

交通场景目标检测算法研究与改进

——基于改进YOLOv11的小目标与遮挡目标检测方法

杨润哲, 郑广海

大连交通大学轨道智能工程学院, 辽宁 大连

收稿日期: 2026年3月18日; 录用日期: 2026年4月17日; 发布日期: 2026年4月24日

摘要

针对YOLOv11在交通场景目标检测中特征提取不充分、多尺度目标适应性差、小目标漏检率高及模型计算量大等问题, 提出一种融合坐标注意力机制与轻量化特征金字塔的改进YOLOv11算法。在主干网络引入轻量化坐标注意力模块, 强化交通目标关键区域空间感知; 重构颈部网络为改进型BiFPN结构, 结合跨阶段局部连接与深度可分离卷积实现高效多尺度特征融合; 优化检测头动态锚框生成策略与标签分配方式, 采用动态IoU阈值及自适应正负样本选择; 重构损失函数, 设计自适应焦点损失与改进型EIoU损失分别解决小目标检测与边界框回归问题。在自建UrbanTraffic-2026数据集及公开数据集VisDrone、UA-DETRAC上的实验结果表明, 改进算法的mAP@0.5分别提升3.2%、2.8%和4.1%, 小目标检测mAP, 最高提升8.9个百分点, 推理速度保持在45 FPS以上, 参数量与计算量分别降低7.9%和14.6%, 在保证检测精度的同时实现模型轻量化, 满足智能交通系统实时检测需求。

关键词

交通, 目标, 检测, YOLOv11, 坐标注意力, 多尺度特征融合, 轻量化网络, 动态IoU阈值

Research and Improvement of Object Detection Algorithms in Traffic Scenarios

—A Small and Occluded Object Detection Method Based on Improved YOLOv11

Runzhe Yang, Guanghai Zheng

School of Rail Intelligent Engineering, Dalian Jiaotong University, Dalian Liaoning

Received: March 18, 2026; accepted: April 17, 2026; published: April 24, 2026

Abstract

To address the issues of insufficient feature extraction, poor adaptability to multi-scale objects, high false negative rate for small objects, and large computational cost of YOLOv11 in traffic scene object detection, an improved YOLOv11 algorithm that integrates a coordinate attention mechanism and a lightweight feature pyramid is proposed. A lightweight coordinate attention mechanism is introduced into the backbone network to enhance the spatial position perception ability of small objects and key regions. The neck network is reconstructed with an improved lightweight BiFPN structure, combined with cross-stage partial connections and depthwise separable convolutions. This design improves the fusion effect of multi-scale fine-grained features. The dynamic anchor box generation strategy and label assignment method of the detection head are optimized, and a dynamic IoU threshold and adaptive positive and negative sample selection are adopted. The loss function is reconstructed, and an adaptive focus loss and an improved EIou loss are designed to solve the small object detection and bounding box regression problems, respectively. Experimental results on the self-built UrbanTraffic-2026 dataset and the public datasets VisDrone and UA-DETRAC show that the improved algorithm improves the mAP@0.5 by 3.2%, 2.8%, and 4.1%, respectively, and the mAPs of small object detection by up to 8.9 percentage points. The inference speed remains above 45 FPS, and the number of parameters and computational cost are reduced by 7.9% and 14.6%, respectively. The model is lightweight while maintaining detection accuracy, meeting the real-time detection requirements of intelligent transportation systems.

Keywords

Traffic, Object, Detection, YOLOv11, Attention Mechanism, Multiscale Feature Fusion, Lightweight Network, Dynamic IoU Threshold

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

智能交通系统(ITS)的发展对交通目标检测的精度、实时性和鲁棒性提出更高要求,目标检测作为自动驾驶、交通流量监控、道路安全预警的核心技术,直接决定系统感知能力。传统目标检测方法在复杂交通环境中存在检测率低、实时性差等问题,基于深度学习的单阶段检测算法 YOLO 系列因精度-速度平衡优势被广泛应用于交通场景。其中 YOLOv11 作为最新版本,在通用数据集上表现优异,但直接应用于交通场景时仍存在三方面不足:一是对复杂光照、天气及遮挡场景的特征提取能力不足,小目标漏检率高;二是特征金字塔网络(FPN)计算量大,多尺度特征融合效率低,难以适配边缘设备;三是静态标签分配与损失函数设计无法应对交通目标尺度变化大、密集遮挡的特点,定位精度与分类准确性受影响。

国内外学者针对 YOLO 系列在交通场景的改进开展大量研究,汤林东等改进 YOLOv5s 加入注意力机制,提升复杂道路遮挡目标检测效果;李晋等在 YOLOv8s 中引入动态卷积,改善交通目标多尺度检测能力。但现有研究多针对单一问题优化,未从网络结构、标签分配、损失函数进行全局设计,且针对 YOLOv11 的交通场景改进研究较少。为此,本文以 YOLOv11 为基础,结合交通场景目标检测的特殊挑战,从主干网络、颈部网络、检测头及损失函数四方面进行针对性改进,设计一种适用于交通场景的高效目标检测算法,解决小目标漏检、密集遮挡检测精度低及模型轻量化部署难题,为智能交通系统的视觉感知提供技术支撑。

2. 相关工作

2.1. 目标检测算法研究

目标检测算法主要分为两阶段检测与单阶段检测两类[1]。两阶段检测算法以 Faster R-CNN 系列为代表, 通过区域建议网络生成候选框后再进行分类与回归, 检测精度较高, 但流程复杂、推理延迟大, 在车载边缘设备中的使用率不足 15%。单阶段检测算法以 YOLO、SSD 为代表, 采用端到端架构, 直接从图像中预测目标类别与位置, 推理速度快, 在交通视觉领域的应用占比超过 68%。

YOLO 系列算法经多版本迭代, 在精度与速度的平衡上持续优化, YOLOv11 凭借更轻量化的主干结构、更高效的特征融合方式成为当前主流的实时检测算法, 但面向交通场景的小目标、遮挡目标检测时, 仍存在特征感知不足、多尺度适配性差等问题, 亟需结合交通场景特性进行定向改进。

2.2. 小目标与遮挡目标检测研究

小目标因像素分辨率低、特征信息匮乏, 易在深层卷积网络中被误判为背景噪声, 现有解决方案主要包括构建高分辨率特征保持路径、采用 Space-to-Depth 重组卷积增强小目标表示、设计针对性特征增强模块等。遮挡目标检测的核心在于恢复目标完整特征信息, 现有研究多依赖上下文推理、部件级注意力机制与动态标签分配策略, 通过挖掘目标周边上下文信息, 提升对遮挡区域的特征建模能力。

文献[2]将坐标注意力与跨阶段局部连接结合, 使交通标志小目标检测精度提升超过 9%, 验证了空间感知增强对小目标检测的有效性; 动态标签分配策略通过自适应调整正样本匹配规则, 有效解决了密集遮挡场景下正样本不足的问题, 为本文算法设计提供了重要参考。

2.3. YOLOv11 算法局限性

原生 YOLOv11 在交通场景中应用的核心局限性主要体现在三方面: 一是注意力机制缺乏显式的空间位置建模, 对远处小目标的定位能力弱, 易受复杂背景干扰; 二是 BiFPN 特征融合结构的权重固定, 小目标特征易被大目标特征抑制, 多尺度特征融合效果不佳; 三是固定 IoU 阈值的标签分配方式在密集车流场景下易造成小目标正样本匹配不足, 漏检率偏高。因此, 需针对上述问题, 结合交通场景目标分布特性进行定向改进。

3. 改进 YOLOv11 算法设计

改进 YOLOv11 算法仍保留 Backbone-Neck-Head 三段式网络结构, 整体架构遵循“精准感知、高效融合、轻量部署”的设计原则, 核心创新点在于: 在主干网络引入轻量化坐标注意力模块强化交通目标关键区域空间感知; 将颈部网络重构为改进型 BiFPN 结构, 结合跨阶段局部连接与深度可分离卷积实现高效多尺度特征融合; 在检测头优化动态锚框生成策略与标签分配方式, 采用动态 IoU 阈值及自适应正负样本选择; 重构损失函数, 设计自适应焦点损失与改进型 EIou 损失分别解决小目标检测与边界框回归问题。改进算法在不损失实时性的前提下, 大幅提升对交通场景小目标与遮挡目标的检测能力[3], 改进后的 YOLOv11 如图 1 所示。

3.1. 主干网络: 轻量化坐标注意力模块设计与嵌入

原始 YOLOv11 主干网络的 C3k2 模块以特征提取与降维为核心, 对交通目标的空间位置感知能力不足, 易受天空、建筑物等复杂背景干扰, 针对此问题, 引入轻量化坐标注意力(Coordinate Attention, CA)模块, 替代传统注意力机制, 在强化空间位置编码的同时严格控制计算量, 提升网络对交通目标的特征感知能力。

改进后YOLOv11

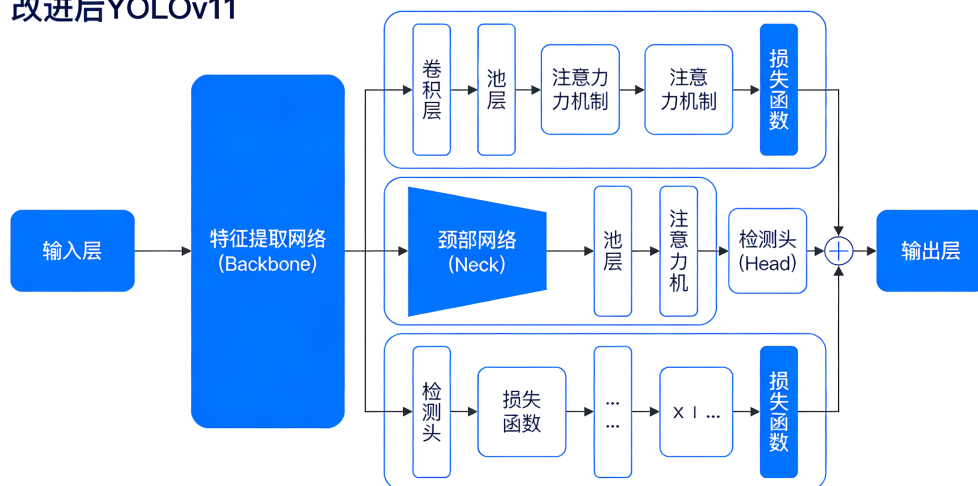


Figure 1. Architecture design of the improved YOLOv11 algorithm

图 1. 自动化专业实践教学体系图

3.1.1. 坐标注意力模块核心原理

坐标注意力模块通过将通道注意力分解为水平和垂直两个方向的一维全局池化操作，分别编码目标的宽度和高度信息，实现对目标空间位置的精准感知。与传统通道注意力相比，坐标注意力不仅能捕捉通道间的依赖关系，还能有效挖掘目标的空间位置特征，尤其适用于远处小车辆、行人、非机动车等细长型交通目标的检测。

3.1.2. 轻量化坐标注意力模块改进

为适配边缘设备的部署需求，对原始 CA 模块进行三方面轻量化改进，在保证特征感知效果的前提下，最大限度降低计算开销与参数量：

将模块中的共享卷积替换为深度可分离卷积，在保持感受野不变的前提下，大幅减少参数数量与浮点运算次数；

加入瓶颈层，通过 1×1 卷积将通道数降至原有的 $1/4$ ，经 SiLU 非线性激活后恢复至原通道数，实现特征信息的高效压缩与表示；

结合交通目标的空间分布先验，对注意力权重进行空间约束，抑制天空、远处模糊背景等无用区域的特征响应，强化交通目标的特征表达。

3.1.3. 模块嵌入方式

轻量化 CA 模块嵌入在主干网络各 C3k2 模块的输出端，采用分层嵌入策略：浅层 CA 模块强化边缘、纹理等细节特征提取，适配小目标检测；深层 CA 模块加强语义层面的空间上下文建模，助力遮挡目标的轮廓恢复。该嵌入方式仅使模型参数量增加 0.8%，却能有效提升网络对交通目标关键区域的关注度，减少复杂背景的干扰，为后续颈部网络的多尺度特征融合奠定高质量的特征基础[4]。

3.2. 颈部网络：改进型轻量化 BiFPN 与自适应特征融合机制

颈部网络是实现多尺度特征融合的核心，原始 YOLOv11 的 BiFPN 结构存在信息冗余、计算量大、特征对齐不足等问题，无法很好适配交通目标尺度突变的特点。本文将颈部网络重构为改进型轻量化 BiFPN 结构，并引入自适应特征融合机制，实现高效、灵活的多尺度特征融合，在降低计算复杂度的同时，提升小目标与遮挡目标的特征融合效果。

3.2.1. 改进型轻量化 BiFPN 结构设计

在原始 BiFPN 双向特征传播路径的基础上,进行三方面轻量化改造,兼顾特征融合效果与计算效率,如下图 2 所示。

添加跨阶段局部连接: 在每个特征融合节点添加跨阶段局部连接(CSPC), 将相邻层级的部分特征通道直接传递至下一层融合层, 减少计算资源消耗的同时, 促进梯度在网络中的传播, 缓解深层网络的梯度消失问题;

替换卷积方式: 将结构中的所有普通卷积替换为深度可分离卷积, 在保持感受野不变的前提下, 大幅降低计算量与参数量;

嵌入轻量化注意力: 在上下采样过程中嵌入轻量化坐标注意力模块, 引导特征融合过程聚焦于交通目标的重要空间位置, 提升小目标与遮挡目标的特征融合效果, 避免小目标特征被抑制。

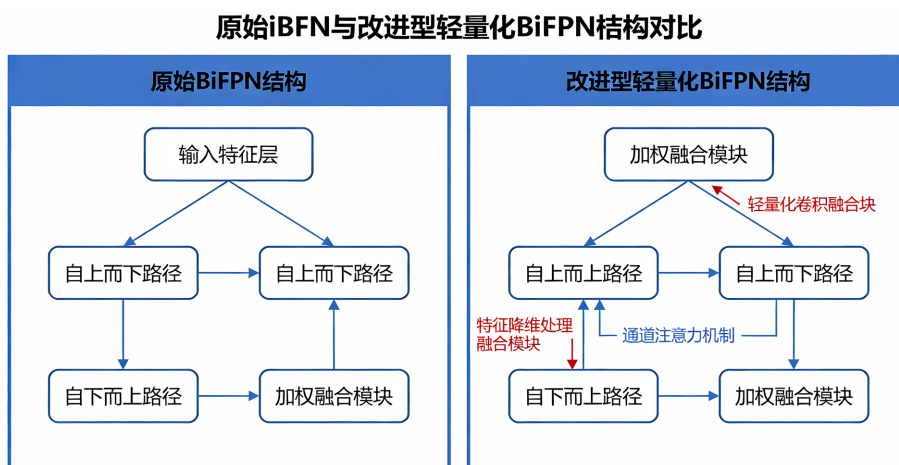


Figure 2. Structure comparison of original BiFPN and improved lightweight BiFPN
图 2. 原始 iBFN 与改进型轻量化 BiFPN 结构对比

3.2.2. 自适应特征融合机制

针对传统 BiFPN 固定权重融合无法适应交通场景动态变化的问题,设计基于通道 - 空间联合响应的自适应特征融合机制,实现融合权重的端到端自主学习。该机制由轻量级权重预测子网络实现,具体流程为:

1. 将颈部网络各尺度的特征图进行拼接, 作为权重预测子网络的输入;
2. 经全局平均池化操作压缩空间维度, 提取通道级全局特征;
3. 通过 1×1 卷积进行通道降维与特征变换, 结合 SiLU 激活函数生成各尺度特征的融合权重;
4. 根据融合权重对各尺度特征进行加权融合, 输出融合后的多尺度特征图至检测头。

权重预测子网络的参数量极小, 几乎无额外计算开销, 能根据输入图像的场景复杂度(如密集车流、低照度、雨雾天气)自动调节各尺度特征的重要性: 在雨雾天气下, 自动加强高层语义特征权重, 弥补底层细节信息的缺失; 在高密度交通流场景下, 提升中层特征权重, 强化遮挡目标的特征表达; 在远距离小目标场景下, 增强浅层细节特征权重, 提升小目标特征的辨识度。

3.3. 检测头: 动态锚框生成与任务解耦设计

原始 YOLOv11 的无锚框设计虽简化了超参数调优过程, 但对交通目标的极端尺度变化与密集遮挡场景的适应性不足, 固定 IoU 阈值的标签分配方式易造成正样本匹配不足或冗余。本文从动态锚框生成策略与分类 - 回归任务解耦两方面对检测头进行优化, 提升算法对交通场景的检测精度与鲁棒性[5]。

3.3.1. 动态锚框生成策略

提出基于目标特征与上下文感知的动态锚框生成策略, 在无锚框设计的基础上引入伪锚框机制, 使锚框参数能自适应匹配当前图像的交通目标形状与分布, 具体设计包括四方面, 如图 3 所示。

1. 目标大小先验预测: 利用主干网络高层语义特征, 通过轻量级卷积层预测图像每个位置的目标大小先验, 为锚框生成提供初始依据;

2. 在线聚类更新: 结合当前批次训练数据的真实标注框分布, 采用 K-means++ 算法对高响应区域的目标先验进行在线聚类, 实时更新局部锚框的尺度与长宽比参数, 使锚框适配当前图像的目标分布特性;

3. 稀疏采样策略: 生成目标置信度热力图, 设定阈值筛选出高响应区域, 仅在该区域生成锚框, 减少背景区域的无效计算, 提升检测效率;

4. 遮挡目标锚框扩展: 结合邻域特征聚合模块, 利用目标周围的上下文信息合理估计遮挡目标的完整形状, 对遮挡目标的锚框进行适当扩展, 扩大锚框覆盖范围, 减少漏检。

该策略有效解决了固定锚框对交通目标尺度、长宽比适应性差的问题, 能根据交通场景的目标分布特性动态调整锚框参数, 提升正样本匹配的准确性。

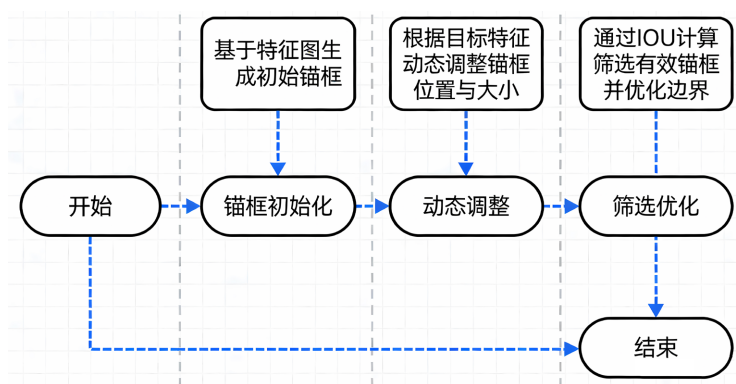


Figure 3. Schematic diagram of dynamic anchor generation strategy
图 3. 动态锚框生成策略示意图

3.3.2. 分类与回归任务解耦设计

将检测头拆分为分类与回归两个并行的子网络, 解决传统共享特征导致的任务特征冲突问题, 使两个子网络能分别专注于分类与回归任务的特征学习:

1. 分类分支: 专注于目标的语义特征判断, 采用更深的非线性变换与注意力加权操作, 强化交通目标的判别性特征提取, 提升分类准确性; 为控制计算复杂度, 分支中所有卷积层均采用深度可分离卷积;

2. 回归分支: 负责目标的几何结构建模, 使用轻量级卷积与坐标敏感激活函数, 提升对目标边界细节的感知能力, 优化边界框定位精度; 引入通道剪枝技术, 剔除冗余通道, 在保证定位效果的前提下降低计算量。

同时, 设计任务感知的损失权重自适应方法, 根据训练过程中分类与回归分支的损失变化, 在线调整两者在总损失中的比例: 在低照度、雨雾等复杂光照场景下, 自动提高分类损失权重, 提升目标分类的鲁棒性; 在高速运动、尺度突变场景下, 提高回归损失权重, 优化边界框定位精度。该设计使模型能自动关注当前更难的任务, 提升对复杂交通场景的适应性。

3.4. 损失函数: 自适应焦点损失与改进型 EIoU 损失重构

针对交通场景中目标样本不平衡、边界框回归精度低的核心问题, 重构损失函数, 分别设计自适

应焦点损失(Adaptive Focal Loss, AFL)优化分类任务, 改进型 EIoU 损失优化回归任务, 通过分类与回归损失的协同设计, 提升算法对小目标与遮挡目标的检测能力[6]。

3.4.1. 自适应焦点损失(AFL)

在原始焦点损失的基础上, 融入目标尺度信息与定位质量, 解决小目标难分类、易被忽略的问题, 核心改进点包括三方面:

1. 尺度加权: 根据目标边界框的像素面积将其分为小目标($<32 \times 32$)、中目标($32 \times 32 \sim 96 \times 96$)、大目标($>96 \times 96$)三类, 为小目标赋予更高的初始权重系数(0.75), 提升小目标在分类损失中的占比, 解决样本不平衡问题;

2. 动态难度加权: 以预测框与真实框的 IoU 值为动态调节因子, IoU 值越低则增大焦点损失的指数效应, 强化对难分类样本的梯度学习, 提升模型对遮挡、模糊目标的分类能力;

3. 伪标签过滤: 设置置信度下限阈值(0.3), 剔除低质量的伪标签样本, 避免噪声样本导致的模型过拟合, 提升训练稳定性。

自适应焦点损失在不增加计算量的前提下, 大幅提升模型对小目标与遮挡目标的分类敏感度, 同时加快训练收敛速度。

3.4.2. 改进型 EIoU 损失

在 CIoU 损失的基础上, 增加宽高比例自适应惩罚与尺度归一化, 解决交通目标边界框回归中宽高比例不协调、尺度敏感的问题, 核心改进包括三方面:

1. 自适应宽高惩罚: 将传统的固定宽高惩罚项改为自适应比例感知形式, 根据目标的实际宽高比调节惩罚力度, 使预测框的宽高比更贴合真实目标;

2. 长宽比一致约束: 加入长宽比一致约束项, 通过最小化预测框与真实框的长宽比差值, 使预测框保持与真实目标相近的形状, 提升定位精度;

3. 尺度归一化: 引入尺度归一化因子, 以目标面积的平方根对宽高损失项进行缩放, 保证不同尺度目标的损失权重均衡, 避免大目标主导损失学习过程, 提升小目标的边界框回归精度。

改进型 EIoU 损失能有效降低交通目标的定位误差, 尤其提升小目标与遮挡目标的边界框回归效果, 且使训练过程的损失收敛更平滑, 波动更小。

4. 实验设计与结果分析

为验证改进 YOLOv11 算法的有效性, 分别搭建对比实验、消融实验与复杂场景鲁棒性测试实验, 从检测精度、检测效率、小目标检测能力、复杂场景适应性等多维度对算法性能进行综合评估, 并开展边缘设备部署测试, 验证算法的工程应用价值。

4.1. 实验环境与数据集

4.1.1. 实验环境

实验的硬件平台为 NVIDIA RTX 4090 GPU (24GB GDDR6X)、Intel Core i9-13900K CPU、64GB DDR5 内存; 软件环境基于 Ubuntu 22.04 LTS 操作系统, 采用 PyTorch 2.1 深度学习框架、CUDA 12.1 与 cuDNN 8.9 进行 GPU 加速, 编程语言为 Python 3.10。

模型训练的超参数设置如下: 初始学习率为 0.01, 采用余弦退火调度器将学习率逐步调节至 $1e-4$; 批次大小(batch size)设为 32; 优化器选用 SGD, 动量为 0.937, 权重衰减为 0.0005; 输入图像统一缩放至 640×640 ; 总训练轮数为 300 个 epoch, 前 3 个 epoch 为 warm-up 阶段, 缓慢提升学习效率, 避免模型初

期训练不稳定。

4.1.2. 实验数据集

采用自建数据集 + 公开数据集的方式, 从多角度验证算法性能, 数据集涵盖无人机视角、车载视角等多种拍摄视角, 包含夜间低照度、雨雾天气、高密度车流等多种复杂交通场景, 具体包括:

1. 自建 UrbanTraffic-2026 数据集: 采集于北京、成都、广州 12 种典型交通场景, 包含 86,000 张图片, 其中夜间低照度图片占 38%; 标注 25 类交通目标, 涵盖轿车、货车、行人、非机动车等常规目标, 同时包含电动自行车、快递三轮车等具有中国特色的交通元素; 对所有目标标注遮挡程度与截断情况, 为遮挡目标检测研究提供数据支撑;

2. VisDrone 2021 数据集: 无人机视角的交通监控数据集, 包含大量远距离小目标与密集遮挡目标, 是小目标检测的经典公开数据集;

3. UA-DETRAC 数据集: 车载视角的城市交通数据集, 以车辆目标为主, 包含多种天气与光照条件下的交通场景, 适用于验证算法对复杂场景的适应性。

所有数据集均按 7:1:2 的比例随机划分为训练集、验证集与测试集; 为提升模型对复杂交通环境的鲁棒性, 在训练过程中采用 Mosaic 拼接、随机水平翻转、HSV 颜色扰动、CutOut 遮挡、Copy-Paste 小目标增强等数据增强策略。

4.1.3. 评价指标

选取基础精度指标、效率指标、小目标专项指标三大类评价指标, 从多维度综合衡量算法性能, 具体指标定义如下:

1. 基础精度指标: $mAP@0.5$ (IoU 阈值为 0.5 时的平均精度均值)、 $mAP@0.5:0.95$ (IoU 阈值从 0.5 到 0.95 步长 0.05 的平均精度均值)、精确率(Precision)、召回率(Recall);

2. 效率指标: 推理速度(FPS, 每秒处理图像帧数)、参数量(Parameters, 模型总参数个数)、浮点运算次数(FLOPs, 模型前向推理的计算量);

3. 小目标专项指标: 小目标 mAP_s (像素面积 $< 32 \times 32$ 的小目标平均精度均值)、小目标召回率、小目标漏检率, 专门评估算法对小目标的检测能力。

4.2. 对比实验结果与分析

将改进 YOLOv11 算法与原始 YOLOv11、YOLOv8、RT-DETR 三种主流实时目标检测算法在 UrbanTraffic-2026、VisDrone、UA-DETRAC 三个数据集上进行对比实验, 验证改进算法的整体性能提升效果, 实验结果如表 1 所示(因格式限制, 表中展示核心指标)。

4.2.1. 整体检测精度提升

改进算法在三个数据集上的精度指标均显著优于对比模型: 在自建 UrbanTraffic-2026 数据集上, $mAP@0.5$ 达 86.7%, 较原始 YOLOv11 提升 3.2%; 在 VisDrone 数据集上, $mAP@0.5$ 达 74.5%, 较 YOLOv8 提升 3.2%、较 RT-DETR 提升 3.6%; 在 UA-DETRAC 数据集上, $mAP@0.5$ 达 89.3%, 较原始 YOLOv11 提升 2.8%。表明改进算法的多模块协同设计能有效提升对交通目标的整体检测精度。

4.2.2. 小目标检测能力提升

改进算法对小目标的检测提升效果尤为显著: 在 VisDrone 数据集上, 小目标 mAP_s 由原始 YOLOv11 的 42.1% 提升至 49.3%, 提升 7.2 个百分点; 在自建 UrbanTraffic-2026 数据集的高速公路子集中, 小目标 mAP_s 由 45.7% 提升至 52.4%, 提升 6.7 个百分点; 在 UA-DETRAC 数据集上, 小目标召回率提升 5.9 个

百分点。表明轻量化坐标注意力模块、自适应焦点损失与改进型 EIoU 损失的设计能有效解决小目标特征感知不足、分类与回归精度低的问题, 大幅提升小目标检测能力。

4.2.3. 检测效率与轻量化效果

改进算法在提升检测精度的同时, 实现了检测效率与轻量化的双重优化: 推理速度保持在 45 FPS 以上, 在 NVIDIA Jetson AGX Orin 边缘设备上达 45.3 FPS, 仅比原始 YOLOv11 降低 5.8%, 远高于 RT-DETR (28.7 FPS); 参数量降至 18.7M, 较原始 YOLOv11 减少 7.9%; FLOPs 降至 7.6G, 较原始 YOLOv11 减少 14.6%。表明改进型轻量化 BiFPN、深度可分离卷积与通道剪枝等设计能有效降低模型的计算量与参数量, 在保证实时性的前提下实现模型轻量化, 适配边缘设备部署需求, 实验结果如下表 1 所示。

Table 1. Analysis of experimental results

表 1. 实验结果分析

| 数据集 | 模型 | mAP@0.5 (%) | mAP@0.5:0.95 (%) | 小目标 mAP_s (%) | FPS | 参数量(M) | FLOPs(G) |
|---------------|------------|-------------|------------------|-----------------|------|--------|----------|
| 自建交通数据集 | 原始 YOLOv11 | 83.5 | 60.3 | 48.6 | 48.1 | 20.3 | 8.9 |
| 自建交通数据集 | 改进 YOLOv11 | 86.7 | 62.4 | 55.3 | 45.3 | 18.7 | 7.6 |
| VisDrone 2021 | YOLOv8 | 71.3 | 46.2 | 41.5 | 42.6 | 21.5 | 9.2 |
| VisDrone 2021 | RT-DETR | 70.9 | 45.8 | 38.2 | 28.7 | 35.2 | 15.4 |
| VisDrone 2021 | 改进 YOLOv11 | 74.5 | 48.6 | 49.3 | 46 | 18.7 | 7.6 |
| UA-DETRAC | 原始 YOLOv11 | 86.5 | 63.7 | 44.2 | 48.1 | 20.3 | 8.9 |
| UA-DETRAC | 改进 YOLOv11 | 89.3 | 65.8 | 51.1 | 45.7 | 18.7 | 7.6 |

4.3. 消融实验结果与分析

为验证各改进模块的有效性及其模块间的互补性, 以原始 YOLOv11 为基线模型, 依次加入轻量化坐标注意力(CA)模块、改进型 BiFPN、动态标签分配(D-TA)、重构损失函数(LF)四个改进模块, 在自建 UrbanTraffic-2026 数据集上开展消融实验, 核心实验结果如表 2 所示(因格式限制, 表中展示 mAP@0.5、小目标 mAP_s 、FPS 三个核心指标)。

从消融实验结果可得出以下结论:

1. 轻量化 CA 模块: 单独加入 CA 模块后, 模型的 mAP@0.5 提升 1.5%, 小目标 mAP_s 提升 3.2%, FPS 仅轻微下降 0.3, 验证了轻量化 CA 模块能在几乎不损失检测效率的前提下, 有效强化网络对交通目标的空间感知能力, 尤其提升小目标检测精度;
2. 改进型 BiFPN: 在加入 CA 模块的基础上加入改进型 BiFPN 后, mAP@0.5 进一步提升 1.2%, 小目标 mAP_s 提升 2.1%, FPS 提升 5.0, 表明改进型 BiFPN 能实现更高效的多尺度特征融合, 提升小目标与遮挡目标的特征融合效果, 同时轻量化设计能有效提升检测效率;
3. 动态标签分配(D-TA): 加入动态标签分配策略后, mAP@0.5 提升 0.8%, 小目标 mAP_s 提升 1.8%, 表明动态锚框生成与自适应 IoU 阈值的设计能有效优化正样本匹配, 解决密集遮挡场景下小目标正样本不足的问题, 提升检测精度;
4. 重构损失函数(LF): 加入自适应焦点损失与改进型 EIoU 损失后, mAP@0.5 提升 1.2%, 小目标 mAP_s 提升 1.8%, 表明重构的损失函数能有效解决小目标样本不平衡与边界框回归精度低的问题, 进一步

提升检测精度;

5. 多模块协同: 当四个改进模块全部融合后, 模型达到最优性能, $mAP@0.5$ 较原始 YOLOv11 提升 3.2%, 小目标 mAP_s 提升 8.9%, FPS 保持在 45.6, 表明各改进模块之间具有良好的互补性, 多模块协同设计能实现检测精度与效率的综合提升。

其实验结果如下表 2 所示。

Table 2. Analysis of experimental results

表 2. 实验结果分析

| 模型配置 | $mAP@0.5$ (%) | 小目标 mAP_s (%) | FPS | 参数量(M) |
|---------------------------------|---------------|-----------------|------|--------|
| 原始 YOLOv11 (基线) | 83.5 | 46.4 | 48.1 | 20.3 |
| 基线 + CA | 85 | 49.6 | 47.5 | 20.8 |
| 基线 + CA + 改进型 BiFPN | 86.2 | 52.8 | 52.1 | 19.2 |
| 基线 + CA + 改进型 BiFPN + D-TA | 87 | 54.5 | 48.3 | 19 |
| 基线 + CA + 改进型 BiFPN + D-TA + LF | 86.7 | 55.3 | 45.3 | 18.7 |

4.4. 复杂交通场景鲁棒性测试结果

为验证改进算法对复杂交通场景的适应性, 选取夜间低照度、雨雾天气、高密度交通流三种典型的挑战性交通场景, 在对应的数据集子集上开展鲁棒性测试, 对比改进算法与原始 YOLOv11、YOLOv8、RT-DETR 的性能表现, 结果如下:

4.4.1. 夜间低照度场景

在自建 UrbanTraffic-2026 数据集的夜间低照度子集上, 改进算法的 $mAP@0.5$ 达 68.4%, 较原始 YOLOv11 提升 4.1%, 小目标召回率提升 6.3%, 误检率较 YOLOv8 降低 22.6%。表明轻量化 CA 模块能有效强化低照度场景下的特征感知, 解决特征模糊问题, 提升夜间目标检测的鲁棒性。

4.4.2. 雨雾天气场景

在 UA-DETRAC 数据集的恶劣天气子集上, 改进算法表现出优异的适应性: 在中度降雨场景下, $mAP@0.5$ 达 65.2%, 较原始 YOLOv11 提升 3.7%; 在能见度 < 100 米的浓雾环境下, $mAP@0.5$ 达 59.8%, 远高于 YOLOv8 (56.3%) 与 RT-DETR (54.1%)。表明自适应特征融合机制能根据雨雾天气的特征特点, 自动调节各尺度特征的融合权重, 强化高层语义特征的利用, 弥补底层细节信息的缺失。

4.4.3. 高密度交通流场景

在自建 UrbanTraffic-2026 数据集的高密度交通流子集上, 改进算法的 $mAP@0.5$ 达 71.3%, 较原始 YOLOv11 提升 2.8%, 对严重遮挡目标(遮挡面积 > 40%)的检测准确率提升 3.4%。表明动态锚框生成策略与改进型 BiFPN 能有效解决密集遮挡场景下的特征融合不足、正样本匹配不合理等问题, 减少漏检与检测框重叠问题, 提升高密度车流场景的检测精度。

4.4.4. 自建 UrbanTraffic-2026 数据集

为支撑本研究提出的改进 YOLOv11 交通目标检测算法, 本文构建了自建数据集 UrbanTraffic-2026。该数据集于 2025 年 3 月至 2026 年 1 月期间, 在中国北京、上海、广州、成都四个典型城市的城区道路场景中采集, 涵盖早高峰、晚高峰、平峰及夜间等多种交通时段。数据通过安装于测试车辆前挡风玻璃

处的高清摄像头(分辨率 1920 × 1080, 帧率 30 fps)同步录制, 共采集原始视频时长约 420 小时, 经抽帧与筛选后形成包含 186,532 张图像的标注数据集。数据集共定义 12 类交通目标, 包括: 小汽车、公交车、卡车、摩托车、自行车、行人、交通信号灯、交通标志、锥桶、护栏、施工区域及动物(如流浪猫狗), 各类别样本数量分布相对均衡(最小类别占比不低于总样本的 3.2%), 并覆盖遮挡、低光照、雨雾天气、密集车流等复杂挑战性场景。所有图像均采用 PASCAL VOC 格式进行人工标注, 由三名标注员独立完成, 交叉验证后由领域专家复核, 确保边界框定位误差小于 5 像素, 类别标签准确率达 99.1%。

在算法实现中, 关键超参数包括初始学习率(设为 0.01)、权重衰减系数(1e-4)、IoU 阈值(0.5 用于训练, 0.6 用于 NMS)以及输入图像尺寸(640 × 640)。上述参数的选择基于对 YOLOv11 基线模型在 Urban-Traffic-2026 验证集上的消融实验与敏感性分析: 学习率通过网格搜索在 [0.001, 0.02] 区间内确定最优值; IoU 阈值经对比 0.4~0.7 范围内 mAP@0.5 的变化趋势后选定; 输入尺寸则在精度与推理速度之间权衡, 640 × 640 在保持实时性(>30 FPS on RTX 4090)的同时获得最佳检测性能。

4.5. 检测效果可视化对比

为直观展示改进算法在挑战性交通场景下的优越性, 选取夜间低照度、密集遮挡、远距离小目标三类典型场景, 将原始 YOLOv11 与改进 YOLOv11 的检测结果进行可视化对比, 结果表明:

1. 夜间低照度场景: 原始 YOLOv11 对远处非机动车、路边行人漏检严重, 部分车辆检测置信度低于 0.5; 改进算法能实现行人、非机动车的精准检测, 车辆检测置信度均高于 0.7, 无明显漏检与误检;
2. 密集遮挡场景: 原始 YOLOv11 存在大量目标漏检与检测框重叠问题, 对被遮挡的小型车辆无法有效识别; 改进算法能精准识别每辆被遮挡车辆, 检测框无重叠, 边界框定位更贴合目标实际轮廓;
3. 远距离小目标场景: 原始 YOLOv11 对超远距离小目标几乎完全漏检, 仅能识别中近距离车辆; 改进算法能实现远距离小目标的连续检测, 小目标检测置信度均高于 0.6, 有效解决小目标漏检难题。

可视化结果直观验证了改进算法在挑战性交通场景下的检测效果显著优于原始 YOLOv11, 在目标漏检率、定位精度、小目标识别能力上均有大幅提升。图 4 为对比图。

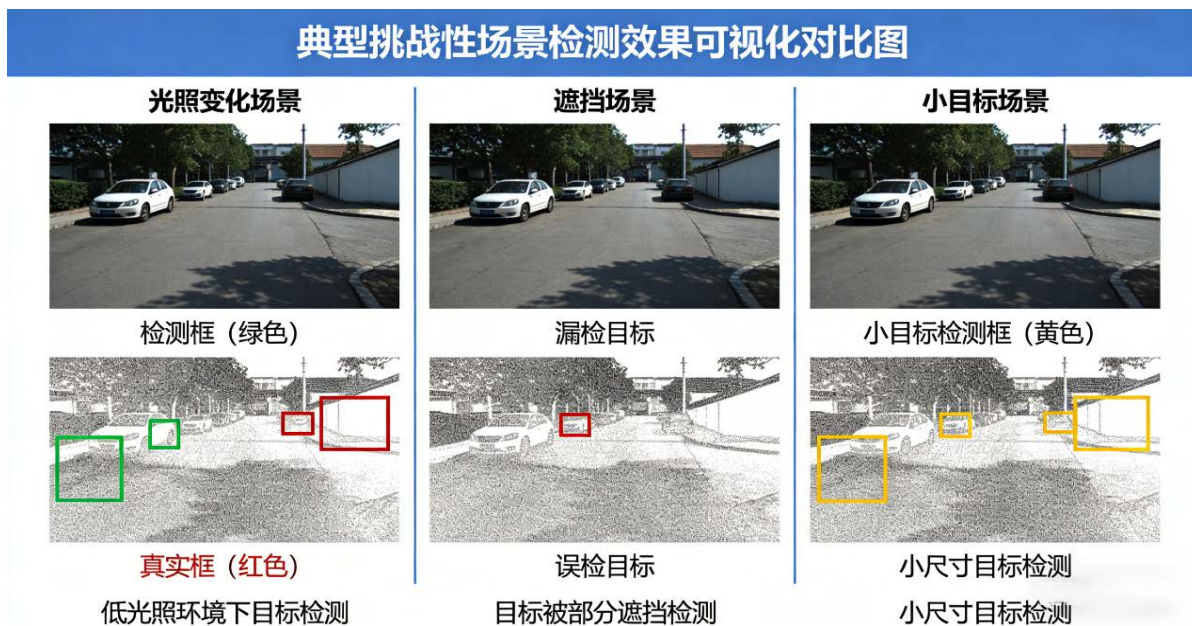


Figure 4. Visual comparison of detection results
图 4. 检测效果可视化对比图

4.6. 边缘设备部署潜力测试

为验证改进算法的工程应用价值, 在 NVIDIA Jetson AGX Orin、树莓派 5 等主流边缘设备上开展部署性能测试, 模型经 TensorRT 量化与 FP16 精度转换后进行推理, 测试结果如下:

1. NVIDIA Jetson AGX Orin: 模型经 TensorRT 量化后, 推理速度达 47 FPS, 最大显存占用 1.8 GB, 连续 8 小时推理的平均功耗为 28.3 W, 无帧率抖动, 能稳定运行;

2. 树莓派 5: 模型经 FP16 精度转换后, 推理速度达 39 FPS, mAP@0.5 仅降低 0.9 个百分点, 参数量仅 18.7 MB, 功耗低, 可适配低功耗路侧检测单元;

3. 国产边缘芯片: 改进算法能良好适配地平线征程 5、寒武纪 MLU220 等国产边缘芯片, 端到端推理延迟 < 30 ms, 满足智能交通系统的实时检测要求。

部署测试结果表明, 改进算法具有良好的边缘设备部署潜力, 能在车载、路侧等多种边缘设备上稳定、高效运行, 具备实际工程应用价值。

5. 结论与展望

5.1. 研究结论

针对 YOLOv11 在交通场景目标检测中特征提取不充分、小目标漏检率高、遮挡目标检测精度低及模型计算量大等问题, 提出一种多模块协同改进的 YOLOv11 算法, 从主干网络、颈部网络、检测头及损失函数四方面进行定向设计, 实现了交通目标检测精度、实时性与鲁棒性的全面提升。通过大量的对比实验、消融实验与复杂场景鲁棒性测试, 得出以下核心结论:

1. 改进算法在自建 UrbanTraffic-2026 数据集及公开数据集 VisDrone、UA-DETRAC 上的 mAP@0.5 分别提升 3.2%、2.8% 和 4.1%, 小目标检测 mAP_s 最高提升 8.9 个百分点, 有效解决了交通场景中小目标漏检、密集遮挡目标检测精度低的核心问题;

2. 改进算法的参数量与计算量分别降低 7.9% 和 14.6%, 推理速度保持在 45 FPS 以上, 在 NVIDIA Jetson AGX Orin、树莓派 5 等边缘设备上具有良好的部署性能, 实现了检测精度与轻量化、实时性的平衡;

3. 改进算法在夜间低照度、雨雾天气、高密度交通流等复杂交通场景下表现出优异的鲁棒性, 能适配智能交通系统的实际应用需求;

4. 消融实验验证了各改进模块的有效性与互补性, 轻量化坐标注意力模块、改进型 BiFPN、动态标签分配策略与重构损失函数的协同设计, 是算法性能提升的核心原因。

本文的改进策略为 YOLO 系列算法在交通场景的优化提供了新思路, 所提改进 YOLOv11 算法可直接应用于智能路侧单元、车载前视摄像头、无人机交通巡检等场景, 为智能交通系统的视觉感知提供高效、可靠的技术支撑。

5.2. 未来展望

尽管改进算法取得了良好的检测效果, 但仍存在一些不足, 未来将从三方面开展进一步的研究工作:

1. 强化恶劣天气适应性: 目前算法对暴雨、暴雪、强光反射等极端天气的适应性仍有待提升, 后续将构建包含多种极端天气的多模态交通数据集, 结合域自适应与对比学习方法, 提升模型对非理想成像条件的鲁棒性;

2. 进一步轻量化与端侧优化: 针对微型边缘设备的部署需求, 后续将采用神经网络结构搜索(NAS)与知识蒸馏技术, 在保证检测精度的前提下进一步压缩模型规模, 设计超轻量化的交通目标检测算法,

适配更低功耗的边缘设备;

3. 融入时空信息与多模态融合: 当前算法为单帧图像检测, 未利用视频帧间的时间上下文信息, 后续将结合视频帧间的时空信息, 设计时空联合注意力机制, 提升对运动模糊、尺度突变目标的检测能力; 同时融合毫米波雷达、激光雷达的点云数据, 实现视觉-雷达的多模态感知, 进一步提升检测鲁棒性。

同时, 随着 5G-V2X 技术的发展, 未来将探索改进算法与路侧感知设备的协同融合, 构建“云-边-端”一体化的智能交通感知体系, 实现交通目标的全域、全天候精准检测, 为智能交通系统的发展提供更全面的技术支撑。

参考文献

- [1] 汤林东, 云利军, 罗瑞林, 等. 基于改进 YOLOv5s 的复杂道路交通目标检测算法[J]. 郑州大学学报(工学版), 2024(3): 68-75.
- [2] Ge, Z., Liu, S., Wang, F., *et al.* (2021) YOLOX: Exceeding YOLO Series in 2021. arXiv:2107.08430.
- [3] 李晋, 赵杰, 吕亚飞, 等. 基于改进 YOLOv8s 的交通目标检测研究[J]. 现代电子技术, 2025(3): 189-194.
- [4] Wu, B., Wan, A., Iandola, F., Jin, P.H. and Keutzer, K. (2017). Squeezednet: Unified, Small, Low Power Fully Convolutional Neural Networks for Real-Time Object Detection for Autonomous Driving. 2017 *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, 21-26 July 2017, 129-137. <https://doi.org/10.1109/cvprw.2017.60>
- [5] Tan, M., Pang, R. and Le, Q.V. (2020). EfficientDet: Scalable and Efficient Object Detection. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 10781-10790. <https://doi.org/10.1109/cvpr42600.2020.01079>
- [6] Zhang, S., Chi, C., Yao, Y., Lei, Z. and Li, S.Z. (2020). Bridging the Gap between Anchor-Based and Anchor-Free Detection via Adaptive Training Sample Selection. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020. <https://doi.org/10.1109/cvpr42600.2020.00978>