

车辆检测中YOLO模型的综合性能评估与实证分析

王少煜¹, 王志刚¹, 王 鑫¹, 刘 鸣¹, 侯典立^{2*}

¹烟台市大数据中心, 山东 烟台

²鲁东大学能源动力与电气工程学院, 山东 烟台

收稿日期: 2025年12月11日; 录用日期: 2026年1月14日; 发布日期: 2026年1月28日

摘要

针对交通路口复杂场景中车辆检测存在的实时性与精度难以兼顾、适应性不足的问题, 本文提出一种融合精度、召回率及检测帧率的加权综合评价指标。基于VisDrone数据集, 系统评估YOLOv5至YOLOv10等主流模型, 并分析图像清晰度提升与掩模策略的优化效果。实验结果表明, YOLOv10n在综合指标得分最高, 达150.41; 提高图像像素使综合性能提升幅度达23.6%, 显著优于单纯增加模型复杂度。研究为智能交通系统中的车辆检测提供了高效且适应性强的技术支持。

关键词

车辆检测, 综合指标, YOLO系列模型, 智能交通

Comprehensive Performance Evaluation and Empirical Analysis of YOLO Models for Vehicle Detection Evaluation

Shaoyu Wang¹, Zhigang Wang¹, Xin Wang¹, Ming Liu¹, Dianli Hou^{2*}

¹Yantai Big Data Center, Yantai Shandong

²School of Energy Power and Electrical Engineering, Ludong University, Yantai Shandong

Received: December 11, 2025; accepted: January 14, 2026; published: January 28, 2026

Abstract

To address the challenges of balancing real-time performance and detection accuracy, as well as

文章引用: 王少煜, 王志刚, 王鑫, 刘鸣, 侯典立. 车辆检测中 YOLO 模型的综合性能评估与实证分析[J]. 交通技术, 2026, 15(1): 102-109. DOI: 10.12677/ojtt.2026.151009

insufficient adaptability in complex traffic intersection scenarios, this study proposes a weighted comprehensive evaluation metric that integrates precision, recall, and detection frame rate. Based on the VisDrone dataset, mainstream models from YOLOv5 to YOLOv10 are systematically evaluated, and optimization strategies such as image clarity enhancement and mask application are analyzed. Experimental results show that YOLOv10n achieves the highest comprehensive score of 150.41. Enhancing image resolution improves overall performance by 23.6%, which is significantly more effective than simply increasing model complexity. This research provides an efficient and adaptable technical solution for vehicle detection in intelligent transportation systems.

Keywords

Vehicle Detection, Composite Metrics, YOLO Series Models, Intelligent Transportation

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

交通路口车辆检测是保障交通安全与提升通行效率的关键，直接影响智能交通系统的可靠性[1]。高空摄像头虽具广域覆盖优势，但远视角下车辆像素占比低、非道路区域干扰强，易导致漏检；现有研究多依赖单一指标评估模型性能，难以兼顾复杂场景中精度与效率的平衡。YOLO 系列模型因轻量化设计被广泛应用于边缘设备，但在多维度性能优化方面仍存在不足[2]。针对上述问题，本文从算法创新性、硬件适配性与场景针对性三方面出发，选取 YOLOv5 至 YOLOv10 等主流模型，提出融合精度、召回率与检测帧率的加权综合评价指标，并基于 VisDrone 数据集系统评估图像清晰度优化效果。

2. YOLO 模型及检测指标

2.1. YOLO 模型分析

YOLO 将目标检测转化为单阶段回归问题，实现了高效的端到端检测[3]。YOLO 模型架构如图 1 所示，YOLO 模型通过多尺度特征提取(骨干网络)、特征融合(颈部网络)及边界框预测(检测头)的架构设计，在精度、实时性与泛化能力之间取得良好平衡。其具备全局上下文感知、动态优化机制和轻量化部署潜力，成为交通路口车辆检测等实时场景的主流方案[4]。

YOLO 系列模型在实时目标检测领域不断演进[5]：YOLOv1 作为首个单阶段检测模型，具备实时性优势，但定位精度低、对小目标不敏感。YOLOv2 在尺度感知阶段引入 Anchor 机制，YOLOv3 采用 FPN 结构提升了多尺度预测能力[6]，YOLOv4 引入路径聚合网络和多种数据增强技术，进一步提升了性能[7]。YOLOv5 采用了 Pytorch 框架，利用 CSP-Net 结构作为骨干网络，提高检测的准确性和速度[8]。YOLOv6 通过参数压缩降低模型规模，但社区支持较弱。YOLOv7 与 YOLOv8 在继承前代核心优势的基础上，通过引入新的网络结构与模块设计，在检测性能和运行速度上实现进一步提升[9]。YOLOv9 算法通过创新的设计优化了目标检测的精度和速度。YOLOv10 通过空间通道解耦下采样消除 NMS (Non-Maximum Suppression, NMS)后处理，显著降低推理延迟，适合轻量化应用[10]。YOLOv11 集成目标跟踪与语义分割模块，支持多任务联合分析；YOLOv12 引入跨阶段局部注意力机制，增强对遮挡目标的检测能力[11]。

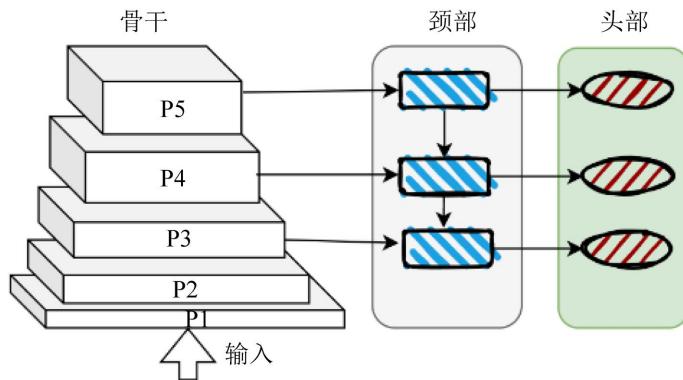


Figure 1. YOLO model architecture
图 1. YOLO 模型架构

2.2. 传统检测指标

评价指标通常需要有一个标准的测试集或验证集进行评估，目的是评估算法的准确性、效率、鲁棒性和泛化能力等方面性能[12]。图像检测性能的评估主要围绕精度、速度和资源消耗等关键指标展开。其中，精度相关指标如下：

1) 精度(Precision, P): 指在所有被预测为正样本的实例中，实际为正样本(True Positive, TP)的比例，即查准率。

$$P = TP / (TP + FP) \quad (1)$$

2) 召回率(Recall, R): 在所有真实正样本中，被模型正确检测到的比例，即查全率。该指标用于衡量模型的漏检率，即模型未能检测到的真实正样本的比例。

$$R = TP / (TP + FN) \quad (2)$$

3) 平均精度(Average Precision, AP): 表示单个类别的精度 – 召回率曲线(PR 曲线)下的面积，用于衡量模型对某一特定类别的检测能力。

$$AP = \int_0^1 p(r) dr \quad (3)$$

4) 所有类别平均精度的均值(Mean Average Precision, mAP)作为目标检测领域最核心的指标，mAP 综合评估了模型对所有类别的检测性能。

5) 速度相关指标：检测帧率(Frames Per Second, FPS)衡量模型实时处理能力的关键指标，表示模型每秒能够处理的图像帧数。单张图像推理所需的浮点运算次数(Floating Point Operations, FLOPs)反映模型计算复杂度的量化指标，用于评估模型在单次推理过程中需要执行的浮点运算数量。FLOPs 越低，模型的计算效率越高，对硬件资源的需求也相对较低。

6) 资源消耗相关指标：参数量指模型中可训练参数的总数量，直接关系到模型的复杂度、内存占用以及部署难度。参数量越大，模型通常具有更强的表达能力，但同时也需要更多的存储空间和计算资源。

2.3. 综合检测指标

当前车辆检测算法在单一性能指标上已取得显著进展，然而如何协同优化多个关键指标仍是一个亟待解决的挑战[12]。在本系统应用场景中，目标检测模型需同时满足三项核心需求：高精度检测、高召回率以确保目标全覆盖，以及满足帧率约束的实时处理能力。现有评估体系存在明显局限：首先，传统平

均精度未考虑类别重要性差异, 难以反映对关键目标的检测效能; 其次, 缺乏对计算效率与精度联合优化的量化机制, 高精度模型常因计算复杂度超标而无法部署于实时系统; 最后, 未针对具体任务场景对精度与召回率进行权衡建模, 难以指导模型在漏检敏感场景下的优化方向。因此, 亟需构建面向多目标协同优化的评估与设计框架。

鉴于上述问题, 考虑到精度、召回率和实时处理能力均为正向影响因子, 且 FPS 易受计算量与模型规模的制约, 综合权衡类别间重要性差异、精度与召回率的相对重要性以及检测速率, 本研究提出了一种加权综合评价指标(Weighted Composite Index, WCI)。

$$WCI = (\alpha P_w + \beta R_w)FPS \quad (4)$$

其中, α 为精度权重系数, β 为召回率权重系数, 且 $\alpha + \beta = 1$ 。权重设定需契合智能交通场景的实际需求: 漏检高风险目标(如闯红灯车辆)可能导致严重安全事故, 故召回率应被赋予不低于精度的优先级。乘性结构显式建模了“高精度但低帧率”模型在边缘部署中的性能折损, 相较于加法融合更能反映实际系统约束。 P_w 为类别加权平均精度:

$$P_w = \sum_{c=1}^N w_c P_c \quad (5)$$

R_w 为类别加权平均召回率为:

$$R_w = \sum_{c=1}^N v_c R_c \quad (6)$$

w_c 、 P_c 表示类别 c 的加权系数和精度, v_c 、 R_c 表示类别 c 的加召回率加权系数和召回率, N 为类别总数, $\sum_{c=1}^N w_c = 1$, $\sum_{c=1}^N v_c = 1$ 。

加权综合指标具有多维度融合能力, 通过调整权重系数 α 和 β , 可以灵活适配不同场景下精度和召回率重要性差异需求, 借助 w_c 和 v_c 权重, 能够突出不同受检测目标重要性差异。FPS 作为乘数纳入考量, 旨在精度与速度间寻找最优平衡, 从而筛选出最合适的模型。本文将综合指标与领域内主流多目标评价方法进行对比, 如表 1 所示:

Table 1. Feature comparison of multi-objective evaluation methods

表 1. 多目标评价方法特性对比

方法	融合方式	类别加权	惩罚低 FPS	适用场景
F1-score + FPS 加权和	加法	否	弱	离线评估
Pareto 前沿分析	非支配排序	是	否	算法研发阶段(无单一排序)
Cost-sensitive 指标	成本矩阵	是	间接	特定业务损失可量化场景
本文综合指标	乘法	是	强	实时边缘部署、召回优先场景

相较于 F1-FPS 加权和、Pareto 前沿分析等方法, WCI 具有三方面优势: (1) 通过加权精度/召回率显式建模类别重要性; (2) 乘性结构天然惩罚低帧率模型, 契合边缘计算约束; (3) 权重参数 α , β 具备明确物理意义, 便于依据业务策略动态调整。WCI 主要适用于实时性敏感、召回优先的在线监控场景。

3. 实验结果与分析

3.1. 测试环境

本研究的实验所用的 PC 配置如下: CPU Intel i9-9900KF、操作系统为 Windows10, 配备双 RTX3060Ti

显卡。为了评估模型性能, 本研究采用了 VisDrone2019 数据集, 该数据集包含 8629 张由无人机拍摄的图像, 其中 6471 张用于训练, 548 张用于验证, 1610 张用于测试[13]。数据集涵盖了多种场景和目标类型, 包括轿车、公交车、行人、车辆和自行车等十类, 分配情况如表 2 所示。

Table 2. Class distribution and weight coefficients for the weighted composite index in VisDrone2019

表 2. VisDrone2019 类别分配及加权综合指标权重系数

编号	名称	数量	n_c/n_{total}	ω_c/v_c
0	行人	79,937	0.23	0.2
1	人	27,057	0.079	0.01
2	自行车	10,477	0.031	0.02
3	汽车	144,865	0.422	0.598
4	货车	24,950	0.073	0.073
5	卡车	12,871	0.038	0.038
6	三轮车	4803	0.014	0.014
7	遮阳篷三轮车	3243	0.01	0.01
8	公共汽车	5926	0.017	0.017
9	摩托车	29,642	0.086	0.02

3.2. 评价参数设置

VisDrone2019 数据集中各类别的数量不仅反映了相应目标的出现频率, 也在一定程度上体现了其对交通状况的影响程度, 这些数据经过归一化处理后, 可作为精度和召回率的权重系数。

$$\omega_c = v_c = n_c / n_{total} \quad (7)$$

其中 n_c 为第 c 类目标的数量, n_{total} 则是所有目标的总数量。鉴于行人、人、自行车和摩托车对交通影响相对较小, 且由于这些目标较小, 导致其检测精度和召回率也偏低, 为了降低这些类别对整体评估的影响, 适当减小了行人和人的加权系数。相比之下汽车对交通的重要性最高, 且相对于公交车、货车和卡车等体积较小, 因此增加了汽车类别精度和召回率系数, 同时确保等式(7)成立。经过上述调整, 最终确定加权综合指标权重系数如表 2 所示。

在智能交通领域, 对于关键目标召回率往往比检测精度更具有重要意义。这尤其体现在一些容易混淆的类别区分上, 例如行人和人的区分、公共汽车、货车和卡车区分、自行车和摩托车的区分以及三轮车和遮阳篷三轮车的区分等, 因此选取 $\alpha = 0.4$ 、 $\beta = 0.6$ 。

3.3. 原始模型对比测试

训练设置如下: 输入图像分辨率为 640×640 , 初始学习率设定为 0.01, 动量参数为 0.937, 权重衰减因子为 0.0005, 每批次处理的图像数量为 16, 整个训练过程迭代 200 次。鉴于随着模型规模的增大, 虽然精度和召回率有所提升, 但实时性会大幅下降, 因此本文主要针对参数量较小的 YOLO 模型版本进行测试, 并采用加权综合指标进行综合评估。多个版本小模型的加权平均精度 P_w 、加权平均召回率 R_w 、每秒帧数 FPS 、加权综合指标 WCI 测试数据如表 3 所示。

相较于基础版本 YOLO11n (nano, n), 其更大规模的变体 YOLO11s (small, s)、YOLO11m (medium, m)、YOLO11l (large, l) 和 YOLO11x (extra large, x) 的加权综合指标分别下降了 1.78%、9.35%、16.75% 和

37.22%。虽然这些模型的精度和召回率有所提升，但实时性却大幅下降。与参数量较小的 n 和 s 系列模型相比，YOLO11 和 YOLO12 版本的综合指标较低，主要受限于较长的后处理时间。例如，YOLO11n 的推理时间为 2.9 ms，而后处理时间却长达 5 ms。而 YOLOv10 版本的综合指标得分最高，这主要得益于其无非极大值抑制的训练策略。YOLOv10n 的推理时间为 3.4 ms，后处理时间仅为 0.2 ms，实现了后处理时间的突破性下降，从而显著提升了实时性。该优势源于其端到端检测架构与高效骨干设计的协同效应：一方面，通过全局 - 局部双路径注意力模块增强关键区域特征表达；另一方面，移除冗余检测头并采用一致匹配策略，彻底消除 NMS 后处理瓶颈。这种“结构 - 流程”联合优化，使其在维持高召回的同时实现业界领先的推理速度。

Table 3. Comprehensive evaluation results of multiple small-scale model variants

表 3. 多个版本小模型综合指标测试数据

YOLO 版本	P_w	R_w	FPS	WCI
v5n	0.5537	0.5835	138.89	78.89
v8n	0.5738	0.5925	135.14	78.81
v10n	0.5717	0.5926	256.41	150.41
11n	0.5680	0.5907	121.95	70.93
12n	0.5864	0.5907	120.48	70.91
v5s	0.6523	0.6163	108.7	68.94
v8s	0.6641	0.6201	109.89	70.56
v10s	0.6390	0.6274	232.56	147.25
11s	0.6522	0.6301	108.7	69.67
12s	0.6629	0.6347	114.94	74.58
11m	0.7167	0.6581	94.34	64.30
11l	0.7300	0.6636	84.75	59.05
1x	0.7464	0.6696	62.89	44.53

3.4. 小目标检测优化测试

在高空俯视视角下，车辆目标通常仅占据极少数像素，导致小目标检测极具挑战性，并显著提升漏检率。在追求模型轻量化的过程中，面临检测精度的下降，在增强特征融合等优化提升方面导致模型的计算参数量和复杂度显著上升[14]。文献[15]通过将原有的 P3~P5 结构修改为 P2~P4，其中 P2 层能够捕获更密集的像素级细节，成功将小目标的 AP 值提升了 15.9%。文献[16]通过超分辨率重建、分块放大或直接输入高分辨率图像等技术来提升目标的像素密度，将目标的像素密度提升 4 倍后，小目标的 AP 提升了 25.5%。

基于上述研究进展，本文对修改检测头为 P2~P4 的 YOLO11n 进行了测试，并同时对 YOLOv10n 和 YOLO11n 进行了高分辨率图像(1280×1280)输入的测试。由于显卡存储空间有限，将批量大小调整为 4，而其他训练参数保持不变。小目标检测的测试数据如表 4 所示。

YOLO11nP234：该模型基于 YOLO11n，通过删除 P5 检测层并增加 P2 检测层进行优化。由于 P2 层的单层计算量是 P5 层的 16 倍，虽然精度和召回率有所提升，但每秒帧数下降了 23.36%，导致其综合指标低于 YOLO11l。11n1280 和 10n1280 模型 YOLO11n 和 YOLOv10n 将输入图像的分辨率从 640×640 提

升至 1280×1280 进行测试。结果显示, 这种分辨率的提升相较于检测头的调整, 带来了更为显著的性能提升。其精度与召回率接近中等规模(m 级)模型, 但 FPS 与综合指标均更优。

Table 4. Test data for small object detection**表 4. 小目标检测测试数据**

YOLO 版本	P_w	R_w	FPS	WCI
11nP234	0.5418	0.6254	93.46	54.54
11n1280	0.7185	0.6954	105.3	74.42
10n1280	0.7053	0.6941	181.8	127.02

4. 结束语

本研究构建了融合精度、召回率与检测帧率的多维评价体系, 并通过优化图像处理流程与部署架构, 提供了一种在精度与效率之间实现有效平衡的车辆检测解决方案。通过策略性地提升图像质量以及进行算法协同优化, 相较于单纯增加模型复杂度, 能够带来更具实践价值的性能提升。然而, 研究受限于单一场景数据、固定硬件平台及有限天气条件。未来工作将聚焦于构建多场景全天候基准数据集、开展边缘端量化部署实验, 并探索多模态融合策略以提升极端环境下的检测鲁棒性。

基金项目

烟台市智慧城市创新实验室科研课题项目 SDGP370600000202302000504。

参考文献

- [1] 秦敏浩, 孙未未. 基于隐状态预测的失真交通信号灯路口控制策略[J]. 计算机工程, 2025, 51(9): 1-13.
- [2] 廖炎华, 鄢元霞, 潘文林. 基于 YOLOv9 的交通路口图像的多目标检测算法[J]. 计算机应用, 2025, 45(8): 2555-2565.
- [3] Jocher, G., et al. (2020) YOLOv5. <https://github.com/ultralytics/yolov5>
- [4] 孟昱煜, 马银宝, 火久元. 基于无人机航拍图像的实时车辆检测算法[J/OL]. 浙江大学学报(工学版), 2025, 1-12. <https://link.cnki.net/urlid/33.1245.T.20250917.1609.008>, 2026-01-16.
- [5] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/cvpr.2016.91>
- [6] Ali, M.L. and Zhang, Z. (2024) The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection. *Computers*, 13, Article 335.
- [7] 张瑶, 张俊三, 马俊朋, 等. CAFR-YOLO: 基于 YOLOv8 的多尺度目标检测算法[J/OL]. 计算机工程, 2025, 1-13. <https://doi.org/10.19678/j.issn.1000-3428.0252821>, 2026-01-16.
- [8] 陈胜宝, 施隆照. YOLO-PD: 轻量级实时行人检测算法[J/OL]. 计算机工程与应用, 2025, 1-15. <https://link.cnki.net/urlid/11.2127.TP.20251104.1240.006>, 2026-01-16.
- [9] 艾强, 冯永安, 王灵超, 等. 融合通道剪枝的轻量化 YOLOv8 交通标志检测算法[J/OL]. 电子测量技术, 2025, 1-14.
- [10] Wang, A., Chen, H., Liu, L.H., et al. (2024) YOLOv10: Real-Time End-to-End Object Detection. *Advances in Neural Information Processing Systems*, 37, 107984-108011.
- [11] Tian, Y.J., Ye, Q.X. and Doermann, D. (2025) YOLOv12: Attention-Centric Real-Time Object Detectors. <https://arxiv.org/abs/2502.12524>
- [12] 游昊, 吕文涛, 叶丹, 等. 基于深度学习的车辆检测方法研究进展[J]. 无线电工程, 2025, 55(2): 230-245.
- [13] Zhu, P., Wen, L., Du, D., Bian, X., Fan, H., Hu, Q. and Ling, H. (2023) Detection and Tracking Meet Drones Challenge. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 45, 12357-12376.

-
- [14] 林树铭, 冯桑, 谭俊楠. 航拍视角下小尺度车辆精确检测方法[J/OL]. 电光与控制, 2025, 1-10. <https://link.cnki.net/urlid/41.1227.tn.20251104.1805.002>, 2026-01-16.
 - [15] Wang, Z., Su, Y., Kang, F., Wang, L., Lin, Y., Wu, Q., *et al.* (2025) Pc-YOLO11s: A Lightweight and Effective Feature Extraction Method for Small Target Image Detection. *Sensors*, **25**, Article 348. <https://doi.org/10.3390/s25020348>
 - [16] 蔡心悦, 周杨, 胡校飞, 等. 基于超分辨率重建的小目标智能检测算法[J]. 激光与光电子学进展, 2023, 60(12): 51-59.