



研究与开发

基于SDL-YOLO的红外图像人车检测算法

黄国潮¹, 常城², 方虹苏², 杨炜²

(1. 浙江交通职业技术学院汽车学院, 浙江 杭州 311112;

2. 长安大学汽车学院, 陕西 西安 710064)

摘要: 为了解决交通安全系统中红外道路行人及车辆目标检测中的漏检、误检及精度较差的问题, 对基础YOLOv8进行改进, 提出了改进SDL-YOLO算法。首先, 通过引入空间深度转换卷积 (SPD-Conv) 以增强多尺度特征捕获, 解决图像分辨率较低及目标检测困难问题; 其后, 引入一种轻量级上采样算子DySample, 通过动态采样策略以优化复杂场景特征图, 改善目标边界模糊及环境遮挡问题; 最后, 将检测头改为轻量化共享卷积检测头 (lightweight shared convolutional detection, LSCD), 以减少模型参数及计算量, 在提升检测精度的同时保证轻量化。与此同时, 在公开红外夜景数据集上使用量化评价指标精确率 P 、召回率 R 、平均精度mAP及每秒浮点操作数 (floating-point operations per second, FLOPS) 对算法进行综合评估。实验结果表明, SDL-YOLO相较原始算法, 其精确率 P 提高1.5%、召回率 R 提高1.5%、平均精度mAP 50和mAP 50-95分别提高1.5%和1.7%, 同时具有较低的模型计算复杂度FLOPS。通过检测效果可视化分析进一步论证了SDL-YOLO算法在提升检测精度、改善漏检及误检方面的有效性。

关键词: 深度学习; 目标检测; 红外图像; YOLOv8; 行人车辆

中图分类号: TP391

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2025176

Infrared image human vehicle detection algorithm based on SDL-YOLO

HUANG Guochao¹, CHANG Cheng², FANG Hongsu², YANG Wei²

1. School of Automotive Engineering, Zhejiang Institute of Communications, Hangzhou 311112, China

2. School of Automobile, Chang'an University, Xi'an 710064, China

Abstract: To solve the problems of missed detection, false detection, and poor accuracy in infrared road pedestrian and vehicle target detection in traffic safety systems, the basic YOLOv8 was improved and an improved SDL-YOLO algorithm was proposed. Firstly, by introducing spatial depth transformation convolution SPD-Conv to enhance multi-scale feature capture, the problem of low image resolution and difficult object detection could be solved. Subsequently, a lightweight upsampling operator DySample was introduced to optimize complex scene feature maps

收稿日期: 2025-02-20; 修回日期: 2025-05-18

通信作者: 杨炜, yw@chd.edu.cn

基金项目: 陕西省自然科学基金青年项目 (No.2017JQ6045)

Foundation Item: Shaanxi Natural Science Foundation Youth Project (No.2017JQ6045)



through dynamic sampling strategies, effectively mitigating target boundary blurring and environmental occlusion issues. Finally, the detection head would be changed to a lightweight shared convolutional detection head LSCD to reduce model parameters and computational complexity, while ensuring lightweight while improving detection accuracy. At the same time, the algorithm was comprehensively evaluated using quantitative evaluation indicators accuracy P, recall R, average accuracy mAP, and FLOPS on the publicly available infrared night scene dataset. The experimental results show that compared with the original algorithm, SDL-YOLO has a 1.5% increase in accuracy P, a 1.5% increase in recall R, and an 1.5% and 1.7% increase in average accuracy mAP50 and mAP50-95, respectively. At the same time, it has lower model complexity FLOPS. The effectiveness of SDL-YOLO algorithm in improving detection accuracy, correcting false positives and false negatives has been further demonstrated through visual analysis of detection results.

Key words: deep learning, object detection, infrared image, YOLOv8, pedestrian and vehicle

0 引言

交通安全系统是保障道路交通安全与预防交通事故的重要工具,其中红外目标检测因不受光照变化影响,被广泛用于交通安全领域^[1-2]。特别是在夜间道路监控环境下,红外图像在人车检测中的应用价值更为突出。目前,实际交通夜间监控场景中,红外摄像头多采用高空固定部署,进行俯视角度的区域监控,用于道路中人车目标的实时检测和跟踪。该类场景具有目标尺度变化大、遮挡严重、对比度低、目标边界模糊等环境特点,对检测算法的精度提出了较高要求。

而传统红外目标检测通常依赖空间滤波、频域滤波^[3]、局部对比度测量及基于序列的图像分割方法^[4]。上述方法对环境依赖性强,面对低分辨率及低对比度的夜间环境时,存在检测精度较低及速度较慢等问题。

随着人工智能技术的进步,深度学习技术在目标检测领域取得了突破性进展^[5],并逐渐应用于红外目标检测。Girshick^[6]提出的区域卷积神经网络(region convolution neural network, R-CNN)首次将深度学习模型引入目标检测。然而,R-CNN在计算效率上存在较大瓶颈,难以在实际场景中有效应用。Sheeny等^[7]使用Fast R-CNN进行红外车辆识别检测,显著提高了检测速度。Redmon等^[8]

提出YOLO (you only look once) 算法,将目标检测转为回归问题,实现了高速检测。近年来,YOLO系列算法已被广泛用于红外目标检测。王晓红等^[9]在YOLOv5模型上引入双向特征金字塔网络(bi-directional feature pyramid network, BiFPN)特征融合机制,提高了检测精度。李阳等^[10]利用卷积下采样层替换低效Focus Layer,并对非极大值抑制(non-maximum suppression, NMS)进行优化以提升检测精度。胡皓等^[11]基于YOLOX-S算法,通过引入结构重参数化模块和AdamW优化算法来提高检测精度。邓天民等^[12]将注意力特征提取模块(AFFM)引入融入网络主干部分,并设计一种多尺度特征融合模块,改进了红外图像中行人的检测效果。赵明等^[13]提出一种跨域融合网络结构,采用红外域及伪可见光域双通道的多尺度特征融合方法,可实现双模态特征有效融合,以提高检测精度。Chen等^[14]对红外目标进行迭代分割,提高了平均精度。

尽管上述方法在提升目标检测精度方面取得了一定进展,但面对低分辨率、低信噪比及复杂背景下目标轮廓模糊红外图像^[15]时,仍存在较为严重的漏检及误检,检测精度仍有待提升。此外,交通安全系统对模型轻量化也提出一定要求,避免过度增加计算资源消耗。

因此,为了满足交通安全系统的实际应用需

求,在保证轻量条件下提升红外图像人车检测的精度,同时解决漏检及误检问题,基于YOLOv8-N模型,针对固定高空监控视角下红外人车检测的实际场景需求,提出一种新型改进算法SDL-YOLO以提升红外图像目标检测精度,同时在红外道路人车检测数据集上对SDL-YOLO进行对比及消融实验以验证其有效性。

1 SDL-YOLO检测算法

YOLOv8作为单阶段目标检测算法中的前沿模型之一,根据应用领域及多样化需求,可分为N、S、M、L及X版本。其中N版本为轻量化算法,具备较低的参数量与计算复杂度,选择将其作为基础模型。

YOLOv8算法相比YOLOv10或YOLOv11等最新算法,由于其经过广泛实践验证,并拥有成熟技术生态,具有更稳定及兼容的优势,可更好地保证模型实际应用稳定性。同时,YOLOv8算法在处理小目标数据集方面具备较强的特征表达能力,更适用于红外图像中目标尺寸普遍较小、特征不明显的问题等挑战。相比Transformer等大模型,YOLOv8在全局特征建模与局部特征提取之间取得了良好平衡,且其计算复杂度较低、推理速度较快,满足实际交通夜间监控场景中对检测实时性与稳定性的要求。综合考虑下,选取YOLOv8作为基础模型进行改进创新,以更好满足实际应用场景需求。

然而,在实际的红外道路监控任务中,特别是在固定高空俯视视角下,受监控设备部署位置和成像特性的限制,检测环境复杂多变,目标图像存在尺寸小、特征弱、背景干扰强等问题。同时,红外图像本身的低分辨率特性也导致目标模糊、细节丢失,影响检测准确性。此外,俯视视角下的人车目标在空间上往往密集排列,且伴随不同程度的遮挡,进一步加大了检测难度。这导致原始YOLOv8架构在该场景下的检测效果存在

明显不足。基于此,提出了SDL-YOLO改进算法,从特征增强、细节重建及检测头轻量化3个核心方向对YOLOv8进行改进,SDL-YOLO网络架构如图1所示(虚线框为改进部分)。

首先,在骨干(Backbone)中引入空间深度转换卷积(space-to-depth convolution, SPD-Conv)^[16],对特征提取及空间细节重建等方面进行全面优化,通过对特征图空间信息的有效压缩与深度增强,提升模型对小尺寸、弱特征目标的感知能力。其次,对颈部(Neck)结构进行改进,引入动态采样(dynamic sampling, DySample)^[17]算子,通过自适应的特征重建策略,强化目标边界和细节信息,缓解遮挡与目标粘连带来的检测干扰。最后,在检测头部分设计轻量化共享卷积检测头(lightweight shared convolution detection, LSCD),实现跨层级共享卷积权重,在确保检测精度的同时降低模型参数与计算开销,以轻量化模型。

1.1 SPD-Conv空间到深度卷积模块

针对红外道路监控图像中普遍存在的分辨率低、目标小、噪声干扰及目标特征不明显等问题,传统YOLOv8网络结构在特征提取过程中,主要依赖步长卷积和池化操作对特征图进行压缩,虽然可降低计算量,但会导致大量细粒度空间信息的丢失,尤其在低分辨率的红外图像中,加剧了小目标特征的丢失,使得人车目标与复杂背景之间的细微差异难以有效区分,影响检测性能。基于此,从特征保留与表达能力出发,本文研究采取一种SPD-Conv,旨在强化小目标与弱特征目标的检测效果,以提升检测精度。

SPD-Conv模块的核心思想是通过空间重组操作,将输入特征图的空间信息重组至通道维度,减少空间压缩带来的细节损失,弱化背景噪声的影响,在保持特征完整性的同时,提升模型对小目标及弱特征目标的感知能力,SPD-Conv工作原理(scale=2)如图2所示。

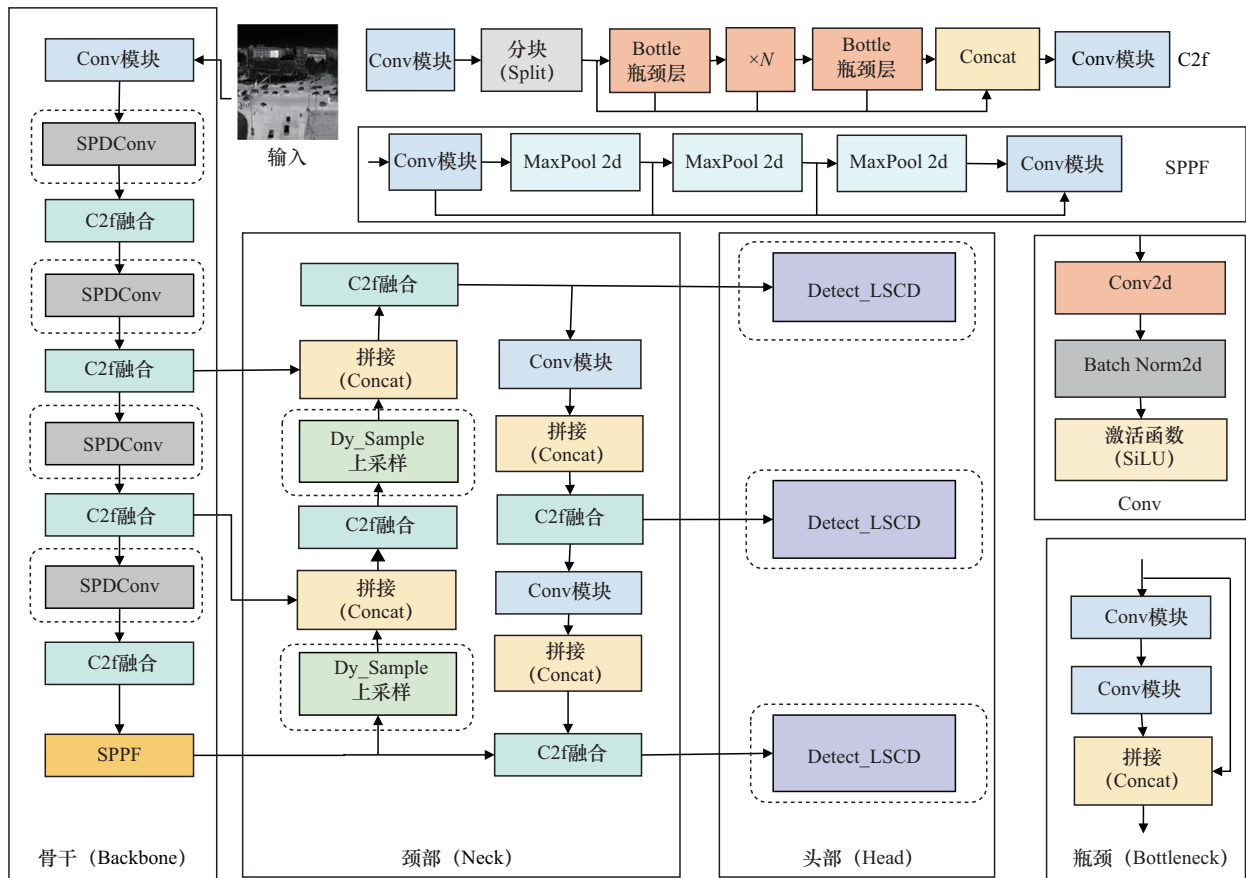


图1 SDL-YOLO网络架构

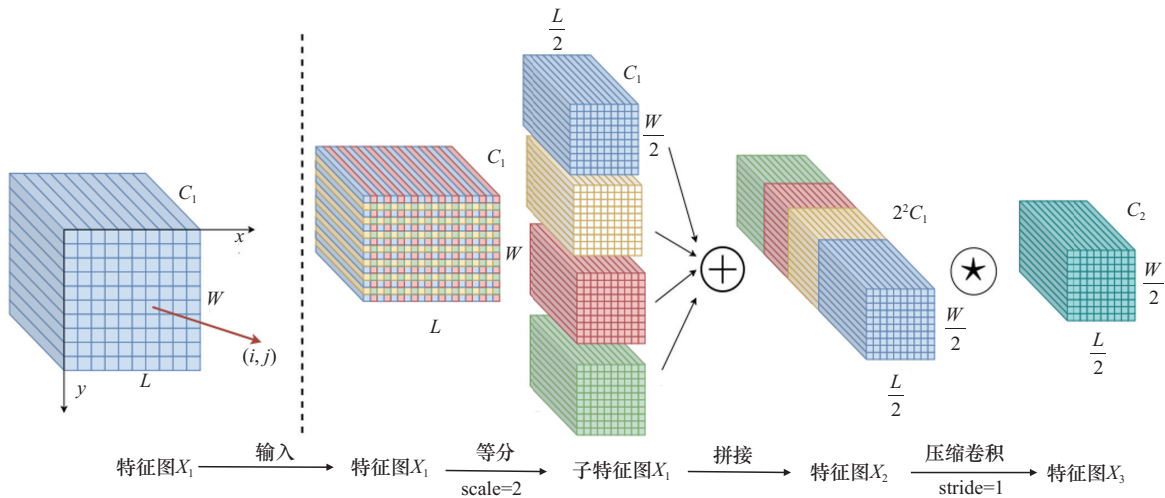


图2 SPD-Conv工作原理(scale=2)

由图2可知，SPD层首先将大小为 $L \times W \times C_1$ 的输入特征图 X_1 (L 代表长度、 W 代表宽度、 C_1 代表通道数)划分为多个子特征图，并在通道维度上重新连接，形成特征图 X_2 。该过程在减少空

间维度的同时扩展了深度维度，可有效保留输入数据的关键空间特征，避免细粒度信息丢失，降低了背景噪声的影响。

其后，为提升特征的表达能力，SPD-Conv

在重组后的特征基础上，使用非步长卷积（stride=1）进行特征提取，避免了步长卷积可能带来的空间信息破坏。通过 C_2 个过滤器，可进一步提炼空间重组后的深度特征，最终输出大小为 $(\frac{L}{2}, \frac{W}{2}, C_2)$ 的结果特征图，有效缓解了小目标与背景特征难以区分的问题。相比传统卷积，该卷积减少了通道数，避免了步长操作带来的空间信息损失。此外，卷积层的学习参数可优化特征表示，增强对细微特征的捕捉能力，以实现低分辨率及小目标的精准检测。

同时为确保 SPD-Conv 模块在训练过程中可有效收敛，同时兼顾检测任务中的分类与定位性能，在此采用分类损失和边界框回归损失的联合优化策略，具体损失函数表达式如下：

$$L = L_{cls} + \lambda L_{bbox} \quad (1)$$

其中， L_{cls} 表示分类损失项，采用交叉熵损失函数对目标类别进行判别； L_{bbox} 表示边界框回归损失项，主要用于指导检测框的准确定位； λ 为权

衡系数，用于平衡分类精度与检测框回归精度的重要性，确保整体检测效果最优。

综合而言，SPD-Conv 模块主要针对目标小、噪声干扰及特征不明显等问题，通过空间信息的有效重组、特征完整性的强化以及小目标细节的精确捕捉，增强了特征表示，提升了红外道路高空俯视视角下人车检测任务的检测精度。

1.2 DySample 动态上采样算子

在处理高空俯视视角下的红外图像，YOLOv8 中传统上采样算子普遍采用固定规则的插值方式（如最近邻插值、双线性插值等）进行特征恢复。然而，这种静态的插值操作未考虑目标本身的空间特征分布，尤其在多目标密集时，易导致特征图的边界模糊与细节丢失，影响检测精度。为此，引入了一种 DySample 轻量动态上采样算子，旨在根据输入特征图的空间与语义特征，动态生成可学习偏移量，实现对重要特征区域的自适应增强，以改善目标边界与细节的恢复效果，DySample 采样流程如图 3 所示。

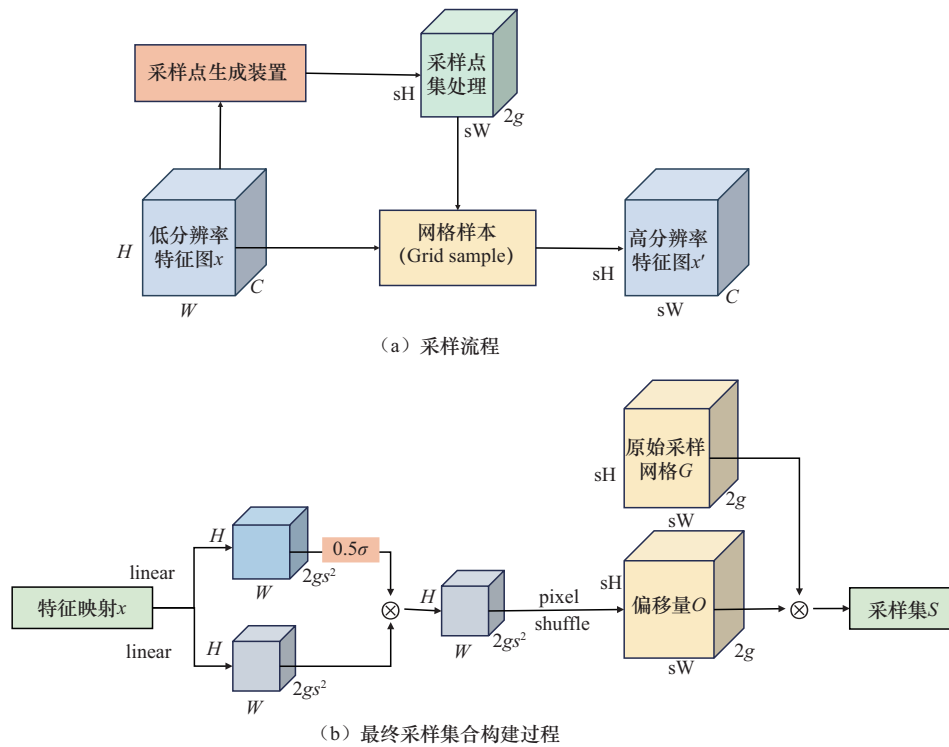


图3 DySample 采样流程



由图3可知,输入低分辨率特征图 x 后,首先通过卷积层计算以生成偏移量 O 。该偏移量计算则由特定线性层进行,其中输入输出通道数分别为 C 和 $2gs^2$,使得特征图中每点均可根据其周围的空间与语义上下文信息进行自适应调整。此外,该偏移量基于特征图局部信息,以帮助模型精确调整采样点位置,并增强细节恢复能力。相比传统上采样方法,DySample不再固定使用步长卷积及池化,其通过动态调整,以确保更精准的特征恢复,偏移量 O 计算式如下:

$$O = \text{linear}(\chi) \quad (2)$$

其中, $\text{linear}(\chi)$ 表示一层线性变换,用于对原始特征图进行特征映射,生成基础的空间偏移量 O 。

与此同时,为进一步增强偏移的表达能, DySample引入了动态范围因子,该动态范围因子由双线性层及sigmoid函数生成,可根据不同输入特征图灵活调整偏移量。经动态调整的偏移量 O 可更精确映射目标位置及细节,并进一步优化特征图分辨率及细节保留,动态调整偏移量计算方式如下:

$$O = 0.5\text{sigmoid}(\text{linear}_1(X)) \cdot \text{linear}_2(X) \quad (3)$$

其中, $\text{linear}_1(X)$ 与 $\text{linear}_2(X)$ 分别为两层线性变换,用于生成动态权重和空间偏移,sigmoid将权重限制在(0,1)区间,通过与 $\text{linear}_2(X)$ 的逐元素相乘,实现对基础偏移的动态调控。

其后,经像素重组以生成大小为 $sH \times sW \times 2g$ 的偏移量 O ,将其叠加至初始采样网格 G 之上,得到最终采样集合 S ,以实现特征图的自适应重采样,有效提升目标的空间表达能力与边界细节还原能力。其计算式如下:

$$S = G + O \quad (4)$$

最后,DySample采用网格采样(grid sampling)操作,将输入特征图 x 按照采样集合 S 进行重采样,生成高分辨率的特征图 x' 。该过程不仅有

效缓解了遮挡及密集场景中目标边界模糊的问题,也增强了红外图像中弱对比度区域的特征恢复能力,为后续检测任务提供了丰富的结构信息。

1.3 轻量型共享卷积LSCD检测头

在前端特征提取与特征重建过程中,由于SPD-Conv与DySample模块的引入增加了模型的整体参数量,而YOLOv8原有的解耦式检测头若继续沿用独立多分支的特征处理方式,易加剧参数量增长。为了在保证检测效果的前提下,合理控制模型的参数增长,同时进一步提升特征利用效率与检测精度,在检测头部分设计了轻量化共享卷积检测头(LSCD),旨在确保检测效果的基础上,有效降低检测头的参数规模,控制整体模型计算量的合理增长,同时提升多尺度特征的利用效率,LSCD架构如图4所示。

由图4可知,3种尺寸特征层(P3、P4及P5)进入检测头后,首先经过 1×1 卷积GN,该卷积采用群组归一化(group normalization, GN)代替传统批量归一化(batch normalization, BN),可将通道数调整为中间层通道数。其后,将所有特征层进行汇聚,通过 3×3 共享卷积进行深度特征提取,实现特征共享与参数复用,以降低模型参数量。此外,检测任务被分解为回归分支及分类分支。

回归分支中,通过 1×1 卷积以预测边界框坐标偏移量,计算式如下:

$$\Delta B = \text{Conv}_{1 \times 1}^{\text{reg}}(Z) \quad (5)$$

其中, ΔB 为预测的边界框偏移量, $\text{Conv}_{1 \times 1}^{\text{reg}}$ 表示用于回归的 1×1 卷积层, Z 为共享卷积后的特征图。

同时引入Scale层,以实现特征尺度动态调整,提升边界框定位精度,计算式如下:

$$B' = \text{Scale}(\Delta B) \odot Z \quad (6)$$

其中, B' 为调整后的边界框, $\text{Scale}(\Delta B)$ 表示对偏移量进行尺度变换, \odot 表示逐元素相乘操作。

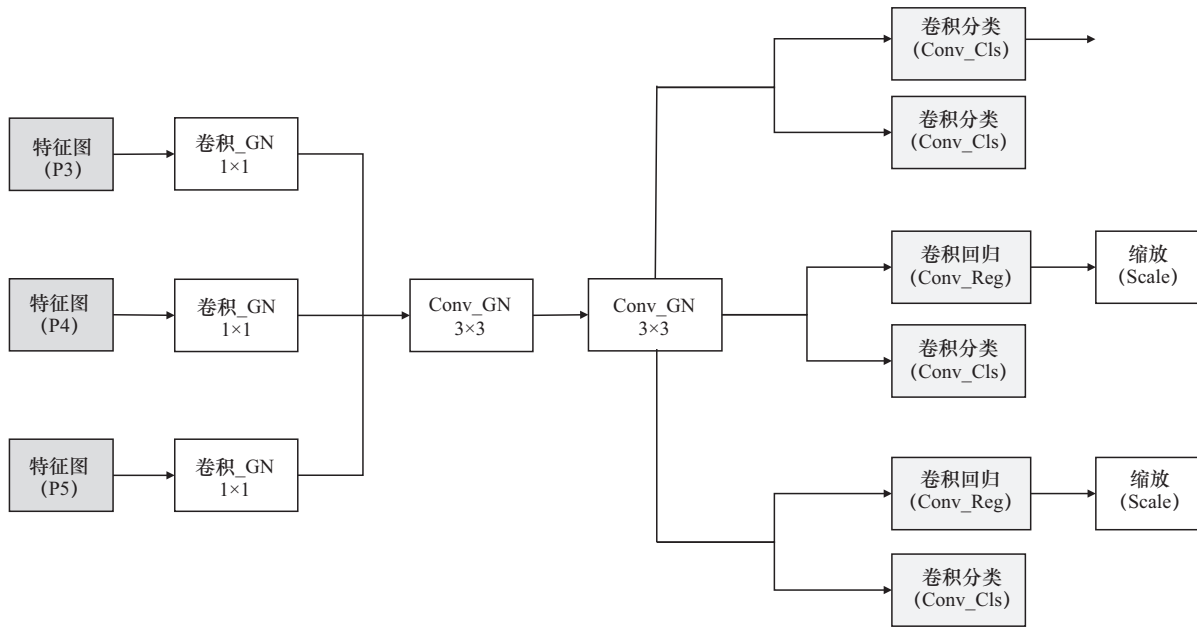


图4 LSCD架构

分类分支中，同样利用 1×1 卷积预测目标类别，具体操作如下：

$$C = \text{Conv}_{1 \times 1}^{\text{cls}}(Z) \quad (7)$$

其中， C 为预测的类别概率分布， $\text{Conv}_{1 \times 1}^{\text{cls}}$ 表示用于分类的 1×1 卷积层。

通过以上的双分支卷积权重独立设计，不仅有效降低了检测头的参数数量，缓解了前端改进模块带来的整体参数增长压力，同时对检测精度的提升起到了积极作用，实现了复杂红外环境下检测精度与模型复杂度之间的合理平衡。

2 实验及结果分析

2.1 环境及参数配置

本文研究采用 PyTorch 深度学习框架进行实验，环境配置：处理器为 i5-10300H，显卡为 GTX1650，显存 4 GB，系统内存 16 GB。软件配置：Python 版本为 3.8，CUDA 版本为 12.1。训练超参数设定：训练次数为 200 轮，批量大小为 16，优化器为 SGD，初始学习率为 0.001。

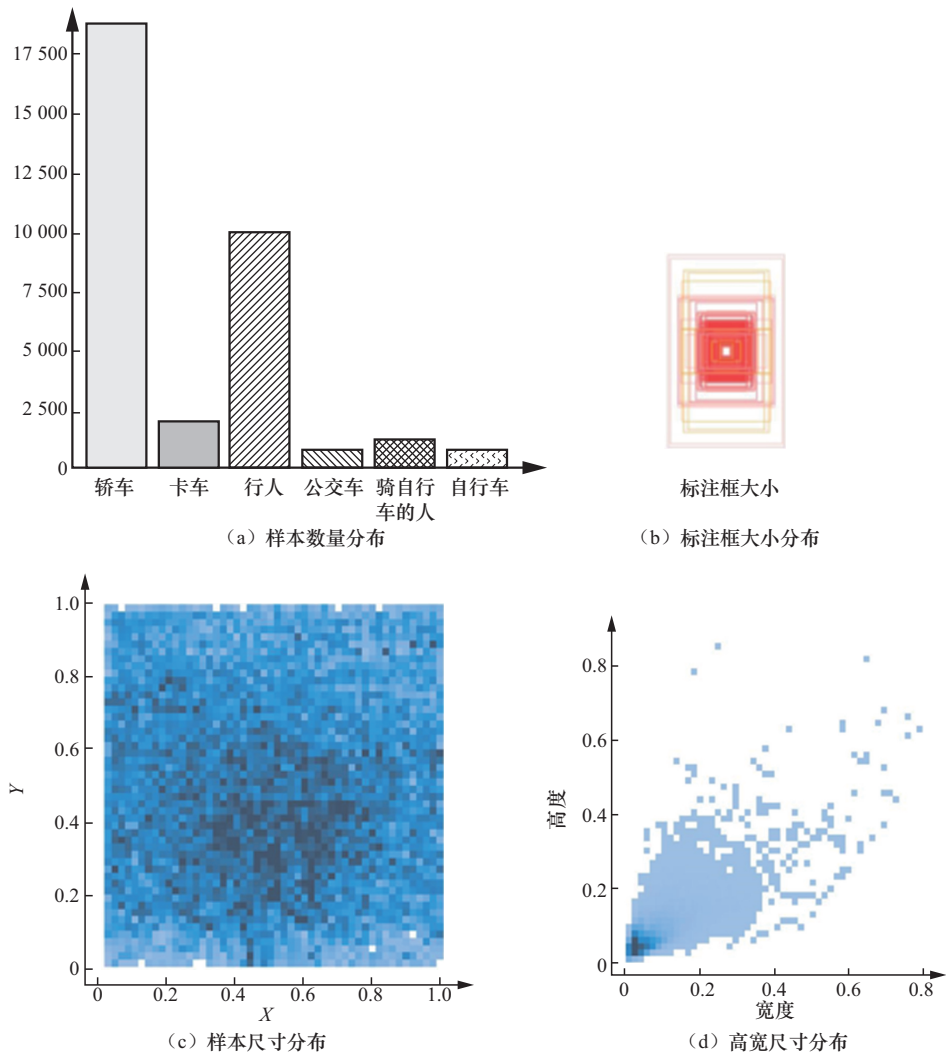
2.2 数据集构建

实验所用红外道路图像数据集源于睿创微纳

(Raytron) 公司所搭建的红外航拍人车数据集，该数据集通过专业红外无人机平台对实际交通夜间监控环境进行模拟，对典型道路场景下的人车目标进行高空航拍采集，图像采集分辨率为 640×512 ，拍摄高度约离地面 10~15 m，包括多视角下图片 2 300 张。

其后，对数据图像使用 labelimg1.8.6 版本进行标注，不同颜色标注框代表不同目标，所生成的对应标注文件包括 6 类目标，分别为轿车 (car)、卡车 (truck)、行人、公交车 (bus)、骑自行车的人 (cyclist) 及自行车 (bike)，数据集分布情况如图 5 所示。

由图 5 (a) 可知，待检目标包括卡车、公交车及轿车等较大目标，同时包含行人、自行车及骑自行车的人等较小目标，该数据集中轿车与行人标注样本最为丰富，其余目标相对较少。在红外图像目标检测任务中，漏检及误检现象主要集中在小目标检测环节，因此数据集中小目标（尺寸小于 0.12）占比最大，如图 5 (b) 所示。同时，图 5 (c) 展现了目标的空间分布规律，表明数据集可对改进算法在不同位置、尺寸目标检测



3
图5 数据集分布情况

效果进行验证。图5(d)则分析了标注框高宽比例分布,表明数据集中目标形状的多样性及复杂性,即数据集在构建过程中充分考虑了不同场景下目标的分布特性,可较为全面地反映改进算法的检测效果。

最后,将数据集按8:1:1比例划分为1840张训练集、230张测试集及230张验证集。

2.3 评价指标

为了评估改进算法的检测效能,选择以下指标进行量化评估。

精确率 (precision, P): 指模型判定为正类

样本中实际为正类的比率,计算方式如下:

$$P = \frac{TP}{(FP + TP)} \quad (8)$$

其中, TP指检测结果与真实标注均为目标的样本数,表示成功检测的正类数量; FP指模型误将负类(如背景)识别为正类(如行人、车辆)的数量。精确率越高,表示模型误检情况越少;

召回率 (recall, R): 指所有实际正类样本中被正确预测为正类的比例,计算方式如下:

$$R = \frac{TP}{(FN + TP)} \quad (9)$$

其中, FN指模型未能检测出的正类样本数量,

即实际存在但未检测出的目标。召回率越高，说明模型漏检情况越少。

此外，以 mAP50 及 mAP50-95 来评估算法平均精度（以 A 代替），计算式如下：

$$A = \frac{1}{n} \sum_{k=1}^{k=n} AP_k \quad (10)$$

其中， n 为类别总数， k 为检测次数，AP 为综合考虑精确率及召回率的平均精度。

mAP50 指交并比 (IoU) 阈值为 0.5 的平均精度，评估模型对目标粗略定位的能力；mAP50-95 则代表 IoU 阈值为 0.95 的平均精度，反映模型在不同定位精度要求下的综合检测性能，评估标准更为严格。

上述指标值越大，表明算法效果越好，漏检

及误检情况越少^[18]。此外，FLOPS 为衡量模型计算复杂度的重要指标，代表模型推理过程中需执行浮点运算的次数。

2.4 实验结果分析

完成 200 轮迭代训练后，改进前后算法损失函数、训练结果及 PR 曲线对比如图 6 所示。

由图 6 中 (a) 可知，网络模型损失函数在初期快速下降，经大约 20 轮迭代后趋于平缓，符合模型训练规律，经大约 3.5 h，SDL-YOLO 网络完成 200 轮迭代训练。此外，由图 6 中 (b) 可知，训练标签标注准确无误，包含各类目标对象。由图 6 中 (c) 及 (d) 可知，在相同召回率基础上，SDL-YOLO 可显著提升对行人、自行车及骑自行车的人等小型目标的检测精确率 P 。对于其余检

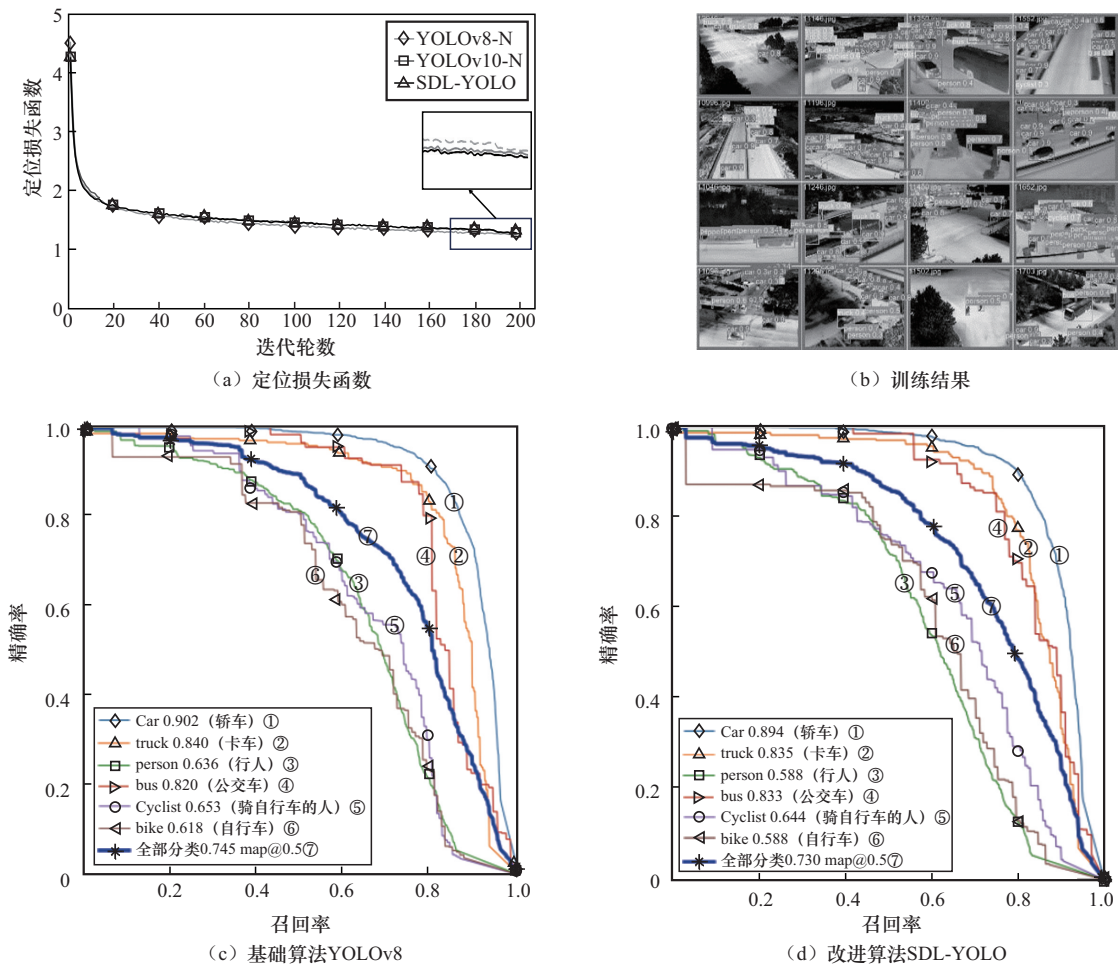


图6 改进前后算法损失函数、训练结果及PR曲线对比



测对象，两模型精确率相近，表明优化后算法在增强特定目标检测能力的同时，并未牺牲整体性能。此外，SDL-YOLO在处理红外复杂环境小目标检测任务时，表现远超基础模型。IoU在0.5阈值下，平均精度mAP相比基础模型提升1.5%。虽然两算法对自行车检测均为薄弱项，但SDL-YOLO对该目标检测效果显著高于基础模型，同时对轿车目标检测精度也保持最高水平。

综合而言，SDL-YOLO通过优化算法结构，有效提升了红外道路环境中高空俯视视角下人车目标的检测精度与鲁棒性，具备应用潜力及实用价值。

2.5 消融实验

为了验证SDL-YOLO在红外道路人车目标检测中各改进模块的独立有效性及组合协同效果，设计了8组消融实验，以评估SPD-Conv、DySample及LSCD 3项改进技术的独立性及协同性。消融实验见表1。

由表1可知，融入SPD-Conv、DySample及LSCD后，SDL-YOLO在所有评估指标上均实现了有效提升。具体而言，其精确率提升1.5%，表明误检率降低，有助于减少系统对非目标区域误判，以提高检测可靠性及用户信任度。同时，其召回率提升1.5%，可确保更多实际目标被正确识别，对于交通夜间监控应用尤为重要，可有效检测潜在目标，以提升整体系统安全性。此外，其mAP50及mAP50-95的提升进一步证明了SDL-YOLO可在不同置信度阈值下进行准确检测，具

备较强的复杂环境适应能力。

协同效应方面，消融实验中的组合模型4（SPD-Conv+DySample）及模型5（SPDConv+LSCD）虽在mAP及召回率上表现出显著提升，但精确率略有下降，表明单独组合某些改进点可能产生负面影响。然而SDL-YOLO通过整合SPD-Conv、DySample及LSCD，可充分发挥各改进点协同作用，以实现全面性能提升，并确保红外复杂道路场景的高效、准确检测。

此外，SDL-YOLO在整体模型计算量上相较于YOLOv8-N有所增加，这种增长主要来源于特征增强与动态采样等模块的引入，以提升复杂红外环境下的检测精度。但得益于LSCD在检测端的轻量化设计，有效缓解了前端改进模块带来的计算压力，使得整体参数与FLOPS的增长幅度被合理控制（计算复杂度仅由8.9 GFLOPS提升至9.5 GFLOPS），在保证检测性能提升的同时，仍具备良好的工程部署价值，满足交通夜间监控场景对检测精度与模型轻量化平衡的实际需求。

综上所述，SDL-YOLO通过多项技术的综合应用，不仅可有效提升目标检测精度及可靠性，还可有效控制计算开销，为交通监控系统中红外复杂道路人车目标检测提供高效且实用解决方案。

2.6 对比实验

为了进一步验证SDL-YOLO在红外道路人车目标检测中的有效性，在相同实验条件下设计了

表1 消融实验

模型	SPD-Conv	DySample	LSCD	精确率	召回率	mAP50	mAP50-95	计算复杂度/GFLOPS
YOLOv8-N	/	/	/	75.6%	66.6%	73.0%	42.5%	8.9
模型1	√	/	/	76.9(+1.3)%	67.6(+1.0)%	74.4(+1.4)%	43.7(+1.2)%	11.6
模型2	/	√	/	77.0(+1.4)%	67.1(+0.5)%	73.7(+0.7)%	43.4(+0.9)%	8.1
模型3	/	/	√	75.7(+0.1)%	66.8(+0.2)%	73.3(+0.3)%	43.0(+0.5)%	6.5
模型4	√	√	/	75.1(-0.5)%	69.0(+2.4)%	74.3(+1.3)%	44.1(+1.6)%	11.5
模型5	√	/	√	74.8(-0.8)%	69.6(+3.0)%	73.7(+0.7)%	43.2(+0.7)%	10.1
模型6	/	√	√	75.6(+0.0)%	69.4(+2.8)%	73.9(+0.9)%	44.1(+1.6)%	6.5
SDL-YOLO	√	√	√	77.1(+1.5)%	68.1(+1.5)%	74.5(+1.5)%	44.2(+1.7)%	9.5

12组对比实验，对比模型实验见表2，旨在全面评估SDL-YOLO性能优势。

表2 对比模型实验

模型	精确率	召回率	mAP50	mAP50-95	计算复杂度/ GFLOPS
Faster-RCNN	43.5%	39.6%	38.7%	14.9%	369.8
RetinaNet ^[19]	75.8%	46.3%	57.9%	32.9%	146.9
SSD ^[20]	75.5%	28.1%	44.3%	23.7%	62.7
CenterNet ^[21]	76.6%	29.6%	55.3%	27.3%	70.2
EfficientDet ^[22]	72.8%	40.5%	51.6%	30.8%	4.7
YOLOv3-tiny ^[23]	68.8%	62.2%	65.9%	37.5%	18.9
YOLOv5-N ^[24]	73.9%	67.0%	71.8%	41.3%	7.1
YOLOv6-N	75.2%	62.2%	69.5%	40.2%	11.8
YOLOv7-tiny ^[25]	72.6%	64.3%	70.5%	37.4%	13.0
YOLOv8-N	75.6%	66.6%	73.0%	42.5%	8.9
YOLOv10-N ^[26]	74.9%	65.2%	72.5%	43.1%	8.2
YOLOv11-N ^[27]	75.7%	68.0%	73.4%	43.4%	6.7
YOLOv12-N ^[28]	74.7%	67.8%	73.5%	43.0%	6.6
SDL-YOLO	77.1%	68.1%	74.5%	44.2%	9.5

由表2可知，双阶段算法Faster-RCNN及RetinaNet由于复杂的网络结构及较高计算资源需求，在红外道路人车目标检测任务中表现不佳。Faster-RCNN精确率、召回率及mAP指标均明显低于单阶段算法，而RetinaNet虽在精确率上有所提升，但其计算复杂度为146.9 GFLOPS，显著高于其他模型，限制了在实时检测中的应用。相比之下，单阶段算法SDL-YOLO以77.1%精确率、68.1%召回率、74.5% mAP50及44.2% mAP50-95，在各项指标上均优于SSD、CenterNet、Efficient-Det及YOLO系列算法，表现出更优的检测效能。

尽管SDL-YOLO在模型计算量上较YOLOv10-N、YOLOv11-N与YOLOv12-N略有提升，但通过三大改进模块的协同作用，此改进算法在各项性能指标均取得最佳效果的同时，计算复杂度仅为9.5 GFLOPS，依然保持在YOLO系列合理的复杂度区间，且远低于传统双阶段检测算法。整体来看，SDL-YOLO在兼顾检测精度与计算复杂度方面，展现出更优的平衡性。

各模型平均精度对比如图7所示。由图7可知，SDL-YOLO在平均精度mAP50及mAP50-95上均处于领先地位，在复杂红外道路场景下具备卓越泛化性及检测精度。

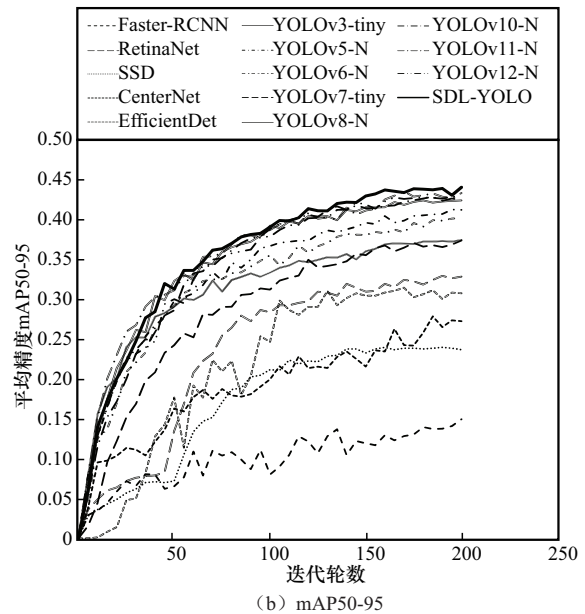
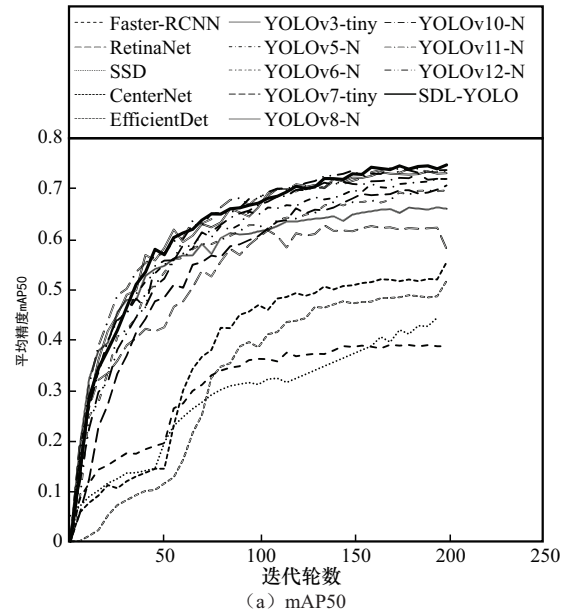


图7 各模型平均精度对比

综合而言，SDL-YOLO通过优化算法结构并引入关键技术，不仅可在目标检测精度上显著优于现有单阶段及双阶段算法，还可在计算复杂度方面保持良好平衡，在实际应用中具备较高效能及实用性。



2.7 检测效果可视化分析

为了进一步直观展示改进算法SDL-YOLO相较于对比算法在检测精度的提升、漏检及误检问题上的改善,从验证集中随机选取一幅典型红外道路图像进行测试分析,检测结果可视化如图8所示。

由图8可知,原始YOLOv8算法在对行人、车辆、自行车等目标进行检测时存在明显的漏检及误检问题。相比之下,改进算法SDL-YOLO通过优化的特征提取与细节重建模块,能够有效减少上述错误检测,以确保目标识别准确性。此外,YOLO系列各模型在检测行人及车辆等小尺寸目标时,普遍存在感知能力不足的缺陷,漏检问题严重。而SDL-YOLO在小尺寸目标检测方面表现出更强的鲁棒性与更低的漏检风险。相较于YOLOv3-tiny、YOLOv7-N、YOLOv8-N、YOLOv10-N等主流模型,SDL-YOLO在整体检测性能上表现更优,且在可视化分析中,其他对比模

型普遍存在不同程度的识别偏差。进一步对比检测置信度与目标定位精度可见,SDL-YOLO在对比图像中的表现同样优于其他模型。

相较而言,SDL-YOLO通过对特征提取、动态采样和检测头的深度优化,能更好地适应红外夜间环境和复杂道路场景,有效提升了目标检测精度,在置信度、定位精度及小目标检测能力上表现最为稳定。

整体而言,SDL-YOLO在多模型对比中,不仅在检测性能方面全面领先,同时在合理控制模型计算复杂度的前提下,达到了最优的轻量化与精度平衡,改善了原始YOLOv8的漏检及误检问题,验证了其在红外道路监控场景中的实用性。

3 结束语

针对YOLOv8在红外道路人车检测中的不足及现代交通夜间监控系统要求,提出了改进算法

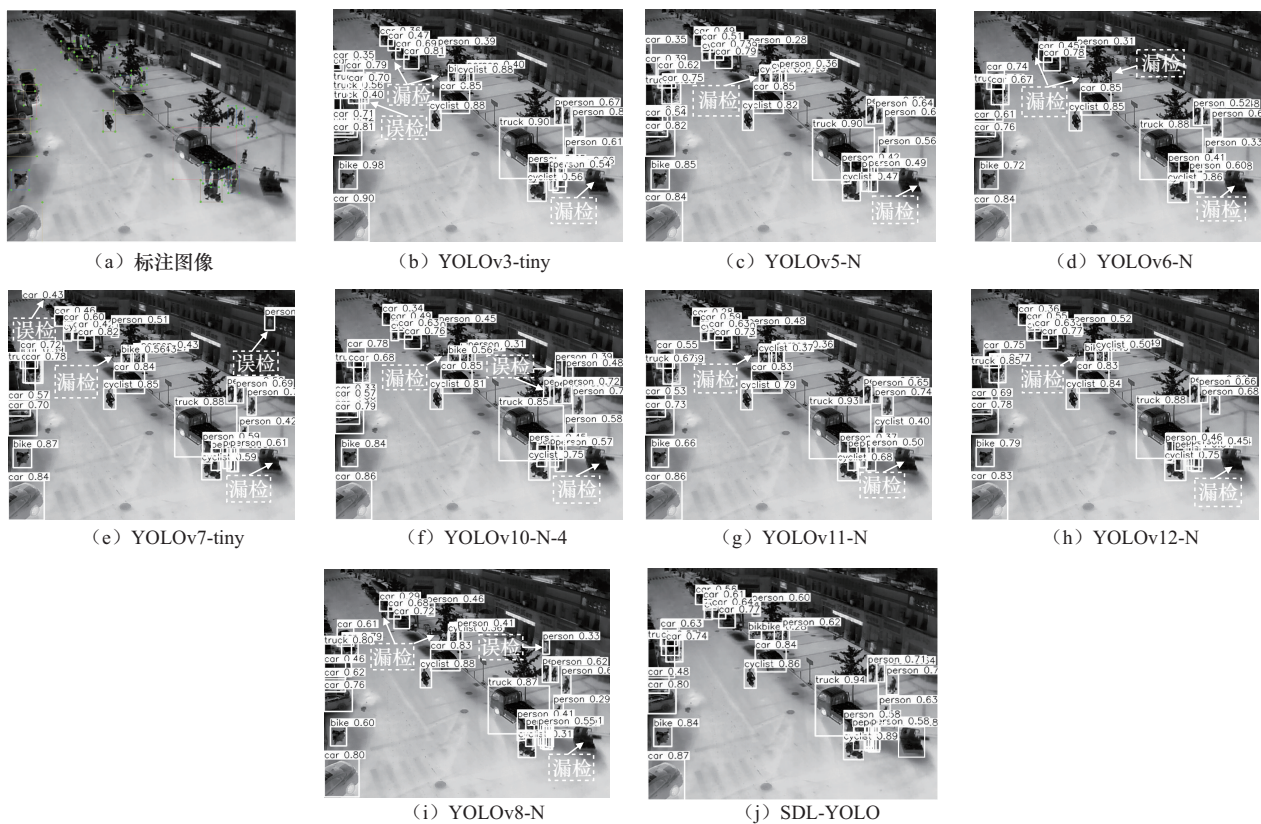


图8 检测结果可视化

SDL-YOLO。通过消融对比实验,其在红外道路检测任务中显著优于基础算法, mAP50-95 提高 1.5%, 达到 44.2%, 精确率及召回率等指标也得到提升, 有效减少了人车目标漏检及误检问题。此外, SDL-YOLO 在整体参数量少量增长的情况下, 仍保持合理的计算复杂度, 实现了检测性能与模型规模的有效平衡, 具备良好的工程适配性与实际应用潜力。

然而, 当前方法仍存在一些不足之处。尽管 SDL-YOLO 在实验中表现优于其他算法, 但其在极端条件下的鲁棒性尚未充分验证, 需要进一步改进模型的适应性。此外, SDL-YOLO 对大规模数据集的训练和推理速度的优化仍有提升空间, 尤其是在资源受限的设备上, 如何进一步减少计算消耗同时维持高精度也是一个具有挑战性的问题。

未来的研究将从以下几方面进行优化和改进。首先, 重点优化网络结构, 探索更高效的特征提取模块, 以实现模型复杂度和性能的最佳平衡, 尤其是在移动端或嵌入式设备上的部署。其次, 针对不同复杂场景中的目标检测鲁棒性问题, 未来将尝试引入多尺度、多模态数据融合的策略, 提高模型在各种环境下的适应性和泛化能力。进一步扩展数据集, 涵盖更多复杂场景(如雨天、雪天等), 将有助于提升模型在实际应用中的可靠性。最后, 在未来的研究中, 还将结合自监督学习或迁移学习等技术, 进一步提高红外图像目标检测的精度和速度, 为复杂红外道路环境中的目标检测提供更加高效、精确的解决方案。

参考文献:

- [1] 周平华, 王峰, 王鹤峰, 等. 基于复杂背景的红外弱小目标识别[J]. 激光杂志, 2022, 43(6): 91-95.
ZHOU P H, WANG F, WANG H F, et al. Infrared weak target recognition based on complex background[J]. Laser Journal, 2022, 43(6): 91-95.
- [2] 金宝根, 吕庆梅. 基于卷积神经网络的红外弱小车辆目标检测方法[J]. 激光杂志, 2024, 45(5): 241-245.
JIN B G, LYU Q M. Infrared small vehicle target detection method based on convolutional neural network[J]. Laser Journal, 2024, 45(5): 241-245.
- [3] ANJU T S, NELWIN RAJ N R. Shearlet transform based image denoising using histogram thresholding[C]//Proceedings of the 2016 International Conference on Communication Systems and Networks (ComNet). Piscataway: IEEE Press, 2016: 162-166.
- [4] 黄毅, 周纯, 刘欣军, 等. 基于YOLOv10的无人机复杂背景下多尺度检测模型[J]. 光通信研究, 2024: 1-8.
HUANG Y, ZHOU C, LIU X J, et al. Multiscale detection model for complex backgrounds in UAV images based on YOLOv10[J]. Study on Optical Communications, 2024: 1-8.
- [5] 肖雨晴, 杨慧敏. 目标检测算法在交通场景中应用综述[J]. 计算机工程与应用, 2021, 57(6): 30-41.
XIAO Y Q, YANG H M. Research on application of object detection algorithm in traffic scene[J]. Computer Engineering and Applications, 2021, 57(6): 30-41.
- [6] ZHANG N, DONAHUE J, GIRSHICK R, et al. Part-based R-CNNs for fine-grained category detection[C]//Proceedings of the 2014 Conference on European Conference on Computer Vision (ECCV). Cham: Springer International Publishing, 2014: 834-849.
- [7] SHEENY M, WALLACE A, EMAMBAKSH M, et al. POL-LWIR vehicle detection: convolutional neural networks meet polarised infrared sensors[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE Press, 2018: 1328-13286.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 779-788.
- [9] 王晓红, 陈哲奇. 基于YOLOv5算法的红外图像行人检测研究[J]. 激光与红外, 2023, 53(1): 57-63.
WANG X H, CHEN Z Q. Research on pedestrian detection in infrared image based on YOLOv5 algorithm[J]. Laser & Infrared, 2023, 53(1): 57-63.
- [10] 李阳, 赵娟, 严运兵. 基于改进型YOLOV5s的热红外道路车辆及行人检测方法[J]. 智能计算机与应用, 2022, 12(6): 33-38.
LI Y, ZHAO J, YAN Y B. Thermal infrared road vehicle and pedestrian detection method based on improved YOLOV5s[J]. Intelligent Computer and Applications, 2022, 12(6): 33-38.
- [11] 胡皓, 郭放, 刘钊. 改进YOLOX-S模型的施工场景目标检测[J]. 计算机科学与探索, 2023, 17(5): 1089-1101.
HU H, GUO F, LIU Z. Object detection based on improved YOLOX-S model in construction sites[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(5): 1089-1101.



- [12] 邓天民, 王丽, 刘旭慧. 基于注意力及特征融合的红外行人检测算法[J]. 重庆理工大学学报(自然科学), 2023, 37(6): 196-203.
DENG T M, WANG L, LIU X H. An infrared pedestrian detection algorithm based on attention and feature fusion[J]. Journal of Chongqing University of Technology (Natural Science), 2023, 37(6): 196-203.
- [13] 赵明, 张浩然. 一种基于跨域融合网络的红外目标检测方法[J]. 光子学报, 2021, 50(11): 1110001.
ZHAO M, ZHANG H R. An infrared object detection method based on cross-domain fusion network[J]. Acta Photonica Sinica, 2021, 50(11): 1110001.
- [14] CHEN Y H, ZHANG G P, MA Y J, et al. Small infrared target detection based on fast adaptive masking and scaling with iterative segmentation[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 7000605.
- [15] 周梦蝶, 黄昶. 复杂背景下改进的红外弱小目标检测[J]. 科学技术与工程, 2023, 23(23): 9999-10007.
ZHOU M D, HUANG C. Improved infrared dim small target detection under complex backgrounds[J]. Science Technology and Engineering, 2023, 23(23): 9999-10007.
- [16] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects[EB]. 2022.
- [17] LIU W, LU H, FU H, et al. Learning to upsample by learning to sample[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 6027-6037.
- [18] 朱开源, 吴佰靖, 高德勇, 等. 基于在线增强和跨尺度特征重建的雾天目标检测[J]. 计算机工程与应用, 2024: 1-13.
ZHU K Y, WU B J, GAO D Y, et al. Target detection in foggy days based on online enhancement and cross-scale feature reconstruction[J]. Computer Engineering and Applications, 2024: 1-13.
- [19] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [20] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Proceedings of the European Conference on Computer Vision (ECCV) 2016. Cham: Springer International Publishing, 2016: 21-37.
- [21] ZHOU X Y, WANG D Q, KRÄHENBÜHL P. Objects as points[EB]. 2019.
- [22] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2020: 10778-10787.
- [23] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB]. 2018.
- [24] WU W T, LIU H, LI L L, et al. Application of local fully convolutional neural network combined with YOLOv5 algorithm in small target detection of remote sensing image[J]. PLoS One, 2021, 16(10): e0259283.
- [25] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2023: 7464-7475.
- [26] WANG A, CHEN H, LIU L, et al. YOLOv10: real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.
- [27] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[EB]. 2024.
- [28] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: attention-centric real-time object detectors[EB]. 2025.

[作者简介]



黄国潮 (1986-), 男, 浙江交通职业技术学院汽车学院副教授, 主要研究方向为智能网联汽车。



常城 (2002-), 男, 长安大学汽车学院硕士生, 主要研究方向为智能网联汽车。



方虹苏 (2001-), 男, 长安大学汽车学院硕士生, 主要研究方向为智能网联汽车。



杨炜 (1985-), 男, 博士, 长安大学汽车学院讲师、硕士生导师, 主要研究方向为智能网联汽车。