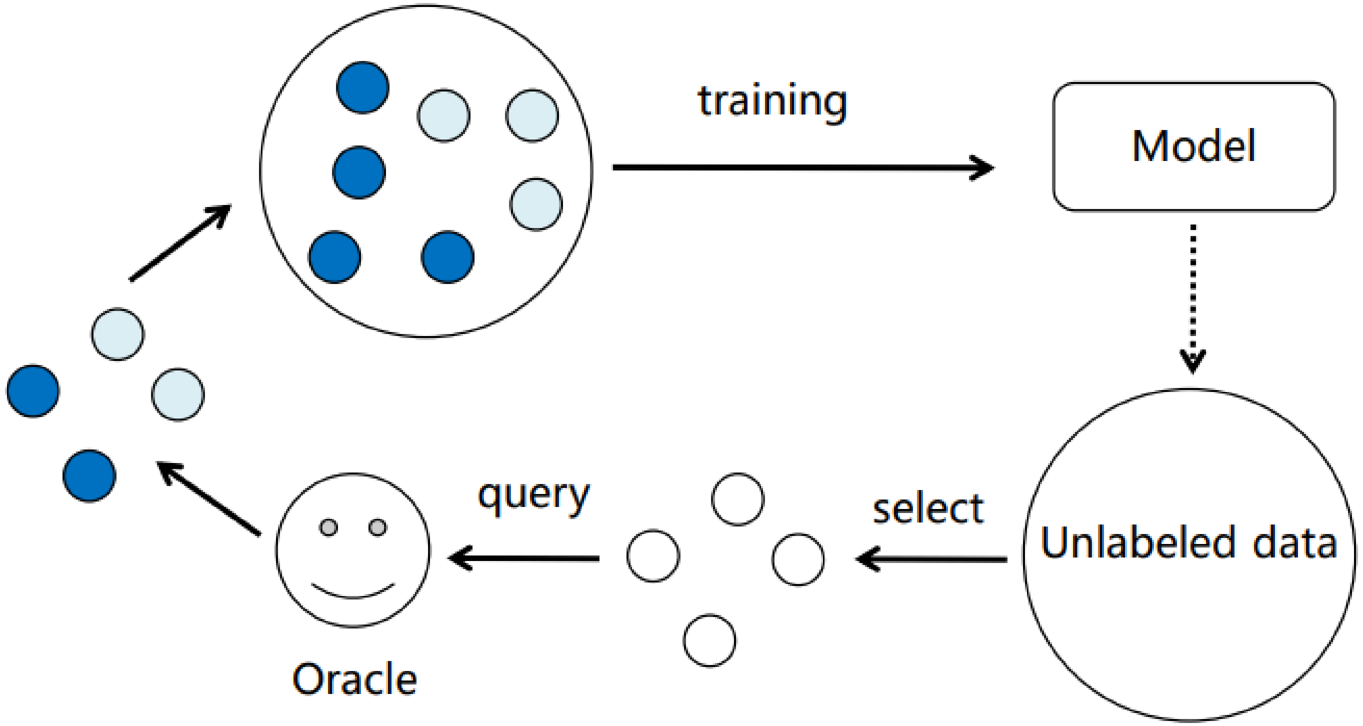
南京航空航天大學
Nanjing University of Aeronautics and Astronautics

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

NIPS 2022

Active Learning

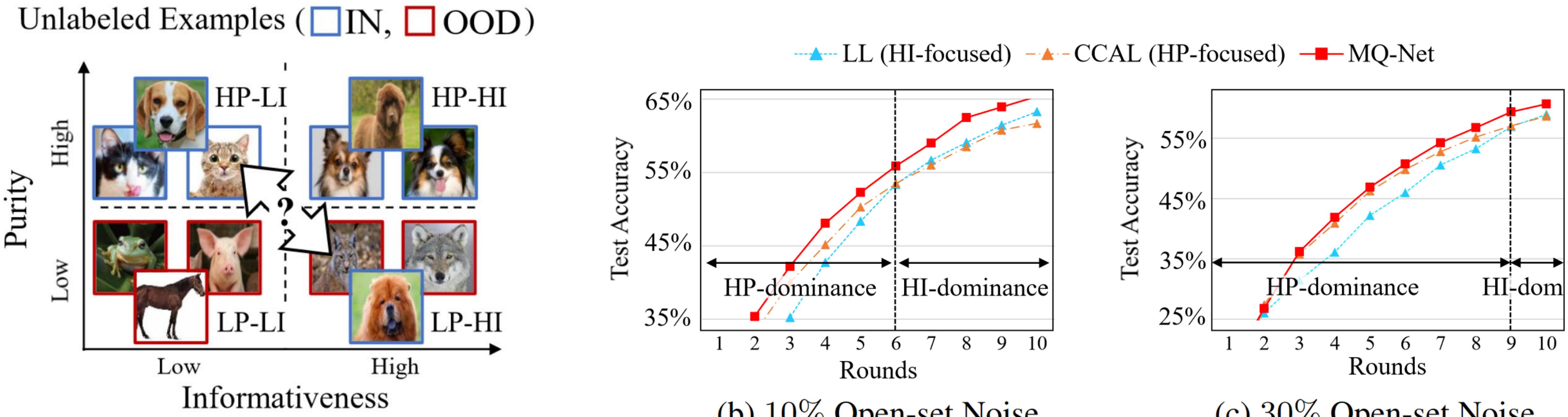


Goal: query less for more.

- Uncertainty-based sampling
  - Least-confidence
  - Margin
  - Entropy
  - MC-dropout
- Diversity-based sampling
  - CoreSet
- Hybird sampling
  - BADGE

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

Motivation



(a) Purity-informativeness Dilemma.

(b) 10% Open-set Noise.

(c) 30% Open-set Noise.

Whether purity needs to be focused on throughout the entire training period?

Meta-Query-Net

➢ The optimal query set:

$$S_Q^* = \underset{S_Q:\, C(S_Q)\le b}{\operatorname{argmin}} \; \mathbb{E}_{(x,y)\in T_{IN}} \Big[\ell_{cls}\big(f(x;\Theta_{S_L\cup S_Q}), y\big) \Big],$$

$$\text{where } \; C(S_Q) = \sum \big(\mathbb{1}_{[x\in X_{IN}]} c_{IN} + \mathbb{1}_{[x\in X_{OOD}]} c_{OOD} \big)$$

➢ MQ-Net:

- $P(x)$: 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

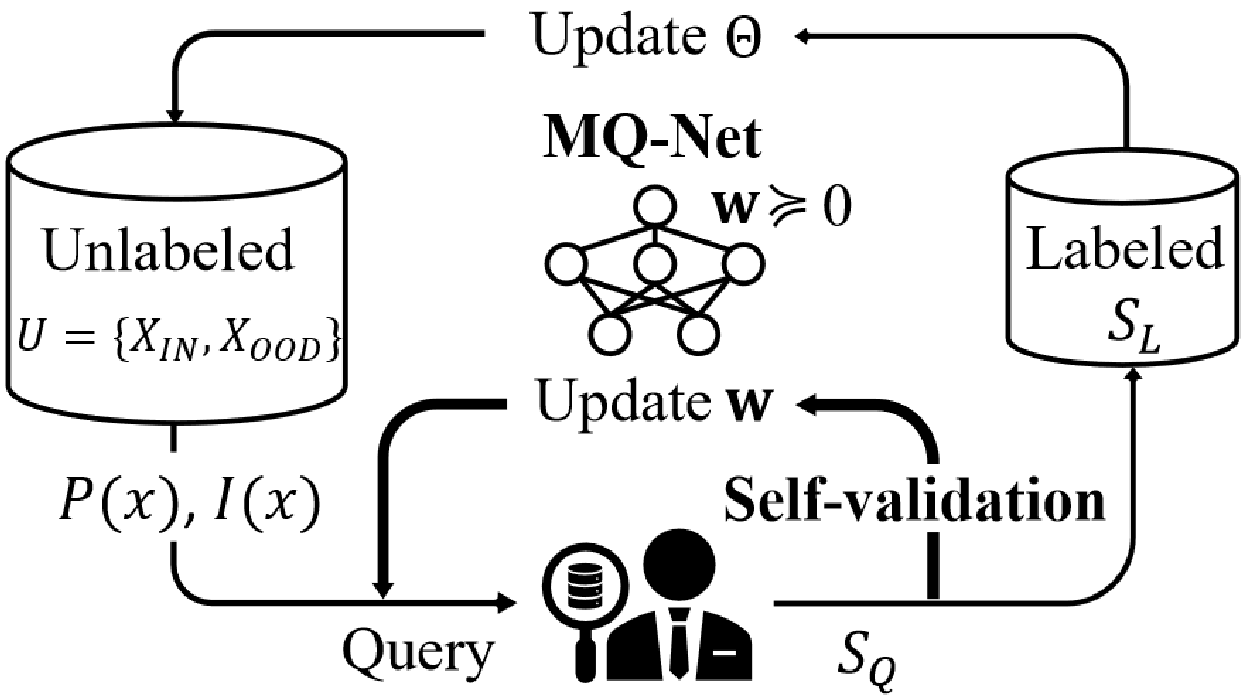


Figure 2: Overview of MQ-Net.

Meta-objective

➢ Target model objective:

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

➢ Meta objective:

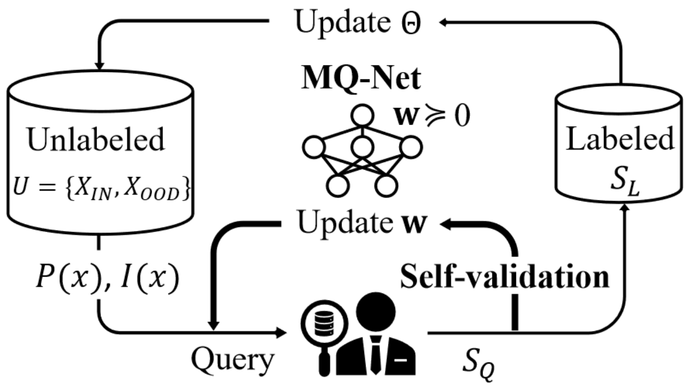


Figure 2: Overview of MQ-Net.

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

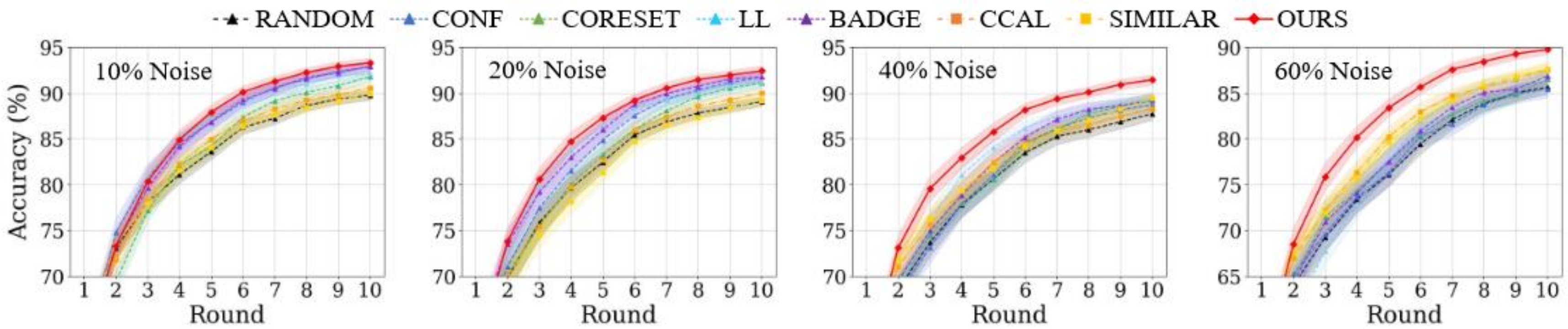
$\Phi(z_{x_i}; w)$ is forced to be higher than $\Phi(z_{x_j}; w)$ if $l_{mce}(x_i) > l_{mce}(x_j)$

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

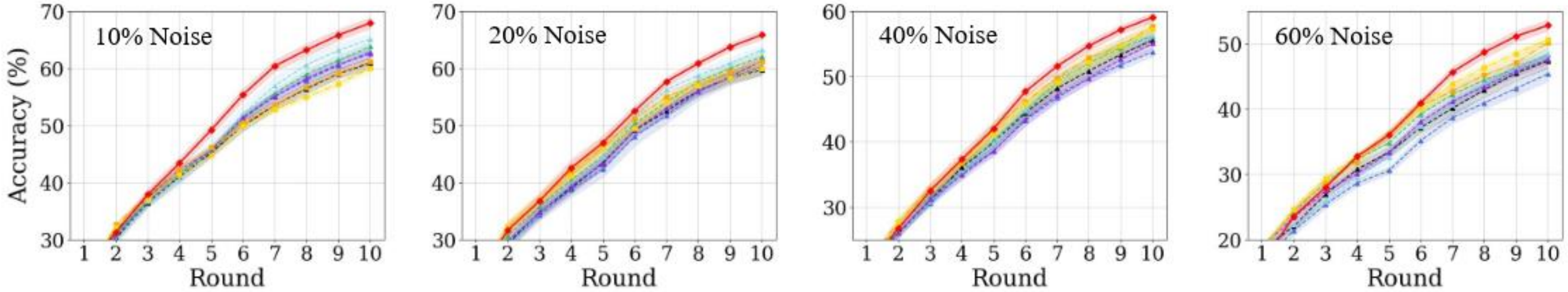
$$s.t.\ \forall x_i, x_j,\ \text{if } \mathcal{P}(x_i) > \mathcal{P}(x_j) \text{ and } \mathcal{I}(x_i) > \mathcal{I}(x_j),\ \text{then } \Phi(z_{x_i}; \mathbf{w}) > \Phi(z_{x_j}; \mathbf{w})$$

Theorem 4.1. *For any MLP meta-model* **w** *with non-decreasing activation functions, a meta-score function* $\Phi(z; \mathbf{w}) : \mathbb{R}^d \to \mathbb{R}$ *holds the skyline constraints if* $\mathbf{w} \succeq 0$ *and* $z(\in \mathbb{R}^d) \succeq 0$*, where* $\succeq$ *is the component-wise inequality.*

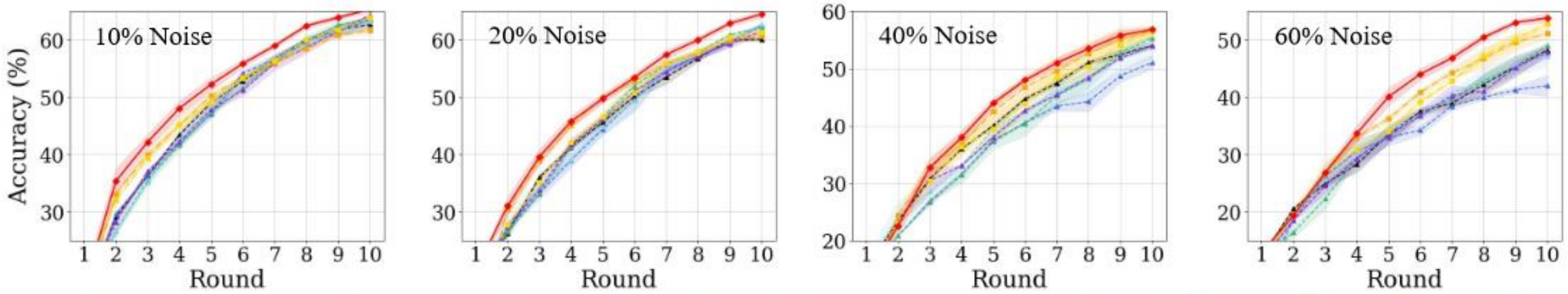
Experiment



(a) Accuracy comparison over AL rounds on CIFAR10 with open-set noise of 10%, 20%, 40%, and 60%.



(b) Accuracy comparison over AL rounds on CIFAR100 with open-set noise of 10%, 20%, 40%, and 60%.



(c) Accuracy comparison over AL rounds on ImgeNet with open-set noise of 10%, 20%, 40%, and 60%.

Experiment

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	92.83	91.72	88.69	85.43	62.84	60.20	53.74	45.38	63.56	62.56	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	89.41	86.95	65.08	64.04	56.27	48.49	63.28	61.56	55.68	47.30
	BADGE	92.80	91.73	89.27	86.83	62.54	61.28	55.07	47.60	64.84	61.48	54.04	47.80
Open-set AL	CCAL	90.55	89.99	88.87	87.49	61.20	61.16	56.70	50.20	61.68	60.70	56.60	51.16
	SIMILAR	89.92	89.19	88.53	87.38	60.07	59.89	56.13	50.61	63.92	61.40	56.48	52.84
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

Experiment

Table 2: Effect of the meta inputs on MQ-Net.

Dataset		CIFAR10 (4:6 split)			
Noise Ratio		10%	20%	40%	60%
Standard AL	BADGE	92.80	91.73	89.27	86.83
Open-set AL	CCAL	90.55	89.99	88.87	87.49
MQ-Net	CONF-ReAct	93.21	91.89	89.54	87.99
	CONF-CSI	**93.28**	**92.40**	91.43	89.37
	LL-ReAct	92.34	91.85	90.08	88.41
	LL-CSI	93.10	92.10	**91.48**	**89.51**

Table 3: Efficacy of the self-validation set.

Dataset		CIFAR10 (4:6 split)			
Noise Ratio		10%	20%	40%	60%
MQ-Net	Query set	**93.10**	**92.10**	**91.48**	**89.51**
	Random	92.10	91.75	90.88	87.65

Table 4: Efficacy of the skyline constraint.

Noise Ratio		10%	20%	40%	60%
MQ-Net	w/ skyline	**93.10**	**92.10**	**91.48**	**89.51**
	w/o skyline	87.25	86.29	83.61	81.67

Experiment

Table 5: Efficacy of the meta-objective in MQ-Net. We show the AL performance of two alternative balancing rules compared with MQ-Net for the split-dataset setup on CIFAR10 with the open-set noise ratios of 20% and 40%.

Dataset	Noise Ratio	Round	1	2	3	4	5	6	7	8	9	10
CIFAR10 (4:6 split)	20%	$\mathcal{P}(x)+\mathcal{I}(x)$	**61.93**	**73.82**	76.16	80.65	82.61	85.73	87.44	88.86	89.21	89.49
		$\mathcal{P}(x)\cdot\mathcal{I}(x)$	**61.93**	71.79	78.09	81.32	84.16	86.39	88.74	89.89	90.54	91.20
		MQ-Net	**61.93**	**73.82**	**80.58**	**84.72**	**87.31**	**89.20**	**90.52**	**91.46**	**91.93**	**92.10**
	40%	$\mathcal{P}(x)+\mathcal{I}(x)$	**59.31**	**72.50**	75.67	78.78	81.70	83.74	85.08	86.48	87.47	88.86
		$\mathcal{P}(x)\cdot\mathcal{I}(x)$	**59.31**	66.37	73.57	77.85	81.37	84.22	86.80	88.04	88.73	89.11
		MQ-Net	**59.31**	**72.50**	**79.54**	**82.94**	**85.77**	**88.16**	**89.34**	**90.07**	**90.92**	**91.48**

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

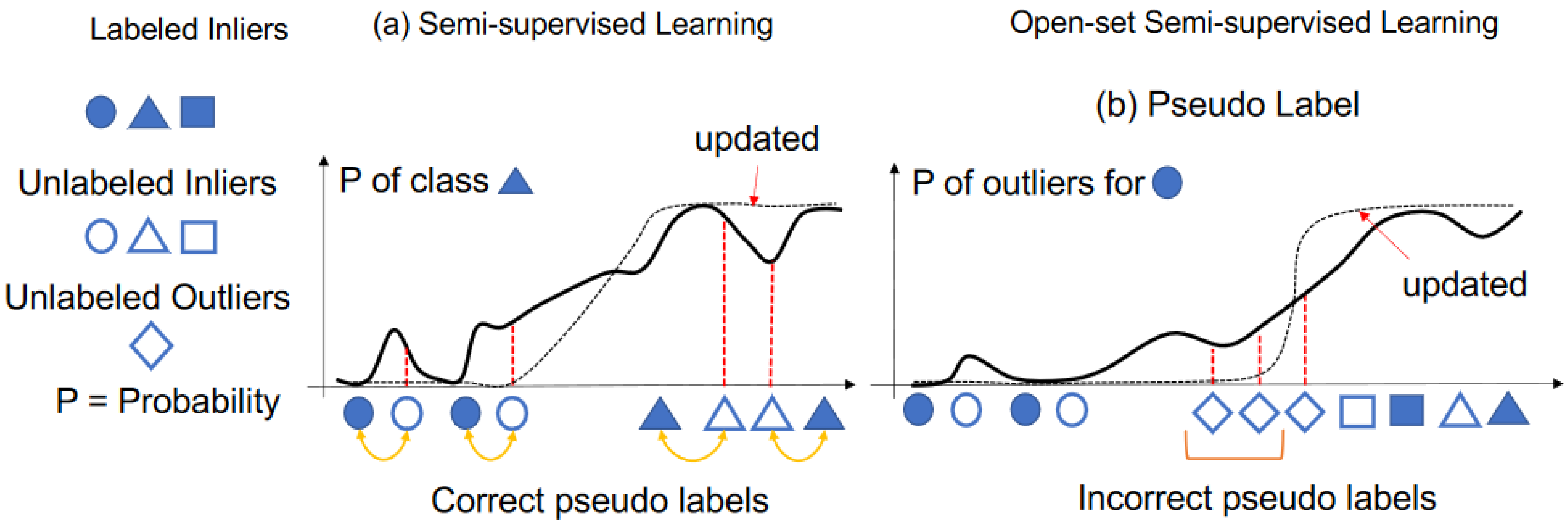
OpenMatch: Open-set Consistency Regularization for Semi-supervised Learning with Outliers

Kuniaki Saito[1] **Donghyun Kim**[1] **Kate Saenko**[1,2]
[1]Boston University [2]MIT-IBM Watson AI Lab

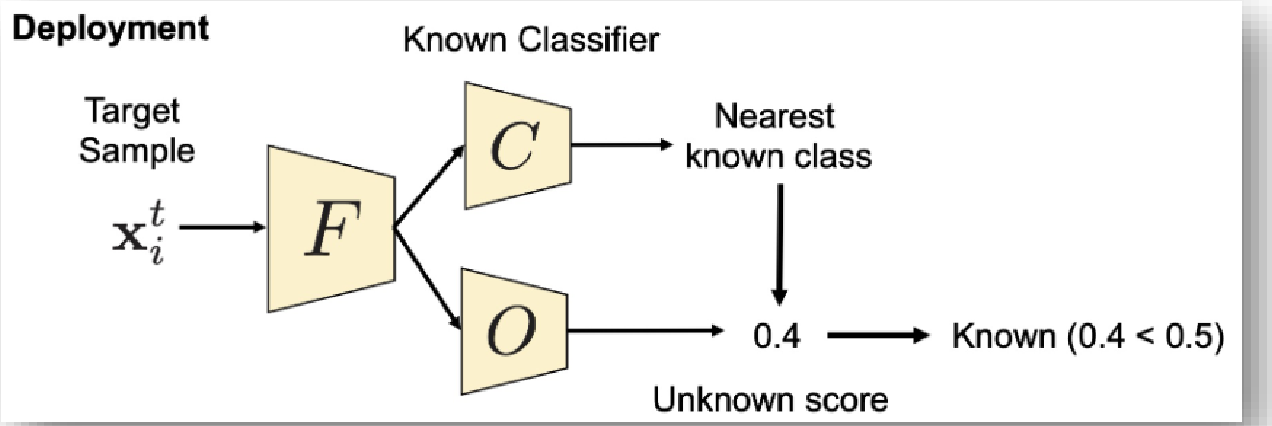
[keisaito,donhk,saenko]@bu.edu

NIPS 2021

Motivation


Labeled Inliers
(a) Semi-supervised Learning
Open-set Semi-supervised Learning
(b) Pseudo Label
updated
P of class
P of outliers for
updated
Unlabeled Inliers
Unlabeled Outliers
P = Probability
Correct pseudo labels
Incorrect pseudo labels

One-vs-All Net



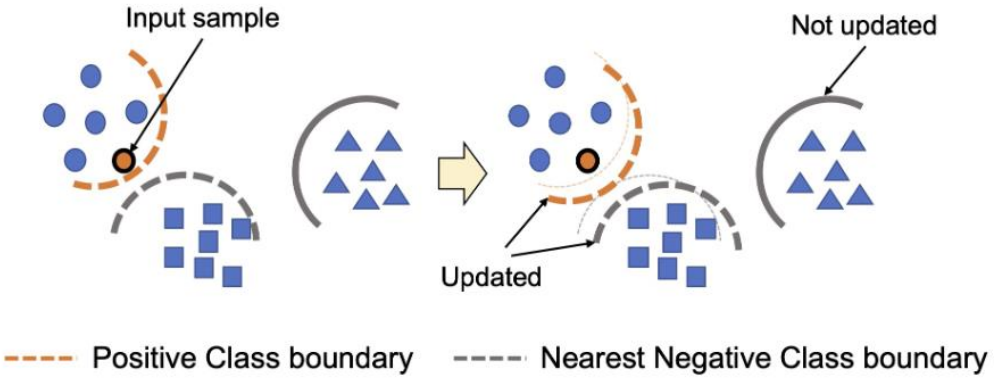
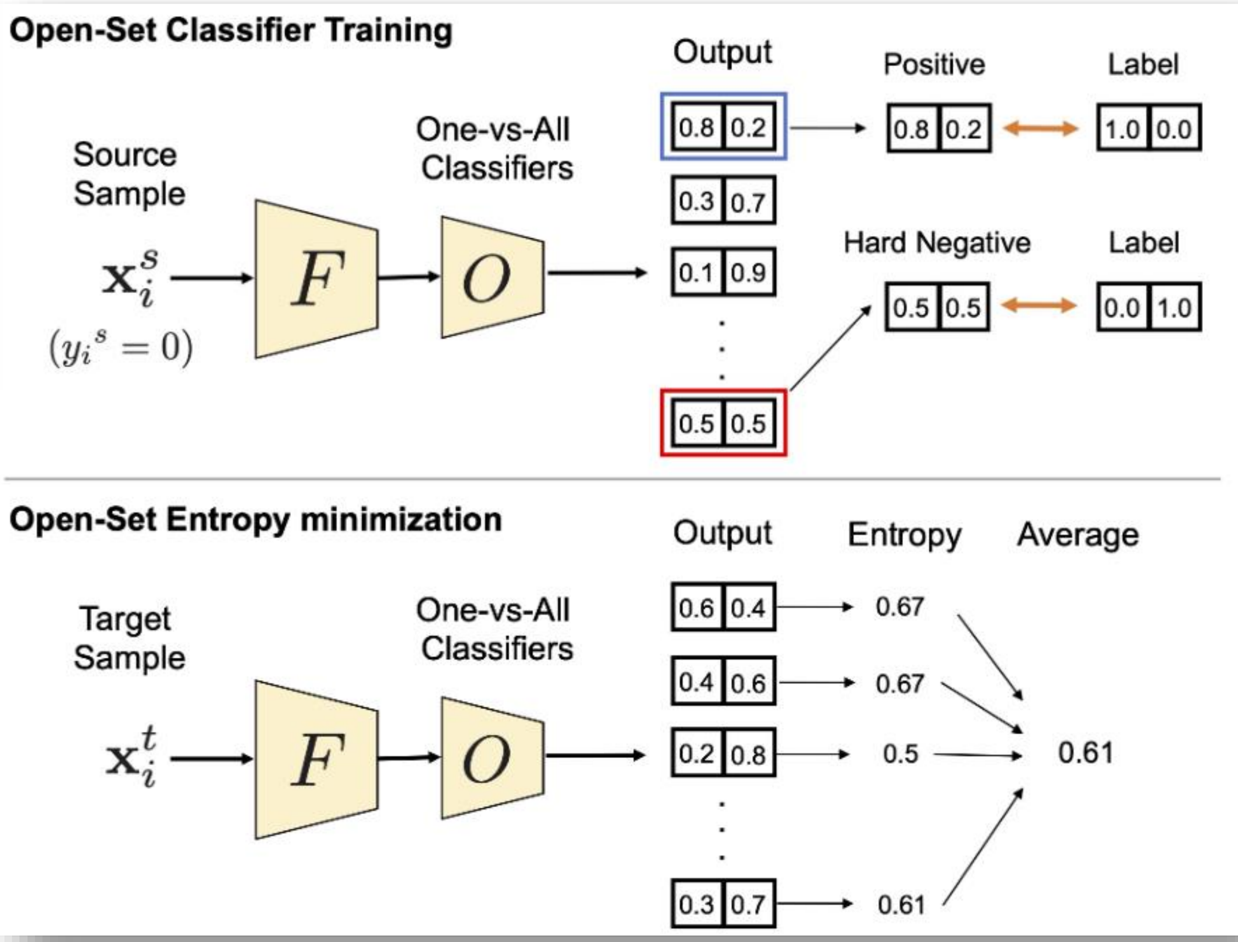


Figure 3: **Overview of the open-set classifier training.**

$$\mathcal{L}_{ova}(\mathcal{X}) := \frac{1}{B}\sum_{b=1}^{B} -\log(p^{y_b}(t=0|x_b)) - \min_{i\neq y_b}\log(p^{i}(t=1|x_b))$$



OpenMatch

Algorithm 1 OpenMatch Algorithm.

1: **Input:** Set of labeled data $\mathcal{S} = \big((x_b, y_b); b \in (1, \ldots, N)\big)$, set of unlabeled data $\mathcal{S}_u = \big(u_b; b \in (1, \ldots, N_u)\big)$, and set of pseudo-inlier data $\mathcal{K} = \emptyset$.
Data augmentation function $\mathcal{T}$. Model parameters w, learning rate η, epoch E_{fix} and $E_{\max}$, iteration $I_{\max}$, trade-off parameters, $\lambda_{em}, \lambda_{oc}, \lambda_{fm}$;

for *Epoch* = *1* *to* $E_{\max}$ **do**
  for *Iteration* = *1* *to* $I_{\max}$ **do**
    2: **Sample** a batch of labeled data $\mathcal{X} \in \mathcal{S}$ and unlabeled data $\mathcal{U} \in \mathcal{S}_u$;
    3: **Compute** $\mathcal{L}_{all} = \mathcal{L}_{sup}(\mathcal{X}) + \lambda_{em}\mathcal{L}_{em}(\mathcal{U}) + \lambda_{oc}\mathcal{L}_{oc}(\mathcal{U}, \mathcal{T})$; // Eq.1, 2 and 3
    if *Epoch* $> E_{fix}$ **then**
      4: **Sample** a batch of pseudo-inliers $\mathcal{I} \in \mathcal{K}$; // Sample pseudo-inliers.
      5: **Compute** $\mathcal{L}_{all} \mathrel{+}= \lambda_{fm}\mathcal{L}_{fm}(\mathcal{I})$; // FixMatch for pseudo-inliers.
    end
    6: **Update** $w = w - \eta\nabla_w L_{all}$; // Update weights
  end
  if *Epoch* $\geq E_{fix}$ **then**
    7: **Update** $\mathcal{K} = \text{Select}(w, \mathcal{D}_u)$; // Detect outliers and select pseudo-inliers.
  end
end

8: **Output:** Model parameters w.

Experiment

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

Dataset	CIFAR10			CIFAR100		CIFAR100		ImageNet-30
No. of Known / Unknown	6 / 4			55 / 45		80 / 20		20 / 10
No. of labeled samples	50	100	400	50	100	50	100	10 %
Labeled Only	35.7±1.1	30.5±0.7	20.0±0.3	37.0±0.8	27.3±0.5	43.6±0.5	34.7±0.4	20.9±1.0
FixMatch [35]	43.2±1.2	29.8±0.6	16.3±0.5	35.4±0.7	27.3±0.8	41.2±0.7	34.1±0.4	12.9±0.4
MTC [44]	20.3±0.9	13.7±0.9	9.0±0.5	33.5±1.2	27.9±0.5	40.1±0.8	33.6±0.3	13.6±0.7
OpenMatch	**10.4**±0.9	**7.1**±0.5	**5.9**±0.5	**27.7**±0.4	**24.1**±0.6	**33.4**±0.2	**29.5**±0.3	**10.4**±1.0

Table 1: Error rates (%) with standard deviation for CIFAR10, CIFAR100 on 3 different folds. Lower is better. For ImageNet, we use the same fold and report averaged results of three runs. Note that the number of labeled samples *per class* is shown in each column.

Dataset	CIFAR10			CIFAR100		CIFAR100		ImageNet-30
No. of Known / Unknown	6 / 4			55 / 45		80 / 20		20 / 10
No. of labeled samples	50	100	400	50	100	50	100	10 %
Labeled Only	63.9±0.5	64.7±0.5	76.8±0.4	76.6±0.9	79.9±0.9	70.3±0.5	73.9±0.9	80.3±1.0
FixMatch [35]	56.1±0.6	60.4±0.4	71.8±0.4	72.0±1.3	75.8±1.2	64.3±1.0	66.1±0.5	88.6±0.5
MTC [44]	96.6±0.6	98.2±0.3	98.9±0.1	81.2±3.4	80.7±4.6	79.4±2.5	73.2±3.5	93.8±0.8
OpenMatch	**99.3**±0.3	**99.7**±0.2	**99.3**±0.2	**87.0**±1.1	**86.5**±2.1	**86.2**±0.6	**86.8**±1.4	**96.4**±0.7

Table 2: AUROC of Table 1. Higher is better. Note that the number of labeled samples *per class* is shown in each column.

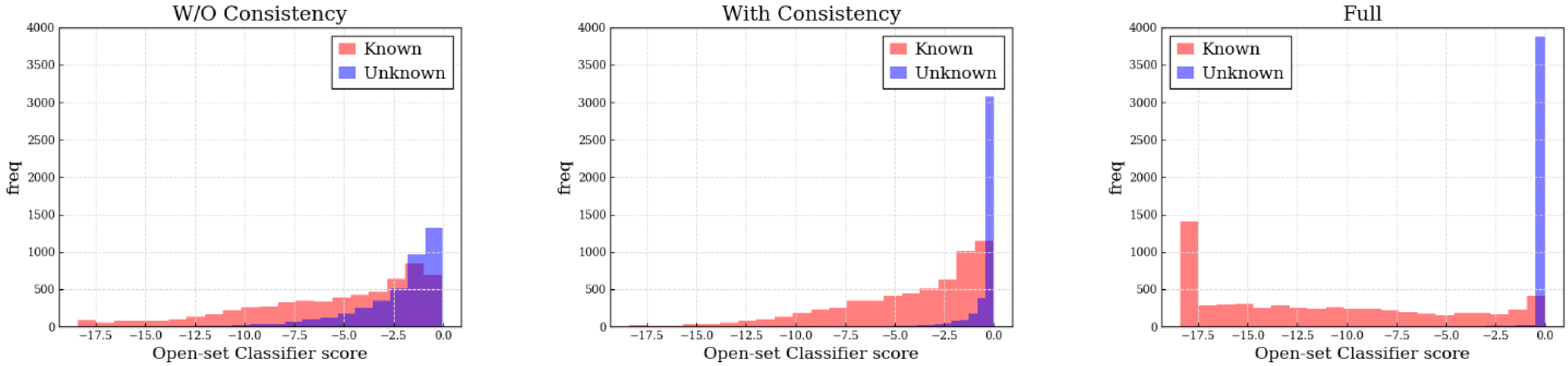
Experiment

Dataset	CIFAR10		CIFAR100		ImageNet-30
No. Known / Unknown	6 / 4		80 / 20		20 / 10
No. Labeled samples	50	400	50	100	10 %
without SOCR	60.5±2.8	75.8±0.8	70.4±0.1	73.2±0.2	81.3±0.4
with SOCR	**81.3**±2.9	**96.8**±0.6	**78.9**±0.1	**85.0**±0.8	**89.3**±0.3

Table 3: Ablation study of our soft consistency regularization (SOCR, $\mathcal{L}_{oc}$). We report AUROC scores (%). In this study, we do not apply FixMatch to pseudo-inliers to see the pure gain from SOCR.

$$\mathcal{L}_{oc}(\mathcal{U},\mathcal{T}) := \frac{1}{\mu B}\sum_{b=1}^{\mu B}\sum_{j=1}^{K}\sum_{t\in(0,1)} |p^j(t|\mathcal{T}_1(u_b)) - p^j(t|\mathcal{T}_2(u_b))|^2$$

Experiment



(a) w/o FixMatch, SOCR.

(b) w/o FixMatch

(c) FULL (OpenMatch)

Figure 3: The histograms of the outlier detector's scores obtained with ablated models. Red: Inliers, Blue: Outliers. From left to right, a model without FixMatch and SOCR, a model without FixMatch, and a model with all objectives. These results show that SOCR ensures separation between inliers and outliers, and FixMatch added to SOCR can further enhance this separation.

THANKS