南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Maintaining Consistent Inter-Class Topology in Continual Test-Time Adaptation

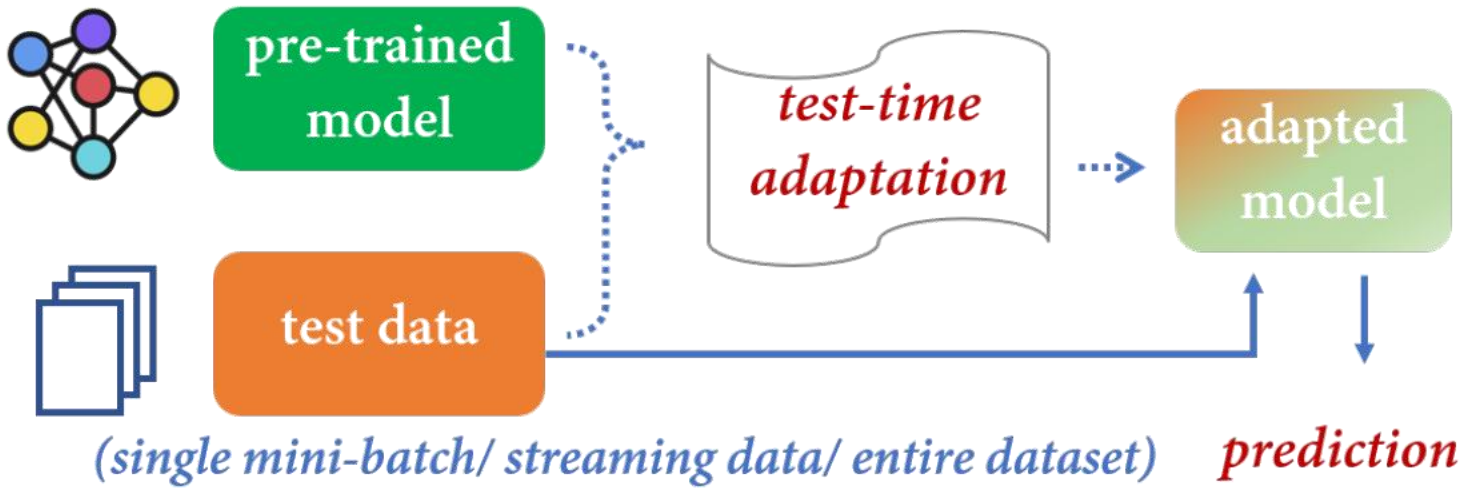
Chenggong Ni[1,2,3] * Fan Lyu[4] * Jiayao Tan[1,2,3] Fuyuan Hu[1,2,3] † Rui Yao[5] Tao Zhou[6]
[1]School of Electronics and Information Engineering, Suzhou University of Science and Technology
[2]Jiangsu Industrial Intelligent and Low-carbon Technology Engineering Center
[3]Suzhou Key Laboratory of Intelligent and Low-carbon Technology Application
[4]New Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences
[5]China University of Mining and Technology, [6]North Minzu University
{cgn, jiayaotan}@post.usts.edu.cn, fuyuanhu@mail.usts.edu.cn
fan.lyu@cripac.ia.ac.cn, ruiyao.contact@gmail.com, zhoutaonxmu@126.com

CVPR 2025

Background

➢ 测试时自适应（Test-Time adaptation）

模型在测试阶段不依赖额外标注数据，仅利用**测试数据自身信息动态调整参数**，从而适配测试数据的分布或语义特征，提升泛化性能。



➢ 三大特点

- **源域数据不可访问：**测试阶段无法访问训练阶段的源域样本或特征，模型只能依赖预训练权重进行自适应更新。
- **无监督设定：**测试数据缺乏真实标签，模型需基于自身预测或伪标签进行优化，易受噪声与不确定性影响。
- **单次样本交互：**模型在流式推理中与每个测试样本仅交互一次，需在有限更新下实现快速且稳定的分布适应。

Background

➢ 持续TTA（Continual Test-Time Adaptation）

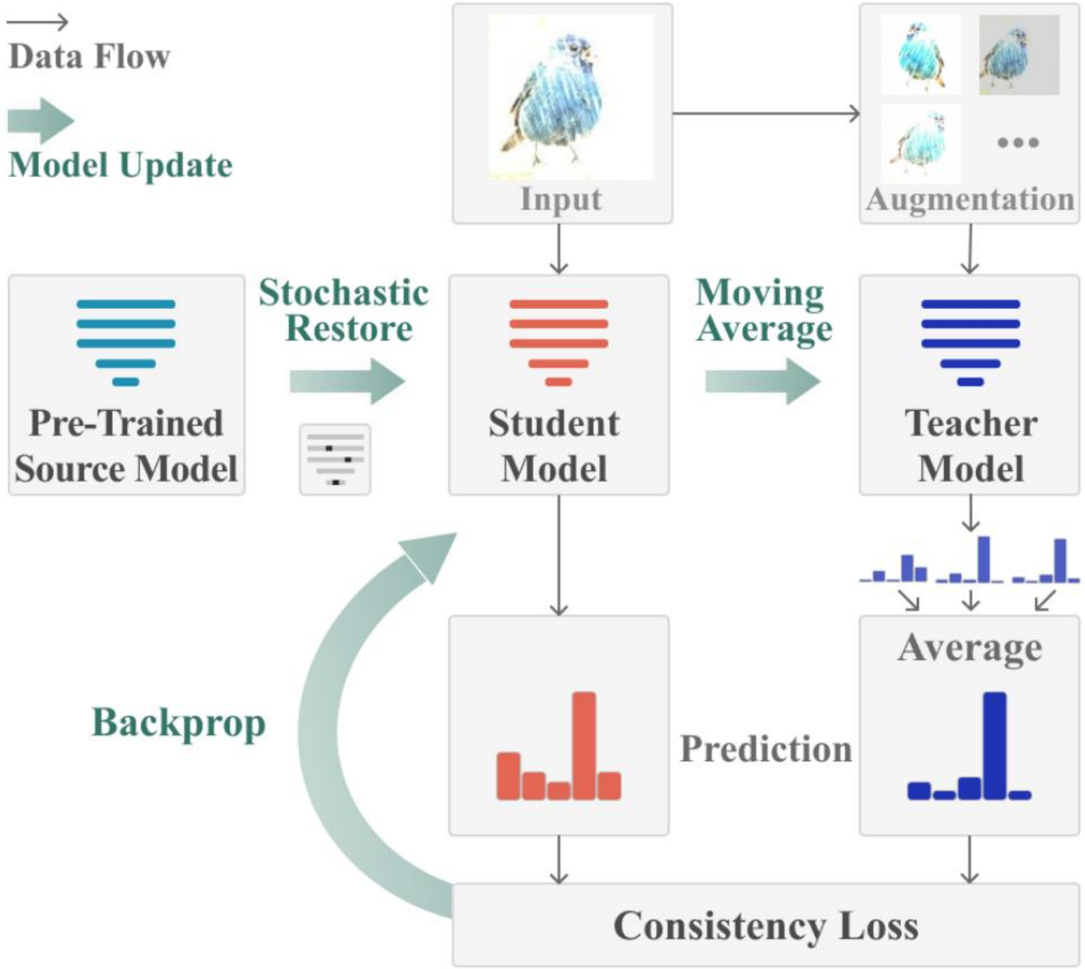
预训练源模型利用新环境未标记数据，**持续适应**动态目标域，缓解**域偏移**导致的性能下降。

➢ 难点

- 在不断变化的环境下，由于分布偏移，伪标签变得更加嘈杂和校准错误。因此，早期预测错误更有可能导致**误差累积**。
- 由于模型在很长一段时间内不断适应新的分布，源域的知识更难保存，导致**灾难性遗忘**。

设定	数据		学习阶段	
	源域	目标域	训练阶段	测试阶段
标准域自适应（DA）	是	固定	是	否
标准测试时训练	是	固定	是（辅助任务）	是
完全测试时自适应（TTA）	否	固定	否（预训练）	是
持续测试时自适应	否	持续变化	否（预训练）	是

Background



(1) Weight-Averaged Pseudo-Labels

$$\mathcal{L}_{\theta_t}(x_t^T) = -\sum_c \hat{y}'^T_{tc} \log \hat{y}^T_{tc}, \tag{1}$$

$$\theta'_{t+1} = \alpha\theta'_t + (1-\alpha)\theta_{t+1}, \tag{2}$$

(2) Augmentation-Averaged Pseudo-Labels

$$\tilde{y}'^T_t = \frac{1}{N}\sum_{i=0}^{N-1} f_{\theta'_t}(\text{aug}_i(x_t^T)), \tag{3}$$

$$y'^T_t = \begin{cases} \hat{y}'^T_t, & \text{if } \text{conf}(f_{\theta_0}(x_t^T)) \geq p_{th} \\ \tilde{y}'^T_t, & \text{otherwise}, \end{cases} \tag{4}$$

(3) Stochastic Restoration

$$M \sim \text{Bernoulli}(p), \tag{7}$$

$$W_{t+1} = M \odot W_0 + (1-M) \odot W_{t+1}, \tag{8}$$

Motivation

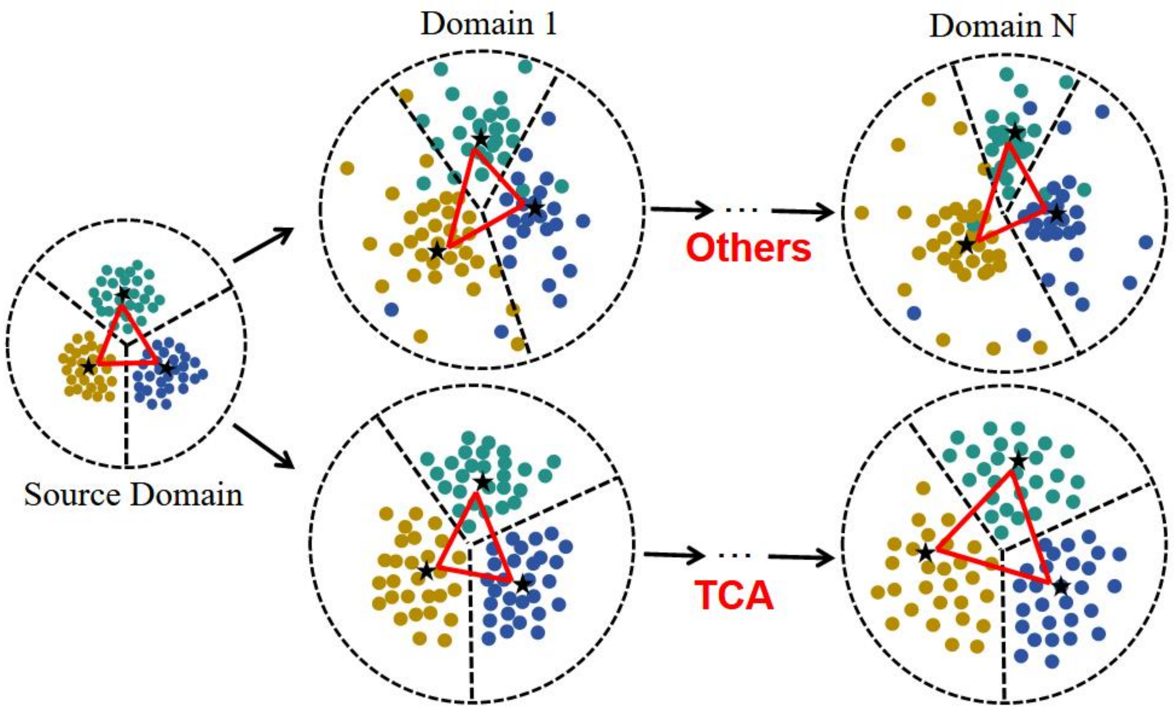
南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

➢ 挑战

误差的累积，这是由模型对未标记数据的**不稳定预测**产生的**错误伪标签**引起的。

➢ 根本原因

- **类间拓扑结构被破坏：**源域中学习的类间稳定拓扑结构被破坏，不同类别之间的特征分布不再保持原有的间隔与关系。
- **类内分布的均匀性变差：**类内特征分布的均匀性恶化，导致分类边界越来越模糊，同一类别的特征越来越散乱。



➢ 创新点

（1）**类间关系：**引入了类间均匀性损失，使用类质心来防止连续域移位期间的拓扑崩溃。
（2）**类内关系：**确保类内特征的合理紧凑性，间接支持类间稳定性。
（3）**批不平衡拓扑加权：**该加权考虑了每个批内的类分布不平衡，进一步优化了质心距离，并确保了类间拓扑的稳定性。

Methods

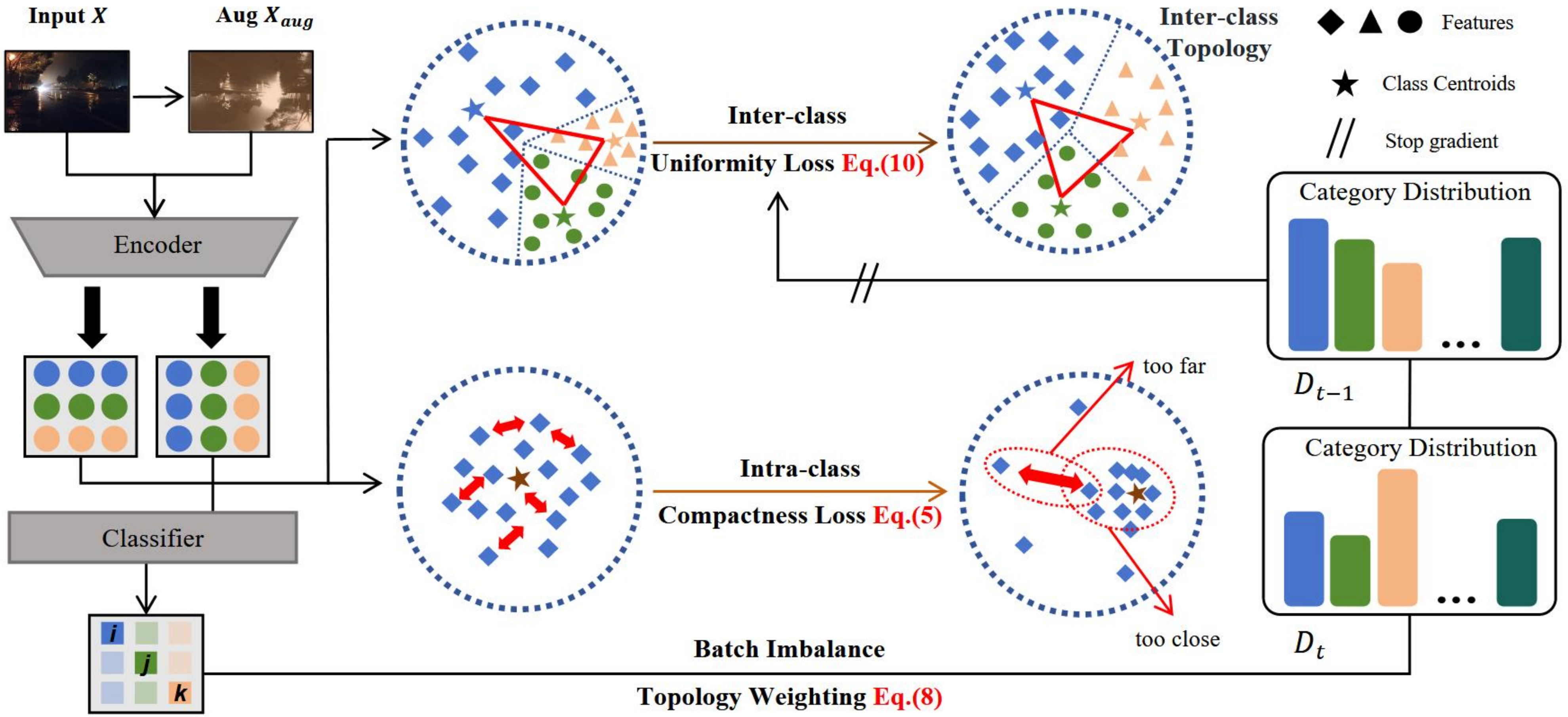


Figure 2. Framework overview. TCA first enhances the uniformity of inter-class feature distribution by leveraging improved pseudo-label prediction to compute pseudo-centroid proxies, thereby promoting inter-class feature alignment. Next, TCA ensures a compact intra-class feature distribution, mitigating imbalances within class representations. Finally, TCA dynamically adjusts the weights of inter-class centroids based on historical prediction distributions, preserving the latent topological relationships between classes.

Methods

南京航空航天大學 NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

➢ 类间拓扑图的构造

类拓扑关系是指特征空间内不同类之间的相对位置和距离。
将**类质心**作为类的代理、**质心的距离**作为类间的代理，构建类间拓扑图。

类间拓扑图定义：
G = <V, E>
质心集合：$V=\{v_1,v_2,v_3,...v_k\}$
质心间的邻域关系边：$E=\{(v_i,v_j) \mid w_{ij}=\|v_i-v_j\|_2\}$
权重w_{ij}：质心v_i和v_j之间的L2距离

$$\hat{y}_i = \operatorname{argmax}(g_s(\operatorname{Aug}(x_i))), \tag{1}$$

$$\mathbf{v}_k^{(t)} = \frac{1}{|\mathcal{I}_k|}\sum_{i=1}^{B}\mathbb{I}(\hat{y}_i = k)\cdot f_s(\operatorname{Aug}(x_i)), \tag{2}$$

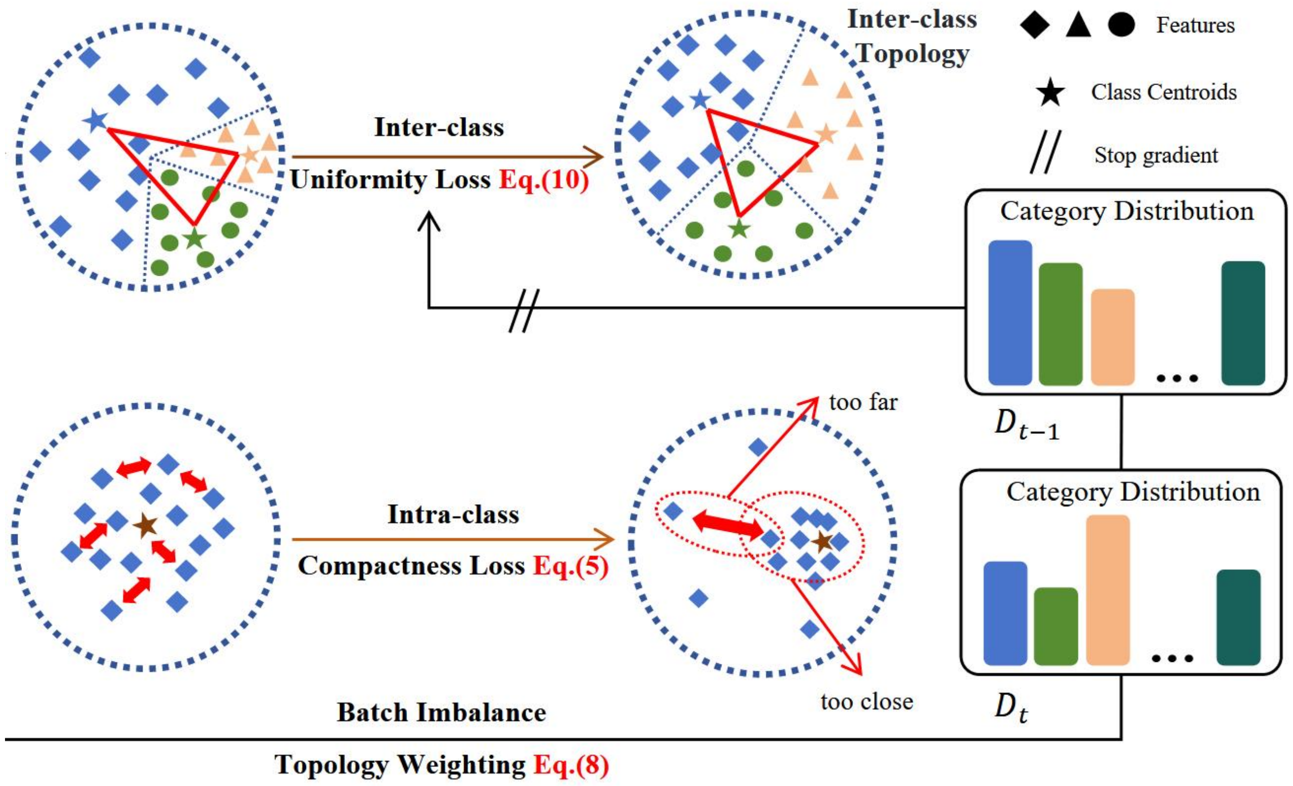
$$\mathbf{v}_k^{(t)} = \alpha\cdot\mathbf{v}_k^{(t-1)} + (1-\alpha)\cdot\mathbf{v}_k, \tag{3}$$

好处：确保在不同批处理大小和不平衡数据分布之间保持稳定的类间拓扑，在某些类缺失时利用历史信息。

Methods

➢ 类拓扑一致性约束

目标：缓解域偏移造成的拓扑结构破坏，避免类间分离度下降和类内混乱，从而防止伪标签失准与后续的误差累积。



①类内聚合

同一类别内的特征分布更紧凑、均匀

$$\mathcal{L}_{\text{intra}} = \log \mathbb{E}_{\mathbf{v}_k \in \mathcal{V}} \left[e^{-t \mathbb{E}_{z_i, z_j \sim \mathbf{v}_k} \left(\| z_i - z_j \|^2 \right)} \right], \quad (5)$$

参数含义

$\mathbb{E}v_k\epsilon\mathcal{V}$：对数据集中的每一个$v_k$取期望

$\| z_i\text{-}z_j\|^2$：两个特征向量的**欧氏距离平方**

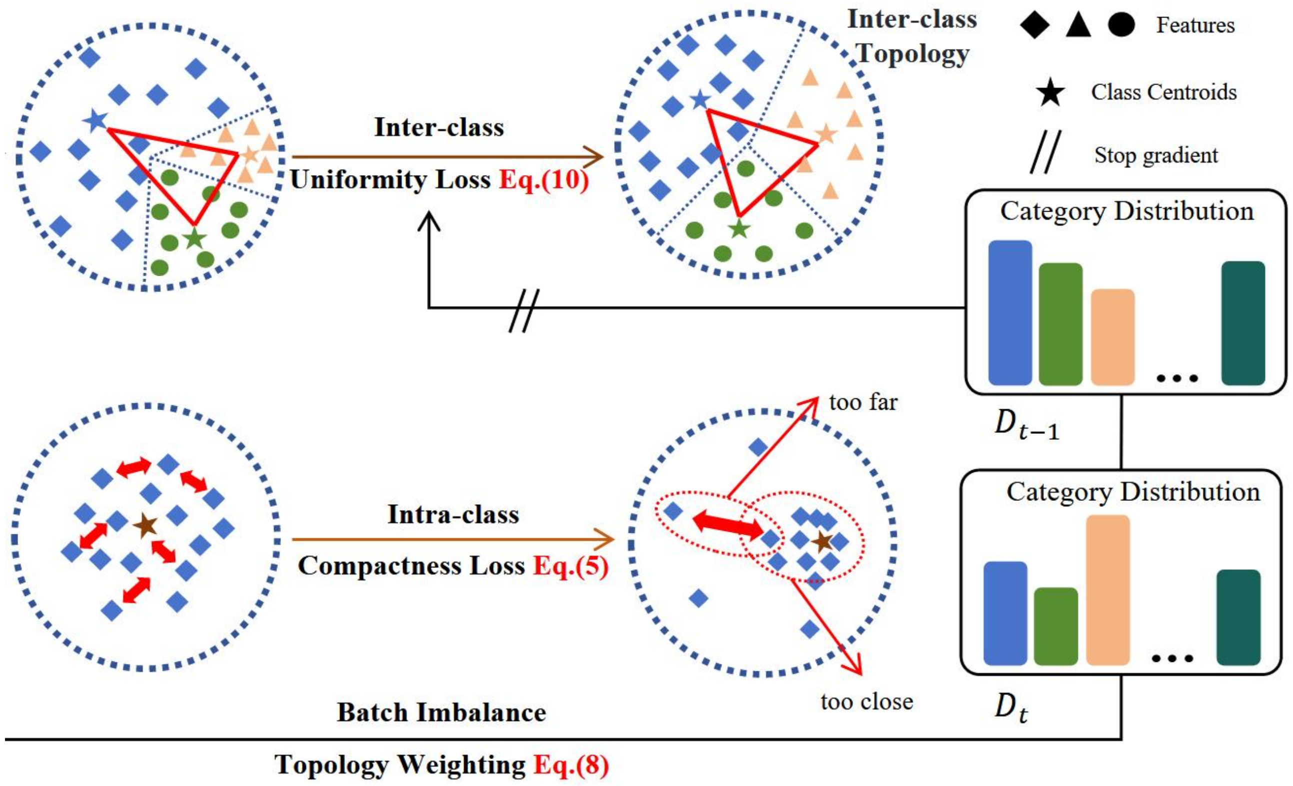
t：固定超参数,用于调节距离对损失的影响强度

$\mathbb{E}_{z_i, z_j \sim \mathbf{v}_k} \left(\| z_i - z_j \|^2 \right)$：对同一类别$v_k$内的所有特征对，计算他们的距离平方期望

Methods

➢ 类拓扑一致性约束

目标：缓解域偏移造成的拓扑结构破坏，避免类间分离度下降和类内混乱，从而防止伪标签失准与后续的误差累积。



②类间均匀

最小化不同类别中心对距离的平均指数衰减，促使类别中心均匀分布

$$\mathcal{L}_{\text{inter}} = \log\left(\frac{2}{|\mathcal{V}|(|\mathcal{V}|-1)}\sum_{\mathbf{v}_i,\mathbf{v}_j\in V,\, i<j} e^{-tw_{ij}^2}\right), \quad (4)$$

参数含义

|V|：类别数量

$\sum_{\substack{v_i,v_j\in\mathcal{V}\\ i<j}} e^{-tw_{ij}^2}$ ：不同类别中心对的高斯势（指数衰减项）

锐化类边界，稳定拓扑权重，提升质心精度，提高拓扑顶点的精度

Methods

➢ 批量不平衡拓扑加权

问题：未考虑实际测试场景中存在的**类间分布批次不平衡问题**（不同批次类别及样本量存在差异），无标签情况下这种不平衡会逐渐破坏类间拓扑结构。

解决办法：把类间拓扑距离和模型输出的类别频率结合起来，用于动态调整拓扑图中各边的权重。

①计算当前 batch 的类别频率 μ_k

$$\mu_k = \frac{\sqrt{|\mathcal{I}_k|}}{\sum_{k=1}^{C}\left(\sqrt{|\mathcal{I}_k|}+\epsilon\right)}, \tag{7}$$

②计算类间边权 μ_{ij}

$$\mu_{ij} = \frac{\mu_i + \mu_j}{2}. \tag{8}$$

③EMA 更新，保留历史信息

$$\mu_{ij}^{(t)} = \beta\mu_{ij}^{(t-1)} + (1-\beta)\mu_{ij}. \tag{9}$$

④将 BTW 融入类间一致性损失

$$\mathcal{L}'_{\text{inter}} = \log\left(\frac{\sum_{\mathbf{v}_i,\mathbf{v}_j\in\mathcal{V},\, i<j}\mu_{ij}\cdot e^{-tw_{ij}^2}}{\sum_{\mathbf{v}_i,\mathbf{v}_j\in\mathcal{V},\, i<j}\mu_{ij}}\right). \tag{10}$$

极端情况

batch=1，BTW仍然可以计算，类间边权不崩溃；
当前batch只有一个类别，类别边权自动反映类别缺失，拓扑结构可维持

Methods

南京航空航天大學 NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

➢ 损失函数

对称交叉熵损失函数：

$$\mathcal{L}_{\text{SCE}} = -\sum_{c=1}^{C} g_t(x) \log g_s(x) - \sum_{c=1}^{C} g_s(x) \log g_t(x).$$

对齐损失函数：$\mathcal{L}_{\text{align}}(f;\alpha) \triangleq -\mathbb{E}_{(x,y)\sim p_{\text{pos}}}\left[\|f(x)-f(y)\|_2^2\right],$

让“正样本对”的特征尽可能对齐

拓扑损失函数：$\mathcal{L}_{\text{topology}} = \mathcal{L}_{\text{inter}} + \mathcal{L}_{\text{intra}}.$

$$\mathcal{L} = \mathcal{L}_{\text{SCE}} + \lambda_1 \cdot \mathcal{L}_{\text{align}} + \lambda_2 \cdot \mathcal{L}_{\text{topology}}.$$

Algorithm 1 Topological Consistency Adaptation (TCA).

Input: Pre-trained model $g = f \circ h$, self-training loss $\mathcal{L}_{\text{SCE}}$, Alignment Loss weight λ_1, Uniformity Loss weight λ_2.

```
for Batch data 𝒳 do
    Augment 𝒳 to construct the class topological graph 𝒢 using Eq. (3)
    if First Batch then
        Compute inter-class uniformity L_inter via Eq. (4)
        Compute intra-class compactness L_intra via Eq. (5)
        return L_topology = L_inter + L_intra
    else
        Compute category weights μ_ij via Eq. (8)
        Update L_inter to L'_inter via Eq. (10)
        return L_topology = L'_inter + L_intra
    end if
    Compute L = L_SCE + λ1 · L_align + λ2 · L_topology
    Model g is updated using L
end for
```

Experiments

	Method	Gau.	shot	imp.	def.	glass	mot.	zoom	snow	fro.	fog	bri.	con.	ela.	pix.	jpeg	Mean
CIFAR10-C	Source only	72.3	65.7	72.9	46.9	54.3	34.8	42.0	25.1	41.3	26.0	9.3	46.7	26.6	58.5	30.3	43.5
	TENT [34]	24.8	20.6	28.6	14.4	31.1	16.5	14.1	19.1	18.6	18.6	12.2	20.3	25.7	20.8	24.9	20.7
	Ada [5]	29.1	22.5	30.0	14.0	32.7	14.1	12.0	16.6	14.9	14.4	8.1	10.0	21.9	17.7	20.0	18.5
	CoTTA [35]	24.3	21.3	26.6	11.6	27.6	12.2	10.3	14.8	14.1	12.4	7.6	10.6	18.3	13.4	17.3	16.2
	RMT [8]	24.0	20.4	25.6	12.6	25.4	14.2	12.2	15.4	15.1	14.1	10.3	13.7	17.1	13.5	16.0	16.7
	DSS [38]	24.1	21.3	25.4	11.7	26.9	12.2	10.5	14.5	14.1	12.5	7.8	10.8	18.0	13.1	17.3	16.0
	BeCoTTA [17]	22.9	**19.1**	26.9	**10.2**	27.5	12.7	10.4	14.7	14.3	12.4	**7.2**	**9.4**	20.9	15.2	20.2	16.3
	TCA	**22.4**	19.2	**23.0**	10.8	**23.2**	**11.6**	**9.9**	**13.1**	**13.1**	**11.8**	7.6	10.7	**16.4**	**12.0**	**15.5**	**14.7**
CIFAR100-C	Source only	73.0	68.0	39.4	29.3	54.1	30.8	28.8	39.5	45.8	50.3	29.5	55.1	37.2	74.7	41.2	46.4
	TENT [34]	**37.2**	**35.8**	41.7	37.9	51.2	48.3	48.5	58.4	63.7	71.1	70.4	82.3	88.0	88.5	90.4	60.9
	Ada [5]	42.3	36.8	38.6	27.7	40.1	29.1	27.5	32.9	30.7	38.2	25.9	28.3	33.9	33.3	36.2	33.4
	CoTTA [35]	40.1	37.7	39.7	26.9	38.0	27.9	26.4	32.8	31.8	40.3	24.7	26.9	32.5	28.3	33.5	32.5
	RMT [8]	40.5	36.1	**36.3**	27.7	**33.9**	28.5	26.4	**29.0**	29.0	32.5	25.1	27.4	28.2	**26.3**	**29.3**	30.4
	DSS [38]	39.7	36.0	37.2	26.3	35.6	27.5	25.2	31.4	30.0	37.8	24.2	26.0	30.0	26.3	31.3	30.9
	BeCoTTA [17]	42.1	38.0	42.2	30.2	42.9	31.7	29.8	35.1	33.9	38.5	27.9	32.0	36.7	31.6	39.9	35.5
	TCA	38.5	36.0	36.6	**25.8**	34.6	**27.2**	**25.1**	30.5	**27.0**	**30.1**	**24.1**	**25.7**	**27.3**	26.6	30.3	**29.7**
ImageNet-C	Source only	97.8	97.1	98.2	81.7	89.8	85.2	78.0	83.5	77.1	75.9	41.3	94.5	82.5	79.3	68.6	82.0
	TENT [34]	81.6	74.6	72.7	77.6	73.8	65.5	**55.3**	61.6	63.0	51.7	38.2	72.1	50.8	47.4	53.3	62.6
	Ada [5]	82.9	80.9	78.4	81.4	78.7	72.9	64.0	63.5	64.5	53.5	38.4	66.7	54.6	49.4	53.0	65.5
	CoTTA [35]	84.7	82.1	80.6	81.3	79.0	68.6	57.5	60.3	60.5	48.3	36.6	66.1	47.3	41.2	46.0	62.7
	RMT [8]	80.2	76.4	74.5	77.1	74.4	66.2	57.6	57.0	**59.1**	48.0	39.1	60.6	47.3	42.5	**43.4**	60.2
	DSS [38]	82.3	78.4	76.7	81.9	77.8	66.9	60.9	**50.8**	60.9	**47.7**	**35.4**	69.0	47.5	**40.9**	46.2	62.2
	BeCoTTA [17]	84.1	74.3	**72.2**	77.4	**71.9**	63.4	**55.1**	57.2	61.2	50.7	36.4	66.1	49.2	45.6	48.4	60.9
	TCA	**78.3**	**71.8**	73.5	**74.4**	73.5	**63.3**	56.5	56.9	59.4	48.1	39.6	**59.6**	**47.2**	42.9	44.7	**59.3**

Table 1. Classification error rate (%) on CIFAR10-to-CIFAR10-C, CIFAR100-to-CIFAR100-C, and ImageNet-to-ImageNet-C. All results are evaluated with the largest corruption severity level 5 in an online manner. We report the performance of our method averaged over 5 runs. Bold text indicates the best.

Experiments

CCC数据集：以从一个域到另一个域的平滑过渡为特征，模仿现实中环境的变化。

Method	CCC-Easy	CCC-Medium	CCC-Hard	Average
CoTTA [35]	14.9±0.88	7.7±0.43	1.1±0.16	7.9
ETA [26]	41.4±0.95	1.1±0.43	0.2±0.05	14.2
EATA [26]	48.2±0.60	35.4±1.02	8.7±0.80	30.8
SANTA [4]	47.8±0.46	32.7±0.80	9.1±0.60	29.9
RDumb [27]	**49.3±0.88**	38.9±1.40	9.6±1.60	32.6
TCA	49.1±0.35	**39.5±0.53**	**10.1±0.22**	**32.9**

小批量数据集的测试效果

	Batchsize	CIFAR10-C	CIFAR100-C	ImageNet-C
CoTTA	1	81.6	85.3	91.2
	50	18.1	38.5	65.3
	100	16.8	35.2	64.9
TCA	1	61.5	66.2	75.4
	50	15.2	32.6	65.2
	100	14.9	30.2	61.4
	150	14.8	30.1	60.1
	200	**14.7**	**29.7**	**59.3**
	300	14.8	29.8	59.4

消融实验

BTW	$\mathcal{L}_{intra}$	$\mathcal{L}_{inter}$	$\mathcal{L}_{align}$	CIFAR10-C	CIFAR100-C	ImageNet-C
-	-	-	-	16.05	32.11	63.14
-	-	-	✓	15.92	31.66	62.83
-	-	✓	-	15.57	31.38	62.08
-	✓	-	-	15.68	31.44	62.41
-	✓	✓	-	15.41	31.01	60.95
✓	✓	✓	-	14.85	29.83	59.51
✓	✓	✓	✓	14.70	29.72	59.31

Table 4. Ablation studies on four components (BTW, $\mathcal{L}_{intra}$, $\mathcal{L}_{inter}$, $\mathcal{L}_{align}$) across CIFAR10-C, CIFAR100-C, and ImageNet-C datasets.

Experiments

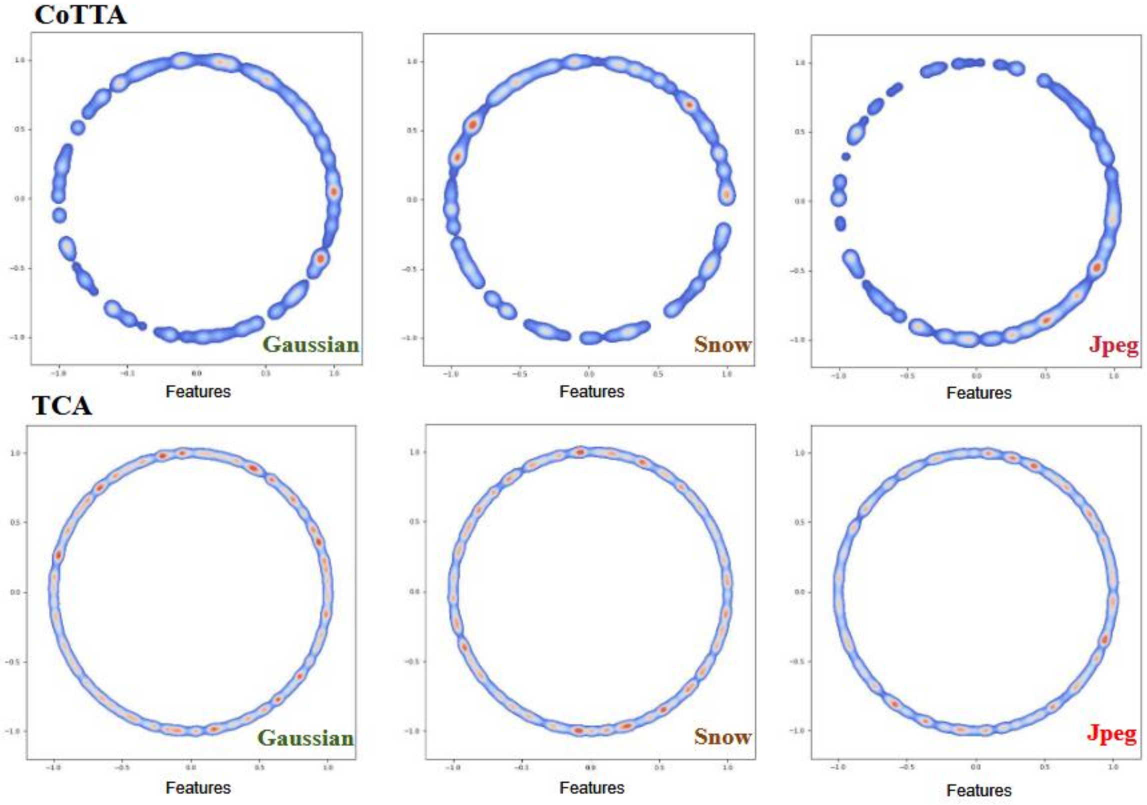


Figure 3. Feature visualizations from ten randomly selected batches of CIFAR10-C under three noise distributions: Gaussian, Snow, and JPEG. The upper panel illustrates the results of CoTTA, while the lower panel showcases those of TCA.

Inter-class Uniformity

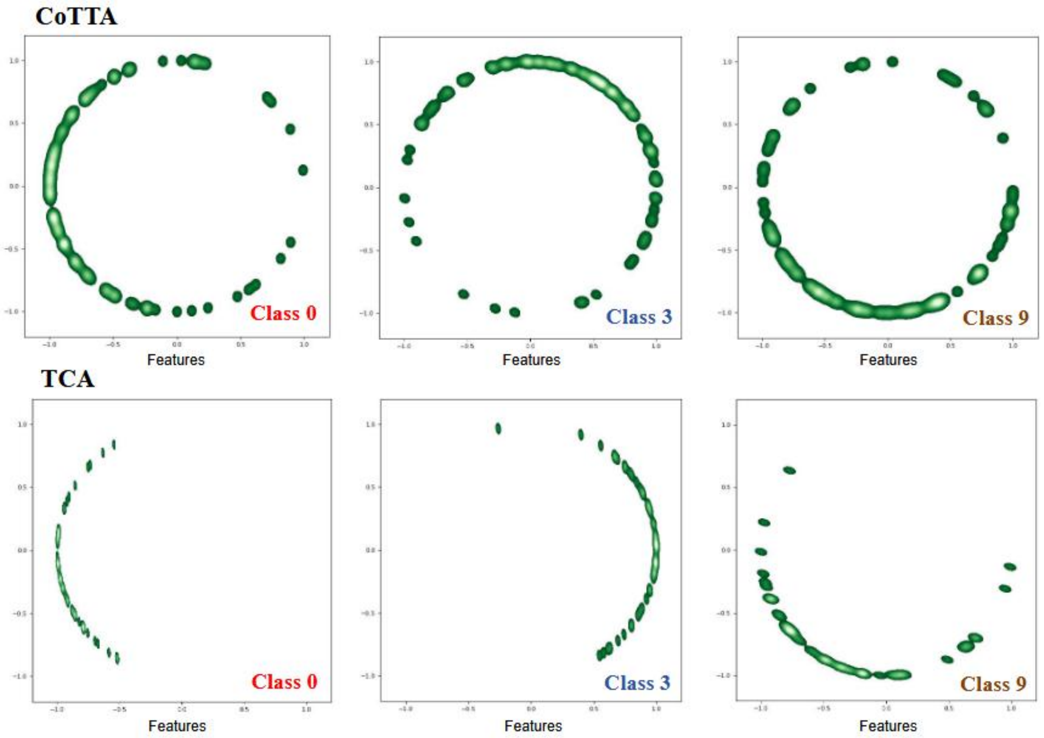


Figure 4. Visualization of intra-class feature distribution in CIFAR-10 under elastic noise conditions involved selecting class 0, class 3, and class 6 (from left to right) and randomly visualizing 10 batches for each. The upper figure illustrates CoTTA, whereas the lower figure depicts TCA.

Intra-class Uniformity

Thanks