



Multi-adversarial Faster-RCNN with Paradigm Teacher for Unrestricted Object Detection

Zhenwei He, Lei Zhang*, Xinbo Gao and David Zhang

指导老师： 张磊

阅读分享汇报人： 张龙港、张军、秘思雨

目录

CONTENTS



研究背景与
动机



研究方法与
内容



研究结果与
讨论



结论与展望



研究背景与动机



研究背景与动机

1 计算机视觉中的目标检测



任务定义：使用给定训练集训练检测器，以完成检测出训练集中包含类别的对象检测

2 传统训练方式带来的问题



Source

Target



研究背景与动机

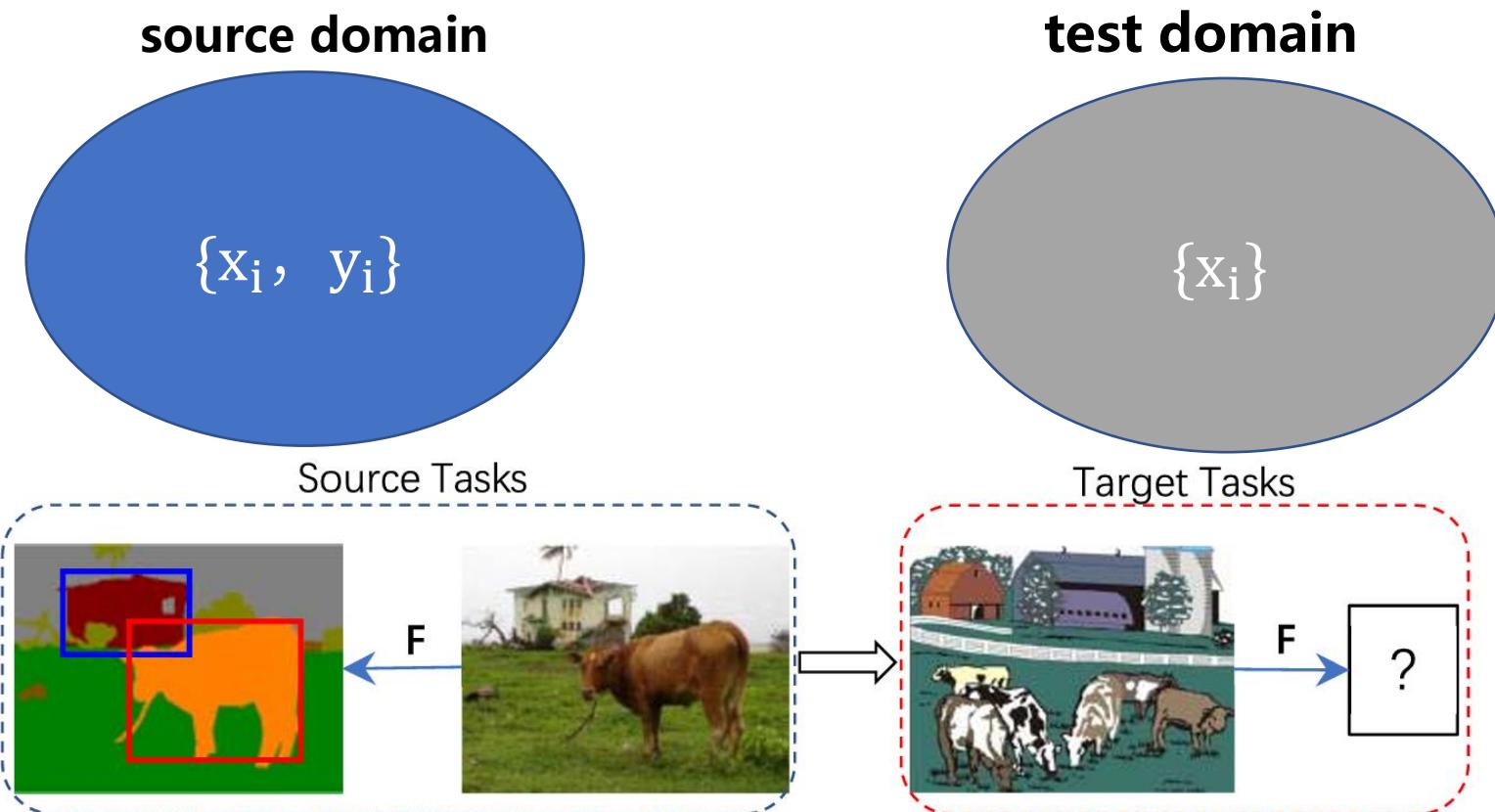
跨域目标检测

传统目标检测存在的问题:

source domain \neq test domain

域适应目标检测

任务定义:



任务目标:



研究背景与动机

跨域目标检测中的理论依据

$$\epsilon_t(h) \leq \epsilon_s(h) + d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{D}_s, \mathcal{D}_t) + \lambda$$

$$\lambda = \epsilon_s(h^*) + \epsilon_t(h^*)$$

$$h^* = \arg \min_{h \in \mathcal{H}} (\epsilon_s(h) + \epsilon_t(h))$$

$d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{D}_s, \mathcal{D}_t)$ 衡量source domain与test domain之间域差异

λ 衡量source domain与test domain之间域相似性,

两个需要考虑的目标

➤ 降低域之间的差异 减小 $d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{D}_s, \mathcal{D}_t)$ 如何减小域间的差异呢?

➤ 保留域之间的相似 减小 λ 如何保证域间的相似呢?



相关工作

域适应目标检测相关工作

任务定义: 领域适应旨在通过减少分布差异来桥接不同的领域或任务

任务方法: 对抗性学习将领域适应视为极小极大问题

相关工作: Self-training based 和 Domain alignment based 两大类

➤ Self-training based : 使用目标pseudo labels训练目标模型

①Cai et al 提出了平均教师来生成伪标签来训练目标域[1]

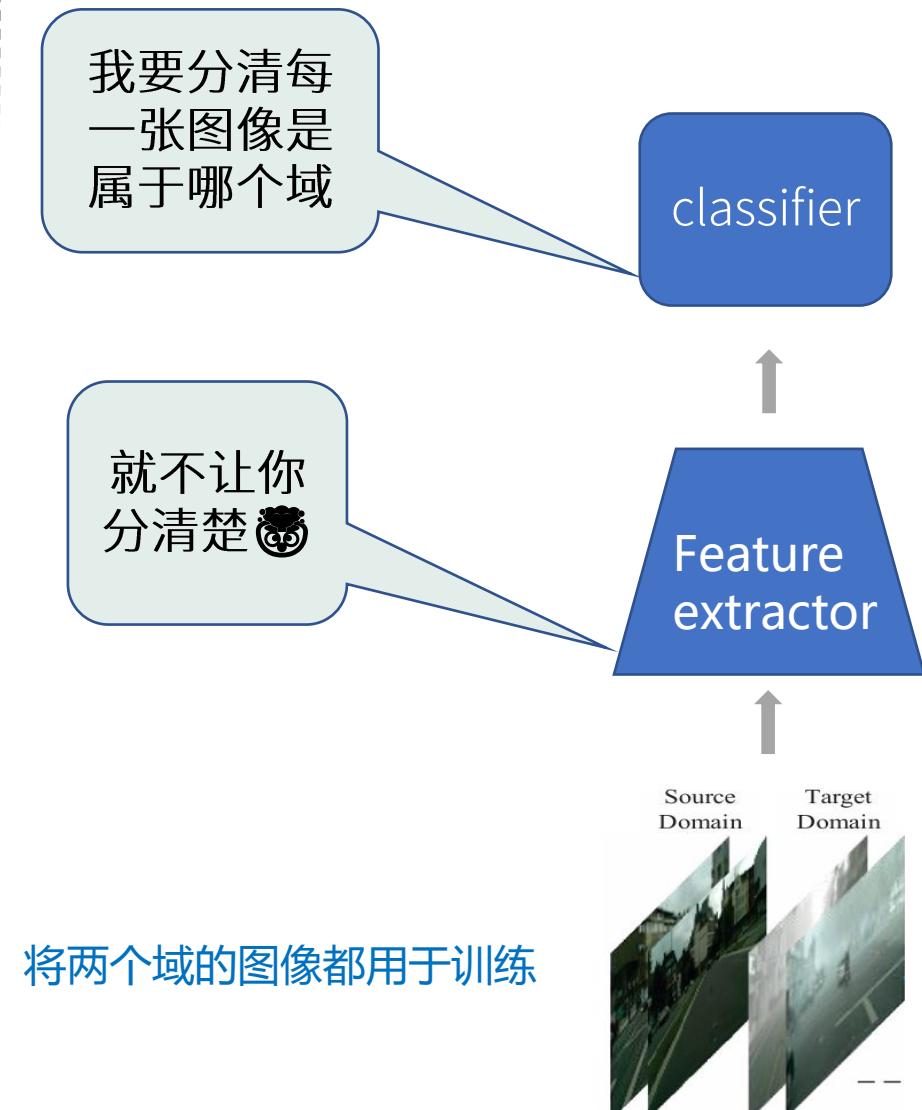
②Khodabandeh et al 利用 Kullback-Leibler 散度改进了伪标签[2]

弊端: 基于 source-only 模型生成的低质量伪标签会降低目标域的性能

➤ Domain alignment based: 学习与域无关的特征以减少域差异

③Chen et al 在图像级别和实例级别都解决了域视差[3]

④Zhu et al 重点关注跨域目标检测器中的 “where to look” 和 “how to align” [4]



[1]Exploring Object Relation in Mean Teacher for Cross-Domain Detection (CVPR,2019)

[2] A Robust Learning Approach to Domain Adaptive Object Detection (ICCV,2019)

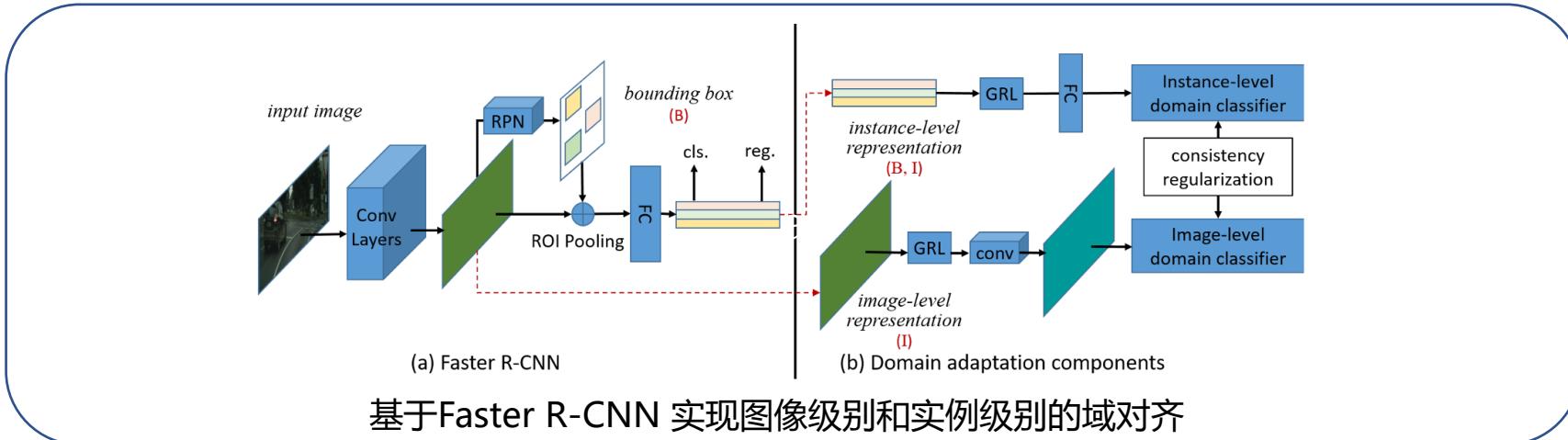
[3] Domain Adaptive Faster R-CNN for Object Detection in the Wild (CVPR,2018)

[4] Adapting Object Detectors via Selective Cross-Domain Alignment (CVPR,2019)

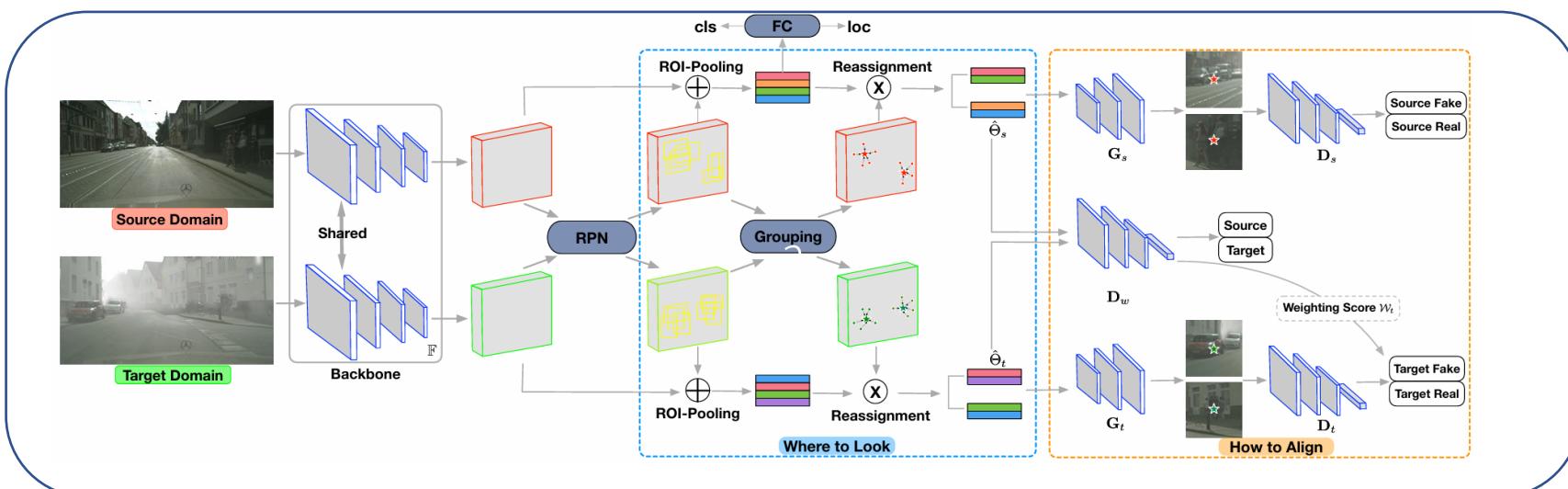


相关工作

- Domain alignment based: 学习与域无关的特征以减少域差异
③Chen et al在图像级别和实例级别都解决了域视差[3]



- ④Zhu et al 重点关注跨域目标检测器中的“where to look” 和 “how to align” [4]





动机

问题与解决方法

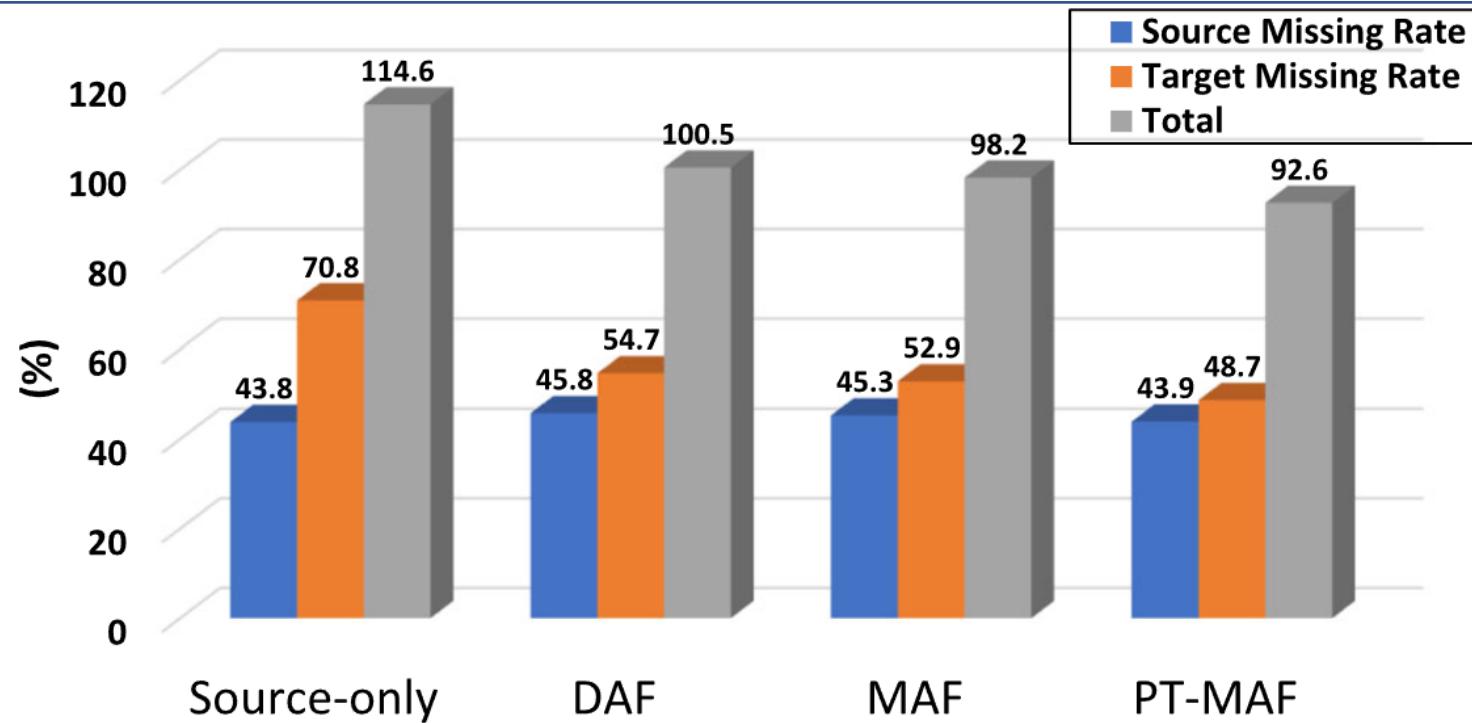
问题: 上述所有模型都将不同domain与对抗性学习结合起来，而忽略了模型的适应性

只关注减小

$d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{D}_s, \mathcal{D}_t)$

忽略

λ





动机

问题与解决方法

思考：一个中心，两个基本点，两手都要抓，两手都要硬！

一个中心：以模型在target domain上的性能为中心

两个基本点：①减小域之间的差异②保证域间相似性



论文中如何实现这两个基本点呢？

域差异？域对齐？如何对齐？对齐什么？

域相似性？什么相似性？如何保证相似性？



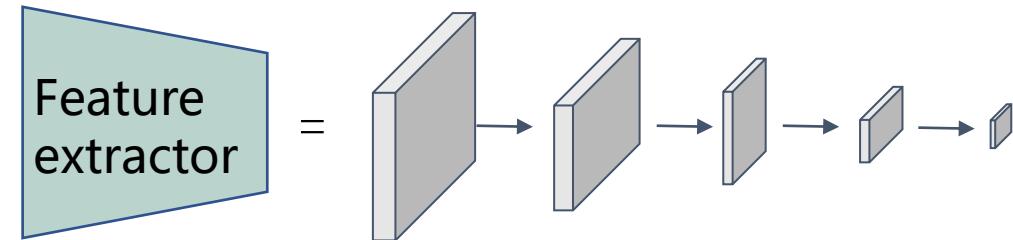


动机

问题与解决方法

解决问题

动机：Tommasi等人揭示了CNN模型倾向捕获数据偏差的偏好[5] → 对于域间的偏见存在于CNN的各个layer之间



方法：

减小域间的差异 (MAF)

- ① domain alignment 在不同layer之间进行，提出**分层域特征对齐 (HDFA)** 模块，此外，考虑到大特征图（例如 125×250 ）会形成很大的训练集，这会显著降低训练效率，我们提出了一种新的**规模缩减模块 (SRM)** 来减少训练样本的数量
- ② 实例级的特征结合不同的局部区域来表征proposal的语义信息，引入**聚合提议特征对齐 (APFA)** 模块来进行实例级的特征对齐

保留域间的相似 提出了带有范式教师的 MAF (PT-MAF)，以保留源知识并减轻范式教师失控的源错误



研究方法

- MAF

- PT-MAF



研究方法 (MAF)

A-distance(衡量域差异)

$$d_A(\mathcal{D}_s, \mathcal{D}_t) = 2 \left(1 - 2 \min_{q \in \mathcal{Q}} \text{err}(q(x)) \right)$$

利用特征表示 f , 最小化域差异

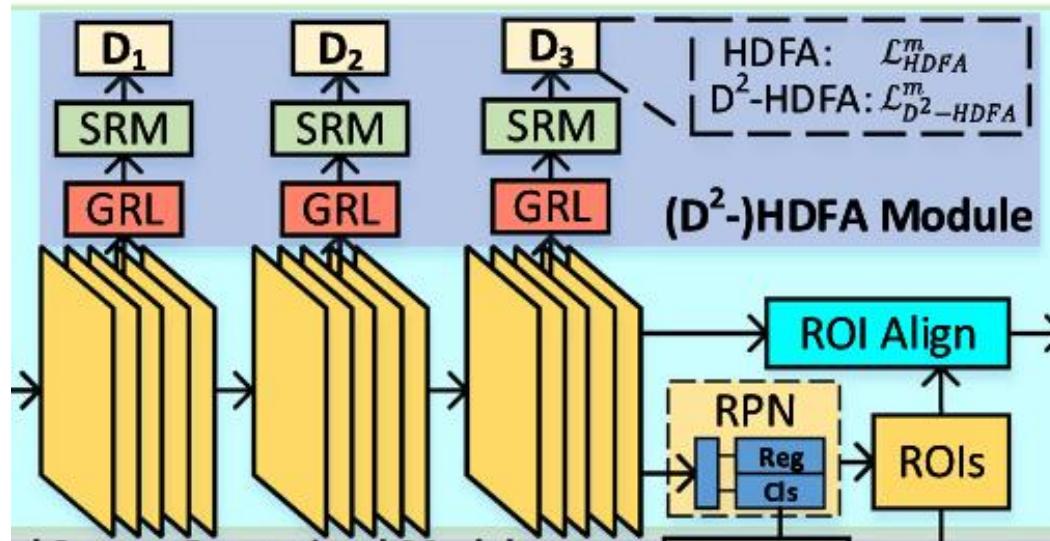
$$\begin{aligned} \min_f d_A(\mathcal{D}_s, \mathcal{D}_t) &\propto \min_f 2 \left(1 - 2 \min_{q \in \mathcal{Q}} \text{err}(q(x)) \right) \\ &\propto \max_f \min_{q \in \mathcal{Q}} \text{err}(q(x)) \end{aligned}$$

最大化最小域分类误差, 这是一个对抗问题, 特征表示 f 和域分类器 q 之间的对抗学习来解决, 从而实现域混淆 (对齐)。



研究方法 (MAF)

分层域特征对齐 (HDFA) - 对齐源域和目标域的特征



第 m 个卷积块中对抗性鉴别器的极小极大学习机制为：

$$\min_{\theta^m} \max_{\mathbf{w}^m} \mathcal{L}_{H DFA}^m$$

$$\mathcal{L}_{H DFA}^m = \sum_{u,v} L_c \left(D^m \left(C^m(I_i, \mathbf{w}^m)^{(u,v)}, \theta^m \right), d_i \right)$$

L_c 交叉熵损失, C^m 表示像素 (u, v) 处的通道特征, d_i 是域标签, 源域标记为 1, 目标域标记为 0

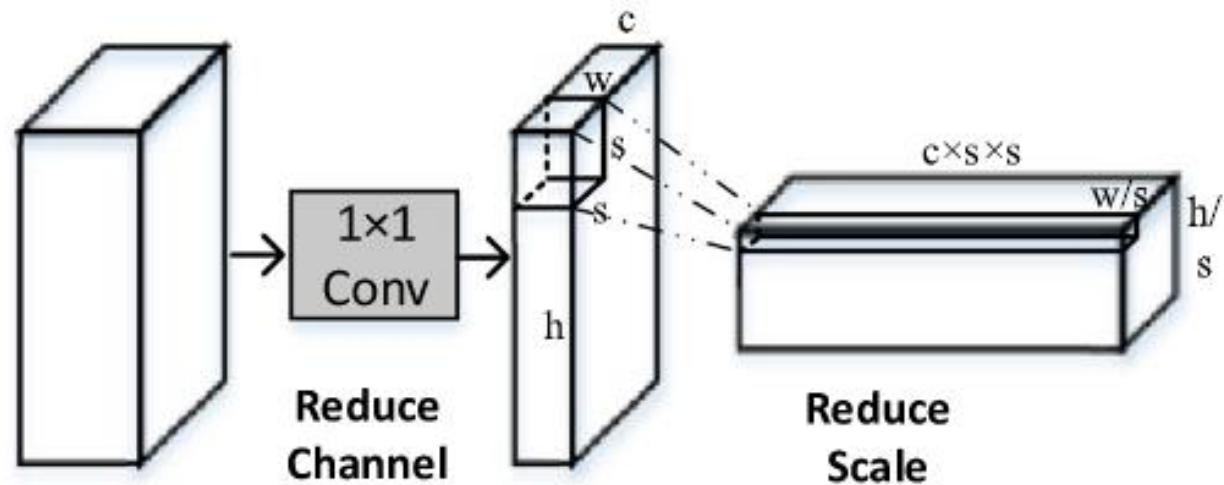
骨干网 w^m 的参数学习使交叉熵损失 L_c 最大化, 而学习鉴别器 θ^m 学习最小化损失函数



研究方法 (MAF)

尺度缩减模块(SRM)

大特征图可能会产生太多的像素级样本，对训练不友好。该模块在不丢失信息的情况下缩小特征映射，提高了训练效率。

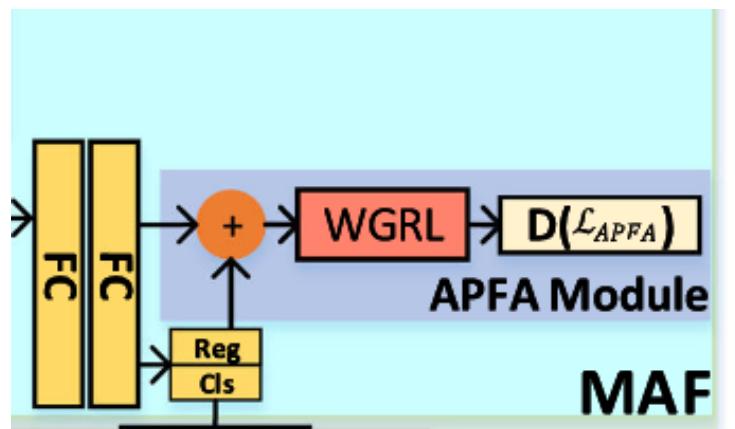


1x1卷积减小通道数，特征图中的 $s \times s$ 相邻像素，为重塑的特征图生成一个新的像素。



研究方法 (MAF)

聚合候选特征对齐 (APFA) - 实现实例级语义对齐



APFA模块的损失函数可以写成：

$$\min_{\theta_p} \max_{\mathbf{w}} \mathcal{L}_{APFA}$$

$$\mathcal{L}_{APFA} = \sum_k L_c \left(D_p(F^k(I_i, \mathbf{w}) \oplus \mathbf{c}^k \oplus \mathbf{b}^k, \hat{\mathbf{c}}_p), d_i \right) \quad F^k \text{是第 } k \text{ 个候选的候选特征}$$

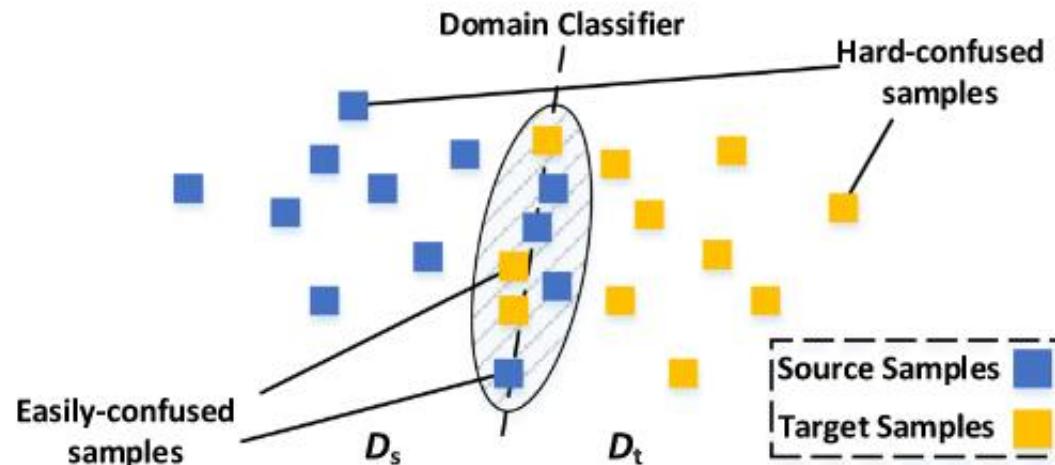
APFA 将候选特征与相应的检测结果（即边界框回归坐标和分类分数）连接起来。这样有两种优势：

- 分类和边界框回归结果分别丰富了类别和位置的语言信息。
- 分类和边界框回归结果提高了特征的判别，以有效地训练判别器。



研究方法 (MAF)

加权梯度反转层 (WGRL) - 放过易混淆的样本，同时惩罚难混淆的样本



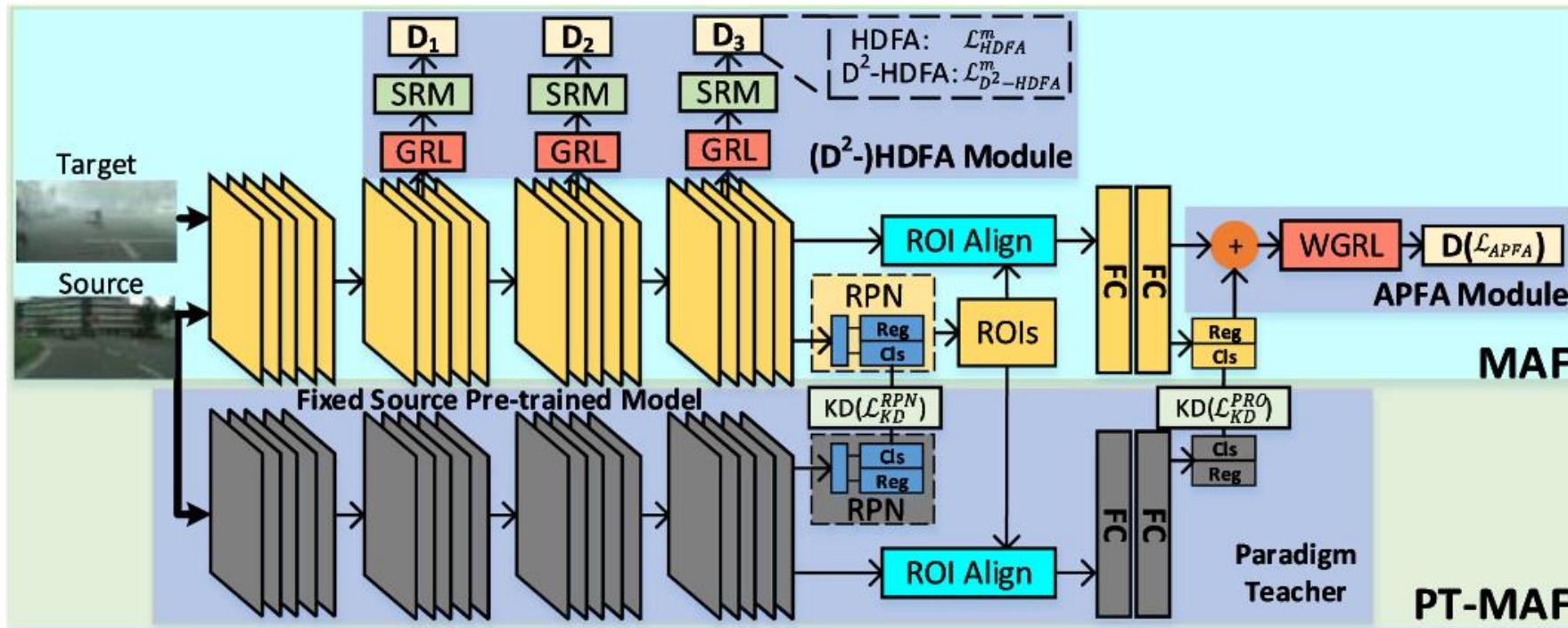
实现高分类分数的样本远离决策边界，因此被视为硬混淆样本。相反，分类分数较低的决策边界附近的样本很容易混淆的样本。

$$G_{wgrl} = -\lambda(d \cdot p + (1-d)(1-p))G \quad \text{候选属于源域的概率为 } p, \text{ 来自目标域的概率为 } 1-p$$

判别器的预测分数用作梯度的权重，鉴别器的置信度越高意味着样本被自动加权，从而可以进一步改进域适应。



研究方法 (PT-MAF)

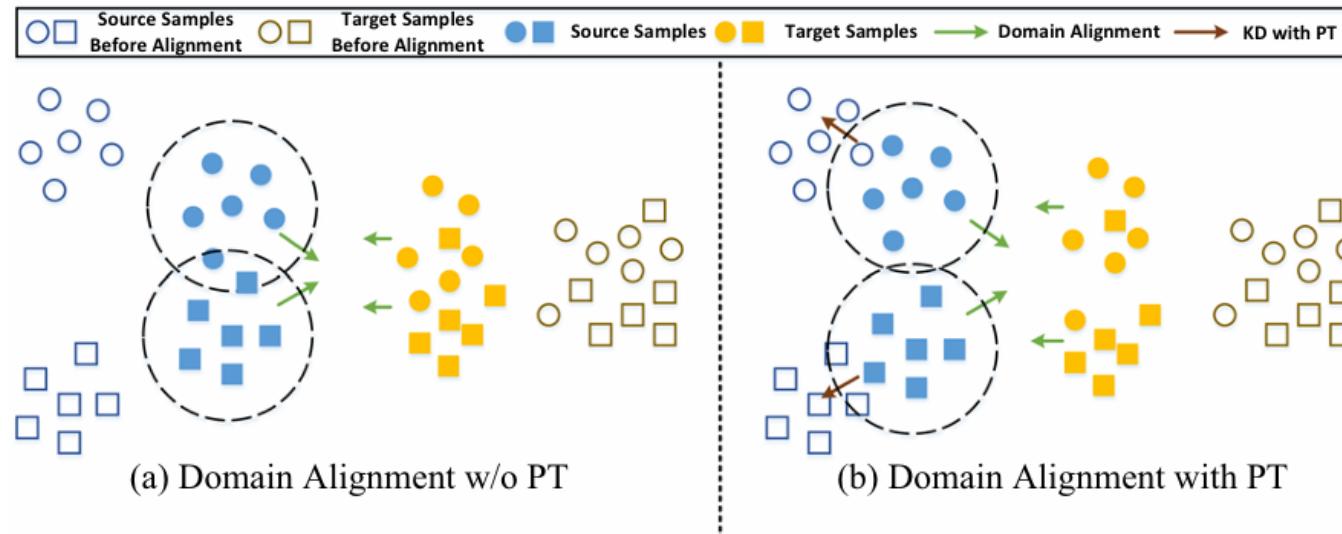


- PT 部分固定权重旨在保留源分布，通过知识蒸馏，MAF 部分从预训练模型中学习源分布，从而可以缓解源崩溃问题并实现更好的适应性。
- 双鉴别器HDFA (D^2 -HDFA)分别对齐前景和背景，缓解中性样本的影响，帮助模型实现更好的特征判别。



研究方法 (PT-MAF)

知识蒸馏 - 缓解源误差崩溃

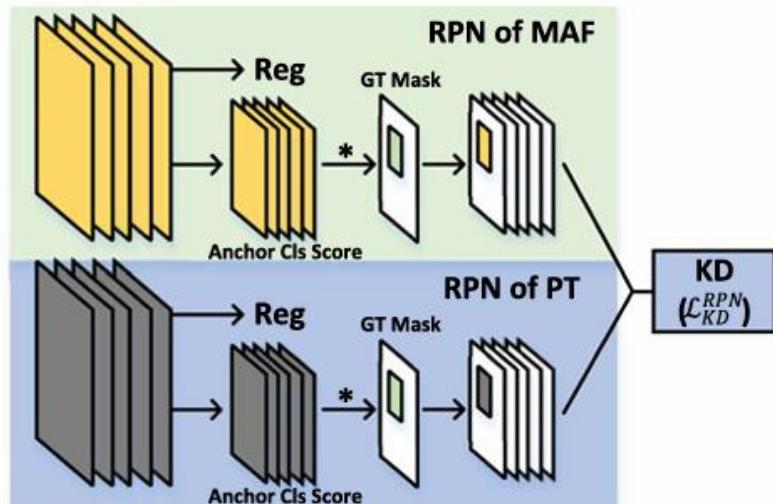


- 图a所示，MAF在域对齐过程中，由于目标标签稀缺，源分布容易偏向不准确的目标分布，使得源误差增大，模型的适应性不能被保留
- 图b所示，通过知识蒸馏，MAF部分从预训练模型中学习源分布。PT部分旨在保留源分布，从而可以缓解源崩溃问题并实现更好的适应性。

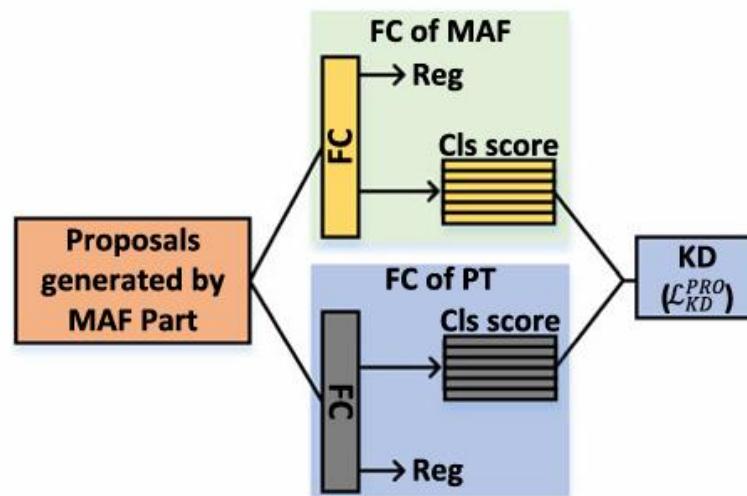


研究方法 (PT-MAF)

知识蒸馏 (RPN和候选分类器) - 缓解源误差崩溃



(a) KD for RPN



(b) KD for Proposal Classifier

- 在RPN中，只有位于ground-truth 框中的前景像素对应锚框分类分数进行知识蒸馏
- 在候选分类器中，候选特征的分类分数进行知识蒸馏



研究方法 (PT-MAF)

知识蒸馏 (RPN和候选分类器) - 缓解源误差崩溃

- 在RPN中进行知识蒸馏

$$p_{al}^{(u,v)} = \frac{\exp(q_{al}^{(u,v)}/T)}{\sum_{l=1}^2 \exp(q_{al}^{(u,v)}/T)}$$

$p_{al}^{(u,v)}$ 软化的每类预测。a 代表第 a 个锚点，而 l 是类别索引（前景或背景）。T 是软化温度参数， $q_{al}^{(u,v)}$ 是第 a 个锚点位置 (u, v) 处的 RPN 分类输出。

- 在候选分类器中进行知识蒸馏

$$p_j^k = \frac{\exp(q_j^k/T)}{\sum_{j=1}^C \exp(q_j^k/T)}$$

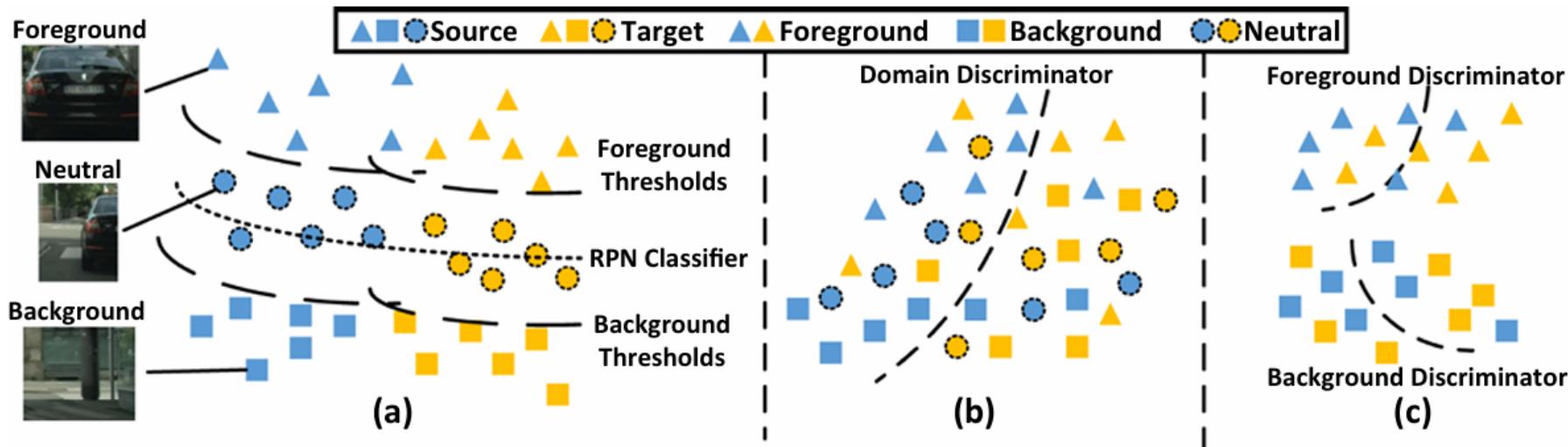
p_j^k 软化的每类输出， q_j^k 是第 k 个候选，属于类别 j 的分类器输出

蒸馏的损失函数选择KL散度进行计算



研究方法 (PT-MAF)

双鉴别器HDFA (D²-HDFA) - 聚焦于前景和背景样本的域对齐



- 图a显示了原始分布，D² -HDFA引入了两个分数阈值(λ_f, λ_b)来过滤中性样本并选择前景和背景。
- 图b中，使用单个域鉴别器，各种样本相互混淆。
- 图c中，得益于双鉴别器，D²-HDFA模块分别对齐前景和背景样本，样本被收集到两个集群中。



研究方法 (PT-MAF)

MAF与PT-MAF损失对比

域对齐损失

MAF

$$\mathcal{L}_{DA} = \mathcal{L}_{APFA} + \sum_{m=3}^5 \mathcal{L}_{HDFA}^m$$

总损失

$$\mathcal{L}_{MAF} = \mathcal{L}_{DET} + \alpha \mathcal{L}_{DA}$$

PT-MAF

$$\mathcal{L}'_{DA} = \mathcal{L}_{APFA} + \sum_{m=3}^5 \mathcal{L}_{D^2-HDFA}^m$$

$$\mathcal{L}_{PT-MAF} = \mathcal{L}_{DET} + \alpha \mathcal{L}'_{DA} + \beta \mathcal{L}_{PT}$$

蒸馏损失

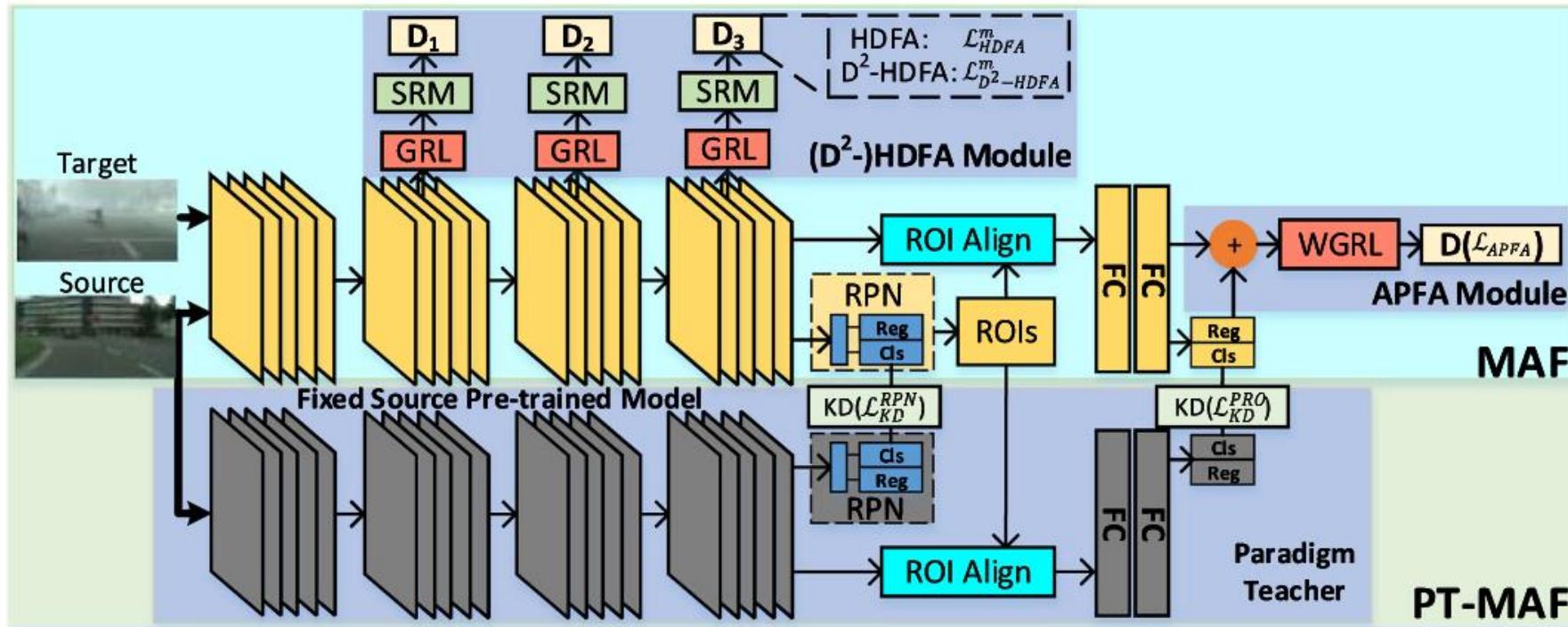
$$\mathcal{L}_{PT} = \mathcal{L}_{KD}^{RPN} + \mathcal{L}_{KD}^{PRO}$$





研究方法 (PT-MAF)

PT-MAF损失对比训练流程



首先，训练一个带有源标记数据的检测器。然后，将预训练模型的参数加载为范式教师，从头开始训练 PT-MAF 模型。最后计算损失（检测损失，域对齐损失，蒸馏损失），获取梯度更新网络参数。



实验与分析



实验设置与数据集选用

- | 设置：源域对象用边界框和类别进行标注，目标域完全不标记。
- | 性能评估：计算目标域上的平均精度均值（mAP）。
- | 预训练：用ImageNet（图像数据库）来预训练网络主干（VGG16或ResNet101），用PT-MAF模型的源训练集预训练 paradigm teacher。

数据集

Cityscapes (城市景观)



Foggy Cityscapes: 模拟有雾天气的图像组成的有雾场景





实验设置与数据集选用

KITTI: 基于自动驾驶平台生成的数据集，用于目标检测任务。



SIM10K: 由《侠盗猎车手V》的引擎生成的模拟数据集，全部图像都被用作我们实验中的源域。使用汽车类别进行训练和测试。



Pascal VOC

VOC09和VOC12合并用于训练，包含数千张手工标注的图像和数十个不同的类别。作为源域数据集。

Clipart and Watercolor

两个数据集作为目标域，Clipart数据集包含1000张图像，20个类别，与Pascal VOC相同。Watercolor有2000张图像，6个类别。对于两个数据集，一半的图像用于训练，而另一半用于测试。



实验——不同天气条件下的目标检测

从Cityscapes到Foggy Cityscapes的检测：

Cityscapes作为源域，Foggy Cityscapes作为目标域；

模型用11个训练周期训练，到第8个时学习率从0.01下降到0.001。模型权重衰减设置为0.0001。 D^2 -HDFA引入的 λ_b 和 λ_f 设置为0.1和0.7。

Table 1 The cross-domain object detection results from Cityscapes to Foggy Cityscapes

Methods	Person	Rider	Car	Truck	Bus	Train	mcycle	bcycle	mAP
SSD	16.5	20.5	26.7	9.4	16.1	9.1	11.4	18.6	14.2
Faster-RCNN	24.1	33.1	34.3	4.1	22.3	3.0	15.3	26.5	20.3
DAF (Chen et al., 2018)	25.0	31.0	40.5	22.1	35.3	20.2	20.0	27.1	27.6
SW (Saito et al., 2019)	29.9	42.3	43.5	24.5	36.2	32.6	30.0	35.3	34.3
Selective Alignment (Zhu et al., 2019)	33.5	38.0	48.5	26.5	39.0	23.3	28.0	33.6	33.8
D&match (Kim et al., 2019b)	30.8	40.5	44.3	27.2	38.4	34.5	28.4	32.2	34.6
MTOR (Cai et al., 2019)	30.6	41.4	44.0	21.9	38.6	40.6	28.3	35.6	35.1
NL /w res101 (Khodabandeh et al., 2019)	35.1	42.2	49.2	30.1	45.3	27.0	26.9	36.0	36.5
CDN (Su et al., 2020)	35.8	45.7	50.9	30.1	42.5	29.8	30.8	36.5	36.6
SCL (Shen et al., 2019)	31.6	44.0	44.8	30.4	41.8	40.7	33.6	36.2	37.9
ATF (He & Zhang, 2020)	34.6	47.0	50.0	23.7	43.3	38.7	33.4	38.8	38.7
GPA (Xu et al., 2020b)	32.9	46.7	54.1	24.7	45.7	41.1	32.4	38.7	39.5
MAF(SSD)	19.2	23.6	28.9	13.9	23.9	11.1	13.6	21.9	17.6
PT-MAF(SSD)	20.1	23.7	35.0	18.0	34.0	17.5	13.4	22.9	20.6
MAF	32.9	44.4	43.0	22.8	36.9	34.2	32.8	36.4	35.4
PT-MAF(MAF+PT+ D^2)	37.3	49.4	52.2	26.7	49.5	34.5	34.9	41.2	40.7

The bold numbers mean the best performance

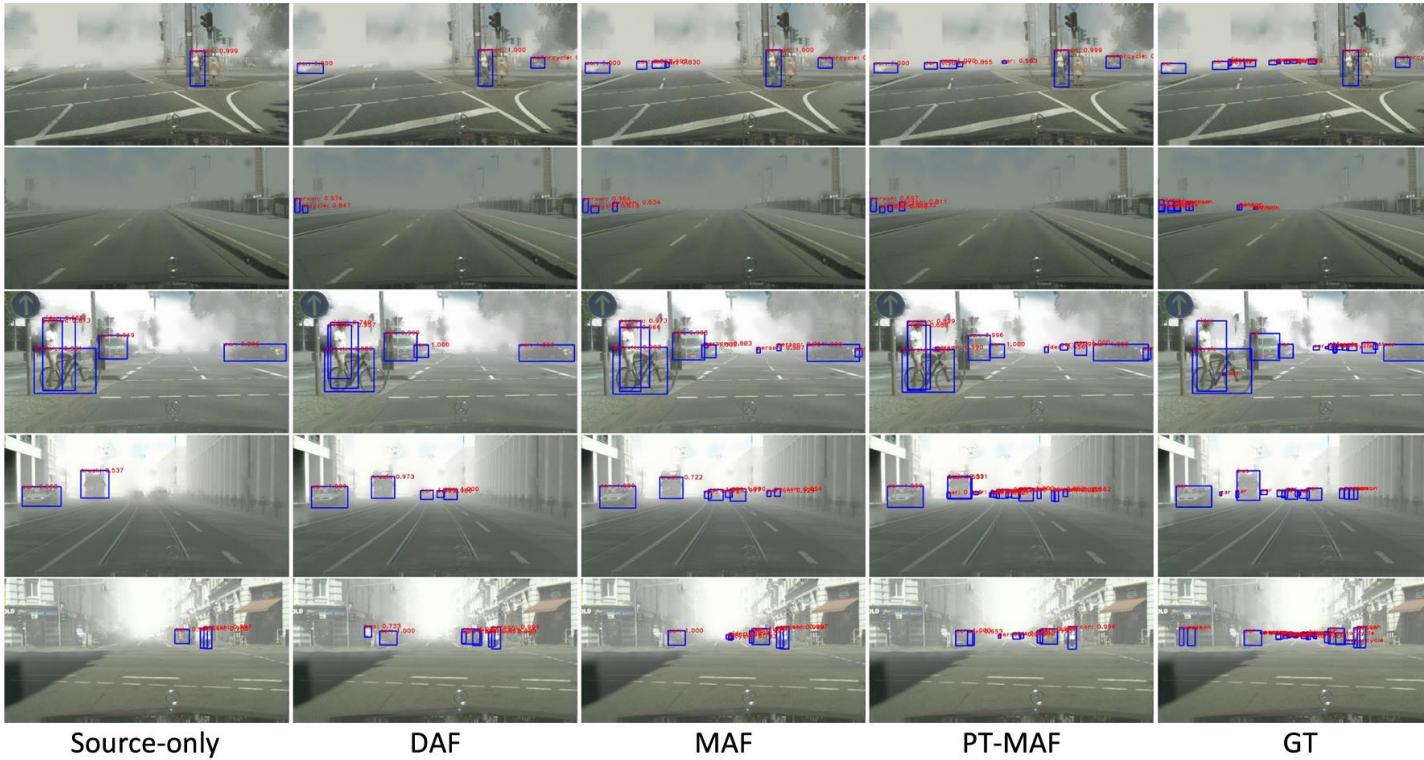
PT-MAF在两个方面提高了性能：

- (1) PT 减少了因不准确的目标分布而引起的源域崩溃。
- (2) D^2 -HDFA 模块增强了不同域之间前景和背景样本的特征对齐。



实验——不同天气条件下的目标检测

Foggy Cityscapes的可视化检测结果： 样本来自Foggy Cityscapes的验证集



Source-only 模型表现最差；

PT-MAF 模型表现最佳。

Fig. 10 Visualized detection results on Foggy Cityscapes by testing source only model, DAF (Chen et al., 2018), MAF, and PT-MAF in each column, resp. The final column shows the ground truth bounding boxes and categories. From the visualization, PT-MAF achieves the best performance



实验——域偏移较大的跨域目标检测

从Pascal VOC到Clipart的检测：

Table 2 The cross-domain detection results from Pascal VOC to Clipart

Methods	Aero	Bike	Bird	Boat	Bottle	Bus	Car	Cat	Chair	Cow	mAP
Faster-RCNN	35.6	52.5	24.3	23.0	20.0	43.9	32.8	10.7	30.6	11.7	
DAF (Chen et al., 2018)	15.0	34.6	12.4	11.9	19.8	21.1	23.2	3.1	22.1	26.3	
BDC-Faster	20.2	46.4	20.4	19.3	18.7	41.3	26.5	6.4	33.2	11.7	
WST-BSR (Kim et al., 2019a)	28.0	64.5	23.9	19.0	21.9	64.3	43.5	16.4	42.2	25.9	
SW (Saito et al., 2019)	26.2	48.5	32.6	33.7	38.5	54.3	37.1	18.6	34.8	58.3	
HTCN (Chen et al., 2020)	33.6	58.9	34.0	23.4	45.6	57.0	39.8	12.0	39.7	51.3	
SCL (Shen et al., 2019)	44.7	50.0	33.6	27.4	42.2	55.6	38.3	19.2	37.9	69.0	
MEAA (Nguyen et al., 2020)	31.3	53.5	38.0	17.8	38.5	69.9	38.2	23.8	38.3	58.1	
MAF	30.1	61.3	30.3	38.7	43.6	44.7	41.9	2.6	39.1	57.3	
PT-MAF	31.1	64.7	32.8	35.6	46.9	36.8	40.8	12.1	42.3	74.0	
	Table	Dog	Horse	Motor	Person	Plant	Sheep	Sofa	Train	tv	mAP
Faster-RCNN	13.8	6.0	36.8	45.9	48.7	41.9	16.5	7.3	22.9	32.0	27.8
DAF (Chen et al., 2018)	10.6	10.0	19.6	39.4	34.6	29.3	1.0	17.1	19.7	24.8	19.8
BDC-Faster	26.0	1.7	36.6	41.5	37.7	44.5	10.6	20.4	33.3	15.5	25.6
WST-BSR (Kim et al., 2019a)	30.5	7.9	25.5	67.6	54.5	36.4	10.3	31.2	57.4	43.5	35.7
SW (Saito et al., 2019)	17.0	12.5	33.8	65.5	61.6	52.0	9.3	24.9	54.1	49.1	38.1
HTCN (Chen et al., 2020)	21.1	20.0	39.1	72.8	63.0	43.1	19.3	30.1	50.2	51.8	40.3
SCL (Shen et al., 2019)	30.1	26.3	34.4	67.3	61.0	47.9	21.4	26.3	50.1	47.3	41.5
MEAA (Nguyen et al., 2020)	14.6	18.1	33.8	88.1	60.3	43.1	19.3	30.1	50.2	51.8	41.1
MAF	30.1	5.9	45.0	55.6	58.5	51.0	15.7	24.5	58.3	31.5	38.5
PT-MAF	37.7	13.1	35.2	61.9	62.8	57.1	51.5	17.4	44.4	42.3	42.0

Pascal VOC 作为源域,

Clipart 作为目标域;

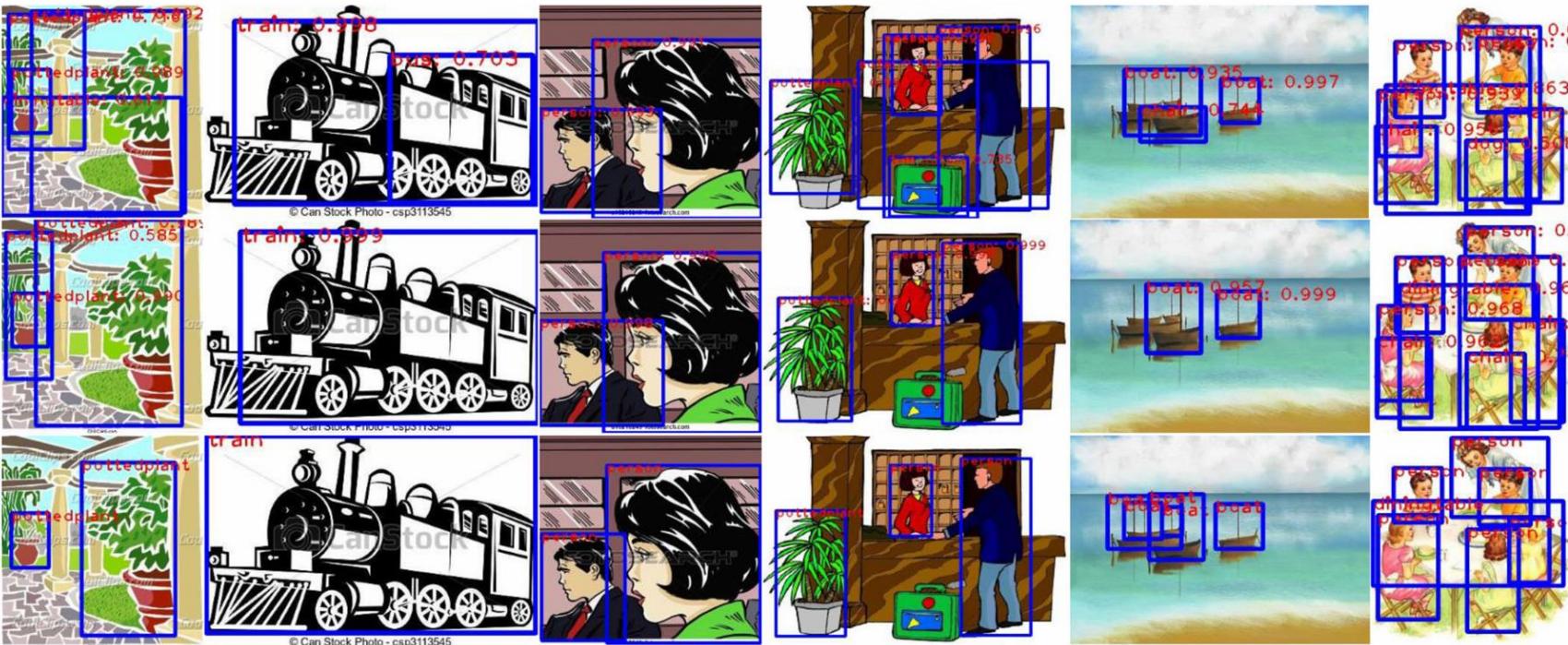
主干网络设置为ResNet101。

The bold numbers mean the best performance



实验——域偏移较大的跨域目标检测

Clipart检测的可视化结果：



MAF 和 PT-MAF 都
展现出优秀的性能。

Fig. 11 The visualized detection results of Clipart. The detection results of MAF and PT-MAF are presented in the first and second rows, respectively. For comparison, we show the ground truth in the last row



实验——域偏移较大的跨域目标检测

从Pascal VOC到Watercolor的检测：

Pascal VOC 作为源域，Watercolor作为目标域，包含六种不同类别

Table 3 The cross-domain detection results from Pascal VOC to Watercolor

Methods	Bike	Bird	Car	Cat	Dog	Person	mAP
Faster-RCNN	68.8	46.8	37.2	32.7	21.3	60.7	44.6
DAF (Chen et al., 2018)	75.2	40.6	48.0	31.5	20.6	60.0	46.0
BDC-Faster	68.6	48.3	47.2	26.5	21.7	60.5	45.5
WST-BSR (Kim et al., 2019a)	75.6	45.8	49.3	34.1	30.3	64.1	49.9
SW (Saito et al., 2019)	82.3	55.9	46.5	32.7	35.5	66.7	53.3
MAF	73.4	55.7	46.4	36.8	28.9	60.8	50.3
PT-MAF	80.8	52.2	48.7	37.5	37.7	64.9	53.6

The bold numbers mean the best performance



实验——不同视点下的目标检测

KITTI 和 Cityscapes 数据集交替用作源域和目标域，模拟不同类型的相机设备或视点引起的域偏移。

只有 Car 类别参与训练和测试，用平均精度 (AP) 评估

Table 4 The cross-domain object detection on the Cityscapes and KITTI

	K → C	C → K
Faster-RCNN	30.2	53.5
DAF (Chen et al., 2018)	38.5	64.1
SW (Saito et al., 2019)	37.9	71.0
MAF	41.0	72.1
PT-MAF	42.3	72.8

MAF和PT-MAF都具有优越的性能。得益于 D²-HDFA 模块，PT-MAF模型在两个任务上都有更好的性能。

The bold numbers mean the best performance



实验——从合成到真实的目标检测

SIM10K作为源域，Cityscapes作为目标域，模拟从合成到真实的应用场景：

选包含 Car 类别的样本参与训练和测试，用平均精度 (AP) 评估

Table 5 The cross-domain object detection from Sim10K to Cityscapes

AP of Car	
Faster-RCNN	34.6
DAF (Chen et al., 2018)	38.9
SW (Saito et al., 2019)	40.1
SCL (Shen et al., 2019)	42.6
MEAA (Nguyen et al., 2020)	42.5
HTCN (Chen et al., 2020)	42.5
MAF	41.1
PT-MAF	43.2

MAF和PT-MAF都表现出良好的性能，其中PT-MAF在所有比较模型中取得了最好的结果。

The bold numbers mean the best performance



分析与讨论——对范式教师（PT）的分析

对SOTAs的普适性：

Table 6 Detection performance of different baseline models by deploying the Paradigm Teacher (PT)

Methods	Person	Rider	Car	Truck	Bus	Train	mcycle	bcycle	mAP
DAF (Chen et al., 2018)	25.0	31.0	40.5	22.1	35.3	20.2	20.0	27.1	27.6
DAF+PT	27.0	31.7	37.5	29.4	40.9	27.5	16.0	26.3	29.5
SW (Saito et al., 2019)	29.9	42.3	43.5	24.5	36.2	32.6	30.0	35.3	34.3
SW+PT	33.4	45.5	46.7	25.0	42.7	31.2	29.4	37.9	36.5
MAF	32.9	44.4	43.0	22.8	36.9	34.2	32.8	36.4	35.4
MAF+PT	34.4	47.5	44.7	26.8	42.9	40.2	32.6	38.3	38.4

The bold numbers mean the best performance

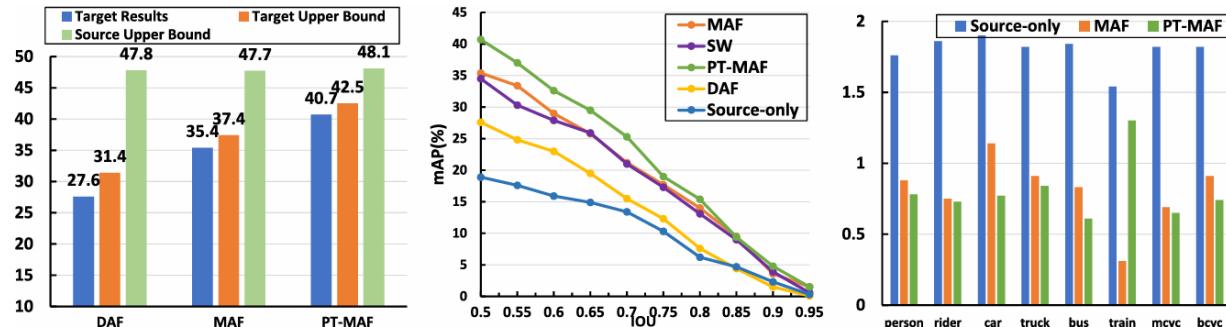
基线模型都提高了大约2%，减轻了源域崩溃问题；

提高了模型的适应性，在目标域预测性能都取得了改进；

证明了所提出的范式教师部分的有效性。



分析与讨论——对范式教师（PT）的分析



(a) The upper bound performance across domains

(b) A-distance

(c) IOU v.s mAP

Fig. 12 Algorithmic analysis. **(a)** shows the upper bound of both source and target domains. **(b)** presents the A-distance across domains w.r.t. different categories. **(c)** is the mAP with different IOU thresholds

目标上限性能分析（图12(a)）：

PT-MAF 可以有效地降低源域崩溃的风险，并在源域和目标域上都表现出最好的上限。

对IOU值的影响（图12(b)）：

所有模型的 mAP 都随着 IOU 阈值的增加而退化，但我们的 PT-MAF (绿线) 在所有比较模型中仍然在 IOU 阈值上实现了最好的性能。



分析与讨论——对范式教师（PT）的分析

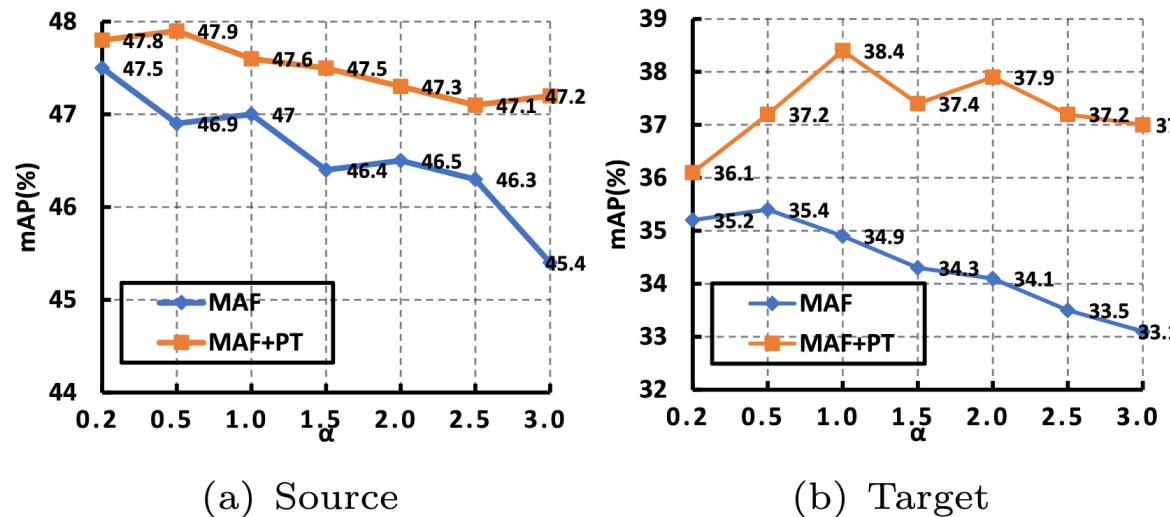


Fig. 13 Comparison of the model with or without PT. We present the performances of both MAF and MAF with the PT stream (MAF+PT). We increase the α in Eq. 12 to search the relation of the source collapse problem and adversarial learning

对抗性学习和源域崩溃的关系：

(1) α 越大，对于MAF，由于源域崩溃，性能不能沿着增加；MAF+PT随着 α 的增加而提高检测性能。

(2) α 过大时，两者都受到源域崩溃影响，但 PT-MAF 受影响程度明显低于 MAF。

表明：PT 模块可以帮助模型缓解源域崩溃问题，它可以同时考虑域差异和域适应性。



分析与讨论——对H DFA 和 D²-H DFA 模块的分析

判别器的分类得分：

$$s_{map} = 1 - 2s_d - 0.5$$

- (1) MAF 对图像的大部分特征进行了对齐，验证了 HDFA 模块的有效性。
- (2) D²-H DFA 模块中的前景和背景判别器实现了不同的侧重点。
- (3) D²-H DFA 实现了更好的前景对齐，PT-MAF 实现了更精确的跨域分布对齐。

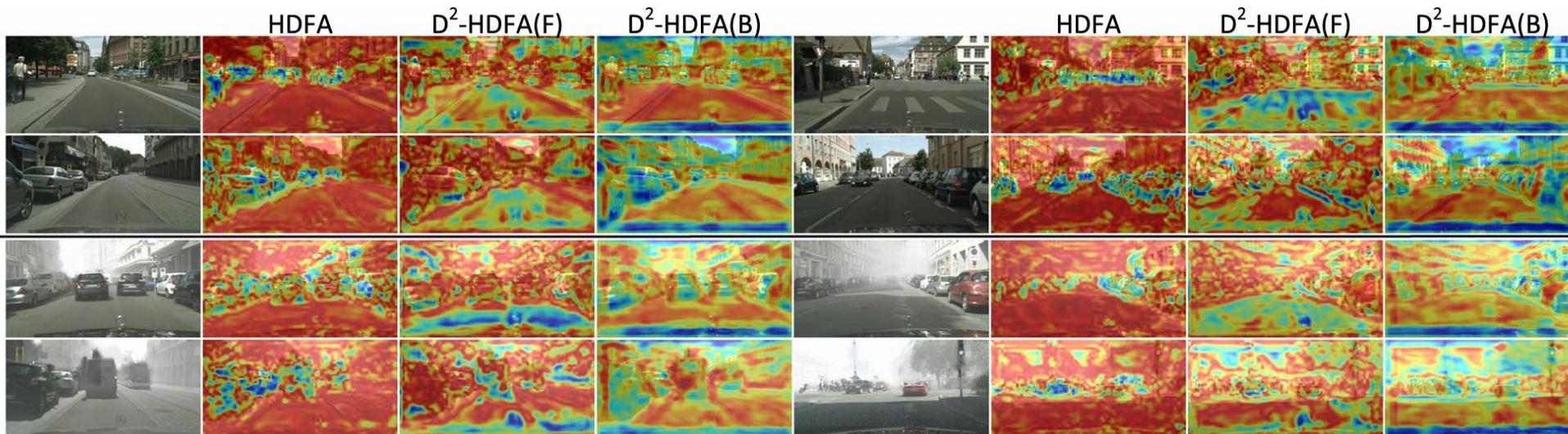


Fig. 14 The heat maps of the discriminator scores of the HDFA and D²-H DFA modules, resp. In the heat maps, the red pixel is recognized to be aligned and its domain discriminator score is close to 0.5. The blue pixel is misaligned and its domain classification score is close to 0 or 1. The first two rows represent the source domain (Cityscapes),

while the last two rows present the results of the target domain (Foggy Cityscapes). The D²-H DFA(F) and D²-H DFA(B) stand for the discriminator scores of foreground and background, respectively. Best viewed in color (Color figure online)



分析与讨论——对HDFA 和 D²-HDFA 模块的分析

不同块上的判别器对模型性能的影响：

所有的块都影响模型的性能，第5个块的域划分对模型的性能影响最大。

实验证明了域对齐在模型层次上的重要性。

Table 7 Analysis of discriminators from the HDFA and D²-HDFA modules

	w/o b3	w/o b4	w/o b5	All bs
MAF	37.9	36.2	33.9	38.4
PT-MAF	39.1	38.9	37.9	40.7



分析与讨论——参数分析

PT的参数分析（表8）：

Table 8 The parameter analysis of T and β

T	2	3	4	5
β				
0.5	39.9	39.6	39.5	38.6
1.0	40.0	40.7	40.1	38.9
1.5	40.2	39.9	39.7	37.1
2.0	39.3	39.8	38.8	38.3

The bold numbers mean the best performance

D²-HDFA的参数分析（表9）：

Table 9 The parameter analysis of λ_b and λ_f

λ_f	0.5	0.7	0.9
λ_b			
0.1	39.7	40.7	40.5
0.3	38.7	39.3	38.4
0.5	38.0	38.9	38.2

The bold numbers mean the best performance

当 $\lambda_b = 0.5$ 和 $\lambda_f = 0.5$ 时，D²-HDFA 模块没有过滤掉任何用于域对齐的样本，在这种情况下模型的性能最差，这证明了通过两个动态阈值来筛选中性样本的有效性。



分析与讨论——特征分析

用t-SNE图对候选框特征进行可视化：

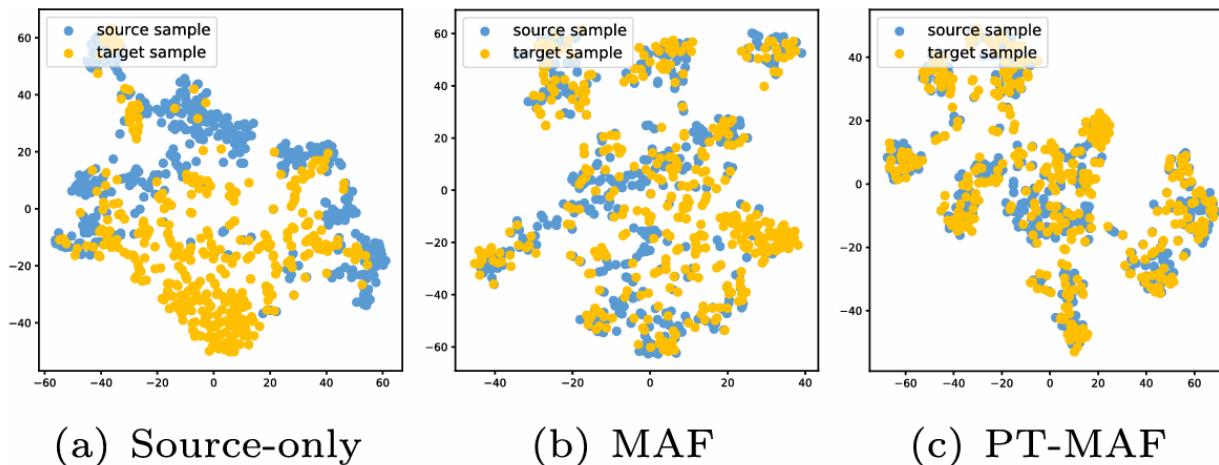


Fig. 15 The t-SNE feature visualization of our models. We present the results of source-only trained Faster-RCNN, MAF, and PT-MAF in the figure. The samples from the source and target domains are shown with blue and yellow colors, respectively (Color figure online)

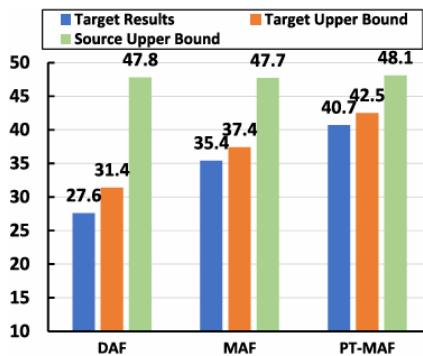
- (1) 从 source-only 模型中提取的源域和目标域特征清晰地分为两部分，即跨域的域差异很大。
- (2) MAF 使用 HDFA 和 APFA 对齐模块，来自两个域的样本都很好地对齐。
- (3) PT-MAF 显示了最好的域混淆可视化结果。

实验证明了所提出的 PT 和 D²-HDFA 模块能够很好地提高域适应性。

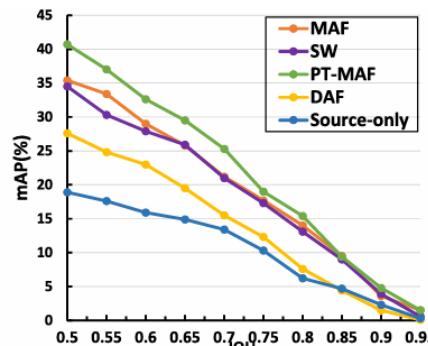


分析与讨论——特征分析

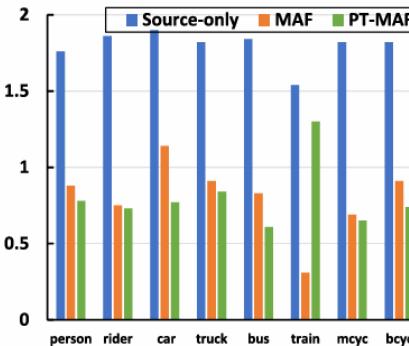
所有类别的 \mathcal{A} – distance (图12(c)) :



(a) The upper bound performance across domains



(b) A-distance across domains



(c) IOU v.s mAP

$$d_{\mathcal{A}} = 2(1 - \varepsilon)$$

- (1) 所提出的MAF实现了更小的 \mathcal{A} – distance。
- (2) 由于 D²-HDFA 关注边缘分布的域对齐, PT-MAF 在除 train 类别外的所有实验中实现了最小的 \mathcal{A} – distance。

Fig. 12 Algorithmic analysis. (a) shows the upper bound of both source and target domains. (b) presents the A-distance across domains w.r.t. different categories. (c) is the mAP with different IOU thresholds



分析与讨论——消融研究

Table 10 Ablation study of MAF and PT-MAF

	WGRL	APFA	HDFA	D ²	PT	mAP(%)
DAF (Chen et al., 2018)						27.8
MAF	✓					30.2
		✓				30.3
	✓	✓				30.9
	✓	✓	✓			35.4
PT-MAF	✓	✓	✓	✓		38.4
	✓	✓	✓		✓	38.8
	✓	✓	✓	✓	✓	40.7

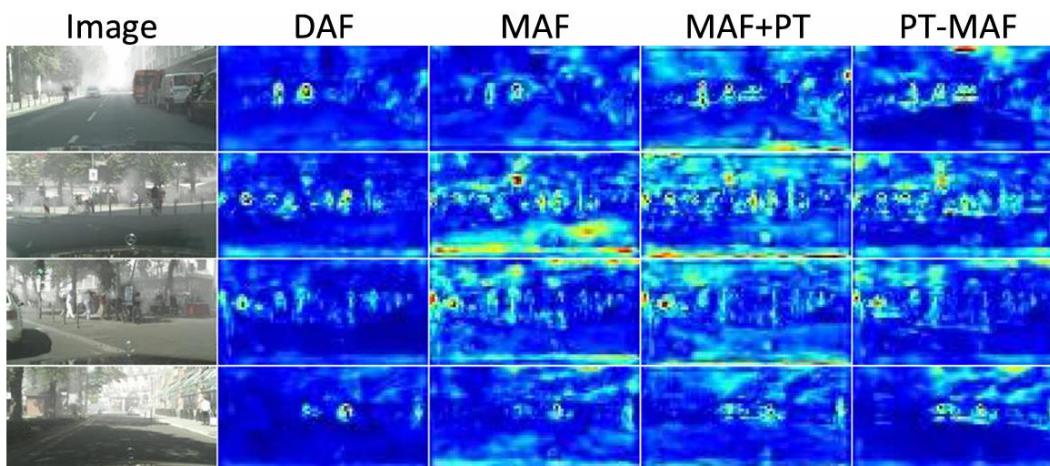


Fig. 16 Comparison of target feature maps. The red pixels in feature maps mean the highest activations, while the blue ones mean the smallest activations (Color figure online)

- (1) 所有模型的特征图在前景部分都被激活。
- (2) DAF 的特征图没有过多激活，这可能是因为特征没有完全对齐。
- (3) MAF 和 MAF+PT 在特征图中有更多激活部分，但一些背景部分也被激活了。
- (4) PT-MAF 取得了最佳性能，其中前景被激活而背景没有被激活。



总结



总结

● 01

提出了一种用于跨域目标检测的多对抗 Faster-RCNN (MAF) , 包括两个模块：HDFA 模块在图像级别表征了域不变特征，APFA 模块在实例级别对难以区分的提议特征进行了对齐。

● 02

为了提高域适应性和模型稳定性，进一步提出了范式教师 MAF (PT-MAF) , 具有两个新贡献：

- (1) 引入范式教师 (PT) 来缓解源域风险崩溃并提高适应性模型的稳定性。
- (2) 引入双鉴别器 HDFA (D^2 -HDFA) 模块，用于更精细地对齐前景和背景，其中中性样本 (难样本) 通过动态阈值进行筛选。

● 03

- (1) 对于特征对齐， D^2 -HDFA 和 APFA 只考虑了分类，这可能是不足的。
- (2) 对于 PT 部分，尽管性能有所提高，但原始源分布是否适合目标域仍然不确定。源域和目标域分布之间的关系值得研究。
- (3) PT-MAF只考虑了两个不同的域。对于实际应用，应面对更多未知的域或环境。



谢谢！