DOI: 10.13382/j.jemi.B2306867

基于半监督域适应的微弱光环境下行人检测研究*

杜运亮 王明甲
(青岛科技大学自动化与电子工程学院 青岛 266100)

摘 要:为了解决可见光图像在微弱光环境下会出现检测性能下降的问题,本文提出了一种半监督域适应的行人检测算法。首先,结合均值教师模型和 YOLOv8 检测器搭建半监督检测网络;其次,使用图像融合算法和风格迁移算法相结合的方式生成伪图像进行伪交叉训练,减少图像之间的域差异问题;最后,将基于 Transform 的混合注意力机制引入主干特征提取网络,在提升图像分辨率的同时进一步提升检测精度。实验结果表明:在 LLVIP 数据集和 KAIST 数据集上,该算法的检测精度分别达到 89.3%和 66.8%,相比 SSDA-YOLO 算法分别高出 7.6%和 19.8%;相比 Efficient Teacher 算法分别高出 4%和 8.7%;相比全监督算法 ICAFusion 分别高出 1.8%和 17.9%。与以往的算法相比,该算法具有更高的检测精度。

关键词:行人检测;微弱光;半监督;域适应;YOLOv8
中图分类号:TP391.4;TN215 **文献标识码**:A **国家标准学科分类代码**:520.2

Research on pedestrian detection in low-light conditions based on semi-supervised domain adaptation

Du Yunliang Wang Mingjia
(College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266100, China)

Abstract: In order to solve the problem that visible light images suffer from performance degradation in low-light conditions, this paper proposes a semi-supervised domain-adapted pedestrian detection algorithm. Firstly, the semi-supervised detection network is built by combining the mean teacher model and the YOLOv8 detector. Secondly, pseudo-images are generated for pseudo-cross-training using a combination of the image fusion algorithm and the style migration algorithm to reduce the problem of domain difference between images. Finally, the hybrid attention mechanism based on Transform is introduced into the backbone feature extraction network, which further improves the detection accuracy while enhancing the image resolution. The experimental results show that the detection accuracy of the algorithm reaches 89.3% and 66.8% on the LLVIP dataset and KAIST dataset, respectively, which is 7.6% and 19.8% higher compared to the SSDA-YOLO algorithm, 4% and 8.7% higher compared to the Efficient Teacher algorithm, and 1.8% and 17.9% compared to the fully supervised algorithm ICAFusion. Compared with previous algorithms, this algorithm has higher detection accuracy.
Keywords: pedestrian detection; low-light; semi-supervised; domain adaptation; YOLOv8

0 引 言

行人检测是利用计算机视觉技术判断图像或视频序列中是否存在行人并给予精确定位,是自动驾驶技术的重要研究方向。随着自动驾驶技术的发展,无人驾驶车辆必须精确感知车辆周围的环境,特别是对行人的感知,以保证行车安全。因此,准确、低延时的行人检测已经成为自动驾驶领域的一个重要问题。

传统行人检测算法在微弱光环境下的检测效果差强人意,研究人员为提高微弱光环境下行人检测算法的性能做出了巨大的努力。目前利用可见光和红外热成像的多光谱行人检测算法已经引起越来越多专家学者的研究关注,因为可见光和红外热成像可以提供更多有价值的轮廓信息,能够提高行人检测算法的准确性和鲁棒性[1]。一些研究机构将多模态行人检测算法采用可见光和红外

收稿日期:2023-09-04 Received Date:2023-09-04
*基金项目:国家自然科学基金(61971253)、山东省自然科学基金(ZR2014FL026)项目资助

热成像并行双流模式。利用可见光和红外热成像信息的互补性,将它们在模型训练前进行图像融合的预训练,然后将融合后的图像输送到行人检测训练网络进行训练,可以获得性能较好的模型[2-3]。文献[4]在 YOLOv3 算法中增加特征尺度,并将 BN 网络层与卷积神经网络层融合,提高红外图像检测精度。文献[5]通过引入 MobileNet 模型增强 Tiny-YOLOv3 的特征提取网络,提升模型高级语义特征提取能力,提高红外图像缺陷检测的速度和准确率。文献[6]设计了一种空间通道混合注意力模块,并在输出检测头处进行反卷积操作扩大输出特征图,提高了红外图像小物体检测的精度。Bao 等[7]从热成像方面入手,提出一种热辅助探测和测距(HADAR)技术,它能利用人工智能解析热信号,从而在黑暗或低能见度的环境中,清晰地识别出物体的温度、材质和纹理,HADAR 技术能够提供更多的物理信息,不仅包括温度,还包括发射率和深度。尽管所提出的结构可以有效地提高微弱光环境下行人检测的精度,但是利用两种模态的信息进行融合并同时提升检测速度的挑战仍然存在。

为了解决上述问题,本文针对微弱光环境的行人检测,提出了一种半监督域适应的行人检测算法,引入目前最先进的 YOLOv8 检测器以提高行人检测性能。

本文的主要工作如下:

1)为了解决微弱光环境下的行人检测问题,提出了一种基于 YOLOv8 的半监督域适应行人检测算法,将 YOLOv8 单级检测器与均值教师(mean teachers,MT)模型的知识蒸馏框架相结合。

2)为了解决红外图像和可见光图像域之间的差异问题,使用 DIVFusion 算法和 CycleGAN 算法相结合的方式生成更高质量的伪图像。

3)为了提高 YOLOv8 检测器的检测精度,在 YOLOv8 网络中增加了基于 Transform 的混合注意力机制(hybrid attention transformer,HAT),提升训练图片的分辨率,激活更多有用像素。

4)在 LLVIP 数据集和 KAIST 数据集中验证了本文所提出方法的有效性。

1　相关工作

1.1　微弱光环境下的行人检测

行人检测在智能驾驶、安全监控和人流量监测等领域都具有非常重要的实际意义,特别是在开发高级驾驶辅助系统和自动驾驶方面,检测行人的能力对于确保研究应用的安全性能至关重要。近年来,微弱光环境下的行人检测一直是专家学者重点研究的领域,文献[8]引入动态缩放损失函数和坐标注意力机制,并设计出 MultiS-C3 模块,提升小尺寸车辆与行人的检测精度。Zhang 等[9]提出了一种称为跨模态交互式注意网络(CAN)的新型多光谱行人检测算法,旨在利用 RGB 和热图像数据之间的相关性来提高检测精度。Zhao 等[10]提出了一种基于红外视觉信息和毫米波雷达数据融合的夜间行人检测方法,该方法使用毫米波雷达图像和红外图像生成多模态目标信息,经过加权处理实现精准的目标定位。

1.2　域适应目标检测

域适应目标检测旨在解决域转移问题,提高源域和目标域之间的相似性。Deng 等[11]设计了一种用于跨域目标检测的无偏教师模型,使用了一种均值教师模型的跨域蒸馏方法,以最大限度地利用教师模型,针对学生模型,通过像素级自适应增加训练样本来减小偏差。Vibashan 等[12]提出类别感知域适应的记忆引导注意算法,将类别信息合并到域适应过程中,采用类别判别器来确保类别感知特征对齐,以学习领域不变判别特征方法。Chen 等[13]将无源无监督域适应问题建模为从噪声标签中学习问题,提出了一种新颖的自监督噪声标签学习方法,使用预生成的标签以及动态自生成的标签有效地微调预训练模型。

1.3　半监督域适应

相比传统的无监督域适应任务中目标域无标签的设定,半监督域适应任务只需要给定少量的标签样本,性能就能够获得大幅度的提升。Zhou 等[14]提出了一种基于半监督域自适应 YOLO 检测的新方法,通过将紧凑的单级检测器 YOLOv5 与域自适应相结合来提高跨域检测性能。Xu 等[15]提出了一个高效教师模型,用于可扩展且有效的一阶段基于锚点的半监督域适应训练,设计了一种新颖的伪标签分配器实现伪标签分配机制,它也可以有效防止师生互相学习机制中出现大量低质量的伪标签,导致检测器出现偏差。Yang 等[16]把半监督域适应分解成无监督域适应和半监督学习两个任务分别进行学习,在两个任务学习过程中,无监督域适应模型得到目标域中无标签数据的伪标签,送给半监督学习任务来更新模型参数,反之亦然,通过这种交叉协同训练学习各自的任务模型参数,确定数据的类别。

2　检测方法

本文提出的方法主要包含 4 个部分:具有用于指导学生模型更新知识蒸馏框架的均值教师模型、用于减轻红外图像和可见光图像域之间差异的图像风格迁移、用于补救图像风格迁移信息丢失的图像融合和用于提高检测精度的 YOLOv8 检测器的改进,如图 1 所示。对于域适应目标检测任务需要解决两个不同域之间的差异问

题,本文解决的是可见光图像和红外图像之间的域适应问题,将容易标注的红外图像作为源域 S ,将不易标注的可见光图像作为目标域 T 。

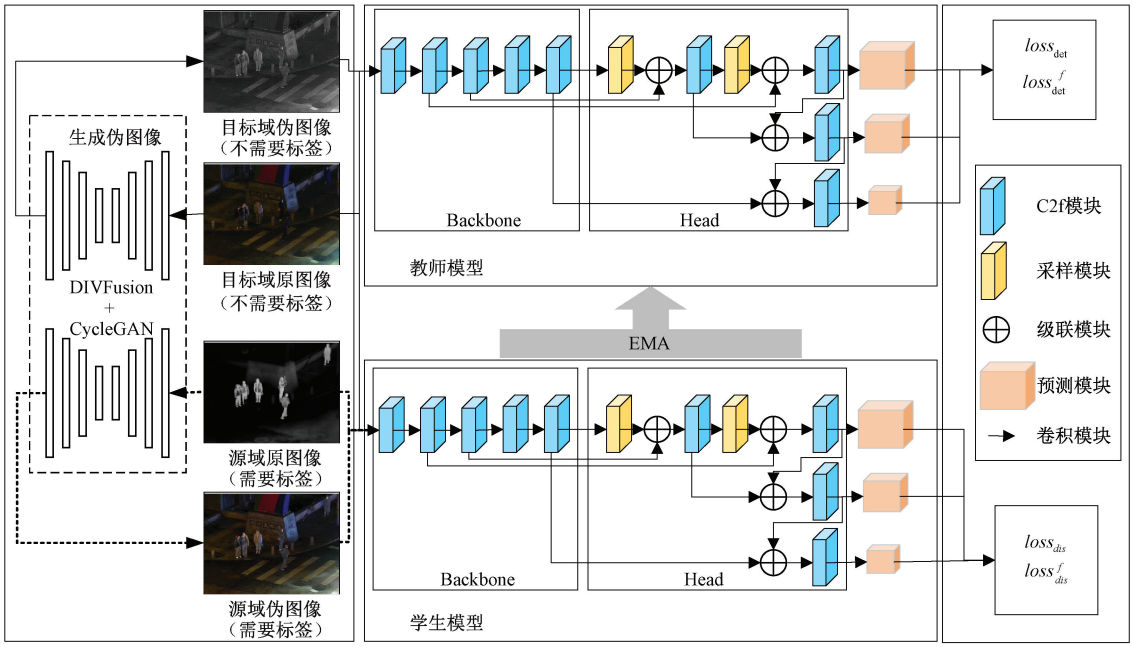


图 1　网络结构图

Fig. 1　Network diagram

2.1　YOLOv8 检测器

YOLOv8 * 融入了许多新的修订和技术,以提高整体的检测精度。YOLOv8 主要有两个部分:Head 和 Backbone。Backbone 模块负责提取不同尺度的多层特征并将这些特征传入 Head 模块进行目标检测,最终输出目标检测信息。本文使用 YOLOv8 作为网络的检测器,对于整个网络来说是至关重要的,决定了网络的检测精度问题。为此本文对原始 YOLOv8 做了部分改进,增加了 HAT 注意力机制[17]。HAT 结合了通道注意力机制和基于窗口的自注意力机制,如图 2 所示,充分利用它们在全局统计信息和强大的局部拟合能力方面的互补优势,目的是从给定的低分辨率图形输入中重建高分辨率图像,激活更多有用像素。

本文使用的是半监督框架,引入 YOLOv8 作为网络检测器,针对源域图像使用监督模型进行训练,则相应的检测损失函数如式(1)所示:

$$loss_{\det} = loss_{box}(S;L^{s}) + loss_{cls}(S;C^{s}) \tag{1}$$

式中:L^{s} 表示源域图像边界框集合,C^{s} 表示源域图像类别标签集合,$loss_{box}$ 表示预测边界框回归损失是分布聚焦(DFL)损失和完全交集(CIoU)损失的结合,$loss_{cls}$ 表示分类损失的二元交叉熵(BCE)损失。

2.2　均值教师模型

MT 模型起初是为图像分类的半监督学习提出的一种方法,它由经典的知识蒸馏结构和两个相同的教师和学生模型组成[18]。对于域适应任务,使用梯度下降优化器在源域中的标记数据中训练学生模型。根据 MT 模型设置,教师模型通过学生模型的指数移动平均权重(EMA)进行更新,如图 1 所示。

将目标域图像分别输入到学生模型和教师模型中进行训练,在知识蒸馏过程中,通过从教师模型预测中获得的高置信度边界框作为伪标签,使学生模型减少在目标域图像中产生的偏差并增强模型的鲁棒性。将输入教师模型的目标域图像 T_t 和输入学生模型的目标域图像 T_s 进行预测,两个模型预测之间的差别可以使用式(2)的蒸馏损失来更新权重:

$$loss_{dis} = loss_{\det}\{T_s, G_{box}[L_{box}(T_t)], G_{cls}[L_{cls}(T_t)]\} \tag{2}$$

式中:$L_{box}(T_t)$ 表示来自教师模型的边界框预测分支,$L_{cls}(T_t)$ 表示来自教师模型的类别预测分支,G_{box},G_{cls} 分别是相应的过滤器。在训练期间使用非极大抑制(NMS)以 IoU 阈值为界来过滤预测边界框,最终提供实例级特征的伪标签给学生模型进行训练。

2.3　伪图像生成

域适应需要减小图像之间的域差异问题,使用伪图

* https://github.com/ultralytics/ultralytics

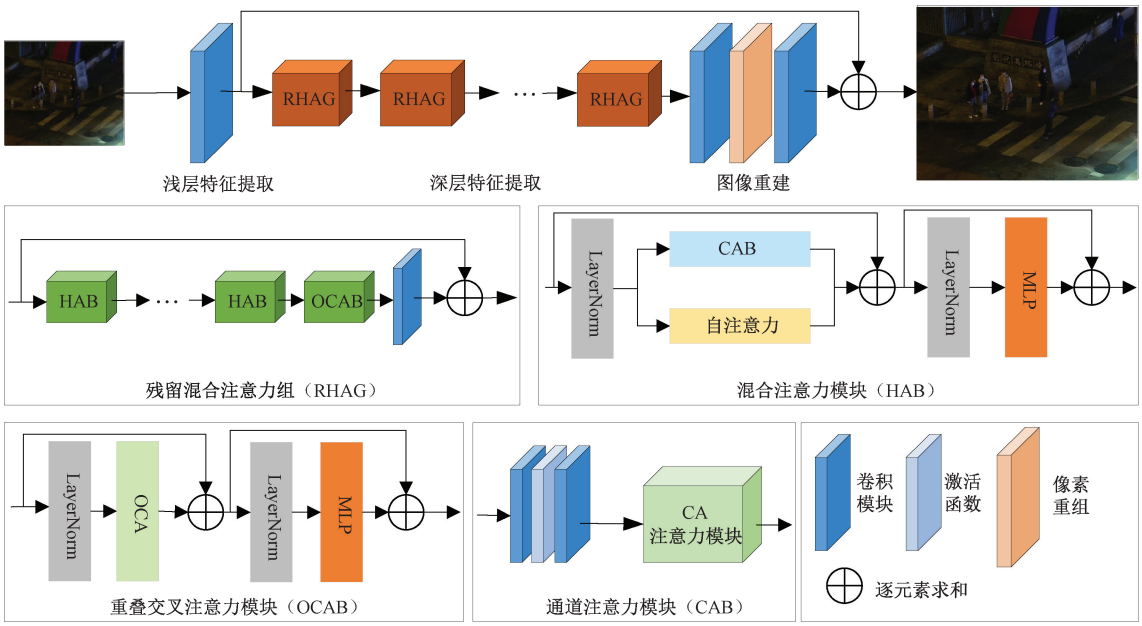


图 2　HAT 结构图[17]

Fig. 2　HAT structure diagram[17]

像进行伪交叉训练来减小域差异问题。本文是针对红外图像和可见光图像之间的域适应问题。由于可见光图像信息缺失严重,而红外图像色彩单一的问题,使用传统的 CycleGAN[19] 算法生成伪图像显然是不合理的。DIVFusion[20] 可以使红外图像信息融合到可见光图像中,为可见光图像提供更多的细节信息;CycleGAN 可实现可见光图像和红外图像之间的转换。本文结合 DIVFusion 和 CycleGAN 进行伪图像的生成,提取 DIVFusion 算法中的场景照明解缠网络(SIDNet)和纹理对比度增强融合网络(TCEFNet),将 SIDNet 和 TCEFNet 与 CycleGAN 算法中的图像输入部分相结合,在图像输入之前经过 SIDNet 和 TCEFNet 进行图像融合的处理,然后将融合后的图像输入到 CycleGAN 网络中进行下一步迁移步骤。分别将源图像和目标图像生成的伪图像称为伪源图像 S_f 和伪目标图像 T_f,生成的部分伪图像如图 3 所示。使用 DIVFusion 网络进行图像融合可以使红外图像和可见光图像之间增加更多的细节信息,可以根据这些细节信息生成更加精确地伪图像;CycleGAN 算法中的判别器网络进行对抗训练,帮助减小伪图像差异之间域差异问题,使生成的图像与原图像之间差异性更小。

2.4　纠正伪图像差异

使用生成的伪图像进行训练会产生跨域之间的偏差。为了减小学生模型因伪图像差异而引起的偏差,使用学生模型训练假源域图像,并使假源域图像与源域图像标签相同。则相应的监督损失函数如式(3)所示:

$$loss^{f}_{\det} = loss_{box}(S_f; B^{s}) + loss_{cls}(S_f; C^{s}) \tag{3}$$

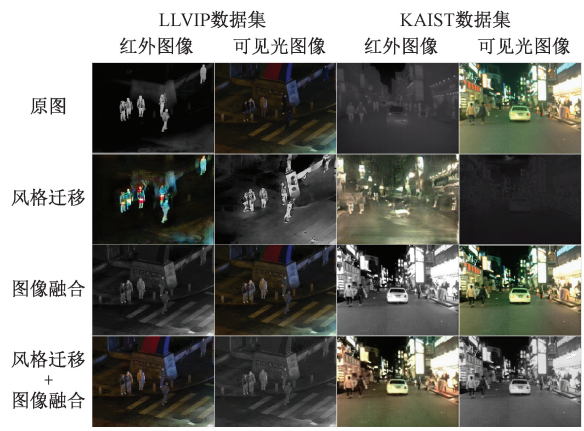


图 3　部分生成伪图像展示

Fig. 3　Partial generation of pseudo-images

对于假目标域图像使用教师模型进行训练。为了使其学习源域的全局图像特征,将假目标域图像替换目标域图像,用于训练学生模型的部分为标记图像保持不变,则新的蒸馏损失可如式(4)所示:

$$loss^{f}_{dis} = loss_{\det}\{T_s, G_{box}[L_{box}(T^{f}_{t})], G_{cls}[L_{cls}(T^{f}_{t})]\} \tag{4}$$

3　实验与讨论

本实验使用的操作系统是 Ubuntu20.04,CPU 型号为 Inter(9) Core(TM) i9-10900K CPU@3.70 GHz,GPU 型号为 GeForce RTX3090,深度学习框架为

Pytorch1.10.0,CUDA11.0 版本搭配 cuDNN8.0.5 版本神经网络加速库使用。

3.1 数据集

本文使用的是红外图像和可见光图像配对的 LLVIP 数据集[21]和 KAIST 数据集[22]。LLVIP 行人数据集包含 30 976 张图像,即 15 488 对,该数据集在时间和空间上严格对齐,图像分辨率为 1 280×1 024,并且数据集中的行人已被标注,所以该数据集不需要进行数据清洗[21]。KAIST 行人数据集总共包括 95 328 张图片,每张图片都包含 RGB 图像和红外图像两个版本,总共包含 103 128 个密集注释[22]。多数使用 KAIST 数据集的论文都经过清洗,因为该数据集是取自视频连续帧图片,相邻图片相差不大,故进行一定程度的清洗[23]。本文主要研究微弱光环境下的行人检测,所以只清洗 KAIST 的夜晚数据集来进行实验,清洗规则为:1)训练集每隔 2 张图片取一张,既每 3 张取一张,并去掉所有不包含任何行人的图片,既选出来的图片中至少包含一个目标,且剔除数据集中只漏出半身或者小于 50 个像素等严重遮挡的图片。2)测试集每隔 19 张取一张,既每 20 张取一张,保留负样本图片。经过上述操作,可以得到 2 846 张训练集图片和 797 张测试集图片。

3.2 评价指标

为了有效评估模型的检测结果,使用精确率(precision,P)、召回率(recall,R)和平均精确率均值(mean average precision,mAP)来评估模型的性能,其中 mAP 选取交并比(intersection over union,IoU)为 0.5 和 0.5~0.95 来评价模型的检测结果,其计算方法如式(5)~(8)所示:

$$P = \frac{TP}{TP + FP} \tag{5}$$

$$R = \frac{TP}{TP + FN} \tag{6}$$

$$AP = \int_0^1 P(R)\,\mathrm{d}R \tag{7}$$

$$mAP = \frac{\sum_{i=1}^{k} AP_i}{k} \tag{8}$$

式中:TP(true positives)为模型成功检测出正样本的数量,FP(false positives)为模型错误的将负样本检测为正样本的数量,FN(false negative)为模型没有检测出的正样本的数量,k 表示类别的数量,AP(average precision)为平均准确率,AP_i 为第 i 个类别的 AP 值。

3.3 对比实验

本文将所提出的方法与 SSDA-YOLO[14]、Efficient Teacher[15]和 ICAFusion[24]做对比实验。SSDA-YOLO 和 Efficient Teacher 算法均是近年来针对半监督域适应领域进行研究,不同的是这两种算法针对的是白天场景中的目标检测。SSDA-YOLO 是一种基于目标检测的域自适应方法,它将源域数据和目标域数据进行混合训练,从而在目标域上实现模型的迁移学习。实验证明,该算法在目标域上取得了较好的性能。Efficient Teacher 是一种高效的知识蒸馏方法,用于改进目标检测模型的性能,通过利用教师模型的知识来指导学生模型的训练。实验证明,该算法在准确率和速度方面都取得了显著的改进。ICAFusion 算法是一种全监督网络,针对微弱光环境进行目标检测研究,该方法采用了一种双交叉注意力变换器的新颖特征融合框架来建模全局特征交互并同时捕获跨模态的互补信息。实验证明,该算法具有优越的性能和更快的推理速度,使其适用于各种实际场景。本文主要研究微弱光环境下红外图像和可见光图像之间的域适应问题,所以使用 SSDA-YOLO、Efficient Teacher 和 ICAFusion 网络进行重新学习来与本文所提出的算法相比较。表 1 和 2 分别展示了在 LLVIP 和 KAIST 数据集中的对比结果。图 4 中展示了使用训练模型分别在 LLVIP 数据集和 KAIST 数据集测试的效果。

表 1 LLVIP 对比结果

Table 1 Results for comparison of LLVIP (%)

方法	P	R	mAP_{50}	$mAP_{50:95}$
仅源域	6.21	2.43	1.98	0.634
SSDA-YOLO	83.4	78.7	81.7	38.2
Efficient Teacher	90.2	81.7	85.3	42.8
ICA-Fusion	95.3	80.6	87.5	43.6
本文	93.7	79.8	89.3	47.2

表 2 KAIST 对比结果

Table 2 Results for comparison of KAIST (%)

方法	P	R	mAP_{50}	$mAP_{50:95}$
仅源域	3.24	1.03	0.312	0.149
SSDA-YOLO	43.3	38.9	47.0	12.6
Efficient Teacher	54.1	49.2	58.1	23.2
ICA-Fusion	85.2	45.0	48.9	22.3
本文	77.7	57.4	66.8	32.9

3.4 消融实验

为了验证本文改进使用图像融合技术和风格迁移技术共同来完成伪图像生成的任务,设计了消融实验,并同时验证增加 HAT 带来的效果。分别在 LLVIP 和 KAIST 数据集中进行消融实验,实验结果如表 3 和 4 所示,图 5 展示了消融实验的训练数据对比。

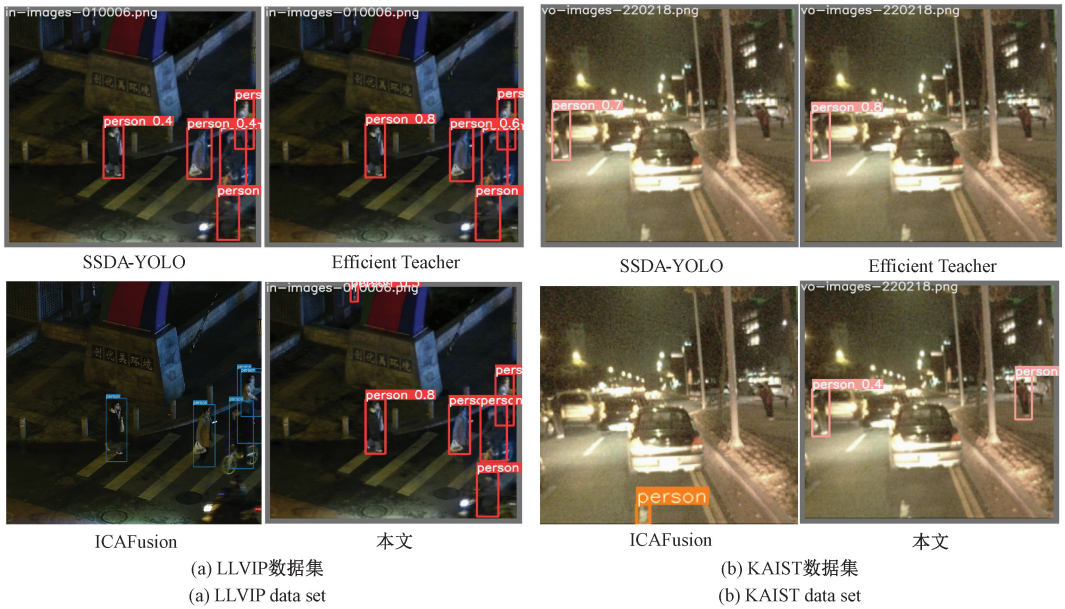


图 4　模型测试展示

Fig. 4　Model test demonstration

表 3　LLVIP 消融结果

Table 3　Results for ablation of LLVIP

方法	DIVFusin	CycleGAN	HAT	P/%	R/%	mAP_{50} /%	$mAP_{50:95}$ /%
MT-YOLOv8	√	×	×	91.2	84.2	85.2	39.3
	×	√	×	84.8	62.4	72.8	36.3
	√	√	×	91.7	85.2	86.2	43.7
	√	×	√	90.8	83.3	86.0	42.4
	×	√	√	84.7	64.6	74.6	38.9
	√	√	√	93.7	86.8	89.3	47.2

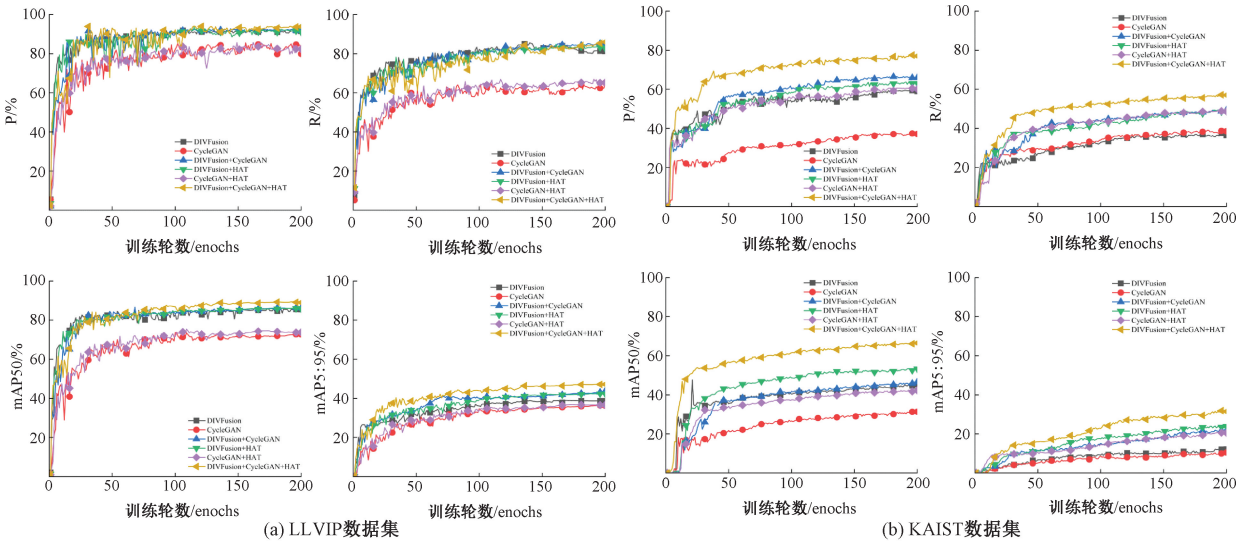
表 4　KAIST 消融结果

Table 4　Results for ablation of KAIST

方法	DIVFusin	CycleGAN	HAT	P/%	R/%	mAP_{50} /%	$mAP_{50:95}$ /%
MT-YOLOv8	√	×	×	59.6	37.7	45.6	12.9
	×	√	×	37.4	39.2	31.4	10.7
	√	√	×	66.6	49.8	46.9	22.6
	√	×	√	63.6	48.9	53.7	23.3
	×	√	√	60.2	49.2	42.3	20.8
	√	√	√	77.7	57.4	66.8	32.9

3.5　实验结果分析

为了全面验证本文方法的有效性，与目前最先进的方法进行定性比较，如表 1 和 2 中数据所示，本文方法相对于 SSDA-YOLO、Efficient Teacher 和 ICAFusion 的在 LLVIP 数据集中 mAP_{50} 分别提高了 7.6%、19.8% 和 1.8%，在 KAIST 数据集中分别提高了 4%、8.7% 和 17.9%；$mAP_{50:95}$ 在 LLVIP 数据集中分别提高了 9%、20.3% 和 3.6%，在 KAIST 数据集中分别提高了 4.4%、9.7% 和 10.6%，提升较为明显。在 P 方面，本文算法相对于两种半监督算法 SSDA-YOLO 和 Efficient Teacher 提升较为明显，但是相对于全监督算法 ICAFusion，本文算法略显不足，下一步需要对本文算法做出进一步的效果提升。如表 3 和 4 中数据所示，使用图像融合加风格迁移的方法能够更好的生成伪图像，并且 HAT 提高图像的分辨率，激活更多有用的像素，进一步提升检测精度。图 4 展示了本文方法相比于 SSDA-YOLO、Efficient Teacher 和 ICAFusion 能够检测出更多的目标。图 5 展示出使用 DIVFusion 和 CycleGAN 生成伪图像、YOLOv8 检



(a) LLVIP数据集　(b) KAIST数据集

(a) LLVIP data set　(b) KAIST data set

图5　训练数据对比

Fig. 5　Comparison of training data

测器和 HAT 注意力机制的有效性。

通过表 1~4 中的数据可以发现,本文提出的算法在 LLVIP 数据集上的表现比 KAIST 数据集更加优异,原因是这两个数据集的图像质量差距较大。LLVIP 数据集分辨率相对于 KAIST 数据集分辨率更高,且 KAIST 数据集是在更加黑暗的场景中拍摄的,有些可见光图像中的目标肉眼也无法发现。

4　结　论

本文针对微弱光环境提出了一种更为先进的基于半监督域适应的行人检测算法,使用了先进的 YOLOv8 目标检测网络作为骨干检测器。本文方法包含 3 个部分:首先,基于知识蒸馏结构,分别使用 YOLOv8 作为 MT 模型的学生网络和教师网络,以构建更加稳定的训练。其次,使用图像融合网络结合风格迁移网络生成伪图像,执行伪交叉训练,以减轻红外域和可见光域之间差异。最后,使用 HAT 注意力机制提升图像分辨率,激活更多有用的像素,提升检测器的检测精度。本文在 LLVIP 数据集和 KAIST 数据集进行实验,最终实验结果证明了本文所提出方法的有效性和优越性,也揭示了采用先进检测器进行域适应研究的必要性。

参考文献

[1] HNEWA M, RADHA H. Multiscale domain adaptive YOLO for cross-domain object detection[C]. 2021 IEEE International Conference on Image Processing (ICIP). IEEE, 2021: 3323-3327.

[2] XU X, LIU W, WANG Z, et al. Towards generalizable person re-identification with a bi-stream generative model[J]. Pattern Recognition, 2022, 132: 108954.

[3] WAGNER J, FISCHER V, HERMAN M, et al. Multispectral pedestrian detection using deep fusion convolutional neural networks[C]. ESANN, 2016, 587: 509-514.

[4] 曹红燕,沈小林,刘长明,等. 改进的 YOLOv3 的红外目标检测算法[J]. 电子测量与仪器学报,2020,34(8):188-194.
CAO H Y, SHEN X L, LIU CH M, et al. Improved infrared target detection algorithm of YOLOv3 [J]. Journal of Electronic Measurement and Instrumentation, 2020,34(8):188-194.

[5] 韩航迪,徐亦睿,孙博,等. 基于改进 Tiny-YOLOv3 网络的航天电子焊点缺陷主动红外检测研究[J]. 仪器仪表学报, 2020,41(11):42-49.
HAN H D, XU Y R, SUN B,et al. Using active thermography for defect detection of aerospace electronic solder joint base on the improved Tiny-YOLOv3 network[J]. Chinese Journal of Scientific Instrument,2020,41(11):42-49.

[6] 谌海云,余鸿皓,王海川,等. 基于改进 YOLOX 的红外目标检测算法[J]. 电子测量技术,2022,45(23):72-81.
SHEN H Y, YU H H, WANG H CH, et al. Object detection algorithm of thermal infrared images based on improved YOLOX [J]. Electronic Measurement Technology,2022,45(23):72-81.

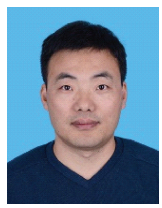
[7] BAO F, WANG X, SURESHBABU S H, et al. Heat-assisted detection and ranging [J]. Nature, 2023,

619(7971): 743-748.
[8] 朱菊香,崔长胜,张赵良,等. 红外道路场景下车辆与行人检测算法[J]. 国外电子测量技术,2023,42(6): 171-179.
ZHU J X, CUI CH SH, ZHANG ZH L, et al. Vehicle and pedestrian detection algorithm in infrared road scene[J]. Foreign Electronic Measurement Technology, 2023, 42(6):171-179.
[9] ZHANG L, LIU Z, ZHANG S, et al. Cross-modality interactive attention network for multispectral pedestrian detection[J]. Information Fusion, 2019, 50: 20-29.
[10] ZHAO W, WANG T, TAN A, et al. Nighttime pedestrian detection based on a fusion of visual information and Millimeter-Wave radar [J]. IEEE Access, 2023.
[11] DENG J, LI W, CHEN Y, et al. Unbiased mean teacher for cross-domain object detection[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 4091-4101.
[12] VIBASHAN V S, OZA P, SINDAGI V A, et al. Mega-CDA: Memory guided attention for category-aware unsupervised domain adaptive object detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 4516-4526.
[13] CHEN W, LIN L, YANG S, et al. Self-supervised noisy label learning for source-free unsupervised domain adaptation[C]. 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2022: 10185-10192.
[14] ZHOU H, JIANG F, LU H. SSDA-YOLO: Semi-supervised domain adaptive YOLO for cross-domain object detection [J]. Computer Vision and Image Understanding, 2023, 229: 103649.
[15] XU B, CHEN M, GUAN W, et al. Efficient teacher: Semi-supervised object detection for YOLOv5[J]. arXiv preprint arXiv:2302.07577, 2023.
[16] YANG L, WANG Y, GAO M, et al. Deep co-training with task decomposition for semi supervised domain adaptation [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 8906-8916.
[17] CHEN X, WANG X, ZHOU J, et al. Activating more pixels in image super-resolution transformer [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 22367-22377.
[18] TARVAINEN A, VALPOLA H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results[J]. Advances in Neural Information Processing Systems, 2017, 30.
[19] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle consistent adversarial networks[C]. Proceedings of the IEEE International Conference on Computer Vision,2017: 2223-2232.
[20] TANG L, XIANG X, ZHANG H, et al. DIVFusion: Darkness-free infrared and visible image fusion[J]. Information Fusion, 2023, 91: 477-493.
[21] JIA X, ZHU C, LI M, et al. LLVIP: A visible-infrared paired dataset for low-light vision[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 3496-3504.
[22] HWANG S, PARK J, KIM N, et al. Multispectral pedestrian detection: Benchmark dataset and baseline[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1037-1045.
[23] LI C, SONG D, TONG R, et al. Multispectral pedestrian detection via simultaneous detection and segmentation[J]. arXiv preprint arXiv:1808.04818, 2018.
[24] SHEN J F, CHEN Y F, LIU Y, et al. ICAFusion: Iterative cross-attention guided feature fusion for multispectral object detection[J]. Pattern Recognition, 2024, 145: 109913.

作者简介

杜运亮,现为青岛科技大学硕士研究生,主要研究方向为深度学习。
E-mail: 2217462012@qq.com

Du Yunliang is now a M. Sc. candidate at Qingdao University of Science and Technology. His main research interest includes deep learning.



王明甲(通信作者),2016 年于华东师范大学获得博士学位,现为青岛科技大学副教授,主要研究方向为计算机视觉与模式识别。
E-mail: mingjiawang@126.com

Wang Mingjia (Corresponding author) received his Ph. D. degree from East China Normal University in 2016. He is now an associate professor in Qingdao University of Science and Technology. His main research interests include computer vision and pattern recognition.