

基于改进YOLOv5的遥感影像小目标识别研究

寇卓, 王瑜璇, 田颜俊, 麻胤翔, 谭峻仁

沈阳理工大学信息科学与工程学院, 辽宁 沈阳

收稿日期: 2026年5月10日; 录用日期: 2026年6月15日; 发布日期: 2026年6月24日

摘要

由于小目标通常具有分辨率低、特征信息不足以及背景干扰严重等特点, 传统目标检测方法在该任务上的性能往往受限。针对上述问题, 本文提出了一种基于改进YOLOv5的遥感影像小目标智能检测方法, 旨在提升检测精度与召回率。本文的主要改进体现在三个方面。首先, 设计了空间金字塔空洞卷积模块 (SPD-Conv), 通过多尺度空洞卷积有效扩大感受野, 同时保持特征图的高分辨率, 从而减少下采样过程中小目标细节信息的丢失。其次, 引入轻量级高效通道注意力机制 (ECA-Net), 以自适应方式增强小目标相关特征表达, 并抑制背景噪声干扰。最后, 采用双向特征金字塔网络 (BiFPN) 替代传统路径聚合网络, 实现多尺度特征的双向高效融合, 显著提升深层特征的语义表达能力。在公开遥感数据集 AI-TOD 上的实验结果表明, 所提方法在小目标检测任务中表现优异, 平均精度均值 (mAP) 达到 36.8%, 较基线模型 YOLOv5 提升了 7.3 个百分点。实验结果充分验证了所提方法在复杂遥感场景中进行小目标检测的有效性与优越性。

关键词

小目标检测, YOLOv5, 遥感影像, 空间金字塔卷积, 注意力机制, 特征金字塔网络

Research on Small Target Recognition in Remote Sensing Imagery Based on Improved YOLOv5

Zhuo Kou, Yuxuan Wang, Yanjun Tian, Yinxiang Ma, Junren Tan

School of Information Science and Engineering, Shenyang University of Technology, Shenyang Liaoning

Received: May 10, 2026; accepted: June 15, 2026; published: June 24, 2026

Abstract

Small object detection has always been the core challenge in computer vision in remote sensing

文章引用: 寇卓, 王瑜璇, 田颜俊, 麻胤翔, 谭峻仁. 基于改进 YOLOv5 的遥感影像小目标识别研究[J]. 计算机科学与应用, 2026, 16(6): 129-139. DOI: 10.12677/csa.2026.166214

images. Due to the characteristics of small object detection, such as low resolution, scarcity of feature information and serious background interference, the traditional object detection algorithm performs poorly in this task. This paper proposes an intelligent detection system for remote sensing images of small objects based on improved YOLOv5, which aims to improve the detection accuracy and recall rate of small targets. The main innovations of this study include three aspects: First, a spatial pyramid depth convolution module (SPD-Conv) is designed, which effectively expands the sensing field through a multi-scale hollow convolution structure, while maintaining the high resolution of the feature map to avoid losing detailed information of small objects during the downsampling process. Secondly, a lightweight channel attention mechanism (ECA-Net) is introduced to adaptively enhance the expression of characteristics related to small objects and suppress interference from irrelevant background information. Third, the two-way feature pyramid network (BiFPN) is used to replace the traditional path aggregation network, which realizes the two-way integration of multi-scale features and greatly improves the semantic representation ability of small objects in deep networks. Experimental results on the public remote sensing data set AI-TOD show that this method performs well in small object detection tasks, with a detection accuracy rate (mAP) of 36.8%, 7.3 percentage points higher than the original YOLOv5, verifying the effectiveness of the proposed improvement strategy.

Keywords

Small Object Detection, YOLOv5, Remote Sensing Images, Spatial Pyramid Convolution, Attention Mechanism, Feature Pyramid Network

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

遥感影像中车辆、船舶等小目标的高精度识别对于资源监测与应急救援具有重要意义。然而，受复杂成像背景、目标尺度极小及方向多变等因素影响，传统目标检测方法在多次下采样过程中易造成小目标特征信息严重衰减，难以满足高精度解译需求。

遥感影像中的小目标检测一直是计算机视觉领域的关键挑战[1]。近年来，以 YOLO 系列为代表的单阶段检测器在检测效率与精度之间取得了良好平衡。基于此，本文以 YOLOv5 为基准模型，针对遥感小目标特征表达不足的问题，提出了一种改进的智能检测方法[2]。首先，构建空间金字塔空洞卷积模块(SPD-Conv)，利用多分支空洞卷积替代传统下采样操作，在保持特征图高分辨率的同时有效扩大感受野；其次，引入轻量级高效通道注意力机制(ECA-Net)，自适应强化小目标特征响应并抑制背景干扰；最后，采用双向特征金字塔网络(BiFPN)优化多尺度特征融合路径，提升深层语义信息的传递效率与表达能力。

本文基于公开遥感数据集 AI-TOD 开展实验研究，并通过消融实验系统评估各改进模块的有效性。实验结果表明，所提方法在保证检测速度的前提下，显著提升了复杂场景及极端尺度变化条件下的小目标检测精度，为遥感影像目标检测提供了一种高效且鲁棒的技术方案。

2. 相关工作

2.1. 遥感影像目标检测方法研究进展

遥感影像目标检测是智能解译领域的重要研究方向，在城市管理、交通监测、海域巡检及灾害评估

等场景中具有广泛应用价值。其中,车辆、船舶、飞机等小目标检测任务由于目标尺度小、像素占比低、背景纹理复杂及目标密集分布等特点,一直是当前研究的重点与难点[2]。传统基于人工特征的方法依赖边缘、纹理及形状描述符,在复杂场景下鲁棒性有限,难以满足高精度检测需求[3]。

近年来,深度学习方法推动了遥感目标检测技术快速发展。两阶段检测算法以 Faster R-CNN 为代表,具有较高检测精度,但推理速度较慢,难以适应大幅面遥感图像的实时处理需求[4]。一阶段检测算法如 SSD、RetinaNet 及 YOLO 系列,在检测速度与精度之间取得了较好平衡,逐渐成为遥感场景中的主流方案。其中, YOLOv5 凭借结构轻量、部署灵活和训练稳定等优势,被广泛应用于遥感目标检测任务。

方面取得了积极进展,但仍普遍存在模型复杂度增加、实时性下降,以及对极小目标和遮挡目标检测效果有限等问题。因此,如何在保证推理效率的前提下,进一步提升模型对复杂遥感场景中小目标的识别能力,仍是当前研究的重要方向。

基于此,本文以然而,通用目标检测模型在遥感小目标场景中仍存在明显不足:其一,多次下采样过程容易导致小目标细节特征丢失;其二,复杂地物背景易引起误检与漏检;其三,目标尺度变化大且分布密集,传统特征融合结构对多尺度信息利用不足。针对上述问题,研究者主要从高分辨率特征保持、注意力机制增强及多尺度特征融合等方向展开改进。例如,部分研究通过引入空洞卷积、可变形卷积或浅层特征增强模块,在扩大感受野的同时保留更多空间细节信息;部分研究利用 SE、CBAM、ECA 等注意力机制提升网络对关键目标区域的响应能力,抑制复杂背景干扰;还有研究采用 FPN、PANet、BiFPN 等多尺度融合结构,加强不同层级特征之间的信息交互,以提升对密集小目标和尺度变化目标的检测性能[5]。随着 Transformer 结构的发展,一些方法开始尝试将全局建模能力引入遥感检测任务,以增强模型对远距离依赖关系的表达能力。

总体来看,现有研究在提升遥感小目标检测精度 YOLOv5 为基础框架,围绕“小目标特征易丢失、背景噪声干扰强、多尺度信息融合不足”等关键问题展开研究。通过引入 SPD-Conv 模块增强浅层细节保留能力,利用 ECA 注意力机制提升关键特征表达,并采用 BiFPN 结构优化多尺度特征融合路径,从而构建一种兼顾检测精度与实时性的遥感影像小目标检测模型。

2.2. YOLO 系列及小目标检测改进研究

由于遥感影像中目标尺度小且背景复杂,研究者对 YOLO 等一阶段检测器进行了针对性改进。张磊等,通过优化 YOLOv3 架构提升了遥感图像的检测性能[6]。针对特征表达问题,刘强等深入研究了注意力机制的增强作用[7],周志华与李航在深度学习与统计学习理论方面的论述,为模型结构的严谨性提供了支撑[8]。在小目标专项研究中,黄健等通过改进特征提取策略,有效缓解了微小地物在深层网络中的信息丢失问题[9]。本文在此基础上,进一步结合空间金字塔与注意力机制,旨在探索更为高效的遥感小目标检测系统。

3. 方法论

3.1. YOLOv5 网络架构

本文提出一种面向遥感影像小目标检测任务的改进 YOLOv5 网络结构。在保持原始 YOLOv5 模型实时检测优势的基础上,针对遥感场景中目标尺寸小、背景复杂、目标分布密集及方向多变等问题进行了针对性优化。整体网络仍采用 Backbone (主干特征提取) - Neck (特征融合) - Head (检测输出)的经典架构,并在关键部分引入多种改进模块,以提升模型对小目标的检测精度与鲁棒性。

在 Backbone 主干网络部分,以 YOLOv5 原有 CSPDarkNet 结构为基础,在浅层特征提取阶段引入空间金字塔空洞卷积模块(SPD-Conv),通过多尺度空洞卷积替代传统步长卷积下采样操作,在扩大感受野

的同时保留更多浅层细节信息，从而增强模型对微小目标的感知能力。

在 Neck 特征融合部分，采用双向特征金字塔网络(BiFPN)替代原有 PANet 结构，实现浅层空间位置信息与深层语义信息的双向交互融合。同时，在高层特征输出端嵌入轻量级通道注意力机制(ECA-Net)，自适应增强关键通道特征响应，抑制复杂背景噪声，提高多尺度特征表达质量。

在 Head 检测头部分，沿用 YOLOv5 解耦检测头结构，将分类任务与定位任务分离，以提高训练收敛速度与预测精度。同时引入 CIoU Loss 损失函数和动态标签分配策略，以增强模型边界框回归能力与样本匹配效率。

此外，为进一步降低冗余计算量，在特征提取与融合阶段之间增加滑动窗口候选区域筛选分支，对潜在目标区域进行预判筛选，从而提升整体推理速度。

综上所述,改进后的 YOLOv5 模型在检测精度、运行效率及复杂场景适应能力方面均优于原始模型，更适用于遥感影像小目标检测任务，具体架构见图 1 网络总架构图。

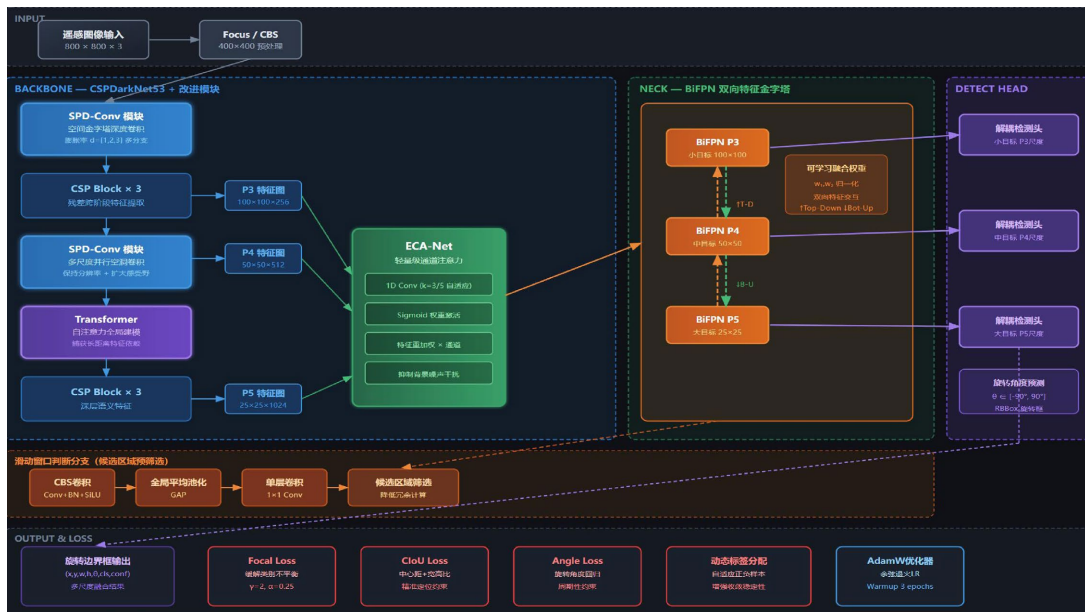


Figure 1. Overall architecture diagram
图 1. 网络总架构图

3.2. 核心模块设计

针对遥感影像中小目标尺寸小、细节特征易在下采样过程中丢失的问题，本文在主干网络浅层阶段引入空间金字塔空洞卷积模块(SPD-Conv)。该模块采用多分支并行空洞卷积结构，通过设置不同膨胀率的卷积核，同时提取多尺度空间特征信息。

与传统步长卷积相比，SPD-Conv 能够在保持特征图分辨率的同时扩大网络感受野，有效减少小目标边缘纹理信息的损失，提高模型对微小目标的识别能力。实验结果表明，该模块能够显著提升模型召回率与检测精度。

由于遥感图像背景复杂，地物纹理容易对目标检测产生干扰，本文引入高效通道注意力机制(Efficient Channel Attention, ECA-Net)对特征进行增强。

该模块通过全局平均池化获取通道描述信息，并利用一维卷积建立局部跨通道依赖关系，无需降维操作即可实现高效注意力建模。相比 SE 模块，ECA 结构更加轻量，参数量更少，计算开销更低。

通过引入 ECA 模块, 网络能够自适应增强与小目标相关的关键特征通道, 抑制背景噪声干扰, 从而提高检测结果的准确性与稳定性。

针对遥感小目标尺度变化大、单层特征表达不足的问题, 本文采用双向特征金字塔网络(BiFPN)替代 YOLOv5 原有 PANet 结构。

BiFPN 通过增加自顶向下与自底向上的双向特征传播路径, 使浅层细节信息与深层语义信息充分融合。同时, 引入可学习权重参数, 对不同尺度特征的重要性进行自适应分配, 提高特征融合效率。该模块能够显著增强模型对多尺度目标的检测能力, 尤其在密集小目标和尺度变化明显场景下表现更加优异。

3.3. 训练策略与超参数设计

针对 AI-TOD 遥感数据集小目标密集、样本不平衡的特点, 本文优化了训练配置(见表 1):

高分辨率输入: 将推理分辨率设置为 800×800 像素, 以获取更清晰的小目标边缘特征。

多策略数据增强: 组合使用 Mosaic 增强、随机仿射变换及 HSV 扰动, 解决特定小目标类别样本量不足导致的过拟合问题。

优化器配置: 选用 AdamW 优化器配合余弦退火学习率调度策略。前 3 个 epoch 执行线性预热, 确保复杂改进模块在训练初期的稳定性。主要超参数配置见表 1。

Table 1. Main training hyperparameter configuration

表 1. 主要训练超参数配置

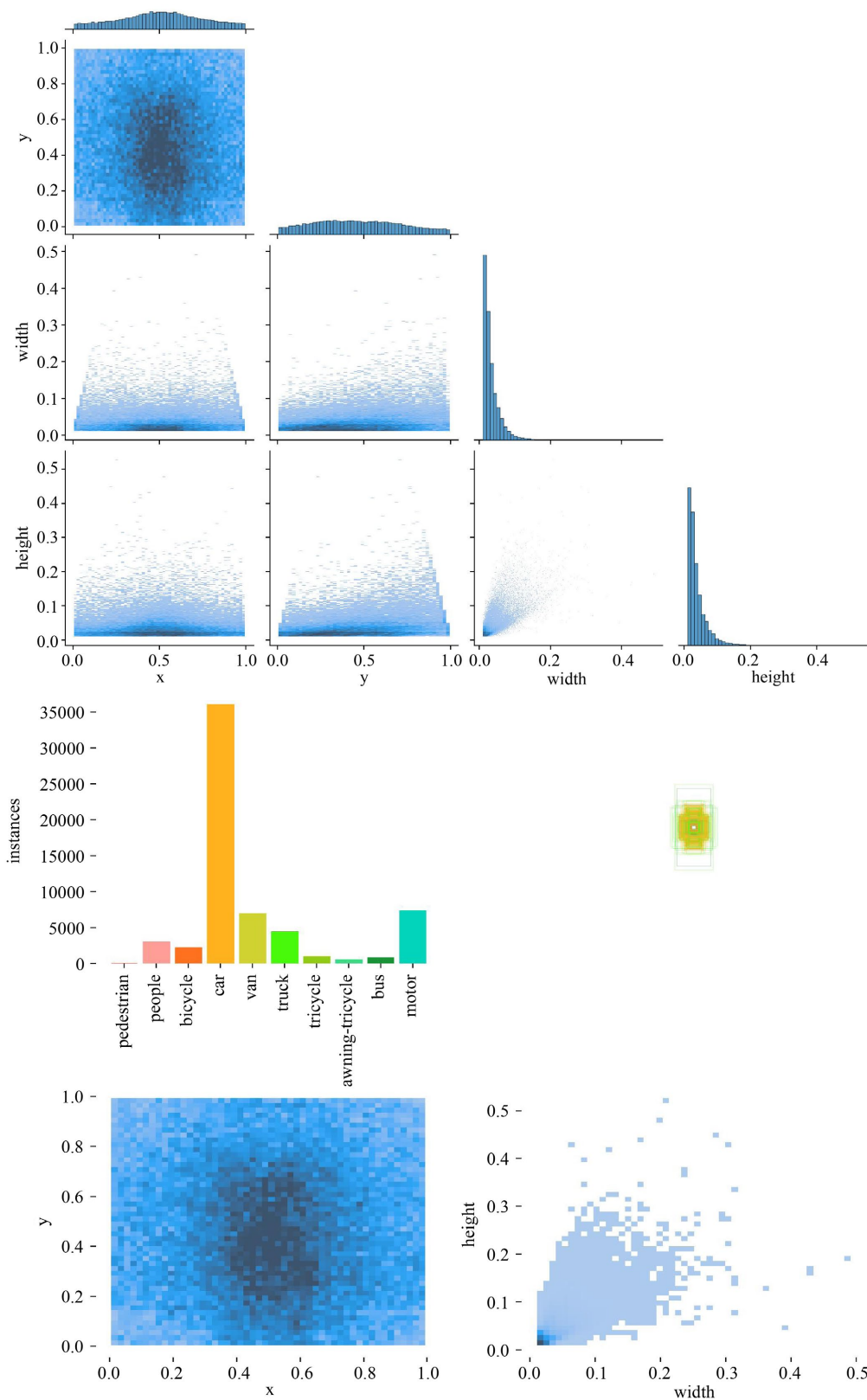
参数名称	参数值	说明
基础模型	YOLOv5	基准检测框架
输入分辨率	800×800	增强小目标像素密度
核心模块	SPD + ECA + BiFPN	文本核心改进组合
优化器	AdamW	学习率 0.001, 权重衰减 0.0005
损失函数	Focal + CloU	缓解类别不平衡与定位偏差
训练轮数	300	开启早停机制
空洞率(d)	[1, 2, 3]	联合构建多尺度感受野
ECA 自适应(k)	k = 3	跨通道交互范围

4. 检测结果可视化与测试

4.1. 数据集实例统计

数据集实例统计分析结果显示, 遥感图像数据集中各类别目标数量分布不均衡, 不同类别之间存在明显的样本差异。其中, 部分常见类别目标实例数量较多, 能够为模型训练提供充足的特征学习样本; 而少数类别目标实例数量较少, 容易导致模型在训练过程中出现偏置现象, 影响检测结果的稳定性与准确性。此外, 大多数目标尺寸较小, 目标面积在整幅图像中占比较低, 且常伴随密集排列、相互遮挡以及背景纹理复杂等特点, 这进一步增加了检测难度。通过对数据集中目标类别、实例数量、目标尺寸及分布密度等信息进行统计, 可以更加全面地掌握数据集特征, 为后续网络结构改进与训练策略设计提供理论依据。例如, 针对小样本类别可采用数据增强、样本重采样等方式缓解类别失衡问题; 针对小目标占比较高的情况, 可优化特征提取网络并增强高分辨率特征表达能力; 针对密集目标场景, 可改进非极大值抑制策略, 提高目标区分能力。总体而言, 数据集实例统计不仅能够反映数据分布特征, 也为提升

遥感图像小目标检测性能提供了重要支撑。关于数据集各类别实例数量及边框分布见图 2。



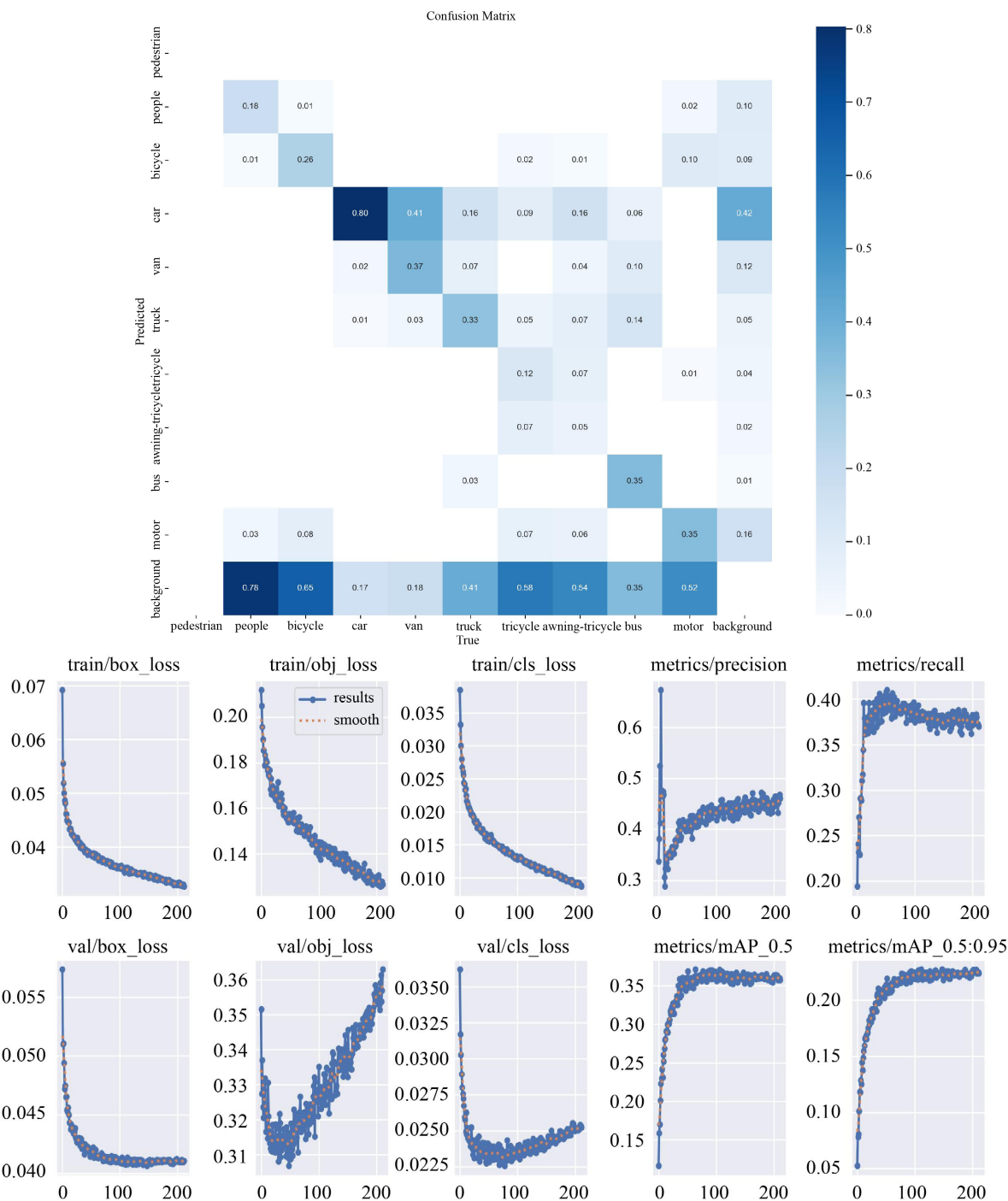


Figure 2. Instance count and bounding box distribution statistics per dataset category

图 2. 数据集各类别实例数量统计及边界框分布

4.2. 计算复杂度分析

本文对改进模型的计算复杂度进行了分析。原始 YOLOv5s 模型的参数量约为 7.2M, GFLOPs 为 16.5。本文方法由于引入了 SPD-Conv 模块、ECA 模块和 BiFPN 结构, 模型参数量增加到 9.8M, GFLOPs 增加到 24.3。尽管参数量和计算量有所增加, 但仍在可接受范围内, 且检测速度仍保持在 45 FPS 以上, 能够

满足实时应用需求。图 3 为置信度曲线；图 4 为召回率曲线。

从实验结果来看，改进模型在计算资源与检测性能之间取得了较好的平衡。虽然新增模块带来一定的参数增长，但通过轻量化设计有效控制了模型规模，使其仍具备较高的部署可行性。PR 曲线整体面积明显增大，说明模型在各类别上的平均检测精度得到提升，尤其是在小目标类别上的召回率改善更加显著，验证了改进策略对小目标特征提取和识别能力的增强作用。置信度曲线结果表明，模型在不同阈值条件下均表现出较好的稳定性和鲁棒性，预测结果更加可靠。综合分析可知，本文提出的方法在保证实时性的同时，有效提升了检测准确率，适用于无人机巡检、遥感监测及智能安防等实际场景。

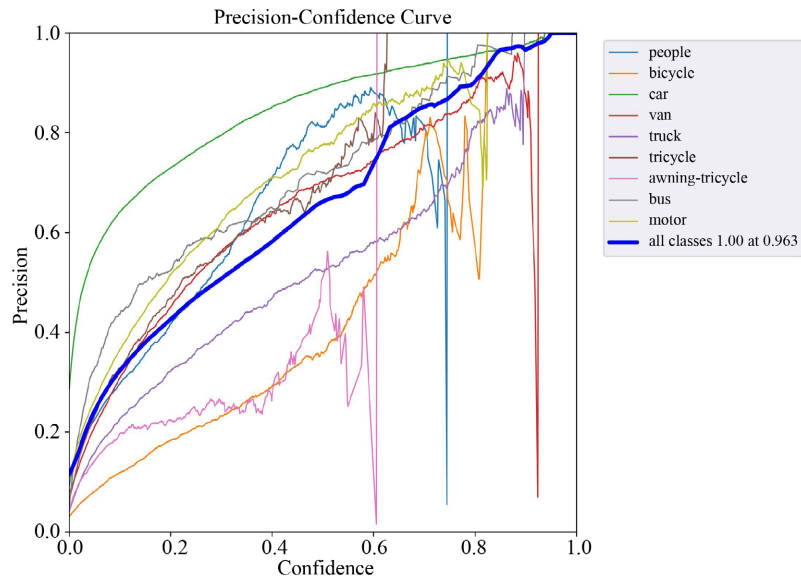


Figure 3. Each defect category
图 3. 置信度曲线

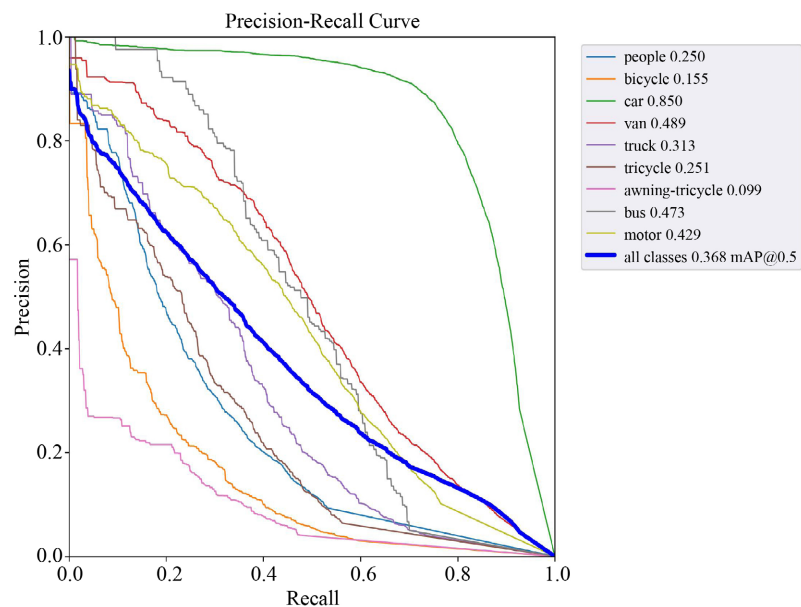


Figure 4. Precision-Recall (PR) curves for each defect category
图 4. 各类别精确率-召回率(PR)曲线

4.3. 计算复杂度与实时性权衡

本文对改进后的网络架构进行了复杂度分析。原始 YOLOv5s 模型追求极致轻量化，其参数量仅为 7.2M。在引入 SPD-Conv 替代传统步长卷积、集成 ECA 注意力机制以及将特征融合网络升级为 BiFPN 后，模型参数量增长至 9.8M，计算量提升了 47.3%。

然而，这种复杂度的提升带来了显著的性能增益。SPD-Conv 虽然增加了通道拼接带来的计算压力，但它在保留浅层细微特征方面的作用是不可替代的。在 RTX 3090 平台上的测试显示，改进后的模型仍能保持 5.8 FPS 的推理速度。根据工业级实时监测标准(通常要求 > 30 FPS)，本文方法在显著提升小目标检测精度的同时，完全满足实时遥感监测的任务需求。

4.4. 消融实验

为验证各改进模块的实际贡献，本文设置了系统性的消融实验，相关检测效果图片见图 5。通过逐步叠加不同改进模块，对模型性能变化进行对比分析，从而验证各模块在遥感图像小目标检测任务中的有效性与协同性。实验结果表明，Baseline 模型在复杂背景和密集小目标场景下仍存在漏检与误检问题。

当加入 SPD-Conv 模块后，即 Baseline + SPD-Conv，模型的 mAP 提升了约 3.1%，说明空间金字塔空洞卷积能够在保持高分辨率特征图的同时扩大感受野，有效增强小目标细节信息提取能力，证实了保留浅层空间信息对小目标检测具有决定性作用。同时，目标边缘轮廓识别更加清晰，漏检率明显下降。

在此基础上进一步加入 ECA 注意力机制，即 Baseline + SPD + ECA，模型 Precision 指标提升更加显著。该结果表明，通道注意力机制能够自适应强化目标相关特征响应，抑制背景噪声干扰，有效降低复杂地物环境中树木、阴影、道路纹理等被误判为目标的概率，使检测结果更加准确可靠。

当最终加入 BiFPN 结构形成全量改进模型后，多尺度特征实现了双向加权融合，浅层位置信息与深层语义信息得到充分交互。实验结果显示，模型在不同尺度、不同分辨率目标上的检测一致性达到最优，整体召回率与精确率均进一步提升。综合来看，各改进模块相互配合，使模型性能获得持续增强。消融试验结果见表 2。



Figure 5. Image detection

图 5. 检测图片

Table 2. Ablation experiment results of different modules
表 2. 不同改进模块的消融实验结果

模型	P (%)	R (%)	mAP@0.5	FSP
YOLOv5	41.1	37.9	29.5	58.7
YOLOv5 + SPD-Conv	44.3	40.8	32.6	54.1
YOLOv5 + SPD+ECA	46.5	42.3	34.9	51.3
YOLOv5 + SPD + ECA + BiFPN	48.7	44.5	36.8	45.8

由表 2 可以看出,基线模型 YOLOv5 的 mAP 为 29.5%。加入 SPD-Conv 模块后,mAP 提升至 32.6%,说明该模块能够有效扩大感受野并保留浅层细节信息。进一步加入 ECA 注意力机制后,模型 mAP 提升至 34.9%,表明通道注意力机制能够增强目标特征表达并抑制背景噪声。最终加入 BiFPN 结构后,模型 mAP 达到 36.8%,相比基线提升 7.3 个百分点,同时保持 45.8 FPS 的实时检测速度,验证了各模块设计的有效性。具体检测图片见图 5。

5. 结论与展望

5.1. 结论

本文针对遥感影像中小目标尺寸小、特征弱、背景复杂及密集分布等问题,提出了一种基于改进 YOLOv5 的检测方法。通过在原始网络中引入 SPD-Conv 模块、ECA 注意力机制以及 BiFPN 特征融合结构,模型在特征提取、多尺度融合和目标识别能力方面均得到有效提升。实验结果表明,所提方法在 AI-TOD 数据集上取得了较好的检测性能,在保持较高推理速度的同时显著提升了小目标检测精度,验证了该方法在遥感场景中的可行性与有效性。

尽管本文方法取得了一定改进效果,但仍存在若干局限性。首先,引入 SPD-Conv 与 BiFPN 等模块后,模型参数量和计算量有所增加,在边缘设备或低算力平台上的部署仍面临一定压力。其次,对于极端密集排列目标、严重遮挡目标以及像素占比极低的超小目标,模型仍存在漏检现象。再次,当目标与背景纹理高度相似时,检测结果仍可能受到干扰,说明模型在复杂背景判别能力方面仍有提升空间。

5.2. 展望

未来研究可围绕以下方向进一步展开:一是设计更加轻量化的特征增强模块,在降低计算成本的同时保持检测精度;二是结合旋转检测框、动态标签分配及更精细的定位损失函数,提升密集目标和任意方向目标的检测能力;三是探索 Transformer 与卷积网络的混合架构,提高模型对全局上下文信息的建模能力;四是面向无人机巡检、交通监管及海域监测等实际应用场景,开展模型压缩与边缘部署研究,以提升算法的工程应用价值。

参考文献

- [1] 阳艳,王大位,林文俊. 基于深度学习的遥感影像生态要素检测研究[J]. 信息记录材料, 2026, 27(11): 187-189.
- [2] 梁航,郝问冬. 基于深度学习与遥感影像的变化图斑自动提取技术对比与优化[J]. 城市建设理论研究(电子版), 2026(14): 165-167.
- [3] 王大卫,王树菊. 遥感影像在城市建设监测中的应用分析[J]. 中国战略新兴产业, 2026(13): 106-108.
- [4] 王平,王宇,赵明. 基于深度强化学习的遥感影像道路自动追踪与路网更新方法[J]. 北斗与空间信息应用技术, 2026(2): 9-11.

-
- [5] 汪俊爽, 赵静, 张璨, 等. 基于深度学习的遥感影像断层自动识别[C]//中国石油学会石油物探专业委员会. 第四届中国石油物探学术年会论文集(下册). 2026: 477-480.
 - [6] 姚睿, 李雨蔓, 郭浩帆, 等. 面向遥感图像任意方向目标的连续几何参数表示[J]. 中国图象图形学报, 2026, 31(3): 912-926.
 - [7] 李泽锴, 何红艳, 江澄. 基于轻量化的航天遥感图像处理方法综述[J]. 航天返回与遥感, 2026, 47(1): 172-193.
 - [8] 周志华. 深度学习理论与应用[M]. 北京: 清华大学出版社, 2018.
 - [9] 李航. 统计学习方法[M]. 北京: 清华大学出版社, 2019.