

RCB-YOLOv5s: 基于改进YOLOv5s算法的高速护栏安装立柱检测

田 哲, 阮久宏, 李国栋*

山东交通学院轨道交通学院, 山东 济南

收稿日期: 2026年7月18日; 录用日期: 2026年8月19日; 发布日期: 2026年8月27日

摘 要

随着高速公路里程的快速增长, 高速公路立柱的自动化检测已成为智能交通基础设施安全巡检与维护的一项重要任务。然而, 高速公路环境极为复杂, 存在强烈的背景干扰以及目标尺度变化显著等问题, 严重限制了传统检测算法的准确性。为应对这些关键的工程挑战, 本文提出了一种基于YOLOv5的改进目标检测算法, 这对于实现可靠、实时的高速公路基础设施监测具有重要意义。首先, 在模型架构中集成了RepVGG模块, 利用结构重参数化技术, 在不降低推理速度的前提下, 增强了对边缘和纹理特征的提取能力。其次, 嵌入了上下文Transformer (CoTAttention)机制以抑制背景噪声, 并采用双向特征金字塔网络(BiFPN)重构网络颈部(Neck), 从而加强了针对远处小尺度目标的多尺度特征融合。在自建的高速公路立柱数据集上进行的大量实验验证了所提方法的优越性。与原始YOLOv5模型相比, 改进算法的平均精度均值(mAP@0.5)提升了4.7%, 精确率(Precision)和召回率(Recall)分别达到了88.4%和85.1%。综上所述, 本研究成功实现了高检测精度与实时效率的有效平衡, 为高速公路基础设施的智能巡检与安全预警系统提供了坚实的技术支撑与实际的工程应用价值。

关键词

目标识别, YOLOv5, 高速立柱, RepVGG, CoTAttention, BiFPN

RCB-YOLOv5s: Detection of Highway Guardrail Installation Poles Based on Improved YOLOv5s Algorithm

Zhe Tian, Jiuhong Ruan, Guodong Li*

School of Rail Transportation, Shandong Jiaotong University, Jinan Shandong

Received: July 18, 2026; accepted: August 19, 2026; published: August 27, 2026

*通讯作者。

文章引用: 田哲, 阮久宏, 李国栋. RCB-YOLOv5s: 基于改进 YOLOv5s 算法的高速护栏安装立柱检测[J]. 计算机科学与应用, 2026, 16(8): 291-302. DOI: 10.12677/csa.2026.168281

Abstract

As the mileage of highways expands rapidly, the automated detection of highway poles has become a crucial task for the safety inspection and maintenance of intelligent transportation infrastructure. However, complex highway environments characterized by strong background interference and significant target scale variations severely limit the accuracy of conventional detection algorithms. To address these critical engineering challenges, this paper proposes an improved target detection algorithm based on YOLOv5, which is of great significance for achieving reliable and real-time highway infrastructure monitoring. First, the RepVGG block is integrated into the model architecture to enhance the extraction of edge and texture features without compromising inference speed through structural re-parameterization. Subsequently, the Contextual Transformer (CoTAttention) mechanism is embedded to suppress background noise, and a Bidirectional Feature Pyramid Network (BiFPN) is employed to reconstruct the network's neck, thereby strengthening multi-scale feature fusion for distant small-scale targets. Extensive experiments conducted on a self-constructed highway pole dataset demonstrate the superiority of the proposed method. Compared with the original YOLOv5 model, the improved algorithm achieves a 4.7% increase in mean average precision (mAP@0.5), with precision and recall metrics reaching 88.4% and 85.1%, respectively. Consequently, this research successfully balances high detection accuracy with real-time efficiency, providing robust technical support and practical engineering value for intelligent inspection and safety early warning systems in highway infrastructure.

Keywords

Object Detection, YOLOv5, Highway Pole, RepVGG, CoTAttention, BiFPN

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

随着我国高速公路里程的不断延伸以及智能交通系统(ITS)的快速发展,交通基础设施的安全巡检与日常维护变得日益关键[1]。作为监控摄像头、指示牌和护栏的重要承载组件,高速公路沿线的护栏立柱在保障路网安全运行方面发挥着至关重要的作用[2]。然而,传统的依靠人工的巡检方式不仅效率低下、劳动强度大,而且极易受到主观人为因素和恶劣天气条件的影响。因此,向基于视觉的自动化巡检系统转型是一项迫切的工程需求[3]。尽管基于深度学习的目标检测技术(特别是 YOLO 系列)在工业部署中已得到广泛应用,但在真实场景中实现高速公路立柱的高精度检测依然面临着严重的瓶颈[4]。这一工程问题的核心制约因素主要体现在三个方面:一是杂乱的户外背景(如植被和行驶中的车辆)以及光照变化引起的严重特征干扰;二是目标尺度变化剧烈导致远处小尺度立柱存在极高的漏检率;三是资源受限的车载边缘设备对实时处理有着极其严格的计算约束[5] [6]。

为了应对上述挑战,大量研究探索了交通场景下基于视觉的目标检测技术;然而,在应对高速公路沿线护栏立柱检测的特定复杂环境时,现有方法暴露出显著的局限性[7]。首先,为了追求边缘端的实时部署,部分研究人员采用了极度轻量化的主干网络(如 MobileNet 或 ShuffleNet) [8]。虽然这些方法提升了推理速度,但却严重削弱了网络提取立柱细微边缘和纹理特征的能力,导致定位精度不佳[9]。其次,在应对复杂背景干扰方面,尽管经常引入传统的空间或通道注意力机制来抑制噪声[10],但它们通常无法同

时挖掘局部与全局的上下文依赖关系, 导致在高度杂乱的道路场景中出现极高的误检率。最后, 为了应对目标尺度的变化, 标准的特征金字塔网络(FPN)被广泛采用。然而, 这些传统结构在融合过程中对所有多尺度特征图一视同仁, 这不可避免地导致特征在深层传播过程中发生严重的信息衰减, 使其在检测远距离小尺度立柱时基本失效。因此, 现有算法无法同时克服背景噪声抑制、远处小目标定位以及边缘端实时推理这三重困境。

为了从根本上解决这些工程瓶颈, 本文提出了一种改进的、基于 YOLOv5 的高速公路护栏立柱高精度检测算法。主要目标是构建一个鲁棒的检测框架, 在卓越的检测精度与实时的计算效率之间取得平衡。具体而言, 本文的主要工作与贡献如下:

(1) 基于 RepVGG 的主干网络重构: 利用 RepVGG 模块对特征提取网络进行了重构。通过引入结构重参数化技术, 该模型在训练阶段保留了多分支架构卓越的特征提取能力, 而在推理阶段则无缝折叠为高效的单路径结构, 从而在不增加推理延迟的前提下强化了立柱轮廓的提取能力。

(2) 嵌入 CoTAttention 机制: 集成了上下文 Transformer (CoTAttention) 模块以深入挖掘全面的局部与全局特征依赖关系。该机制能够精准地引导网络将注意力集中在立柱目标上, 有效抑制由复杂背景噪声引发的误检。

(3) 基于 BiFPN 的颈部(Neck)优化: 采用双向特征金字塔网络(BiFPN)来优化多尺度特征融合过程。通过高效的跨尺度连接和自适应的加权融合策略, 该结构缓解了深层特征衰减, 并显著增强了对远处小尺度立柱的感知能力。

在自建的高速公路立柱数据集上进行的大量实验结果表明, 所提算法的平均精度均值(mAP@0.5)达到了 88.4%, 较基线 YOLOv5 模型提升了 4.7%, 同时精确率(Precision)和召回率(Recall)分别达到了 88.4% 和 85.1%。综上所述, 本研究为高速公路基础设施中智能巡检与安全预警系统的部署提供了高度鲁棒且极具实用价值的算法基础。

2. 改进 YOLOv5 算法

2.1. 引入 RepVGG 模块

在高速公路场景中, 将目标检测模型部署于终端设备时, 通常面临着如何平衡强大的特征提取能力与实时推理速度的挑战。为了解决这一固有的权衡问题, 本文在 YOLOv5 架构中引入了 RepVGG 模块, 取代了原有的传统卷积块。RepVGG 的核心原理是“结构重参数化(Structural Re-parameterization)”, 这是一种策略性地将模型的训练阶段与推理阶段解耦的技术。这种设计范式明确考虑了资源可用性的差异: 模型在训练期间通常可以获得丰富的计算资源, 而部署在实际现场的终端设备则在内存容量和计算能力方面受到严格的限制。

在训练阶段, RepVGG 借鉴了 ResNet 的残差网络拓扑结构, 构建了多分支架构以促进更优的特征表示。具体而言, 在保留流线型 VGG 风格主干网络的同时, 该网络融合了一个复合结构, 其中包括一个 3×3 卷积分支、一个 1×1 卷积残差分支以及一个恒等映射分支。这种多分支配置融合了 VGG 和 ResNet 两者的表征优势, 从而使模型能够更充分地挖掘并整合特征图中的深层语义信息。针对高速公路立柱检测任务, 这种增强的特征提取能力使得网络能够在复杂杂乱的道路背景下, 精确勾勒出不同空间尺度下立柱的边缘轮廓和纹理细节。网络的主干正是基于这些残差结构构建的, 见图 1 所示。

在训练完成并向实际部署过渡的过程中, RepVGG 执行结构重参数化过程, 将复杂的训练时模型等效转换为精简的推理时模型。在该转换过程中, 原有的搭配批归一化(BN)层的 3×3 卷积层、搭配 BN 层的 1×1 卷积层以及独立的 BN 分支, 被分别独立转化为新的 3×3 卷积层。随后, 这三个新生成的卷积

核被相加合并为一个单一且统一的 3×3 卷积层。经过这一重参数化步骤后,用于实际推理的网络架构便折叠为一条完全由堆叠的 3×3 卷积层和 ReLU 激活函数组成的单路径结构。这种极简的布局不仅表现出高度的计算并行性和适度的内存占用,而且在保持检测精度的同时显著加快了前向推理速度。因此,该设计充分满足了实际工程部署中对严格实时响应的要求。

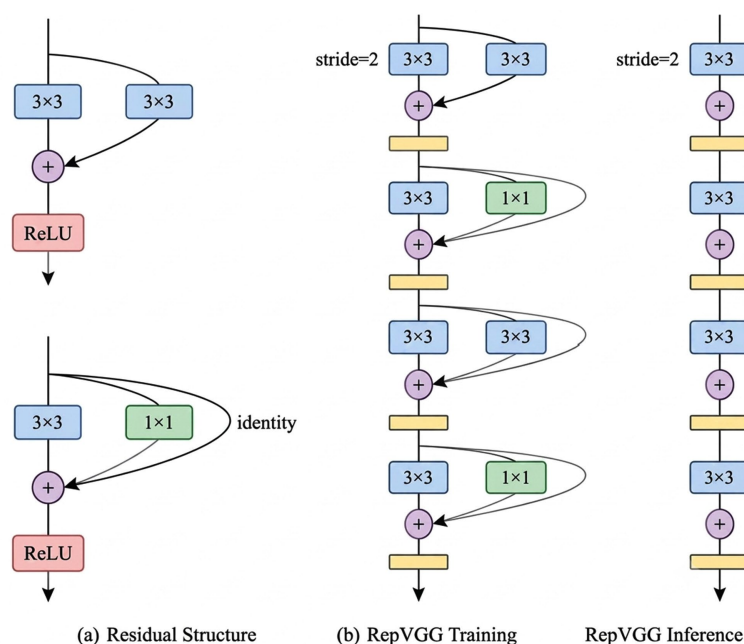


Figure 1. Structure diagram of RepVGG block

图 1. RepVGG 模块结构图

2.2. CoTAttention 机制

在复杂的高速公路监控环境中,立柱目标检测所处的背景往往杂乱且极其复杂。此外,大跨度距离和目标遮挡等挑战通常会导致立柱边界模糊,从而降低了网络提取特征的有效分辨率。为了有效抑制冗余背景特征的干扰并进而提高检测精度,本文在 YOLOv5 架构中引入了 CoTAttention(上下文 Transformer 注意力)模块。该模块在一个统一的架构框架内集成了上下文信息挖掘与自注意力学习。通过全面挖掘显著特征的上下文信息以对特征间的依赖关系进行建模,CoTAttention 促进了整个输入图像上的自适应注意力分配,从而显著增强了特征图的视觉表征能力。CoTAttention 模块的结构示意图见图 2 所示。

在具体的特征处理流程中,设输入特征图为 X ,其维度包括高度 H 、宽度 W 和通道数 C 。首先,CoTAttention 对输入特征执行 $k \times k$ 分组卷积操作,以获得键(Key)变量 K 。这一步骤赋予了 K 编码局部上下文信息的表征能力。随后,通过特征拼接操作,将查询(Query)变量 Q 与富含局部上下文信息的 K 进行拼接。然后,将拼接后得到的张量依次通过两个 1×1 卷积层,以计算注意力矩阵 A 。

与仅对查询 Q 和键 K 之间的关系进行建模的传统自注意力机制不同,CoTAttention 中的注意力矩阵 A 是通过查询信息与局部上下文表征之间的交互生成的。换言之,该机制通过局部上下文建模的引导,有效增强了自注意力过程的表达能力。此后,将计算得到的注意力矩阵 A 与值(Value)变量 V 相乘,生成动态上下文建模的特征图。最后,CoTAttention 模块将静态建模的局部上下文特征与动态建模的全局上下文特征进行融合,作为最终输出。这种架构设计有效地整合了多维特征数据,加深了模型对立柱特征的整体理解,从而大幅减少了在复杂环境中对小尺度或边界模糊目标的漏检情况。

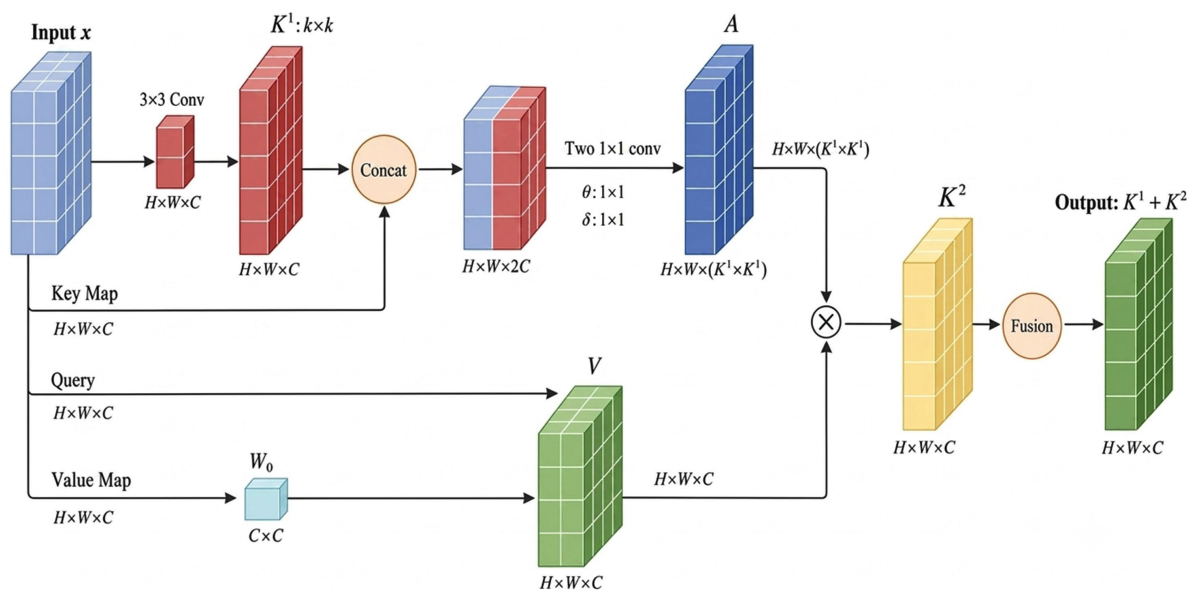


Figure 2. Structural diagram of the CoTAttention module

图 2. CoTAttention 模块的结构示意图

2.3. BiFPN 模型结构

在高速公路监控场景中，由于距离摄像头的远近不同，立柱目标表现出极大的空间尺度跨度，而道路背景固有的复杂性进一步加剧了对目标特征提取的干扰。YOLOv5 原始的颈部(Neck)组件采用了特征金字塔网络(FPN)与路径聚合网络(PANet)的组合来促进多尺度特征融合。尽管这种包含自顶向下和自底向上两种传播方式的双向路径增强了特征传输，但它隐含地假设所有输入特征图在拼接过程中的贡献是相等的。因此，这种设计不仅引入了一定程度的计算冗余，而且未能突出最显著的特征。为了解决上述局限性，本文引入了双向特征金字塔网络(BiFPN)来重构模型的 Neck 特征融合层。各种主流 FPN 结构的架构配置见图 3 所示。

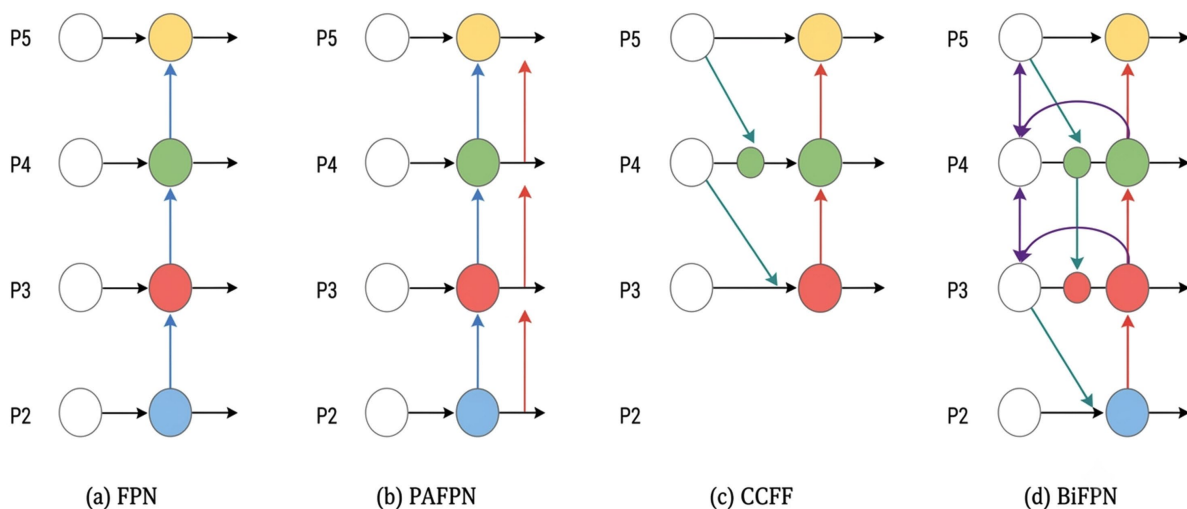


Figure 3. Architectural configurations of various prevailing FPN structure

图 3. 各种主流 FPN 结构的结构图

与传统的 FPN 和 PANet 架构相比, BiFPN 在网络拓扑结构和融合策略两方面都进行了关键的优化。首先, 在拓扑结构方面, BiFPN 移除了仅有单一输入边且不参与特征融合的中间节点。这种剪枝操作在保持表征能力的同时, 减少了多余的参数量和计算开销。其次, 在位于同一层级的原始输入节点和输出节点之间引入了一条额外的跳跃连接(skip connection)。这种高效的双向跨尺度连接机制, 不仅加速了自顶向下深层语义信息与自底向上浅层空间定位线索的双流传播, 而且有效缓解了细粒度特征(例如位于远距离的小尺度立柱边界)在深层网络中的衰减。

BiFPN 的核心原理是通过加权求和的方式进行特征融合, 并辅以双向融合路径。具体而言, 其融合方法涉及加权特征融合, 即将不同分辨率的特征图与可学习的权重进行整合, 而不是进行简单的直接相加:

$$P_i = \frac{\sum_j w_j \cdot \text{UpSample}(P_j)}{\sum_j w_j + \varepsilon} \varepsilon \quad (1)$$

或者

$$P_i = \frac{\sum_j w_j \cdot \text{UpSample}(P_j)}{\sum_j w_j + \varepsilon} \varepsilon \quad (2)$$

其中 w_j 是每个特征图的权重系数, ε 是一个小常数(以防止被零除以)。通过这种方式, 每层的特征根据其重量被融合。与 PAN 类似, BiFPN 也设有自上而下和自下而上的通路, 在每一层, 特征通过 hted 方法与相邻层融合:

$$P_i^{td} = \frac{w_i^{td} \cdot (P_{i+1} + \text{UpSample}(P_{i+1}))}{w_i^{td} + 1} \quad (3)$$

或者

$$P_i^{bu} = \frac{w_i^{bu} \cdot (P_{i-1} + \text{DownSample}(P_{i-1}))}{w_i^{bu} + 1} \quad (4)$$

通过自上而下和自下向上路径的加权融合, BiFPN 能够更灵活地结合不同分辨率的特征, 增强了融合效果。

2.4. 算法改进的协同作用与理论分析

在设计 RCB-YOLOv5s 算法时, 本文并非将各个改进模块进行简单的线性堆砌, 而是针对高速公路立柱检测的特定物理环境与特征传递逻辑, 构建了深度互补的特征处理闭环, RCB-YOLOv5s 算法结构见图 4 所示。

首先, 在特征提取阶段, 鉴于高速公路环境的复杂性, 立柱的物理轮廓极易与杂乱的背景植被或车辆相混淆。本文引入的 RepVGG 模块凭借其在训练阶段的多分支残差拓扑结构, 在网络浅层和中层有效捕捉并强化了立柱精细的边缘和纹理特征。然而, 这种细粒度特征的增强不可避免地也会放大背景噪声的干扰。

为此, 本算法在特征提取的关键链路中无缝嵌入了 CoTAttention 机制。该机制发挥了至关重要的上下文感知作用, 它接收来自 RepVGG 提取的富含细节的特征图, 通过计算局部与全局的注意力矩阵, 精准地抑制了树木遮挡、光照突变等背景噪声, 使得模型的计算资源高度聚焦于立柱目标本身。

经过 CoTAttention 提纯后, 网络获得了高质量、低噪声的增强语义特征。但由于立柱目标存在巨大

的空间尺度跨度，不同尺度特征图的重要程度在深层传播中存在显著差异。传统的 PANet 等权重融合策略容易导致远处小尺度立柱的定位线索被淹没。此时，BiFPN 结构的引入补齐了最后一块短板。BiFPN 的自适应加权融合策略能够动态评估这些高质量特征的贡献度，结合高效的双向跨尺度连接，完美契合了处理增强特征的需求。这确保了即便是在远距离的小尺度立柱特征，也能在深层网络中被完整保留。

综上所述，从 RepVGG 的“细粒度特征提取”，到 CoTAttention 的“上下文噪声抑制”，再到 BiFPN 的“自适应多尺度融合”，三者相互反哺，从理论底层保障了改进算法在复杂非结构化道路上的高精度与强鲁棒性。

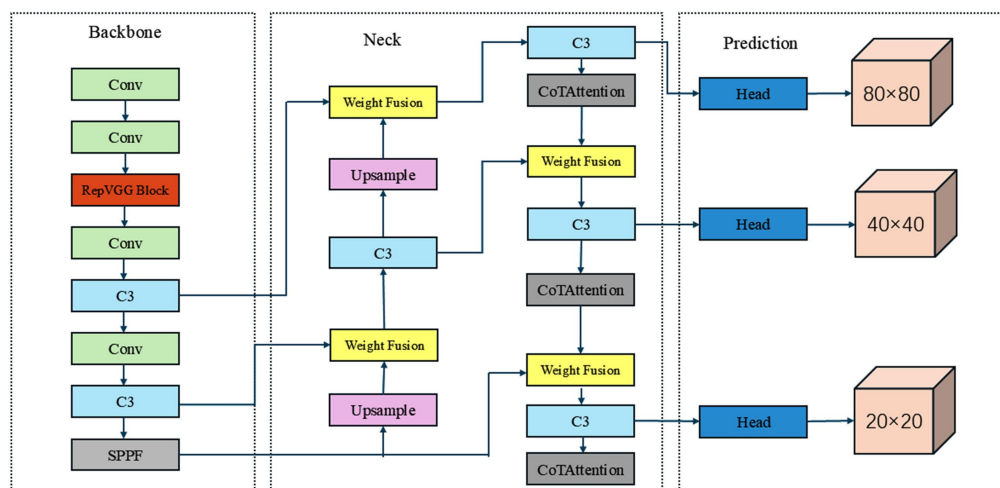


Figure 4. Structure of the RCB-YOLOv5s algorithm

图 4. RCB-YOLOv5s 算法结构

3. 高速公路立柱检测实验

3.1. 实验环境与参数

本实验环境配置在一台运行 Windows 10 操作系统的高性能计算机上，该计算机配备了 NVIDIA GeForce RTX 4060 GPU (10 GB 显存)、主频为 2.50 GHz 的第 12 代 Intel(R) Core(TM) i5-12600 处理器，以及 8 GB 内存。采用的深度学习框架为 PyTorch 1.13.1，并结合 CUDA 11.7 并行计算平台与 Python 3.7 协同使用。在训练参数配置方面，模型共训练了 300 个轮次(epochs)，批处理大小(batch size)设定为 16。初始学习率设定为 0.001，动量参数(momentum)设定为 0.937，权重衰减系数(weight decay)设定为 0.0005。此外，本研究采用 Mosaic 数据增强技术作为图像增强策略。

3.2. 实验数据集



Figure 5. Representative images of highway poles under diverse weather and illumination conditions

图 5. 不同天气与光照条件下的高速公路立柱代表性图像

实验所用数据集为本研究课题组采集的 1000 张国内高速公路立柱图片, 包括正常天气下立柱、雨天立柱、雪天立柱、夜晚立柱和阴天立柱 5 类, 高速公路立柱图片见图 5 所示。将所有照片分为训练、验证和测试数据集, 各数据集图片数之比为 8:1:1。

3.3. 评价指标

目标立柱检测试验的评价指标为检测精度和平均检测时间, 检测精度指标包括准确率 p 、召回率 r 及平均检测精度均值 E_{mAP} 。

$$\text{准确率 } p = \frac{N_{TP}}{N_{TP} + N_{FP}} \times 100\% \quad (5)$$

式中: N_{TP} 为算法模型成功识别图像特征并确认目标立柱处于正确状态的样本数, N_{FP} 为算法模型未能检测到目标立柱或在检测过程中未准确识别目标立柱的样本数。

$$\text{召回率 } r = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100\% \quad (6)$$

式中 N_{FN} 为算法模型错误识别目标立柱的样本数。

$$\text{目标立柱模糊的平均检测精度 } E_{AP_i} = \int_0^1 p(r) dr \quad (7)$$

$$\text{平均检测精度均值 } E_{mAP} = \sum_{i=0}^N \frac{E_{AP_i}}{N} \quad (8)$$

式中 N 为目标立柱模糊总数。

3.4. 消融实验

为验证改进后算法的检测性能, 探究 RepVGG Block 模块、CoTAttention 注意力机制和 BiFPN 模块对 YOLOv5 算法的改进效果, 设计 8 组消融实验方案, 选取召回率 r 、平均检测精度均值 e 和平均检测时间 t 作为评价 8 组消融实验方案的指标, 消融实验方案及结果见表 1 所示。

Table 1. Ablation experimental configurations and corresponding outcomes

表 1. 消融实验配置及相应结果

算法模型	YOLOv5	RepVGG	CoTAttention	BiFPN	$r/\%$	$E_{mAP}/\%$	t/ms
1	√	×	×	×	80.8	83.7	6.8
2	√	√	×	×	83.1	87.5	6.2
3	√	×	√	×	82.8	85.3	8.1
4	√	×	×	√	83.5	85.0	8.4
5	√	√	√	×	82.2	87.7	7.4
6	√	√	×	√	81.9	86.5	8.7
7	√	×	√	√	82.0	87.2	8.6
8	√	√	√	√	85.1	88.4	7.1

注: “√”表示包含该模块; “×”表示未包含该模块。

从表 1 可以看出, 在 YOLOv5 算法中单独引入 CoTAttention 机制或 BiFPN 结构, 均能提升召回率 (R) 和平均精度均值 (mAP); 然而, 这些性能增益也伴随着平均检测延迟的增加。此外, 同时集成

CoTAttention 机制和 BiFPN 结构进一步提升了检测精度,但代价是推理时间的进一步延长。相比之下,将 RepVGG 模块引入 YOLOv5 框架,不仅有效提高了召回率和精度,还显著降低了平均检测延迟。

具体而言,与算法模型 1 (YOLOv5 算法)相比,改进的 YOLOv5 算法召回率提高了 4.3%,平均精度均值(mAP)提高了 4.7%,而平均检测延迟仅增加了 0.3 毫秒。值得注意的是,召回率和精度的提升主要归功于 CoTAttention 机制和 BiFPN 结构的采用,而 RepVGG 模块则有效缓解了通常由注意力机制和特征金字塔网络引发的延迟增加。

因此,针对高速公路护栏立柱检测任务,改进的 YOLOv5 算法不仅在检测精度上实现了显著提升。

3.5. 算法改进分析

两种算法的平均精度均值(mAP@0.5)对比见图 6 所示。两种算法的 mAP 曲线在迭代 200 次后均逐渐收敛。RCB-YOLOv5s 算法的 mAP@0.5 达到了 88.4%,相比于 YOLOv5s 基线模型达到的 83.7%,提升了 4.7%。这一提升表明,多尺度特征融合与注意力机制的集成,使得改进算法能够在复杂的道路条件和不同的尺度下,更加准确地定位护栏立柱。

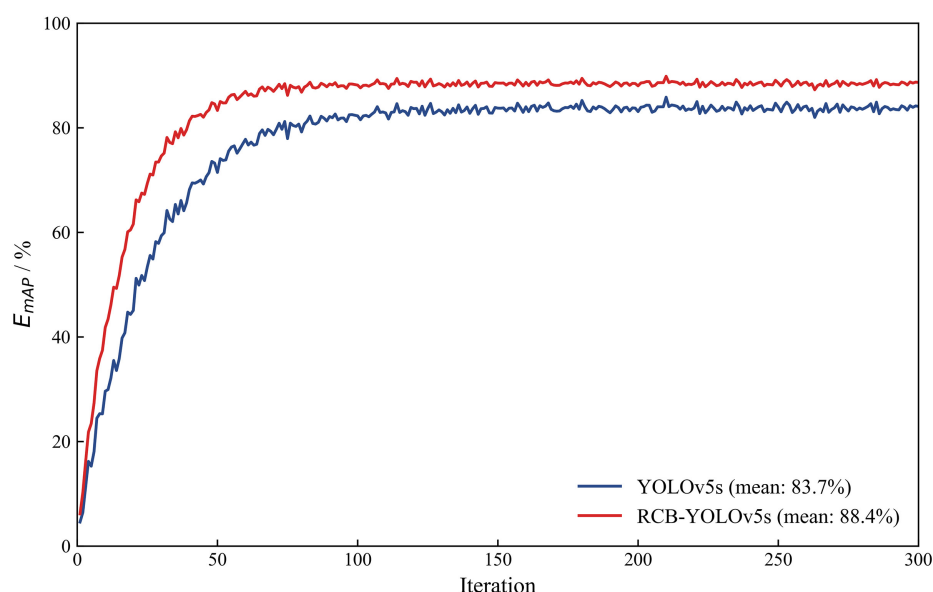


Figure 6. Comparison of mean Average Precision (mAP@0.5) between the baseline YOLOv5s and the proposed RCB-YOLOv5s

图 6. YOLOv5s 与本文所提 RCB-YOLOv5s 的平均精度均值(mAP@0.5)对比

召回率对比见图 7 所示。如图所示,RCB-YOLOv5s 算法的召回率在训练初期迅速上升,很快超越了 YOLOv5s 算法,并在 200 次迭代后逐渐收敛。RCB-YOLOv5s 算法的召回率达到了 85.1%,较基线模型的 80.8%提升了 4.3%。这表明,改进后的算法能够更有效地提取目标特征,克服了树木遮挡和光照突变等干扰,从而显著降低了远距离小尺度立柱的漏检率。

单帧平均检测延迟的对比如下。尽管由于引入了额外模块导致 RCB-YOLOv5s 模型的网络深度有所增加,但其平均检测延迟仅从 6.8 ms 轻微增加至 7.1 ms (仅延长了 0.3 ms),这主要归功于 RepVGG 模块的结构重参数化能力。因此,本文所提算法在显著提升检测精度和召回率的同时,依然保持了极佳的推理速度,完全满足护栏安装机器人视觉导航的实时性要求。

定性检测结果对比见图 8 和图 9 所示。如图所示,RCB-YOLOv5s 算法能够在不同的距离跨度、多样

的光照条件以及杂乱的背景下精准地检测出护栏立柱。其检测置信度明显高于 YOLOv5s 算法, 展现出更优异的环境适应性和鲁棒性。

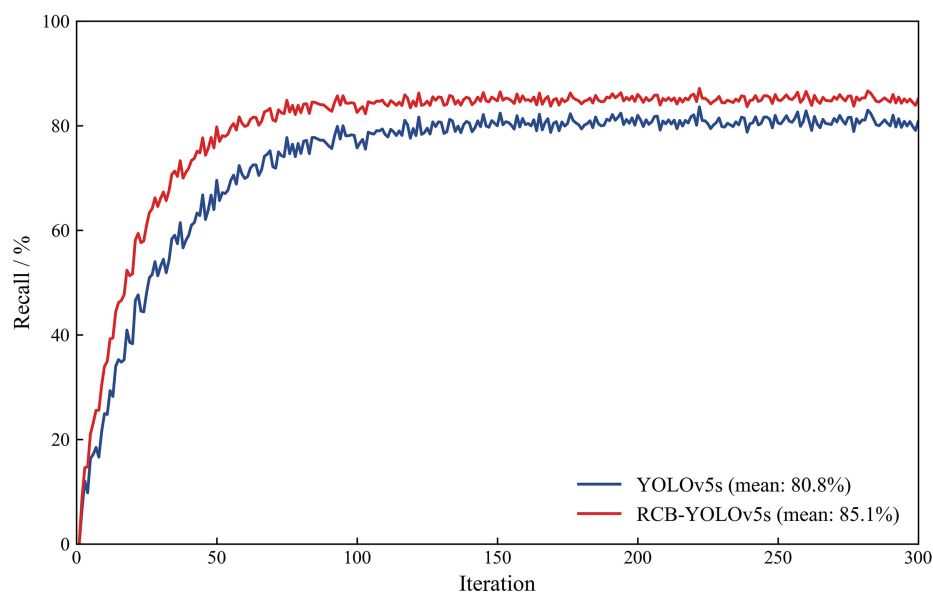


Figure 7. Comparison of recall curves between the baseline YOLOv5s and the proposed RCB-YOLOv5s
图 7. YOLOv5s 与 RCB-YOLOv5s 的召回率曲线对比



Figure 8. Qualitative detection results of the baseline YOLOv5s algorithm
图 8. YOLOv5s 算法检测结果



Figure 9. Qualitative detection results of the proposed RCB-YOLOv5s algorithm
图 9. RCB-YOLOv5s 算法检测结果

3.6. 不同算法对比分析

Table 2. Comparative detection results of different algorithms

表 2. 不同算法的对比检测结果

算法模型	$r/\%$	$E_{mAP}/\%$	t/ms
ssd	77.5	71.2	33.5
YOLOv3	78.4	77.4	25.8
YOLOv4	78.6	77.1	17.6
YOLOv5	80.8	83.7	6.8
RCB-YOLOv5s	85.1	88.4	7.1

为了进一步验证改进的 YOLOv5 算法在高速公路护栏立柱检测任务中的适用性,本研究在相同的实验平台和数据集下进行了对比实验。参与评估的算法包括改进的 YOLOv5 算法、SSD、YOLOv3、YOLOv4 以及基线 YOLOv5。对比检测结果汇总见表 2。

从表 2 可以看出,改进的 YOLOv5 算法的召回率(R)分别比 SSD、YOLOv3、YOLOv4 和基线 YOLOv5 高出 7.6%、6.7%、6.5%和 4.3%。相应地,平均精度均值(E_{mAP})分别提升了 17.2%、11.0%、11.3%和 4.7%。在推理效率方面,改进的 YOLOv5 算法的平均检测延迟相比于 SSD、YOLOv3 和 YOLOv4 分别减少了 26.4 ms、18.7 ms 和 10.5 ms;然而,相对于基线 YOLOv5,仅延长了 0.3 ms。

改进的 YOLOv5 算法在 YOLOv5 基线的基础上引入了 CoTAttention 机制和 BiFPN 结构。这些改进增强了模型的特征提取与多尺度融合能力,从而大幅提升了高速公路立柱的检测精度。然而,它们同时也增加了模型的计算复杂度,导致推理延迟延长。反之,同时集成的 RepVGG 模块有效地提升了网络的推理效率,抵消了额外的延迟开销。这些改进的综合作用使得平均检测延迟相比原始 YOLOv5 算法仅有 0.3 ms 的微小增加。因此,综合考虑高速公路护栏立柱检测任务的实时性要求,以及召回率、检测精度和推理延迟等指标,改进的 YOLOv5 算法展现出了显著的综合优势。

4. 结论

针对高速公路护栏安装机器人视觉导航中,由于树木遮挡、光照剧变等复杂户外环境导致的立柱目标检测精度低、远距离小目标易漏检等难题,本文在 YOLOv5s 框架中引入了 RepVGG 重参数化模块、CoTAttention 机制和 BiFPN 特征金字塔结构,由此构建了 RCB-YOLOv5s 算法。改进后的算法显著增强了模型在复杂背景下的特征提取与多尺度护栏立柱定位能力,从而满足了移动底盘在非结构化道路上的实时导航与精准对接需求,为高速公路护栏的自动化施工提供了可靠的视觉技术支撑。基于综合的实验分析,得出以下主要结论:

1) 消融实验结果表明,RCB-YOLOv5s 算法中融入的 CoTAttention 机制与 BiFPN 模块,有效增强了网络的全局上下文感知与跨层多尺度特征融合能力,从而大幅提升了召回率和平均精度均值(E_{mAP})。同时,借助结构重参数化技术,RepVGG 模块的引入不仅提高了边缘特征提取的效率,还有效抵消了网络结构加深对平均检测延迟产生的不利影响。

2) 多算法对比实验结果表明,在复杂的高速公路场景下,RCB-YOLOv5s 算法在召回率、平均精度均值(E_{mAP})和检测速度等综合评价指标上,均显著优于 Faster R-CNN、SSD 以及 YOLOv3、YOLOv4 等主流目标检测算法,展现出更优越的环境适应性与鲁棒性。

3) 综合消融实验与对比实验的结果,与原始 YOLOv5s 基线模型相比,RCB-YOLOv5s 算法的召回

率提升了 4.3% (由 80.8% 提升至 85.1%), 平均精度均值(E_{mAP})提升了 4.7% (由 83.7% 提升至 88.4%)。值得注意的是, 其单帧平均检测延迟仅微增了 0.3 ms。因此, 综合考虑高速公路护栏立柱检测任务对精度与实时性的严格要求, 部署本文所提的 RCB-YOLOv5s 算法进行视觉特征提取与导航定位, 在实际工程应用中具有显著的优势。

参考文献

- [1] Veres, M. and Moussa, M. (2019) Deep Learning for Intelligent Transportation Systems: A Survey of Emerging Trends. *IEEE Transactions on Intelligent Transportation Systems*, **21**, 3152-3168. <https://doi.org/10.1109/tits.2019.2929020>
- [2] Spencer, B.F., Hoskere, V. and Narazaki, Y. (2019) Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. *Engineering*, **5**, 199-222. <https://doi.org/10.1016/j.eng.2018.11.030>
- [3] Koch, C., Georgieva, K., Kasireddy, V., Akinci, B. and Fieguth, P. (2015) A Review on Computer Vision Based Defect Detection and Condition Assessment of Concrete and Asphalt Civil Infrastructure. *Advanced Engineering Informatics*, **29**, 196-210. <https://doi.org/10.1016/j.aei.2015.01.008>
- [4] 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>
- [5] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv:2004.10934.
- [6] Zou, Z., Chen, K., Shi, Z., Guo, Y. and Ye, J. (2023) Object Detection in 20 Years: A Survey. *Proceedings of the IEEE*, **111**, 257-276. <https://doi.org/10.1109/jproc.2023.3238524>
- [7] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., *et al.* (2017) MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv:1704.04861.
- [8] Ma, N., Zhang, X., Zheng, H. and Sun, J. (2018) ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 122-138. https://doi.org/10.1007/978-3-030-01264-9_8
- [9] Deng, B.L., Li, G., Han, S., Shi, L. and Xie, Y. (2020) Model Compression and Hardware Acceleration for Neural Networks: A Comprehensive Survey. *Proceedings of the IEEE*, **108**, 485-532. <https://doi.org/10.1109/jproc.2020.2976475>
- [10] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>