模式分析与机器智能
工业和信息化部重点实验室
MIIT Key Laboratory of
Pattern Analysis & Machine Intelligence

ParNeC 模式识别与神经计算研究组
PAttern Recognition and NEural Computing

IOMatch: Simplifying Open-Set Semi-Supervised Learning with Joint Inliers and Outliers Utilization

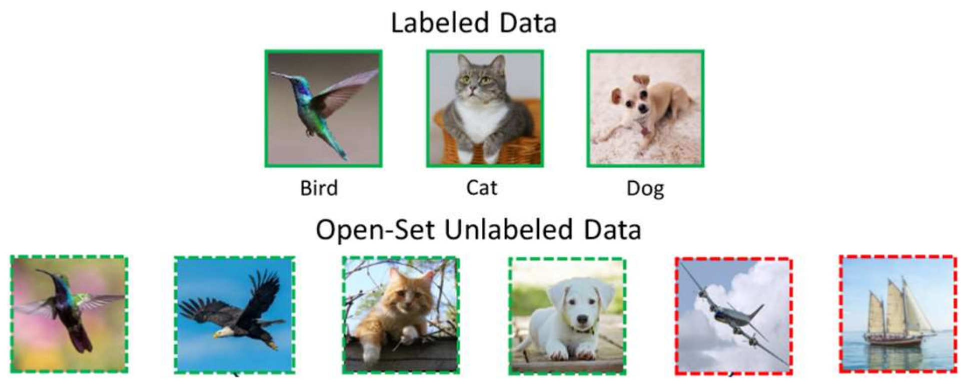
Zekun Li[1] Lei Qi[2] Yinghuan Shi[1,*] Yang Gao[1]
[1] Nanjing University [2] Southeast University

ICCV 2023

Background

Semi-Supervised Learning

Semi-supervised learning (SSL) is a classical machine learning paradigm that attempts to improve a model's performance by utilizing unlabeled data in addition to insufficient labeled data. With a tiny fraction of labeled data, advanced deep SSL methods can achieve the performance of fully supervised methods in some cases, such as image classification and semantic segmentation.

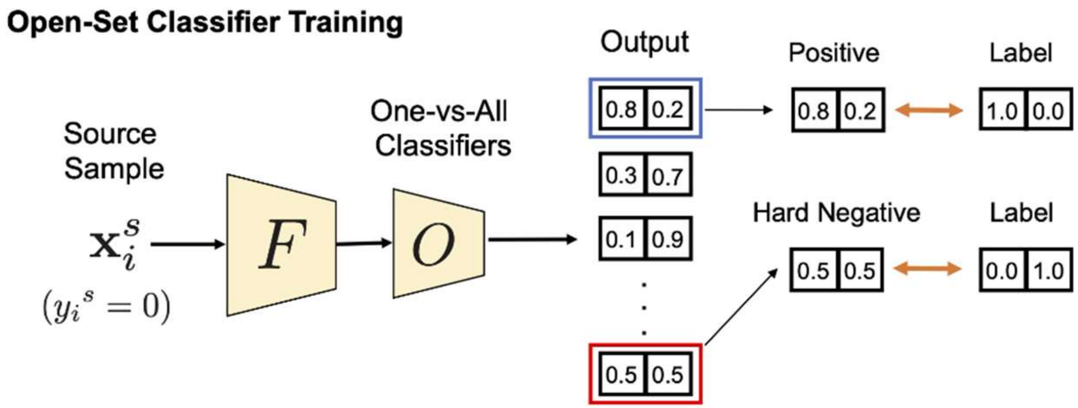


Open-Set Semi-Supervised Learning

Most existing SSL methods rely on the fundamental assumption that labeled and unlabeled data share the same class space. However, it is usually difficult, even impossible, to collect such a unlabeled data set in many real-world applications since we can not manually examine the massive unlabeled data. Therefore, a more challenging scenario arises, where unseen-class outliers not belonging to any of the labeled classes exist in the unlabeled data. Such setting is called Open-Set Semi-Supervised Learning (OSSL)

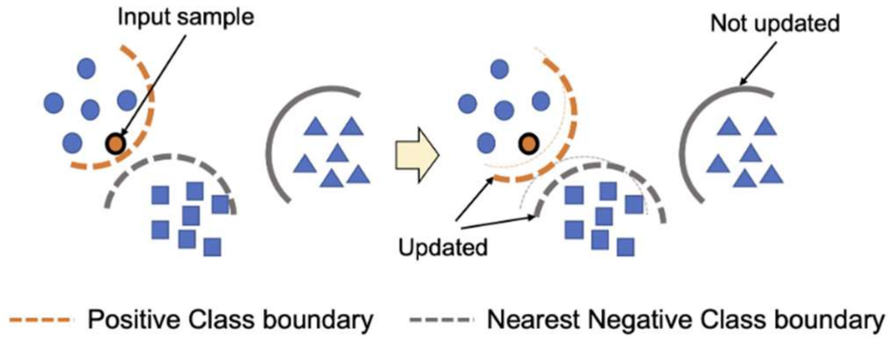
Related Work

One-vs-All Network



Hard Negative Classifier Sampling

$$\mathcal{L}_{ova}(\mathbf{x}^s, y^s) = -\log(p(\hat{y}^{y^s}|\mathbf{x}^s)) - \min_{j \neq y^s} \log(1 - p(\hat{y}^j|\mathbf{x}^s))$$



Related Work

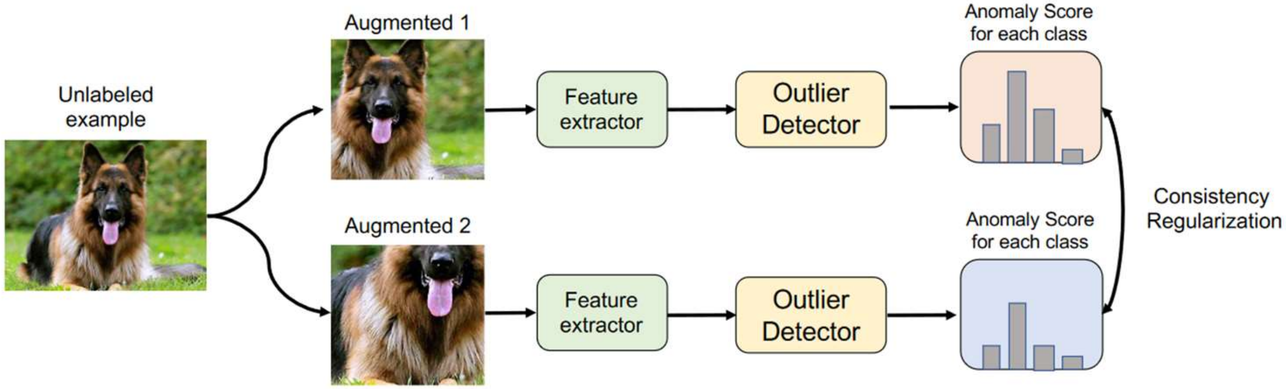
ParNeC 模式识别与神经计算研究组 PAttern Recognition and NEural Computing

OpenMatch

detect-and-filter strategy

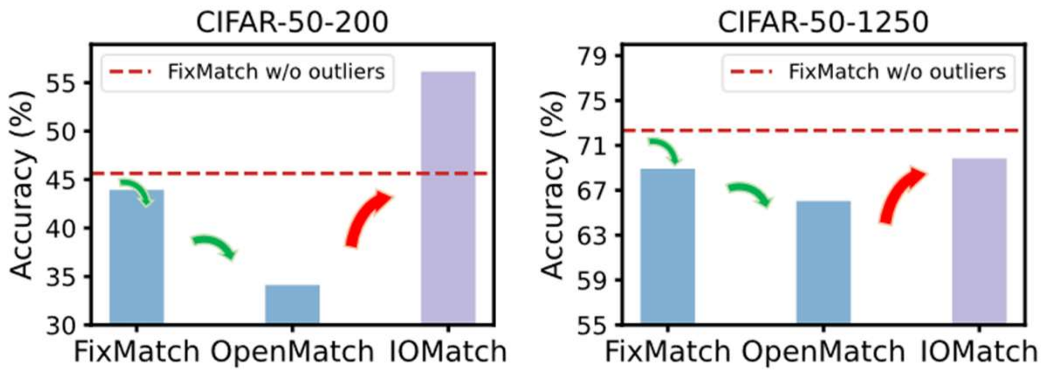
$$\mathcal{L}_{ova}(\mathcal{X}) := \frac{1}{B}\sum_{b=1}^{B} -\log(p^{y_b}(t=0|x_b)) - \min_{i \neq y_b} \log(p^{i}(t=1|x_b)).$$

identify the unlabled sample $u_b$as an outlier if $p^{\hat{y}}(t=0|u_b) < 0.5$, where $\hat{y} = \arg\max_j C(F(u_b))$



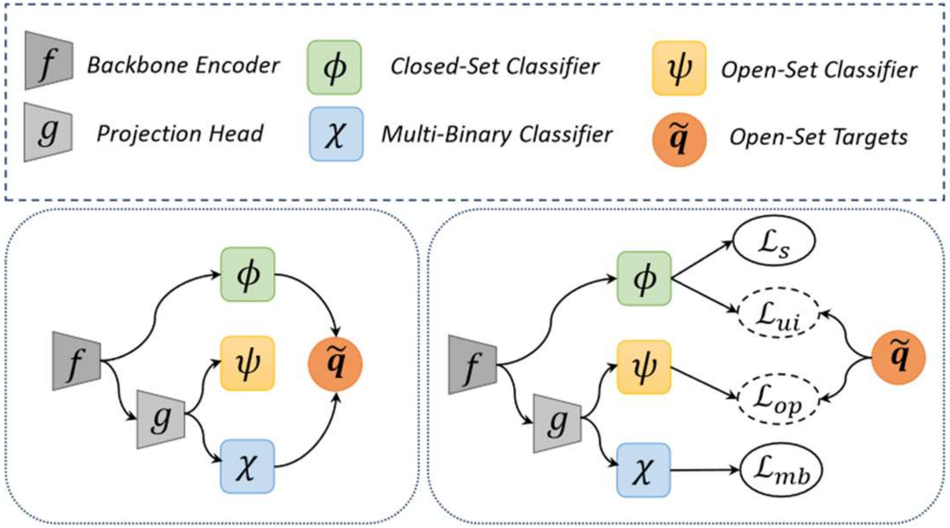
Motivation

ParNeC 模式识别与神经计算研究组 PAttern Recognition and NEural Computing



The motivation of our work comes from a surprising fact in open-set semi-supervised learning tasks: An unreliable outlier detector can be more harmful than outliers themselves, because it will wrongly exclude valuable inliers from subsequent training. For this issue, we consider a unified paradigm for utilizing openset unlabeled data, even when it is difficult to distinguish exactly between inliers and outliers, and thus we propose IOMatch.

Method



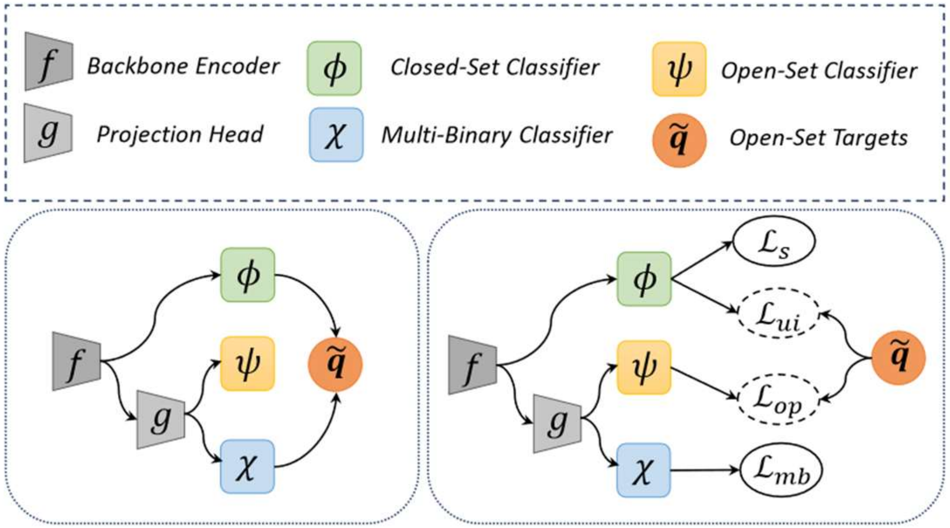
(a) Produce Open-Set Targets (b) Optimize Whole Framework

$$\chi = \{\chi_k : k \in (1, \ldots, K)\}$$

$$\boldsymbol{o}_{i,k} = \chi_k(\boldsymbol{z}_i) \in \mathbb{R}^2$$

where $\boldsymbol{o}_{i,k} = (o_{i,k}, \bar{o}_{i,k})$ and $o_{i,k} + \bar{o}_{i,k} = 1$

Method



(a) Produce Open-Set Targets (b) Optimize Whole Framework

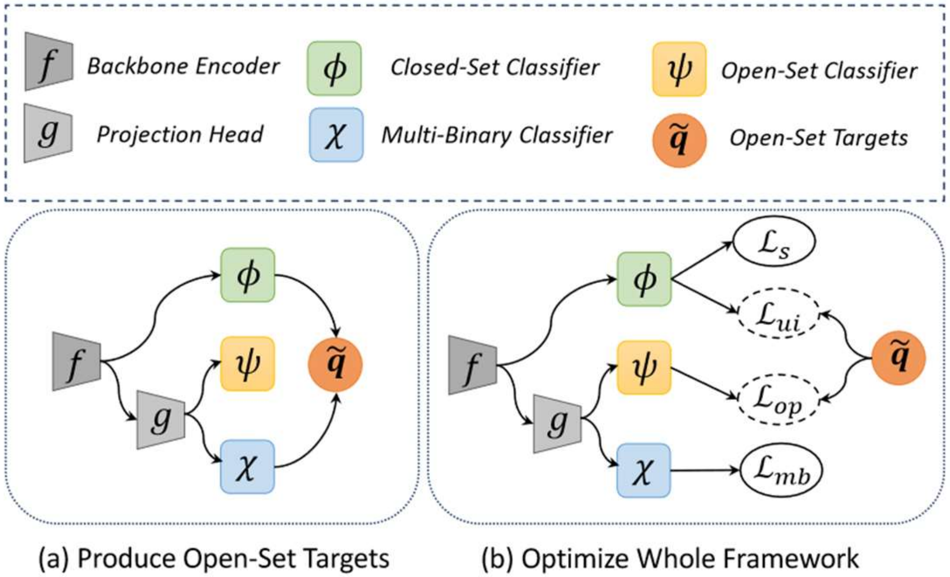
$$\widetilde{\boldsymbol{p}}_i = \mathrm{DA}(\phi(\boldsymbol{h}_i^w))$$

$$\text{for } 1 \le k \le K \qquad \widetilde{q}_{i,k} = \widetilde{p}_{i,k} \cdot o^w_{i,k}$$

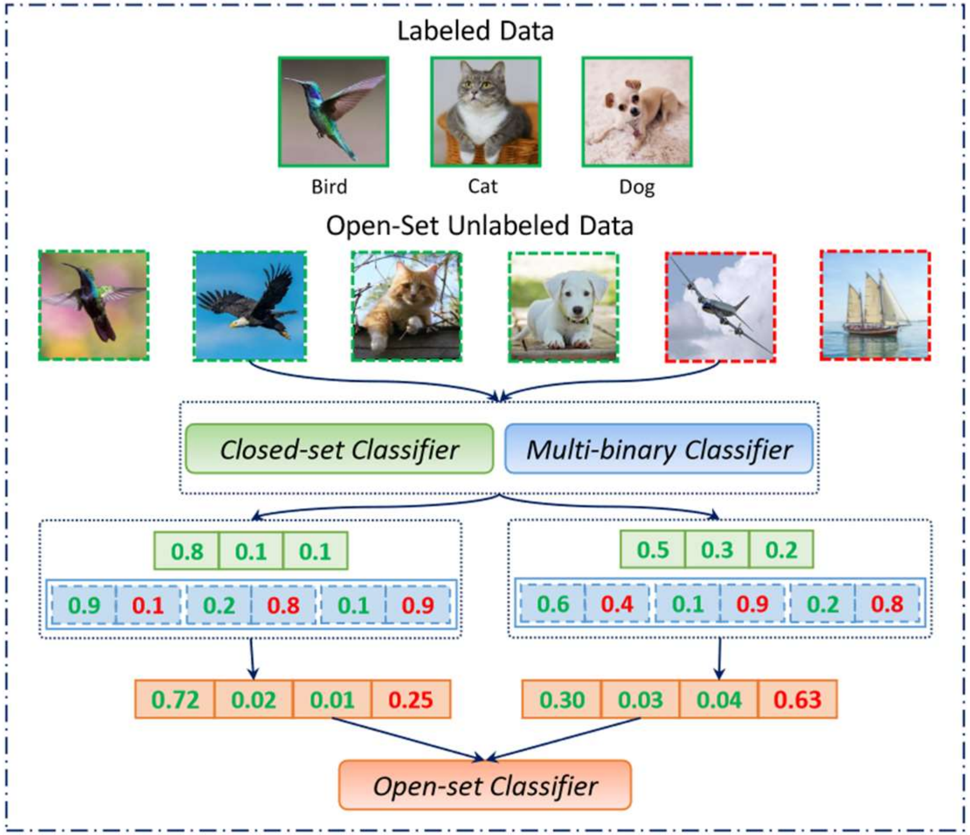
$$\mathcal{S}_i = 1 - \sum_{j=1}^{K} \widetilde{q}_{i,j} = \sum_{j=1}^{K} \widetilde{p}_{i,j} \cdot \bar{o}^w_{i,j}.$$

$$\widetilde{q}_{i,k} = \begin{cases} \widetilde{p}_{i,k} \cdot o^w_{i,k} & \text{if} \quad 1 \le k \le K; \\ \sum_{j=1}^{K} \widetilde{p}_{i,j} \cdot \bar{o}^w_{i,j} & \text{if} \quad k = K+1. \end{cases}$$

Method

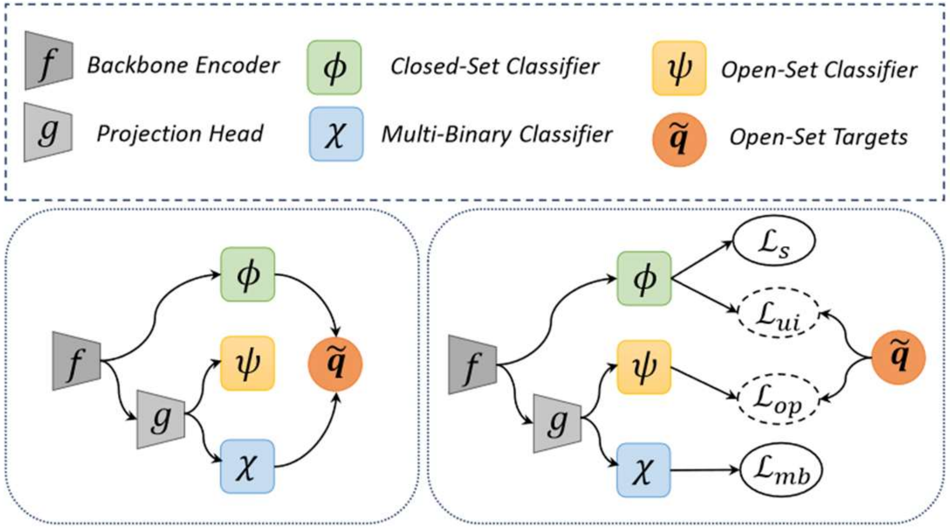


Backbone Encoder
Closed-Set Classifier
Open-Set Classifier
Projection Head
Multi-Binary Classifier
Open-Set Targets
(a) Produce Open-Set Targets
(b) Optimize Whole Framework



Labeled Data
Bird
Cat
Dog
Open-Set Unlabeled Data
Closed-set Classifier
Multi-binary Classifier
0.8 0.1 0.1
0.9 0.1 0.2 0.8 0.1 0.9
0.5 0.3 0.2
0.6 0.4 0.1 0.9 0.2 0.8
0.72 0.02 0.01 0.25
0.30 0.03 0.04 0.63
Open-set Classifier

Method

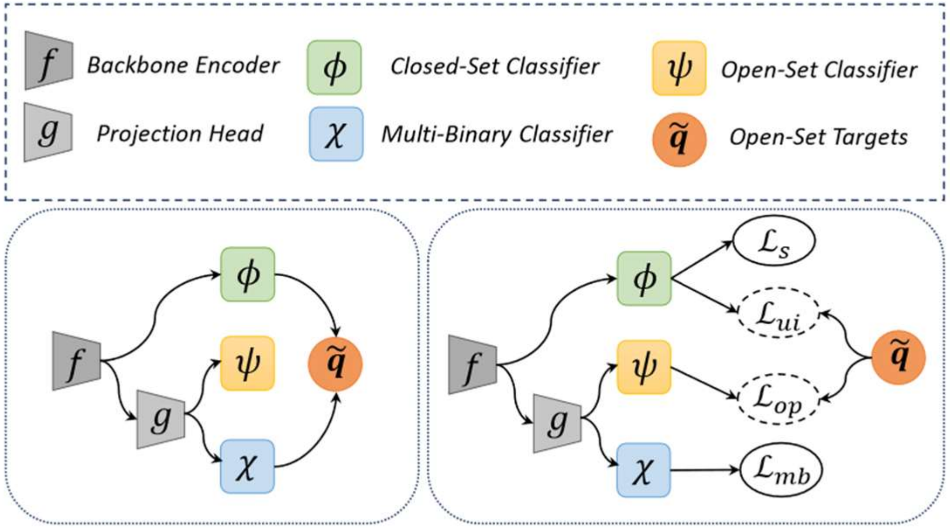


(a) Produce Open-Set Targets (b) Optimize Whole Framework

hard negative classifier sampling

$$\mathcal{L}_{mb}(\mathcal{X}) = \frac{1}{B}\sum_{i=1}^{B}\Big(-\log(o_{i,y_i}) - \min_{k\neq y_i}\log(\bar{o}_{i,k})\Big)$$

Method



(a) Produce Open-Set Targets (b) Optimize Whole Framework

$$\mathcal{L}_{op}(\mathcal{U}) = \frac{1}{\mu B} \sum_{i=1}^{\mu B} \mathbb{1}(\max_{k}(\widetilde{q}_{i,k}) > \tau_q) \cdot \mathrm{H}(\widetilde{\boldsymbol{q}}_i, \boldsymbol{q}_i^s),$$

$$\mathcal{L}_{ui}(\mathcal{U}) = \frac{1}{\mu B} \sum_{i=1}^{\mu B} \mathcal{F}(\boldsymbol{u}_i) \cdot \mathrm{H}(\widetilde{\boldsymbol{p}}_i, \boldsymbol{p}_i^s).$$

$$\mathcal{F}(\boldsymbol{u}_i) = \mathbb{1}(\max_k(\widetilde{p}_{i,k}) > \tau_p) \cdot \mathbb{1}(\mathcal{S}_i < 0.5).$$

$$\mathcal{L}_{overall} = \mathcal{L}_s + \lambda_{mb}\mathcal{L}_{mb} + \lambda_{ui}\mathcal{L}_{ui} + \lambda_{op}\mathcal{L}_{op}.$$

Experiment

Closed-set classification accuracy (%) on the seen-class test data of CIFAR-10/100 with varying seen/unseen class splits and labeled set sizes.

Dataset			CIFAR-10		CIFAR-100					
Class split (Seen / Unseen)			6 / 4		20 / 80		50 / 50		80 / 20	
Number of labels per class			4	25	4	25	4	25	4	25
Standard SSL	MixMatch [3]	NeurIPS'19	43.08 ± 1.79	63.13 ± 0.64	28.13 ± 5.06	51.28 ± 1.45	26.97 ± 0.46	56.93 ± 0.84	28.35 ± 0.83	53.77 ± 0.97
	ReMixMatch [2]	ICLR'20	72.82 ± 1.81	87.08 ± 1.12	36.02 ± 3.56	61.83 ± 0.81	37.57 ± 1.54	65.80 ± 1.33	40.64 ± 2.97	62.90 ± 1.07
	FixMatch [28]	NeurIPS'20	81.58 ± 6.63	92.94 ± 0.80	46.27 ± 0.64	66.45 ± 0.74	48.93 ± 5.05	68.77 ± 0.89	43.06 ± 1.21	64.44 ± 0.51
	CoMatch [20]	ICCV'21	86.08 ± 1.08	92.57 ± 0.47	43.53 ± 3.01	66.82 ± 1.37	43.17 ± 0.55	67.85 ± 1.17	37.89 ± 1.22	62.04 ± 0.08
	FlexMatch [41]	NeurIPS'21	73.34 ± 4.42	86.44 ± 3.72	37.93 ± 4.49	62.68 ± 2.02	44.10 ± 1.88	68.98 ± 0.94	43.44 ± 2.40	64.34 ± 0.64
	SimMatch [43]	CVPR'22	79.84 ± 4.76	90.07 ± 2.44	36.93 ± 5.72	67.23 ± 1.13	51.53 ± 2.02	69.71 ± 1.44	50.32 ± 2.57	**65.68 ± 1.43**
	FreeMatch [34]	ICLR'23	79.26 ± 4.11	92.27 ± 0.15	45.18 ± 8.36	64.62 ± 0.79	50.26 ± 1.92	68.57 ± 0.27	47.34 ± 0.57	64.41 ± 0.55
Open-Set SSL	UASD [7]	AAAI'20	35.25 ± 1.07	56.42 ± 1.34	29.78 ± 4.28	53.78 ± 0.67	29.08 ± 1.44	54.24 ± 1.10	26.41 ± 2.16	50.33 ± 0.62
	DS^{3}L [10]	ICML'20	39.09 ± 1.24	51.83 ± 1.06	19.70 ± 1.98	41.78 ± 1.45	21.62 ± 0.54	47.41 ± 0.61	20.10 ± 0.48	40.51 ± 1.02
	MTCF [39]	ECCV'20	49.15 ± 6.12	74.42 ± 2.95	32.58 ± 3.36	55.93 ± 1.66	35.35 ± 2.39	57.72 ± 0.20	25.40 ± 1.20	54.59 ± 0.49
	T2T [16]	ICCV'21	73.89 ± 1.55	85.69 ± 1.90	44.23 ± 2.27	65.60 ± 0.71	39.31 ± 1.16	68.59 ± 0.92	38.16 ± 0.59	63.86 ± 0.32
	OpenMatch [25]	NeurIPS'21	43.63 ± 3.26	66.27 ± 1.86	37.45 ± 2.67	62.70 ± 1.76	33.74 ± 0.38	66.53 ± 0.54	28.54 ± 1.15	61.23 ± 0.81
	SAFE-STUDENT [14]	CVPR'22	59.28 ± 1.18	77.87 ± 0.14	34.53 ± 0.67	58.07 ± 1.40	35.84 ± 0.86	62.75 ± 0.38	34.17 ± 0.69	57.99 ± 0.34
	IOMatch	**Ours**	**89.68 ± 2.04**	**93.87 ± 0.16**	**53.73 ± 2.12**	**67.28 ± 1.10**	**56.31 ± 2.29**	**69.77 ± 0.58**	**50.83 ± 0.99**	64.75 ± 0.52

Experiment

Open-set classification balanced accuracy (%) on the open-set test data of CIFAR-10/100, which consist of samples from all the seen and unseen classes.

	Dataset		CIFAR-10		CIFAR-100					
	Class split (Seen / Unseen)		6 / 4		20 / 80		50 / 50		80 / 20	
	Number of labels per class		4	25	4	25	4	25	4	25
Open-Set SSL	UASD [7]	AAAI'20	17.10 ± 0.32	36.01 ± 0.22	10.50 ± 0.83	26.96 ± 0.53	6.92 ± 0.55	32.23 ± 0.54	5.77 ± 0.21	27.61 ± 1.15
Open-Set SSL	DS3L [10]	ICML'20	30.89 ± 0.33	40.45 ± 0.77	12.56 ± 1.21	34.35 ± 0.41	12.14 ± 0.39	35.17 ± 0.48	11.10 ± 1.27	29.09 ± 0.31
Open-Set SSL	MTCF [39]	ECCV'20	33.35 ± 7.21	46.13 ± 0.54	8.12 ± 2.10	26.60 ± 3.66	4.13 ± 0.37	38.36 ± 0.29	1.46 ± 0.17	30.75 ± 0.52
Open-Set SSL	T2T [16]	ICCV'21	50.57 ± 0.38	61.10 ± 0.39	17.17 ± 1.37	37.18 ± 0.60	12.74 ± 2.66	44.24 ± 0.42	34.23 ± 0.57	51.41 ± 0.96
Open-Set SSL	OpenMatch [25]	NeurIPS'21	14.37 ± 0.05	20.35 ± 3.50	8.77 ± 2.84	39.89 ± 1.16	7.00 ± 0.02	49.75 ± 1.08	6.30 ± 0.87	44.83 ± 0.62
Open-Set SSL	SAFE-STUDENT [14]	CVPR'22	45.27 ± 0.36	52.78 ± 0.64	15.94 ± 1.07	28.83 ± 0.46	23.98 ± 0.88	46.71 ± 1.74	29.43 ± 0.66	50.48 ± 0.61
	IOMatch	**Ours**	**75.08 ± 1.92**	**78.96 ± 0.08**	**45.94 ± 1.70**	**58.52 ± 0.48**	**46.36 ± 1.93**	**60.78 ± 0.71**	**39.96 ± 0.95**	**54.39 ± 0.38**

Experiment

ParNeC 模式识别与神经计算研究组 PAttern Recognition and NEural Computing

Evaluation	Closed-Set		Open-Set	
Labeled ratio	1%	5%	1%	5%
FixMatch	52.52 ± 3.82	78.55 ± 1.46	–	–
CoMatch	62.92 ± 0.90	79.17 ± 0.42	–	–
SimMatch	64.15 ± 0.94	80.23 ± 0.53	–	–
T2T	63.70 ± 0.83	78.87 ± 0.49	48.81 ± 0.88	58.51 ± 0.41
OpenMatch	56.35 ± 3.35	73.90 ± 1.05	21.80 ± 1.90	57.25 ± 0.76
SAFE-STUDENT	58.38 ± 2.34	75.85 ± 0.99	44.08 ± 2.09	55.25 ± 1.46
IOMatch	**69.18** ± 1.68	**81.43** ± 0.78	**57.71** ± 2.69	**73.94** ± 0.99

Table 3. Close-set and open-set accuracy (%) on ImageNet-30 with the class split of 20/10. We report the mean with standard deviation over 3 runs of different random seeds.

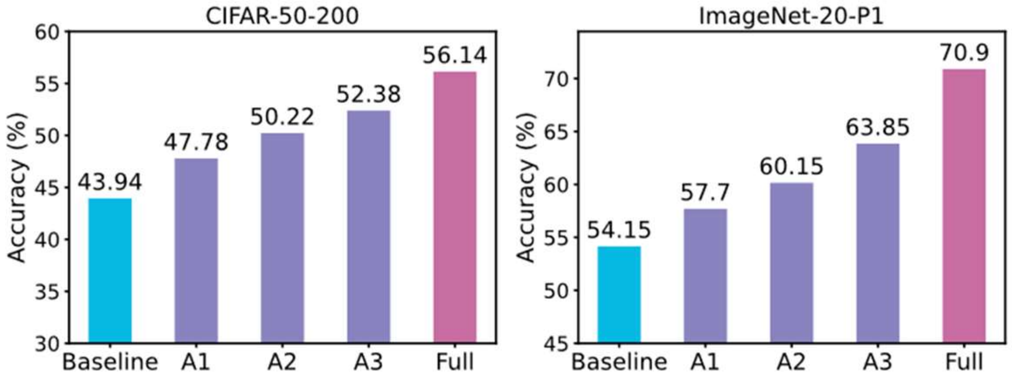


Figure 5. Ablation results on different combinations of learning objectives. "A1", "A2", and "A3" stand for the frameworks optimized with $\{\mathcal{L}_s, \mathcal{L}_{ui}\}$, $\{\mathcal{L}_s, \mathcal{L}_{mb}, \mathcal{L}_{ui}\}$, $\{\mathcal{L}_s, \mathcal{L}_{mb}, \mathcal{L}_{op}\}$, respectively. We compare the performance with FixMatch ("Baseline") and the full version of IOMatch ("Full").

Experiment

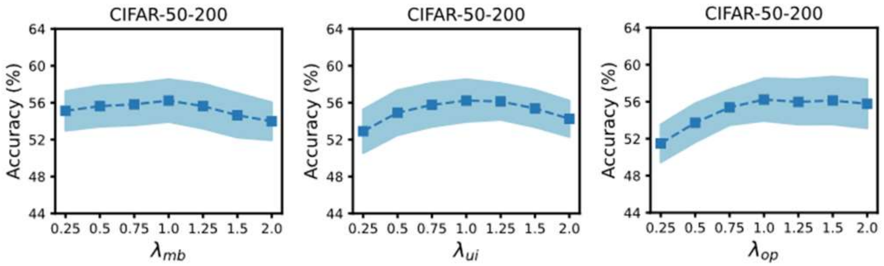


Figure 6. Performance with different values of each weight (*i.e.*, λ_{mb}, λ_{ui} and λ_{op}). It is shown that setting all the weights to 1 is a simple yet appropriate choice.

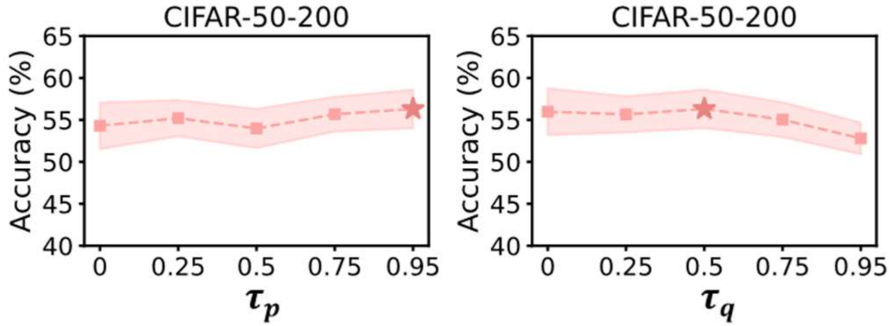


Figure 7. We vary the confidence thresholds, τ_p and τ_q, respectively. The set $\{\tau_p = 0.95, \tau_q = 0.5\}$ gives the best performance.

Experiment

Table 4. Closed-set classification accuracy (%) of several methods in the standard SSL setting (presented in the column of "SSL") compared to the performance in the OSSL setting.

Task	CIFAR-50-200		CIFAR-50-1250	
Setting	OSSL	SSL	OSSL	SSL
FixMatch	43.94	45.64	68.92	72.74
SimMatch	49.98	51.76	69.70	**73.66**
OpenMatch	37.60	39.16	66.54	67.80
IOMatch	**56.14**	**55.94**	**69.84**	73.28

Table 5. Closed-set classification accuracy (%) of IOMatch extended with auxiliary self-supervised learning objectives.

Dataset	CIFAR100			
Class split	50 / 50		80 / 20	
Number of labels	4	25	4	25
IOMatch	56.14	69.84	49.89	64.28
w/ Contrastive	57.08	70.80	50.25	65.92
w/ Rotation	**58.92**	**71.54**	**50.90**	**66.50**

Thanks