

基于改进 YOLOv10 和无人机影像的树种识别

韩誉, 刘浩然, 林文树*

(东北林业大学 机电工程学院, 哈尔滨 150040)

摘要: 精准高效的森林树种识别是智慧林业实现的关键。传统的地面调查方法存在效率低、成本高等问题, 而基于机器学习的树种识别方法通常依赖于大量的特征提取和先验知识。为解决这些问题, 提出一种基于改进 YOLOv10 的无人机影像树种识别算法, 通过轻量化网络设计和注意力机制增强, 实现边缘设备的高效部署, 为森林资源数字化管理提供技术支持。选取东北地区 5 种常见的树种落叶松 (*Larix gmelinii*)、黄檗 (*Phellodendron amurense*)、胡桃楸 (*Juglans mandshurica*)、榆树 (*Ulmus pumila*) 和水曲柳 (*Fraxinus mandshurica*) 构建无人机影像数据集, 通过采用轻量化卷积 (Ghost) 重构主干网络以降低计算复杂度。此外, 在特征融合层引入卷积块注意力模块 (Convolutional Block Attention Module, CBAM), 通过通道和空间双维度的特征校准增强细粒度特征提取能力。通过双向跨尺度连接 (BiFPN) 优化多尺度特征融合, 采用对称交并比损失函数 (Symmetric Intersection over Union Loss, SIoU) 改进边界框回归速度。最后于 Jetson Nano 嵌入式平台进行模型部署验证。改进后的 YOLOv10 模型在验证集上达到 91.5% 的查准率 (Precision) 和 77.5% 的 mAP@0.5, 分别较基线模型提升 4.5% 和 3.8%。部署实测平均推理速度为 43.5 FPS, 较基线模型提升 35.5%, mAP@0.5 达 75.7%。结果表明, YOLOv10 改进算法通过轻量化架构和优化多尺度特征提取, 在保持实时性的同时提升复杂林区场景的树种识别精度。该算法为无人机林业调查提供可嵌入式部署的解决方案, 特别适用于林冠层重叠度高、光照条件多变的实际作业环境。

关键词: 树种识别; 深度学习; 无人机; YOLO; 目标检测

中图分类号: S771.8

文献标识码: A

DOI: 10.7525/j.issn.1006-8023.2025.05.006

Tree Species Identification Based on Improved YOLOv10 and UAV RGB Imagery

HAN Yu, LIU Haoran, LIN Wenshu*

(College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin 150040, China)

Abstract: Efficient and accurate tree species identification is critical for the realization of smart forestry. Traditional field survey methods are low efficiency and high cost, while machine learning-based tree species identification approaches often rely on extensive feature extraction and prior knowledge. To address these issues, a tree species identification algorithm based on improved YOLOv10 for UAV imagery is proposed in this paper. The improved architecture integrates lightweight network design and attention mechanisms to enable efficient edge device deployment, providing technical support for digital forest resource management. A UAV imagery dataset was developed for five common tree species (*Larix gmelinii*, *Phellodendron amurense*, *Juglans mandshurica*, *Ulmus pumila*, and *Fraxinus mandshurica*) in North-east China. The backbone network was reconstructed using lightweight convolution (Ghost) for computational complexity reduction. The convolutional block attention module (CBAM) was introduced in the fusion layer to strengthen fine-grained feature extraction through channel and spatial dual dimensional feature calibration. Multi-scale feature fusion was optimized through bidirectional cross-scale connections (BiFPN), while bounding box regression efficiency was improved using a structured intersection over union (SIoU) loss function. Final deployment validation was conducted on

收稿日期: 2025-03-05

基金项目: 黑龙江省自然科学基金联合引导项目 (LH2020C049)。

第一作者简介: 韩誉, 硕士研究生。研究方向为基于深度学习与无人机影像的树种识别。E-mail: hanyu@nefu.edu.cn

* **通信作者:** 林文树, 博士, 教授。研究方向为林业遥感。E-mail: linwenshu@nefu.edu.cn

引文格式: 韩誉, 刘浩然, 林文树. 基于改进 YOLOv10 和无人机影像的树种识别 [J]. 森林工程, 2025, 41(5): 922-935.

HAN Y, LIU H R, LIN W S. Tree species identification based on improved YOLOv10 and UAV RGB imagery [J]. Forest Engineering, 2025, 41(5): 922-935.

the Jetson Nano embedded platform. The improved YOLOv10 model achieved 91.5% precision and 77.5% mAP@0.5 on the validation set, showing improvements of 4.5% and 3.8% compared to the baseline model, respectively. In practical deployment, the model achieved an inference speed of 43.5 FPS, 35.5% faster than the baseline model, with mAP@0.5 of 75.7%. Results showed that, the improved YOLOv10 algorithm successfully balances identification accuracy and real-time performance in complex forest environments through lightweight architecture and multi-scale feature optimization. The solution demonstrates particular effectiveness in scenarios with dense canopy overlap and variable illumination, offering an embeddable solution for UAV forestry surveys.

Keywords: Tree species identification; deep learning; UAV; YOLO; object detection

0 引言

树种的精确识别可为林业管理和森林生态保护提供重要的数据支撑。传统树种分类方法依赖人工样地调查,存在效率低、成本高和难以覆盖复杂地形等问题。随着遥感技术的发展,卫星和机载遥感平台等被广泛用于大规模的林草监测。其中,无人机因具有低成本、高灵活性及能够提供更高分辨率影像的特点而被广泛应用^[1-3]。

在遥感数据类型中,高光谱影像虽然可以提供更丰富的光谱信息,但需要进行数据降维和波段优选等预处理,过程较为繁琐;激光雷达虽然能提供精确的三维结构信息,但受限于高昂的设备和处理成本^[4-5]。相反,无人机影像(red green blue, RGB)仅需搭载常规的数字相机即可获得,并可利用现有的技术手段进行快速处理,使其成为树种分类中较为便捷、高效的遥感数据类型。目前,树种识别大多数依赖于机器学习算法,周湘山等^[6]运用支持向量机进行树种识别,该方法的整体精度为60%。王俊杰等^[7]运用随机森林进行无人机多光谱树种识别,总体精度为70.83%。姚扬^[8]运用随机森林进行多源遥感数据树种识别,精度达到了77.96%。虽然取得了较高的识别精度,但需要具备大量的先验知识。

而近年来深度学习技术在图像处理和模式识别领域取得了显著的成就,相比于传统的机器学习方法,深度学习能够通过学习大量数据自动提取复杂的特征,无需手动设计特征提取器,这使得深度学习在树种识别上表现出更高的准确性和鲁棒性^[9]。卷积神经网络(convolutional neural networks, CNN)^[10]、深度残差网络(deep residual network, ResNet)^[11]和深度信念网络(deep belief network, DBN)^[12]等多种深度学习算法被应用于对森林树种进行高效、精确的分类与识别。Knauer等^[13]采用混合卷积神经网络(CNN)和随机森林分类器对树种进行分类。尽管这种方法在准确性方面表现良好,但其对计算资源的要求较高,限制其在资源受限环境中的应用。周治焯^[14]运用了无人机近地遥感影像结合深度学习算法FC-DenseNet模型,对长

沙市的城市绿化进行了识别与分析,提出了不同飞行高度下的采集策略,但识别精度仍有很大提升空间。Chen等^[15]则采用了一种结合传统图像特征与深度学习的混合算法进行树种识别。该方法虽然通过结合光谱和纹理特征提高了识别准确性,但在样本数量有限时,算法的性能提升仍然受限。

在众多深度学习算法中,单阶段检测算法如YOLO^[16](you only look once)因其实时性和高效性而被广泛应用。与R-CNN(region-based convolutional neural networks)^[17]、SSD(single shot multiBox detector)^[18]和FasterR-CNN(faster region-based convolutional neural networks)^[19]等两阶段检测算法相比,YOLO通过单次前向传播即可完成图像中目标的定位和分类任务,显著提高了检测速度。Xu等^[20]基于YOLO开发的无人机多光谱树种分类算法在树种识别任务中展现出显著优势,但在树冠密度高的地区仍存在单棵树边界分割不清楚的情况。黄丽明等^[21]采用YOLO算法在无人机影像异常变色木识别中得到了良好的应用,但对于较小的异常变色木树冠,识别时容易产生漏检。Han等^[22]提出的LUFFD-YOLO模型在无人机RGB影像的轻量级森林火灾检测中展现出较好的跨数据集泛化能力,但其网络轻量化设计存在特征表达能力受限的问题。随着YOLO算法架构的持续优化,YOLOv10通过引入效率与精度协同设计策略^[23],使其更加轻量化,展现了在精细树种识别方面的潜力。

综上,现有树种识别研究仍存在3方面局限。1)林区复杂背景下树冠纹理相似性导致特征区分度不足;2)无人机影像中树木尺度差异显著,现有特征金字塔难以有效融合多尺度信息;3)模型轻量化与精度难以平衡。针对上述问题,本研究提出基于YOLOv10的改进算法。1)采用Ghost卷积(ghost convolution, Ghost-Conv)重构主干网络,通过特征重用机制降低计算复杂度;2)在特征融合层嵌入卷积块注意力模块(convolutional block attention module, CBAM),增强树冠纹理的细粒度特征提取;3)构建双向特征金字塔网络(bidirectional feature pyramid network, BiFPN),优化多尺度树冠的特征交互;4)引入对称交并比损失函数(struct-

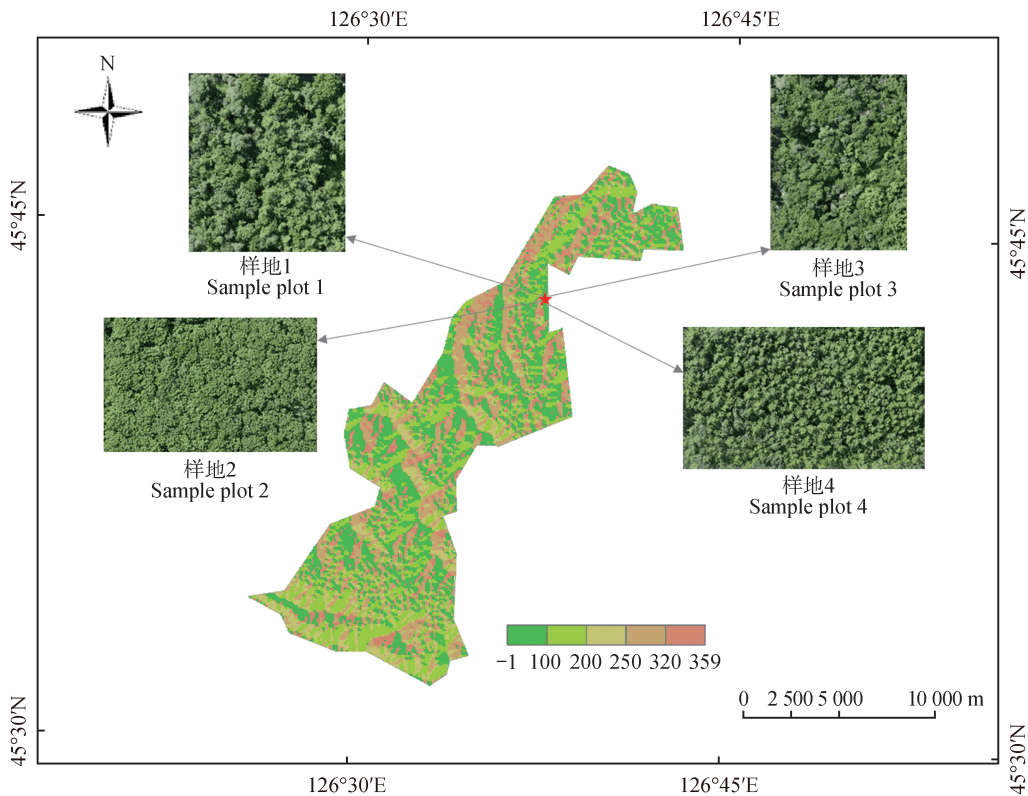
ted intersection over union, SloU)改进倾斜树冠的定位精度。通过算法轻量化设计和优化多尺度特征融合,在Jetson Nano嵌入式平台实现部署,为森林资源数字化管理提供可落地的技术方案。

1 数据与方法

1.1 研究区域与数据采集

研究区域位于黑龙江省哈尔滨市林业示范基地,如图1所示,地理坐标范围为 $45^{\circ}43'05.83''$ — $45^{\circ}43'15.09''$ N、 $126^{\circ}37'22.63''$ — $126^{\circ}37'28.49''$ E,海拔136~148 m^[24]。

本研究选取4块典型样地,总面积约为20 550 m²,占林场航拍区域(4.8 hm²)的42.8%,涵盖针阔混交林(样地1)、胡桃楸林(样地2)、黄檗林(样地3)、落叶松林(样地4),面积分别为4 550、5 250、3 750、7 000 m²。主要研究树种包括落叶松(*Larix gmelinii*)、黄檗(*Phellodendron amurense*)、胡桃楸(*Juglans mandshurica*)、榆树(*Ulmus pumila*)和水曲柳(*Fraxinus mandshurica*)等。其中,样地1混交林包含落叶松、黄檗、榆树、胡桃楸和水曲柳等树种,其他样地均为纯林,仅包含少量其他树种。



*审图号:GS(2024)0650号

*Drawing number:GS(2024)0650

图1 研究区域

Fig. 1 Study area

采用大疆御2 RTK无人机(Hasselblad L1D-20c相机,28 mm焦距)于2023年9月1日夏秋季节更迭期间进行航拍,参数设置为相对航高130 m(地面分辨率1.2 cm/px)、航速5 m/s、航向207°、旁向/航向重叠率75%/80%。光照条件为晴天无云,太阳高度角52°,影像采集时间UTC+8 15:00。通过航拍整个林场矩形区域获取359张原始影像后,经Pix4Dmapper生成数字正射影像(DOM),并通过千寻RTK(平面精度1 cm)获取单木三维坐标(采样频率1 Hz,10万元均值),通过Arc-

GIS实现亚像素级地理配准(平面偏差均值1.2 cm)。

1.2 数据集制作

研究采用亚米级正射影像构建深度学习专用数据集,其制作流程涵盖空间裁剪、分层抽样、精准标注与增强优化4个阶段。通过ArcGIS对4块样地进行空间网格划分(单元尺寸7.68 m×7.68 m),以50%重叠滑动窗口生成640×640像素样本图像。为平衡模型泛化能力与空间独立性,实施双层抽样策略:第一层按样地类型(针阔混交林/黄檗林/胡桃楸林/落叶松林)进行7:3

比例划分,第二层在样地内部设置 20 m 空间缓冲带消除训练与验证集地理重叠。最终形成包含 1 000 张训练图像和 500 张测试图像的数据集,确保同一样木不会同时出现在 2 个子集(经空间坐标哈希验证)。

标注工作依托 Roboflow 协同标注平台,以千寻 RTK 定位数据为空间基准实施三级质量控制。首先由 3 名林业工程专业人员独立标注树冠矩形框(bounding box),随后通过交叉计算标注者间 IoU 值(0.89±0.03)筛选争议样本,最终结合 RTK 三维坐标反演(平面残差 1.2 cm,高程残差 2.8 cm)对重叠树冠进行空间拓扑仲裁。

在此基础上,通过多尺度数据增强策略提升模型

泛化能力。首先对训练集随机施加±30°旋转和水平翻转(概率 50%),模拟无人机不同航向角度的拍摄差异;随后在 HSV 色彩空间内调整亮度(±15%)、饱和度(±20%)和色调(±5%),以应对光照条件变化对树种色彩特征的影响。对增强后的数据集进行人工筛选,最终构建的数据集包含训练集 2 000 张、验证集 500 张、7 394 个检测框,其中榆树(347 框)和水曲柳(351 框)样本量显著低于其他类别,仅占落叶松样本量的 10.9%和 11.0%,见表 1。针对类别不平衡对模型训练的潜在偏倚,本研究提出自适应权重调整方法。在损失函数设计中,本研究采用自然对数函数对样本量进行非线性变换,构建类别权重,计算公式为

表 1 检测目标分布与权重分配

Tab. 1 Distribution of detection targets and class weights

树种 Species	训练集/个 Training set	验证集/个 Validation set	总检测框/个 Total detection box	权重 w_c
黄檗 <i>P. amurense</i>	899 (15.2%)	165	1 064	0.97
落叶松 <i>L. gmelinii</i>	2 623 (44.3%)	566	3 189	0.42
胡桃楸 <i>J. mandshurica</i>	1 988 (33.6%)	455	2 443	0.55
水曲柳 <i>F. mandshurica</i>	109 (1.8%)	242	351	1.52
榆树 <i>U. pumila</i>	299 (5.1%)	48	347	1.53
总计 Total	5 918	1 476	7 394	—

$$w_c = \ln \left(\frac{\sqrt{N}}{n_c} \right) \quad (1)$$

式中: w_c 为权重; N 为总检测框数; n_c 为 c 类别检测框数量; $\ln(\cdot)$ 为自然对数函数,用于压缩权重范围。

1.3 研究方法

1.3.1 YOLOv10 算法改进

YOLOv10 算法的核心架构基于单阶段目标检测的理念,通过将目标检测任务转化为回归问题,直接从图像中预测边界框和类别^[23]。本研究以 YOLOv10n 为基线模型,采用 GhostConv、CBAM 注意力机制、BiFPN 双向特征金字塔网络和 Siou 损失函数进行改进,改进的模型网络结构如图 2 所示,图 2 中红色框线为改进部分。输入图像为 640×640 分辨率,RGB 三通道。

1) Ghost 卷积^[25]。在无人机影像树种识别任务中,高分辨率图像处理与边缘设备部署需求之间难以平衡。传统卷积操作通过 $k \times k$ 卷积核生成 C_{out} 通道特

征图,其复杂度计算公式为

$$O_{Conv} = H \times W \times C_{in} \times C_{out} \times k^2 \quad (2)$$

式中: H 、 W 分别为特征图尺寸的长和宽; O_{Conv} 为计算复杂度; k 为卷积核尺寸; C_{in} 、 C_{out} 分别为输入、输出通道数(维度)。

当处理无人机拍摄的 4 000×3 000 像素级影像时,YOLOv10 原始主干网络的计算量达到 32.7 GFLOPs,难以满足 Jetson Nano 等嵌入式设备的实时性要求。在 YOLOv10 的主干网络中引入 GhostConv(图 2)。通过特征重映射机制实现计算量压缩,具体分为 2 阶段。第一阶段为基础特征提取,采用深度可分离卷积生成内在特征;第二阶段为 Ghost 特征生成,先通过 1×1 卷积进行线性变换,扩展特征通道;最终输出通过通道拼接获得完整特征图。

2) CBAM 注意力机制^[26]。基于无人机影像中林冠层遮挡严重、小目标纹理模糊的特性,本研究在

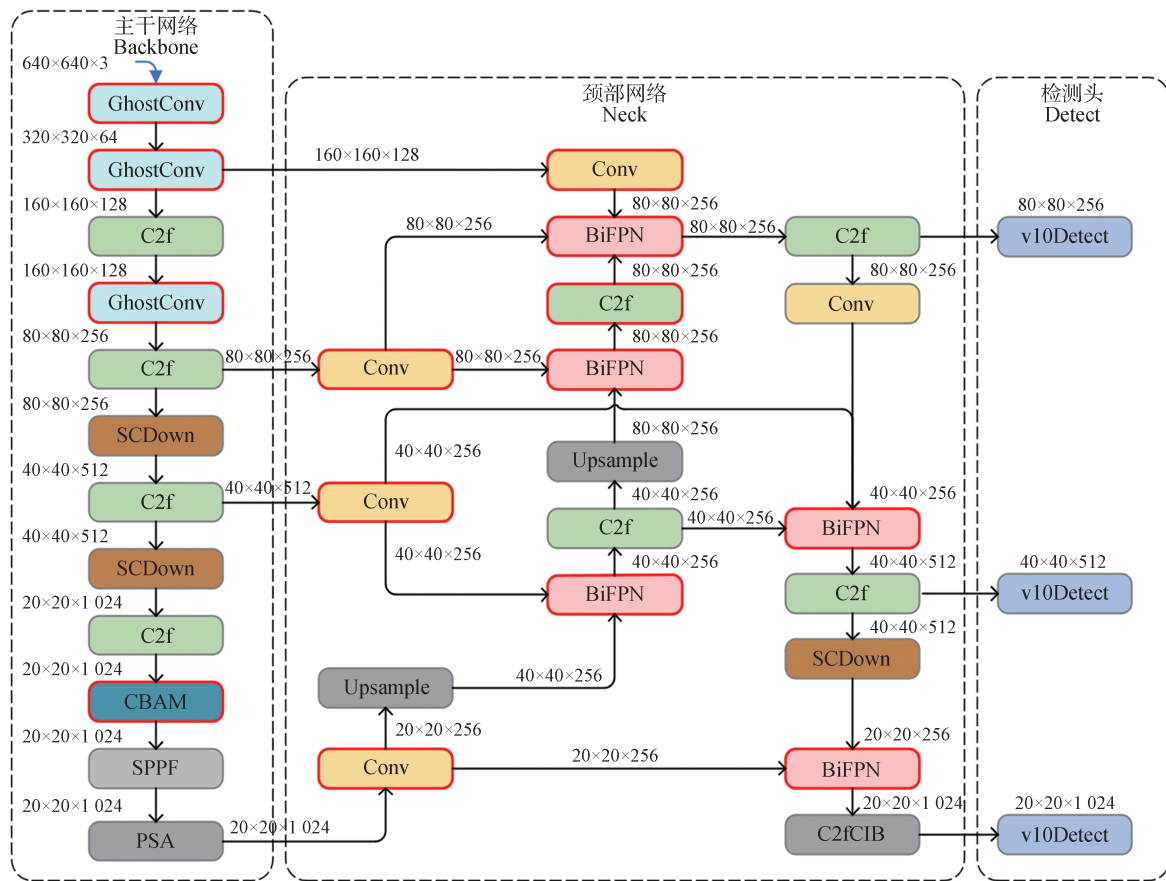


图2 改进的YOLOv10网络结构

Fig. 2 Improved YOLOv10 network architecture

YOLOv10主干网络末端(Pre-SPPF阶段)嵌入改进的通道与空间注意力模块(CBAM)。由图2可知,CBAM模块位于主干网络SPPF层前,输入为C2f模块输出的1024维高层特征,输出流向为校准后的特征依次输入SPPF(空间金字塔池化)和PSA(金字塔注意力模块),最终传递至BiFPN进行跨尺度融合。

3)BiFPN双向特征金字塔网络^[27]。针对林区无人机影像中树种多尺度分布的特性,本研究在YOLOv10检测头前端设计双向跨尺度特征金字塔网络(BiFPN),通过自适应权重融合与双向信息流增强多层次特征的表征能力。由于BiFPN的跨层连接需要特征图在空间和通道维度对齐,故同时引入多个Conv将不同层级的通道数统一到相同维度,使多尺度特征可以逐元素相加或拼接。

4)Siou损失函数^[28]。针对林区无人机影像中树冠不规则、树木倾斜导致的边界框回归偏差问题,本研究采用Siou损失函数替代传统CIou损失,通过引入角度惩罚项与动态形状约束,优化检测框的方向敏感性与几何匹配精度。

1.3.2 基于改进YOLOv10的嵌入式部署

为满足林区实时监测需求,本研究基于Jetson Nano嵌入式平台完成改进YOLOv10模型的部署优化。通过轻量化模型结构与硬件加速技术结合,实现无人机影像树种识别的边缘端高效推理。

具体技术流程如图3所示。首先,将PyTorch训练的改进模型通过开放神经网络交换(open neural network exchange, ONNX)中间表示格式进行标准化封装,消除框架依赖性。在Jetson Nano嵌入式平台部署TensorRT推理引擎,实施三级计算图优化策略。将Backbone网络中Conv-BN-SiLU复合结构合并为单节点计算单元,减少GPU内核调用次数。对BiFPN特征金字塔层保留FP16精度以防止小目标特征丢失,对Head检测层实施INT8量化^[29](校准算法采用基于KL(Kullback-Leibler)散度的熵最小化方法),生成校准表时通过500张验证集图片统计激活值分布。采用CUDA Graph固化计算图拓扑结构,消除动态形状推理的上下文切换开销,通过Zero-Copy内存管理器^[30]建立CPU-GPU地址映射直通通道,使数据预处理阶段的内存拷贝延迟降低。

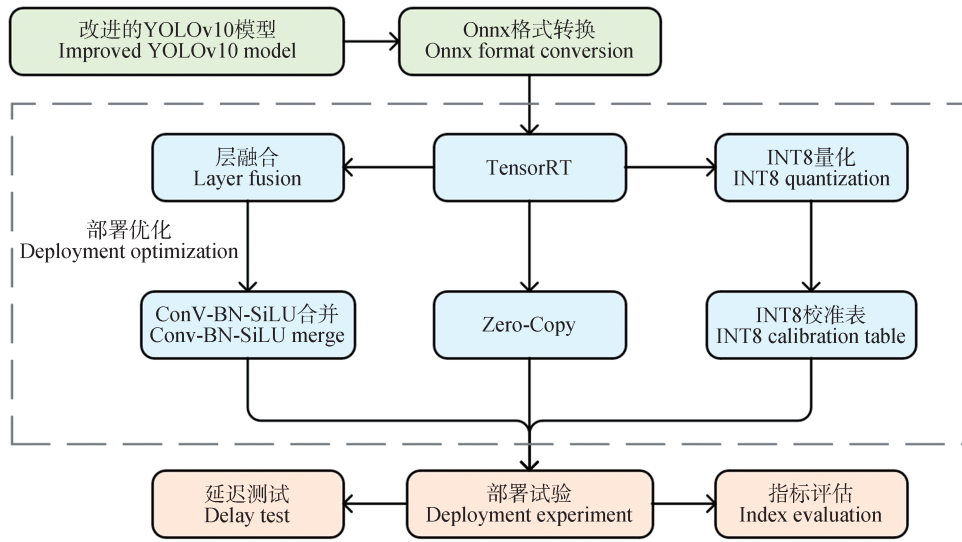


图 3 部署流程图

Fig. 3 Deployment flowchart

1.3.3 评价指标

本研究选择查准率 (precision, 式中记为 P)、查全率 (recall, 式中记为 R)^[31]、平均精度均值 (mAP@0.5, 式中记为 m_{AP})^[32] 和算法推理速度 (frames per second, FPS) 作为评价指标, 查准率和查全率计算公式为

$$P = \frac{T_p}{T_p + F_p} \quad (3)$$

$$R = \frac{T_p}{T_p + F_N} \quad (4)$$

式中: T_p 为预测框与真实框 IoU ≥ 阈值, 且分类正确; F_p 为预测框与真实框 IoU < 阈值, 或分类错误; F_N 为未被检测到的真实框, 阈值为 0.5。

A_p 为平均精度, 是对某一类别的查准率-查全率曲线下的面积计算。计算公式为

$$A_p = \int_0^1 P(r) dR \quad (5)$$

式中: R 为查全率; dR 为积分变量。

m_{AP} 为平均精度均值 (mean average precision, mAP), 是所有类别 A_p 的平均值, 可以用来评估模型对所有类别的检测性能, mAP@0.5 为交并比 (IoU) 阈值为 0.5 时的平均精度均值。计算公式为

$$m_{AP} = \frac{1}{N} \sum_{i=1}^N \int_0^1 A_{p_i} \quad (6)$$

式中: N 为类别数; A_{p_i} 为第 i 类的平均精度。

2 结果与分析

2.1 试验设置

本试验使用的工作站试验平台的软硬件环境为

GPU NVIDIA RTX 3090、CPU intel W2295、Memory 64GB、Windows11、Cuda 12.1、Python 3.11.5、Pytorch 2.2.0, 对比试验和消融试验均在此环境下进行。试验期间工作站保持不间断状态。部署试验硬件平台选用 Jetson Nano 4 GB Developer Kit, 搭载 128 核 NVIDIA Maxwell GPU 及 4 核 ARM Cortex-A57 CPU。软件环境部署 JetPack 5.1.2 SDK, 集成 CUDA 11.4、cuDNN 8.6 及 TensorRT 8.6, 支持 PyTorch 模型至 TensorRT 引擎的转换。

试验采用 AdamW 优化器, 初始学习率设为 3×10^{-4} , 配合余弦退火调度 (cosine annealing) 策略, 最小学习率衰减至 1×10^{-6} 。批量大小 (batch size) 设置为 32, 输入图像分辨率调整为 640×640 。训练过程持续 400 个 epoch, 前 3 个 epoch 进行线性热身 (linear warm-up), 权重衰减系数为 0.05。损失函数由分类损失 (BCEWithLogitsLoss)、定位损失 (SIoU) 和置信度损失 (focal loss) 加权组成, 权重比为 1:0.7:0.5。早停机制 (patience) 设置为 50。

2.2 不同模型性能评估

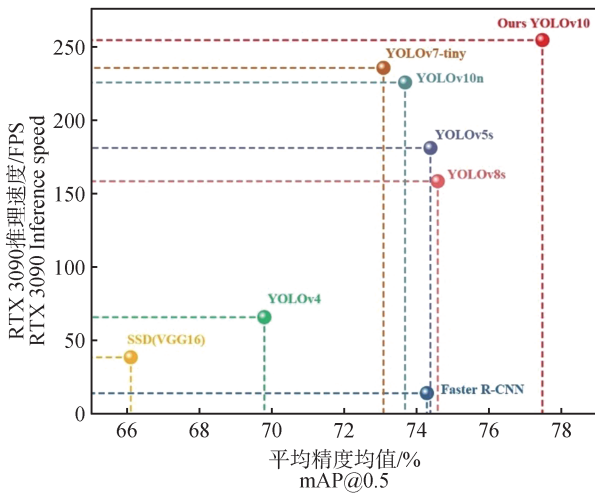
为系统评估改进 YOLOv10 模型的综合性能, 本研究在自建无人机树种数据集上对比了 Faster R-CNN、SSD 和 YOLO 系列等主流检测模型, 并从精度、速度、轻量化及能效 4 个维度展开量化分析, 不同模型性能数据见表 2, Jetson Nano 推理速度为部署环境下的指标, 其余各项指标皆在工作站环境中得到, 其中 Faster R-CNN 因显存占用过高 (>4 GB), 未能在 Jetson Nano 完成推理测试。

表2 不同模型性能对比
Tab. 2 Performance comparison of different models

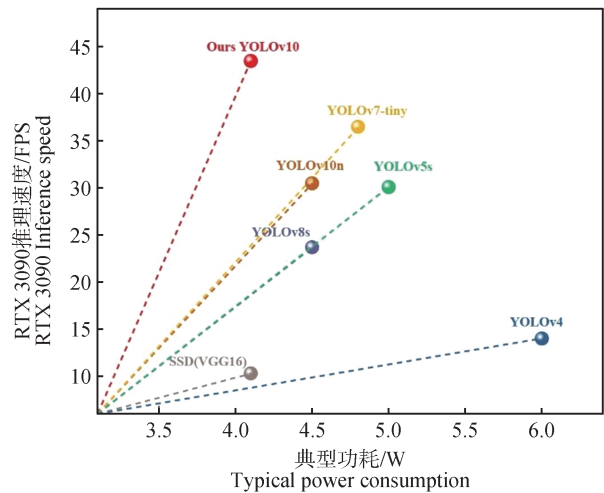
模型 Models	查准率/% Precision	查全率/% Recall	平均精度均值/% Mean average precision (mAP@0.5)	参数量/M Parameters	浮点运算数量/G FLOPs	RTX 3090 推理速度/FPS RTX 3090 inference speed	Jetson Nano推理速度/ FPS Jetson Nano inference speed
Faster R-CNN	88.1	68.4	74.3	136.20	287.4	14.2	—
SSD (VGG16)	83.5	66.7	66.1	26.80	62.7	38.5	10.3
YOLOv4	84.2	68.9	69.8	65.20	112.3	65.8	14.0
YOLOv5s	85.4	70.2	74.4	7.20	16.4	180.6	28.3
YOLOv7-tiny	87.2	70.5	73.1	6.00	13.7	235.0	36.5
YOLOv8s	87.5	71.8	74.6	10.50	26.9	158.0	23.7
Yolov10n	87.0	72.7	73.7	2.70	8.2	225.0	30.5
Improved YOLOv10	91.5	75.0	77.5	2.58	8.1	253.8	43.5

根据表2绘制不同模型平均检测精度与推理时间散点图,如图4(a)所示。改进模型在RTX 3090 GPU上mAP@0.5与推理速度(FPS)的散点图中显著位于右上方,表明其具有高精度与高实时性。改进模型较同家族的YOLOv8s分别提升4个百分点和60.6%的推理速度,同时超越YOLOv7-tiny和YOLOv5s等其他主流模型,体现了无NMS架构与轻量化设计的协同增效作用。图4(b)为不同模型能效比,能耗数据为基于3次重复试验标准差得到,误差范围为±5%,直线斜率为模型能效比(FPS/典型功耗)。改进模型在4.1W的典型

功耗下可稳定运行43.5 FPS,其能效比(10.6 FPS/W)优于YOLOv5s(6.0 FPS/W)和SSD(2.5 FPS/W)。从模型复杂度视角分析,改进YOLOv10以2.58 M参数量成为对比模型中最为紧凑的架构,较YOLOv5s和YOLOv8s分别减少64.2%和75.4%的参数规模。其参数量仅为两阶段模型Faster R-CNN的1.9%。综合试验结果表明,改进YOLOv10在精度、速度、轻量化及能效4个关键指标上均展现显著优势,更加适用于林业复杂场景的树种识别。



(a) 不同模型平均检测精度与推理速度
(a) Average detection accuracy and inference speed for different models



(b) 不同模型能效比
(b) Energy efficiency ratio of different models

图4 不同模型性能对比

Fig. 4 Performance comparison of different models

2.3 不同改进模块的效果分析

为验证各模块对模型性能的协同优化机制,本研究通过消融试验对比GhostConv轻量化卷积、CBAM注意力机制、BiFPN跨尺度融合三者的组合效应,见表3。

由表3可知,原始YOLOv10基线模型(试验组1)在轻量化特性下达到73.7%的平均精度均值与225 FPS的推理速度,但多尺度的查全率存在瓶颈。

表 3 消融试验结果
Tab. 3 Results of ablation experiment

试验组 Experimental groups	模块组合 Combination of modules	查准率(P)/% Precision	查全率(R)/% Recall	平均精度均值 (mAP@0.5)/% Mean average precision(mAP@0.5)	参数量/M Parameters	浮点运算数量/G FLOPs	推理速度/FPS Inference speed
1	Baseline (原始 YOLOv10)	87.0	72.7	73.7	2.70	8.2	225.0
2	Baseline+BiFPN	86.4	73.1	74.1	2.55	8.4	195.3
3	Baseline+CBAM	85.0	72.0	75.2	2.76	8.2	198.7
4	Baseline+GhostConv	89.9	71.3	77.2	2.47	7.9	228.4
5	Baseline+SIOU	87.0	72.5	74.1	2.7	8.2	226.5
6	Baseline+GhostConv+CBAM	87.5	73.5	75.8	2.75	8.1	210.2
7	Baseline+GhostConv+BiFPN	90.1	73.8	76.3	2.51	8.2	228.6
8	Baseline+CBAM+BiFPN	88.2	74.2	75.8	2.59	8.5	205.4
9	Baseline+GhostConv+CBAM+BiFPN	91.2	74.5	76.9	2.58	8.1	235.1
10	全模块组合 (Ours)	91.5	75.0	77.5	2.58	8.1	253.8

通过引入 BiFPN(试验组 2)改善多尺度特征的提取且不增加参数量,查全率较基线模型提升 0.4%,参数量减少 5.6%,但因其多尺度连接结构增加了计算流复杂性,在未配合轻量化设计时推理速度下降了 13.2%。而单独添加 CBAM(试验组 3)虽增强特征选择能力(mAP@0.5 提升 1.5%),却因通道与空间注意力计算引入额外内存延迟,参数量提高 2.2% 导致 FPS 下降 11.7%,表明注意力机制需与轻量化设计协同使用。同时为增强多尺度融合时的特征选择能力,将 CBAM 注意力机制与 BiFPN 融合(试验组 8),进一步提高查全率,较基线模型提升 1.5%,这种折中设计在保证特征表达能力的同时实现了轻量化,但推理速度仍然低于基线模型,需要进行轻量化改进。SIOU 损失函数的单独试验表明(试验组 5),其通过角度惩罚项与

动态形状约束机制,提升了倾斜目标的定位精度,mAP@0.5 提升 0.4%,推理速度略微提升 1.5 FPS。单独引入 GhostConv 提高轻量化特性时(试验组 4),参数量减少 8.5%,GFLOPs 降低 3.7%,mAP@0.5 提升 3.5%。GhostConv 与 CBAM 的联合使用时(试验组 6)在除推理速度以外各方面性能均有小幅度提升。GhostConv 与 BiFPN 的联合使用时(试验组 7),GFLOPs 与基线模型相近,将 Precision 提高到 90% 以上,且 FPS 提高 1.6%。集成 GhostConv、CBAM 与 BiFPN 时(试验组 9),获得了全方位的性能提升。

对比改进模型与原始 YOLOv10n 的行归一化混淆矩阵可知,基线模型中落叶松查准率最高,主要误检为黄檗,如图 5 所示。榆树查准率最低,漏检率高达 33%,这与样本量较少有关。改进模型使漏检样本总数减少 41%。

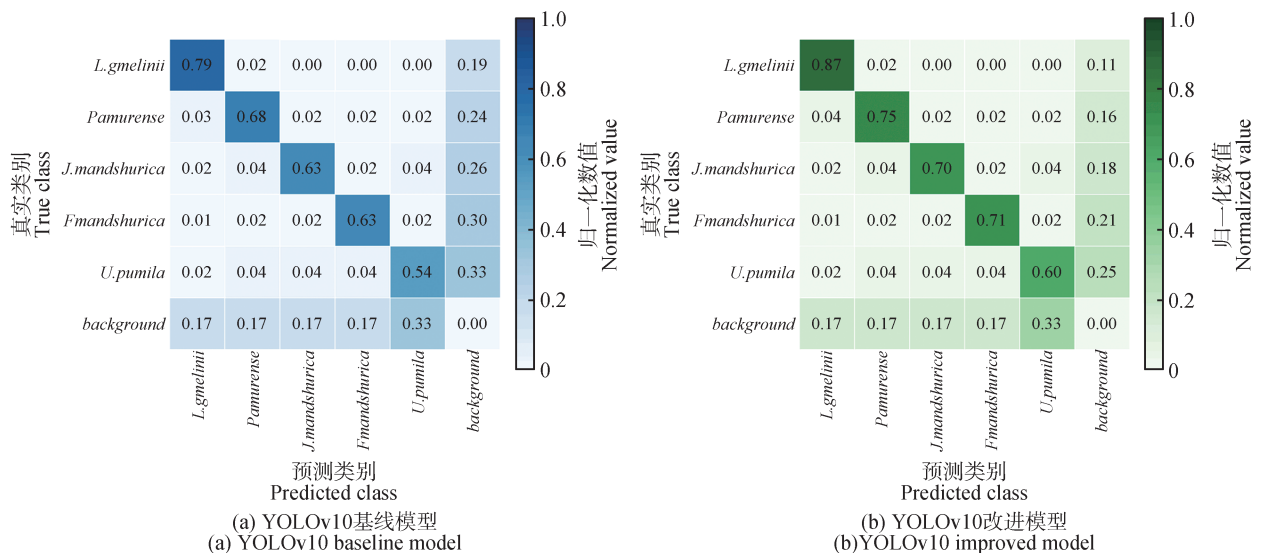


图 5 行归一化混淆矩阵

Fig. 5 Row normalized confusion matrix

全模块组合(Ours)通过集成 GhostConv、CBAM 与 BiFPN 和 Siou,以 2.58 M 参数量达到 77.5% 平均精度均值与 253.8 FPS,较基线模型提升 5.2% 的平均精度均值与 12.8% 推理速度。结果表明, Siou 并非孤立提升单指标,而是通过几何约束重构与现有模块形成正向反馈。其角度与尺度联合优化特性,使 GhostConv 的轻量化优势、CBAM 的注意力聚焦能力、BiFPN 的多尺度适应性得到更充分释放。性能雷达图中清晰地展示了各模型在不同指标下的分布情况,如图 6 所示,图 6 中 GFLOPs 和参数量(parameters)的值越小,表示性能

越好。因此,在雷达图中,对 GFLOPs 和参数量的坐标轴进行了反转处理。在误差允许范围内,全模块组合模型在 P、R、mAP@0.5 和 FPS 4 个评价指标上全面领先,这一结果印证了模块协同设计的必要性: GhostConv 降低卷积计算密度使模型轻量化, CBAM 通过通道权重校准抑制复杂背景干扰筛选特征,而 BiFPN 则增强了对重叠目标的多尺度适应性, Siou 优化了角度与尺度特性。全模块模型的 FPS-mAP 能效比(FPS/GFLOPs=31.3)较基线(27.4)提升 14.2%,证明多模块联合优化可突破轻量化模型的性能边界。

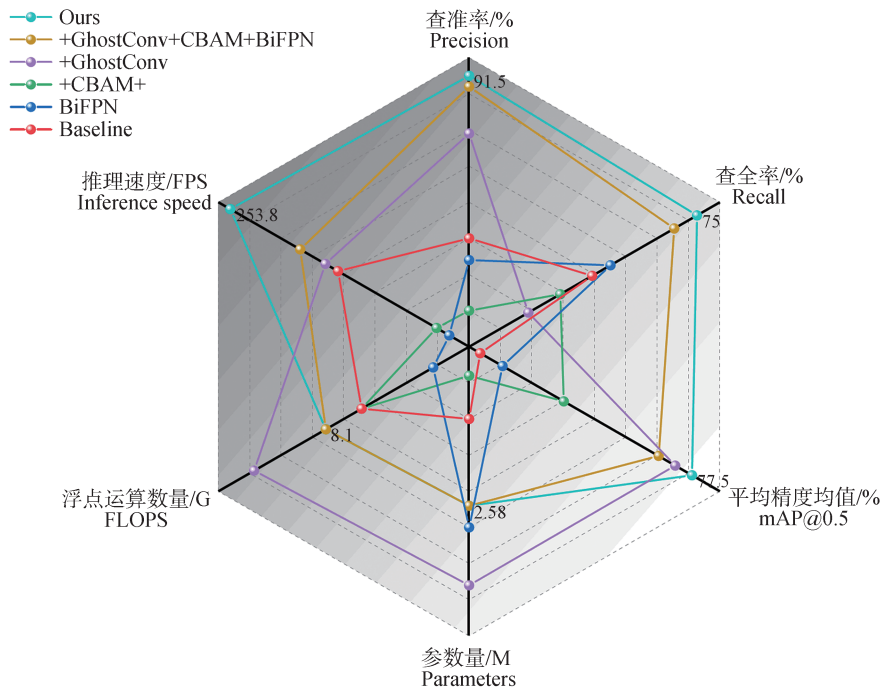


图 6 模块性能雷达图

Fig. 6 Module performance radar diagram

2.4 模型部署试验分析

为验证改进 YOLOv10 模型在林业边缘计算场景下的实用性,基于 NVIDIA Jetson Nano 4 GB 平台开展优化试验。试验结果见表 4,通过 TensorRT INT8 量化(试验组 3)和层融合技术(试验组 4),最终实现推理速

度 41.2 FPS 至 43.5 FPS 的提升,较 FP32 基线(试验组 1)提升 373.6%~400%,同时内存占用压缩 79.2%,达成边缘设备实时检测的技术目标。INT8 量化(试验组 3)使模型内存占用从 FP16(试验组 2)的 683 MB 压缩至 327 MB(降幅 52.1%)。

表 4 部署试验结果

Tab. 4 Results of deployment experimental

实验组 Experimental groups	优化策略 Optimization strategy	推理速度/FPS Inference speed	内存占用/MB Memory occupancy	平均功耗/W Average power consumption
1	FP32	8.7	1 245	6.8
2	FP16	22.4	683	5.2
3	INT8	41.2	327	4.6
4	Ours	43.5	259	4.1

为验证量化部署稳定性,由图7(a)可知,INT8 量化策略使 C2fCIB Stage5 层的 99.6% 的激活值集中在 $[-0.5, 0.5]$, 形成显著截断效应。相较于 FP32 的正态分布(KS 检验 $p=0.32$), 量化后的分布与原始分布重叠面积为 35.7%, 通过蒙特卡罗积分计算的信息熵损失率 $\Delta H=4.2\%$ 。这种可控的信息损失提升硬件并行计算效率。图 7(b) 的检测框置信度分布进一步揭示模型优化效果。FP32 模型置信度呈宽幅波动(0.65~

0.91, 标准差 0.21), 并出现 2 个 0.48、0.53 异常低值; 而 INT8 模型置信度范围收窄至 0.74~0.87(标准差 0.17), 异常值完全消除。通过图 8 分析发现, 异常值主要源于 2 种场景, 光照突变导致归一化误差(占 85%, 如图 8(a) 所示); 树冠晃动引发纹理断裂(占 15%, 如图 8(b) 所示)。量化策略通过动态范围约束算法, 使模型在复杂环境下保持稳定输出。

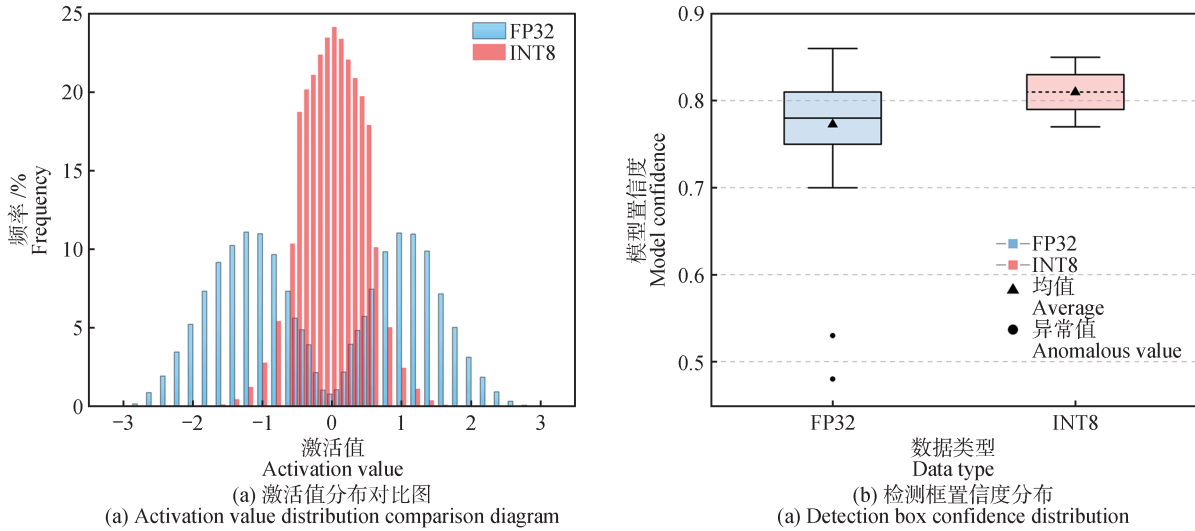


图 7 模型优化效果

Fig. 7 Model optimization effect

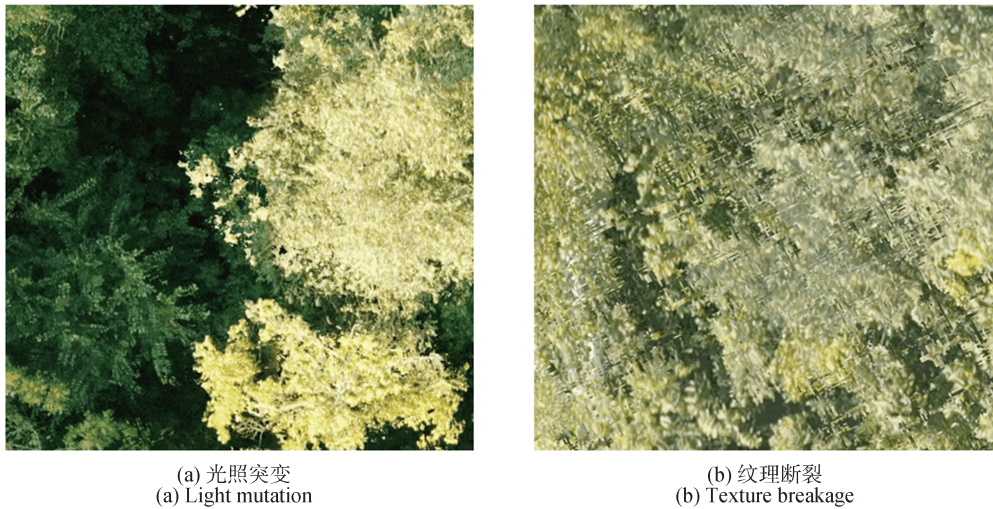


图 8 异常值主要场景

Fig. 8 Outliers in the main scenario

层融合与内存优化策略(试验组 4)带来性能的进一步提升。由图 9 可知, 通过 Conv-BN-SiLU 复合层融合, 计算节点数量减少 32.7%, 最大延迟层 C2fCIB Stage5 的执行时间从 11.2 ms 压缩至 3.7 ms, 降幅 66.9%。这种结构优化使 FPS 提升 5.6%, 同时峰值内存占用下降 20.8%。图 10 的功耗监测曲线表明, 优化

后的推理过程功率波动范围从 ± 1.2 W 缩窄至 ± 0.3 W, 这是由于内存访问模式优化使 DDR(double data rate) 突发请求频率降低。最终在试验组 4 实现平均功耗 4.1 W, 较基线模型功耗降低 39.7%, 且设备温度稳定在 48 ± 2 °C(持续运行 1 h 测试)。

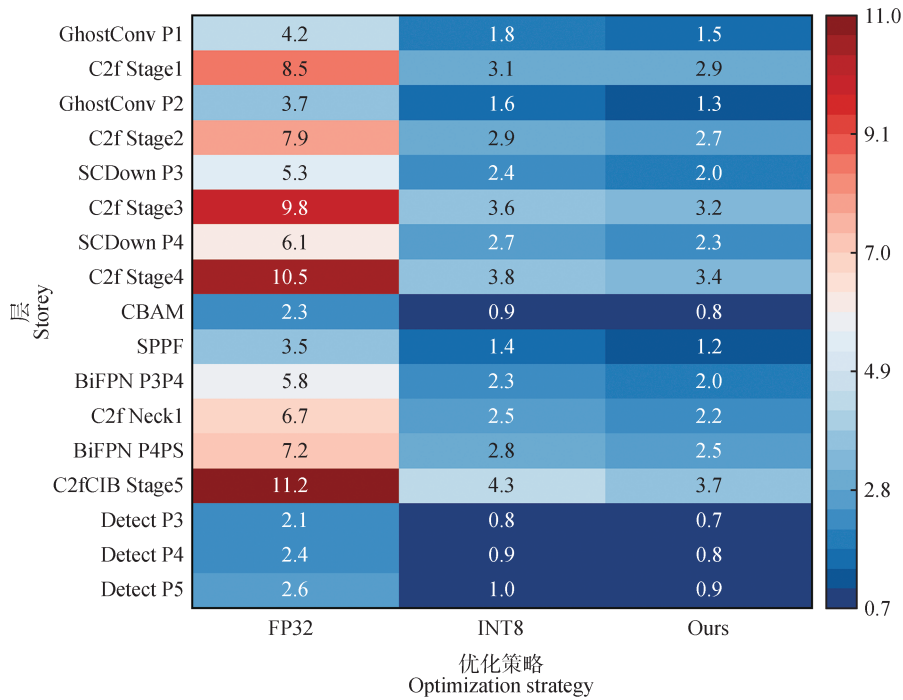


图9 层执行时间热力图

Fig. 9 Layer execution time heatmap

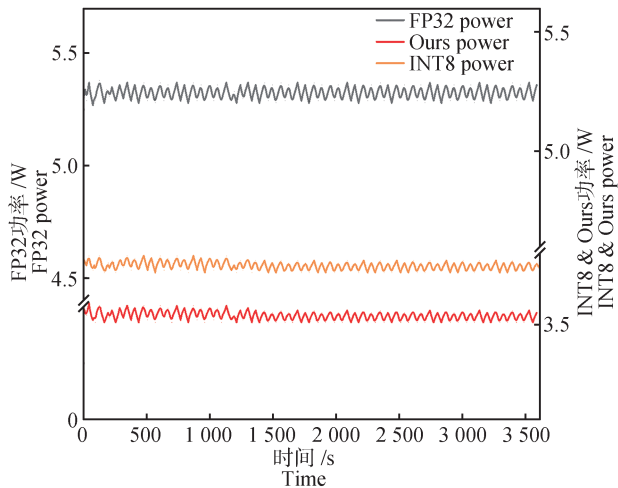


图10 实时功耗监测曲线

Fig. 10 Real-time power consumption monitoring curve

由表5可知,改进后的YOLOv10在NVIDIA 3090 GPU端达到77.5%的平均精度均值,优于主流检测模型。经过INT8量化及层融合优化后(试验组4),部署在Jetson Nano上的模型仍保持75.7%的平均精度均值,较原始YOLOv10n(73.7%)提升2.0%,且推理速度达到43.5 FPS,满足林业巡检实时性需求。

检测结果如图11所示,图中“luo”“huang”“qiu”“yu”和“shui”分别表示落叶松(*L. gmelinii*)、黄檗(*P. amurense*)、胡桃楸(*J. mandshurica*)、榆树(*U. pumila*)和水曲柳(*F. mandshurica*)。由图11可知,与基线模型

表5 部署前后指标变化

Tab. 5 Indicator changes before and after deployment

指标 Index	查准率(P)/% Precision	查全率(R)/% Recall	平均精度均值/% Mean average precision(mAP@0.5)
部署前 Pre-deployment	91.5	75.0	77.5
部署后 Post-deployment	90.8	74.2	75.7

相比,改进后的模型在胡桃楸树冠重叠较为严重的密集区域具有更好的检测效果,并且在黄檗、水曲柳和落叶松等树种的检测中也有更高的置信度。

3 讨论

1)对比试验中,双阶段算法Faster-RCNN其复杂的网络结构查准率高于所有算法,但其136.2 M的参数量与框架依赖性导致显存占用远超边缘设备极限,故难以满足部署要求。SSD的精度瓶颈是基于VGG16的特征提取层感受野固定,难以适应树木尺度多变、枝叶遮挡严重的复杂场景,导致对小目标的漏检率较高。而YOLO系列中早期版本(如v3、v4)通过引入CSPNet和多尺度预测提升检测精度,但参数量与计算成本限制了边缘部署能力,v5—v8系列逐步强化轻量化设计,直至v10版本引入通过端到端架构后,参数量在对比

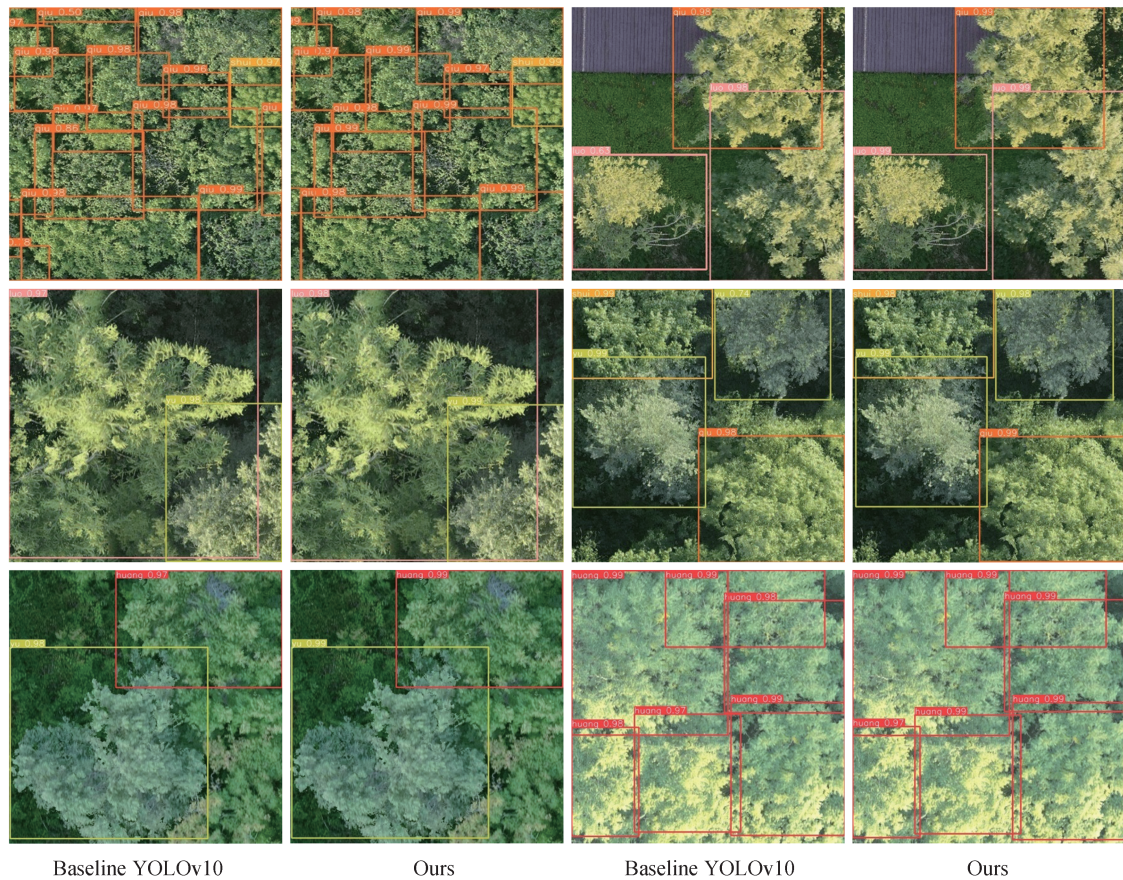


图 11 改进模型可视化检测结果

Fig. 11 Visualization detection results of the improved model

试验中达到最小。

2) 消融试验中, BiFPN 的跨尺度交互虽提升查全率, 但其多分支融合结构导致特征图维度跃升, 在未引入 GhostConv 时, 特征通道的线性增长使计算流带宽需求骤增, 引发 GPU 显存访问延迟, 最终导致 FPS 下降。而 GhostConv 的稀疏卷积机制通过跨通道特征复用, 以此降低主干网络的计算密度推动 FPS 增长。CBAM 单独引入时, 通道注意力通过动态校准通道特征提高了检测平均精度均值。但空间注意力模块应用于高分辨率特征图会产生较大计算开销, 导致单帧推理时延增加。而在联合 GhostConv 后起到了轻量化卷积减少计算冗余的作用。全模块组合的性能大幅提升源于计算路径的深度协同, BiFPN 在高层特征融合时依赖大感受野特征, 而 CBAM 的空间注意力恰好强化了此类特征的边缘梯度, 二者形成互补。GhostConv 则通过动态通道剪枝, 将 BiFPN 中冗余的低响应特征分支剔除, 使跨尺度融合的计算流宽度减少, 最终在参数量持平条件下实现 FPS 提升。

3) 部署试验中, INT8 量化将 FP32 的浮点运算替换

为 8bit 整数运算, 故单层计算密度提升从而提升 FPS。量化与层融合的协同降低了平均功耗, 是由于 INT8 运算将每个 MAC 操作的能耗降低, 而层融合通过减少数据搬运次数, 使 DRAM 功耗占比下降。

4) 尽管模型总体表现良好, 但在一些特定情况仍存在较大的误差。对于不同树种, 落叶松 mAP@0.5 最高(为 79.2%), 主要是由于其树冠呈规则圆锥形, 轮廓清晰。针叶排列紧密, 形成均匀的纹理, 在影像中边缘锐利、对比度高, 便于算法识别。榆树 mAP@0.5 最低(为 75.4%), 则是其主要存在于混交林中, 影像中边缘模糊纹理细节多变, 增加了识别难度, 且其叶片与水曲柳、黄檗等阔叶树相似, 易导致分类混淆。对于不同样地, 混交林的 mAP@0.5 最低(为 72.1%), 主要是多树种混杂导致树冠边界模糊, 遮挡严重, 目标检测困难。胡桃楸林的 mAP@0.5 较低(为 76.3%)是单个树冠覆盖面积广, 可能与其他树冠边缘重叠, 导致检测框定位不精准。对于季节适应性, 数据于温带气候地区的夏秋更迭期间采集, 此时存在轻微叶绿素降解树木变色的情况。通过结合本研究的 HSV 色彩扰动数据增强

策略使得模型在季节适应性上有一定的泛化能力,但模型对夏季完全叶色稳定期或深秋枯黄期的泛化能力仍需多时相数据验证。

4 结 论

本研究对 YOLOv10 算法进行改进,提出一种高效、准确的无人机 RGB 影像树种识别方法,并进行边缘设备环境下的试验验证。通过对算法结构和训练策略的优化,改进后的算法在树种识别任务达到了 91.5% 的查准率和 77.5% 的平均精度均值,优于原版 YOLOv10 和其他主流目标检测算法。在 NVIDIA Jetson Nano 设备上部署时,算法的平均精度均值为 75.7%,平均推理速度为 43.5 FPS,虽然相比部署前性能有所下降,但仍然高于原版 YOLOv10 算法。这一研究结果提高了树种识别的自动化水平,增强了其在复杂自然环境中的适用性,具有潜在的应用前景,特别是在森林资源动态监测、生态监测 and 环境保护领域。但模型仍存在不足,这与训练数据的不平衡、树种数据季节泛化性不足以及图像质量等因素有关。为进一步提升模型的识别精度和鲁棒性,后续的改进可以着重于以下方面,比如,优化数据集、增加多时相数据验证以及改进图像预处理方法,将算法推广到其他林地类型。

参 考 文 献

- [1] 赵勋,岳彩荣. 基于遥感数据的林地变化检测方法研究[J]. 林业资源管理,2019(1):101-108,135.
ZHAO X, YUE C R. Study on forest change detection method based on remote sensing data[J]. Forest Resources Management, 2019(1):101-108,135.
- [2] 吴雪琼,覃先林,周汝良,等. 森林覆盖变化遥感监测方法研究进展[J]. 林业资源管理,2010(4):82-87.
WU X Q, QIN X L, ZHOU R L, et al. Progress of study on forest cover change detection by using remote sensing technique[J]. Forest and Grassland Resources Research, 2010(4):82-87.
- [3] 陈龙伟,周小成,李传昕,等. 基于 UNet-ResNet14 半监督学习的无人机影像森林树种分类[J]. 农业工程学报,2024,40(1):217-226.
CHEN L W, ZHOU X C, LI C X, et al. Classification of tree species based on UNet-ResNet14 semi-supervised learning using UAV images[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024,40(1):217-226.
- [4] 吴晓明. 改进的混合空洞卷积神经网络高光谱影像树种分类算法[D]. 廊坊:北华航天工业学院,2023.
WU X M. Improved hybrid hollow convolutional neural network for tree species classification in hyperspectral images [D]. Langfang: North China Institute of Aerospace Engineering, 2023.
- [5] 陈健昌,陈一铭,刘正军. 激光点云深度学习的树种识别研究[J]. 遥感信息,2022,37(2):105-111.
CHEN J C, CHEN Y M, LIU Z J. Tree species identification based on laser point cloud deep learning[J]. Remote Sensing Information, 2022,37(2):105-111.
- [6] 周湘山,朴虹奕,周杰,等. 基于机器学习算法的无人机高光谱树种分类方法研究[J]. 能源与环境,2023(4):134-137.
ZHOU X S, PIAO H Y, ZHOU J, et al. Research on classification method of drone hyperspectral tree species based on machine learning algorithms[J]. Energy and Environment, 2023(4):134-137.
- [7] 王俊杰,张思媛,滕鹏程. 基于无人机多光谱和 HMLS 的森林树种识别[J]. 黑龙江科技大学学报,2023,33(5):774-778.
WANG J J, ZHANG S Y, TENG P C. Forest tree species identification based on UAV multispectral and HMLS[J]. Journal of Heilongjiang University of Science & Technology, 2023,33(5):774-778.
- [8] 姚扬. 基于多源遥感数据的城市植被优势树种分类[D]. 昆明:云南大学,2022.
YAO Y. Classification of dominant tree species of urban vegetation based on multi-source remote sensing data[D]. Kunming: Yunnan University, 2022.
- [9] NEZAMI S, KHORAMSHAHI E, NEVALAINEN O, et al. Tree species classification of drone hyperspectral and RGB imagery with deep learning convolutional neural networks[J]. Remote Sensing, 2020,12(7):1070.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [C]//Advances in Neural Information Processing Systems 25 (NIPS 2012). Curran Associates Inc., 2012.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA, June 27-30, 2016. IEEE, 2016:770-778.
- [12] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006,313(5786):504-507.
- [13] KNAUER U, VON REKOWSKI C S, STECKLINA M, et al. Tree species classification based on hybrid ensembles of a convolutional neural network (CNN) and random forest classifiers[J]. Remote Sensing, 2019,11(23):2788.
- [14] 周治煊. 基于 FC-DenseNet 的无人机 RGB 影像城市树种识别[D]. 长沙:中南林业科技大学,2023.
ZHOU Z X. Identification of urban tree species in UAV

- RGB images based on FC-DenseNet[D]. Changsha: Central South University of Forestry and Technology, 2023.
- [15] CHEN C, JING L, LI H, et al. Individual tree species identification based on a combination of deep learning and traditional features [J]. *Remote Sensing*, 2023, 15(9):2301.
- [16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA, June 27-30, 2016. IEEE, 2016:779-788.
- [17] GIRSHICK R. Fast R-CNN [C]//2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, December 07-13, 2015. IEEE, 2015:1440-1448.
- [18] WEI L, DRAGOMIR A, DUMITRU E, et al. SSD: single shot MultiBox detector [C]// Computer Vision - ECCV 2016. Springer, Cham, 2016:21-37.
- [19] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6):1137-1149.
- [20] XU S C, WANG R R, SHI W, et al. Classification of tree species in transmission line corridors based on YOLO v7 [J]. *Forests*, 2023, 15(1):61.
- [21] 黄丽明,王懿祥,徐琪,等. 采用YOLO算法和无人机影像的松材线虫病异常变色木识别[J]. *农业工程学报*, 2021, 37(14):197-203.
- HUANG L M, WANG Y X, XU Q, et al. Recognition of abnormally discolored trees caused by pine wilt disease using YOLO algorithm and UAV images [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2021, 37(14):197-203.
- [22] HAN Y H, DUAN B C, GUAN R X, et al. LUFFD-YOLO: A lightweight model for UAV remote sensing forest fire detection based on attention mechanism and multi-level feature fusion [J]. *Remote Sensing*, 2024, 16(12):2177.
- [23] WANG A, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. *arXiv preprint arXiv:2405.14458*, 2024.
- [24] 汪永英,孟琳,韩冬荟,等. 城市森林物种多样性——以哈尔滨城市林业示范基地为例[J]. *东北林业大学学报*, 2017, 45(3):34-38.
- WANG Y Y, MENG L, HAN D H, et al. Species diversity in urban forest—A case study of City Forestry Demonstration Base in Harbin [J]. *Journal of Northeast Forestry University*, 2017, 45(3):34-38.
- [25] HAN K, WANG Y, TIAN Q, et al. GhostNet: More features from cheap operations [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA, June 13-19, 2020. IEEE, 2020:1577-1586.
- [26] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]//Computer Vision - ECCV 2018. Munich, Germany: Springer, 2018:3-19.
- [27] TAN M X, PANG R M, LE Q V. EfficientDet: Scalable and efficient object detection [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA, August 05, 2020. IEEE, 2020:10778-10787.
- [28] ZHORA G. SiLU Loss: More powerful learning for bounding box regression [J]. *arXiv preprint arXiv:2205.12740*, 2022.
- [29] DETTMERS T, LEWIS M, BELKADA Y, et al. LLM.int8(): 8-bit matrix multiplication for transformers at scale [J/OL]. *arXiv preprint arXiv:2208.07339*, 2022.
- [30] MCCANNE S, JACOBSON V. The BSD packet filter: A new architecture for user-level packet capture [C]//UseNIX Winter Conference on Usenix Winter Conference. USENIX Association, 1993.
- [31] 杨堃,范习健,薄维昊,等. 基于视觉加强注意力模型的植物病虫害检测[J]. *南京林业大学学报(自然科学版)*, 2023, 47(3):11-18.
- YANG K, FAN X J, BO W H, et al. Plant disease and pest detection based on visual attention enhancement [J]. *Journal of Nanjing Forestry University (Natural Sciences Edition)*, 2023, 47(3):11-18.
- [32] 孙丰刚,王云露,兰鹏,等. 基于改进YOLOv5s和迁移学习的苹果果实病害识别方法[J]. *农业工程学报*, 2022, 38(11):171-179.
- SUN F G, WANG Y L, LAN P, et al. Identification of apple fruit diseases using improved YOLOv5s and transfer learning [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022, 38(11):171-179.