【工程与技术】

基于改进 YOLOv8 的复杂道路场景目标检测算法

董　凤[1],张　鑫[1],孔　林[2],张　涛[1]

(1.信阳学院 理工学院,河南 信阳 464000;
2.浙江犀重新能源汽车技术有限公司,浙江 金华 321000)

摘　要:为改善复杂道路场景中小目标漏检、误检问题,提出了一种基于改进 YOLOv8 的复杂道路场景目标检测算法。首先,在特征提取网络中引入 RFCAConv 模块,以提升模型对关键特征信息的抓取能力;其次,为丰富深层次网络中的上下文信息,在聚合网络中构建了上下文感知模块 CAM;最后,将原本的损失函数 CIoU 替换为 Inner-CIoU,以增强模型的泛化能力。实验结果表明:在 New_BDD 公开数据集上,本算法模型检测精度 mAP 可达 77.6%,推理速度 FPS 达 98 帧/s,优于其他主流算法。同时,在图像检测结果的可视化实验中,该算法模型在面对多种复杂道路场景的图像实例中有效改善了原有模型的漏检和误检等情况。

关键词:目标检测算法;小目标;YOLOv8;RFCAConv;上下文感知;Inner-CIoU

中图分类号:TP 391.4　　**文献标识码**:A　　DOI:10.13486/j.issn.2097-4973.2026.02.009

深度学习目标检测技术发展迅猛,已在智能交通、工农业检测、安防监控等领域得到广泛应用。在智能交通领域,随着自动驾驶技术的迅速发展,如何精准识别和实时定位车辆、行人和道路标志等交通目标信息已成为环境感知技术中的核心目标[1]。因此,目标检测技术的不断发展和优化对推动交通领域的创新和进步具有重要意义。

主流目标检测算法主要有一阶段算法(One-Stage)和二阶段算法(Two-Stage),两者的区别主要在于二阶段算法基于候选框,而一阶段算法基于回归。二阶段检测算法虽检测精度较高,但检测速度相对较慢,难以满足工程应用中实时检测的需求,典型算法主要有基于区域的卷积神经网络(Regions with Convolution Neural Network,R-CNN),如 Faster R-CNN[2] 和 Mask R-CNN[3] 等;而基于回归的一阶段检测算法在保持检测速度的同时,具备较高的检测精度和定位准确性,经典算法主要有单一的多锚框检测器(Single Shot multibox Detector,SSD)[4] 和 YOLO(You Only Look Once)系列[5-7] 等。YOLO 系列凭借简易的框架体系和高检测速度一直受到业内人士的青睐,YOLO 家族也因此不断壮大,其中,YOLOv8 是 Ultralytics 官方推出的升级版本,在精度、速度、部署适配等方面都更先进,也逐渐成为工业项目落地中的主流选择。

此前已有不少学者展开对复杂道路环境下目标检测的研究。薛阳等针对车辆行驶过程中易受光照影响的问题,提出一种车辆检测算法 TSA-YOLO,通过设计注意力模块,增强对运动目标的关注,有效提升

收稿日期:2025-08-02

基金项目:信阳学院校级科研项目(2024-XJLYB-009)

第一作者简介:董　凤(1997—　),女,河南信阳人,讲师,硕士,主要从事深度学习目标检测研究。
E-mail:1773041460@qq.com

了车辆在光照干扰下的目标定位精度[8]；霍华等为缓解复杂场景下的行人检测网络不够轻量化的问题，通过优化网络模型，提出一种轻量高效的行人检测算法 SCD-YOLO[9]；孙海明等为实现无人驾驶汽车对交通标志的精准识别，基于 YOLOv4 提出了一种改进的交通标志图像识别算法 YOLO-slim[10]。以上研究主要是将车辆、行人以及交通标志单独检测，检测目标单一，忽略了在道路驾驶环境下“车-人-物”的统一性，需整合多种检测目标算法才能部署到车载单元，否则既增大了对移动端计算资源的消耗，也难以实现自动驾驶在复杂道路交通环境下的智能决策。鉴于此，本文选用 YOLOv8 为基准算法，基于真实驾驶场景的数据集 BDD-100K，开展面向车辆、行人以及交通标注复杂道路场景的目标检测算法研究。

1 模型方法

YOLOv8 模型延续了 YOLO 架构的经典设计，主要由输入端（Input）、骨干（BackBone）网络、颈部（Neck）网络以及检测头预测端（Prediction）四部分组成。在输入端，YOLOv8 保留了 YOLO 系列的 Mosaic 数据增强、Anchor 自适应以及图片缩放等数据预处理操作，便于为模型适配不同的输入分辨率。骨干网络也称特征提取网络，主要用于特征提取，由标准卷积（Conv）模块、C2f 模块及 SPPF（Spatial Pyramid Pooling Fast）模块搭建而成；C2f 中包含了多个残差瓶颈结构，有利于信息的深度传递；SPPF 采用空间金字塔池化在不同尺度的特征图上提取特征，从而使模型能更好地捕捉不同尺寸目标的信息；颈部网络也称特征聚合网络，主要对 Backbone 提取到的特征信息进行深度聚合，使用 FPN＋PANet 的聚合网络结构，FPN 自顶向下的信息流和 PANet 自底向上的信息流对错交互地将高层语义信息和底层位置信息从不同主干层对不同检测层进行参数聚合，以此增强网络特征融合能力。检测头预测端位于整个目标检测流程的末尾阶段，为不同尺寸的检测目标提供了 3 个不同尺寸的特征图，同时采用解耦头（decoupled-head）[11]结构，将分类和回归任务分离，从而提高检测精度。

2 改进算法

本文提出的算法基于 YOLOv8 改进而来，其模型结构如图 1 所示。在骨干网络部分，通过引入一个基于感受野注意力机制的 RFCAConv 模块，来提高模型对小目标关键特征信息的解析能力。在颈部网络部分，设计了一个上下文感知模块 CAM，并将其放置于特征提取网络和特征聚合网络之间，为特征聚合网络注入丰富的特征语义的深层特征信息。此外，为进一步提高模型对小目标的泛化能力，引入了基于辅助边界框的损失函数 Inner-CIoU。本文所提出的模型可在降低计算量和参数量的同时，提高检测精度，特别适用于复杂道路场景下的目标检测。

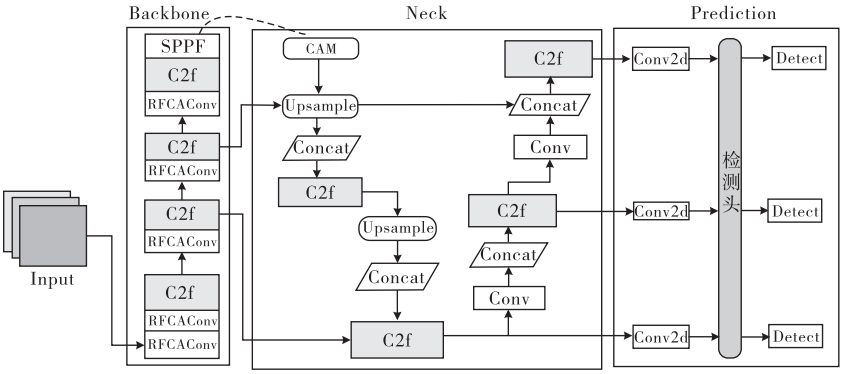

图 1 改进 YOLOv8 模型结构图

2.1 RFCAConv 模块

受到注意力机制可使模型聚焦于显著特征的启发，本文将 RFCAConv（Receptive Field Channel Attention Convolution）[12]安插于骨干特征提取网络。RFCAConv 是一种结合了感受野注意力机制的卷

积模块，通过引入注意力机制动态生成卷积参数，使得卷积核的参数能根据输入特征图的不同空间位置和通道进行自适应调整，从而缓解原有标准卷积核特征信息不足的问题，进而增强模型对小目标特征信息的捕获能力，能够捕获图像中更广泛的上下文信息和长距离依赖关系，RFCAConv 结构示意图如图 2 所示。

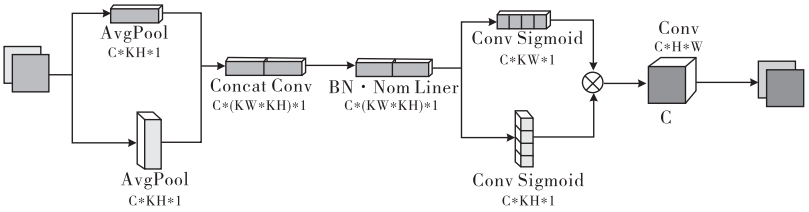

图 2　RFCAConv 结构示意图

RFCAConv 卷积模块首先对输入特征图通过平均池化操作进行局部感受野划分，提取每个感受野区域内的空间特征，为后续生成自适应的卷积参数提供重要的上下文指导；其次通过引入的一个参数生成网络动态及一系列非共享卷积权重张量，这些非共享卷积权重会根据不同的感受野区域内的特征信息自适应地匹配不同的权重系数，为模型提供最适宜的特征提取器，最大限度地保证卷积操作后的输出特征图中涵盖着最丰富的上下文信息，从而在一定程度上弥补原有卷积对图像特征信息提取不足和小目标特征图易丢失等问题；再次借助注意力机制计算感受野内每个位置的权重，进而生成位置注意力图，并将编码后的位置注意力图与原始输入特征图进行逐点相乘，从而突出关键区域的特征响应，抑制无关背景信息；最后对加权后的特征图进行标准卷积运算，输出最终的增强特征图，进一步提取高层语义特征。

RFCAConv 模块具有三大优势：其一，较坐标注意力，RFCAConv 通过使用非共享参数的方式，在应对大卷积核时能更有力地处理由于特征信息不足导致的关注度不够的问题；其二，较常规卷积，RFCAConv 对感受野中每个特征的重要性均有不同程度的关注，能较大程度地规避特征提取过程中的信息丢失；其三，较自注意机制，RFCAConv 是一种更加轻量化的卷积操作。面对复杂道路场景中远距离的小目标，RFCAConv 的引入使算法模型在享有大卷积核广阔感受野的同时，获得了远比常规卷积更丰富、更鲁棒的特征信息，通过非共享参数使卷积核参数内容自适应，能更有效地聚焦并捕捉这些小目标信息，进而提升检测模型对小目标检测的精确度和效率。

2.2　上下文感知模块 CAM

受 ResNet[13] 残差结构中“shortcut”连接的启发，本文设计了一种多分支膨胀卷积网络架构。该模块通过并行部署不同扩张率的膨胀卷积层，实现了对特征空间中远近距离信息的有效捕捉。在轻微增加计算量的前提下，该设计模块将浅层特征信息传递至深层网络空间，显著增强了深层特征的上下文表征能力。基于这一特性，将该模块命名为上下文感知模块(Context-Aware Module，CAM)，其结构如图 3 所示。

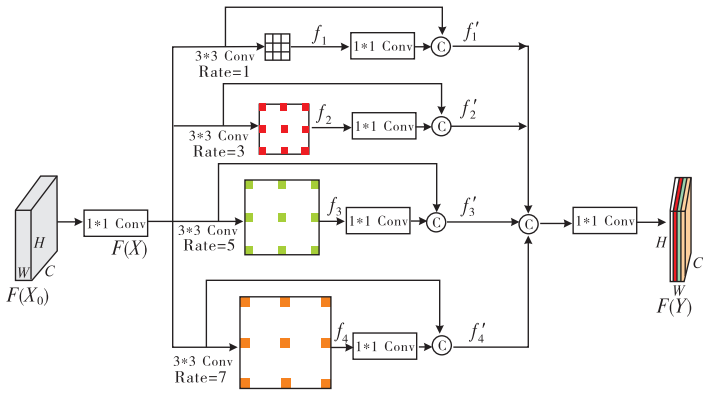

图 3　环境上下文感知模块结构图

对于一个输入特征图 $F(X_0)\in R^{H*W*C}$，首先采用四个并行膨胀卷积(扩张率分别为 1、3、5、7)对其进

行卷积操作；然后与输入特征图做短路连接以进行初步融合，初步融合后，浅层特征图通过短路路径将细节信息直传给深层特征图，以对深层特征空间捕获的高级语义信息做进一步的补充和完善，进而弥补相邻信息间的不连续性，同时减少计算参数；最后对四个分支 $f_i'(X)\in R^{H*W*C}$ 特征图做通道拼接(Concat)以汇聚多分支之间的上下文信息。如此一来，CAM 模块通过四个不同扩张率的残差膨胀卷积分支，可得到尺寸相同但感受野不同的特征图 $F(Y)\in R^{H*W*C}$。CAM 模块凭借多分支锯齿状的膨胀卷积结构能循序渐进地捕获图像的不同区域信息，极大地提高了网络模型的特征表征能力。

2.3 Inner-IoU 损失函数

YOLOv8 的定位损失函数沿用了 YOLOv5 的 CIoU 损失函数，其公式为

$$L_{\mathrm{CIoU}}=1-IoU+\frac{\rho^2(b,b^{gt})}{c^2}+av,IoU=\frac{|B\bigcap B^{gt}|}{|B\bigcup B^{gt}|}。$$

在实际自动驾驶场景中，车辆和行人这两类检测目标极易出现流量较大的情形。由于各目标之间特征相似且高度重叠，CIoU 仅通过增添新损失项来加速模型收敛，忽略了 IoU 损失项本身的局限性，无法对上述复杂道路场景进行自适应调整，以致存在较高的误检率和漏检率。鉴于此，本文将原有的 CIoU 函数替换为 Inner-IoU[14] 损失函数。Inner-IoU 损失函数的相关计算公式如下：

$$b_l^{gt}=x_c^{gt}-\frac{w^{gt}r}{2},b_r^{gt}=x_c^{gt}+\frac{w^{gt}r}{2},b_t^{gt}=y_c^{gt}-\frac{h^{gt}r}{2},b_b^{gt}=y_c^{gt}+\frac{h^{gt}r}{2},$$

$$b_l=x_c-\frac{wr}{2},b_r=x_c^{gt}+\frac{wr}{2},b_t=y_c-\frac{hr}{2},b_b=y_c^{gt}+\frac{hr}{2},$$

$$i=[\min(b_r^{gt},b_r)-\max(b_l^{gt},b_l)]\times[\min(b_b^{gt},b_b)-\max(b_t^{gt},b_t)],$$

$$u=w^{gt}h^{gt}r^2+whr^2-i,IoU^{\mathrm{inner}}=\frac{i}{u}。$$

式中：b_l^{gt}、b_r^{gt}、b_t^{gt} 和 b_b^{gt} 分别代表真实框的上下左右四个边界坐标，b_l、b_r、b_t 和 b_b 分别代表预测框的上下左右四个边界坐标；x_c^{gt} 和 y_c^{gt} 分别表示真实框和内部真实框的中心点，x_c 和 y_c 分别表示预测框和内部预测框的中心点；w^{gt} 和 h^{gt} 分别表示真实框的宽和高，w 和 h 分别表示预测框的宽和高；r 为尺度因子比，其范围通常在[0.5,1.5]。

Inner-IoU 损失函数通过引入尺度因子比 r 对辅助边界框大小进行调节，如图 4 所示。该损失函数通过使用较小尺度的辅助边界框来计算 IoU，有助于更好地回归高 IoU 样本，进而加速收敛过程；反之，使用更大尺度的辅助边界框计算 IoU，则能够有效促进 IoU 样本的回归过程。

本文将 Inner-IoU 损失应用于现有 CIoU 的边界框回归损失函数，可用公式表示如下：

$$L_{\text{inner-CIoU}}=L_{\mathrm{CIoU}}+IoU-IoU^{\mathrm{inner}}。$$

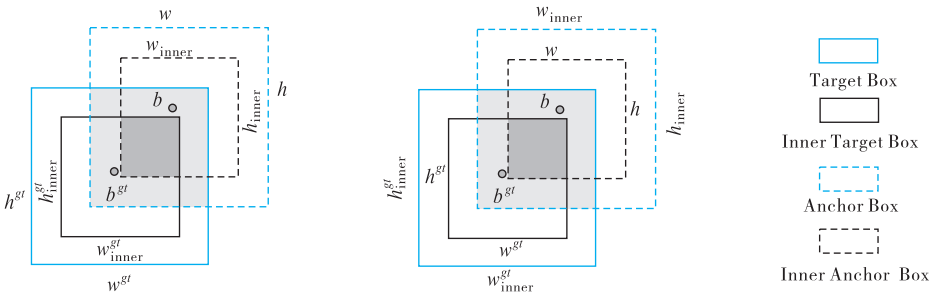

图 4 Inner-IoU 示意图

3 实验环境及设备

3.1 实验环境与参数设置

本实验采用的操作系统为 Windows 10，硬件配置为 Gen Intel(R) Core(TM) i7-11800H 处理器，搭

载 NVIDIA GeForce RTX 3060 显卡，显存 12 GB。软件环境为 CUNA 11.8、Pytthon 3.8 和 Pytorch 2.0.0。为加速模型训练，将 Moscaic 数据增强设置为开启状态，在模型训练时，神经网络的初始学习率为 0.01，动量为 0.937，权重衰减系数为 0.000 5，批量大小设置为 8，迭代轮次为 300 轮。

3.2　数据集准备

本实验选用源于真实驾驶场景的 BDD100K 数据集[15]，该数据集涵盖的驾驶场景丰富，数量庞大，标注信息清楚，图 5 为该数据集中的几个路况场景的实例图。原 BDD100K 数据集共有 10 万张图像，基于实验设备和训练时间的综合考量，本文选择的训练集、验证集和测试集的图片数量依次为 7 000、2 000、1 000。

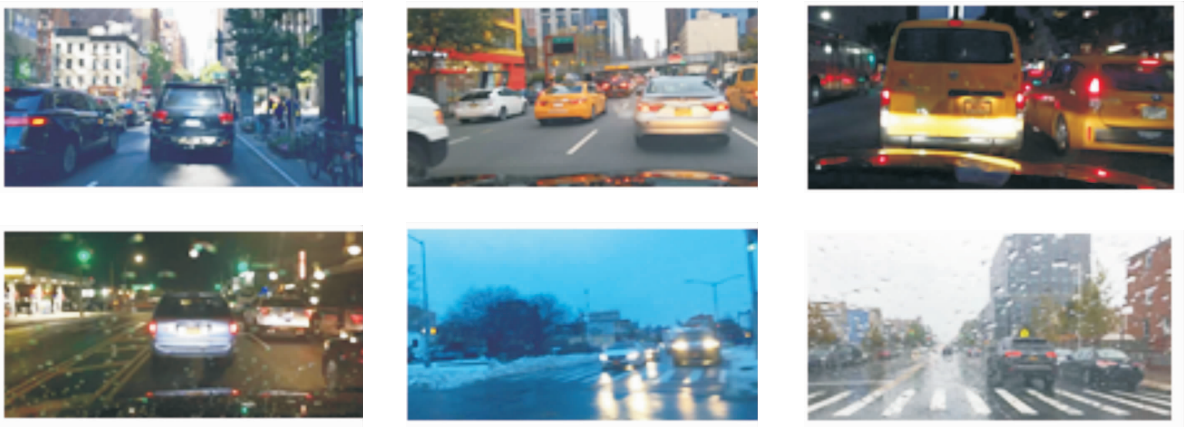

图 5　数据集实例图

原数据集标注的 10 类目标分别是公共汽车、交通灯、交通标志、行人、自行车、卡车、汽车、火车、骑行者、摩托车。考虑到在实际交通道路场景中，影响行车安全的外在因素主要是车辆、行人、交通标志、交通灯等常规物体，故用 Python 脚本将卡车、公共汽车、摩托车三类兼并到汽车类别，并滤除罕见类火车。最终优化后的数据集一共包含汽车、行人、交通标志及交通灯四类，并将新的数据集命名为 New_BDD。表 1 即为 New_BDD 数据集中各类别标签数目的统计情况。

表 1　New_BDD 数据集各类别样本统计情况

类别	图片/张	汽车/辆	行人/位	交通标志/个	交通灯/个
训练集	7 000	72 983	30 158	54 627	42 356
验证集	2 000	22 456	20 287	10 598	9 868

3.3　评价指标

本文采用精确率(P)、召回率(R)、平均精度(AP)及平均精度均值(mAP)为算法模型评估指标，各指标计算公式如下：

$$P=\frac{TP}{TP+FP},R=\frac{TP}{TP+FN},\ AP=\int_0^1 P\mathrm{d}R,mAP=\frac{1}{c}\sum_{i=1}^{c}AP_i\ 。$$

式中：TP、FP 分别代表被模型正确预测、错误预测的正样本数，FN 为被模型错误预测的负样本数，i 是缺陷类别，c 为缺陷类别数。本文主要采用 AP、mAP 以及 FPS 作为评估指标。AP 表示各类交通标志目标的平均精度，mAP 表示 $IoU=0.5$ 时的平均精度均值，FPS 用来衡量检测速度。

4　实验结果与分析

4.1　消融实验

为系统评估各改进组件的有效性，本文基于 YOLOv8 基准模型设计了消融实验。如表 2 所示，采用渐进式优化策略，依次引入 RFCAConv 模块、CAM 模块及 Inner-CIoU 函数，并定量分析各阶段模型的

AP 和 mAP 变化。由表 2 可知，随着各改进模块的依次融入，模型的检测精度呈现阶梯式提升。其中，实验 1 是原基准算法，其 mAP 仅为 68.8%，交通灯和交通标志两个小型目标的检测精度均在 70%以下。实验 2 将实验 1 中的骨干网络的卷积模块全都更换为 RFCAConv 模块后，小尺度目标交通灯和交通标志的 AP 分别提高了 2.7 和 1.0 个百分点。这是因为 RFCAConv 模块增强了模型对小目标和复杂特征的提取能力，从而提升了检测精度。实验 3 单独引入 CAM 模块，较之实验 1，mAP 提高了 6.2 个百分点，其中交通灯和交通标志类别的 AP 值分别提高了 6.9 和 5.8 个百分点。性能提升主要归因于 CAM 模块的多分支膨胀卷积结构，该设计通过扩大感受野并融合多尺度特征，有效增强了网络对上下文语义信息的提取能力，从而显著改善了小目标的特征表征质量。实验 4 将基准网络中的损失函数替换为 Inner-CIoU，Inner-CIoU 通过辅助边框计算 IoU 损失，进一步优化目标边界框的定位损失函数，在不同尺度下优化模型，能够更加精准地调整和优化目标检测结果，mAP 提升至 75.8%充分证明了在模型优化过程中，选择更加合适的损失函数对提升模型性能的重要性。实验 5 是将 RFCAConv 模块和 CAM 模块共同引入模型中，实验 6 则是同时采用 CAM 模块和 Inner-CIoU 损失函数。两组对照实验结果表明，相较于前面三组单个改进措施的引入实验，两个改进点共同作用也可使模型的 mAP 和各小目标的检测精度得到一定提升，这从侧面证明了本文提出的改进策略的可行性。

表 2 消融实验

实验序号	实验模块			$AP(IoU=0.5)$/%				mAP/%
	RFCAConv	CAM	Inner-CIoU	汽车	行人	交通标志	交通灯	
1				79.8	60.7	66.1	68.4	68.8
2	√			82.2	63.1	68.8	69.4	70.9
3		√		85.5	67.4	73.0	74.2	75.0
4			√	86.2	68.1	73.8	75.0	75.8
5	√	√		86.8	68.6	74.2	75.6	76.4
6		√	√	87.1	68.9	74.8	76.0	76.9
（本文算法）	√	√	√	87.7	70.5	75.6	76.4	77.6

本文方法将 RFCAConv 模块、CAM 模块以及 Inner-CIoU 损失函数依次加入基准算法。实验结果表明，mAP 比基准模型提升 8.8 个百分点，且各检测类别的精度均获得显著提升，这一结果充分验证了本文算法在复杂道路场景下的目标检测中的有效性，尤其在小目标检测方面展现出明显的性能优势。

4.2 对比实验

为充分验证本文改进算法的性能优势，将其与多种主流目标检测算法在相同训练环境下进行对比实验，以全面评估算法的综合性能，包括二阶段检测算法（Faster R-CNN）、一阶段检测算法（SSD）、同系列高性能模型（YOLOv5、YOLOv7、YOLOv9、YOLOv10）及基准算法 YOLOv8。实验结果如表 3 所示。

表 3 各算法在 New_BDD 数据集上的性能对比

算法模型	P/%	R/%	mAP/%	FPS/(帧·s^{-1})	算法模型	P/%	R/%	mAP/%	FPS/(帧·s^{-1})
Faster R-CNN	71.8	60.4	63.5	31	YOLOv8	78.9	63.7	68.8	93
SSD	70.2	58.3	58.8	36	YOLOv9	79.2	64.1	69.9	96
YOLOv5	72.4	58.8	60.9	89	YOLOv10	79.9	65.6	72.5	101
YOLOv7	75.6	60.1	63.4	90	本文算法	82.6	68.4	77.6	98

本文算法的检测精度 mAP 明显优于其他算法。与二阶段算法 Faster R-CNN 相比，mAP 提高了 14.1 个百分点；与一阶段 SSD 相比，mAP 提高了 18.8 个百分点；与同系列 YOLOv5、YOLOv7 相比，

mAP 分别提高 16.7、14.2 个百分点；与算法 YOLOv9、YOLOv10 相比，*mAP* 依次提高了 7.7、5.1 个百分点。相比之下，本文算法在复杂道路场景检测任务中的检测精度最佳。本文算法的检测速度 *FPS* 为 98 帧/s，仅次于最新模型 YOLOv10 的检测速度 101 帧/s，但优于其他算法的推理速度，并达到了嵌入式车载移动端设备实时检测的速度要求。

4.3　可视化分析

为更直观地展现本文所提改进策略的有效性，在 New_BDD 测试集上对改进前后两个算法展开可视化分析，从而进一步验证改进策略的显著作用。图 6～9 为 YOLOv8 和本文算法在闹区、逆光、晚间以及雨天等不同复杂场景下的检测效果图。其中，左图为 YOLOv8 算法的检测效果图，右边则为本文提出的改进算法的检测效果图。

图 6　闹区检测效果对比图

图 7　逆光检测效果对比图

图 8　晚间检测效果对比图

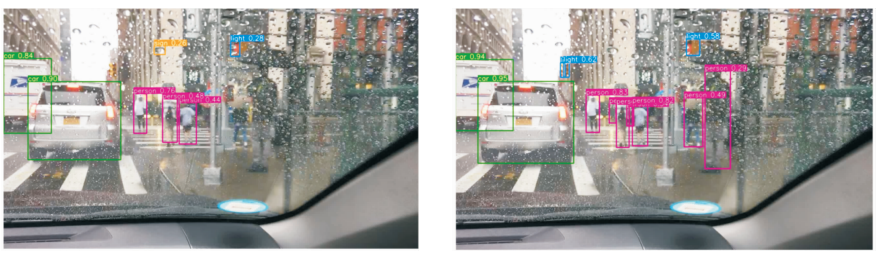

图 9　雨天检测效果对比图

从这四组对比图可以清晰地看出，无论是在车流量和人流量较密集的闹区、还是在逆光环境下以及在检测视觉受阻的夜晚和雨天等复杂的道路检测环境下，YOLOv8 算法漏检率较高，特别是远处尺寸较小的交通指示牌漏检严重。对比之下，本文算法对以上复杂的道路场景，能检测出更多的目标，有较强的抗

干扰性和泛化性。

5 结论

为缓解复杂道路检测场景中小目标检测效果欠佳的问题，本文以 YOLOv8 为模板算法，选用基于感受野注意力机制的 RFCAConv 卷积模块对模板算法的骨干网络进行更新，以提升模型对关键特征信息的关注度；为将深层次的特征语义信息有效传递到颈部网络，在颈部网络的前端搭建了一个上下文感知模块 CAM，以增强模型对小目标的关键特征信息的提取能力；最后将原损失函数替换为 Inner-CIoU 损失，以增强模型泛化能力。在 New_BDD 数据集上，通过消融实验和对比实验验证了本文算法既能提高模型的检测精度，又能达到较高的检测速度。通过可视化分析可知，本文算法在多种复杂的道路场景下改善了小目标漏检情况，大大降低了小目标漏检率。

参考文献：

[1] 寻永恒，王旭成，杨兴平. 智能网联汽车环境感知技术研究[J]. 汽车知识，2025，25(8)：31-33.

[2] REN S，HE K，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.

[3] HE K，GKIOXARI G，DOLLÁR P，et al. Mask R-CNN[C]//2017 IEEE international conference on computer vision. Venice：IEEE，2017：2980-2988.

[4] LIU W，ANGUELOV D，ERHAN D，et al. Ssd：single shot multibox detector[C]//LEIBE B，MATAS J，WELING M(eds). Computer Vision-ECCV 2016. Cham：Springer，2016：21-37.

[5] 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. Las Vegas：IEEE，2016：779-788.

[6] REDMON J，FARHADI A. YOLO9000：better，faster，stronger[C]//2017 IEEE conference on computer vision and pattern recognition. Honolulu：IEEE，2017：6517-6525.

[7] REDMON J，FARHADI A. Yolov3：an incremental improvement[PP/OL]. V1.（2018-04-08）[2025-07-23]. https：//arxiv. org/abs/1804. 02767.

[8] 薛阳，徐笑，卢秋红，等. 针对光照干扰的 TSA-YOLO 车辆检测算法[J]. 计算机工程与设计，2025，46(7)：2021-2027.

[9] 霍华，邢晓钰. 复杂场景下的轻量化行人检测算法研究[J/OL]. 计算机应用与软件(2025-07-04)[2025-08-02]. https：//link. cnki. net/urlid/31. 1260. tp. 20250703. 1757. 002.

[10] 孙海明，付世超. 基于改进 YOLOv4-Tiny 的交通标志图像识别算法研究[J]. 计算机应用与软件，2025，42(5)：164-170.

[11] SONG G，LIU Y，WANG X. Revisiting the sibling head in object detector[C]//2020 IEEE/CVF conference on computer vision and pattern recognition. Seattle：IEEE，2020：11560-11569.

[12] ZHANG X，LIU C，YANG D，et al. RFAConv：innovating spatial attention and standard convolutional operation[PP/OL]. V4.（2023-04-18）[2025-07-23]. https：//arxiv. org/abs/2304. 03198.

[13] HE K，ZHANG X，REN S，et al. Deep residual learning for image recognition[C]. 20116 IEEE conference on computer vision and pattern recognition. Las Vegas：IEEE，2016：770-778.

[14] ZHANG H，XU C，ZHANG S. Inner-IoU：more effective intersection over union loss with auxiliary bounding box[PP/OL]. V4.（2023-11-14）[2025-07-23]. https：//arxiv. org/abs/2311. 02877.

[15] YU F，CHEN H F，WANG X，et al. BDD100K：a diverse driving dataset for heterogeneous multitask learning[C]//2020 IEEE/CVF Conference on computer vision and pattern recognition，Seattle：IEEE，2020：2633-2642.

Object detection algorithm for complex road scenes based on improved YOLOv8

DONG Feng[1]，ZHANG Xin[1]，KONG Lin[2]，ZHANG Tao[1]

(1. *School of Science and Technology*，*Xinyang University*，*Xinyang* 464000，*China*；
2. *Zhejiang Xizhong New Energy Automobile Technology Co.*，*Ltd.*，*Jinhua* 321000，*China*)

Abstract：To address the issues of missed and false detection of small objects in complex road scenes，this paper proposes a complex road scene object detection algorithm based on YOLOv8. Firstly，a RFCAConv module is introduced into the feature extraction network to enhance the model's ability to capture key feature information；secondly，to enrich the contextual information in deeper networks，a Context-Aware Module (CAM) is constructed in the aggregation network；finally，the original CIoU loss function is replaced with Inner-CIoU to improve the model's generalization ability. Experimental results show that on the New_BDD public datasets，the proposed algorithm achieves a detection precision (*mAP*) of 77.6% and an inference speed (*FPS*) of 98 frames per second，outperforming other mainstream algorithms. Additionally，in the visualization experiments of image detection results，the proposed model effectively reduces missed and false detection in various complex road scene image instances compared to the original model.

Keywords：object detection algorithm；small object；YOLOv8；RFCAConv；context awareness；Inner-CIoU

(责任编辑：王新亮)

引用格式　董凤，张鑫，孔林，等. 基于改进 YOLOv8 的复杂道路场景目标检测算法[J]. 山东航空学院学报，2026，43(2)：73-81.

DONG F，ZHANG X，KONG L，et al. Object detection algorithm for complex road scenes based on improved YOLOv8[J]. Journal of Shandong University of Aeronautics，2026，43(2)：73-81.