

跨模态渐进式知识迁移SAR目标检测技术

赵国威, 蒋嘉庆, 董刚刚*

(西安电子科技大学雷达信号处理全国重点实验室, 陕西西安 710071)

摘要: 随着光/SAR传感器的投入使用, 如何挖掘异源数据信息、提高遥感图像解译与信息获取效率, 是当前亟待解决的科学问题. 对此, 论文提出了一种基于光/SAR(Synthetic Aperture Radar)图像渐进式知识迁移的跨模态目标检测方法(Progressive Crossmodal Knowledge Transfer, PCKT), 充分挖掘可见光图像丰富的纹理细节信息, 改善SAR图像目标检测性能. 第一, 针对可见光遥感与SAR截然不同的成像机理导致模态差异大、信息融合难的问题, 提出从可见光图像到SAR图像的跨域生成技术, 利用可见光图像语义信息引导构建伪SAR中间域, 实现光学辐射特征与SAR散射特性的语义融合; 第二, 针对可见光与SAR图像之间存在的语义鸿沟、跨模态学习效果差的问题, 设计光/SAR多尺度特征对齐学习策略, 实现可见光源域-伪SAR中间域语义特征、伪SAR中间域-SAR目标域散射特性分布的高效对齐, 形成分阶段的跨模态知识迁移学习框架; 第三, 针对光/SAR模态离群样本对跨模态学习的不利影响, 提出基于质量感知的动态权重分配机制, 根据域分类器置信度衡量动态量化合成样本的可靠性, 在训练过程中优先学习高置信度样本, 从而抑制离群中间域图像产生的不利影响. 最后, 利用SpaceNet6、SSDD(SAR Ship Detection Dataset)、HRSID(High-Resolution SAR Images Dataset)等典型异源数据进行了大量跨模态学习实验验证, 结果表明所提方法能充分挖掘可见光图像的重要信息, 改善SAR图像目标检测效果, 与源域无迁移方法相比, 论文所提出方法在SpaceNet同源知识迁移结果提升平均检测精度为21.5个百分点, 与次优基线方法相比平均检测精度提升为3.3个百分点, 很好地验证了跨模态知识迁移在SAR目标检测任务中的可行性.

关键词: 合成孔径雷达; 目标检测; 知识迁移; 对抗学习; 多尺度对齐; 伪标签学习

基金项目: 国家自然科学基金(No.61971324, No.62525105)

中图分类号: TN95 **文献标识码:** A **文章编号:** 0372-2112(2025)09-3147-16

电子学报URL: <http://www.ejournal.org.cn> **DOI:** 10.12263/DZXB.20250417

Cross-Modal SAR Target Detection via Progressive Knowledge Transfer

ZHAO Guo-wei, JIANG Jia-qing, DONG Gang-gang*

(National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an, Shaanxi 710071, China)

Abstract: The wide application of electro-optical sensors presented the urgent need of cross-modal learning between optical and SAR image. In this paper, a new cross-modal synthetic aperture radar (SAR) target detection method via progressive knowledge transfer was proposed. First, a new generative technique from the optical image to SAR image was presented. The immediate domain composed of the generated pseudo SAR images can be formed accordingly. The semantic discrepancies between SAR backscattering imaging mechanism and the passive optical radiation imagery can be bridged. The optical radiation features with SAR scattering characteristics can be fused effectively. Second, a dual-stage domain adaptation strategy composed of the multi-scale feature alignment skill was presented. The semantic components between the optical source domain and the intermediate domain can be aligned through the multi-scale feature learning trick initially. The scattering distribution alignment between the intermediate domain and the SAR target domain can be then achieved. Third, a quality-aware dynamic weighting strategy was presented to mitigate the impact of outlier samples in the intermediate domain. It was capable of adjusting the contributions of synthetic data based on confidence metrics dynamically. Finally, multiple rounds of experiments were pursued on SpaceNet6, SSDD (SAR Ship Detection Dataset), and HRSID (High-Resolution SAR Images Dataset) datasets. The results proved the advantages of proposed method. The improvement of 21.5 per-

centage points was achieved compared to the source-only learning method. Likewise, the improvement of 3.3 percentage points was achieved in comparison to the state-of-the-art. These results confirm the viability of electro-optical-to-SAR knowledge transfer for enhancing cross-modal target detection.

Key words: synthetic aperture radar; target detection; knowledge transfer; adversarial learning; multi-scale alignment; pseudo-label learning

Foundation Item(s): National Natural Science Foundation of China (No.61971324, No.62525105)

1 引言

随着大量可见光/SAR(Synthetic Aperture Radar)传感器的投入使用,如何挖掘异源传感器的多模态数据,是当前亟待解决的科学问题^[1].与红外、可见光等传感手段相比,合成孔径雷达能够全天时、全天候对观测地域实施高分辨率成像,因此它广泛应用海域、地面、空中等典型目标的检测与识别^[2,3].随着高分辨率SAR图像数据的不断积累,对图像的自动化解译(如目标检测识别)显得十分迫切.在早期研究中,SAR目标检测主要通过恒虚警率(Constant False Alarm Rate, CFAR)技术实现^[4],如图1所示,其核心思想是利用滑动窗口估计杂波功率水平,并通过与特定杂波统计模型计算出的阈值进行比对实现目标检测.

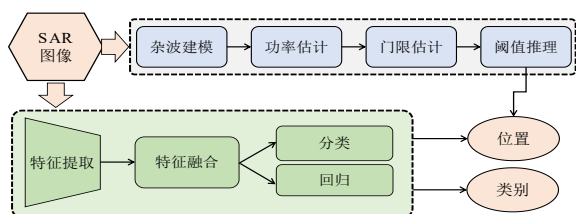


图1 基于统计建模与基于先验数据的SAR目标检测简明流程

然而,基于经验模型的检测方法高度依赖于对表征杂波分布的统计模型,在实际复杂的非均匀场景中,很难选择合适的统计分布模型对杂波进行描述,导致检测模型失配^[5-7].因此,近年来的研究更多聚焦于历史数据驱动的目标检测技术,代表性方法包括R-CNN算法族^[8,9]、YOLO(You Only Look Once)系列^[10]、SSD(Single Shot multibox Detector)算法^[11]等.例如,Lin等人^[12]提出了通道注意力机制以增强特征提取能力,改善舰船检测效率;Zhao等人^[13]设计了膨胀注意力模块以优化检测器性能;Kang等人^[14]提出结合上下文信息的SAR船舶检测方法;夏勇等人^[15]通过巴氏距离替代交并比(Intersection of Union, IoU)提升小目标检测精度;谢兆哲等人^[16]则采用矩阵特征值分解优化SAR船舶检测.尽管这些方法取得了一些进展,但其检测性能依赖于大规模的标注数据覆盖完备的样本分布空间.然而,SAR图像因独特的成像机制,成像结果遍布着固有的斑点噪声,图像信噪比低、可读性差,导致标注过

程耗时且准确性难以保证^[17,18],若在稀疏样本分布空间下进行训练,很容易导致检测模型陷入过拟合^[19,20].因此,如何挖掘异源传感器的多模态数据,提高遥感图像解译与信息获取效率就具有十分重要的研究意义.

与SAR后向散射成像机理不同,可见光遥感的成像结果“所见即所得”,数据可获取性高、可读性强且标注更便捷容易^[21].因此,能否设计从可见光图像到雷达图像的知识迁移模型,利用标注的可见光遥感数据解决无任何先验标注条件下的SAR图像目标检测,是论文拟解决的关键核心问题.相关的研究可大致分为三类^[22,23].一是基于自训练伪标签的技术路线,其核心是借助源域预训练的初始检测器,对目标域未标注数据进行预测,筛选高置信度的检测结果作为伪标签,并通过迭代训练将伪标签与源域标注数据联合优化模型.Khodabandeh等人^[24]提出了一种三阶段训练策略来处理伪标签噪声问题.Kim等人^[25]提出了一种基于伪标签的自训练策略,专注于一级目标检测模型.French等人^[26]采用教师-学生互学方法,使用教师模型生成的伪标签来训练学生模型,以解决噪声错误标签的问题.Sohn等人^[27]交替使用弱增强数据的选定伪标签,从强增强的未标注数据中学习.Li等人^[28]提出了一种称为自适应教师的教师-学生框架,将教师-学生模型与对抗性特征对齐相结合.针对低质量伪标签的噪声对检测器性能的限制,Cao等人^[29]提出了一个通用的对比均值教师框架(Contrastive Mean Teacher, CMT),将平均教师自我训练和对比学习进行协同对齐,最大限度地提高有用信息的学习效率.在此基础上,Bai等人^[30]通过设计不同的标签分配策略将领域特征融入对比学习,其中源自同一图像的域内特征,以及具有相同域标签的样本视为正对,提取细粒度的特征表示,实现对比领域均值教师学习.针对现有方法忽视在检测模型的解码器结构内的实例级特征对齐,Chen等人^[31]提出一种域自适应检测变换器,用于无监督域自适应目标检测,利用类似原型对齐弥合对象检测任务和域自适应任务之间的差距,以类感知的方式有效地对齐跨域特征.二是基于真伪对抗学习的技术方案,核心思路是引入对抗博弈的思想实现跨域生成与真伪辨别,借助梯度反转机制实现不同数据在特征空间的精准对齐.Chen等人^[32]在目标检测中引入梯度反转层以适配跨域特征.

Saito 等人^[33]将全局层级的传统二元交叉熵损失替换为焦点损失,以实现弱对齐,同时采用局部强对齐来减少源域和目标域数据之间的领域偏移. Zhang 等人^[34]在域自适应过程中对目标兴趣区域对应的区域进行了特征对齐,并选择性地对源域和目标域数据的特征进行了对齐. Belal 等人^[35]提出一种基于原型学习的均值教师策略(Prototype-based Mean Teacher, PMT),该方法使用类原型而不是特定域子网来编码特定域信息,其中的原型是通过对比损失来学习的,从而在不同域中对齐相同类别,并将不同类别尽可能分开. 三是基于异源图像转换的技术路线,其核心在于将域差异问题定位于样本层面,主要针对跨域图像在色彩分布、纹理特征及几何结构等视觉属性上的不一致性设计适配机制. Aruda 等人^[36]利用图像转换模块,专注解决夜间和白天采集数据之间的域偏移问题. Chen 等人^[37]使用图像转换模块创建“插值图像”,帮助模型填补领域之间的分布差异. Hsu 等人^[38]提出了一种基于图像级转换策略的渐进式域适应方法. Shi 等人^[39]使用 CycleGAN (Cycle-consistent Generative Adversarial Network)生成过渡域,以减少 SAR 图像的外观差异. 然后将源域和过渡域结合形成一个新的源域,最终在新的源域和目标域之间进行领域适应. Du 等人^[40]提出一种从白天到夜晚的零样本域自适应目标检测方法 (Day-Night Domain Adaptation, DAI-Net),首先设计反射表示学习模块,根据光照不变性策略学习图像中基于 Retinex 的光照不变量,然后引入交换重组的相干过程,通过执行两次连续的图像分解并引入重新合成相干损失来改进图像分解过程,提高在暗夜条件下的目标检测效果.

然而,现有研究多集中于光学模态之间的适配迁移. 与可见光成像相比,SAR 特殊成像机理导致图像视觉表征以几何纹理和介电特性为主导,成像结果反映了目标的纹理、结构与材质等信息,图像缺乏色彩与亮度等,二者在特征空间存在着显著的语义差异^[41-43],若将通用的视觉方案直接应用于可见光与 SAR 图像之间的知识迁移,会造成模型偏向光学域特征,忽视了 SAR 图像特有的电磁散射特性,导致跨域知识迁移效果受限^[44]. 针对这些问题与挑战,本文提出了一种基于光电模态渐进式知识迁移的 SAR 零样本目标检测技术. 首先构建伪 SAR 中间域,设计光/SAR 模态迁移模型,将光学图像转换为 SAR 风格的伪中间域图像,将其作为可见光模态(源域)与 SAR 模态(目标域)的桥梁. 伪 SAR 中间域图像在纹理上逼近真实 SAR 数据,同时又能保留可见光图像的结构信息,从而将复杂的跨域检测任务分解为两个子任务. 在此基础上,设计对抗学习实现源域与中间域的多尺度特征对齐,在过渡的特征空间完成中间域与目标域的特征对齐,整体架构如图 2 所

示. 由于中间域图像的生成过程无监督信息可供参考,合成结果可能存在偏离中间域分布的风险,对此进一步设计加权损失函数用于调节不同样本在域适应过程中的贡献度,其权重根据合成图像特征分布与目标域的距离动态确定,若距离目标域更近则赋予该合成样本更高权重,从而抑制异常样本的影响. 最后通过伪标签重训练目标域数据以增强特征判别性,在检测器训练阶段引入数据增强和互学习机制,抑制噪声伪标签的负面影响.

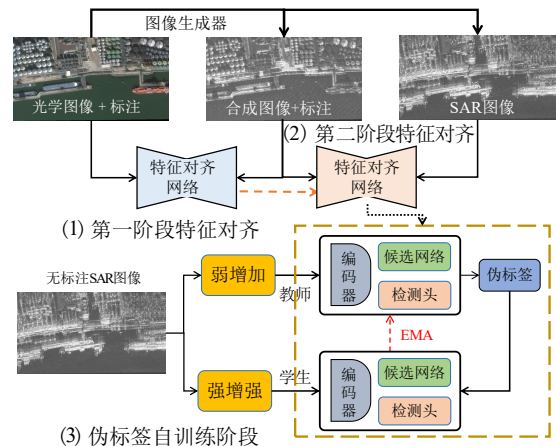


图2 跨模态渐进式知识迁移SAR目标检测算法总体流程

概括起来,本文的主要创新与贡献包括:

(1) 针对可见光与 SAR 图像域截然不同的成像机理形成的语义鸿沟,设计分阶段渐进式的跨模态知识迁移机制. 通过构建基于风格解耦的跨模态生成模型,在光学图像与 SAR 数据间建立伪 SAR 中间域,将复杂域适应任务分解为两个简化子任务,显著缩小跨模态域间差异.

(2) 针对合成图像无参考信息可能存在的域偏移、影响知识迁移效果,提出加权损失度量量化样本的域适应贡献度,在域对齐过程中实现动态加权任务损失度量,通过特征分布距离度量有效消除中间域异常样本对学习效果产生的不利影响.

(3) 针对无标注 SAR 数据容易导致目标检测器向源域偏移的问题,提出融合数据增强与互学习的伪标签生成方法,有效提升目标域的特征表征能力.

(4) 利用多组异源多模态数据集进行了实验验证,结果表明本文方法相较于传统的基线方法展现出显著的优势,与原始方法相比检测的平均精度提升 21.4 个百分点,与基线方法相比平均精度提升 3.7 个百分点,实验结果也充分证明了跨模态知识迁移在 SAR 零样本目标检测任务中的可行性.

2 方法

针对现有 SAR 目标检测方法依赖于完备的样本分布空间,通用的模态转换方法容易丢失雷达图像目标电磁散射特性,论文提出一种基于跨模态渐进式知识迁移的 SAR 目标检测技术.首先,通过设计光/SAR 模态迁移模型引入伪 SAR 中间域,将其作为可见光模态与 SAR 模态之间的连接桥梁.然后,分别设计对抗学习策略实现源域与中间域特征对齐,构造全局关联与局部匹配完成中间域与目标域对齐,如图 2 所示.在此基础上,设计加权损失函数,动态调整不同样本的域适应贡献度,优化对齐效果.最后,利用伪标签学习优化 SAR 图像特征,引入强-弱数据增强与互学机制抑制不可靠伪标签对检测器训练的负面影响.

2.1 构建中间域与任务分解

由于光学图像和 SAR 图像之间存在着显著的语义差异,域间隙较大,现有跨模态学习方法难有效对齐^[29,40].对此,论文提出一种基于图像跨模态生成的中间域构建方法,将原有间隙较大的对齐任务分解为两个间隙较小的子任务对齐.首先,设计从可见光图像到 SAR 图像的智能生成方法,据此构建伪 SAR 图像中间

域,分别与 SAR 模态(目标域)与可见光模态(源域)连接起来,下文将源域、中间域和目标域分别表示为 S 、 I 和 T ,从而知识迁移任务分解两个步骤: $S \rightarrow I$ 和 $I \rightarrow T$,如图 3 所示.

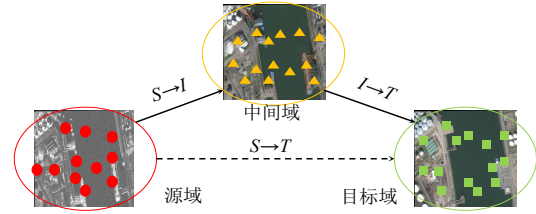


图3 分阶段知识迁移方法流程

2.1.1 风格解耦转换

通用模态转换方法用于光/SAR 数据时,其解耦能力有限,对内容(地形结构)与风格(光学颜色、纹理)的分离效果不好,容易产生风格泄露或内容扭曲.对此,本文提出一种双路径风格编码器,实现跨域风格特征与内容特征的高效解耦,抑制斑点噪声与强杂波的传递,同时增强风格的可控性,生成符合目标域分布的伪 SAR 图像,所提的风格解耦与编码整体框架如图 4 所示.

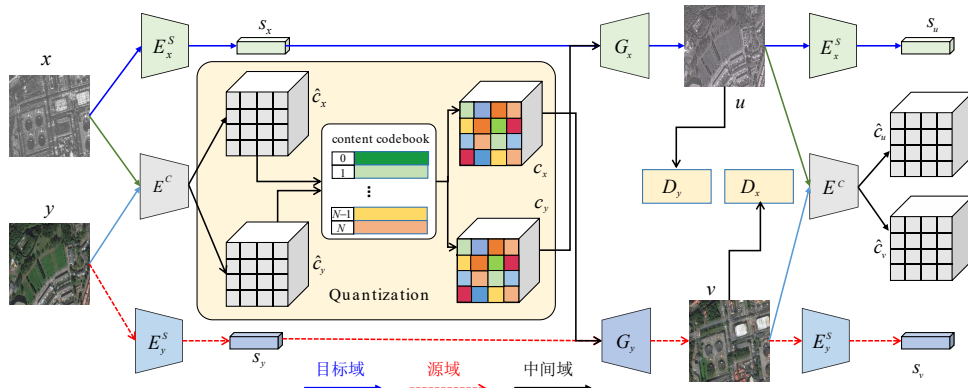


图4 双路径风格编码解耦的模态转换框架

对于输入图像 x, y , 首先通过共享内容编码器 E^c 提取跨域内容表示:

$$c_x = E^c(x), \quad c_y = E^c(y) \quad (1)$$

其中, c_x, c_y 表示内容特征, 同样采用与源/目标域相关的特定风格编码器 E_x^s, E_y^s 从输入图像中提取风格特征:

$$s_x = E_x^s(x), \quad s_y = E_y^s(y) \quad (2)$$

其中, s_x, s_y 表示风格特征, 风格编码器通过对抗训练学习域特异性分布.

然后设计生成器 G_x 和 G_y 对风格编码获取的抽象特征与内容编码的抽象特征进行解码, 使用组合交换的内容表示和样式特征, 以合成转换后的图像 $u \in X$ 和 $v \in Y$:

$$u = G_x(c_y, s_x), \quad v = G_y(c_x, s_y) \quad (3)$$

其中, 生成器的输入分别是内容特征 $\{c_x, s_y\}$ 与风格特征 $\{c_y, s_x\}$. 为提高学习效果, 设计判别器 D_x 和 D_y 进行真伪鉴别, 形成域对抗损失 L_{adv} , 最终用于模型训练的目标函数可表示为

$$L_{X \rightarrow Y}^{adv} = \mathbb{E}_{y \sim P_{data}(Y)} [\log D_Y(y)] + \mathbb{E}_{x \sim P_{data}(X)} [\log (1 - D_Y(G_Y(cx, sy)))] \quad (4)$$

$$L_{Y \rightarrow X}^{adv} = \mathbb{E}_{x \sim P_{data}(X)} [\log D_X(x)] + \mathbb{E}_{y \sim P_{data}(Y)} [\log (1 - D_X(G_X(cy, sx)))] \quad (5)$$

$$L_1^{style} = \left\| E_x^s(G_x(c_y, s_x)) - s_x \right\| + \left\| E_y^s(G_y(c_x, s_y)) - s_y \right\| \quad (6)$$

在此基础上进一步使用潜在内容回归损失提升学习效率:

$$L_1^{\text{content}} = \left\| E^c(G_X(c_y, s_x)) - c_y \right\| + \left\| E^c(G_Y(c_x, s_y)) - c_x \right\| \quad (7)$$

综上所述,总的损失函数可被定义如下:

$$L_D = L_{\text{adv}} \quad (8)$$

$$L_{E^c, Z, E^s, G} = -\lambda_{\text{adv}} L_{\text{adv}} + \lambda_1^{\text{content}} L_1^{\text{content}} + \lambda_1^{\text{style}} L_1^{\text{style}} \quad (9)$$

其中, L_D 用于训练鉴别器; $L_{E^c, Z, E^s, G}$ 用于训练编码器和生成器; $\lambda_{\text{adv}}, \lambda_1^{\text{content}}, \lambda_1^{\text{style}}$ 是用于控制对抗损失、内容损失与风格损失的权重参数.

2.1.2 知识迁移流程

模态迁移合成结果在视觉风格和纹理上近似 SAR 图像, 具有高对比度和粗糙纹理等特点, 它既复制了目标域 SAR 图像的散射特性, 也保留了光学图像固有的一些视觉语义特征, 最终形成介于两种模态之间的混合特征分布. 通过引入中间域, 可将困难特征对齐任务分解为两个简单子任务, 即从源域到中间域和从中间域到目标域. 具体而言, 在 $S \rightarrow I$ 阶段, 首先在含有标注的源域 S 与不含标注的中间域 I 进行特征对齐. 由于 S 和 I 之间存在潜在的内容相似性, 本阶段主要是实现像素级结构与几何特征的对齐. 在完成从源域到中间域的特征对齐后, 中间域也从源域继承其标注信息, 然后继续进行中间域 I 与目标域 T 之间的特征对齐. 在第二阶段, 模型重点根据第一阶段学习到的结构不变特征, 调整外观和光学特征等. 总的来说, 我们的分阶段适应方法通过将复杂任务拆分为两个可管理的子任务, 从而简化学习过程, 各子任务分别解决复杂任务的部分内容.

为了更直观地展示中间域效果, 利用 t-SNE (t-distributed Stochastic Neighbor Embedding)^[45] 将 SpaceNet6 数据集迁移结果映射到二维空间, 如图 5 所示(绿色、紫色、红色特征点分别代表源域、中间域及目标域特征), 可以看出中间域的特征分布位于 SAR 域和光学域之间, 有效弥合光/SAR 模态数据之间的语义鸿沟. 在此基础上, 引入 MMD (Maximum Mean Discrepancy distribution) 分布和 CORAL (Correlation Alignment for Domain Adaptation) 距离两种经典的域间隙距离度量^[46], 如表 1 所示, 对源域、中间域和目标域之间的距离进行量化描述, 可以看出, 通过图像生成方法构建的中间域, 有效缓解了直接从源域到目标域之间的语义鸿沟, 从源域到中间域的距离以及从中间域到目标域之间的距离均明显小于从源域到目标域的距离, 也验证了论文所提出的利用生成的伪 SAR 图像构建中间域的合理性.

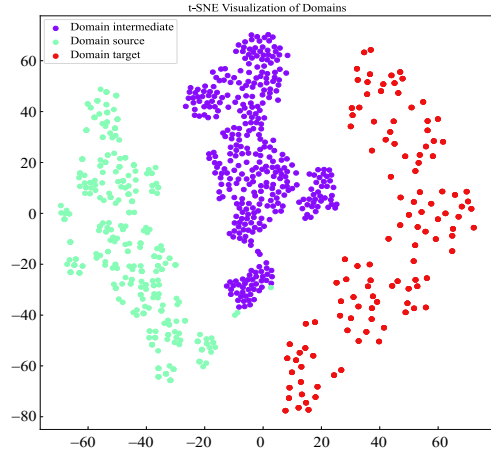


图 5 SpaceNet6 数据集特征可视化

表 1 样例数据的 MMD 和 CORAL 域间距离度量结果

参数	MMD	CORAL
$S \rightarrow T$	0.142 6	95.920 9
$S \rightarrow I$	0.141 6	84.614 6
缩小	0.001 0	11.306 3
$I \rightarrow T$	0.141 4	29.230 5
缩小	0.001 2	66.690 4

2.2 特征对齐与加权损失

2.2.1 多尺度层级特征对齐

针对光/SAR 图像的显著差异, 本文提出全局弱特征对齐与局部强特征对齐的学习策略, 提高特征对齐效果, 整体学习框架如图 6 所示, 其中 $F1$ 和 $F2$ 分别表示局部级和全局级特征提取器, FC 表示全连接层, D_l 和 D_g 分别是局部和全局域分类器, GRL 表示梯度反转层. 通过基于对抗学习的特征对齐网络和域判别网络实现特征对齐, 其中弱对齐主要针对全局特征空间, 而强对齐则应用于局部特征空间.

假设含有标注的图像表示为 x^a , 其标注 y^a , 它们均来自一组含有标注的数据集 $\{X_a, Y_a\}$. 将未被标注的图像记为 x^b , 它是从一组未被标注的数据集 X_b 中采样获

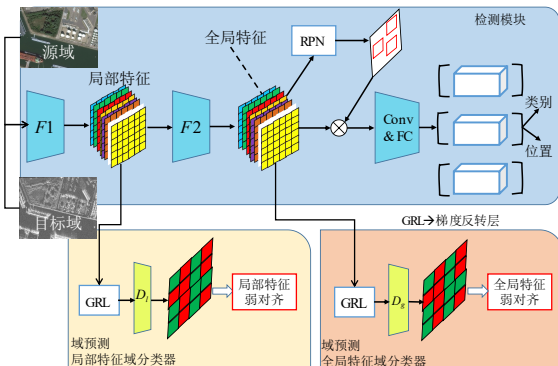


图 6 多尺度层级特征对齐网络结构

得. 全局域分类器 D_g 预测的域标签表示为 d , 对于含有标注的域, 设置为 1, 对于未被标注的域设置为 0. 网络 R 包含区域提议网络 (Region Proposal Network, RPN) 以及其他附加组件. 检测损失的目标可以描述为

$$L_{\text{cls}}(F, R) = -\frac{1}{n} \sum_{i=1}^n L_{\text{det}}(R(F(x_i)), y_i) \quad (10)$$

其中, L_{det} 表示所有检测损失; n_a 表示特征对齐过程中含有标注的图像数量. 若定义全局弱对齐域分类器的损失函数表示为 L_{global} :

$$L_{\text{global}_a} = -\frac{1}{n_a} \left(1 - D_g(F(x_i^a))^y \log(D_g(F(x_i^a))) \right) \quad (11)$$

$$L_{\text{global}_b} = -\frac{1}{n_b} D_g(F(x_i^a b))^y \log(D_g(F(x_i^b))) \quad (12)$$

$$L_{\text{global}}(F, D_g) = \frac{1}{2} (L_{\text{global}_a} + L_{\text{global}_b}) \quad (13)$$

其中, n_b 表示未被标注图像的数量. 仅采用对齐全局特征无法有效缩小两个不同域之间的间隙. 因此进一步引入全局特征对齐. 局部域分类器 D_l 侧重于提取细粒度的局部特征差异, 与全局特征对齐形成互补. 特征提取模块 F 可细分为 F_2 和 F_1 . 其中 F_1 输出的局部特征图 (尺寸为 $W \times H$) 被输入到局部域分类器 D_l 中. 这种局部对齐的损失函数表示为 L_{loc} , 定义为

$$L_{\text{loc}_a} = \frac{1}{n_a HW} \sum_{i=1}^{n_a} \sum_{w=1}^W \sum_{h=1}^H D_l(F_1(x_i^a))_{wh}^2 \quad (14)$$

$$L_{\text{loc}_b} = \frac{1}{n_b HW} \sum_{i=1}^{n_b} \sum_{w=1}^W \sum_{h=1}^H (1 - D_l(F_1(x_i^b)))_{wh}^2 \quad (15)$$

$$L_{\text{loc}}(F_1, D_l) = \frac{1}{2} (L_{\text{loc}_a} + L_{\text{loc}_b}) \quad (16)$$

对抗损失 $L_{\text{adv}}(F, D)$ 可以表示为

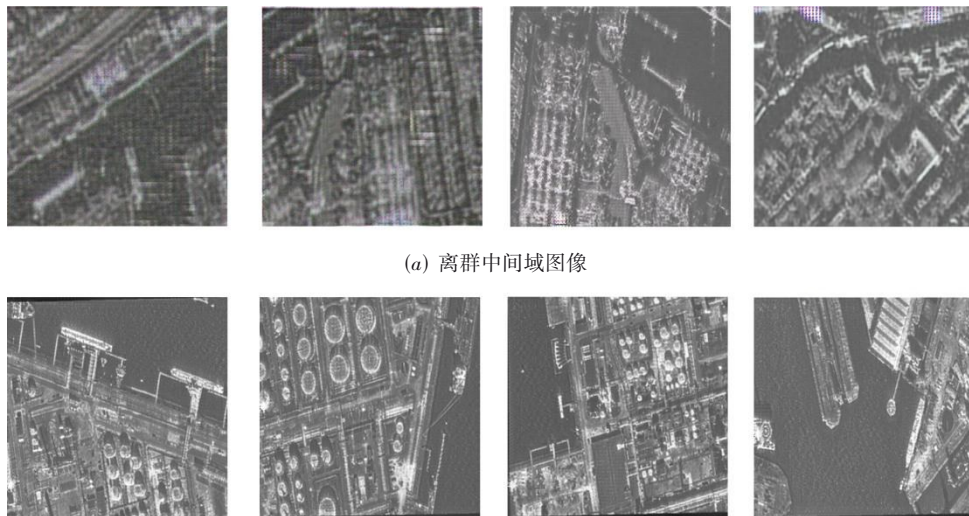
$$L_{\text{adv}}(F, D) = L_{\text{loc}}(F_1, D_l) + L_{\text{global}}(F, D_g) \quad (17)$$

2.2.2 加权任务损失

由于在中间域伪 SAR 图像的生成过程中没有标记监督信息可供参考, 可能造成合成结果偏离中间域分布的风险, 如局部失真、噪声干扰、纹理偏差等, 无法保留源图像的细节信息, 或是没有充分采用 SAR 图像风格, 如电磁散射特性扭曲等. 离群样本可能会导致模型混淆, 产生不正确的跨域特征对齐, 部分样例如图 7 所示. 此外, 从图 5 的 t-SNE 特征分布也可以看出, 部分紫色的中间域特征点距离源域和目标域均很远处, 因此离群样本是不可避免. 为了缓解这种现象, 抑制离群样本的不利影响, 论文提出一种样本重要性评估策略, 通过评估合成图像特征相对于源域或目标域中特征的距离, 分配重要性权重. 具体而言, 通过域鉴别器的预测输出来计算权重, 特征分布离未被标注域越远, 对它的重视程度就越低, 其中重要性权重定义为

$$w_i = \frac{\mu P_S + (1 - \mu) P_T}{P_S + P_T} \quad (18)$$

其中, P_S 和 P_T 分别表示合成图像属于源域或目标域的可能性. 在第一阶段, 核心任务是完成源域 (可见光图像) 到合成中间域 (伪 SAR 图像) 的特征对齐, 学习并匹配两者之间的场景内容语义结构、几何形状等特征, 因此参数 μ 近似为 1; 在第二阶段, 核心任务转变为跨域对齐, 通过引入中间域使其到目标域 (SAR 图像) 的特征进行对齐, 学习并匹配两者之间的纹理模式、噪声特性、散射机制等特征, 此时权重参数演化, 它更改为 0. 该权重作用于检测损失函数表现为



(a) 离群中间域图像

(b) 正常中间域图像

图7 中间域离群样本示例

$$\tilde{L}_{\text{cls}}(F, R) = -\frac{1}{n} \sum_{i=1}^n w_i L_{\text{det}}(R(F(x_i)), y_i) \quad (19)$$

目标函数更新为

$$\max_D \min_{F, R} \tilde{L}_{\text{cls}}(F, R) + \lambda L_{\text{adv}}(F, D) \quad (20)$$

其中,参数 λ 用于调整检测和对抗损失之间的平衡.

2.2.3 自训练伪标签

SAR图像缺乏标注信息,检测模型在训练过程中往往受到光学图像的影响偏向光学域,这可能导致检测器对源域数据的偏离,影响SAR图像目标检测性能.对此,提出伪标签学习策略,对目标域图像进行伪标注,利用伪标签进一步训练检测模型.尽管SAR图像的伪标签样本无法保证完全准确地被标注,但它们仍完全来自SAR域,这使得检测器能够直接学习到SAR图像更多更具辨别力的特征.

伪标签学习策略包括教师模型和学生模型两个关键要素,它们通过两阶段特征对齐过程训练的检测器进行初始化,其中教师模型对输入SAR图像进行弱增强,如水平翻转和裁剪等,学生模型则对输入的SAR图像进行强增强,如颜色抖动和补丁丢弃等.教师模型生成的伪标签指导学生的训练,学生模型反过来使用指数移动平均(Exponential Moving Average, EMA)^[47]方法更新教师模型.更新公式定义为

$$\theta_t \leftarrow \alpha \theta_t + (1 - \alpha) \theta_s \quad (21)$$

其中, θ_t 和 θ_s 分别表示教师和学生模型的参数.

为提高伪标签的准确性,为教师模型的边界框预测设置了置信阈值 δ .置信阈值有助于滤除误检.此外,我们还利用非最大抑制(Non Maximum Suppression, NMS)来消除冗余的边界框预测.因此,在获得从教师模型导出的伪标签后,可以利用这些伪标签来进一步优化学生模型,损失函数定义如下:

$$L_{\text{unsup}}(X_t, \hat{Y}_t) = L_{\text{cls}}^{\text{pn}}(X_t, \hat{Y}_t) + L_{\text{cls}}^{\text{oi}}(X_t, \hat{Y}_t) \quad (22)$$

其中, \hat{Y}_t 表示伪标签.最终总的目标函数是

$$\max_D \min_{F, R} \tilde{L}_{\text{cls}}(F, R) + \lambda L_{\text{adv}}(F, D) + \theta L_{\text{unsup}}(X_t, \hat{Y}_t) \quad (23)$$

其中, λ 和 θ 是用于调整相应损失权重的超参数,如何确定该组超参数目前并没有科学的解决方案,本文主要根据经验进行设定,即通过多组实测验证确定最优值.

2.3 总体流程

针对光/SAR图像之间的语义差异,本文提出了一种渐进式知识迁移学习方案,整体过程可概括为三个步骤:(1)源域和中间域之间的初始特征对齐;(2)中间域和目标域之间的精细对齐;(3)自训练伪标签学习阶段.首先,通过提出的跨域转换模型将光学图像转换为SAR风格表示,称其为中间域伪SAR图像;然后构建两阶段特征对齐任务.

第一阶段,使用对抗学习将源域和中间域之间的

多尺度特征进行对齐,建立一个过渡特征空间.

第二阶段,在第一阶段对齐的基础上,将学习结果进一步扩展到中间域和目标域,进行特征对齐.

此外,为了解决合成图像中部分离群图像的影响,本文提出对任务损失进行线性加权的办法,根据与源域或目标域的特征分布的远近程度来赋予其不同的权重;最后启动自训练伪标签学习,对于经过两阶段特征对齐过程训练的检测器,使用其对目标域图像进行预测推理,生成样本的伪标签继续进行训练,使用强弱增强和伪标签学习的方法来最小化噪声伪标签对检测器训练的影响.

3 实验结果与分析

为了验证论文提出的渐进式跨模态知识迁移SAR目标检测技术的合理性与有效性,论文利用多级实测数据进行实验验证.首先介绍实验设置与数据,然后阐述实验采用的评价指标,最后对实验结果进行全面分析与讨论.

3.1 实验设置

3.1.1 数据集介绍

为全面评估算法性能,本文利用SpaceNet6^[48]、SSDD(SAR Ship Detection Dataset)^[49]和HRSID(High-Resolution SAR Images Dataset)^[50]公开数据集进行验证,任务说明与实验细节如表2所示.

表2 实验所定义的不同任务类型

任务	可见光模态(源域)	SAR模态(目标域)	迁移方向
任务一	SpaceNet6 光学图像	SpaceNet6 SAR图像	同源迁移
任务二	SpaceNet6 光学图像	SSDD SAR图像	异源迁移
任务三	SpaceNet6 光学图像	HRSID SAR图像	异源迁移

SpaceNet6数据集同时包含光学域和SAR域两个模态的图像,因此论文选择SpaceNet6中的光学图像作为知识源域,SpaceNet6中的SAR图像、SSDD数据和HRSID数据分别作为知识迁移的目标域.从SpaceNet6数据集中选取包含港口和海洋场景的1500对图像,并对其中舰船目标进行标注,训练集与测试集按7:3的比例划分.

SpaceNet6是典型的多模态遥感数据集,聚焦于多模态图像转换、融合、地物分类和变化检测研究.该数据集包含配对的Capella Space SAR影像(X波段, VV/VH极化, 0.5 m分辨率)与Maxar光学图像(多光谱, 0.3 m分辨率),通过精确的时空匹配与坐标校准,确保SAR与光学数据在空间上完全对齐,并经过辐射定标、地形校正等标准化预处理流程.

SSDD是面向SAR舰船检测的开源数据集,它涵盖RadarSat-2、TerraSAR-X、Sentinel-1等多源卫星数据,包含1160张SAR图像和2456个舰船实例,覆盖不同分

分辨率(1~15 m)、极化方式(HH、HV、VV、VH)及场景(远海、近岸、港口等)。

HRSID是专为高分辨率SAR舰船检测与语义分割设计的数据集,包含5 604张1 000×1 000像素的图像和16 951个舰船实例,分辨率覆盖0.5~3.0 m,场景包括开阔海域、港口及复杂干扰环境(如海杂波),支持目标检测、实例分割等任务。

3.1.2 实验环境及参数配置

论文实验基于PyTorch框架进行验证,所有实验在相同的环境下进行训练和测试。所提方法 λ 、 θ 分别设置为0.2、1.0,置信阈值 δ 被设置为0.7,EMA的权重平滑参数被设置为0.999 6。模型的训练过程采用200轮次,前120个轮次的学习率固定为0.000 2,后80个轮次线性衰减到0。所有模型使用Adam优化器进行模型训练,其指数衰减因子 β_1 设置为0.5, β_2 设置为0.999。

为更全面、更客观地验证所提方法的有效性,论文分别从同源知识迁移SAR目标检测与异源知识迁移SAR目标检测两个维度组织实验,与基线方法进行比较,并使用“源域无迁移”作业下限基准,它是指直接将可见光图像训练的目标检测网络应用于目标域SAR图像,进行性能测试,即不执行任何迁移处理。

3.2 同源知识迁移实验

所谓“同源知识迁移”实验是指作为源域的可见光图像与作为目标域的SAR图像均来自相同的数据集SpaceNet6,将可见光模态数据作为知识源,迁移至SAR图像,用于目标检测任务,详见表3所示的任务一。

表3 SpaceNet6同源模态迁移的实验结果(任务一) 单位:%

Method	AP	Precision	Recall	F1-score
源域无迁移	32.1	42.1	36.0	38.8
DA-Faster ^[32]	39.2	43.8	39.4	41.5
SWDA ^[33]	43.3	44.5	42.3	43.4
AT ^[28]	49.4	46.8	49.8	48.2
HTCN ^[51]	42.7	43.2	45.7	44.4
D-Adapt ^[52]	49.7	46.7	45.1	45.6
CMT ^[29]	48.9	45.4	47.1	46.2
DAI-Net ^[40]	50.3	48.2	49.5	48.8
PMT ^[35]	48.6	47.5	47.5	47.3
PCKT	53.6	56.5	50.9	53.5

3.2.1 与基线方法的对比实验

为验证论文所提方法的有效性与合理性,本节执行迁移任务一(SpaceNet6数据可见光图像到SAR图像)进行实验,并与当前各种最先进的方法进行比较,具体结果如表3所示。对实验结果从三个层面进行分析:

(1)与源域无迁移方法的结果比较。源域无迁移是

指利用可见光图像训练目标检测模型,然后直接应用于SAR图像,寻找感兴趣目标和区域,可以看出在没有任何知识迁移技巧的情况下,SAR图像目标检测结果为32.1%,该结果也表明了可见光模态与SAR模态之间存在着显著的语义差异,直接将源域的学习结果应用于目标域效果并不理想。

(2)与基线方法结果的比较与分析。论文复现了5种经典的跨模态知识迁移学习方法,分别是发表于CVPR2018、CVPR2019、CVPR2022、ICIR2022等著名学术会议的DA-Faster(Domain Adaptation Faster-RCNN)、SWDA(Strong-Weak Distribution Alignment)、AT(Adapt Teacher)、HTCN(Hierarchical Transferability Calibration Network)和D-Adapt,其中DA-Faster通过对抗学习实现特征对齐,得到的检测AP(Average Precision)为39.2%,SWDA分别对全局和局部特征构建了强对齐和弱对齐两种策略,实现目标检测AP为43.3%,HTCN方法采用Cycle GAN对抗博弈的思路最小化像素级差异,实现目标检测AP为42.7%,D-adapt通过采用解耦自适应的特征对齐方法实现知识迁移,最终的目标检测AP为49.7%,AT提出的自适应教师模型是利用领域对抗学习和弱-强数据增强来解决领域差距,其目标检测AP为49.4%。可以看出,所提方法最多提升检测率达到21.5个百分点,与次优方法相比也提高了3.3个百分点,实验结果很好地验证所提方法的合理性。

(3)与最新方法结果的比较与分析。同样论文也复现并比较了3种最新的跨模态学习方法,分别是发表于CVPR2023、CVPR2024学术会议的DAI-Net、PMT和CMT,其中DAI-Net通过使用零样本光学域自适应来改进低光物体检测,目标检测率为50.3%,但基于特征对齐的方法主要侧重于缩小域间隙,而不是针对未标注的目标域来优化目标检测网络。CMT使用伪标签和对比损失优化对象级特征,消除了对目标域标注的需求,其目标检测AP为48.9%。PMT方法使用类原型对领域特定信息进行编码,取得的目标检测AP为48.6%。可以看出,由于存在较大的域间隙,最新提出的三种方法在可见光和SAR图像之间的知识迁移有效性有限,本文所提方法分别提高了目标检测AP为3.3、4.7和5.0个百分点。与所有基线方法相比,论文所提方法利用光/SAR生成技术引入伪SAR中间域,有效地缩小可见光图像和SAR图像之间的域间隙。同时,利用伪标签学习策略进一步提升异源数据对标域图像的学习效率,有效地应对检测器的域偏移现象。论文所提方法的目标检测AP为53.6%,明显优于所有基线方法。同样通过Precision、Recall和F1-score等指标也能够看出论文所提方法与对比方法相比有了较大提升。

3.2.2 消融实验

为了弥合光/SAR 模态数据的语义鸿沟,论文提出了一种渐进式跨模态学习方法,主要创新与贡献是包括渐进式任务分解(Progressive Task Decomposition, PKT)、多尺度特征对齐(Multi-scale Feature Alignment, FDA)和自训练伪标签学习(Self-training Pseudo-label learning, STPL)三种策略,实现从可见光遥感图像到 SAR 图像的知识迁移,提高无标签信息条件下的目标检测效果.为了验证所提方法各模块及不同策略的有效性,论文利用 SpaceNet6 数据集进行同源知识迁移(任务一)消融实验,实验结果如表 4 所示.

表 4 消融实验结果

方法	PKT	FDA	STPL	AP/%
源域无迁移				32.1
PCKT	√			35.9
		√		42.6
			√	34.7
	√	√		51.9
	√		√	39.3
		√	√	48.8
	√	√	√	53.6

自训练伪标签学习是利用目标检测模型对目标域 SAR 图像进行推理获取伪标注,进一步提高模型学习效率,解决检测器对源域数据的偏离影响.从实验结果来看,引入自训练伪标签学习策略的目标检测 AP 为 34.7%,检测性能提升了 2.6 个百分点,表明自训练伪标签学习能更充分地挖掘异源数据的重要信息,提高知识迁移效率.

同样将分阶段知识迁移、多尺度特征对齐、自训练伪标签学习分别协同用于模型训练,也能显著提升目标检测的检测效果,如 PKT 与 FDA 的组合效果目标检测率为 51.9%,PKT 与 STPL 的协同学习效果目标检测率为 39.0%,而 FDA 与 STPL 组合后的学习效果为 48.8%,也表明多种策略间存在显著的协同效应.

此外,进一步利用 SpaceNet6 数据集同源知识迁移

实验对合成图像的不同使用方法进行对比,验证其对检测结果的影响,实验结果如图 8 所示,其中“Source only”是指仅利用源域的可见光图像进行训练,然后对目标域的 SAR 图像进行检测.Synthetic Augment 将图像跨域转换作为一种数据增强手段,将源域与转换后的合成图像合并在一起训练,然后对目标域图像进行检测;Synthetic Augment DA 表示将源域图像与合成图像组合在一起与目标域间直接进行特征对齐.从实验结果可以看出,合成图像的各种使用方法均在不同程度上提升目标检测性能,但所提方法的协同效果最好,这是因为所提方法将各种改进策略整合在同一个学习框架内,从而有效提升协同效果.

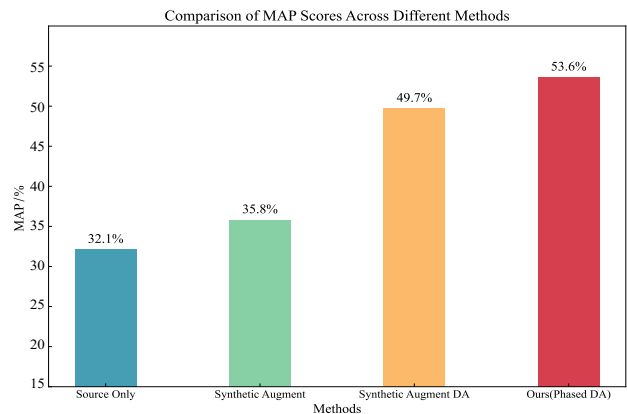


图 8 合成图像各种应用方法对检测结果的影响

3.2.3 域间隙可视化展示

为了更好地展示模态知识迁移效果,本节对引入中间域后,源域、中间域和目标域图像进行可视化投影,结果图 9 所示.可以看出,中间域图像表现出从光学到 SAR 风格的转变.然而,中间域图像保留了特定的光学特性,特别是在物体的几何和结构方面仍然与目标域存在部分差异.总体而言,中间域伪 SAR 图像特征代表光学图像和 SAR 图像之间的中间状态,可作为源域与目标域之间的一个桥梁,将单阶段难以有效对齐的问题分解为两阶段的简单子任务.

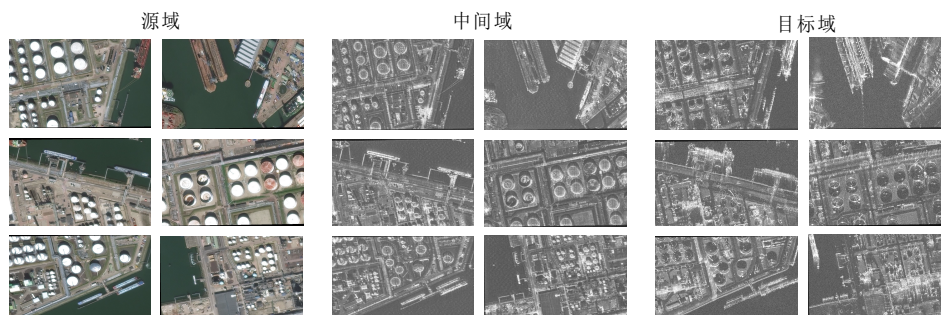


图 9 源域、中间域和目标域的部分样例数据

为了进一步说明源域、目标域和中间域之间的域间隙,我们使用t-SNE将特征分布进行可视化展示,结果如图10所示,其中第一、二、三行分别表示SpaceNet6可见光图像与SpaceNet6、SSDD和HRSID SAR图像在特征对齐前后的特征分布情况,每行三幅子图分别表示源域、中间域和目标域特征分布.从图10(a)和图10(b)可以看出中间域(紫色特征点)的特征分布位于源域(绿色特征点)和目标域(红色特征点)之间.通过引入中间域图

像,可显著降低源域和目标域之间的特征对齐难度.

图10(c)显示了多尺度特征对齐阶段后源域和目标域的特征分布,通过与图10(a)对比可以看出,在特征对齐之前,源域与目标域的特征分布得很远,将其分辨出的难度很低.图10(c)显示,在进行多尺度特征对齐后,源域与目标域的特征分布变得更相近,难以有效区分,从而验证了所提的多尺度特征对齐方法能实现不同域特征分布的有效对齐.

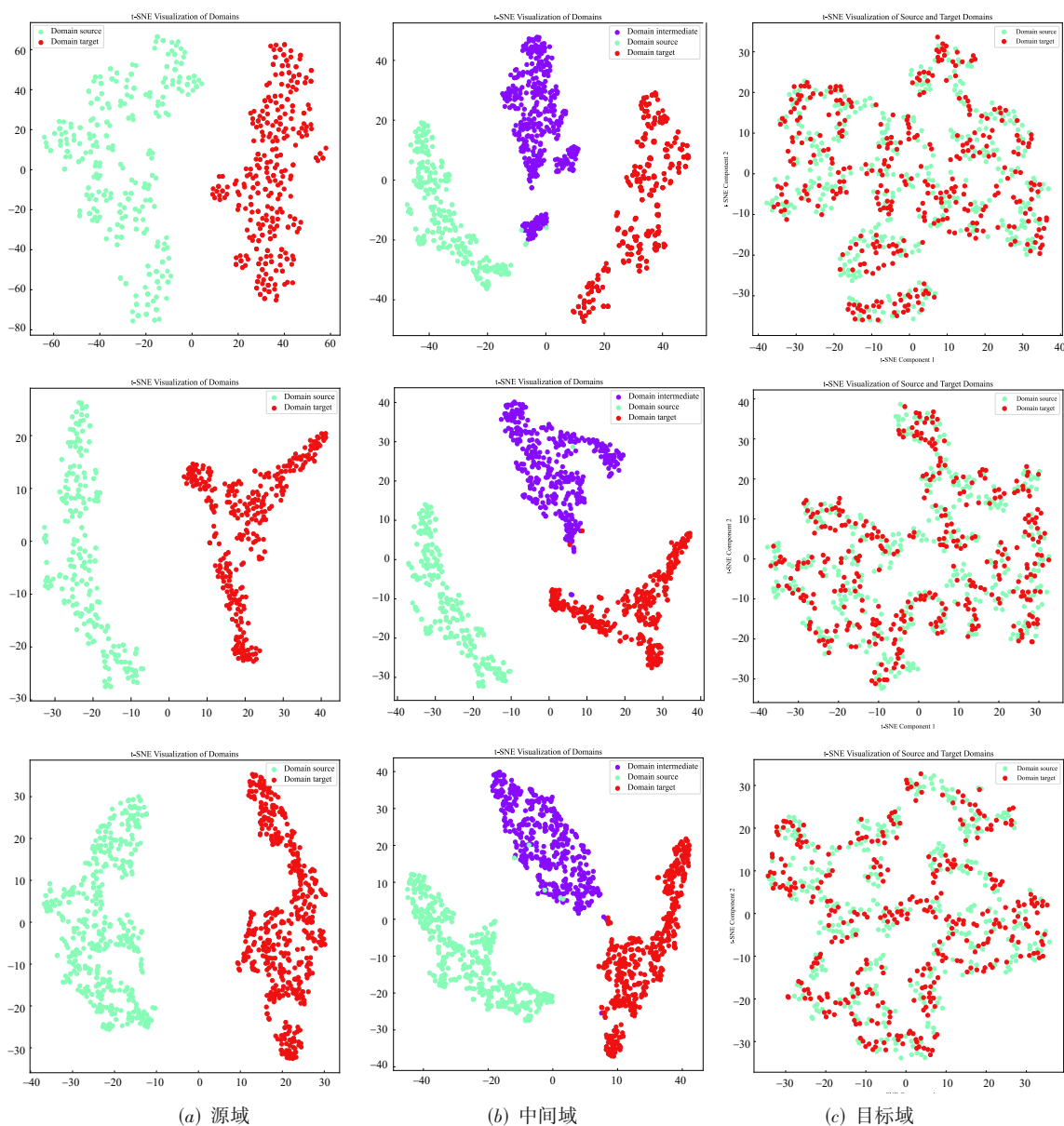


图10 多域数据t-SNE特征可视化分布

3.2.4 自适应权重参数

自适应加权任务损失对于所提方法的学习效果起着十分重要的作用,为进一步验证其对于处理合成图像质量不平衡问题的有效性,本节利用SpaceNet6数据

集(任务一)针对权重参数的变化情况进行了一组同源知识迁移实验,其中各种固定权重和自适应加权任务损失对目标检测结果的影响如表5所示.

可以看出,由于所提方法根据每个合成图像与源

域、目标域分布之间的距离来计算其权重,从而有效抑制低质量离群数据对学习效果的影响,实验结果证明该方法明显优于其他固定权重方法的检测结果.同时,与不使用任何权重(即权重等于1)相比,所提方法也可将检测AP提高了1.9个百分点.

表5 权重参数对检测性能影响

权重	0.8	0.9	1.0	1.1	1.2	PCKT
AP/%	49.8	50.3	51.7	51.2	50.8	53.6

3.2.5 伪标签学习效果

为了证明自训练伪标签策略的有效性,本节利用SpaceNet6数据集(任务一)进行多组实验,进一步对比分析自训练伪标签学习前后的性能差异.如图11所示,伪标签策略显著提升了预测框的准确性和mIoU指标.具体而言,图11(a)的对比结果表明,经过自训练伪标签阶段优化的预测结果在边界框定位精度上表现出明显优势.该策略使检测器能够更准确地捕捉目标的空

间特征,有效降低了虚警率和漏检率.同样的,图11(b)展示了自训练伪标签策略对预测框mIoU性能的显著提升.可以看出,经过伪标签优化后的预测边界框与真实目标的空间对齐精度明显提高.实验结果也进一步验证了本文所提方法能够生成更准确的检测结果.

在此基础上,我们进一步组织实验对伪标签学习前后的性能差异进行对比分析,实验结果如图12所示.可以看出,利用知识迁移策略训练的检测器,对目标域图像进行预测推理生成的伪标签,可显著提升预测框的准确性和mIoU指标.具体而言,图12(a)的对比结果表明,经过自训练伪标签学习优化后的预测结果,在边界框定位精度上表现出明显优势,检测器能更准确地捕捉目标的空间特征,有效降低了虚警率和漏检率.同样图12(b)展示了自训练伪标签策略对预测框mIoU性能的显著提升,经过伪标签学习优化后的预测边界框与真实目标的空间对齐精度明显提高,也验证了伪标签学习方法能够生成更准确的检测结果.

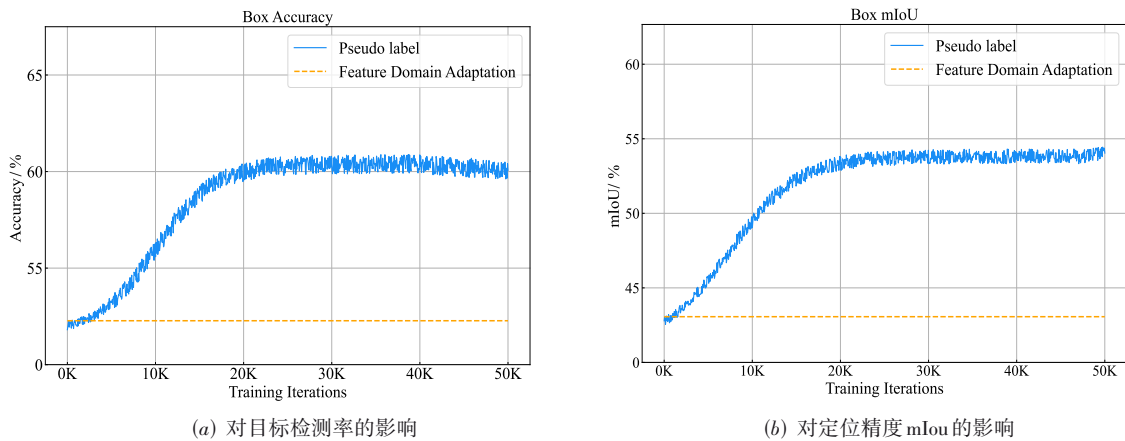


图11 伪标签学习效果

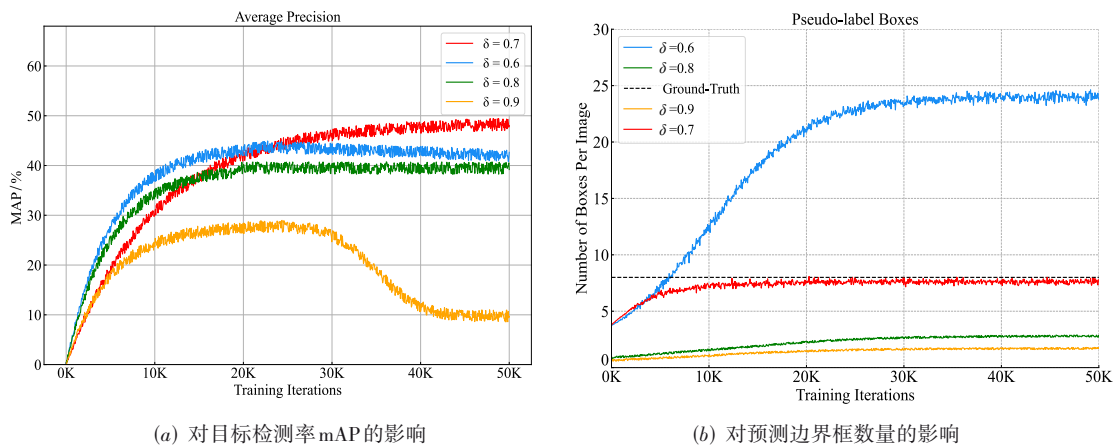


图12 置信度阈值 δ 的取值讨论

3.3 异源知识迁移实验

“异源知识迁移”实验是指作为源域的可见光图像与作为目标域的SAR图像均来自不同的数据集,即使用SpaceNet6可见光模态数据作为知识源,将其迁移至其他SAR图像数据集,实现零样本条件下的SAR目标检测任务,详见表2所示的“任务二”和“任务三”。显然,与同源知识迁移任务相比,异源知识迁移的挑战和难度更大。

从SpaceNet6可光图像到SSDD SAR图像的知识迁移结果如表6所示。可以看出,与同源知识迁移实验结果相比,异源知识迁移的目标检测性能均有不同程度的下降,这也表明异源知识迁移的难度远高于同源知识迁移。但与基准方法相比,所提方法仍取得了很好的检测效果,与“源域无迁移”基线方法相比,所提方法的目标检测AP提升了21.4个百分点。现有最新方法的结果与“任务一”相比有些变化,DAI-Net在“任务一”的同源知识迁移中效果优于其他基线方法,但在异源知识迁移实验结果却逊于CMT和PMT,表明该方法的表现并不稳健,而自训练伪标签学习方法的效果更优。与现有基线方法相比,所提方法的目标检测AP比次优方法提升了2.6个百分点,比源域方法提升了21.4个百分点,很好地验证了所提方法知识保留效率。

表6 SpaceNet6与SSDD异源模态迁移实验结果(任务二) 单位:%

方法	AP	Precision	Recall	F1-score
源域无迁移	29.4	31.8	33.5	32.6
DA-Faster	35.7	38.8	39.4	39.1
SWDA	39.7	44.6	42.3	43.4
HTCN	38.1	43.6	39.3	41.3
D-Adapt	45.4	44.1	47.3	45.6
AT	43.1	45.4	42.5	43.9
CMT	48.2	48.7	45.1	46.8
DAI-Net	45.2	47.7	48.6	48.2
PMT	47.7	48.2	48.4	48.3
PCKT	50.8	52.3	48.5	50.3

为进一步验证论文提出方法的异源知识迁移效率,我们组织了从SpaceNet6可光图像到HRSID SAR图像的知识迁移实验,结果如表7所示。可以看出,与SSDD数据集的迁移实验结果类似,整体的目标检测效果低于同源知识迁移结果,DAI-Net的表明弱于CMT和PMT伪标签学习,但本文提出的方法仍明显优于基线方法与现有的主流方法,也验证了本文方法的异源知识迁移合理性与先进性。

3.4 知识迁移目标检测结果演示

为了直观地展示本文方法的实际效果,本节将论文所提方法与多种主流知识迁移方法对于部分样例

表7 SpaceNet6与HRSID异源模态迁移实验结果(任务三) 单位:%

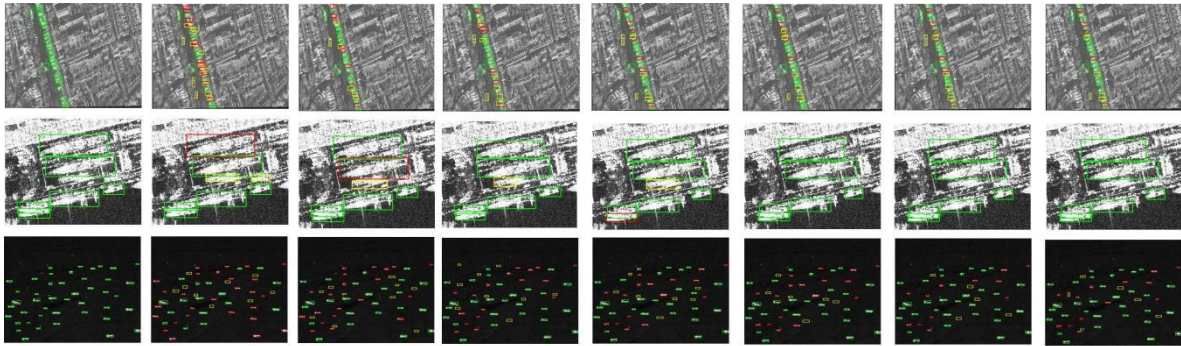
方法	AP	Precision	Recall	F1-score
源域无迁移	31.7	32.9	29.7	31.2
DA-Faster	35.1	36.0	38.1	37.0
SWDA	36.9	37.5	37.8	37.6
HTCN	38.5	41.1	38.2	39.6
D-adapt	42.1	45.3	46.0	45.6
AT	43.7	46.1	45.9	46.0
CMT	47.2	48.1	46.6	47.3
DAI-Net	44.1	43.9	45.7	44.7
PMT	44.6	45.9	45.5	45.7
PCKT	48.1	47.2	46.8	47.0

图像的目标检测结果进行可视化对比,结果如图13所示。从图13可以看出,可见光图像和SAR图像之间存在着显著的目标特性差异,因此对跨模态知识迁移提出了很大挑战。显著的特征差异导致DA-Faster方法中虚警和漏检率增加,D-Adapt通过应用边界框自适应提高准确性并减少了漏检,但虚警方面仍需改善;自适应老师模型(AT)通过强、弱特征对齐方法学习光学和SAR数据的共同特征,提高了对SAR数据信息的挖掘,较好地改善了漏检现象。DAI-Net通过利用Retinex理论和光照不变性策略提高了目标检测率,同样CMT和PMT有效地利用未标记数据中的有用信息,增强了模型的泛化能力,显著减少了虚警。与之相比,本文方法PCKT(Progressive Crossmodal Knowledge Transfer)对目标检测率、虚警率和漏检率各方面均取得了显著提升,也验证了渐进式分阶段知识迁移的合理性。

为进一步验证所提方法实际效果,将PCKT与经典的恒虚警率目标检测方法CFAR进行比较,它是基于杂波统计建模所推导出的理论最优检测器,改进方法如单元均值恒虚警(Cell-Average Constant False Alarm Rate, CA-CFAR)、双参数恒虚警(Two-Parameter Constant False Alarm Rate, TP-CFAR)、两阶段恒虚警(Two-Stage-SP-CFAR)等,实验结果如图14所示。可以看出,与传统方法相比,论文所提方法显著优于恒虚警率方法族的检测性能,无论是远海还是近岸,论文提出的方法虚警和漏警情况更少。通过从可见光图像到SAR图像的分阶段知识迁移,可学习目标更丰富的细节信息,显著改善了检测算法对复杂杂波背景条件下的鲁棒性和可靠性。

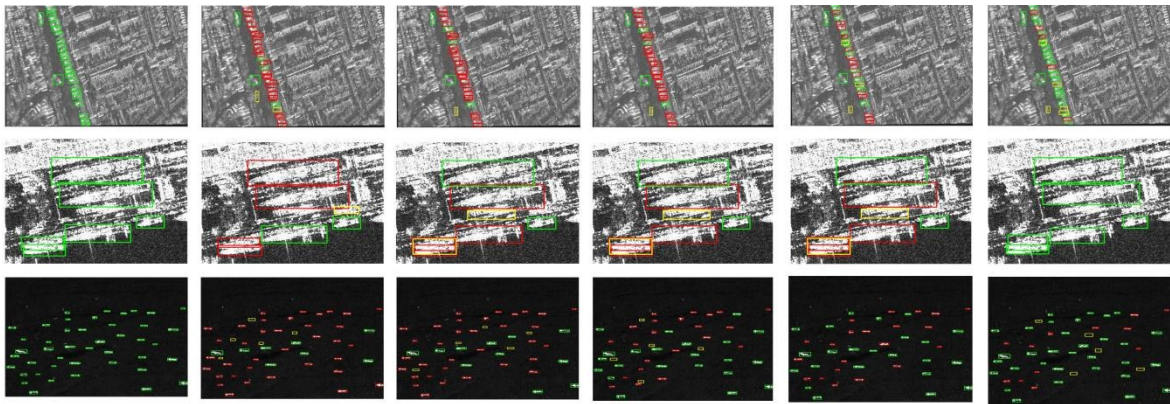
4 结论

针对可见光与SAR图像之间显著的模态差异影响知识迁移效果的问题,论文提出了一种融合分阶段知识迁移与自训练伪标签优化的SAR图像无监督的目标



注:从上向下分别是SpaceNet6、SSDD、HRSID检测结果;各行从左至右分别是真实标注、DA-faster、D-Adapt、AT、CMT、DAI-Net、PMT和PCKT的结果;图中绿色框表示正确判决、黄色框表示虚警(假阳判决)、红色框表示漏检测(假阴判决)。

图13 跨模态知识迁移目标检测结果可视化



注:从上向下分别是SpaceNet6、SSDD、HRSID检测结果;各行从左至右分别是真实标注、CA-CFAR、TP-CFAR、IS-CFAR、Two-Stage-SP-CFAR、PCKT的结果;图中绿色框表示正确判决、黄色框表示虚警(假阳判决)、红色框表示漏检测(假阴判决)。

图14 与传统方法的比较

检测算法. 通过构建“光学域→中间域→SAR域”的渐进式迁移框架,设计了多尺度特征对齐模块与加权任务损失平衡机制,有效缓解了跨模态域间特征分布差异. 引入自训练伪标签学习策略,缓解目标检测器对源域数据的偏离. 最后,本文利用SpaceNet6、SSDD和HRSID等数据集,开展了同源知识迁移和异源知识迁移的实验验证. 结果表明,本文所提方法均取得优异的目标检测性能. 究其原因,所设计的分阶段知识迁移策略能显著改善域间特征对齐效果,提出的多尺度特征对齐策略可显著提升小目标检测精度,通过t-SNE可视化分析可知SAR目标域与可见光源域的特征分布更加聚集,自训练优化策略可使伪标签mIoU大幅提升. 因此,所提方法为解决稀疏样本分布条件下的SAR图像解译问题提供了有效的技术支撑.

参考文献

- [1] 孙琦钰, 赵超强, 唐漾, 等. 基于无监督域自适应的计算机视觉任务研究进展[J]. 中国科学: 技术科学, 2022, 52(1): 26-54.
SUN Q Y, ZHAO C Q, TANG Y, et al. A survey on unsupervised domain adaptation in computer vision tasks[J]. Scientia Sinica (Technologica), 2022, 52(1): 26-54. (in Chinese)
- [2] 郭倩, 王海鹏, 徐丰. SAR图像飞机目标检测识别进展[J]. 雷达学报, 2020, 9(3): 497-513.
GUO Q, WANG H P, XU F. Research progress on aircraft detection and recognition in SAR imagery[J]. Journal of Radars, 2020, 9(3): 497-513. (in Chinese)
- [3] 方澄, 管方恒, 李天驰, 等. 端到端全复数域SAR目标分类神经网络[J]. 电子学报, 2024, 52(7): 2449-2460.
FANG C, GUAN F H, LI T C, et al. End-to-end full complex-valued domain SAR target classification neural network[J]. Acta Electronica Sinica, 2024, 52(7): 2449-2460. (in Chinese)
- [4] DE MAIO A, AUBRY A. Radar detection, performance analysis, and CFAR techniques[C]//2019 IEEE Radar Conference. Piscataway: IEEE, 2019: 1-120.
- [5] 罗会兰, 陈鸿坤. 基于深度学习的目标检测研究综述[J]. 电子学报, 2020, 48(6): 1230-1239.
LUO H L, CHEN H K. Survey of object detection based

- on deep learning[J]. *Acta Electronica Sinica*, 2020, 48(6): 1230-1239. (in Chinese)
- [6] 刘颖, 刘红燕, 范九伦, 等. 基于深度学习的小目标检测研究与应用综述[J]. *电子学报*, 2020, 48(3): 590-601.
LIU Y, LIU H Y, FAN J L, et al. A survey of research and application of small object detection based on deep learning[J]. *Acta Electronica Sinica*, 2020, 48(3): 590-601. (in Chinese)
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 580-587.
- [8] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2016: 1440-1448.
- [9] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 779-788.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multiBox detector[M]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [12] LIN Z, JI K F, LENG X G, et al. Squeeze and excitation rank faster R-CNN for ship detection in SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(5): 751-755.
- [13] ZHAO Y, ZHAO L J, XIONG B L, et al. Attention receptive pyramid network for ship detection in SAR images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 2738-2756.
- [14] KANG M, JI K F, LENG X G, et al. Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection[J]. *Remote Sensing*, 2017, 9(8): 860.
- [15] 夏勇, 田西兰, 常沛, 等. 基于深度学习的复杂沙漠背景 SAR 目标检测[J]. *雷达科学与技术*, 2019, 17(3): 305-309, 318.
XIA Y, TIAN X L, CHANG P, et al. SAR target detection in complex desert background images based on deep learning[J]. *Radar Science and Technology*, 2019, 17(3): 305-309, 318. (in Chinese)
- [16] 谢兆哲, 程永强, 吴昊, 等. 基于 Toeplitz 矩阵特征值分解的 SAR 图像舰船目标检测方法[J]. *信号处理*, 2023, 39(3): 496-504.
XIE Z Z, CHENG Y Q, WU H, et al. Ship target detection method in SAR imagery based on eigenvalue decomposition of the toeplitz matrix[J]. *Journal of Signal Processing*, 2023, 39(3): 496-504. (in Chinese)
- [17] 程旭东, 史彩娟, 高炜翔, 等. 面向域自适应目标检测的一致无偏教师模型[J]. *图学学报*, 2025, 46(1): 114-125.
CHENG X D, SHI C J, GAO W X, et al. Consistent and unbiased teacher model research for domain adaptive object detection[J]. *Journal of Graphics*, 2025, 46(1): 114-125. (in Chinese)
- [18] 张鸿杰, 李昂, 过洁, 等. 基于风格迁移及对抗加权的开放集域自适应[J]. *计算机科学技术学报*, 2023, 38(3): 644-658.
- [19] 胡逸雯, 刘鑫, 匡翠方, 等. 基于深度学习的自适应光学技术研究进展及展望[J]. *中国激光*, 2023, 50(11): 1101009.
HU Y W, LIU X, KUANG C F, et al. Research progress and prospect of adaptive optics based on deep learning[J]. *Chinese Journal of Lasers*, 2023, 50(11): 1101009. (in Chinese)
- [20] 刘建伟, 孙正康, 罗雄麟. 域自适应学习研究进展[J]. *自动化学报*, 2014, 40(8): 1576-1600.
LIU J W, SUN Z K, LUO X L. Review and research development on domain adaptation learning[J]. *Acta Automatica Sinica*, 2014, 40(8): 1576-1600. (in Chinese)
- [21] 周洋, 韩冰, 高新波, 等. 基于注意力机制和循环域三元损失的域自适应目标检测[J]. *自动化学报*, 2024, 50(11): 2188-2203.
ZHOU Y, HAN B, GAO X B, et al. Domain adaptive object detection based on attention mechanism and cycle domain triplet loss[J]. *Acta Automatica Sinica*, 2024, 50(11): 2188-2203. (in Chinese)
- [22] 胡杰, 徐博远, 熊宗权, 等. 基于多尺度掩码分类域自适应网络的跨域目标检测算法[J]. *汽车工程*, 2022, 44(9): 1327-1338.
HU J, XU B Y, XIONG Z Q, et al. Cross-domain object detection algorithm based on multi-scale mask classification domain adaptive network[J]. *Automotive Engineering*, 2022, 44(9): 1327-1338. (in Chinese)
- [23] 李威, 王蒙. 基于渐进多源域迁移的无监督跨域目标检测[J]. *自动化学报*, 2022, 48(9): 2337-2351.
LI W, WANG M. Unsupervised cross-domain object detection based on progressive multi-source transfer[J]. *Acta Automatica Sinica*, 2022, 48(9): 2337-2351. (in Chinese)

- [24] KHODABANDEH M, VAHDAT A, RANJBAR M, et al. A robust learning approach to domain adaptive object detection[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 480-490.
- [25] KIM S, CHOI J, KIM T, et al. Self-training and adversarial background regularization for unsupervised domain adaptive one-stage object detection[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2019: 6091-6100.
- [26] FRENCH G, MACKIEWICZ M, FISHER M. Self-ensembling for visual domain adaptation[EB/OL]. (2018-09-23)[2025-03-25]. <https://arXiv.org/abs/1706.05208>.
- [27] SOHN K, BERTHELOT D, CARLINI N, et al. Fixmatch: Simplifying semi-supervised learning with consistency and confidence[J]. *Advances in Neural Information Processing Systems*, 2020, 33: 596-608.
- [28] LI Y J, DAI X L, MA C Y, et al. Cross-domain adaptive teacher for object detection[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 7571-7580.
- [29] CAO S C, JOSHI D, GUI L Y, et al. Contrastive mean teacher for domain adaptive object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 23839-23848.
- [30] BAI Y F, WU Y Q, ZHU B, et al. Contrastive-domain mean teacher for domain adaptive object detection[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2025: 3575205.
- [31] CHEN L, HAN J H, WANG Y P. DATR: Unsupervised domain adaptive detection transformer with dataset-level adaptation and prototypical alignment[J]. *IEEE Transactions on Image Processing*, 2025, 34: 982-994.
- [32] CHEN Y H, LI W, SAKARIDIS C, et al. Domain adaptive faster R-CNN for object detection in the wild[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 3339-3348.
- [33] SAITO K, USHIKU Y, HARADA T, et al. Strong-weak distribution alignment for adaptive object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 6949-6958.
- [34] ZHANG H, LUO G Y, CAO Y, et al. Scale-disentangled and uncertainty-guided alignment for domain-adaptive object detection[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2024, 25(12): 19507-19521.
- [35] BELAL A, MEETHAL A, ROMERO F P, et al. Multi-source domain adaptation for object detection with prototype-based mean teacher[C]//2024 IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2024: 1266-1275.
- [36] ARRUDA V F, PAIXÃO T M, BERRIEL R F, et al. Cross-domain car detection using unsupervised image-to-image translation: From day to night[C]//2019 International Joint Conference on Neural Networks. Piscataway: IEEE, 2019: 1-8.
- [37] CHEN M, RADFORD A, CHILD R, et al. Generative pretraining from pixels[C]//Proceedings of the 37th International Conference on Machine Learning. New York: ACM, 2020: 1691-1703.
- [38] HSU H K, YAO C-H, TSAI Y H, et al. Progressive domain adaptation for object detection[C]//2020 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2020: 738-746.
- [39] SHI Y, DU L, GUO Y C. Unsupervised domain adaptation for SAR target detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 6372-6385.
- [40] DU Z P, SHI M, DENG J K. Boosting object detection with zero-shot day-night domain adaptation[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 12666-12676.
- [41] 姚琴凤, 宁永香, 杜孙稳. 基于域自适应神经网络的光学与合成孔径雷达遥感图像变化检测[J]. *激光与光电子学进展*, 2024, 61(18): 255-264.
- YAO Q F, NING Y X, DU S W. Change detection of optical and synthetic aperture radar remote sensing images based on a domain adaptive neural network[J]. *Laser & Optoelectronics Progress*, 2024, 61(18): 255-264. (in Chinese)
- [42] 季鼎承, 蒋亦樟, 王士同. 基于域与样例平衡的多源迁移学习方法[J]. *电子学报*, 2019, 47(3): 692-699.
- JI D C, JIANG Y Z, WANG S T. Multi-source transfer learning method by balancing both the domains and instances[J]. *Acta Electronica Sinica*, 2019, 47(3): 692-699. (in Chinese)
- [43] 余游, 冯林, 王格格, 等. 一种基于伪标签的半监督少样本学习模型[J]. *电子学报*, 2019, 47(11): 2284-2291.
- YU Y, FENG L, WANG G G, et al. A few-shot learning model based on semi-supervised with pseudo label[J]. *Acta Electronica Sinica*, 2019, 47(11): 2284-2291. (in Chinese)
- [44] TIAN Z Z, WANG W, ZHOU K, et al. Weighted pseudo-labels and bounding boxes for semisupervised SAR target detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, 17: 5193-

- 5203.
- [45] VAN DER MAATEN L, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9(86): 2579-2605.
- [46] SUN B C, SAENKO K. Deep CORAL: Correlation alignment for deep domain adaptation[M]//Computer Vision-ECCV 2016 Workshops. Cham: Springer International Publishing, 2016: 443-450.
- [47] LITTLE TODD D. The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2: Statistical Analysis[M]. Oxford: Oxford University Press, 2013: 458-485.
- [48] SHERMEYER J, HOGAN D, BROWN J, et al. SpaceNet 6: Multi-sensor all weather mapping dataset[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2020: 768-777.
- [49] ZHANG T W, ZHANG X L, LI J W, et al. SAR ship detection dataset (SSDD): Official release and comprehensive data analysis[J]. Remote Sensing, 2021, 13(18): 3690.
- [50] WEI S J, ZENG X F, QU Q Z, et al. HRSID: A high-resolution SAR images dataset for ship detection and instance segmentation[J]. IEEE Access, 2020, 8: 120234-120254.
- [51] CHEN C Q, ZHENG Z B, DING X H, et al. Harmonizing transferability and discriminability for adapting object detectors[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 8866-8875.
- [52] JIANG J, CHEN B, WANG J, et al. Decoupled adaptation for cross-domain object detection[C]//The Tenth International Conference on Learning Representations. Virtual Event: ICLR, 2022: 1-16.

作者简介



赵国威 男, 2000年3月出生于山东省枣庄市. 现为西安电子科技大学信息与通信工程专业硕士研究生. 主要研究方向为雷达信号处理、目标检测识别.
Email: gwzhao0321@163.com



蒋嘉庆 男, 2002年10月出生于河北省任丘市. 现为西安电子科技大学信息与通信工程专业硕士研究生. 主要研究方向为SAR干扰抑制与目标检测.
E-mail: 15613785520@163.com



董刚刚 男, 1982年11月出生于陕西省西安市. 现为西安电子科技大学教授、博士生导师. 2012年和2016年获得国防科技大学工学硕士与博士学位. 主要研究方向为雷达成像、目标检测识别、遥感图像解译等.
E-mail: dongganggang@nudt.edu.cn