

基于改进 YOLOv8 的乡村无信号灯交叉口多目标检测方法

邹佳鑫¹, 李冰彦², 鄢然¹, 邓鑫楠¹, 罗建¹, 廖杨^{1,*}

1. 西华大学, 汽车与交通学院, 四川 宜宾, 644000

2. 西华大学, 文学与新闻传播学院, 四川 宜宾, 644000

摘要: 针对乡村无信号灯交叉口交通参与者类型复杂、目标尺度差异明显以及遮挡和背景干扰较强等问题, 本文提出一种改进 YOLOv8n 的轻量化多目标检测方法。该方法在保持 YOLOv8n 单阶段检测框架的基础上, 引入 EfficientViT2、DySample、EMA 注意力机制和 MPDIoU 回归损失, 分别从特征提取、多尺度融合、关键区域响应和边界框定位等方面提升模型对复杂乡村交通场景的适应能力。基于自建乡村无信号灯交叉口数据集, 本文开展模型对比实验和消融分析。结果表明, 最终组合模型在 Precision、Recall、mAP@0.5、mAP@0.5:0.95 和 F1-score 上均优于基线模型, 其中 mAP@0.5 和 mAP@0.5:0.95 分别达到 0.771 和 0.532, 较 YOLOv8n 分别提升 0.104 和 0.113。研究表明, 所提方法能够有效提升复杂乡村无信号灯交叉口场景下的多目标检测性能, 可为乡村道路交通安全预警提供视觉感知支持。

关键词: 交通安全; 乡村无信号灯交叉口多目标检测; 改进 YOLOv8; 轻量化目标检测

Multi-Object Detection at Rural Unsignalized Intersections Based on an Improved YOLOv8

Jiaxin Zou¹, Bingyan Li², Ran Yan¹, Xinnan Deng¹, Jian Luo¹, Yang Liao^{1,*}

1. School of Automotive and Transportation Engineering, Xihua University, Yibin, Sichuan 644000, China

2. School of Literature and Journalism, Xihua University, Yibin, Sichuan 644000, China

Abstract: To address the challenges of diverse traffic participants, large target-scale variations, occlusion, and complex background interference at rural unsignalized intersections, this study proposes a lightweight multi-object detection method based on an improved YOLOv8n. Without changing the original one-stage detection framework, EfficientViT2, DySample, EMA attention, and MPDIoU regression loss are introduced to enhance feature extraction, multi-scale fusion, key-region response, and bounding-box localization. Comparative experiments and ablation analyses are conducted on a self-constructed rural unsignalized intersection dataset. The results show that the final combined model outperforms the baseline YOLOv8n in Precision, Recall, mAP@0.5, mAP@0.5:0.95, and F1-score. Specifically, mAP@0.5 and mAP@0.5:0.95 reach 0.771 and 0.532, respectively, improving by 0.104 and 0.113 over YOLOv8n. The results indicate that the proposed method can effectively improve multi-object detection performance in complex rural intersection scenarios and provide visual perception support for rural traffic safety warning.

Keywords: Traffic safety; Multi-object detection at rural unsignalized intersections; Improved YOLOv8; Lightweight object detection

随着道路网络向城乡地区延伸, 乡村无信号灯交叉口安全风险日益突出。该类交叉口基础设施和交通组织相对薄弱, 缺乏统一信号控制与明确通行优先规则, 车辆、行人及其他交通主体主

要依赖经验判断和相互避让通行；在视距受限、交通流量波动或驾驶行为不规范条件下，交通冲突随机性较强，事故风险较高^[1,2]。

与城市道路相比，乡村无信号灯交叉口具有更明显的混合交通特征。场景中除机动车和行人外，还可能出现牲畜、农用车辆等非标准交通主体，目标外观、运动方式和尺度分布差异较大；同时，植被、建筑物遮挡以及逆光、低照度等因素会削弱目标纹理和边界信息，使小尺度、弱显著和遮挡目标更难稳定检出，对检测模型的特征表达、多尺度感知和边界定位能力提出更高要求^[2,3,4]。

近年来，基于深度学习的目标检测方法已广泛应用于交通监控和辅助驾驶等场景。YOLO 系列单阶段检测模型兼具检测速度和部署效率，YOLOv8 进一步优化了网络结构与训练策略，具备较好的实时检测基础^[5,6,7]。但现有模型多面向城市道路或标准交通环境，直接用于乡村无信号灯交叉口时，仍存在复杂背景下特征表达不足、尺度变化中细节损失和边界框回归稳定性不够等问题^[7,8,9,11]。

为此，本文以 YOLOv8n 为基线，面向乡村无信号灯交叉口多目标检测任务进行改进：引入 EfficientViT2 增强骨干网络语义表征能力，采用 DySample 改善多尺度特征融合中的细节保留，嵌入 EMA 注意力机制强化关键区域响应，并采用 MPDIoU 提升边界框定位稳定性^[7,9,10,11]。最后基于自建数据集开展对比实验与消融分析，验证所提方法的有效性。

1 面向乡村无信号灯交叉口的多目标检测算法

1.1 基线模型与总体思路

本文以 YOLOv8n 作为基线模型开展改进。YOLOv8n 属于轻量化单阶段目标检测模型，在保持较高推理效率的同时具备一定检测精度，适合乡村道路安全预警等资源受限场景^[7]。结合乡村无信号灯交叉口交通主体混杂、目标尺度差异明显、遮挡和背景干扰较强等特点，本文在不改变 YOLOv8n 原有检测框架和输出形式的基础上，从特征提取、特征融合、检测头和边界框回归损失四个方面进行改进。

改进模型总体结构如图 1 所示。输入图像首先经 Backbone 提取多层次特征，随后在 Neck 中完成多尺度特征融合，最终由 Head 输出目标类别、置信度和边界框位置。本文在 Backbone 中引入 EfficientViT2 结构^[10]，在 Neck 上采样位置采用 DySample 动态上采样^[9]，在检测头多尺度分支中嵌入 EMA 注意力机制^[11]，并在训练阶段采用 MPDIoU 回归损失^[7,12]，以提升复杂乡村交通场景下小尺度目标、弱显著目标和遮挡目标的检测稳定性。

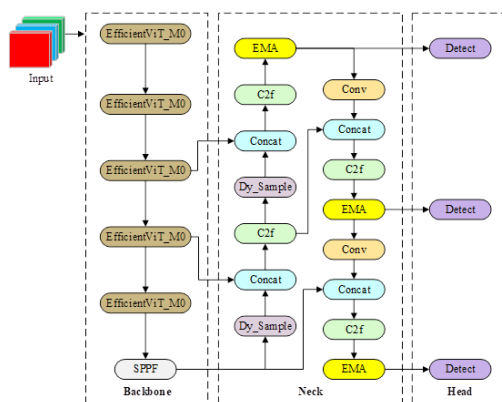


图 1 YOLOv8 基线模型的总体结构示意图

Fig. 1 Schematic diagram of the overall structure of YOLOv8 baseline model

1.2 改进网络结构设计

乡村无信号灯交叉口中，植被、建筑物和光照变化等因素容易削弱目标边界与纹理信息，导致

小尺度目标、弱显著目标和遮挡目标检出率下降。针对复杂背景下特征表达不足的问题,本文在骨干网络中引入 EfficientViT2 结构,替换部分特征提取单元^[10]。该结构通过多尺度线性注意力增强全局上下文建模能力,在保持轻量化特征提取的同时加强局部细节与全局语义信息的协同表达,从而提高复杂背景条件下的特征判别能力。

在多尺度特征融合阶段,YOLO 类模型通常依赖特征金字塔结构完成不同尺度特征的信息传递^[13]。但固定插值上采样在目标尺度变化较大或边界较弱时,容易造成局部细节损失。为此,本文在 Neck 的上采样位置采用 DySample 动态上采样替代固定插值操作^[9],使采样位置能够根据局部结构自适应调整,从而改善尺度变换过程中的细节保留能力,增强小尺度目标和弱边界目标的特征表达。

在检测头阶段,融合特征中同时包含有效目标信息和背景噪声,若缺乏特征筛选机制,容易影响分类与边界框回归结果。本文在检测头多尺度分支中嵌入 EMA 注意力模块^[11],通过跨空间多尺度注意力增强关键区域和关键尺度响应,并抑制复杂背景中的冗余干扰。总体而言,EfficientViT2、DySample 和 EMA 分别对应复杂背景下语义表达不足、尺度变化下细节损失以及多目标混杂下关键特征响应不充分等问题,三者共同提升模型在乡村无信号灯交叉口场景中的多尺度感知能力和检测稳定性。

1.3 MPDIoU 回归损失与训练目标

乡村无信号灯交叉口中,目标尺度差异较大,且遮挡、弱光和背景干扰容易造成边界信息不稳定。为提升模型在边界模糊和尺度变化条件下的回归性能,本文采用 MPDIoU 作为边界框回归损失^[7,12]。模型训练目标可表示为:

$$L = \lambda_{box} L_{MPDIoU} + \lambda_{cls} L_{cls} + \lambda_{dfl} L_{DFL}. \quad (1)$$

其中, λ_{box} 、 λ_{cls} 和 λ_{dfl} 为各损失项权重, L_{MPDIoU} 为边界框回归损失, L_{cls} 为分类损失, L_{DFL} 为分布式回归损失。

MPDIoU 在 IoU 项基础上进一步引入预测框与真实框角点之间的几何距离约束,可表示为:

$$L_{MPDIoU} = 1 - IoU(B, B^{gt}) + \frac{d_1^2 + d_2^2}{w^2 + h^2}. \quad (2)$$

$L_{MPDIoU} = 1 - IoU + \frac{d_1^2}{C} + \frac{d_2^2}{C}$ 式中, d_1 和 d_2 分别表示预测框与真实框对应左上角点、右下角点之间的欧氏距离, C 为尺度归一化项。该损失能够同时约束重叠区域和角点位置,有助于提升尺度差异较大或边界模糊目标的定位稳定性。

2 实验设计与结果分析

2.1 实验设置与数据集说明

为验证改进模型在乡村无信号灯交叉口复杂交通场景中的检测效果,本文在统一数据集、统一训练流程和统一评估标准下开展实验。实验数据来源于实际采集的乡村无信号灯交叉口固定视角视频,经抽帧和人工标注后构建图像数据集。数据集中包含行人、机动车、牲畜等交通参与者,覆盖遮挡、逆光、背景杂波和目标尺度差异等典型情况,能够反映乡村无信号灯交叉口混合交通场景的实际特征。

本文数据集共包含 2614 张图像,其中训练集 2091 张,验证集 523 张。各组实验采用相同的数据划分和评估流程,以减少数据划分差异对模型性能比较的影响。实验以 YOLOv8n 为基线模型^[7],并分别构建数据增强模型、EfficientViT2 改进模型^[10]、DySample 改进模型^[9]、EMA 注意力改进模型

[11]、MPDIoU 回归损失改进模型^[7,12]、SIoU 损失函数对比模型^[14]以及最终组合模型。

图 2 展示了自建数据集中的部分样例图像。由图可见，乡村无信号灯交叉口场景中存在多类交通参与者混行、目标尺度差异明显和背景干扰较强等问题，可为复杂乡村交通环境下的多目标检测模型验证提供数据基础。

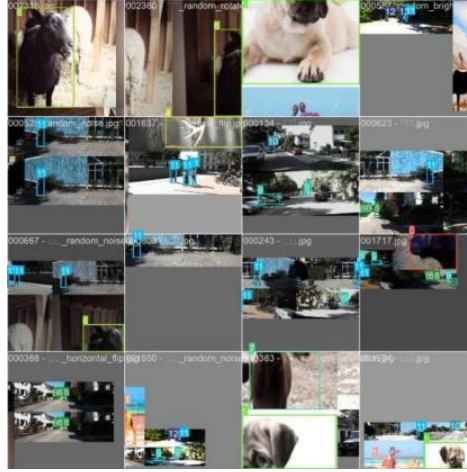


图 2 自建数据集部分样例图像

Fig. 2 Partial sample images of self-constructed dataset

2.2 评价指标

为从误检、漏检、综合精度和定位质量等方面评价模型性能，本文采用 Precision、Recall、F1-score、mAP@0.5 和 mAP@0.5:0.95 作为评价指标。其中，Precision 用于衡量预测结果中正确检测目标所占比例，Recall 用于衡量真实目标被成功检出的比例，F1-score 用于综合反映 Precision 与 Recall 的平衡关系。设 TP、FP 和 FN 分别表示真阳性、假阳性和假阴性，则：

$$Precision = \frac{TP}{TP+FP}, Recall = \frac{TP}{TP+FN}. \quad (3)$$

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}. \quad (4)$$

$$AP_c = \int_0^1 P_c(R) dR. \quad (5)$$

mAP 用于衡量模型在各类别上的平均检测精度，可表示为：

$$AP_c = \int_0^1 P_c(R) dR. \quad (6)$$

式中， C 为类别数， AP_c 为第 c 类目标的平均精度。mAP@0.5 表示 IoU 阈值为 0.5 时的平均精度均值，mAP@0.5:0.95 表示在多个 IoU 阈值下计算得到的平均精度均值，对边界框定位质量要求更高。

2.3 模型对比实验

表 1 列出了不同模型在乡村无信号灯交叉口数据集上的检测结果。对比模型包括 YOLOv8n 基线模型、数据增强模型、单项结构改进模型、回归损失对比模型以及最终组合模型。通过在相同数据集和统一评估流程下进行横向比较，可分析各项改进策略对模型检测性能的影响。

表 1 不同模型在乡村无信号灯交叉口数据集上的检测性能对比

Table 1 Performance comparison of different models

模型	P(精确率)	R(召回率)	mAP0.5	mAP0.5:0.95	F1-score
Yolov8n	0.676	0.626	0.667	0.419	0.64
dataset_aug(数据增强)	0.759	0.625	0.695	0.455	0.68
Yolov8n+ EfficientViT2	0.754	0.698	0.737	0.489	0.72
Yolov8n+DySample	0.769	0.634	0.701	0.455	0.69
Yolov8n+EMA 注意力	0.71	0.678	0.711	0.459	0.69
Yolov8n+MPDIoU	0.752	0.627	0.698	0.459	0.68
Yolov8n+SIOU	0.719	0.582	0.643	0.359	0.64
yolov8-EfficientViT2-DySample-EMA-MPDIoU	0.816	0.682	0.771	0.532	0.767

由表 1 可知, 基线 YOLOv8n 的 Precision、Recall、mAP@0.5、mAP@0.5:0.95 和 F1-score 分别为 0.676、0.626、0.667、0.419 和 0.640。最终组合模型对应指标分别达到 0.816、0.682、0.771、0.532 和 0.767, 均为各组模型中的最优结果。与基线模型相比, 最终组合模型的 Precision、Recall、mAP@0.5、mAP@0.5:0.95 和 F1-score 分别提升 0.140、0.056、0.104、0.113 和 0.127, 说明所提多策略联合改进能够同时改善目标检出能力、定位质量和综合检测性能。

为更直观地展示不同模型在 mAP 指标上的变化趋势, 本文根据表 1 中的 mAP@0.5 和 mAP@0.5:0.95 结果绘制图 3。由图 3 可知, EfficientViT2 单项改进对 mAP 提升较明显, 表明增强骨干网络特征表达有助于复杂背景下的目标识别; SIOU 在本文数据集上表现不稳定, mAP 指标低于基线模型。最终组合模型在 mAP@0.5 和 mAP@0.5:0.95 上均取得最高值, 说明 EfficientViT2、DySample、EMA 与 MPDIoU 的联合改进具有更好的综合效果。

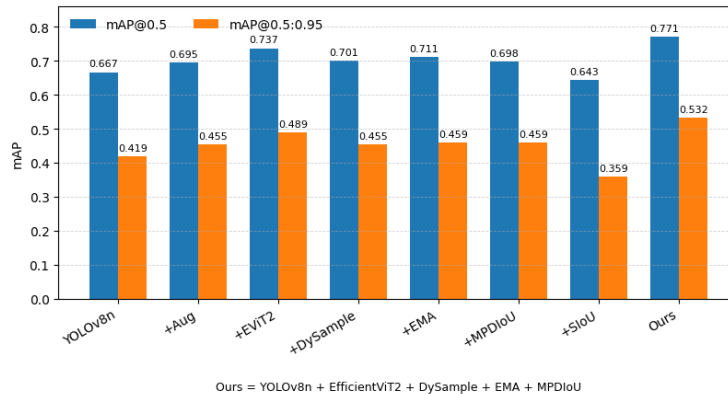


图 3 各模块对 mAP 值的提升趋势曲线图

Fig. 3 mAP performance changes of the model before and after introducing different improvement modules

2.4 消融实验分析

为分析各改进策略的独立贡献, 本文结合表 1 对数据增强、EfficientViT2、DySample、EMA 和 MPDIoU 进行消融分析。仅引入数据增强后, 模型 mAP@0.5 由 0.667 提升至 0.695, mAP@0.5:0.95 由 0.419 提升至 0.455, 说明数据增强能够改善模型对复杂背景和外观变化的适应能力。引入 EfficientViT2 后, 模型 mAP@0.5 和 mAP@0.5:0.95 分别提升至 0.737 和 0.489, 单项提升最为明显, 说明骨干网络语义表征能力对复杂乡村交通场景下的目标检测具有重要影响^[10]。DySample、EMA 和 MPDIoU 也分别从尺度特征融合、关键区域响应和边界框定位等方面带来一定增益^[7,9,11,12]。

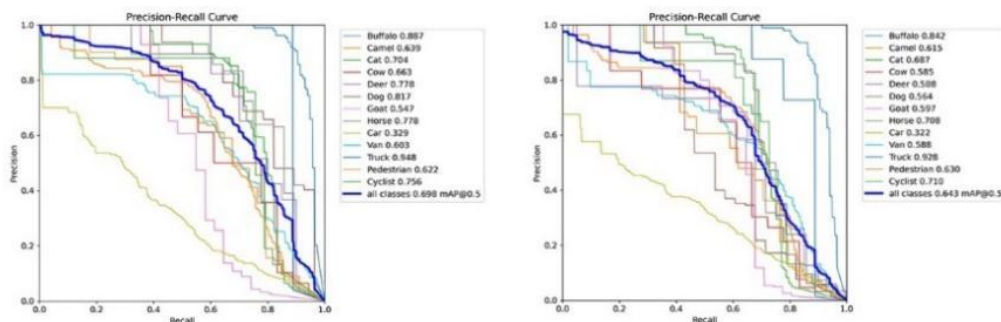


图 4 Siou 与 MPDIoU 回归损失的性能对比

Fig. 4 Training performance changes of the model under different regression loss functions

3 讨论

实验结果表明,针对乡村无信号灯交叉口复杂交通场景,仅依靠 YOLOv8n 基线模型难以充分应对小尺度目标、遮挡目标、复杂背景和边界信息不稳定等问题。由表 1、图 3 和图 4 可知,最终组合模型整体优于基线模型和单项改进模型,说明特征提取增强、动态特征融合、注意力筛选和回归损失优化能够形成协同作用。其中, EfficientViT2、DySample 和 EMA 分别对应复杂背景下的语义表达、多尺度细节保留和关键区域响应, MPDIoU 则进一步改善边界框定位稳定性^[7,9,10,11,12]。这些改进与乡村交叉口交通主体混杂、目标尺度差异明显和背景干扰较强的场景特征相匹配,因此能够提升多目标检测性能。

在回归损失方面,图 4 显示 MPDIoU 在 mAP@0.5 和 mAP@0.5:0.95 上均优于 Siou,说明在本文数据集条件下,引入角点距离约束的回归方式更有利于提升边界框定位稳定性^[7,12,14]。这与乡村交通目标边界易受遮挡、弱光和背景杂波影响的特点有关。对于固定视角下的交叉口监测任务,定位质量不仅影响检测评价指标,也会进一步影响后续交通冲突识别、目标跟踪和安全预警效果,因此 MPDIoU 的引入具有一定工程意义。但本文数据主要来自固定视角场景,对夜间、雨雾、多季节和边缘设备部署性能的验证仍需进一步补充。

4 结论

(1) 本文面向乡村无信号灯交叉口混合交通场景,提出了一种改进 YOLOv8n 的多目标检测方法。该方法在骨干网络、特征融合、检测头和回归损失四个方面进行优化,引入 EfficientViT2、DySample、EMA 和 MPDIoU,以增强复杂背景下的特征表达、多尺度融合和边界框定位能力;

(2) 基于自建乡村无信号灯交叉口数据集的实验结果表明,最终组合模型取得最优检测效果。与基线 YOLOv8n 相比, mAP@0.5 和 mAP@0.5:0.95 分别提升 0.104 和 0.113, Precision、Recall 和 F1-score 也同步提高,说明所提方法能够有效改善复杂乡村交通场景下的多目标检测性能;

(3) 消融实验表明,各改进模块具有不同作用侧重: EfficientViT2 主要增强复杂背景下的语义表征能力, DySample 改善尺度变换中的细节保留, EMA 强化关键区域响应, MPDIoU 提升边界框定位稳定性。后续研究仍需进一步扩展夜间、雨雾和多季节数据,并结合边缘设备部署与交通安全预警任务开展验证。

参考文献

- [1] 张名芳, 马艳华, 马 勇. 无信号交叉口人车冲突严重程度影响因素分析[J]. 中国安全科学学报, 2023, 33(8): 190.
- [2] 赵 丹, 马社强, 张雨萌, 等. 农村公路交叉口交通事故特征关联性与风险因素分析[J]. 中国安全科学学报, 2020, 30(07): 146-151.
- [3] 华 夏, 王新晴, 王 东, 等. 基于改进 SSD 的交通大场景多目标检测[J]. Acta Optica Sinica, 2018, 38(12): 1215003.

- [4] 曲立国, 张 鑫, 卢自宝, 等. 基于改进 YOLOv5 的交通标志识别方法[J]. 光电工程, 2025, 51(6): 240055-1-240055-13.
- [5] 岑伟迪, 江佳玲, 黄 勃, 等. YOLO 十年演进: 从实时检测先锋到多任务智能前沿[J/OL]. 武汉大学学报(理学版), 1-16[2025-10-17].
- [6] 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.
- [7] Ma S, Xu Y. Mpdious: a loss for efficient and accurate bounding box regression[J]. arXiv preprint arXiv:2307.07662, 2023.
- [8] 马 鸽, 李洪伟, 严梓维, 等. 基于多注意力的改进 YOLOv5s 小目标检测算法[J]. 工程科学学报, 2024, 46(9): 1647-1658.
- [9] 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.
- [10] Cai H, Li J, Hu M, et al. Efficientvit: Multi-scale linear attention for high-resolution dense prediction[J]. arXiv preprint arXiv:2205.14756, 2022.
- [11] Ouyang D, He S, Zhang G, et al. Efficient multi-scale attention module with cross-spatial learning[C]//ICASSP 2023-2023 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2023: 1-5.
- [12] Li X, Wang W, Wu L, et al. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection[J]. Advances in neural information processing systems, 2020, 33: 21002-21012.
- [13] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.
- [14] Gevorgyan Z. SIOU loss: More powerful learning for bounding box regression[J]. arXiv preprint arXiv:2205.12740, 2022.

基金项目: 2025 年省级大学生创新创业训练计划资助项目“基于两段式红外-摄像头的乡村交叉口智联预警装置”(项目编号: S202510623132)。

¹**第 1 作者简介:** 邹佳鑫(2004-), 女, 本科在读, 研究方向: 交通运输。E-mail: 2213246403@qq.com。

***通讯作者简介:** 廖杨(1999-), 女, 硕士研究生, 助理实验师, 研究方向: 交通行为。E-mail: liaoyang@mail.xhu.edu.cn。