

基于改进 YOLO11n 的林分小径木识别方法

孔硕, 秦昌浩, 张杨, 赵辉*

(东北林业大学 机电工程学院, 哈尔滨 150040)

摘 要: 针对复杂林分中小径木树干细长、易遮挡且与背景纹理相似导致检测困难的问题, 在自建小径木图像数据集基础上, 提出基于 YOLO11n 构建轻量化改进模型——LDM-YOLO, LDM 是由 LSKA (large separable kernel attention)、DySample 动态上采样与 MBConv-Detect 首字母缩写。方法上, 在 Backbone 中引入双卷积跨阶段部分网络瓶颈-大核可分离卷积注意力模块 (cross-stage partial network bottleneck with 2 convolutions-large separable kernel attention, C3k2-LSKA), 增强细长树干及其上下文特征的表征能力; 在 Neck 中采用 DySample 动态上采样, 细化多尺度特征对齐; 在检测头回归分支中使用移动倒置残差卷积-检测结构 (mobile inverted bottleneck convolution-detection head, MBConv-Detect), 以降低参数量和计算量。试验结果表明, 在自建小径木数据集上, 提出的 LDM-YOLO 模型的 mAP50 和 mAP50-95 分别较 YOLO11n 提升 3.1 和 2.3 个百分点, GFLOPs 从 6.3 G 降至 5.1 G, 在保持推理开销较低的前提下取得更高的检测精度。对比 Faster R-CNN、YOLOv5n、YOLOv8n 和 YOLOv10n 等代表算法, LDM-YOLO 在相近或更低计算复杂度下取得了更优的综合检测性能。研究证明所提出模型在小径木检测任务中能够较好地兼顾精度与轻量化, 为后续面向资源受限平台的林业智能装备应用提供了有益的参考。

关键词: 小径木检测; YOLO11; 大核注意力; 动态上采样; 轻量化; 特征增强; MBConv; 林下复杂场景

中图分类号: S758

文献标识码: A

DOI: 10.7525/j.issn.1006-8023.2026.03.013

Improved YOLO11n-Based Recognition Method for Small-Diameter Trees in Forest Stands

KONG Shuo, QIN Changhao, ZHANG Yang, ZHAO Hui*

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

Abstract: To address the detection difficulty of small-diameter tree trunks in complex forest stands—characterized by slender shapes, frequent occlusion, and high similarity to background textures—this study proposes a lightweight improved model, termed LDM-YOLO, built upon YOLO11n using a self-constructed small-diameter tree image dataset. The acronym LDM is derived from LSKA (large separable kernel attention), DySample dynamic upsampling, and MBConv-Detect. Specifically, a C3k2-LSKA module (cross-stage partial network bottleneck with two convolutions integrated with large separable kernel attention) is introduced into the Backbone to enhance the representation of slender trunks and their contextual features. In the Neck, DySample dynamic upsampling is adopted to refine multi-scale feature alignment. Moreover, an MBConv-Detect structure (mobile inverted bottleneck convolution - detection head) is applied to the regression branch of the detection head to reduce parameter count and computational cost. Experimental results on the self-built dataset show that LDM-YOLO improves mAP50 and mAP50 - 95 by 3.1 and 2.3 percentage points, respectively, compared with YOLO11n, while reducing GFLOPs from 6.3 G to 5.1 G, thereby achieving higher detection accuracy under low inference overhead. Compared with representative methods such as Faster R-CNN, YOLOv5n, YOLOv8n, and YOLOv10n, LDM-YOLO delivers superior overall detection performance at comparable or lower compu-

收稿日期: 2025-12-08

基金项目: 黑龙江省应用技术与开发计划项目 (GA21A403); 2025 年高教强省专项-产教融合研究生联合培养基地建设资助项目 (41502123)。

第一作者简介: 孔硕, 硕士研究生。研究方向为视觉识别。E-mail: 2864553125@qq.com

***通信作者:** 赵辉, 博士, 副教授。研究方向为机器视觉。E-mail: zhaohv@163.com

引文格式: 孔硕, 秦昌浩, 张杨, 等. 基于改进 YOLO11n 的林分小径木识别方法 [J]. 森林工程, 2026, 42(3): 570-582.

KONG S, QIN C H, ZHANG Y, et al. Improved YOLO11n-based recognition method for small-diameter trees in forest stands [J]. Forest Engineering, 2026, 42(3): 570-582.

tational complexity. The results demonstrate that the proposed model better balances accuracy and lightweight design for small-diameter tree detection, providing a useful reference for subsequent forestry intelligent equipment applications on resource-constrained platforms.

Keywords: Small-diameter tree detection; YOLO11; large-kernel attention; dynamic upsampling; lightweight; feature enhancement; MBConv; complex understory scenes

0 引言

在现代可持续林业管理中,对林分结构的优化至关重要。林分中普遍存在胸径较小、树干细长的幼树及下层细小树干(本研究统称为“小径木”),这类个体生长势相对较弱但数量较多,通常被作为选择性砍伐与抚育作业的主要对象,用以调节林木密度、改善林分整体结构,进而促进优势目标木的健壮生长^[1]。然而,传统的人工巡查与样地调查方法效率低下、成本高昂,且难以满足大面积、高精度的监测需求,已成为制约现代林业向自动化、智能化发展的关键瓶颈。近年来,计算机视觉与深度学习技术的飞速发展为解决这一难题提供了全新的技术路径^[2]。

以卷积神经网络(convolutional neural network, CNN)为核心的深度学习模型,能够从海量图像数据中自动学习和提取高维、鲁棒性的特征,摆脱了传统图像处理对人工设计特征的依赖。单阶段目标检测算法YOLO系列^[3]凭借端到端高速检测和较高准确率,已成为林业树干检测的主流选择。例如,Badgujar等^[4]采用YOLOv5对天然牧场环境中的东红杉树干进行检测,经过迁移学习训练的模型在测试中达到约87.8%的精度、84.3%的召回率和88.9%的平均精度,能够识别不同粗细的树干并有效区分树干与树枝。在密林场景中, Da Silva等^[5]比较了多种算法,结果表明YOLOv7及其变体YOLOR在大规模林地树干数据集上取得了最高约90%的F1分数,而且在图形处理器(graphics processing unit, GPU)上仍可实现毫秒级的实时检测。针对更复杂的应用需求, Liu等^[6]在YOLOv7中融入注意力机制并改进损失函数,用于油茶林收获机器人视觉中的树干识别,使模型平均精度提升至89.2%、召回率提高到94%,检测性能显著优于原始YOLOv7模型。Zhang等^[7]在YOLOv8基础上引入金字塔特征增强网络、Transformer模块和坐标注意力机制,提出了PCC-YOLO金字塔增强网络、上下文Transformer模块、坐标-通道融合注意力机制YOLO(a YOLOv8-based model integrating PNet, CoT-Net, and Coord-SE attention, PCC-YOLO)树干检测模型,在果园树干数据集上实现了82.6%的mAP和约143帧/s的检测速度,表现出卓越的精度与效率平衡。Zhang等^[8]

提出了面向林地场景的轻量级YOLOv11-TrunkLight模型,引入新型高效骨干网络和可变形注意力模块,有效提升了树干目标的检测性能。杨灿等^[9]提出空间感知与多尺度特征增强自适应网络YOLO(spatial-aware and FastScaleNet YOLO, SSFYOLO),通过空间感知模块(spatial-aware module, SAM)、多尺度特征增强网络FastScaleNet与加权IoU机制,显著增强了模型在复杂林地小目标检测中的精度与稳定性。胡峻峰等^[10]针对森林移动机器人算力受限与光照变化大的问题,提出基于YOLOv7-tiny的轻量级检测方法,引入部分通道卷积高效层聚合网络(partial efficient layer aggregation network, P-ELAN)与 α -完整交并比(alpha-complete Intersection over union, alpha-CIoU),提升推理速度同时保持检测精度,适配边缘设备部署。在农业无人机低空作业场景下,徐相华等^[11]提出双坐标注意力-单阶段目标检测模型(dual coordinate attention you only look once, DCA-YOLO),引入动态通道注意力与轻量化主干网络,实现复杂背景下对病虫害树木的高效精准识别。此外,针对与树干结构高度关联的树木年轮信息提取任务,王品博等^[12]提出基于YOLOv8的DCW-YOLOv8)可扩张残差注意力模块-轻量级通用上采样算子-动态非单调聚焦机制边界框损失函数的YOLOv8模型(dilation-wise residual-content-aware re-assembly of features-weighted interpolation of sequential evidence for intersection over union, DCW-YOLOv8),实现了年轮实例分割与宽度测量,为林木生长特征自动化识别提供了新思路。

尽管现有研究成果已在树干检测领域取得显著进展,但在处理林分复杂场景,尤其是小径木目标时仍存在挑战。一方面,小径木树干直径较小、形态差异明显,易受枝叶遮挡、光照变化和背景纹理干扰影响,深度学习模型对细小目标的识别鲁棒性仍显不足;另一方面,高精度检测模型多依赖高算力硬件,对于森林巡检机器人、便携式巡护设备或移动端部署而言,其实时性和能耗需求难以同时满足。基于上述问题,为同时提升小径木检测精度与模型可部署性,本研究在YOLO11n的基础上构建改进模型LDM-YOLO:在Backbone中引入双卷积跨阶段部分网络瓶颈-大核可分离卷积注意力模块(cross-stage partial network bottle-

neck with 2 convolutions-large separable kernel attention, C3k2-LSKA)以增强细长树干及其上下文信息建模能力;在 Neck 中采用 DySample 动态上采样缓解特征对齐不准问题;并在检测头中设计轻量化移动倒置残差卷积-检测结构(mobile inverted bottleneck convolution-detection head, MBConv-Detect),以在降低参数量和计算量的同时,提升对小径木树干边界和内部纹理的刻画能力,LDM 是由 LSKA (large separable kernel attention)、DySample 动态上采样与 MBConv-Detect 首字母缩写。

1 LDM-YOLO 模型构建

1.1 YOLO11 网络模型

YOLO11 是 Ultralytics 团队于 2024 年发布的新一代单阶段目标检测模型,提供纳米级模型(nano, n)、小型模型(small, s)、中型模型(medium, m)、大型模型(large, l)、超大型模型(extra-large, x)多种规模以兼顾检测精度与计算开销,整体网络仍由 Backbone、Neck 和 Head 构成。Backbone 以改进的 C3k2 模块替代 C2f, 结合跨阶段部分连接思想与多分支卷积/特征拼接,在控制参数量的同时增强特征重用与梯度传播;末端引入简化空间金字塔池化快速模块(simplified spatial pyramid pooling fast, SPPF)与 C2PSA 模块以获取多尺

度上下文并强化空间注意,从而提升复杂背景和小目标的表征能力。Neck 采用改进的路径聚合网络-特征金字塔网络结构(path aggregation network-Feature pyramid network, PAN-FPN),通过上采样、下采样及特征拼接实现多尺度特征融合,并在各尺度分支继续使用 C3k2 以优化跨层信息传递。Head 采用无锚设计,在多尺度特征图上直接进行回归与分类预测,由 C3k2 与标准卷积(conv-batchnorm-SiLU, CBS)组合,对各尺度输出边界框参数、置信度和类别概率,实现高效端到端检测。

1.2 LDM-YOLO 模型

在基线网络 YOLO11n 的基础上,本研究结合小径木“细长、易遮挡、与背景相似”的特点,从主干网络、特征融合和检测头 3 个层面对模型进行了有针对性地改进:在 Backbone 中将部分 C3k2 单元替换为 C3k2-LSKA 模块,以增强对细长树干及其上下文信息的建模能力;在 Neck 中用 DySample 动态上采样替代原有固定插值,上采样过程中更好地对齐多尺度特征;在解耦检测头中设计轻量化 MBConv-Detect 结构,在控制参数量和计算量的同时提升对小径木边界和内部纹理的刻画能力。整体网络结构见图 1,改进后的 LDM-YOLO 在保持推理速度基本不变的前提下,使模型在遮挡和密集林下场景中的小径木检测精度得到进一步提升。

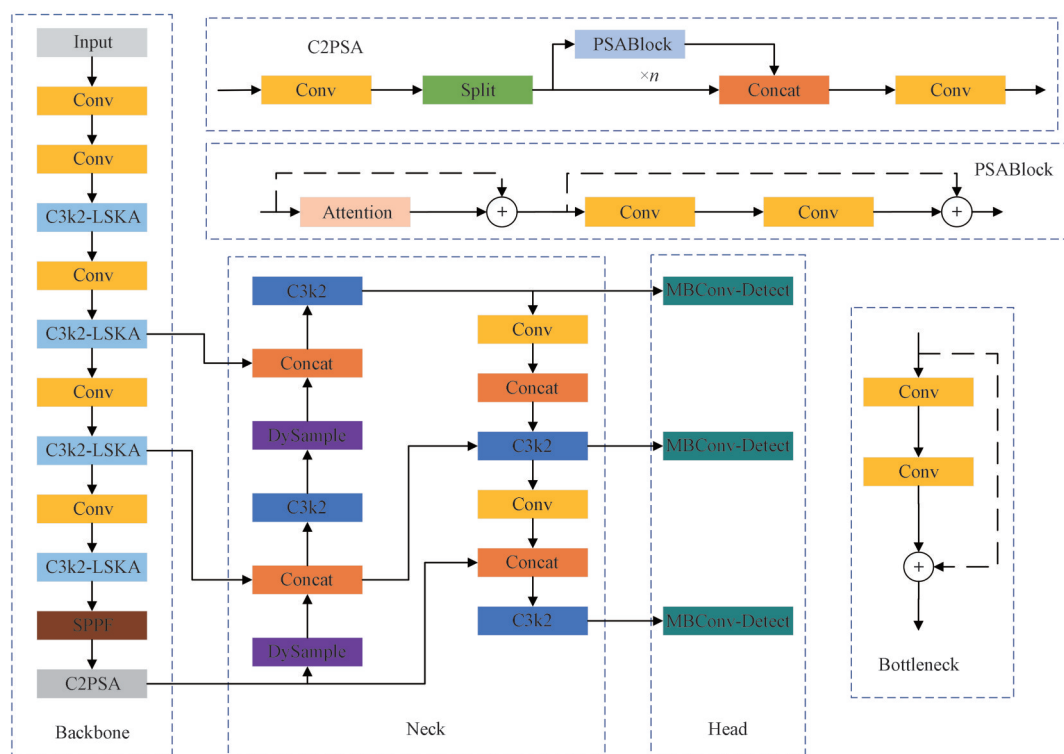


图 1 LDM-YOLO 整体结构

Fig. 1 Overall architecture of LDM-YOLO

1.2.1 LSKA 模块

小径木目标在图像中呈现出极端的纵向细长形态,且常被林下灌木或枝叶沿树干方向不规则遮挡,导致目标特征在空间上呈现非连续的断裂状态。传统的小卷积核由于感受野受限,难以捕获树干的长距离上下文依赖,极易将断裂的树干误判为背景噪声或多个独立目标。为此,本研究在YOLO11n主干网络中引入大核可分离卷积注意力(large separable kernel attention, LSKA)模块^[13],对部分C3k2基本单元进行替换,

构建C3k2-LSKA结构。LSKA通过大核卷积模拟人类视觉对细长目标的整体感知机制,能够在更大的空间范围内聚合上下文信息,从而有效跨越局部遮挡区域,将被打断的树干特征重新关联,增强模型对细长目标形态的完整性感知。其整体结构见图2。LSKA本质上是一种“大核+可分离深度卷积+空间注意力”的结构,通过将二维 $k \times k$ 大核分解为级联的 $1 \times k$ 与 $k \times 1$ 一维核,在近似保持大感受野的同时有效降低参数量和GFLOPs。

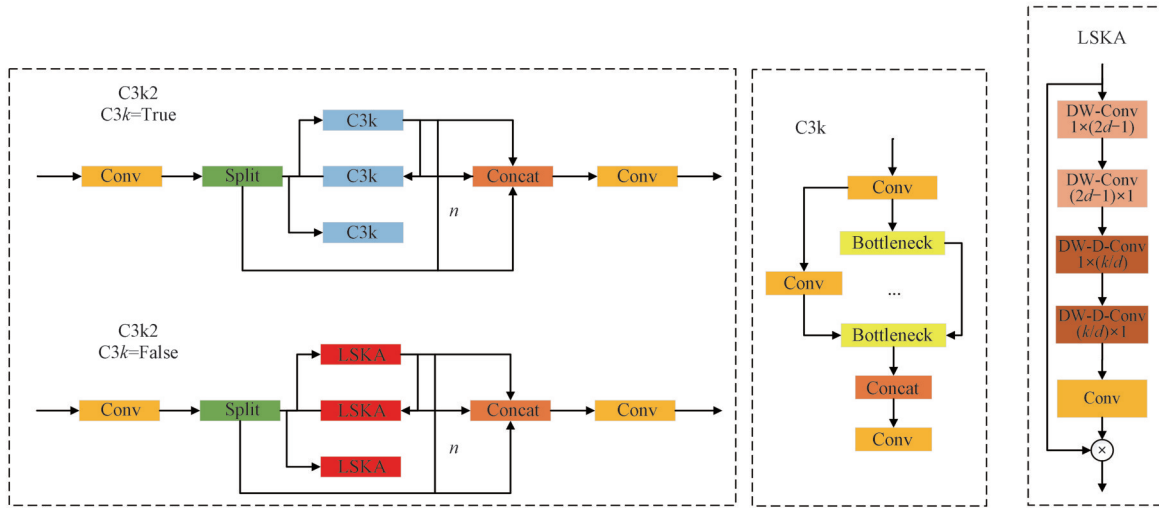


图2 C3k2-LSKA 结构

Fig. 2 Structural diagram of the C3k2-LSKA module

设输入特征为 F , LSKA首先采用一对水平方向和垂直方向的深度卷积提取局部上下文,再进一步结合空洞深度卷积建模长程依赖关系。以第 c 个通道为例,其计算过程可表示公式为

$$\tilde{Z}^c = W_{(2d-1) \times 1}^c * (W_{1 \times (2d-1)}^c * F^c). \quad (1)$$

$$Z^c = W_{k/d \times 1}^c * (W_{1 \times k/d}^c * \tilde{Z}^c). \quad (2)$$

$$A^c = W_{1 \times 1}^c * Z^c. \quad (3)$$

$$\hat{F}^c = A^c \odot F^c. \quad (4)$$

式中: c 为通道索引; F^c 表示第 c 个通道特征; $*$ 表示卷积运算; $\odot$ 表示Hadamard逐元素乘积; $W_{a \times b}^c$ 为第 c 个通道对应的卷积核,尺寸为 $a \times b$; $\tilde{Z}^c$ 为经水平和垂直深度卷积后的中间特征; Z^c 为进一步结合空洞深度卷积后的特征; A^c 为生成的注意力权重; $\hat{F}^c$ 为加权后的输出特征; k 为最大感受野尺寸参数; d 为空洞率。

上述过程等价于使用一个大核深度卷积生成空间注意力权重 A^c ,再与特征 F^c 相乘实现自适应特征重标定。由于二维大核被分解为4个一维核,LSKA的参数量和计算复杂度随 k 近似线性增长,相比直接使用 $k \times$

k 深度卷积或传统大核注意力(large kernel attention, LKA)^[14]的二维大核,可显著降低GFLOPs和显存占用。

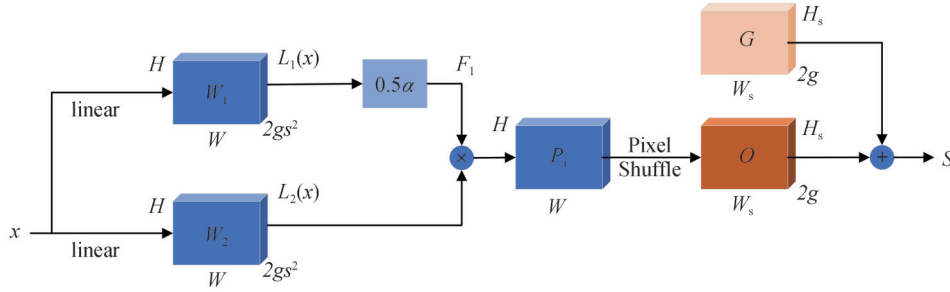
在具体实现上,对于浅层特征,将C3k2中的残差块替换为LSKA注意力单元,利用大核可分离卷积,该模块能够在较高分辨率上有效捕获树干的纵向延展与弯曲形态;对于深层特征,考虑到通道数显著增加、计算成本更为敏感,则保留原有的C3k2结构,仅在浅层引入LSKA,以在检测精度与推理效率之间取得平衡。整体而言,C3k2-LSKA既保持了C3k2良好的梯度传递和特征复用能力,又通过大核可分离注意力机制增强了对细长小径木目标的形状建模与全局上下文感知能力。

1.2.2 动态上采样模块 DySample

在小径木检测场景中,树干往往细长、弯曲且尺度差异明显,简单的最近邻或双线性上采样容易造成细窄树干被模糊、断裂,难以兼顾粗壮树干与细小枝干的结构细节。在特征金字塔的融合过程中,小径木目标由于像素占比极低,对上采样算法的插值误差极为敏

感。传统的最近邻或双线性插值采用固定的采样核,容易造成细微的树干纹理被平滑模糊,甚至导致细小目标在深层特征映射回高分辨率时发生结构丢失。为此,引入动态上采样模块 DySample^[15]替代固定插值。DySample 的核心优势在于其内容感知机制,能够根据

小径木树干的纹理分布自适应调整采样点位置,相当于沿着树干的边缘与骨架进行精准重采样。这种机制能有效避免细小目标边缘的模糊化,最大限度地保留小径木在复杂背景下的高频结构信息。DySample 的整体结构见图 3。



x 为输入特征图; H 、 W 分别为输入特征图的高度与宽度; linear 表示线性映射; W_1 、 W_2 为两分支可学习权重; 输出分别为 $L_1(x)$ 与 $L_2(x)$, 其通道维为 $2gs^2$ (2 为 x/y 两方向偏移, g 为分组数, s 为上采样倍率, s^2 对应 $s \times s$ 子像素位置); α 为缩放系数, 0.5α 为缩放因子; F_1 为该分支产生的缩放信息; P_i 为中间特征; PixelShuffle (像素重排) 后得到的偏移为 O , 其空间尺寸为 $H_s \times W_s$, 其中: H_s 、 W_s 为上采样后的高度和宽度; G 为原始网格; S 为最终采样集合。

x denotes the input feature map; H and W are the height and width of the input feature map, respectively; linear indicates a linear projection; W_1 and W_2 are learnable weights of the two branches; outputs are $L_1(x)$ and $L_2(x)$, respectively, with a channel dimension of $2gs^2$ (2 for x/y offsets, g for the number of groups, s for the upsampling scale, and s^2 for $s \times s$ sub-pixel locations); α is the scaling coefficient, 0.5α is the scaling factor, and F_1 denotes the scaling information generated by this branch; P_i denotes the intermediate feature; the offset after PixelShuffle is O with a spatial size of $H_s \times W_s$, where H_s and W_s denote the height and width after upsampling, respectively; G denotes the base grid; and S is the final sampling set.

图 3 DySample 结构

Fig. 3 Structure diagram of the DySample module

对于输入特征图 x , DySample 采用双分支线性变换分别建模“动态缩放”和“静态分布”两类信息。2 条支路的线性映射可表示为

$$L_1(X) = W_1 X + b_1. \quad (5)$$

$$L_2(X) = W_2 X + b_2. \quad (6)$$

式中: $L_1(X)$ 更关注局部高频、变化剧烈的区域; $L_2(X)$ 则描述整体较平滑的静态特征分布; b_1 与 b_2 为 2 条线性映射的偏置项, 用于为分支输出提供可学习的常数平移, 使线性变换的输出不必经过原点, 从而调节特征响应的基线水平并提高映射的表达能力和训练稳定性; 为控制动态偏移量的幅度, 对动态分支进行抑制后与静态分支融合, 得到中间特征 P_i , 其计算公式为

$$F_1 = 0.5\alpha \odot L_1(X). \quad (7)$$

$$P_i = F_1 + L_2(X). \quad (8)$$

这一设计一方面利用 $L_1(X)$ 强调边缘和纹理等细节, 另一方面通过 $L_2(X)$ 保持整体结构的稳定性, 两者在 P_i 中实现互补, 有利于在后续上采样过程中避免偏移量过大导致采样位置发散。随后, DySample 依托像素重排 (PixelShuffle) 完成亚像素级上采样, 将通道维度的冗余映射为更高的空间分辨率。由中间特征 P_i

得到的静态偏移可表示公式为

$$O = \text{PixelShuffle}(P_i, s). \quad (9)$$

式中: O 为由 PixelShuffle 将中间特征 P_i 重排到高分辨率空间后得到的偏移。

为获得最终采样位置集合, 将原始网格 G 与偏移 O 相加得到采样集合 S , 其计算公式为

$$S = O + G. \quad (10)$$

在 LDM-YOLO 中, 将预测头 $P4 \rightarrow P3$ 、 $P5 \rightarrow P4$ 这 2 条上采样路径的固定上采样算子替换为 DySample, 以在保持轻量级结构的前提下, 使网络能够根据小径木目标的空间分布自动调整采样位置。

1.2.3 基于 MBConv 的轻量解耦检测头

由图 4 可知, 原始 YOLO11n 检测头采用解耦结构: 上分支由 2 层 3×3 卷积和 1 层 1×1 卷积构成, 用于回归边界框; 下分支由交替堆叠的 3×3 深度可分离卷积 (DWConv) 和 1×1 卷积构成, 用于类别预测。该设计能够在一定程度上实现回归与分类的任务解耦, 但回归分支仍大量使用普通卷积, 参数量和计算量相对较大, 对细长、弯曲的小径木边界的建模能力仍有进一步提升空间。

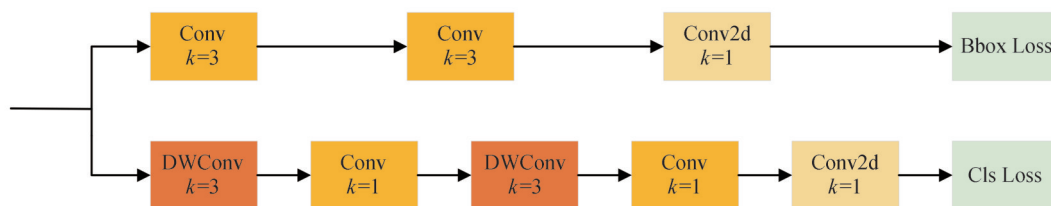


图4 YOLO11n检测头结构

Fig. 4 Architecture of the YOLO11n detection head

为在保持检测头轻量化的同时增强回归分支对小径木目标的表征能力,本研究引入EfficientNet系列中的MBConv结构^[16],对YOLO11n的检测头进行改造。MBConv采用“扩展—深度卷积—压缩”的倒残差设计:首先通过 1×1 卷积在通道维度上对特征进行扩展,随后使用 3×3 深度可分离卷积提取空间细节,并在中间

嵌入SE通道注意力模块以实现特征重标定,最后通过 1×1 卷积完成通道压缩,并与输入进行残差连接,见图5。该结构一方面利用深度可分离卷积显著降低参数量和GFLOPs,另一方面借助SE与残差连接提升有效感受野和梯度传递能力,兼顾了轻量化与表达能力。

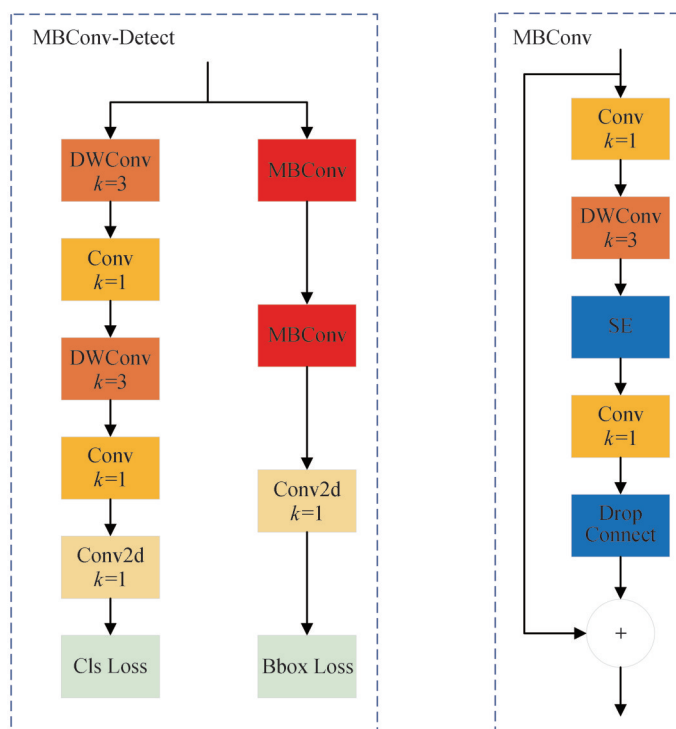


图5 MBConv-Detect与MBConv结构

Fig. 5 Architectures of the MBConv-Detect and MBConv modules

基于上述思想,本研究构建了MBConv-Detect检测头。在林下复杂环境中,小径木树皮纹理往往与背景灌木高度相似,导致特征辨识度低。MBConv结构中的SE注意力机制能够对特征通道进行自适应重标定,抑制背景噪声通道的响应,增强包含树干纹理特征的通道权重,从而在特征层面实现“去伪存真”。同时,倒残差结构在保持低参数量的同时维持了深层特征的非线性表达能力。在回归分支中引入MBConv,不仅大幅降低了检测头的计算负载,更通过特征重标定机制提

升了对微弱纹理目标的边界回归精度,实现了轻量化与高精度的双重优化。

2 材料与设置

2.1 小径木图像数据集构建

本研究于2025年9—11月在东北林业大学城市林业示范基地($45^{\circ}43'25.50''\text{N}$, $126^{\circ}37'51.05''\text{E}$)开展多批次实地采集,使用配备高分辨率摄像头的智能手机(iPhone 15)进行地面近景拍摄。采集过程中重点覆盖

生长弯曲、胸径较小和密集成丛等典型小径木个体,并兼顾不同光照条件与背景复杂度,以增强数据集对实际作业场景的代表性。对原始图像进行人工筛选,剔除模糊、严重遮挡及目标轮廓不清晰等低质量样本,最终保留 1 600 张高质量小径木图像作为有效原始样本。为了明确任务边界并保证试验的可复现性,本研究结合林业幼林抚育需求与计算机视觉任务特性,制定了小径木视觉检测标准:定义检测对象为地径或胸径在 1~4 cm 的林下幼树及灌木状细小林木;视觉特征层面,仅对图像中纹理特征可辨且未被枝叶完全遮挡的树干区域进行标注。对于远处背景中像素宽度过小或因运动模糊导致轮廓严重缺失的极微小目标,视为背景噪声不予标注,以降低标签的不确定性对模型训练

的干扰。使用 LabelImg 对全部图像逐张标注,将小径木树干作为唯一检测目标,以矩形边界框进行标记,并按 YOLO 格式生成标注文件,类别统一命名为“小径木(sdtree)”。

为提升模型在不同成像条件下的鲁棒性与泛化能力,对原始样本实施数据增强,包括水平翻转以模拟视角变化,图像锐化以强化树干边缘与纹理细节,以及适度调整饱和度和色调以模拟不同光照与环境色温。经增强后,数据集规模由 1 600 张图像扩展至 3 044 张图像,并按 8:1:1 的比例随机划分为训练集、验证集和测试集,用于后续模型训练与性能评估。部分小径木图像样本见图 6。



图 6 小径木图像样本

Fig. 6 Sample images of small-diameter trees

2.2 试验环境与参数设置

本研究的试验硬件配置为 Intel Core i9-13900HX 处理器、32 GB 内存、NVIDIA GeForce RTX 5070 Ti 独立显卡,操作系统为 Windows 11。软件环境采用 Python 3.11.13、PyTorch 2.8 与 CUDA 12.8,目标检测算法基于 Ultralytics YOLO 框架实现。训练时,采用官方

提供的 YOLO11n 预训练权重作为初始化参数,以加快收敛并提升特征表达能力。输入图像尺寸为 640×640 像素, batchsize 设为 16, 训练轮数为 150, 初始学习率为 0.01, 优化器采用随机梯度下降(stochastic gradient descent, SGD), 其余超参数设置遵循 Ultralytics YOLO 的默认配置。

2.3 评估指标

本研究采用精度(precision, P)、召回率(recall, R)、平均精度(average precision, AP, 式中记为 A_p)和平均精度均值(mean average precision, mAP, 式中记为 m_{AP})评价各模型在小径木检测任务中的综合性能,并同时给出 mAP50 与 mAP50-95。精度表示被判为正类的样本中真正例的比例,召回率表示所有真正例中被正确检出的比例,二者均可由混淆矩阵中的 True Positive(式中记为 T_p)、False Positive(式中记为 F_p)和 False Negative(式中记为 F_n)计算得到。

在目标检测中,AP 表征某一类别在不同召回率水平下的整体检测性能,定义为精度-召回率(P - R)曲线在召回率区间上的面积;mAP 为各类别 AP 的平均值,是衡量模型总体检测效果的核心指标。本研究中, mAP50 对应 IoU 阈值为 0.5 时的 mAP, mAP50-95 则在 IoU 为 0.5-0.95(步长 0.05)下分别计算 AP 后取平均,用以综合反映不同重叠要求下的检测性能。为刻画模型复杂度与推理开销,本研究同时报告参数量(params)、浮点运算量(GFLOPs)及模型权重大小(model size),分别用于衡量可学习参数规模、单次前向推理的近似计算量以及模型的存储占用。

本研究中上述评价指标的计算公式为

$$P = \frac{T_p}{T_p + F_p} \quad (11)$$

$$R = \frac{T_p}{T_p + F_n} \quad (12)$$

$$A_p = \sum_{k=1}^K P_k \Delta R_k \quad (13)$$

$$m_{AP} = \frac{1}{N_{cls}} \sum_{i=1}^{N_{cls}} AP_i \quad (14)$$

式中: P_k 为第 k 个采样点的精度; ΔR_k 为相邻 2 点召回率增量; N_{cls} 为类别数; i 为类别索引; AP_i 为第 i 类目标的平均精度。

3 结果与分析

3.1 模型训练收敛性分析

为了验证改进模型在训练过程中的收敛性能与泛化能力,本研究记录了 YOLO11n 基线模型与 LDM-YOLO 改进模型的训练损失与验证损失变化曲线,见图 7。由图 7(a)的训练损失曲线可以看出, LDM-YOLO 在训练初期的下降速度快于基线模型,且在整个训练周期内始终保持更低的损失值,表明引入的 C3k2-LSKA 模块有效加速了网络对小径木特征的提取与拟合。由图 7(b)的验证损失曲线可见, 2 种模型均在训练至 80 轮左右进入相对平稳的收敛阶段,且未出现验证损失持续上升的过拟合现象。值得注意的是,在第 140 轮处,由于训练策略关闭了 Mosaic 数据增强, 2 条曲线均出现了预期的短暂波动并迅速回落,进一步证明了模型对真实数据分布的适应能力。相比基线模型, LDM-YOLO 的验证损失更低,说明改进后的模型在非结构化林业场景中具有更强的鲁棒性。

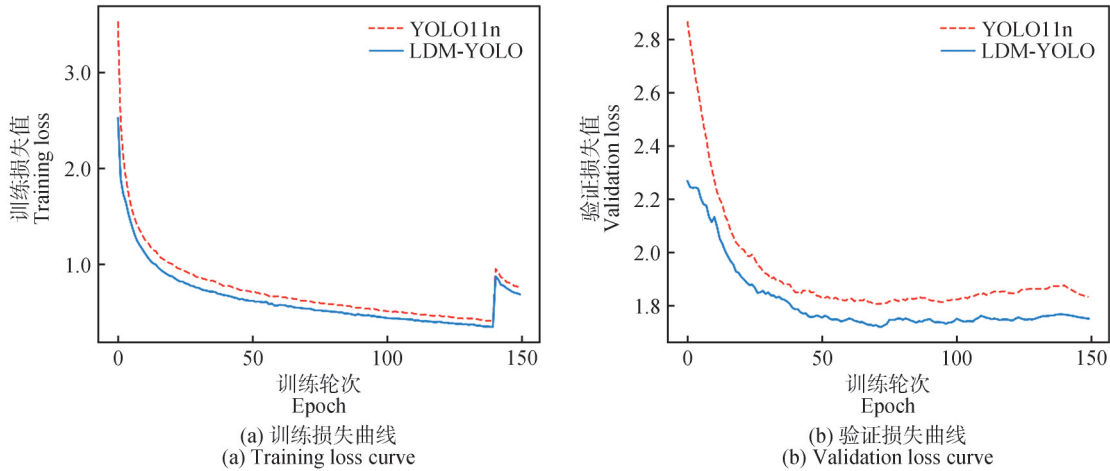


图7 训练损失与验证损失对比曲线

Fig. 7 Comparison curves of training loss and validation loss

3.2 消融试验

为系统评估 3 种改进模块对 YOLO11n 在小径木检测任务上的作用,现对表 1 中的消融试验结果进行分析。表 1 中第 1 组为未引入任何改进模块的

YOLO11n 基线模型,第 2 至第 6 组在此基础上按列打“√”表示启用了相应模块。与基线相比,单独加入 C3k2-LSKA(第 2 组)和 DySample(第 3 组)均能显著提升检测精度: P 分别提高至 75.9% 和 77.9%, mAP50 提

升到 77.6% 和 77.1%, mAP50-95 也增加了 1.8 和 2.7 个百分点,说明大卷积核注意力与动态上采样均有利于在复杂林分背景下强化小径木目标的特征表达与空间细节;引入 MBConv-Detect(第 4 组)时, mAP 基本保

持不变,而 GFLOPs 由 6.3 G 降至 5.1 G、参数量由 2.58 M 降至 2.26 M、权重由 5.2 MB 缩减到 4.6 MB,表明该检测头在保证精度的前提下有效实现了模型轻量化。

表 1 消融试验
Tab. 1 Ablation experiment

组 Group	C3k2-LSKA	DySample	MBConv-Detect	准确率/% Precision	召回率/% Recall	平均精度均值/% mAP50	高平均精度均值/% mAP50-95	浮点运算量/G GFLOPs	参数量/M Params	权重/MB Model size
1				71.7	71.5	74.3	36.9	6.3	2.58	5.2
2	√			75.9	71.2	77.6	38.7	6.2	2.57	5.2
3		√		77.9	69.1	77.1	39.6	6.3	2.59	5.2
4			√	73.2	68.2	74.2	37.9	5.1	2.26	4.6
5	√	√		76.6	71.6	78.0	39.3	6.3	2.59	5.2
6	√	√	√	74.9	72.0	77.4	39.2	5.1	2.27	4.6

组合试验中, C3k2-LSKA 与 DySample 同时启用(第 5 组)使 mAP50 达到 78.0% 的最高值,验证了二者在全局建模与上采样细节方面的互补性;在此基础上进一步加入 MBConv-Detect(第 6 组)后, *P*、*R* 以及各项 mAP 指标仍整体优于基线,同时保持 GFLOPs 的 5.1 G 和 2.27 M 参数量,实现了检测性能与模型复杂度之间较为理想的折中,表明三种模块在 Backbone、Neck 与 Head 这三个层面形成协同作用,使改进后的 YOLO11n 在小径木检测任务上兼顾了精度提升与轻量化需求。

3.3 对比试验

为进一步评估所提出 LDM-YOLO 模型在小径木检测任务中的综合性能,本研究选取 Faster R-CNN^[17]、YOLOv5n、YOLOv8n、YOLOv10n 以及未改动的 YOLO11n 作为对比模型,在相同数据集划分、输入分辨率和训练策略下进行试验,结果见表 2。由表 2 可以看出, LDM-YOLO 在所有模型中取得了最高的精度和召回率, mAP50 和 mAP50-95 也分别达到 77.4% 和 39.2%, 相较于 YOLO11n 分别提升 3.1 和 2.3 个百分点。

表 2 对比不同的模型在数据集集中的表现
Tab. 2 Comparison of the performance of different models in the dataset

模型 Model	准确率/% Precision	召回率/% Recall	平均精度均值/% mAP50	高平均精度均值/% mAP50-95	浮点运算量/G GFLOPs	参数量/M Params	权重/MB Model size	推理时间/ms Time	检测速度/(帧·s ⁻¹) FPS
Faster R-CNN	70.8	70.9	76.6	34.9	133.92	41.3	158.1	29.8	34
YOLOv5n	71.8	68.2	74.1	32.8	5.8	2.18	4.4	5.1	195
YOLOv8n	72.2	70.0	73.0	34.6	6.8	2.68	5.4	5.6	180
YOLOv10n	69.7	61.7	68.2	32.8	8.2	2.69	5.5	6.9	140
YOLO11n	71.7	71.5	74.3	36.9	6.3	2.58	5.2	6.3	158
LDM-YOLO	74.9	72.0	77.4	39.2	5.1	2.27	4.6	4.8	208

与两阶段检测器 Faster R-CNN 相比,在显著降低计算复杂度的前提下优势更为明显: Faster R-CNN 的 GFLOPs 与参数量高达 133.92 G 和 41.3 M, 而 LDM-YOLO 仅为 5.1 G 和 2.27 M, 却在 mAP50 与 mAP50-95 上分别高出 0.8 和 4.3 个百分点,说明基于 C3k2-LSKA、DySample 与 MBConv-head 的改进结构能够在

复杂林分场景中更有效地表征小径木目标。与轻量化单阶段检测器相比, LDM-YOLO 在保持类似参数规模的同时也表现出更为优越的精度表现。相较于 YOLOv8n, 本模型在 mAP50 和 mAP50-95 上分别提升了 4.4 和 4.6 个百分点。这主要归因于 YOLOv8n 采用的标准 C2f 模块和最近邻上采样在处理极细树干时,

容易因感受野受限和插值锯齿效应导致纹理细节丢失,而LDM-YOLO的LSKA大核注意力和DySample动态上采样机制则能更精准地保留细长目标的边缘结构。此外,针对YOLOv10n的试验结果显示,尽管其GFLOPs高达8.2 G,但mAP50却比LDM-YOLO低9.2个百分点。深入分析可知,YOLOv10n虽然引入了PSA大核卷积与非极大值抑制(non-maximum suppression, NMS)策略,但其通用特征提取结构在应对纹理匮乏的小径木时出现了计算资源错配现象,大量算力被消耗在背景特征的冗余计算上,未能像LDM-YOLO一样实现对细长前景目标的算力聚焦。

针对资源受限平台的部署需求,本研究在统一640×640输入分辨率下进行了推理速度测试。LDM-YOLO的单张推理耗时为4.8 ms,对应的检测速度约为208 帧/s。与基线模型YOLO11n相比,推理时间减少了约1.5 ms;与Faster R-CNN相比,推理速度有显著优势。

综合来看,LDM-YOLO在保证模型轻量化的前提下,实现了在小径木数据集上的最高精度和较高的定位准确度,验证了所提出的3种结构改进在实际小径木检测场景中的有效性与实用价值。

为进一步探究改进模型在具体类别上的识别优势,本研究绘制了YOLO11n与LDM-YOLO在测试集上的归一化混淆矩阵,见图8。混淆矩阵的对角线数值代表模型对该类别的识别准确率(recall),数值越高表明漏检越少。由图8可知,基线模型YOLO11n对小径木(sdtree)类别的识别率为0.72,这意味着有28%的真实小径木目标被误判为背景。而改进后的LDM-YOLO模型将小径木的识别率提升至0.77,漏检率降低了5个百分点。这一提升主要得益于DySample动态上采样与MBConv-Detect检测头的引入,有效增强了模型对细小、被遮挡树干目标的特征对齐与边界感知能力,使其在纹理复杂的林下背景中能更精准地分离出前景目标。

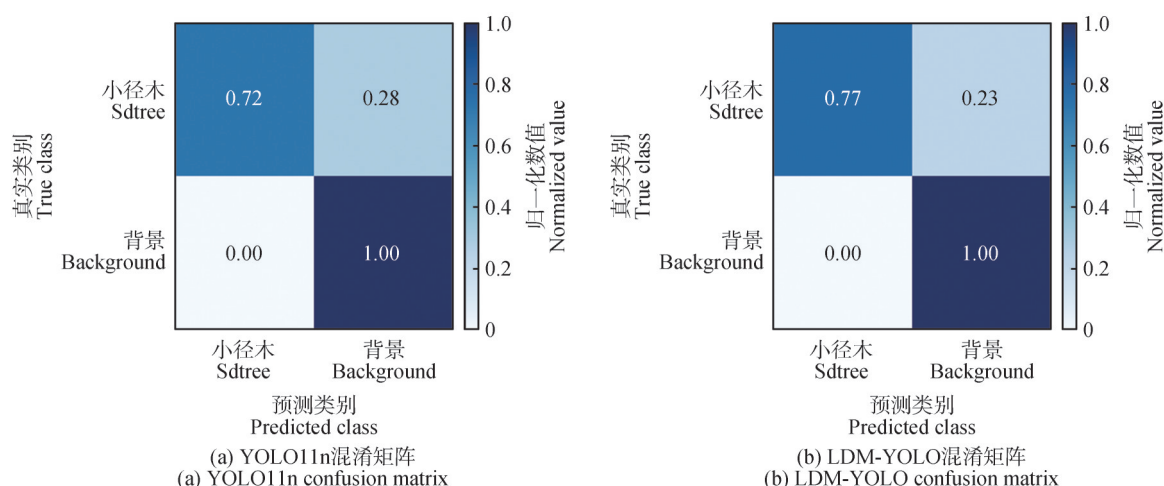


图8 模型混淆矩阵对比

Fig. 8 Comparison of model confusion matrices

3.4 结果可视化分析

为了直观揭示模型在复杂林分环境下对小径木目标的关注区域,本研究采用了一种由检测先验驱动的特征可视化方法。该方法首先在推理阶段提取网络关键层的深层特征响应,并将其映射为空间激活强度;然后利用非线性增强技术突出显著特征区域的对比度;最后结合模型预测的边界框信息引入空间位置约束,对原始激活图进行加权融合。该处理策略能够有效滤除非目标区域的背景噪声干扰,从而更具针对性地呈现模型对小径木目标主体特征的聚焦效果。

对3幅典型场景进行热力图可视化,见图9。其中,第1行为原始图像,第2行为YOLO11n的热力图结

果,第3行为LDM-YOLO的热力图结果,颜色越接近红色表示响应越强。从第1列、第2列可以看到,2种模型都能在目标树干区域形成纵向高响应带,但LDM-YOLO的高响应更加连续,热点集中在树干主体及其分叉处,背景灌木和地面叶片区域基本被压制为冷色;相比之下,YOLO11n的响应分布略显发散,部分能见度较低的小径木只呈现出零散的弱响应。在第3列(复杂背景、遮挡较重)中差异更为明显:YOLO11n在目标位置几乎没有明显高响应,整体呈冷色,容易造成漏检;而LDM-YOLO仍能在细长树干上形成清晰的纵向暖色条带,说明改进后的网络在林下杂乱环境中,对细小、背景相近的小径木具有更强的聚焦能力和特征分离能力。

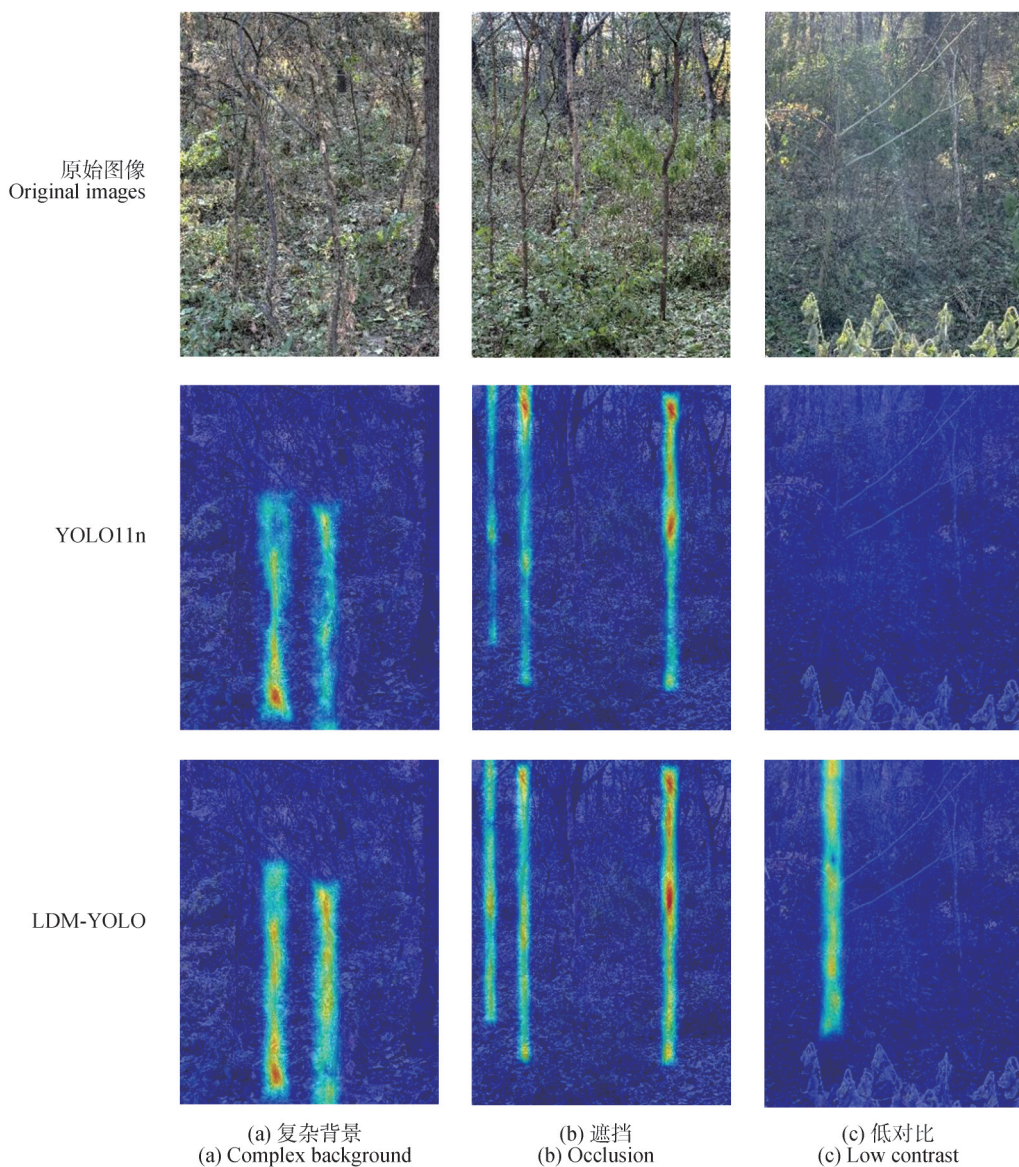


图9 热力图可视化

Fig. 9 Heat map visualization

为了对比不同模型在真实林下环境中的检测表现,本研究选取了“简单场景、遮挡场景和密集场景”3幅代表性图像进行可视化,见图10。整体来看,YOLOv5n、YOLOv8n、YOLOv10n、YOLO11n和LDM-YOLO均能在3种场景下检测出主要小径木目标,但置信度和框的位置精度存在明显差异:YOLOv5n、YOLOv8n在遮挡和密集场景中部分目标的置信度偏低,YOLOv10n在密集场景中出现了小径木漏检现象。相比之下,YOLO11n与LDM-YOLO的检测框更贴合树干轮廓,其中LDM-YOLO在3种场景下均表现出更高且更稳定的置信度,说明改进模型在复杂背景、遮挡和多目标条件下具有更好的检测可靠性。

4 结论

本研究针对复杂林分环境中小径木“细长、易遮挡、与背景相似”等特征带来的检测难题,在YOLO11n的基础上构建了轻量化改进模型LDM-YOLO。通过在Backbone中引入C3k2-LSKA模块,在Neck中采用DySample动态上采样,以及在检测头回归分支中使用MBConv-Detect结构,模型在多尺度特征表达、细长树干边界分析和参数压缩等方面得到协同优化。试验及对比结果表明,LDM-YOLO在mAP50和mAP50-95上均较基线模型与典型对比算法取得了更优性能,同时降低了GFLOPs和参数量。这证明了本研究提出的改

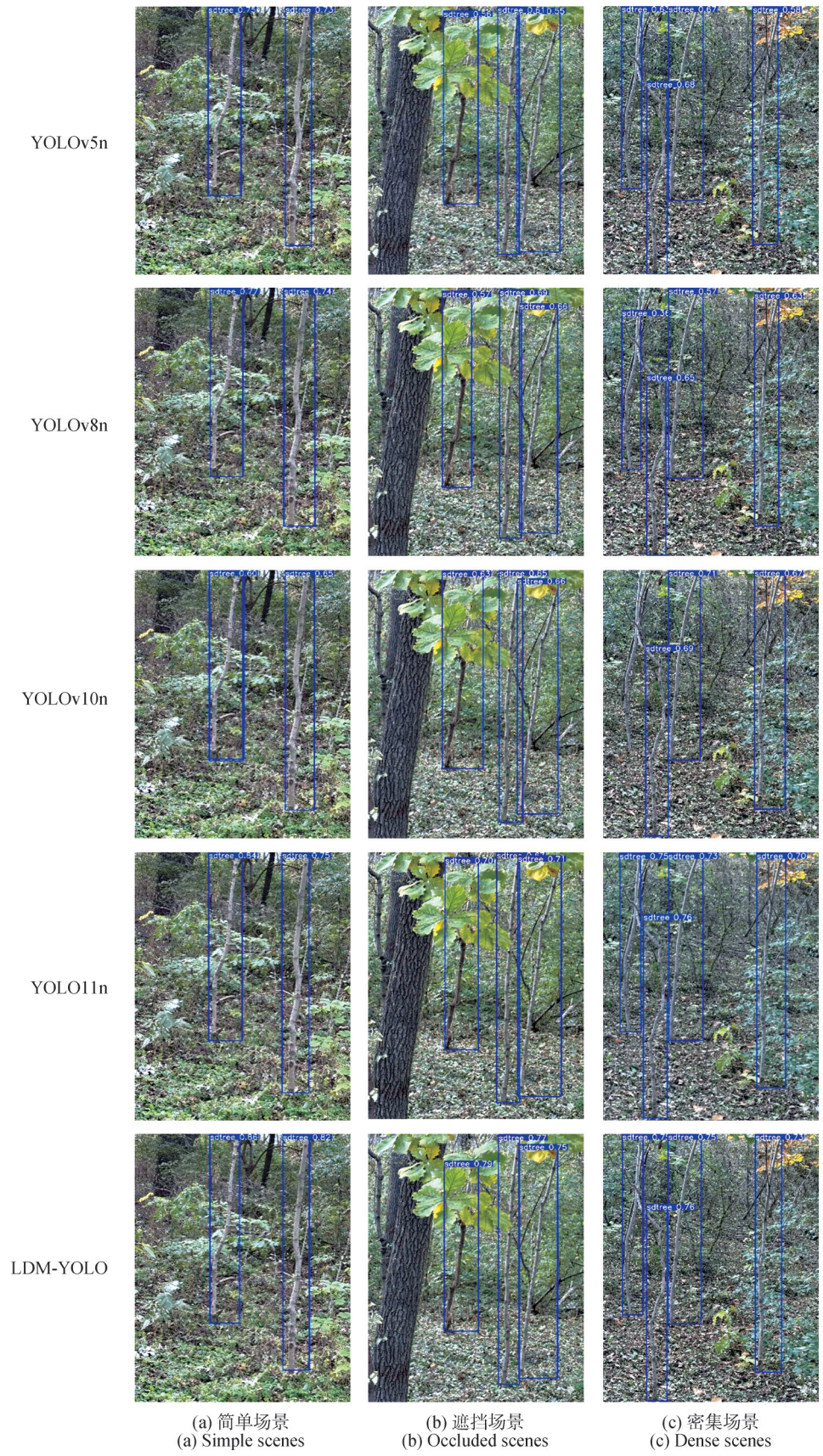


图 10 不同模型检测结果
Fig. 10 Detection results of each model

进策略能够更有效地聚焦细长、被遮挡及密集分布的小径木目标,较好地兼顾了检测精度与模型轻量化需求。

尽管 LDM-YOLO 在本试验中实现了精度与速度的良好平衡,但本研究仍存在一定局限性。自建数据集(3 044 张)规模偏小且场景采集范围相对集中,主要反映了特定林分与季节的特征,限制了模型在异构植被或极端气候下的泛化能力。面向广泛场景部署时,还需采用迁移学习或微调策略以增强模型的环境适应性。热力图分析显示,在极度遮挡或树皮纹理与背景高度相似的极端场景下,仅依赖 RGB 特征仍难以完全避免漏检。这表明可见光感知在复杂林下环境中存在物理瓶颈。

针对上述不足,未来工作可进一步扩展数据集场景与季节范围,引入多类别林木与障碍物检测;同时考虑引入深度信息或多光谱数据进行多模态融合,利用三维空间与光谱差异进一步提升检测稳定性。综上所述,LDM-YOLO 模型具有一定的实用价值,能够为后续在林业移动机器人、便携式巡护设备等资源受限平台上开展小径木智能识别提供可靠的算法基础与参考。

参 考 文 献

- [1] 白宇,杨华,温静,等. 基于邻近木的林分结构多样性研究[J]. 北京林业大学学报,2020,42(6):52-58.
BAI Y, YANG H, WEN J, et al. Study on forest structure diversity based on the neighbourhood trees [J]. Journal of Beijing Forestry University, 2020, 42(6): 52-58.
- [2] ZHAO Z Q, ZHENG P, XU S, et al. Object detection with deep learning: A review [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(11): 3212-3232.
- [3] 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. June 27-30, 2016, Las Vegas, NV, USA, IEEE, 2016: 779-788.
- [4] BADGUJAR C, FLIPPO D, GUNTURU S, et al. Tree trunk detection of eastern red cedar in rangeland environment with deep learning technique [J]. Croatian Journal of Forest Engineering, 2023, 44(2): 357-368.
- [5] DA SILVA D Q, DOS SANTOS F N, FILIPE V, et al. Edge AI-based tree trunk detection for forestry monitoring robots [J]. Robotics, 2022, 11(6): 136.
- [6] LIU Y, WANG H R, LIU Y H, et al. A trunk detection method for *Camellia oleifera* fruit harvesting robot based on improved YOLOv7 [J]. Forests, 2023, 14(7): 1453.
- [7] ZHANG S, SUN Y, QIAO Y, et al. PCC-YOLO: A trunk detection algorithm for orchard robots based on improved YOLOv8 [J]. Agriculture, 2025, 15(16): 1786.
- [8] ZHANG S Q, ZHENG Y B, BI R G, et al. Trunk detection in complex forest environments using a lightweight YOLOv11-TrunkLight algorithm [J]. Sensors, 2025, 25(19): 6170.
- [9] 杨灿,范习健,张九于. SSFYOLO: 一种面向复杂森林场景的树干检测算法[J]. 北京林业大学学报, 2025, 47(2): 132-142.
YANG C, FAN X J, ZHANG J Y. SSFYOLO: A trunk detection algorithm for complex forest scenarios [J]. Journal of Beijing Forestry University, 2025, 47(2): 132-142.
- [10] 胡峻峰,朱昊,黄晓文,等. 基于深度学习的森林移动机器人树干检测[J]. 森林工程, 2024, 40(4): 109-114.
HU J F, ZHU H, HUANG X W, et al. Trunk detection for forest mobile robots based on deep learning [J]. Forest Engineering, 2024, 40(4): 109-114.
- [11] 徐相华,周德靖,余超然,等. 基于DCA-YOLO的受病虫害侵染树木农业无人机低空检测模型[J]. 农业机械学报, 2025, 56(10): 479-491.
XU X H, ZHOU D J, YU C R, et al. Improved model for low altitude detection of trees infected by pests and diseases using agricultural drones based on DCA-YOLO [J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(10): 479-491.
- [12] 王品博,王雨晨,王双永,等. 基于改进YOLOv8的树木年轮实例分割及信息检测方法[J]. 木材科学与技术, 2025, 39(3): 7-18.
WANG P B, WANG Y C, WANG S Y, et al. Study on tree ring instance segmentation and information detection based on improved YOLOv8 [J]. Chinese Journal of Wood Science and Technology, 2025, 39(3): 7-18.
- [13] LAU K W, PO L M, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN [J]. Expert Systems With Applications, 2024, 236: 121352.
- [14] AZAD R, NIGGEMEIER L, HUTTEMANN M, et al. Beyond self-attention: Deformable large kernel attention for medical image segmentation [C]//2024 IEEE/CVF Winter Conference on Applications of Computer Vision. 2024: 1287-1297.
- [15] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample [C]//2023 IEEE/CVF International Conference on Computer Vision. 2023: 6027-6037.
- [16] TAN M X, LE Q V. EfficientNetV2: Smaller models and faster training [C]//2021 International Conference on Machine Learning. 2021: 10096-10106.
- [17] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]//IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE, 2015, 28: 1137-1149.