中南大學
CENTRAL SOUTH UNIVERSITY

Unified Entropy Optimization for Open-Set Test-Time Adaptation

Accepted by **CVPR 2024**

Authors

Zheng-Qing Gao

Master Student

Pattern Analysis and Learning Group

National Laboratory of Pattern Recognition (NLRR)

Institute of Automation,Chinese Academy of Sciences

Email: gaozhengqing2021@ia.ac.cn

Authors

Xu-Yao Zhang

关注

Institute of Automation, Chinese Academy of Sciences
在 nlpr.ia.ac.cn 的电子邮件经过验证
Pattern Recognition Machine Learning OCR

标题	引用次数	年份
Deep direct regression for multi-oriented scene text detection W He, XY Zhang, F Yin, CL Liu International Conference on Computer Vision (ICCV)	454	2017
Robust classification with convolutional prototype learning HM Yang, XY Zhang, F Yin, CL Liu IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3474-3482	435	2018
ICDAR 2013 Chinese handwriting recognition competition F Yin, QF Wang, XY Zhang, CL Liu International Conference on Document Analysis and Recognition (ICDAR)	412	2013
Prototype augmentation and self-supervision for incremental learning F Zhu, XY Zhang, C Wang, F Yin, CL Liu IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 5871-5880	406	2021
Drawing and recognizing Chinese characters with recurrent neural network XY Zhang, F Yin, YM Zhang, CL Liu, Y Bengio IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI) 40 (4), 849-862	405	2018
Convolutional prototype network for open set recognition HM Yang, XY Zhang, F Yin, Q Yang, CL Liu IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)	216	2022
Class-incremental learning via dual augmentation F Zhu, Z Cheng, XY Zhang, C Liu Advances in Neural Information Processing Systems 34, 14306-14318	186	2021
Hybrid CNN and dictionary-based models for scene recognition and domain adaptation GS Xie, XY Zhang, S Yan, CL Liu IEEE Trans. Circuits and Systems for Video Technology (CSVT)	173	2016
Towards robust pattern recognition: A review XY Zhang, CL Liu, CY Suen Proceedings of the IEEE 108 (6), 894-922	148	2020

Pilora: Prototype guided incremental lora for federated class-incremental learning H Guo, F Zhu, W Liu, XY Zhang, CL Liu European Conference on Computer Vision, 141-159	2	2025
Towards trustworthy dataset distillation S Ma, F Zhu, Z Cheng, XY Zhang Pattern Recognition 157, 110875	10	2025
DESIRE: Dynamic Knowledge Consolidation for Rehearsal-Free Continual Learning H Guo, F Zhu, F Zeng, B Liu, XY Zhang arXiv preprint arXiv:2411.19154		2024
Happy: A Debiased Learning Framework for Continual Generalized Category Discovery S Ma, F Zhu, Z Zhong, W Liu, XY Zhang, CL Liu arXiv preprint arXiv:2410.06535		2024
Modalprompt: Dual-modality guided prompt for continual learning of large multimodal models F Zeng, F Zhu, H Guo, XY Zhang, CL Liu arXiv preprint arXiv:2410.05849	1	2024
Breaking the limits of reliable prediction via generated data Z Cheng, F Zhu, XY Zhang, CL Liu International Journal of Computer Vision, 1-27	1	2024
Enhancing Outlier Knowledge for Few-Shot Out-of-Distribution Detection with Extensible Local Prompts F Zeng, Z Cheng, F Zhu, XY Zhang arXiv preprint arXiv:2409.04796	1	2024
PASS++: A Dual Bias Reduction Framework for Non-Exemplar Class-Incremental Learning F Zhu, XY Zhang, Z Cheng, CL Liu arXiv preprint arXiv:2407.14029		2024
Open-world machine learning: A review and new outlooks F Zhu, S Ma, Z Cheng, XY Zhang, Z Zhang, CL Liu arXiv preprint arXiv:2403.01759	16	2024
Federated Class-Incremental Learning with Prototype Guided Transformer H Guo, F Zhu, W Liu, XY Zhang, CL Liu arXiv preprint arXiv:2401.02094	9	2024
RCL: Reliable Continual Learning for Unified Failure Detection F Zhu, Z Cheng, XY Zhang, CL Liu, Z Zhang Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern …	1	2024

02

Introduction

深度学习模型在训练和测试数据分布一致时表现优异，但现实中测试数据与训练数据分布不一致，会产生**分布偏移**。

协变量偏移：输入特征的边缘分布$P(x)$发生变化，但条件分布$P(y|x)$保持不变。 $P_{train}(x) \neq P_{test}(x), P_{train}(y|x) = P_{test}(y|x)$

语义偏移：标签空间 $\boldsymbol{Y}$ 发生变化，即测试集中可能出现训练集中未见的未知类（未出现在训练分布中的标签）。

$Y_{train} \subseteq Y_{test}, P_{train}(y|x) \neq P_{test}(y|x)$

- $P(y)$和 $P(y|x)$可能受到新类的引入或现有类的分布调整的影响。
- 通常$P(x)$ 也会变化，因为新增的类通常伴随新的特征模式。

标签偏移：标签的边缘分布$P(y)$发生变化，但条件分布 $P(x|y)$保持不变。$P_{train}(y) \neq P_{test}(y), P_{train}(x|y) = P_{test}(x|y)$

$P(x) = \sum_y P(x|y)p(y)$ $P(y|x) = \frac{P(x|y)P(y)}{P(x)}$

Introduction

协变量偏移和语义偏移两种情况的交叉下有四种情况：

ID：无协变量偏移下的已知类

OOD：协变量偏移下的已知类

Covariate-shifted ID (csID):协变量偏移下的已知类

Covariate-shifted OOD(csOOD):协变量偏移下的未知类

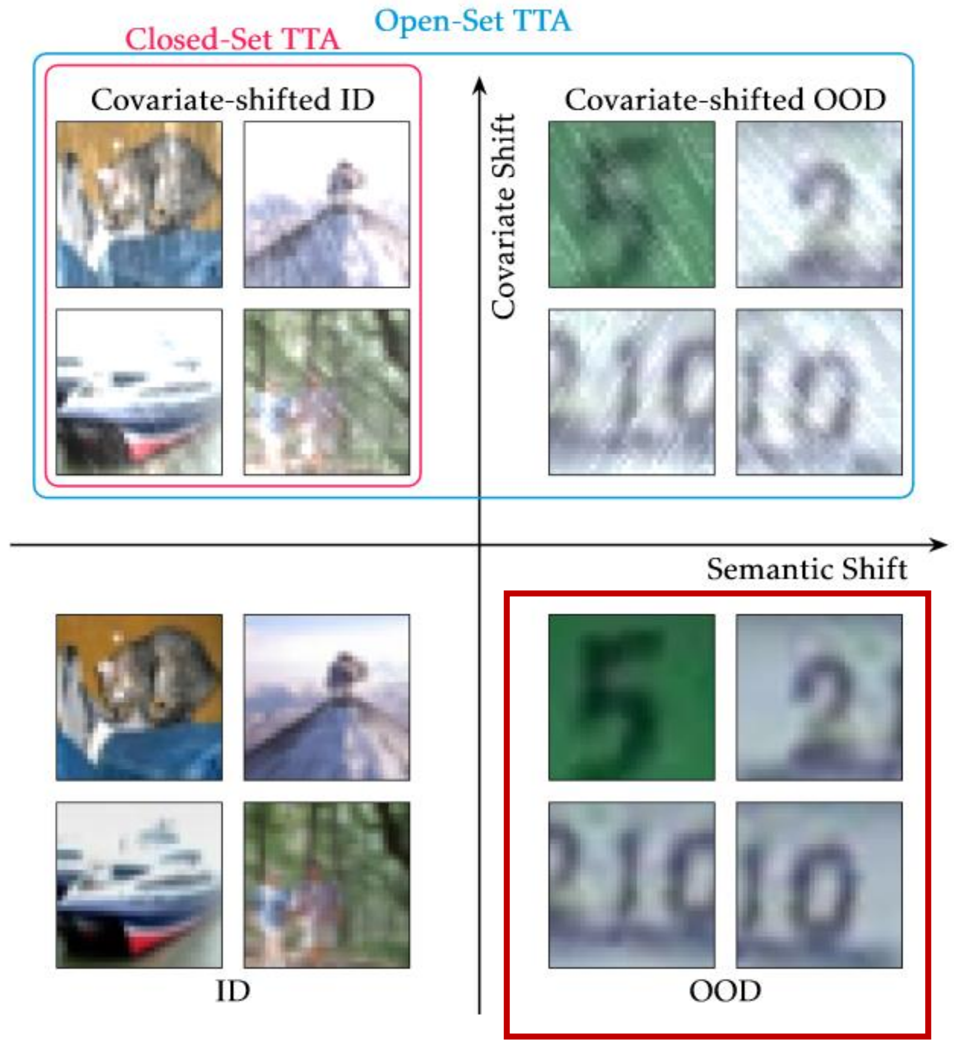
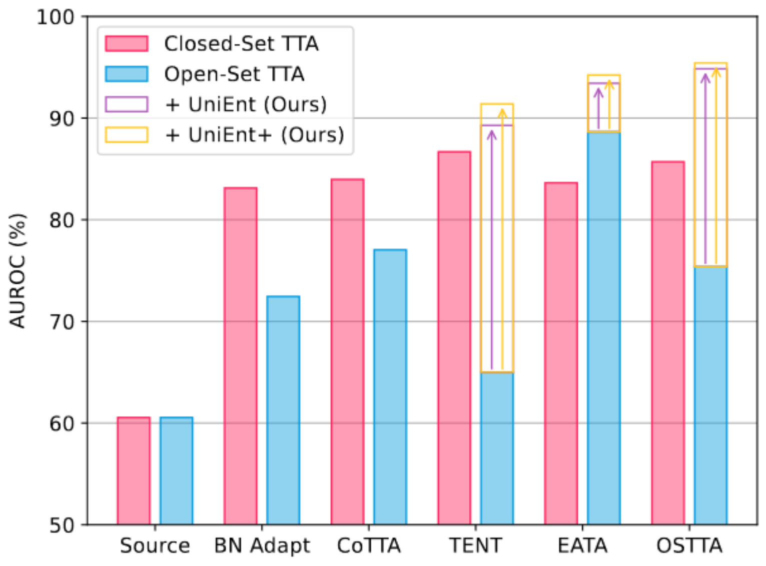


Figure 2. Comparison between closed-set TTA and open-set TTA.

Introduction

Test-Time Apdataion(TTA):通过在推理阶段仅利用未标注的目标域数据动态调整模型参数，以适应测试分布的变化，从而提升模型的泛化性能。现有TTA方法大多通过校准BN层统计信息，使用熵最小化策略更新模型。



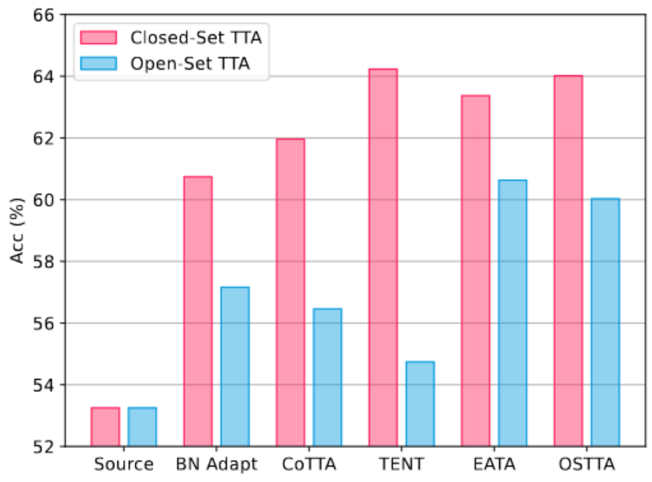
然而，大多数现有的TTA方法只专注于解决协变量漂移，而忽略了语义漂移

在协变量偏移和语义偏移同时存在的场景下，现有TTA方法的性能显著下降

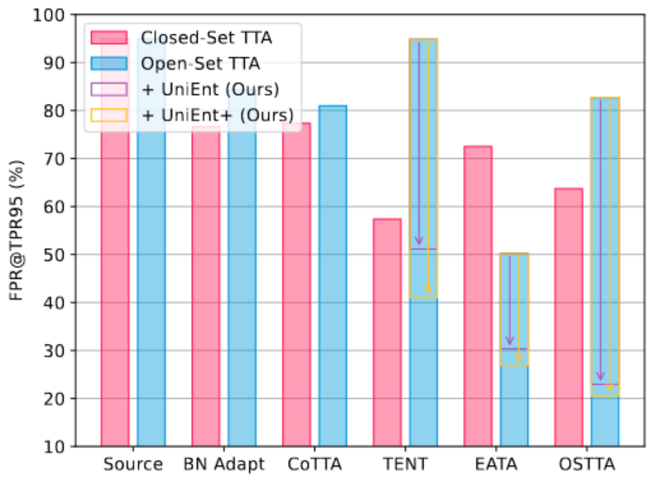
- 降低了csID的分类准确率
- csOOD的检测性能也受损

原因：

- 引入未知类的样本会导致模型对归一化统计量的错误估计，从而导致模型参数的更新不可靠
- 对未知类样本的熵最小化迫使模型输出可信的预测，从而破坏模型的置信度并导致模型区分已知类和未知类的能力下降



(a) Acc↑



(b) FPR@TPR95↓

Method

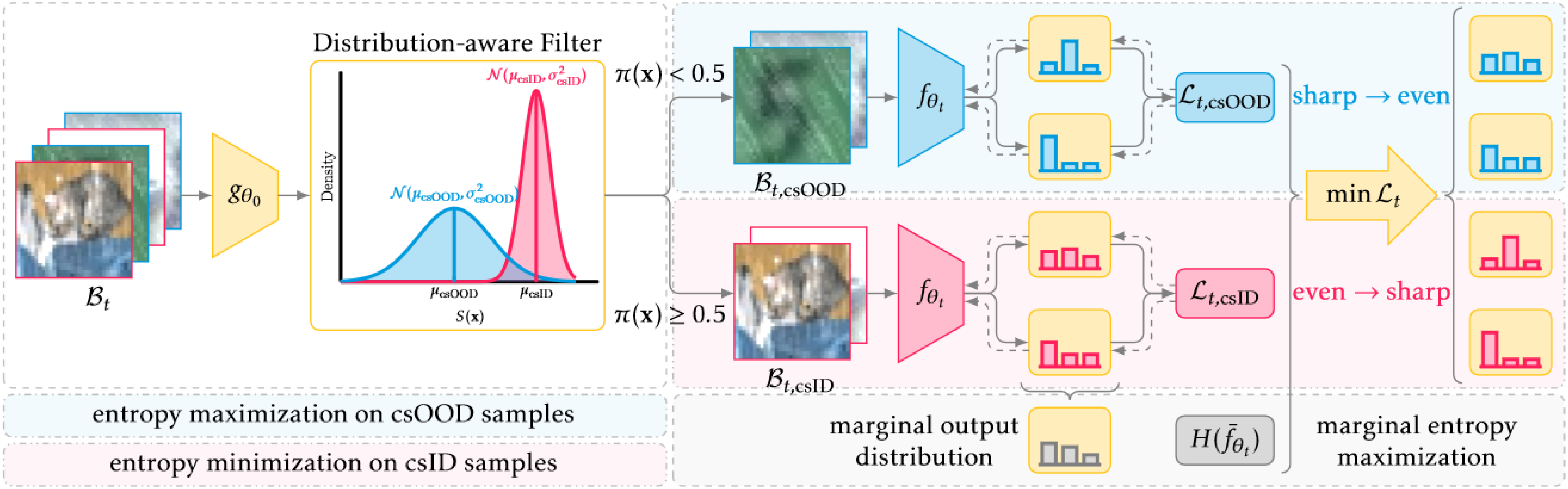


Figure 3. Illustration of the unified entropy optimization (UniEnt) framework. At timestamp t, mini-batch $\mathcal{B}_t$ may contain samples from csID and csOOD. First, we filter csOOD samples by csOOD score $S(\mathbf{x})$. Then, we perform entropy minimization for csID samples and entropy maximization for csOOD samples, we also adopt marginal entropy maximization to pervent model collapse. After optimization, we can yield better classification and detection performance tradeoff.

中南大學 CENTRAL SOUTH UNIVERSITY

Distribution-aware Filter

协变量偏移下的开放数据集 $\left\{ \begin{array}{l} \text{协变量偏移下的已知类csID} \\ \text{协变量偏移下的未知类csOOD} \end{array} \right\}$ $P_{open} = \pi P_{csID} + (1-\pi)P_{csOOD}$

相似性度量 $S(x) = v(\max_{c} \frac{g_{\theta_0}(x) \cdot p_c}{\|g_{\theta_0}(x)\| \|p_c\|})$

$g_{\theta_0}(x)$ 使用源域预训练模型的特征提取器对样本x提取的特征向量

p_c 训练数据集中第c类的原型，该类别的代表性特征向量

$v(\cdot)$ 将结果进行最小-最大归一化到 [0, 1] 区间

$S(x)$ 评估样本 x 的特征向量，找出它与哪个类别的原型最为相似，并测量二者的相似程度

Method

Distribution-aware Filter

$[S(x_1), S(x_2), S(x_3), \ldots, S(x_n)]$

区间划分：在[0,1]区间内划分若干个小区间

计数：对于每一个小区间，统计所有测试样本中落在该区间内的 $S(x)$ 值的数量

归一化为概率密度：将每个小区间的样本计数除以总的样本数量，得到每个小区间的相对频率。再将上述相对频率除以小区间的宽度，使得最终的高度代表的是概率密度

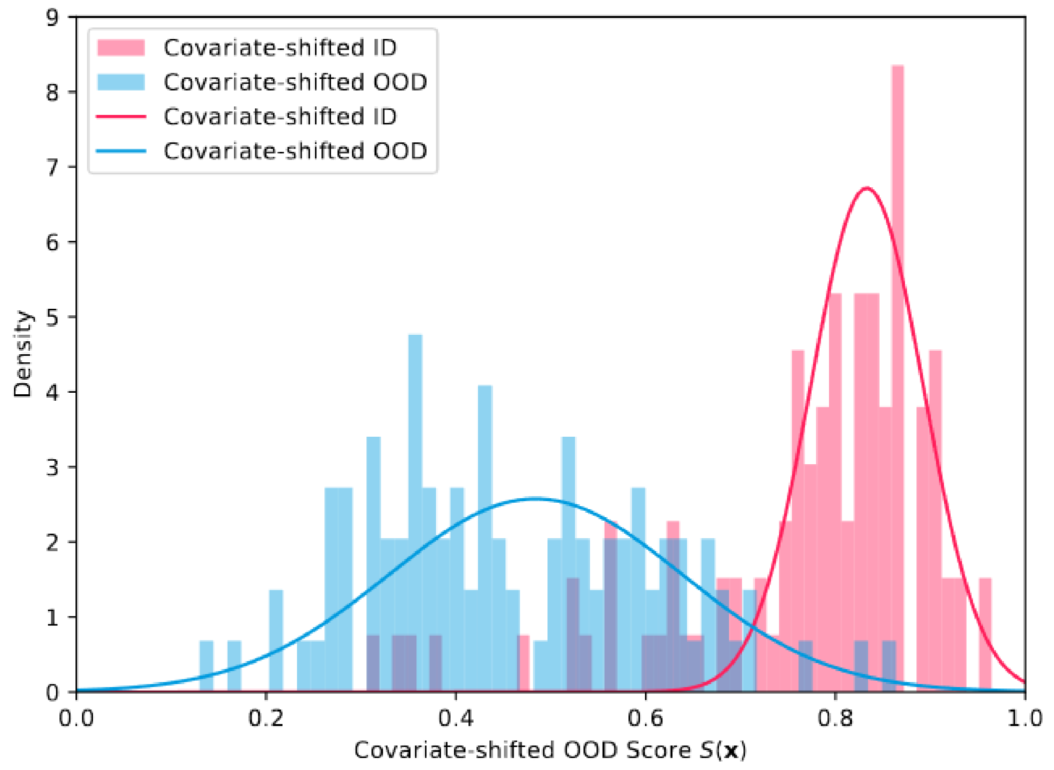


Figure 5. The csOOD score $S(\mathbf{x})$ presents a bimodal distribution.

$S(x)$ 呈双峰分布

- 协变量偏移下的已知类csID
  - *方差更小的分布*
- 协变量偏移下的未知类csOOD
  - *方差更大的分布，OOD的多样性更强，与模型训练时所见的数据模式差异较大*

Distribution-aware Filter

$$P(x)=\pi\cdot N(x|\mu_{csID},\sigma^2_{csID})+(1-\pi)\cdot N(x|\mu_{csOOD},\sigma^2_{csOOD})$$

1.**初始化**：

- csID分量的均值μ_{csID}和方差σ^2_{csID}
- csOOD分量的均值μ_{csOOD}和方差σ^2_{csOOD}
- 混合系数π，整个混合高斯分布下csID分量的概率

2. **E 步**：

计算后验概率$\pi(x)$：对于每一个测试样本 x，使用当前的 参数来计算它属于 csID 分量的概率$\pi(x)$

$$\pi(x)=\frac{\pi\cdot N(x|\mu_{csID},\sigma^2_{csID})}{\pi N(x|\mu_{csID},\sigma^2_{csID})+(1-\pi)N(x|\mu_{csOOD},\sigma^2_{csOOD})}$$

3.**M步**

更新模型参数：基于 E 步计算出的$\pi(x)$，重新估计混合高斯分布的参数

$$\pi=\frac{1}{N}\sum_{i=1}^{N}\pi(x_i)$$

$$\mu_{csID}=\frac{\sum_{i=1}^{N}\pi(x_i)x_i}{\sum_{i=1}^{N}\pi(x_i)},\mu_{csOOD}=\frac{\sum_{i=1}^{N}(1-\pi(x_i))x_i}{\sum_{i=1}^{N}(1-\pi(x_i))}$$

$$\sigma^2_{csID}=\frac{\sum_{i=1}^{N}\pi(x_i)(x_i-\mu_{csID})^2}{\sum_{i=1}^{N}\pi(x_i)},\sigma^2_{csOOD}=\frac{\sum_{i=1}^{N}(1-\pi(x_i))(x_i-\mu_{csID})^2}{\sum_{i=1}^{N}(1-\pi(x_i))}$$

$$\boldsymbol{B_{t,csID}=\{x|x\in B_t\wedge\pi(x)\geq 0.5\}}$$

$$\boldsymbol{B_{t,csOOD}=\{x|x\in B_t\wedge\pi(x)<0.5\}}$$

Entropy Optimization

闭集TTA

$$\min_{\theta_t} L_t = \frac{1}{\|B_t\|}\sum_{x\epsilon B_t} H\left(f_{\theta_t}(x)\right) - \lambda H(\bar{f}_{\theta_t})$$

无监督交叉熵

$$H\left(f_{\theta_t}(x)\right) = -\sum_{c=1}^{C} f_{\theta_t}^{c}(x) log f_{\theta_t}^{c}(x)$$

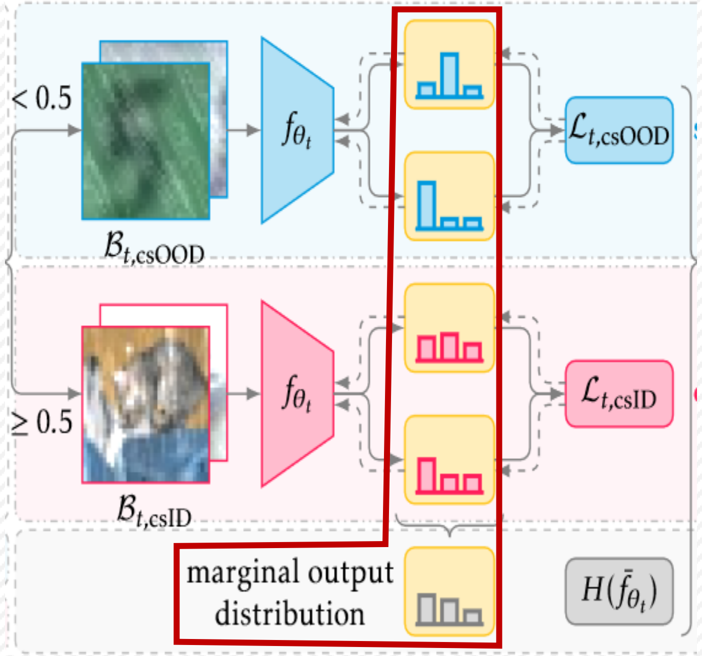
平均输出概率分布

$$\bar{f}_{\theta_t} = \frac{1}{\|B_t\|}\sum_{x\epsilon B_t} f_{\theta_t}(x)$$

边缘熵

$$H\left(\bar{f}_{\theta_t}\right) = -\sum_{c=1}^{C} \bar{f}_{\theta_t}^{c}(x) log \bar{f}_{\theta_t}^{c}(x)$$

最小化样本熵的平均值，从而使得模型对这些样本的预测更加确定



最大化边际熵确保模型不会过度集中在某些特定类别上，从而保持模型的泛化能力，避免过度拟合已知类别，防止模型崩溃

Unified Entropy Optimization-UniEnt

传统的熵优化方法通过最小化样本熵增强模型对样本预测的确定性，但是这也会导致对未知类别的预测过于自信，错误地将未知类别样本分类为某个已知类别，从而降低系统的整体可靠性。

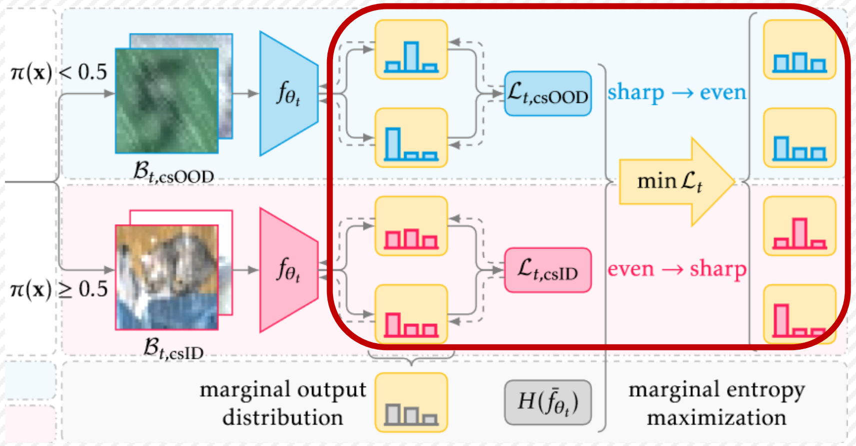
$$L_{t,csID} = \frac{1}{\|B_{t,csID}\|} \sum_{x \epsilon B_{t,csID}} H\left(f_{\theta_t}(x)\right)$$

最小化已知类别样本预测的熵，使得模型对已知类别的预测更加集中确定

$$L_{t,csOOD} = \frac{1}{\|B_{t,csOOD}\|} \sum_{x \epsilon B_{t,csOOD}} H\left(f_{\theta_t}(x)\right)$$

最大化未知类别样本预测的熵，使得模型对未知类别的预测更加均匀，降低模型对未知类别的过度自信

$$\min_{\theta_t} L_t = L_{t,csID} - \lambda_1 L_{t,csOOD} - \lambda_2 H(\bar{f}_{\theta_t})$$



Unified Entropy Optimization-UniEnt+

硬性地将样本分为已知类和未知类会引入噪声

$$\min_{\theta_t} L_t = \frac{1}{\|B_t\|}\sum_{x\epsilon B_t}\pi(x)H\left(f_{\theta_t}(x)\right)$$
$$-\lambda_1 \frac{1}{\|B_t\|}\sum_{x\epsilon B_t}(1-\pi(x))H\left(f_{\theta_t}(x)\right)$$
$$-\lambda_2\, H(\bar{f}_{\theta_t})$$

通过更精细地区分已知类别和未知类别样本，可以减少噪声的影响

Experiments

Method	CIFAR-10-C				CIFAR-100-C				Average			
	Acc$\uparrow$	AUROC$\uparrow$	FPR@TPR95$\downarrow$	OSCR$\uparrow$	Acc$\uparrow$	AUROC$\uparrow$	FPR@TPR95$\downarrow$	OSCR$\uparrow$	Acc$\uparrow$	AUROC$\uparrow$	FPR@TPR95$\downarrow$	OSCR$\uparrow$
Source [54]	81.73	77.89	79.45	68.44	53.25	60.55	94.98	39.87	67.49	69.22	87.22	54.16
BN Adapt [33]	84.20	80.40	76.84	72.13	57.16	72.45	84.29	47.10	70.68	76.43	80.57	59.62
CoTTA [46]	85.77	85.89	72.40	77.26	56.46	77.04	80.96	48.95	71.12	81.47	76.68	63.11
TENT [44]	79.38	65.39	95.94	56.73	54.74	65.00	94.79	42.24	67.06	65.20	95.37	49.49
+ UniEnt	**84.31** (+4.93)	92.28 (+26.89)	36.74 (-59.20)	80.32 (+23.59)	**59.07** (+4.33)	89.28 (+24.28)	51.14 (-43.65)	56.26 (+14.02)	**71.69** (+4.63)	90.78 (+25.59)	43.94 (-51.43)	68.29 (+18.81)
+ UniEnt+	84.03 (+4.65)	**93.18** (+27.79)	**32.74** (-63.20)	**80.62** (+23.89)	58.58 (+3.84)	**91.39** (+26.39)	**41.09** (-53.70)	**56.36** (+14.12)	71.31 (+4.25)	**92.29** (+27.09)	**36.92** (-58.45)	**68.49** (+19.01)
EATA [35]	80.92	84.32	71.66	72.63	**60.63**	88.64	50.18	57.24	70.78	86.48	60.92	64.94
+ UniEnt	84.31 (+3.39)	**97.15** (+12.83)	**13.25** (-58.41)	82.99 (+10.36)	59.75 (-0.88)	93.42 (+4.78)	30.36 (-19.82)	57.99 (+0.75)	72.03 (+1.26)	95.29 (+8.81)	21.81 (-39.12)	70.49 (+5.55)
+ UniEnt+	**85.18** (+4.26)	96.97 (+12.65)	14.28 (-57.38)	**83.67** (+11.04)	59.71(-0.92)	**94.23** (+5.59)	**26.87** (-23.31)	**58.19** (+0.95)	**72.45** (+1.67)	**95.60** (+9.12)	**20.58** (-40.35)	**70.93** (+6.00)
OSTTA [27]	**84.44**	72.74	77.02	65.17	**60.03**	75.37	82.75	51.35	**72.24**	74.06	79.89	58.26
+ UniEnt	82.46 (-1.98)	96.20 (+23.46)	16.37 (-60.65)	80.51 (+15.34)	58.69 (-1.34)	94.84 (+19.47)	22.95 (-59.80)	57.28 (+5.93)	70.58 (-1.66)	95.52 (+21.47)	19.66 (-60.23)	68.90 (+10.64)
+ UniEnt+	84.30 (-0.14)	**97.38** (+24.64)	**11.56** (-65.46)	**82.91** (+17.74)	58.93 (-1.10)	**95.42** (+20.05)	**20.59** (-62.16)	**57.69** (+6.34)	71.62 (-0.62)	**96.40** (+22.35)	**16.08** (-63.81)	**70.30** (+12.04)

Table 1. Results of different methods on CIFAR benchmarks. $\uparrow$ indicates that larger values are better, and vice versa. All values are percentages. The **bold** values indicate the best results, and the underlined values indicate the second best results.

Experiments

Method	Tiny-ImageNet-C			
	Acc↑	AUROC↑	FPR@TPR95↓	OSCR↑
Source [54]	22.29	53.79	93.41	16.29
BN Adapt [33]	37.00	61.06	90.90	28.50
TENT [44]	28.96	49.78	95.96	19.02
+ UniEnt	37.23 (+8.27)	**63.92** (+14.14)	89.72 (-6.24)	**30.18** (+11.16)
+ UniEnt+	**37.31** (+8.35)	63.83 (+14.05)	**89.12** (-6.84)	30.12 (+11.10)
EATA [35]	37.09	57.55	93.22	27.91
+ UniEnt	37.54 (+0.45)	**64.34** (+6.79)	**89.23** (-3.99)	30.59 (+2.68)
+ UniEnt+	**38.65** (+1.56)	62.30 (+4.75)	90.88 (-2.34)	**30.95** (+3.04)
OSTTA [27]	**37.29**	55.66	94.34	**27.74**
+ UniEnt	33.72 (-3.57)	**62.69** (+7.03)	89.67 (-4.67)	26.63 (-1.11)
+ UniEnt+	34.47 (-2.82)	61.28 (+5.62)	**89.56** (-4.78)	26.65 (-1.09)

Table 2. Results of different methods on Tiny-ImageNet-C.

Experiments

Method	$\mathcal{L}_{t,\text{csID}}$	$\mathcal{L}_{t,\text{csOOD}}$	CIFAR-10-C				CIFAR-100-C			
			Acc↑	AUROC↑	FPR@TPR95↓	OSCR↑	Acc↑	AUROC↑	FPR@TPR95↓	OSCR↑
TENT [44]	✗	✗	79.38	65.39	95.94	56.73	54.74	65.00	94.79	42.24
	✓	✗	**85.04** (+5.66)	81.80 (+16.41)	68.89 (-27.05)	73.57 (+16.84)	**59.30** (+4.56)	86.09 (+21.09)	63.65 (-31.14)	55.55 (+13.31)
	✓	✓	84.31 (+4.93)	**92.28** (+26.89)	**36.74** (-59.20)	**80.32** (+23.59)	59.07 (+4.33)	**89.28** (+24.28)	**51.14** (-43.65)	**56.26** (+14.02)
EATA [35]	✗	✗	80.92	84.32	71.66	72.63	**60.63**	88.64	50.18	57.24
	✓	✗	**85.53** (+4.61)	82.94 (-1.38)	67.95 (-3.71)	74.85 (+2.22)	60.46 (-0.17)	88.53 (-0.11)	54.30 (+4.12)	57.26 (+0.02)
	✓	✓	84.31 (+3.39)	**97.15** (+12.83)	**13.25** (-58.41)	**82.99** (+10.36)	59.75 (-0.88)	**93.42** (+4.78)	**30.36** (-19.82)	**57.99** (+0.75)
OSTTA [27]	✗	✗	84.44	72.74	77.02	65.17	**60.03**	75.37	82.75	51.35
	✓	✗	**84.86** (+0.42)	84.96 (+12.22)	62.66 (-14.36)	75.84 (+10.67)	58.95 (-1.08)	90.62 (+15.25)	44.79 (-37.96)	56.50 (+5.15)
	✓	✓	82.46 (-1.98)	**96.20** (+23.46)	**16.37** (-60.65)	**80.51** (+15.34)	58.69 (-1.34)	**94.84** (+19.47)	**22.95** (-59.80)	**57.28** (+5.93)

Table 3. Ablation study on CIFAR benchmarks. We investigate the effectiveness of $\mathcal{L}_{t,\text{csID}}$ and $\mathcal{L}_{t,\text{csOOD}}$ in Eq. (8) for UniEnt.

Method		0.1	0.2	0.5	1.0	Δ
TENT [44]	+ UniEnt	(59.09, 89.11, 51.68, 56.20)	(59.07, 89.28, 51.14, 56.26)	(58.92, 89.59, 50.16, 56.22)	(58.76, 89.95, 48.92, 56.21)	(0.33, 0.84, 2.76, 0.06)
	+ UniEnt+	(58.64, 91.18, 41.79, 56.34)	(58.58, 91.39, 41.09, 56.36)	(58.41, 91.68, 40.22, 56.33)	(58.12, 91.89, 39.68, 56.13)	(0.52, 0.71, 2.11, 0.23)
EATA [35]	+ UniEnt	(59.50, 93.34, 30.72, 57.72)	(59.75, 93.42, 30.36, 57.99)	(59.37, 92.56, 34.98, 57.40)	(59.58, 93.82, 28.29, 57.97)	(0.38, 1.26, 6.69, 0.59)
	+ UniEnt+	(59.73, 93.47, 30.25, 58.00)	(59.81, 93.88, 27.84, 58.17)	(59.71, 94.23, 26.87, 58.19)	(59.62, 93.47, 30.37, 57.91)	(0.19, 0.76, 3.50, 0.28)
OSTTA [27]	+ UniEnt	(58.85, 93.89, 26.59, 57.14)	(58.82, 94.32, 24.94, 57.24)	(58.69, 94.84, 22.95, 57.28)	(57.88, 94.80, 23.51, 56.51)	(0.97, 0.95, 3.64, 0.77)
	+ UniEnt+	(59.25, 94.19, 24.62, 57.54)	(59.15, 94.84, 22.29, 57.69)	(58.93, 95.42, 20.59, 57.69)	(58.20, 95.65, 20.12, 57.06)	(1.05, 1.46, 4.50, 0.63)

Table 4. Performance of UniEnt and UniEnt+ with varying λ_1 on CIFAR-100-C. The values in the table are presented as (Acc, AUROC, FPR@TPR95, OSCR). Δ is the difference between the maximum and minimum values when λ_1 take different values. Smaller Δ values represent better robustness.

Experiments

Method		0.1	0.2	0.5	1.0	Δ
TENT [44]	+ UniEnt	(59.44, 87.02, 60.32, 55.93)	(59.07, 89.28, 51.14, 56.26)	(58.09, 92.87, 33.24, 56.23)	(56.62, 94.53, 25.26, 55.24)	(2.82, 7.51, 35.06, 1.02)
	+ UniEnt+	(59.19, 87.95, 57.31, 56.04)	(58.58, 91.39, 41.09, 56.36)	(56.71, 94.57, 25.02, 55.34)	(53.13, 94.93, 24.19, 52.01)	(6.06, 6.98, 33.12, 4.35)
EATA [35]	+ UniEnt	(60.54, 88.14, 55.48, 57.15)	(60.06, 89.45, 50.99, 57.16)	(59.75, 93.42, 30.36, 57.99)	(58.26, 95.07, 22.18, 57.02)	(2.28, 6.93, 33.30, 0.97)
	+ UniEnt+	(60.35, 89.49, 50.20, 57.44)	(60.51, 91.03, 42.50, 58.02)	(59.71, 94.23, 26.87, 58.19)	(59.03, 95.28, 21.20, 57.81)	(1.48, 5.79, 29.00, 0.75)
OSTTA [27]	+ UniEnt	(58.69, 94.84, 22.95, 57.28)	(56.63, 95.43, 21.02, 55.46)	(49.85, 93.77, 32.12, 48.59)	(43.89, 91.19, 47.50, 42.41)	(14.80, 4.24, 26.48, 14.87)
	+ UniEnt+	(59.15, 94.84, 22.29, 57.69)	(57.55, 95.82, 18.91, 56.43)	(50.31, 94.09, 30.05, 49.11)	(43.66, 91.78, 43.35, 42.28)	(15.49, 4.04, 24.44, 15.41)

Table 5. Performance of UniEnt and UniEnt+ with varying λ_2 on CIFAR-100-C. Δ is the difference between the maximum and minimum values when λ_2 take different values.

Experiments

Method	2	4	6	8	10	Δ
Source [54]	70.84	69.28	69.32	69.18	68.44	2.40
BN Adapt [33]	72.56	72.48	72.52	72.44	72.14	0.42
TENT [44]	49.51	48.29	51.74	49.53	50.97	3.45
+ UniEnt	78.71	78.39	78.28	78.13	77.82	0.89
+ UniEnt+	78.65	78.23	78.23	78.07	77.68	0.97

Table 6. OSCR of UniEnt and UniEnt+ on CIFAR-10-C under different number of unknown classes.

Method	0.2	0.4	0.6	0.8	1.0	Δ
Source [54]	40.00	40.03	39.98	39.92	39.87	0.16
BN Adapt [33]	49.91	49.55	48.92	47.97	47.10	2.81
TENT [44]	47.68	44.12	44.06	42.90	42.16	5.52
+ UniEnt	56.84	57.48	57.13	56.77	56.26	1.22
+ UniEnt+	57.15	57.59	57.24	56.88	56.33	1.26

Table 7. OSCR of UniEnt and UniEnt+ on CIFAR-100-C under different ratios of csOOD to csID samples.

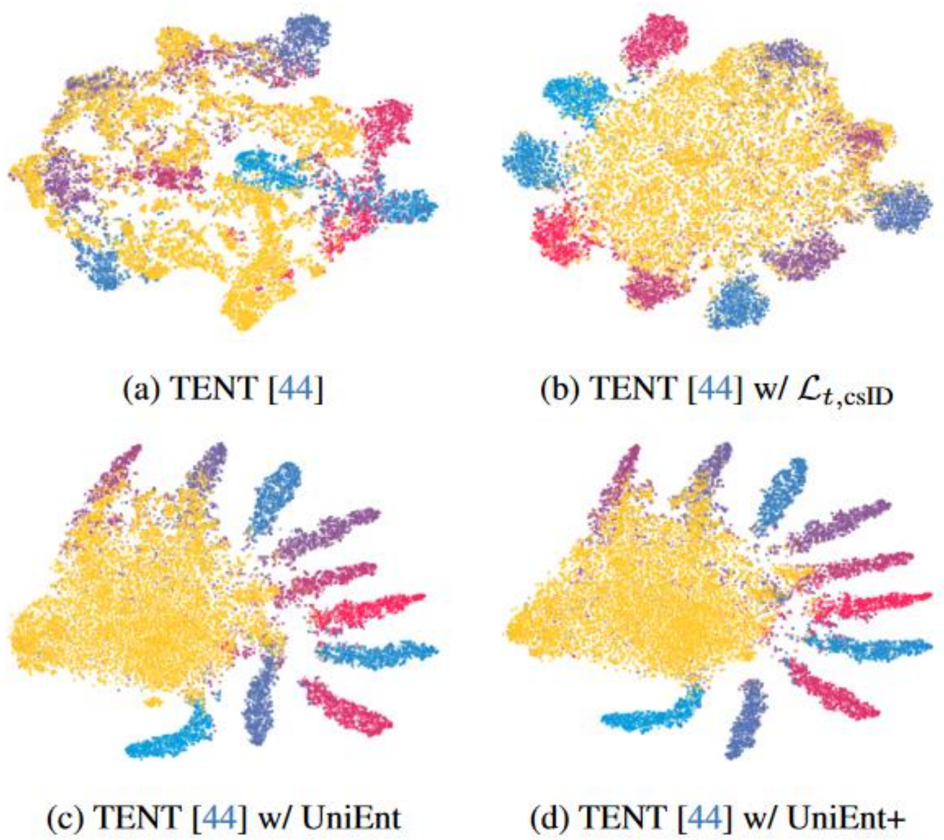


Figure 6. T-SNE visualization on CIFAR-10-C test set with SVHN-C as csOOD. red → blue denotes csID samples and yellow denotes csOOD samples.