

基于改进 YOLOv8 的夜间车辆检测研究

郭俊, 郭金霆[†], 张世龙

(盐城工学院汽车工程学院, 江苏 盐城 224051)

摘 要:针对夜间光照不足、小目标易遮挡等导致检测性能下降的问题,提出了改进模型 DFF-YOLOv8。该模型在 YOLOv8 基础上引入 Dysample 动态采样模块,自适应调整特征采样率以增强弱光小目标感知;融合 ASFF 特征融合机制,实现多尺度信息高效整合;并采用 Focaler IoU 损失函数动态优化回归精度。基于 BDD100K 夜间数据集实验结果显示,DFF-YOLOv8 在 mAP@50、mAP@50:95、 P 、 R 上分别提升 4.8%、3.8%、1.9%、4.1%,显著改善夜间低光环境下的检测性能。

关键词:夜间光照不足;动态采样;特征融合;损失函数;YOLOv8

中图分类号:TP391

文献标识码:A

Research on Nighttime Vehicle Detection Based on Improved YOLOv8

GUO jun, GUO Jinting[†], ZHANG Shilong

(School of Automotive Engineering, Yancheng institute of Technology, Yancheng, Jiangsu 224051, China)

Abstract: To address the decline in detection performance caused by insufficient illumination and occlusion of small targets at night, this paper proposes an improved model DFF-YOLOv8. Based on YOLOv8, the model introduces the Dysample dynamic sampling module, which adaptively adjusts the feature sampling rate to enhance the perception of small targets under low light. The ASFF feature fusion mechanism is integrated to achieve efficient multi-scale information aggregation, while the Focaler IoU loss function is adopted to dynamically optimize regression accuracy. Experimental results on the BDD100K nighttime dataset show that DFF-YOLOv8 improves mAP@50, mAP@50:95, precision, and recall by 4.8%, 3.8%, 1.9%, and 4.1%, respectively, significantly enhancing detection performance in low-light nighttime environments.

Key words: insufficient nighttime illumination; dynamic sampling; feature fusion; loss function; YOLOv8

近年来,随着机动车数量激增,夜间行车安全问题日益突出。受光照不足与能见度低影响,夜间场景对自动驾驶系统感知提出更高要求,尤其在远距离、弱光条件下小目标识别精度显著下降。因此,提升模型在夜间复杂环境中的检测准确性与响应速度成为研究重点。

现有夜间车辆检测方法主要基于深度学习目标检测框架,分为两阶段与单阶段检测器。前者如

Faster R-CNN^[1]、Mask R-CNN^[2] 和 R-FCN^[3] 等精度高但速度慢;后者如 YOLO^[4]、SSD^[5]、RetinaNet^[6] 等具备实时检测优势。针对夜间特性,研究者提出多种改进方案: Xu 等^[7] 引入 DCN 与 SABL 提升边界定位; Yi 等^[8] 基于 YOLOX 融合坐标注意力与多尺度特征增强小目标检测; Wu 等^[9] 在 YOLOv7 中集成 GhostNetV2、FReLU 与 Wise-IoU,实现性能与效率提升。

收稿日期:2025-07-31

基金项目:盐城应用基础研究计划资助项目(YCBK2025015)

作者简介:郭俊(1986—),男,安徽安庆人,博士,副教授,研究方向:机器视觉。

[†] 通信联系人, E-mail: GJTq965@163.com

尽管精度与速度有所改善,但极暗、模糊场景下仍存在特征表达不足、漏检率高的问题。Transformer 架构(如 DETR)虽具潜力,但在夜间检测中的优化仍有限。此外,针对光照干扰与运动模糊等噪声问题,尚需探索如对抗训练等使鲁棒性增强的策略。

为解决夜间漏检与低检测率问题,提出了 DFF-YOLOv8 模型,以提升复杂场景下的车辆识别能力。本模型基于 YOLOv8^[10] 框架,引入动态上采样、特征融合与改进损失函数,增强特征提取与回归精度。主要贡献如下:

(1) 引入 DySample 动态上采样模块,自适应调整特征采样以保留小目标空间细节,提升边界敏感性;

(2) 在检测头中融合 ASFF 特征融合机制,自适应加权多尺度特征,强化关键特征表达,提升小目标与遮挡检测能力;

(3) 采用 Focaler-IoU 损失替代 CIOU,通过非线性权重映射聚焦难分样本,提高回归稳定性;

1 改进 YOLOv8 模型

1.1 动态上采样器 (DySample)

在 YOLOv8 模型中,采用最近邻插值方法进行特征上采样。这种方法虽然加快了采样速度,但是也漏掉了深层特征中的信息,导致模型对图像细节的捕获能力降低。为提升模型的特征提取能力,引入了 DySample 方法^[11],如图 1 所示。DySample 是一种高效的动态上采样方法,通过点采样动态调整采样位置,取代了传统的插值或卷积方法。其关键优势在于能够根据图像内容的特定特征智能选择采样点。该方法不仅避免了额外的计算负担,还通过精确采样保留了图像中的关键信息,从而提升了模型在夜间车辆识别方面的准确性。

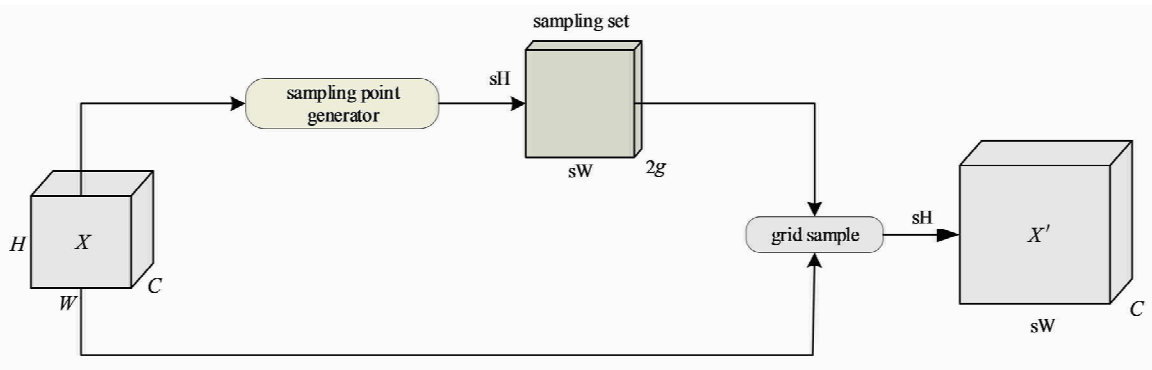


图 1 DySample 上采样过程

网格采样(Grid Sampling)是 DySample 的核心操作之一,具体来说,给定一个大小为 $C \times H_1 \times W_1$ 的特征图 X 和一个大小为 $H_2 \times W_2 \times 2$ 的采样点集 S ,其中 S 的第一维表示 x 和 y 坐标,网格采样函数使用 S 中的位置对假设的双线性插值 X 进行重新采样,得到大小为 $C \times H_2 \times W_2$ 的特征图 X' 这一过程定义为:

$$X' = \text{grid-sample}(X, S) \quad (1)$$

DySample 提供了一种保留动态上采样优势同时降低计算成本的方法。通过简化上采样过程,该方法减少了模型的复杂性,同时保持了对图像特征的敏感性,满足了实时目标检测任务的需求。这对于提升夜间车辆检测,提高检测速度具有重要意义。

1.2 自适应空间特征融合机制

为了提升模型在复杂背景下的多尺度识别能力,尤其是小目标检测精度,本文改进了 YOLOv8 的检测头结构,引入自适应空间特征融合(Adap-

tive Spatial Feature Fusion, ASFF) 机制,并增加了一条小目标检测分支。该设计允许模型根据不同空间位置的语义信息,自主选择各尺度特征的贡献比例,从而缓解传统特征金字塔中多层特征语义不一致所导致的信息冗余和融合失效问题。

ASFF^[12] 的工作流程如图 2 所示,主要包含两个阶段:特征尺度对齐与权重驱动的空间融合。具体来说,原始来自不同尺度(Level 1、Level 2、Level 3)的特征图在空间尺寸上存在差异,因此需通过上采样或下采样操作将其统一到相同的分辨率。上采样部分先进行 1×1 卷积压缩通道数后采用插值重建分辨率,下采样部分则通过步长为 2 的 3×3 卷积或最大池化实现特征压缩,确保三者融合前维度一致。

统一尺度后,ASFF 对每个空间位置定义一组融合权重,用以衡量该位置在不同层级特征中的信息有效性。该过程通过 1×1 卷积提取融合权重图,并使用 Softmax 函数^[13] 进行归一化,确保各尺

度权重之和为 1,表示为:

$$\alpha_{ij} + \beta_{ij} + \gamma_{ij} = 1 \tag{2}$$

最终,各层特征图按照该位置的权重进行逐元素加权叠加,生成融合后的输出特征。ASFF_1, ASFF_2, ASFF_3 表示应用了 ASFF 机制的不同层级的特征融合。以 ASFF_3 为例,其计算过程可表示为:

$$\text{ASFF}_3(i,j) = \alpha_{ij} \cdot x_{ij}^1 + \beta_{ij} \cdot x_{ij}^2 + \gamma_{ij} \cdot x_{ij}^3 \tag{3}$$

其中, $x_{ij}^1, x_{ij}^2, x_{ij}^3$ 分别代表来自三个不同尺度的特征图在 (i,j) 位置的响应值。融合后的特征图具有更高的空间自适应能力和语义表达一致性,为后续的检测任务提供了更稳定的语义支撑。

为进一步提升模型对小尺度目标的感知能力,本文在 YOLOv8 原有的三尺度检测头(对应于深层特征图 P3、P4、P5)基础上,新增了一条专门针对小目标的输出分支。原有的三条检测分支主要依赖于较低分辨率的深层特征图,这些特征图具备更强的语义抽象能力,因此更适合中、大目标的检

测。然而,这种低分辨率特征图在检测小目标时容易丢失边缘细节,导致夜间弱光和远距离场景下的小目标漏检率较高。相比之下,新增的小目标检测分支以浅层网络的高分辨率特征图(P2 层)作为输入,并在融合中低层特征信息的基础上保留了更多细粒度的空间结构信息,从而能够更精准地表征小尺度车辆目标的纹理与边缘特征。

新分支与原分支的区别主要在于:(1)输入不同:前者依赖深层语义特征,后者基于浅层高分辨率特征;(2)检测对象不同:原分支检测中大目标,小目标分支专注微小目标;(3)融合策略不同:小目标分支结合 ASFF 输出,更注重细节保留与空间定位精度。

ASFF 通过逐位置、逐通道的动态加权融合,充分利用多尺度特征的细粒度信息,提升了模型对尺度变化与目标形态差异的适应性。结合新增的小目标检测层,模型在夜间低亮度、远距离及遮挡场景下表现出更优的检测与定位性能。

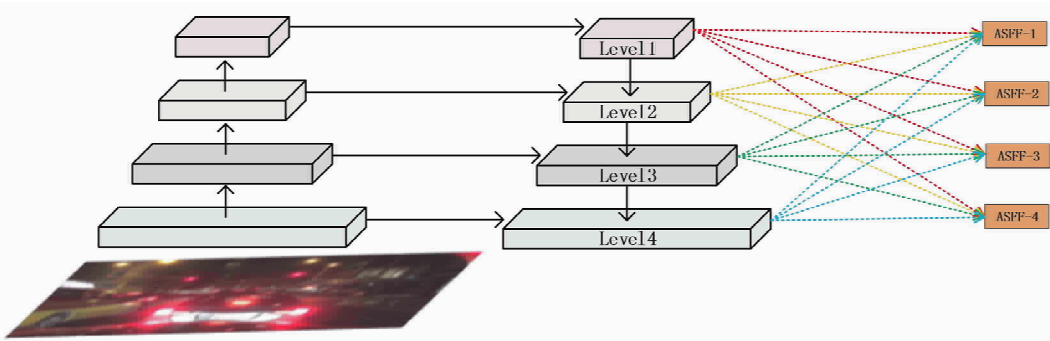


图 2 ASFF 工作流程

1.3 损失函数

在夜间车辆检测任务中,由于环境光照条件差、目标尺寸小、遮挡和噪声干扰严重,导致检测样本中存在大量定位困难的目标。在此类复杂场景下,YOLOv8 原始的 CIoU(Complete IoU)损失函数虽然在 IoU 基础上引入了中心点距离和宽高比惩罚项,从而提升了边界框回归的精度和收敛速度,但其优化策略对样本的难易程度缺乏区分处理。在训练过程中,CIoU 对所有回归样本赋予相同的权重,易使模型在梯度更新时过度聚焦于易回归样本,导致难样本(如夜间模糊、小尺度车辆)回归效果欠佳,进而影响整体检测性能。

为进一步增强模型在夜间复杂交通环境下的定位能力并缓解重叠检测框问题,采用 Focaler-IoU^[14] 损失函数以替代原有的 CIoU 机制。Foca-

ler-IoU 通过引入线性区间映射机制,对 IoU 值进行重构,并根据样本的重叠程度动态调节损失权重,从而实现了针对不同难度样本的自适应关注。该机制显著增强了模型在夜间复杂场景下对低质量预测框的回归能力,有效缓解了目标框重叠、定位偏差等问题,提升了边界框回归的鲁棒性与精度。Focaler-IoU 值的定义如下:

$$\text{IoU}_{\text{focaler}} = \begin{cases} 0, & \text{IoU} < d \\ \frac{\text{IoU} - d}{u - d}, & d \leq \text{IoU} \leq u \\ 1, & \text{IoU} > u \end{cases} \tag{4}$$

其中,IoU 是预测框与真实框的标准交并比, d,u 是线性区间的下限和上限超参数,满足 $0 \leq d < u \leq 1$,用于控制对样本难度的响应范围; $\text{IoU}_{\text{focaler}}$ 是经过线性重映射后的 IoU 值,用于构造新的损失函数。其损失定义如下:

$$L_{\text{Focaler-IoU}} = 1 - \text{IoU}_{\text{focaler}} \tag{5}$$

将 Focaler-IoU 应用于 YOLOv8 的回归分支后,模型在训练过程中能够更加聚焦于那些回归误差较大的困难样本,避免在样本分布不均的夜间场景中陷入对易样本的过拟合。通过动态调整梯度反馈,该方法不仅提升了边界框的位置回归精度,还在训练早期提供了更有效的优化方向,有助于加快收敛速度。实验结果表明,采用 Focaler-IoU 的检测模型在夜间复杂环境中对小目标、弱边界车辆的检测能力显著优于传统 CIOU 损失,验证了其在提升模型鲁棒性与实用性方面的有效性。

1.4 DFF-YOLOv8 整体结构

根据以上改进策略,设计出了一种名为 DFF-YOLOv8 的模型用于夜间道路车辆检测。其结构如图 3 所示,首先,将 YOLOv8 中原始的上采样模块替换为动态上采样(Dysample),模型能够根据局部特征自适应调整插值权重,从而在夜间车辆检测中更有效地保留边缘细节与小目标信息,显著提升弱光环境下的检测精度与鲁棒性。其次,将检测头引入自适应空间特征融合机制(ASFF),并增加一个小目标检测层,有助于显著提升模型在多尺度目标感知上的表现。两者协同作用,有效缓解了传统检测头在复杂场景下存在的特征表达不足与小目标漏检问题,整体增强了检测网络的精度与鲁棒性。最后,采用 Focaler-IoU 损失函数替代原始的 CIOU 损失函数,聚焦难检测样本并动态调整权重,增强了模型对多样化检测任务的适应性。

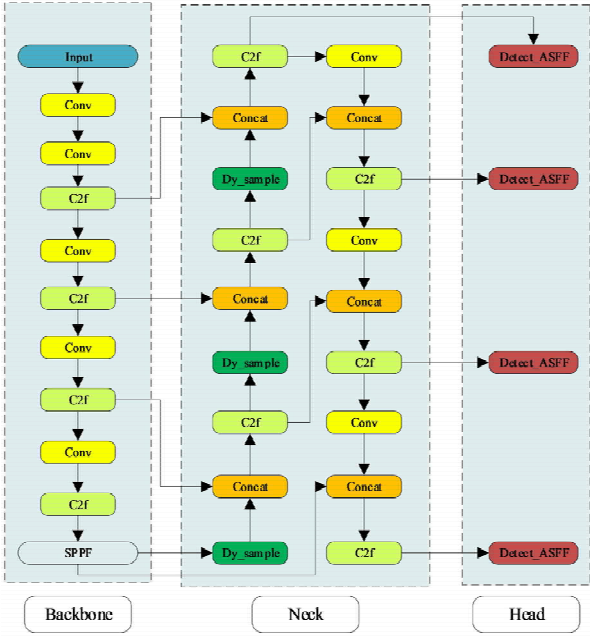


图 3 DFF-YOLOv8 网络结构

2 实验结果及分析

2.1 数据集

本实验所使用的数据集来源于公开的数据集 BDD100k。BDD100K(Berkeley DeepDrive 100K)是目前最具代表性的大规模自动驾驶视觉感知数据集之一,由加州大学伯克利分校发布,旨在支持对复杂交通场景下的多任务视觉学习研究。由于该数据集涵盖了美国不同城市、天气、光照和时间条件下采集的道路场景,因此筛选出 4500 张所需要的夜间车辆场景。其中,按照 7 : 2 : 1 的划分比例将其划分为训练集,验证集以及测试集。该划分策略有助于保障模型训练的充分性及评估结果的客观性。

2.2 实验环境

本实验是在本地电脑上进行训练的,操作系统为 Windows 11, CPU 为 i5-12400F, 显卡为 NVIDIA GeForce RTX 4060。搭建基于 CUDA 12.1 的 Pytorch 2. 1 深度学习框架,编程语言采用 Python 3. 10,如表 1 所示。

实验训练参数 epoch 大小为 200, Batch size 大小设置为 8, 初始图像大小为 640 × 640, 学习率动量为 0. 937, 权重衰减系数为 0. 0005, 采用 SGD 优化器并设置初始学习率为 0. 01。

表 1 实验环境

名称	环境配置
CPU	16 core,i5-12400F
GPU	NVIDIA GeForce RTX 4060
操作系统	Windows 11
深度学习框架	PyTorch 2. 1 Python 3. 10
Cuda 版本	12. 1

2.3 评价指标

为了评估所提出的模型的有效性,本研究利用了一套全面的指标,包括精确度(P),召回率(R), F1 值(F1), 平均精确度(AP), 平均精度(mAP), GFLOPs(每秒千兆浮点运算)和 FPS(每秒帧数)。Precision 计算所有阳性检测中正确识别的阳性检测的比例,而 Recall 测量正确识别的实际阳性实例的比例。精确度和召回率的定义如下:

$$p = \frac{TP}{TP + FP} \tag{6}$$

$$R = \frac{TP}{FN + TP} \tag{7}$$

其中, TP 表示真阳性示例的数量, FP 表示假

阳性示例的数量, FN 表示假阴性示例的数量, TN 表示真阴性示例的数量。

AP(即平均精度)是指测量目标类在不同阈值级别上的平均精度,定义如下:

$$AP = \int_0^1 P(R)dr$$

(8)

算法在识别和分类目标方面的精度使用检测精度的度量进行评估,特别是通过平均精度(mAP),并具有 0.5 的交集 (IOU) 阈值。该准确性指标的定义如下:

$$mAP = \frac{1}{n} \sum_{i=1}^n AP(i)$$

(9)

鉴于本研究的数据集仅关注车辆类别,因此设置 $n=1$,其中 n 表示不同类别的总数。检测速度以每秒帧数(FPS)为单位进行测量,FPS 反映了算法的帧速率或每秒处理的图像数量。此指标的计

算在式(10)中指定:

$$FPS = \frac{1000}{\delta \cdot \beta \cdot NMS}$$

(10)

其中, δ 是图像预处理时间, β 表示推理速度,NMS 表示后处理时间。

2.4 实验结果

2.4.1 消融实验

为验证各改进模块在夜间车辆检测任务中的有效性与贡献,本文设计了消融实验以评估其对模型性能的实际影响。在保证参数配置条件不变的情况下,得到的实验结果如表 2 所示,原始 YOLOv8 的 map 为 67.7%,而经过改进之后的 DFF-YOLOv8 在 map 的表现上提升了 4.8 个百分点。同时,在 mAP0.5:0.95, P , R 等评价指标上分别提升了 3.8%,1.9%和 4.1%。

表 2 消融实验结果

模型	Dysample	FASFFHead	FocalerIoU	mAP/%	mAP0.5 : 0.95/%	P	R
YOLOv8n				67.7	37.2	76.6	58.3
D- YOLOv8	✓			68.6	38.3	77.2	59.3
F- YOLOv8		✓		71.2	37.9	77.3	60.8
F- YOLOv8			✓	68.3	37.4	76.2	58.8
DF- YOLOv8	✓	✓		71.3	40.2	78.1	60.8
DF- YOLOv8	✓		✓	68.3	38.4	76.3	59.2
FF- YOLOv8		✓	✓	71.9	40.1	77.1	61.8
DFF- YOLOv8	✓	✓	✓	72.5	41.0	78.5	62.4

2.4.2 对比实验

为了验证所提出方法的有效性,在 BDD100K 数据集上与主流目标检测算法进行了充分的对比实验,包括 YOLOv8 原始版本、YOLOv9、YOLOv10、YOLOv11、YOLOv12、SSD 以及 Faster R-CNN 等。在实验中,为保证对比的公平性,所有模型均采用相同的训练数据集划分、训练轮数以及优化器参数等设置,确保在相同实验条件下进行对比,如表 3 所示。结果表明,本文提出的 DFF-YOLOv8 改进模型在精度方面均远优于其他主流检测模型,进一步验证了所设计模块在提升检测性能方面的优势。

2.5 可视化分析

为进一步直观展示所提出改进模型在目标检测任务中的性能表现,对模型的检测结果进行了可视化分析,如图 4 所示。通过对测试集中部分图像的预测输出进行标注展示,可以观察到原始 YOLOv8n 在近距离与远距离暗光环境检测中,均存在漏检的情况。在近距离亮光检测方面,其没有

改进后的 DFF-YOLOv8 识别率高且改进后的检测算法目标边界框定位更加精确。与原始 YOLOv8 模型相比,本文模型在对暗光环境下目标的识别能力有明显提升,验证了所引入的 FAS-FF 模块与小目标检测头在特征提取与融合方面的有效性。可视化结果进一步印证了量化指标中精度提升的真实性和可靠性。

表 3 对比实验

模型名称	输入尺寸	mAP/%	mAP0.5 : 0.95/%
SSD	300 * 300	39.1	18.7
Fast R-CNN	600×600	37.2	19.4
YOLOv8n	640×640	67.7	37.2
YOLOv9t	640×640	68.3	38.4
YOLOv10n	640×640	66.8	36.6
YOLOv11n	640×640	67.6	37.1
YOLOv12n	640×640	65.4	36.8
DFF-YOLOv8	640×640	72.5	41.0

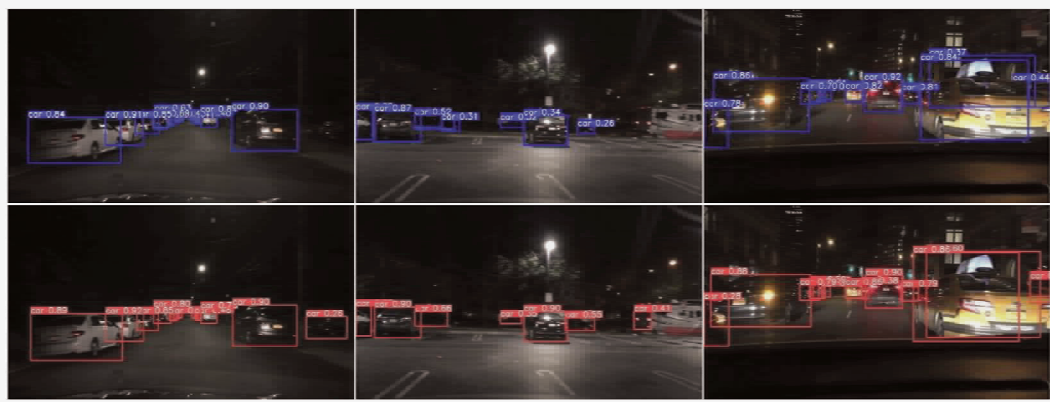


图 4 可视化检测结果对比图

3 结 论

针对夜间车辆检测中存在的低照度、强干扰、小目标识别困难等问题,提出了一种改进的目标检测模型 DFF-YOLOv8。在保持模型整体结构简洁的基础上,通过特征增强与损失函数优化,提高了检测性能。实验验证显示,该方法在夜间复杂环境下具备良好的实用性和鲁棒性,具体结论如下:

(1)本模型引入了 Dysample 模块与 FASFF 融合策略,有效增强了对多尺度特征的提取与上下文信息的利用能力,特别是在弱光与远距离条件下,对小目标的识别效果提升显著。

(2)在损失函数设计方面,采用 Focaler-IoU 替代传统的 CIoU,提高了模型对难以定位目标的关注程度。该策略使得检测模型在回归阶段更加稳定,并显著降低了漏检率。

(3)在 BDD100K 数据集上测试,改进模型在 $mAP@0.5$ 上达到 72.5%,在 $mAP@0.5:0.95$ 上达到 41.0%,相较于原始 YOLOv8n 提升了 4.8% 和 3.8%,验证了各模块改进的有效性。

(4)进一步对比 YOLOv9t、YOLOv10 和 SSD 等主流方法,DFF-YOLOv8 在准确率方面表现稳定,尤其在复杂夜间场景中具备更强的特征表达能力和适应性,能够更好地完成夜间车辆及行人目标的检测任务。

综上,DFF-YOLOv8 模型通过结构设计与损失优化,在不显著增加计算开销的前提下,实现了夜间目标检测性能的全面提升,为后续在实际交通监控与智能驾驶等领域的部署提供了可行的技术基础。

参考文献

[1] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal net-

wlks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137—1149.

[2] HE K, GKIOXARI G, DOLLAR P, et al. Mask r-cnn[C]// Proceedings of the IEEE International Conference on Computer vision, 2017: 2961—2969.

[3] DAI J, LI Y, HE K, et al. R-FCN: object detection via region-based fully convolutional networks[C]//Proceedings of the 30th International Conference on Neural Information Processing System, 2016: 379—387.

[4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779—788.

[5] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2016: 21—37.

[6] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980—2988.

[7] XU Y, CHU K, ZHANG J. Nighttime vehicle detection algorithm based on improved faster-rcnn[J]. IEEE Access, 2023, 12: 19299—19306.

[8] YI K, LUO K, CHEN T, et al. An improved YOLOX model and domain transfer strategy for nighttime pedestrian and vehicle detection [J]. Applied Sciences, 2022, 12 (23): 12476.

[9] WU Y, TANG Y, YANG T. An improved nighttime people and vehicle detection algorithm based on YOLO v7[C]//2023 3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE). IEEE, 2023: 266—270.

[10] SOHAN M, THOTAKURA S R, CHIRRA V R R. A review on yolov8 and its advancements[C]//International Conference on Data Intelligence and Cognitive Informatics. Springer, Singapore, 2024: 529—545.

[11] LIU W, LU H, FU H, et al. Learning to upsample by learning to sample[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 6027—6037.

[12] LIU S, HUANG D, WANG Y. Learning spatial fusion for single-shot object detection[J]. ArXiv Preprint ArXiv, 2019: 1911.09516.

[13] 刘海莹, 乔瑞秀, 陈刚, 等. 一种高效的 Softmax 函数计算方法及硬件电路[J]. 微电子学与计算机, 2024, 41(2): 91—100.

[14] ZHANG H, ZHANG S. Focaler-iou: more focused intersection over union loss[J]. ArXiv Preprint ArXiv, 2024: 2401.10525.