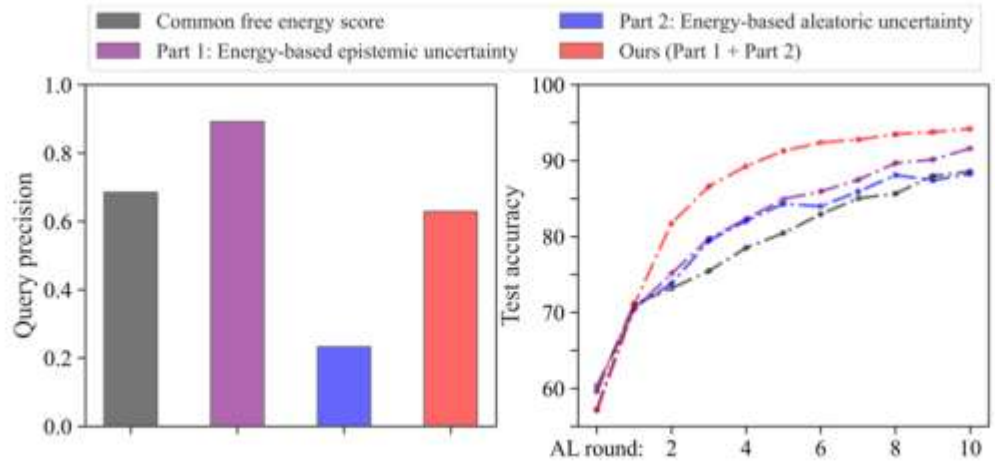
- 对于基于能量的模型（EBM）
  - 将 logits 用作多分类情形下的概率分布表示，得到：

$$p(y|x) = \frac{e^{-E(x,y)}}{\int_y e^{-E(x,y)}} = \frac{e^{f_y(x)}}{\sum_{c=1}^{C} e^{f_c(x)}} = \frac{e^{-E(x,y)}}{e^{-E(x)}}$$

  - 自由能量（Free energy）分数：$E(x) = -\log \sum_{c=1}^{C} e^{-E(x,c)}$
  - 在 EBM 框架下，样本 x 的概率密度可表示为：

$$p(x) = \frac{e^{-E(x)}}{\int_x e^{-E(x)}} = \frac{\int_y e^{-E(x,y)}}{\int_x \int_y e^{-E(x,y)}} = \frac{e^{-E(x)}}{\mathcal{Z}}.$$

- 存在问题：**能量分数直接应用于开集主动学习时性能不佳**



CVPR'25: EAOA

- 基于能量的认知不确定性 ⇨ 检测器

$$EU\left(x\right) = E_{kno}\left(x\right) - E_{unk}\left(x\right)$$
$$= -\log\sum_{c=1}^{C} e^{-E(x,c)} + \log\left(1 + e^{-E(x,C+1)}\right).$$

- 基于能量的数据不确定性 ⇨ 目标模型

$$AU(x) = E(x) - E_{secondary\ classes}(x)$$
$$= -\log\sum_{c=1}^{C} e^{-E(x,c)} + \log\left[\sum_{c=1}^{C} e^{-E(x,c)} - e^{-E(x,y_{max})}\right]$$

- 数据驱动的结构概率
  - 获取未标记池中样本在标记样本池中的逆K紧邻节点
  - 基于贝叶斯准则，x_i^U 在类别 y 上的概率可以近似表示为：

$$p\left(y|x_i^U\right) = \frac{\#\ \text{of Arrows}_{(x_j^L,y)}}{\sum_{c=1}^{C+1}\#\ \text{of Arrows}_{(x_j^L,c)}} = \frac{e^{-E\left(x_i^U,y\right)}}{e^{-E\left(x_i^U\right)}}$$

CVPR'25: EAOA

- 对于未标记样本 x，
  - 认知不确定性得分： $\widetilde{EU}(x) = GMM(EU_L(x)) \odot GMM(EU_D(x))$
  - 数据不确定性得分： $\widetilde{AU}(x) = GMM(AU(x))$

- 目标驱动的自适应采样策略
  - 首先挑选认知不确定性最低的 kb 个样本，以保证闭集属性
  - 进一步挑选数据不确定性最高的 b 个样本构造查询集

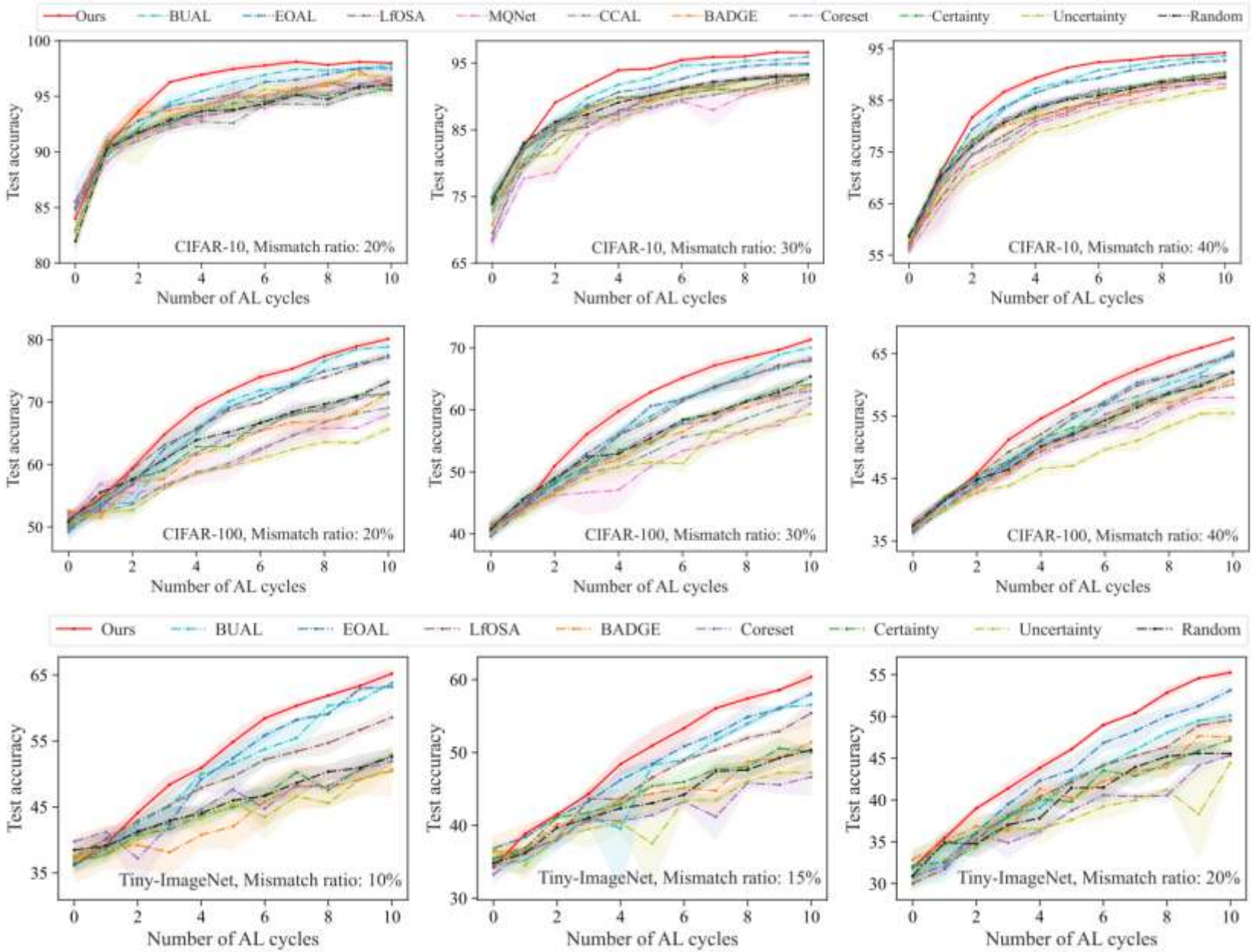
$$k_{t+1} = \begin{cases} k_t + a & \text{if } rP - tP > z, \\ k_t - a & \text{if } tP - rP > z, \\ k_t & \text{if } |tP - rP| \leq z, \end{cases}$$

rP 上轮查询查准率
tP 目标查询查准率

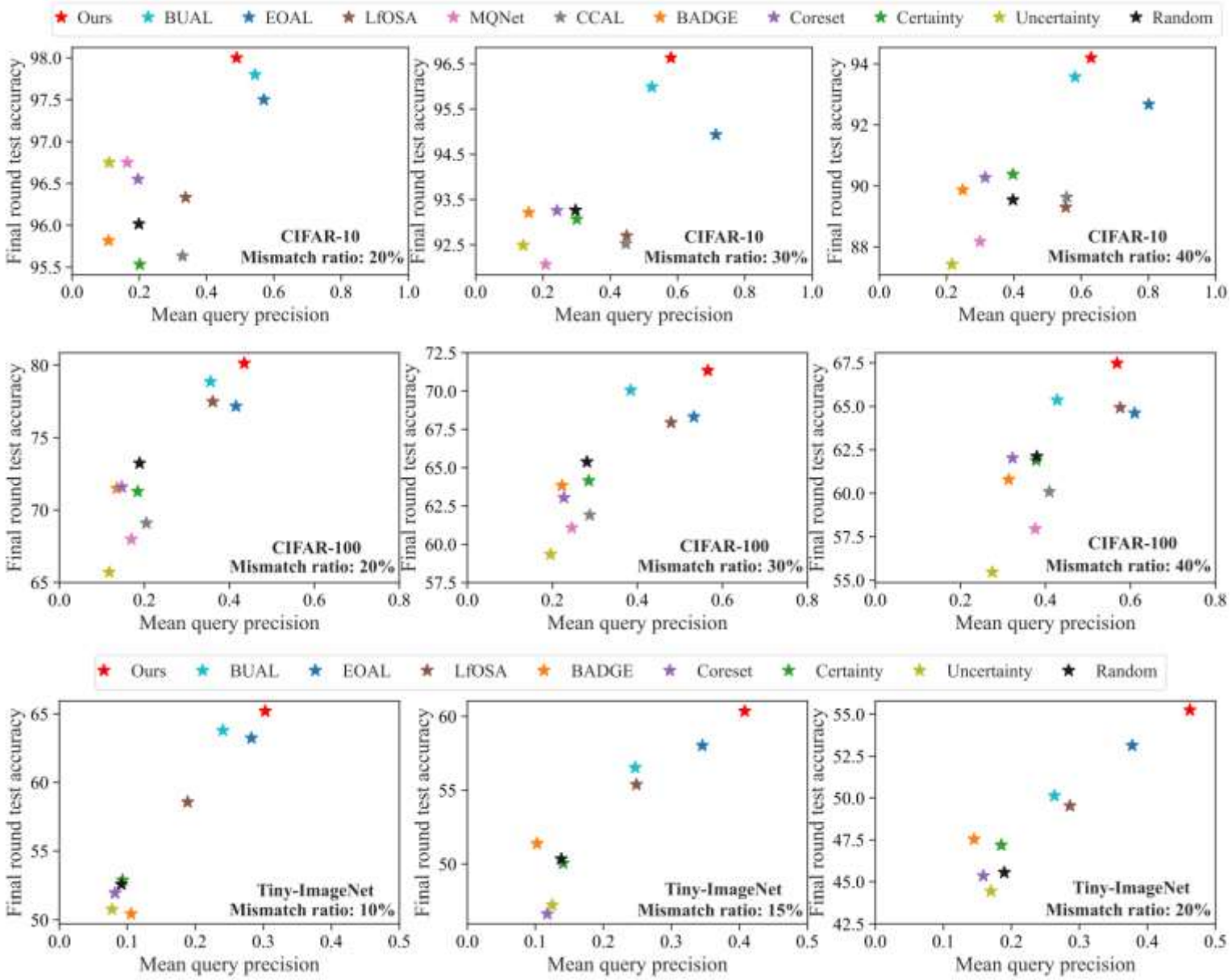
- 目标模型及检测器训练
  - 检测器采用交叉熵损失和能量损失训练
  - 目标模型采用交叉熵损失训练

$$\mathcal{L}_{energy}^{x_i} = \begin{cases} (\max(0, E_{kno}(x_i) - m_{kno}))^2 & \text{if } x_i \in \mathcal{D}_L^{kno} \\ (\max(0, m_{unk} - E_{kno}(x_i)))^2 & \text{if } x_i \in \mathcal{D}_L^{unk} \end{cases}$$

CVPR'25: EAOA


Ours
BUAL
EOAL
LfOSA
MQNet
CCAL
BADGE
Coreset
Certainty
Uncertainty
Random
Test accuracy
Number of AL cycles
CIFAR-10, Mismatch ratio: 20%
CIFAR-10, Mismatch ratio: 30%
CIFAR-10, Mismatch ratio: 40%
CIFAR-100, Mismatch ratio: 20%
CIFAR-100, Mismatch ratio: 30%
CIFAR-100, Mismatch ratio: 40%
Tiny-ImageNet, Mismatch ratio: 10%
Tiny-ImageNet, Mismatch ratio: 15%
Tiny-ImageNet, Mismatch ratio: 20%

CVPR'25: EAOA


Ours
BUAL
EOAL
LfOSA
MQNet
CCAL
BADGE
Coreset
Certainty
Uncertainty
Random
Final round test accuracy
Mean query precision
CIFAR-10
Mismatch ratio: 20%
CIFAR-10
Mismatch ratio: 30%
CIFAR-10
Mismatch ratio: 40%
CIFAR-100
Mismatch ratio: 20%
CIFAR-100
Mismatch ratio: 30%
CIFAR-100
Mismatch ratio: 40%
Tiny-ImageNet
Mismatch ratio: 10%
Tiny-ImageNet
Mismatch ratio: 15%
Tiny-ImageNet
Mismatch ratio: 20%

Evidence Conflict Sampling for Open-set Active Learning

Kun-Peng Ning[1] · Hai-Jian Ke[1] · Jia-Yu Yao[1] · Yu-Yang Liu[1] · Yong-Hong Tian[1,2] · Li Yuan[1,2]

IJCV'26: ECO

$$P(Y|\boldsymbol{x}) = P(Y|x^1, x^2, ..., x^m) = \frac{P(Y)\prod_j P(x^j|Y)}{P(x^1, x^2, ..., x^m)}$$

$$\Downarrow$$

$$Y = \arg\max_{y\in\mathcal{Y}} P(\boldsymbol{y}) \prod_j P(x^j|\boldsymbol{y}).$$

$$\Downarrow$$

$$P(x^j|\boldsymbol{y}) = \frac{1}{\sqrt{2\pi\sigma^{j2}}} e^{-\frac{(x^j-\mu^j)^2}{2\sigma^{j2}}} \qquad \mu^j = \frac{1}{N_L}\sum_{i=1}^{N_L} x_i^j, \quad \sigma^j = \sqrt{\frac{1}{N_L}\sum_{i=1}^{N_L}(x_i^j - \mu^j)^2}.$$

$$\Downarrow$$

$$\mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{y}} \triangleq \{x^j | \arg\max_{y\in\mathcal{Y}} P(x^j|\boldsymbol{y})\}$$

$$\Downarrow$$

$$E_{\boldsymbol{y}}(\boldsymbol{x}) = \prod_{x^j\in\mathcal{C}_{\boldsymbol{x}}^{\boldsymbol{y}}} e^{P(x^j|\boldsymbol{y})} \qquad \hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}\in\mathcal{D}_U} \prod_{\boldsymbol{y}\in\mathcal{Y}} E_{\boldsymbol{y}}(\boldsymbol{x})$$

IJCV'26: ECO

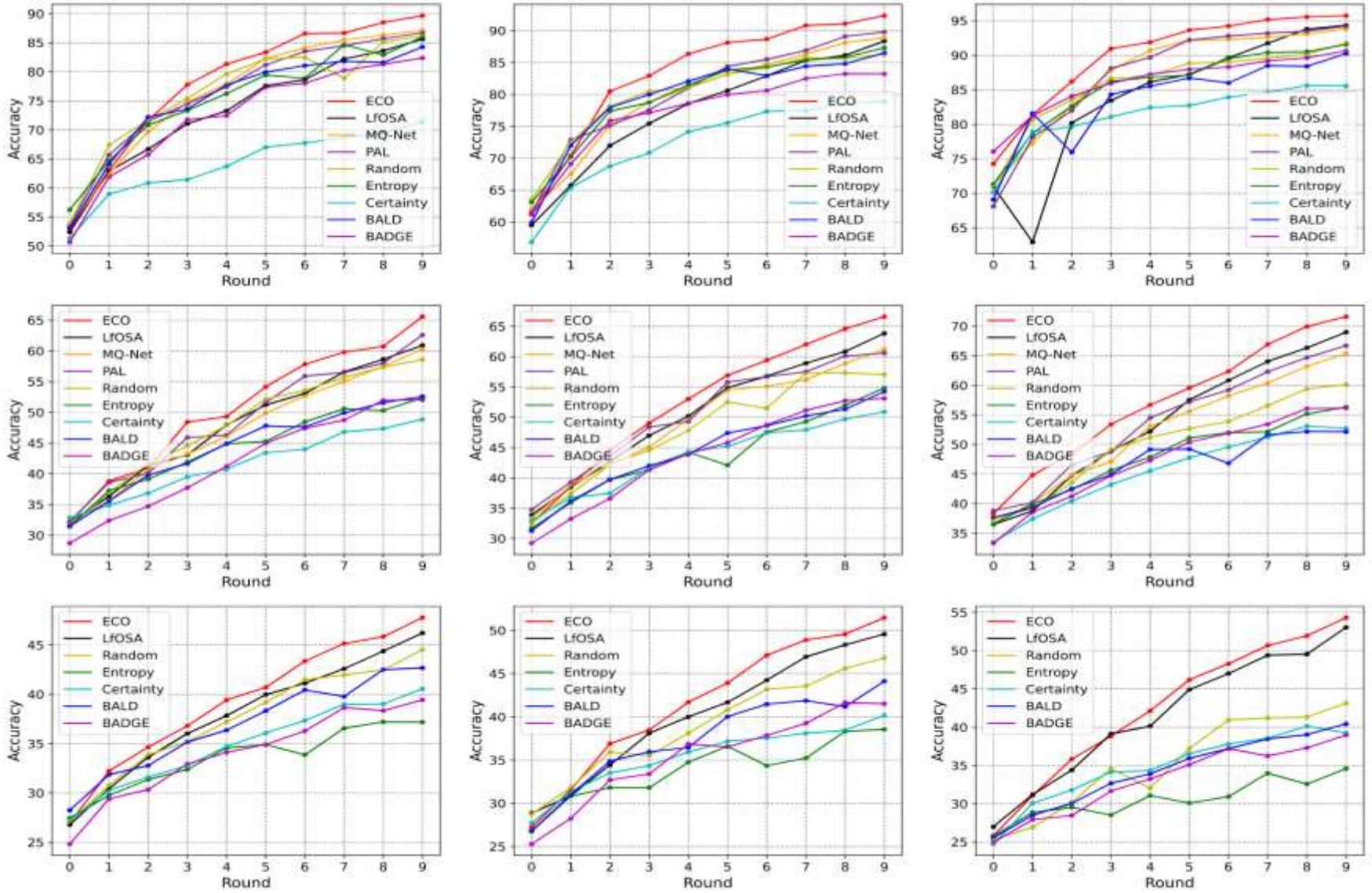


Fig. 3 Comparison of testing accuracy(%) with different methods on CIFAR10 (first row), CIFAR100 (second row) and Tiny-ImageNet (third row) under 50% (first column), 60% (second column) and 70% (third column) openness ratios

CVPR'26: E2OAL

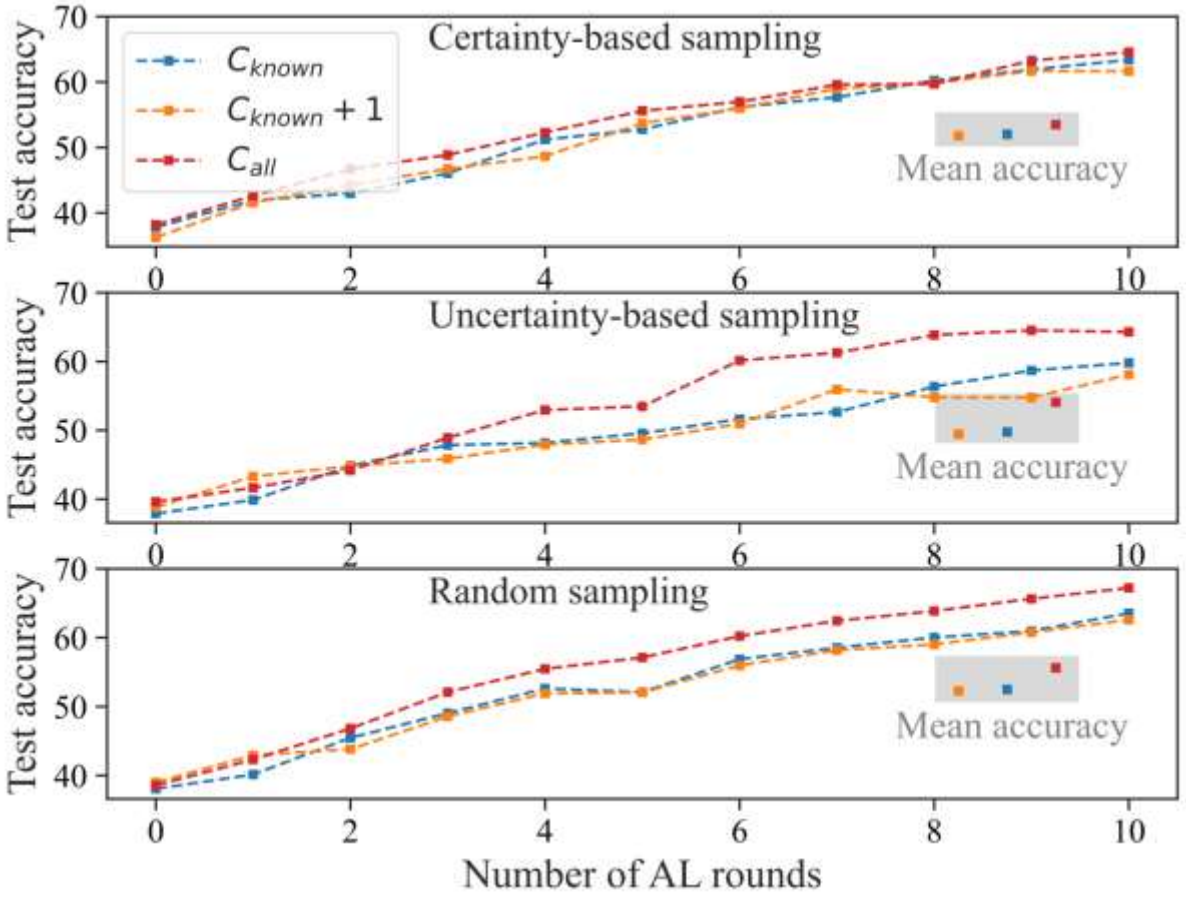
南京航空航天大学
Nanjing University of Aeronautics and Astronautics

Revisiting Unknowns: Towards Effective and Efficient Open-Set Active Learning

Chen-Chen Zong, Yu-Qi Chi, Xie-Yang Wang, Yan Cui, Sheng-Jun Huang*
Nanjing University of Aeronautics and Astronautics
Nanjing, 211106, China
{chencz,chiyuqi,xieyang,cuiyan,huangsj}@nuaa.edu.cn

CVPR'26: E2OAL

- Existing methods often rely on separately trained OOD detectors, which incur substantial training overhead, and further overlook the potential of labeled unknowns as valuable supervision for enhancing known-class learning.



CVPR'26: E2OAL

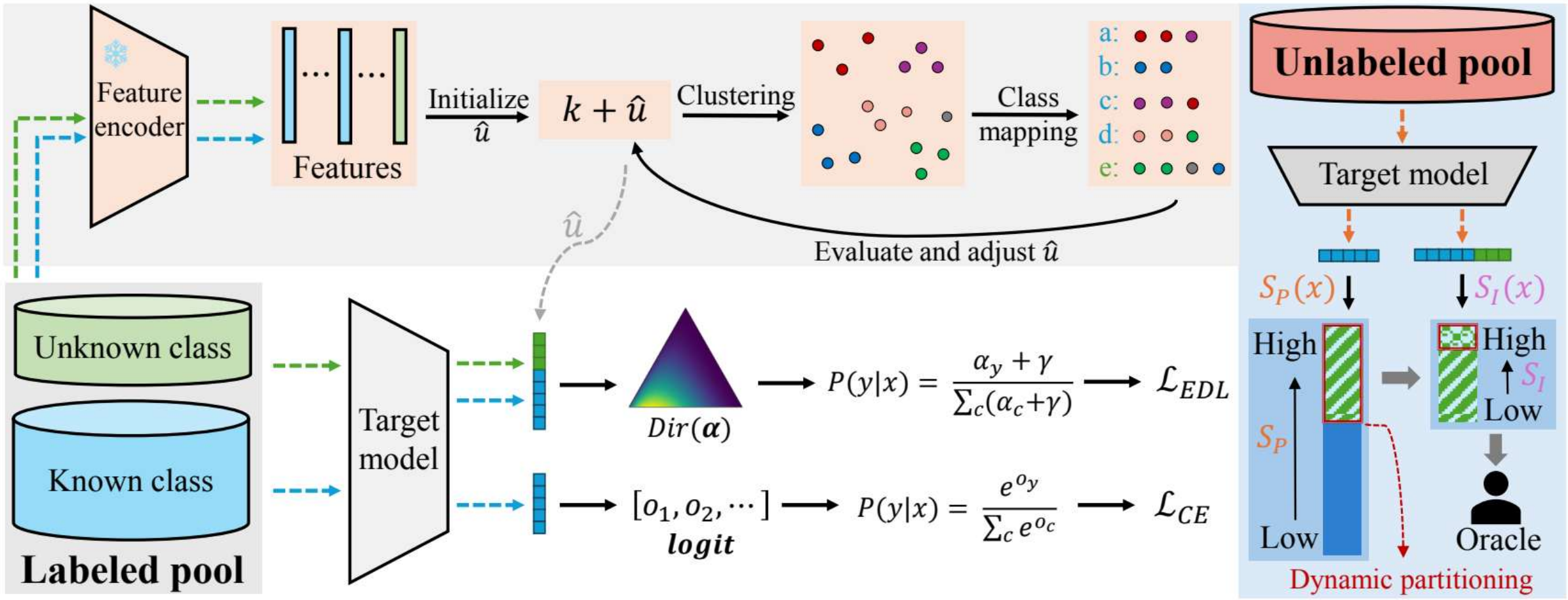


Figure 2. Overview of the proposed E^{2}OAL framework. Each AL round consists of two stages: (1) Adaptive class estimation and calibration-aware training, where latent unknown classes are discovered via label-guided clustering and incorporated into model learning through Dirichlet-based auxiliary supervision; (2) Flexible two-stage query selection, where a high-purity candidate pool is first constructed using a purity score guided by a target query precision, followed by informativeness-driven sample selection.

CVPR'26: E2OAL

Algorithm 1 The adaptive class estimation algorithm

Input: Labeled data pool $\mathcal{D}_L = \{(x_i, y_i)\}$, known class count k, and upper limit $\hat{u}_{max}$

Output: Estimated number of unknown classes $\hat{u}$

Extract CLIP features $\{f_i\}$ for all $x_i \in \mathcal{D}_L$
 # Ternary search for optimal unknown class count
Initialize bounds: $l \leftarrow k+1, r \leftarrow \hat{u}_{max}$
while $r - l > 2$ **do**
  $m_1 \leftarrow \lfloor l + \frac{r-l}{3} \rfloor, \quad m_2 \leftarrow \lfloor r - \frac{r-l}{3} \rfloor$
  $s_{m_1} \leftarrow \text{EVALUATE}(m_1, \{f_i\}, k)$
  $s_{m_2} \leftarrow \text{EVALUATE}(m_2, \{f_i\}, k)$
  if $s_{m_1} < s_{m_2}$ **then**
    $l \leftarrow m_1$
  else
    $r \leftarrow m_2$
  end if
end while
$\hat{u} \leftarrow \arg\max_{m \in \{l, l+1, \ldots, r\}} \text{EVALUATE}(m, \{f_i\}, k)$

Function EVALUATE(m, $\{f_i\}$, k):

$C \leftarrow k + m$
Perform K-Means clustering on $\{f_i\}$ into C clusters
Match the $k+1$ clusters to $k+1$ classes—including all k known classes and a unified unknown class—using the Hungarian algorithm [15]
Assign the remaining $(C - (k+1))$ clusters to the unknown class as well
Compute class-wise F1 scores $\{\text{F1}_c\}_{c=1}^{k+1}$
return $\prod_{c=1}^{k+1} \text{F1}_c$

CVPR'26: E2OAL

Calibrated probability

logits $[0, 5, 0, 0]$ and $[-5, 0, -5, -5]$

roughly 0.88 ⇨ $P(y|x) = \frac{e^{o_y} + \gamma}{\sum_{c=1}^{k+\hat{u}} (e^{o_c} + \gamma)}$ ⇨ 0.60 and 0.38

Evidential Deep Learning (EDL)

$$P(y|x) = \mathbb{E}_{\boldsymbol{p} \sim \mathrm{Dir}(\boldsymbol{\alpha})}[p_y] = \frac{\alpha_y}{\sum_{c=1}^{k+\hat{u}} \alpha_c} \longrightarrow \boldsymbol{\alpha} = \frac{g(\boldsymbol{o})}{\gamma} + 1$$

$$\mathcal{L}_{\mathrm{EDL}} = \mathcal{L}_{\mathrm{NLL}} + \mathcal{L}_{\mathrm{KL}}$$

$$\mathcal{L}_{\mathrm{NLL}} = -\log P(y|x) = -\log \frac{\alpha_y}{\sum_{c=1}^{k+\hat{u}} \alpha_c}$$

$$\mathcal{L}_{\mathrm{KL}} = \mathrm{KL}\big(\mathrm{Dir}(\tilde{\boldsymbol{\alpha}}) \,\|\, \mathrm{Dir}(\mathbf{1})\big)$$

$$\tilde{\boldsymbol{\alpha}} = \boldsymbol{y} + (1 - \boldsymbol{y}) \odot \boldsymbol{\alpha}$$

CVPR'26: E2OAL

- Purity metric

$$S_{\text{purity}}(x) = o_{(1)} - o_{(2)} = \max_{c \in \mathcal{C}_k} o_c - \max_{c \in \mathcal{C}_{\hat{u}}} o_c$$

- Informativeness metric

$$S_{\text{info}}(x) = \text{JS}\big(\mathbf{p} \parallel \mathbf{u}\big) \cdot \text{JS}\big(\mathbf{p} \parallel \mathbf{p}^{\max}\big), \tag{7}$$

where $\mathbf{p}$ is the predicted probability vector from the primary head, $\mathbf{u}$ is the uniform distribution, and $\mathbf{p}^{\max}$ is the one-hot encoding of the most confident class in $\mathbf{p}$.

- Flexible two-stage selection scheme
  - a high-purity candidate pool is first constructed using a purity score guided by a target query precision.

$$\hat{p}^*_{t+1} = \begin{cases} \max\big(\min(\hat{p}^*_t + (p^* - \bar{p}^*_t), 0), 1\big) & \text{if } t > 0, \\ p^* & \text{if } t = 0. \end{cases}$$

  - then selecting the most informative samples for annotation.

CVPR'26: E2OAL

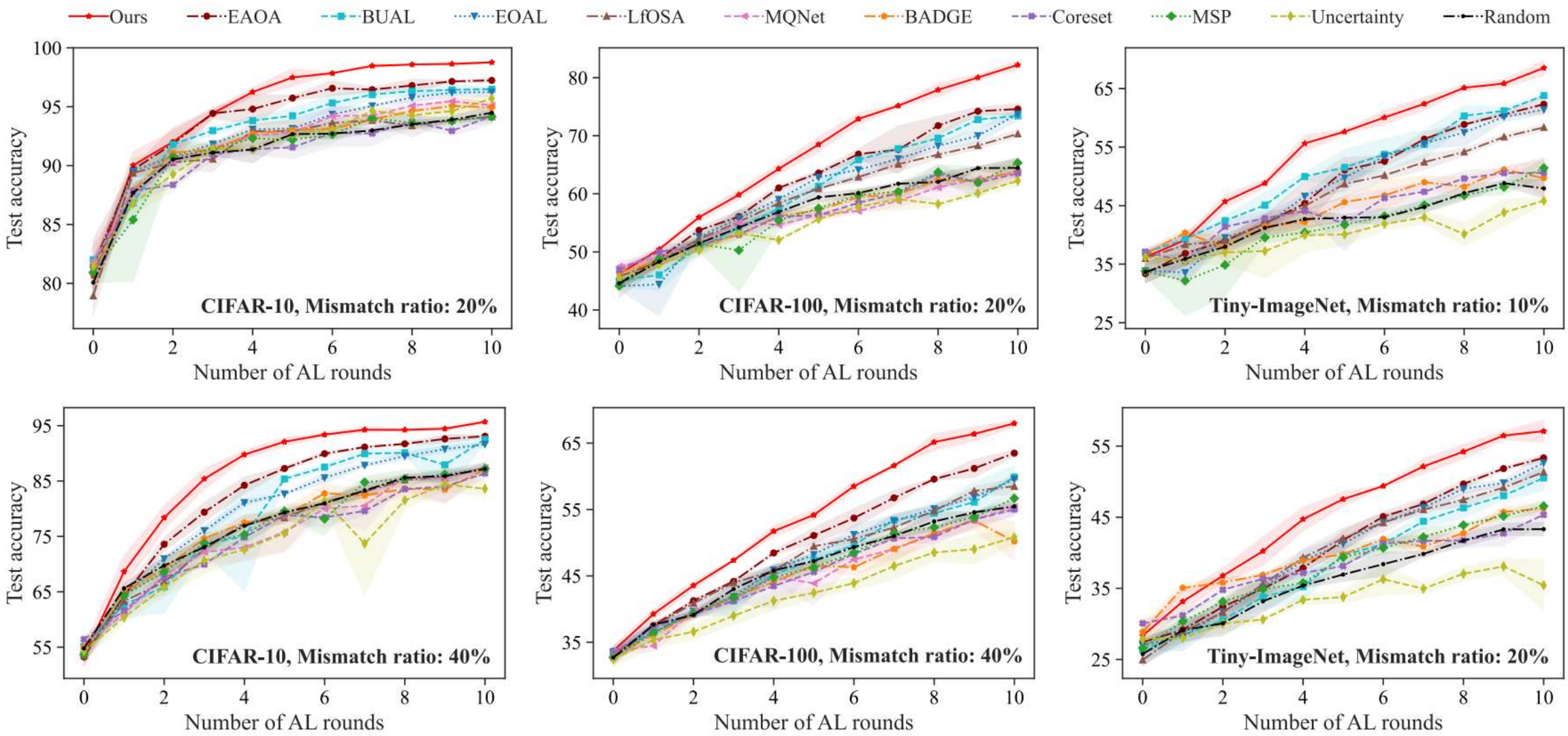


Figure 3. Test accuracy across AL rounds under varying mismatch ratios on CIFAR-10/100 and Tiny-ImageNet.

CVPR'26: E2OAL

Dataset	CIFAR-10			CIFAR-100			Tiny-ImageNet		
Mismatch ratio	20%	30%	40%	20%	30%	40%	10%	15%	20%
Random	94.48	91.11	87.18	64.45	58.48	55.48	47.93	46.00	43.32
Uncertainty [18]	95.70	89.77	83.61	62.25	53.52	50.83	45.83	43.40	35.43
Coreset [28]	94.20	89.56	86.38	63.53	56.62	55.00	50.60	47.33	45.35
BADGE [2]	94.95	90.91	87.12	64.00	56.49	50.20	49.70	48.16	46.23
MSP [11]	94.15	91.51	87.21	65.33	58.69	56.68	51.43	47.78	46.57
LfOSA [20]	94.15	90.91	87.43	70.32	62.49	58.49	58.37	54.78	51.33
MQNet [21]	95.12	89.39	87.42	63.70	53.52	55.44	-	-	-
EOAL [27]	96.23	93.64	91.63	73.73	63.69	59.55	61.40	56.13	52.65
BUAL [42]	96.48	95.04	92.52	73.43	63.73	59.89	63.80	56.09	50.52
EAOA [41]	97.23	95.88	93.09	74.60	67.14	63.49	62.33	57.31	53.33
Ours*	97.33	95.94	93.13	75.90	67.54	63.85	64.23	60.44	54.73
↑ over best baseline (%)	0.10	0.06	0.04	1.30	0.40	0.36	1.15	3.13	1.40
Ours	**98.77**	**97.52**	**95.69**	**82.20**	**72.10**	**67.98**	**68.53**	**64.02**	**57.10**
↑ over best baseline (%)	1.44	1.64	2.60	7.60	4.96	4.49	4.73	6.71	3.77

Table F1. Final-round test accuracy (%) of all methods under varying mismatch ratios on CIFAR-10, CIFAR-100, and Tiny-ImageNet. "Ours*" denotes a variant of our method where the target classifier is trained independently without leveraging labeled unknowns. The best result in each setting is highlighted in bold, while the second best is underlined. Due to the poor performance and high training cost of MQNet, we do not include it on Tiny-ImageNet, and thus mark it with "–".

CVPR'26: E2OAL

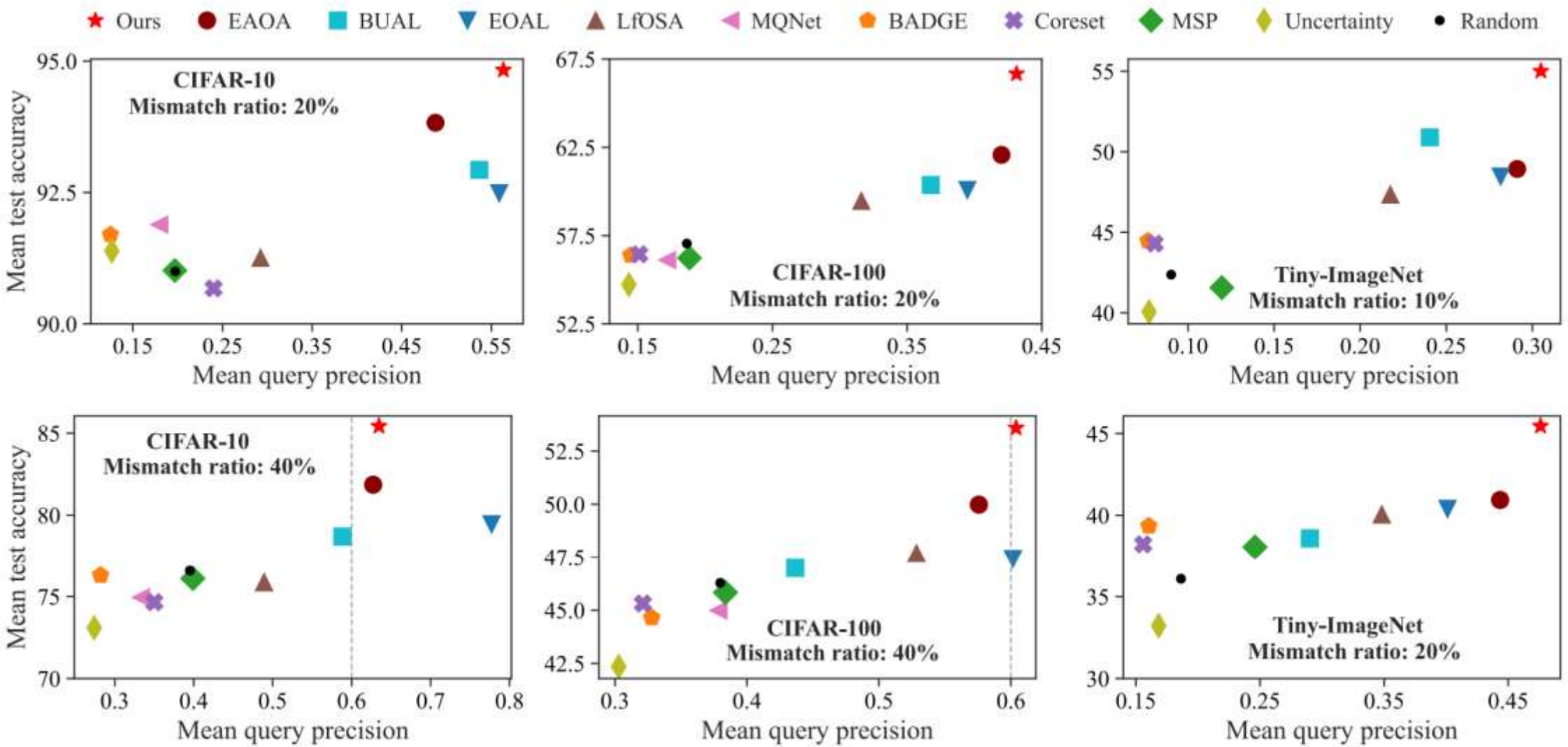


Figure 4. Mean query precision vs. mean test accuracy across rounds under varying mismatch ratios on CIFAR-10/100 and Tiny-ImageNet.

CVPR'26: E2OAL

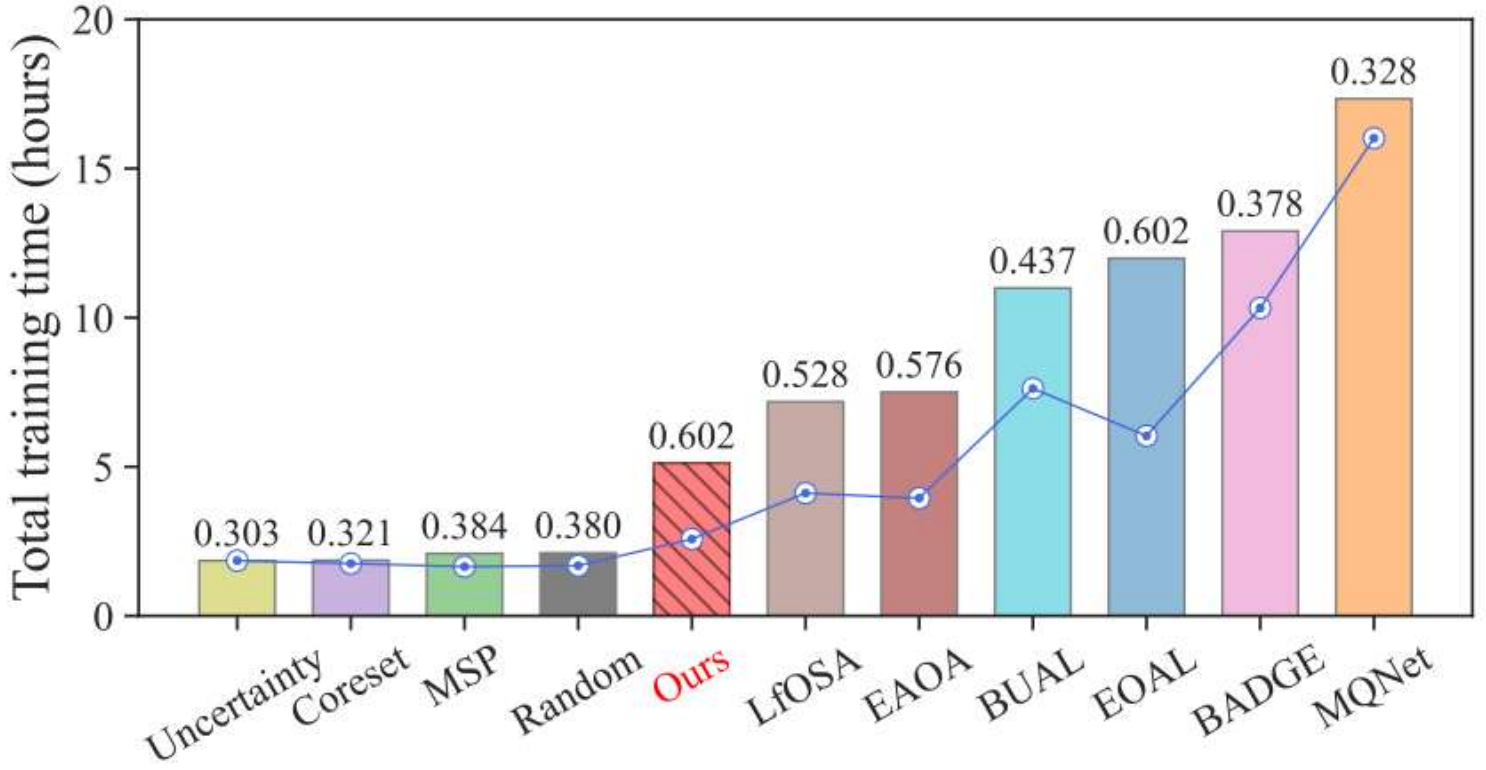


Figure 5. Total training time (hours) on CIFAR-100 under a 40% mismatch ratio. Bars indicate actual training time with the average query precision annotated; the dashed line shows the approximate projection assuming a linear relationship between query precision and time, aligned to the precision level of "Uncertainty".

THANKS