

基于改进YOLO11的木材端面识别模型设计

张小波, 曾子荣*, 廖彩霞

(江西环境工程职业学院 汽车机电学院, 江西 赣州 341000)

摘要: 天然木材端面存在不规则纹理与缺陷特征, 木材端面识别定位属于一个难点问题。为提高木材端面的检测精度, 同时减少模型参数量、提高模型运行速度、方便移动端部署, 对YOLO11模型进行改进, 构建更加适用于原木检测的端到端深度学习模型。首先, 使用飞浆轻量级CPU卷积神经网络(PaddlePaddle lightweight CPU convolutional neural network, PP-LCNet)替换YOLO11网络结构的骨干网络, 减少模型参数量, 扩大感受野, 提升模型大目标检测精度; 其次, 在颈部网络中加入无参注意力机制简易注意力模块(Simple attention module, SimAM), 自适应增强关键特征, 抑制冗余信息, 增强小目标识别能力; 最后, 引入归一化Wasserstein距离损失函数(normalized Wasserstein distance, NWD), NWD更适合测量极小目标间的相似性, 进一步提高对木材端面识别的准确率和精度。试验结果表明, 同比基准模型, 改进版模型具有更高的端面识别精度, mAP@0.5提升2.65%, mAP@0.95提升5.29%, 浮点计算数下降15.15%, 在原木木材材积检测领域有着较好的应用价值。

关键词: 原木木材; 端面识别; 深度学习; YOLO改进; 目标检测

中图分类号: S776

文献标识码: A

DOI: 10.7525/j.issn.1006-8023.2026.01.007

Wood Log End Recognition Model Design Based on Improved YOLO11

ZHANG Xiaobo, ZENG Zirong*, LIAO Caixia

(School of Automotive Mechanics and Electronics, Jiangxi Environmental Engineering Vocational College, Ganzhou 341000, China)

Abstract: Natural wood end surfaces exhibit irregular textures and defect features, making end surface recognition and localization a challenging problem. To enhance detection accuracy while reducing model parameters and improving computational efficiency for mobile deployment, this study proposes an improved end-to-end deep learning model tailored for log detection by enhancing the YOLO11 architecture. Firstly, the PP-LCNet backbone is adopted to replace the original YOLO11 backbone, effectively reducing the number of parameters, expanding the receptive field, and improving large target detection precision. Secondly, a parameter-free attention mechanism, SimAM, is integrated into the neck network to adaptively emphasize critical features and suppress redundant information, thereby enhancing small target recognition capabilities. Finally, the normalized Wasserstein distance (NWD) loss function is introduced, which is more suitable for measuring similarity between extremely small targets, further improves the accuracy and precision of wood end surface identification. Experimental results demonstrate that the improved model achieves higher end surface recognition accuracy compared to the baseline model, the improved model improves 2.65% and 5.29% on the mAP@0.5 and mAP@0.95 metrics, and FLOPs are decreased by 15.15%. It has good application value in the field of log volume measurement.

Keywords: Log timber; end-surface recognition; deep learning; YOLO enhancement; object detection

0 引言

森林资源采伐与运输环节中,原木楞堆作为临时存储与转运的核心载体,其材积计量长期依赖人工检

尺手段。这种传统方法不仅存在劳动强度大、人力成本高的固有缺陷,更易因操作人员主观判断差异导致测量误差(通常误差率高达8%~12%),且单根原木检测耗时超过2 min,难以满足现代林业规模化运营需

收稿日期:2025-05-19

基金项目:江西省教育厅科学技术研究项目(GJJ2205420);江西省教育厅科学技术研究项目(GJJ2205418)。

第一作者简介:张小波,硕士,副教授。研究方向为机器人视觉。E-mail:109805585@qq.com

*通信作者:曾子荣,硕士,讲师。研究方向为深度学习物体检测。E-mail:376849136@qq.com

引文格式:张小波,曾子荣,廖彩霞.基于改进YOLO11的木材端面识别模型设计[J].森林工程,2026,42(1):65-77.

ZHANG XB,ZENG ZR,LIAO CX. Wood log end recognition model design based on improved YOLO11[J]. Forest Engineering,2026,42(1):65-77.

求。随着机器视觉与深度学习技术的突破性发展,基于多视角图像采集与三维重建的非接触式测量技术为行业革新提供了技术路径。然而,原木断面形态的多样性(如椭圆、裂痕和节疤等异形结构)与作业场景的光照波动、枝叶遮挡等干扰因素,对图像处理算法的鲁棒性提出了严苛要求。

目前,对于原木端面的识别技术主要基于传统图像处理技术和深度学习物体检测技术。在传统图像处理算法方面,赵亚凤等^[1]将遗传算法结合同态滤波用于原木端面图像的分割,先用遗传算法对原木端面图像进行增强,再利用同态滤波进行分割,得到图像分割处理的试验结果。唐浩等^[2]通过色差值聚类将原木图像分割为原木截面、孔隙及背景,去除背景干扰提取原木端面;采用逐级开运算与改进分水岭算法,对端面进行分割计数。陈广华等^[3]提出基于最大熵阈值分割的区域标识算法,通过设定动态阈值,实现对原木端面与背景的精确分割。林耀海等^[4]提出基于圆弧的原木轮廓识别方法,先分析圆弧在数字化图像中的性质,找出圆弧边缘;然后定义验证模型,判断识别结果,这种方法利用部分轮廓边缘识别原木端面。在基于深度学习的算法应用方面,林耀海等^[5]使用YOLOv3-tiny模型对原木端面图像进行目标检测,首先计算得到端面对应的目标区域,然后计算边缘并去除目标区域噪声后,利用Hough变换圆检测算法计算原木端面的准确直径。余平等^[6]针对目标密集且具有大量小目标的问题,提出融合BiFPN和YOLOv5s的密集型原木端面检测方法,添加了一个小目标检测层,融合简化版的BiFPN,并加入跨尺度连接线,以保留更多深层的语义信息,取得了不错的检测效果。胡笑天等^[7]以单次多框检测器(Single shot multiBox detector, SSD)网络为基础网络,改进了特征融合方式,增加了感受野,提高了对小目标的特征提取能力;加入融合多尺度卷积核和空洞卷积的模块,引入能够卷积块注意力模块(convolutional block attention module, CBAM),增强了特征识别能力。李佳雨等^[8]提出了two-stage目标检测算法结合边缘拟合算法的原木检尺径算法,目标检测算法部分采用两阶段的框架,第一阶段扫描图像生成提议框,第二阶段将提议框分类,并生分割掩膜。曾小山等^[9]在YOLOv3主干网络基础上,采用数据增强、特征融合和损失函数等优化手段,对模型进行了改进,提出了端到端深度学习模型YOLO-RW,能够对原木材积图像准确识别和定位。

目前的检测算法还存在一些不足,一是对于遮挡与重叠干扰的识别不足,木材楞堆中端面常因交叉堆放形成遮挡(部分遮挡率>50%),传统霍夫圆检测算法

对此失效,而深度学习模型虽引入CBAM注意力机制强化特征聚焦,但密集遮挡时仍易丢失目标轮廓信息,导致漏检率上升;二是小目标检测困难,混楞堆中小端面目标(直径<10 cm)仅占图像面积0.1%~1%,主流检测器(如YOLOv8)浅层特征提取不足,小目标的检测精度不到85%;三是实时性与精度平衡难于平衡,轻量模型通过深度可分离卷积等技术压缩参数量,但降低了模型对于形变目标的鲁棒性。

新版的端到端模型YOLO11相较于前期模型,在精度和速度方面都有更好的表现。本研究基于YOLO11模型,对其骨干网络(backbone)、颈部(neck)和损失函数部分分别进行改进,进一步提升模型性能,增强其速度和精度。

1 YOLO模型介绍

YOLO(you only look once)是目标检测领域的里程碑算法,其发展历程体现了效率与精度的平衡优化。2016年,YOLOv1首创单阶段检测思想,将检测转化为回归问题^[10],实现45帧/s的实时检测,但定位精度和召回率较低。2017年,YOLOv2引入锚框机制,采用批归一化和高分辨率分类器,在保持速度的同时将mAP提升至76.8%。YOLOv3通过残差网络Darknet-53和多尺度预测,显著提升小物体检测能力,形成3个特征层的金字塔结构^[11]。2020年的YOLOv5重构网络架构,采用Focus切片操作和自适应锚框计算,大幅提升训练效率和部署便利性^[12]。YOLOv8(2023)引入可变形卷积和Transformer混合架构,支持分类分割等多任务,在COCO数据集达到53.1% mAP^[13]。

YOLO11作为YOLO系列新版本,在模型轻量化与效率优化上实现显著突破。其核心特点包括:1)参数精简与计算效率提升。采用C3K2模块的并行卷积设计,结合动态稀疏计算,使YOLO11m模型在COCO数据集上以比YOLOv8m少22%的参数量实现更高mAP;2)架构创新。引入快速空间金字塔池化(spatial pyramid pooling, SPFF)和C2PSA混合注意力机制,通过多尺度特征融合与通道加权策略,增强小目标检测能力。其模型结构如图1所示。YOLO11还完成了多任务集成,通过统一架构支持检测、分割和姿态估计等5类任务。

2 基于YOLO11的改进模型

YOLO11网络结构由输入端Input、骨干网络Backbone、多尺度特征融合网络Neck和检测头Head组成。本研究提出一种基于YOLO11算法改进的检测模型,使用飞浆轻量级CPU卷积神经网络(Paddle paddle

lightweight CPU convolutional neural network, PP-LCNet) 替换 YOLO11 的主干网络,在颈部网络特征输出部

分添加了简易注意力模块(Simple attention module, SimAM),改进模型的网络结构如图 2 所示。

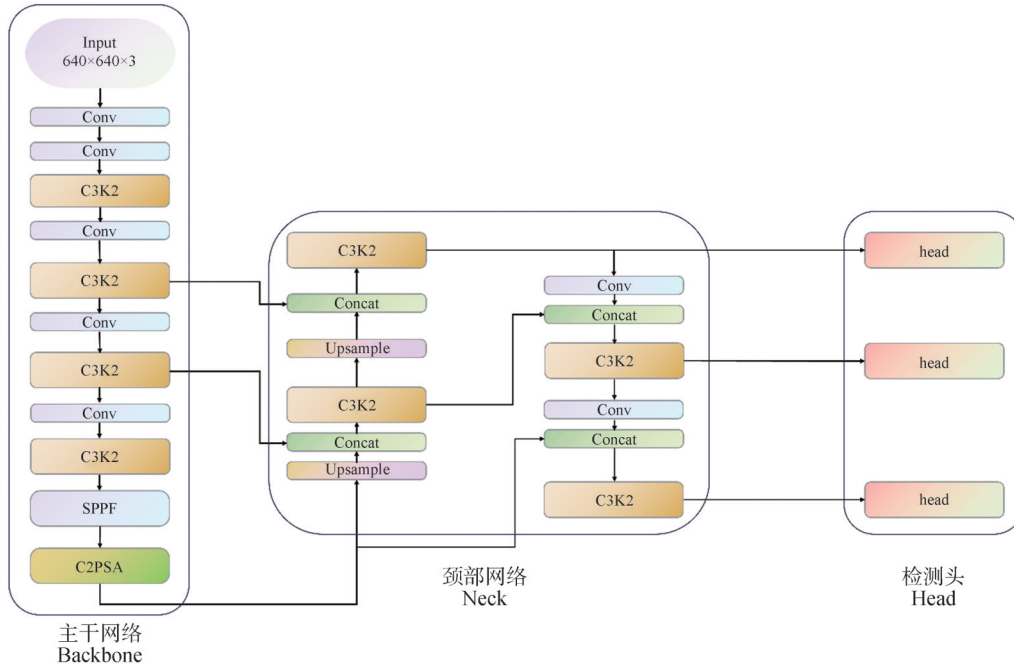


图 1 YOLO11 模型结构
Fig. 1 Overall structure of the YOLO11

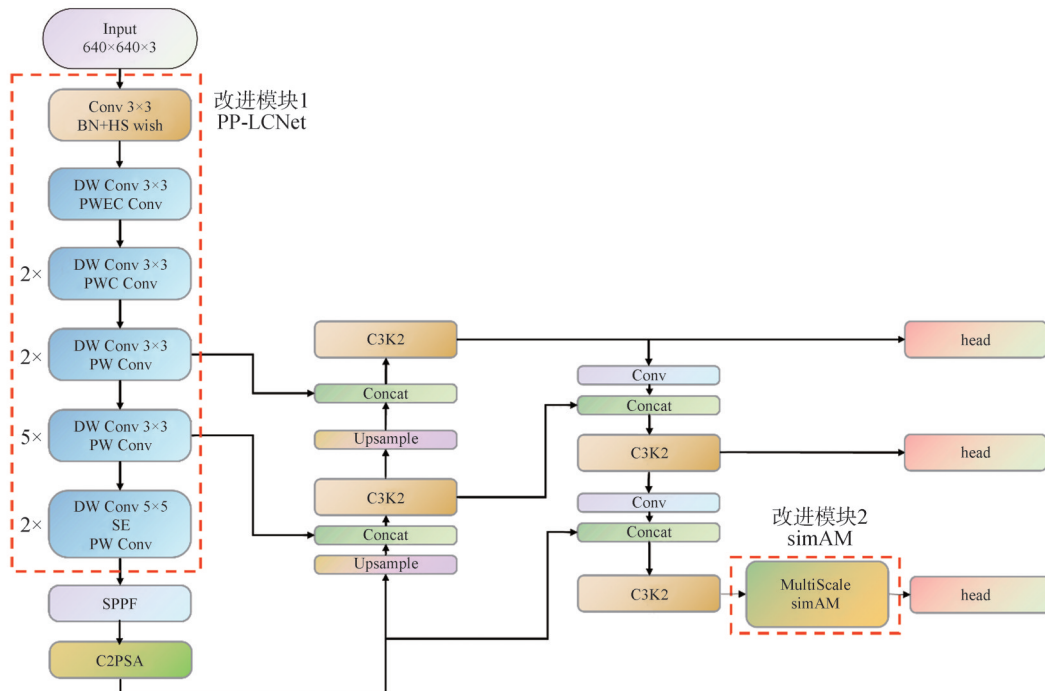


图 2 改进模型整体结构图
Fig. 2 Overall structure of the improved algorithm

2.1 YOLO11主干网络优化

首先对 YOLO11 的主干部分进行改进,引入 PP-LCNet 替换 YOLO11 的主干网络,在实现轻量化的同时提升检查精度。PP-LCNet 具有轻量化的特点^[14],其结

构如图 3 所示。通过深度可分离卷积(depthwise separable convolution, DSC)替代标准卷积,减少参数。通过引入通道注意力机制(squeeze-and-excitation, SE)模块,增强重要特征通道的权重,提升小模型精度,在

保持高精度的同时显著降低算力需求。SE模块由压缩、激励、重标定3个核心步骤组成。首先对输入特征图进行全局平均池化(global average pooling, GAP),将每个通道的二维空间特征压缩为一个实数,目的是将空间维度的信息压缩为通道级的全局描述,为后续权重计算提供全局上下文。激励是权重增强的核心,通过多层感知机(multi-layer perceptron, MLP)学习通道间的依赖关系,生成每个通道的权重。重标定将激励生成的通道权重应用到原始特征图上,实现通道权重的动态调整。

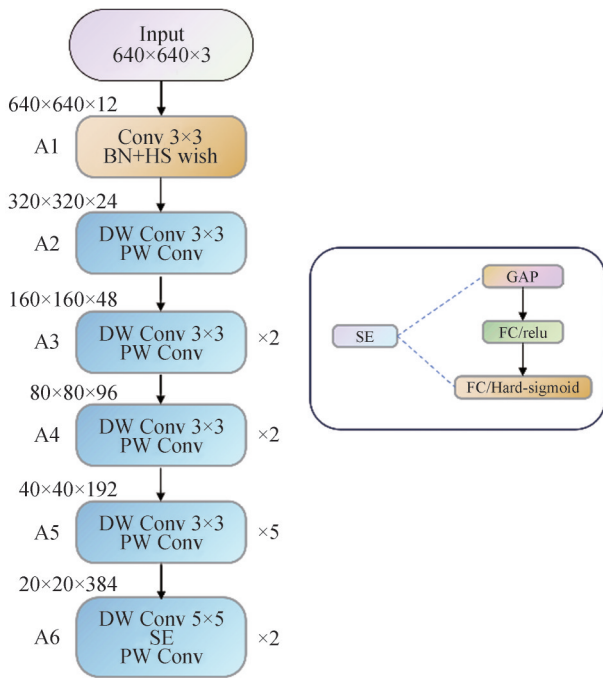


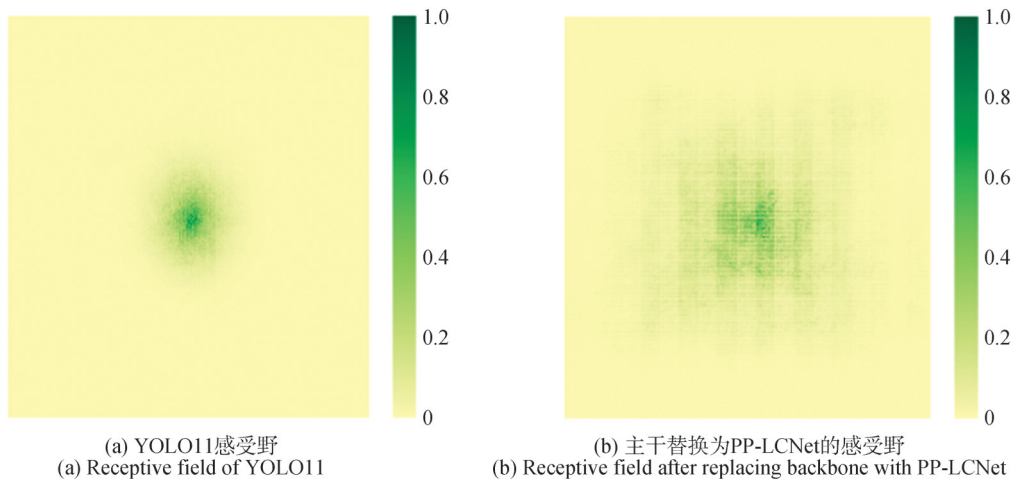
图 3 PP-LCNet网络结构

Fig. 3 PP-LCNet network architecture

PP-LCNet采用深度可分离卷积替代普通卷积操作,构建轻量级网络结构。深度可分离卷积是一种高效的卷积方法,通过将标准卷积分解为深度卷积和点对点卷积,减少了计算量和参数数量^[15]。深度卷积(depthwise convolution)使用 $K \times K$ 卷积核对每个输入通道单独进行空间卷积,输出通道数与输入通道数相同。逐点卷积(pointwise convolution)使用 1×1 卷积核组合深度卷积后的通道,调整通道数至目标维度 M 。普通卷积的参数量为 $K \times K \times C \times M$ (C 为输入通道数; K 为卷积核大小; M 为输出通道数),深度可分离卷积参数量为 $K \times K \times C + C \times M$,模型的参数量因此大幅减小。PP-LCNet使用H-Swish激活函数提升性能,同时几乎不增加推理时间^[16]。PP-LCNet在网络尾部添加SE模块和大卷积核,平衡准确度和速度。SE注意力模块通过轻量级的全局池化和全连接操作,实现了高效的通道注意力建模。

改进后的骨干结构可分为6个部分(A1—A6)。A1是普通3x3卷积进行特征提取,并使用了H-Swish激活函数,主要是从输入图像中提取低层次特征。A2—A5由深度可分离卷积DepthSepConv组成,目的在于减少模型参数和网络计算量。A2—A5的卷积核大小为3x3,A6模块采用5x5的卷积核。

对主干替换后的模型进行感受野可视化,由图4可知,通过引入5x5大核卷积,可显著扩大感受野,帮助模型感知更广的上下文信息,提升对全局特征的捕捉能力。大核卷积能有效整合边缘、纹理等跨区域特征,通过连接不同层级的特征图,促进低层细节与高层语义的融合增强模型对微小变化的鲁棒性。



(a) YOLO11感受野
(a) Receptive field of YOLO11

(b) 主干替换为PP-LCNet的感受野
(b) Receptive field after replacing backbone with PP-LCNet

图 4 模型感受野对比

Fig. 4 Receptive field comparison

2.2 SimAM 注意力机制

通道注意力为 1D 注意力,对不同通道区别对待,对所有位置同等对待。而空间注意力是 2D 注意力,对不同位置区别对待,对所有通道同等对待,这限制了学习更多辨别线索的能力。已有的注意力机制如瓶颈注意力模块(bottleneck attention module, BAM)、CBAM^[17]等分别将空间注意力与通道注意力进行并行或串行组合;而人脑的 2 种注意力往往是协同工作,所以需要一种统一权值的注意力模块。为更好地实现注意力,需要评估每个神经元的重要性。

基于成熟的神经科学理论, Yang 等^[18]提出 SimAM 注意力机制,这是一种全 3D、加权且无参数的注意力机制,其原理如图 5 所示。 H 和 W 为特征图的高和宽, X 为输入的特征图。与其他现有的注意力机制相比, SimAM 考虑了空间和通道因素之间的相关性,并且可以高效地为特征的映射生成现实的 3D 权重,无需额外的参数。具体计算过程为

$$e_i^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (1)$$

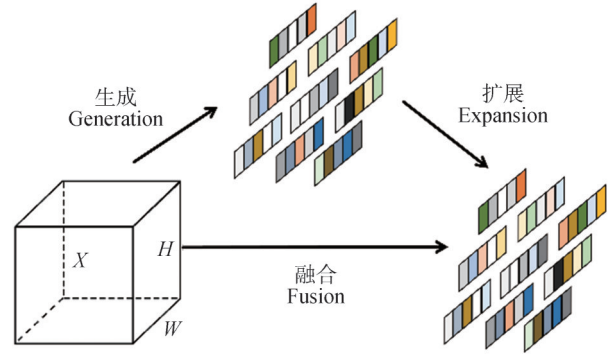
$$\hat{\mu} = \frac{1}{M} \sum_{i=1}^M X_i \quad (2)$$

$$\hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (X_i - \hat{\mu})^2 \quad (3)$$

式中: e_i^* 为每个通道上的能量函数; M 为各通道神经元个数; $\hat{\mu}$ 为 X 中每个通道上的均值; $\hat{\sigma}^2$ 为 X 中每个通道上的方差; t 为输入特征的目标神经元; λ 为超参数。神经元的能量 $1/e_i^*$ 越低,其与周围的神元元的差值越大,即其重要性越大。添加注意力机制的输出特征公式为

$$Y = \text{Sigmoid}\left(\frac{1}{e_i^*}\right) \odot X \quad (4)$$

SimAM 注意力机制通过统计量计算权重,无需可学习的参数,每个神经元的权重由全局特征分布决定,因此全局感知能力强。又因为采用闭式解,从而有效解决过度优化迭代问题,大大降低计算的复杂度。



H 和 W 为特征图的高和宽; X 为输入的特征图。

H and W are the width and height of the feature map; X is the input feature map.

图 5 SimAM 注意力机制原理图

Fig. 5 SimAM attention architecture

本研究采用多尺度 SimAM 以加强特征提取,首先对输入特征图进行多尺度下采样,尺度分别为 2 和 4,接着对每个尺度计算 SimAM 注意力图,然后将注意力图上采样回原始尺寸,对各尺度的注意力图计算平均值,再使用 Sigmoid 激活函数得到最终注意力权重,再将注意力权重与原始特征图相乘,从而得到最终的输出结果,具体过程如图 6 所示。通过以上过程,进一步对各尺度进行了特征融合,增强网络的特征提取能力。

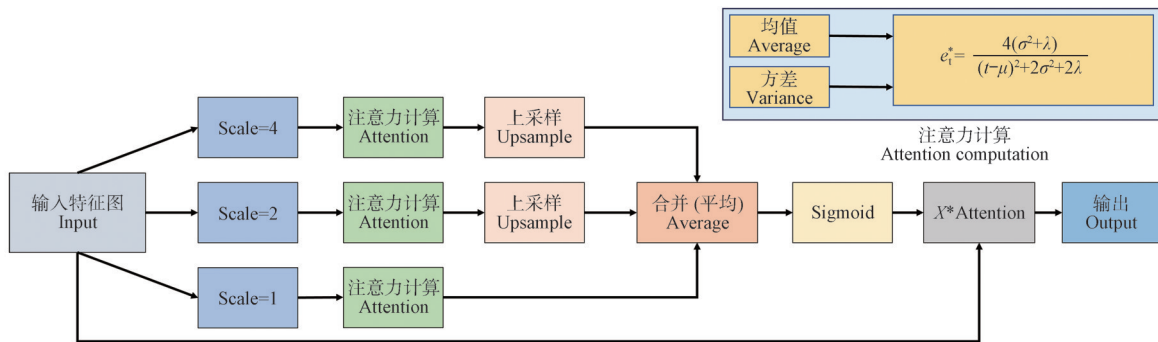


图 6 多尺度 simAM 注意力计算过程

Fig. 6 Multi-scale simAM attention computation process

在本研究中,木材截面的边缘与周围环境具有明显的差异性,在 YOLO11 的特征输出层中引入 SimAM,可以进一步增强差异性,增强网络提取边缘特征的能力,并去除背景冗余信息,降低背景复杂信息对于对木材截面检测的干扰。

2.3 归一化 Wasserstein 距离损失函数 (normalized Wasserstein distance, NWD)

传统目标检测常用交并比 (IoU, 式中记为 I_{o_u}) 衡量预测框与真实框的重叠程度,但在小目标 (绝对像素面积小于 16×16 像素) 场景存在明显短板,如图 7 所

示。检测一个微小的木材截面(约5×5像素),初始状态下 $A \cap B=16, A \cup B=24, I_{ou}=0.667$ 。检测框的中心点偏差2个像素后, $A \cap B=4, A \cup B=46, I_{ou}=0.087, I_{ou}$ 骤降,导致训练信号剧烈波动。而对于大的目标端面,初

始状态下 $A \cap B=441, A \cup B=527, I_{ou}=0.837$,检测框的中心点偏差2个像素后 $A \cap B=361, A \cup B=607, I_{ou}=0.594$,随着锚框的移动, I_{ou} 的变化平缓,不会导致训练结果的剧烈波动。

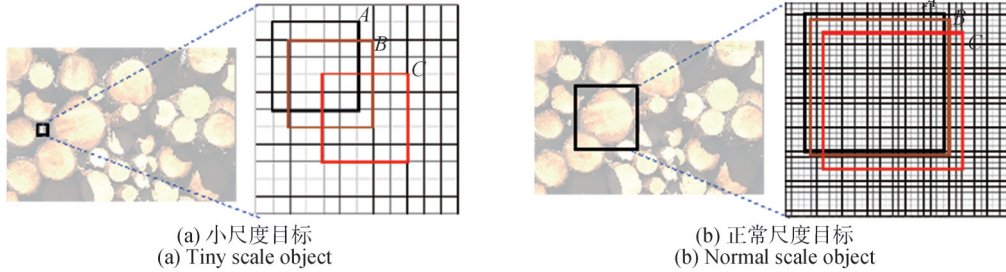


图7 小目标与大目标IoU变化对比

Fig. 7 Comparison of IoU variation between small and large targets

为解决小目标检测中IOU剧烈波动的问题,Zhou等^[19]提出了NWD损失函数。NWD损失函数是一种针对小目标检测优化的损失函数,通过建模边界框为高斯分布并利用Wasserstein距离度量分布相似性,显著提升了微小物体的检测精度^[20]。对于极小目标,边界框中通常包含一些背景像素,因为大部分的真实目标不是严格的矩形。在这些边界框中,前景像素和背景像素分别集中在边界框的中心和边界。为更好地描述边界框中不同像素的权重,边界框可以被建模为2个二维的高斯分布,其中中心像素有最高的权重,然后从中心到边界像素的重要性递减。

NWD损失函数的关键优势在于其尺度不变性,相较于IoU对尺寸敏感的特性,NWD对不同尺度的边界框相似性度量更稳定。NWD通过分布相似性评估边界框关系,能有效解决小目标检测中因遮挡或稀疏分布导致的标签分配问题^[21]。同时,在计算复杂度方面,因为损失计算仅涉及均值、平方差和指数运算,计算的复杂度较低。

NWD损失的具体计算过程如下。水平边界框为

$$R = (C_x, C_y, \omega, h) \quad (5)$$

式中, (C_x, C_y) 、 ω 、 h 分别为中心坐标、宽度和高度,边界框的中心像素权重最高,像素的重要性从中心向边界递减。将边界框建模为二维高斯分布 $N(\mu, \Sigma)$, μ 和 Σ 分别为高斯分布的均值向量与协方差矩阵,定义为

$$\mu = \begin{bmatrix} C_x \\ C_y \end{bmatrix}, \Sigma = \begin{bmatrix} \frac{\omega^2}{4} & 0 \\ 0 & \frac{h^2}{4} \end{bmatrix} \quad (6)$$

边界框A和B之间的相似度可以转换为2个高斯分布之间的距离,二阶高斯分布 $\mu_1 = N(m_1, \Sigma_1), \mu_2 =$

$N(m_2, \Sigma_2)$,其之间的Wasserstein距离 W 定义为

$$W_2^2(\mu_1, \mu_2) = \|m_1 - m_2\|_2^2 + T_r \left(\sum_1 + \sum_2 - 2 \left(\sum_1^{1/2} \sum_2^{1/2} \right)^{1/2} \right) \quad (7)$$

归一化为

$$N_{WD}(\mu_1, \mu_2) = \exp \left(-\frac{\sqrt{W_2^2(\mu_1, \mu_2)}}{T} \right) \quad (8)$$

式中: T 为与数据集相关的常数; m_1 为第1个边界框的均值向量; m_2 为第2个边界框的均值向量; T_r 为矩阵对角线元素的和; N_{WD} 为归一化Wasserstein距离的相似度量。

将NWD用作损失函数 L_{NWD} 为

$$L_{NWD} = 1 - N_{WD}(\mu_1, \mu_2) \quad (9)$$

NWD对于极小目标检测有以下优势:1)尺度不变性;2)对位置偏移的平滑性;3)测量2个互不相交或相互包含边界框的相似度。

3 算法试验与分析

3.1 数据集与预处理

本研究木材堆放图像的采集地点为江西省赣州市某木材检验所,初始采集图片600张。图片包括各种木材类型、形状、大小和颜色,涵盖不同的光照条件,还有部分木材因为保管时间过长,存在腐蚀、裂纹等缺陷,形状特征不规则且不完整,各类原木的占比见表1。使用labellmg工具对木材端面人工标记,如图8所示。

表1 各类原木图像占比

Tab. 1 The proportion of images of various types of logs

				%
端面完整 End face integrity	形状不规则 Irregular shape	腐蚀 Corrosion	裂纹 Crack	
62.3	21.6	6.7	9.4	

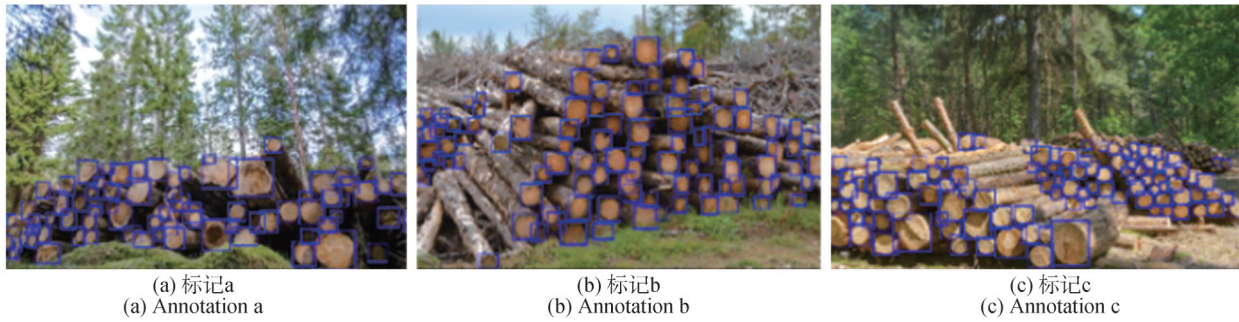


图 8 样本数据集标记
Fig. 8 Annotation of datasets

为降低过拟合风险,并提升模型适应性、强化模型的泛化能力、提升数据分布的多样性,对数据集首先进行增强处理。第一是基本的几何变换,通过随机水平/垂直翻转(概率 0.5)、旋转($\pm 30^\circ$ 内随机旋转)、缩放(比例 0.5~1.5)、裁剪(保留 $\geq 70\%$ 目标)和透视变换($X、Y$ 方向剪切分量 < 0.3),模拟目标在不同视角、位置和尺度的变化,增强模型对目标形变和遮挡的鲁棒性。第二是颜色空间调整,包括对比度增强(明度调整幅度 $\pm 40\%$)、色彩抖动(亮度抖动、对比度抖动 $\pm 20\%$)等,使模型适应不同光照条件和环境干扰。

此外,还采用了高级增强策略,第一是 Mosaic,将 4 张图像拼接训练,增加目标密度和背景多样性;第二

是 MixUp 策略,通过线性融合两图及其标注,提升小目标检测能力;第三是随机擦除,用于模拟遮挡场景,防止模型过度依赖局部特征。以上策略对小目标具有良好的优化,因为平衡了数据集中不同尺寸目标的分布,缓解小目标检测中的特征丢失问题。通过数据增强,得到图片数据集共 2 400 张,如图 9 所示。目标尺寸和形状分布,如图 10 所示。由图 10 可以看出,目标大小在 0.03~0.5,分布均匀,0.03~0.2 的小目标占比大于 70%,保证了训练模型的鲁棒性和有效性。图片像素统一为 640 \times 640,然后按 7:2:1 的比例划分为训练集、验证集和测试集。



图 9 原木数据增强

Fig. 9 Log image data augmentation

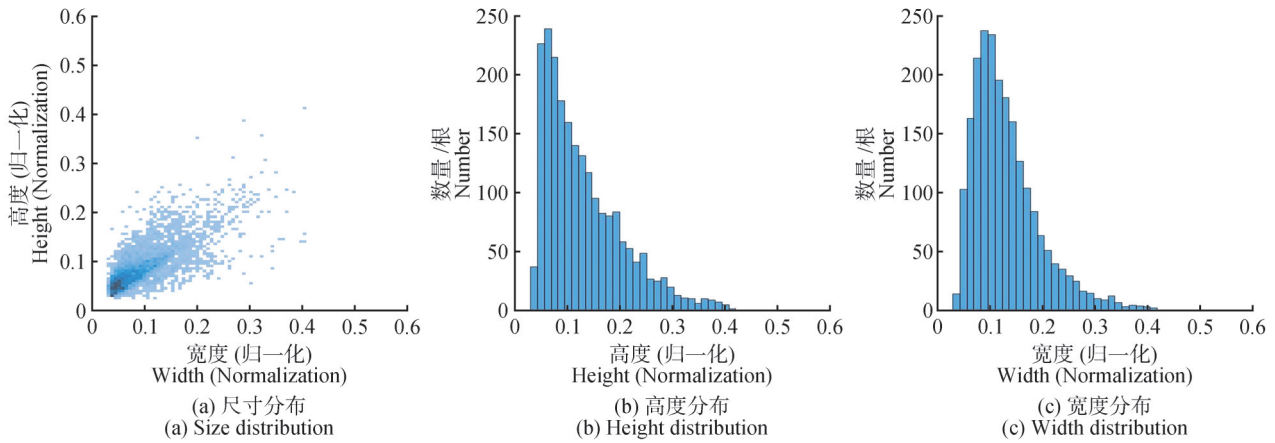


图10 木材端面尺寸分布

Fig. 10 Size distribution diagram of wood end faces

试验软硬件配置如下:操作系统为64位 windows11, GPU: NVIDIA GeForce RTX2080Ti 11 G 显存, CPU: Intel(R)Core(TM)i7-8700K, 深度学习框架为 Py-Torch2.0, 编程语言为 python3.9。试验 epochs 为 200, batchsize 为 8, 采用小批量随机梯度下降。imagesize 为 640×640, 初始学习率为 0.01, 权重衰减系数为 0.0005。

3.2 评价指标

采用3个指标来精确评价模型的性能,即平均精度均值(mean average precision, mAP)、召回率(Recall 式中记为 R_{recall})和精确率(Precision, 式中记为 P_{recision})来衡量模型在不同检测评价函数阈值下的平均精度。 M_{AP} 是不同类别精度的平均值,定义为不同召回率下的平均精度。其计算方式为

$$M_{\text{AP}} = \frac{1}{n} \sum_{i=1}^n A_p(i) \quad (10)$$

式中: n 为类别数; $A_p(i)$ 为第*i*类的平均精度。该指标综合评估模型对所有类别的检测能力,值越高越好。

Precision 为正确检测的正样本数占所有真实正样本数的比例,主要用于评估模型对漏检(false negative)

的敏感度,尤其在密集目标场景中至关重要。 T_p 为将正样本正确预测为正样本的数量, F_p 为将负样本错误预测为正样本的数量。精确率公式为

$$P_{\text{recision}} = \frac{T_p}{T_p + F_p} \quad (11)$$

Recall 为正确检测的正样本数占所有预测为正样本数的比例,用于评估模型对误检(false positive)的控制能力。 F_N 为将正样本错误预测为负样本的数量。召回率公式为

$$R_{\text{ecall}} = \frac{T_p}{T_p + F_N} \quad (12)$$

3.3 消融试验

为验证本研究改进模块对于模型检测能力提升效果,在自制的数据集上将不同模块对于模型检测性能的影响进行消融试验。以 YOLO11n 为基线模型,针对该模型逐个添加改进模块,首先应用更改 Backbone 为 PP-LCNet(模块1),然后在大目标检测层添加 SimAM 注意力机制(模块2),最后引入 NWD 损失(模块3)。通过相同的试验条件评估不同模块对检测效果的影响,相关消融试验结果见表2。

表2 消融试验数据

Tab. 2 Dissolution experiment results

模型 Model	平均精度均值@0.5/ mAP@0.5	平均精度均值@0.95/ mAP@0.95	精确率/% Precision	召回率/% Recall	权重文件大小 Wsz	浮点计算数/G GFlops
baseline model(YOLO11n)	90.56	70.35	90.11	82.74	5.23	6.6
PP-LCNet	91.14	72.16	91.23	83.15	4.76	5.6
simAM	92.69	74.00	92.35	85.70	5.23	6.6
NWD	92.55	73.90	91.33	85.29	5.23	6.6
PP-LCNet+simAM	92.94	75.47	92.79	86.21	4.76	5.6
PP-LCNet+simAM+NWD	93.21	75.64	92.59	86.98	4.76	5.6

由表 2 可知,单独使用 PP-LCNet 替换主干网络后, $mAP@0.5$ 提升 0.58%, $mAP@0.95$ 提升 1.81%, 精确度提升 1.12%, 召回率提升 0.41%, 而模型参数量下降 8.99%, 浮点计算数下降 15.15%。主干网络替换的主要目的是降低模型复杂度, 提升模型在计算资源受限情况下的适用性, 同时模型性能也取得不错的提升。单独加入 SimAM 模块, $mAP@0.5$ 提升 2.13%, $mAP@0.95$ 提升 3.65%, 精确度提升 2.24%, 召回率提升 2.96%, 因为 SimAM 为无参注意力机制, 所以参数量没有变化。该模块对于模型性能的提升显著, 因其增强模型重要神经元的作用及凸显木材截面的边界。单独使用 NWD 替代 CIoU 损失, $mAP@0.5$ 提升 1.99%, $mAP@0.95$ 提升 3.55%, 精确度提升 1.22%, 召回率提升 2.55%, 该模块主要针对小目标检测能力较弱的问题, 对模型性能的提升也较为显著。

在多个模块联合改进中, 同时使用 PP-LCNet+

SimAM 模块后, $mAP@0.5$ 提升 2.38%, $mAP@0.95$ 提升 5.12%, 精确度提升 2.68%, 召回率提升 3.47%, 表明这 2 个模型可以协同工作, 通过共同作用, 对模型性能完成了综合提升。同时应用 PP-LCNet+SimAM+NWD 损失模块后, $mAP@0.5$ 提升 2.65%, $mAP@0.95$ 提升 5.29%, 精确度提升 2.48%, 召回率提升 4.24%, 相比不使用 NWD 损失, 精确度方面略有下降, $mAP@0.5$ 、 $mAP@0.95$ 和召回率都有明显提升, 证明改进的有效性。综上试验结果, 同时应用 3 种改进模块能够改进检测性能, 符合轻量级网络设计, 并对木材小目标检测任务是有效的。

由图 11 可知, 改进的模型相比基准模型 YOLO11n 具有更快的收敛速度, 且在各个训练轮次中, 精度始终高于基准模型, 没有出现大幅波动的情况, 说明改进后模型的整体性能明显优于原模型。

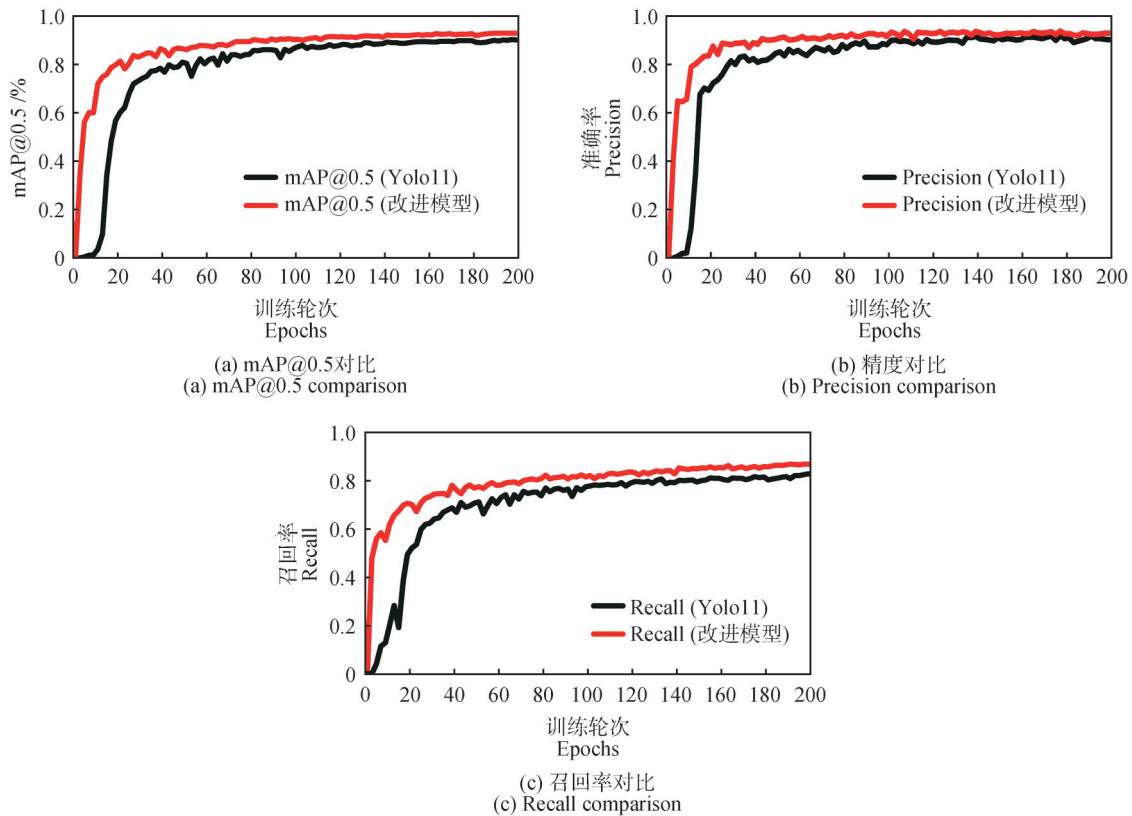


图 11 模型训练曲线对比

Fig. 11 Model training curves comparison

3.4 注意力对比试验

为验证无参注意力模块 SimAM 在本模型中的作用, 在 YOLO11n 网络的相同位置添加卷积块注意力模块 CBAM、高效多尺度注意力模块 (Efficient multi-scale attention, EMA)、可变形大核注意力模块 (Deformable large kernel attention module, DLKA) 及混合局部通

道注意力模块 (Mixed local channel attention module, MLCA), 将这 4 种常见注意力机制与 SimAM 进行对比试验。试验结果见表 3。

由表 3 可知, CBAM 通过串联通道注意力和空间注意力, 对 mAP 提升较大; EMA 采用多尺度并行分支 (1×1 和 3×3 卷积) 和跨空间交互^[22], 小幅增大参数量,

表3 不同注意力机制对比试验
Tab. 3 Comparison experiment of different attention mechanisms

注意力机制 Attention mechanisms	平均精度均值@0.5/ mAP@0.5	召回率/ Recall	浮点计算数/G GFlops
CBAM	92.13	85.5	6.6
EMA	91.87	84.8	6.9
DLKA	91.96	85.1	13.3
MLCA	92.02	85.1	6.6
SimAM	92.69	85.7	6.6

mAP提升有限;DLKA使用大核卷积(捕获全局上下文)和可变形采样(动态适应形变)^[23]具有不错的mAP提升,但是计算量大幅增加;MLCA通过局部池化(提取细节)和全局池化(建模长距依赖)^[23],计算量增加很小,mAP提升较为明显。相较而言,SimAM注意力通过能量函数得出三维注意力权重,在不增加参数的基础上,mAP和召回率提升幅度最大,表明该机制在本研究中的有效性。

表4 主流算法对比数据

Tab. 4 Comparative experiments of different algorithms

模型 Model	平均精度均值@0.5/ mAP@0.5	平均精度均值@0.95/ mAP@0.95	精确率/ Precision	召回率/ Recall	浮点计算数 GFlops	权重文件大小 Wsz
Faster R-CNN	89.20	73.40	87.90	78.70	189.6	235.70
YOLOv3-tiny	85.30	64.90	86.90	79.40	14.3	19.20
YOLOv5n	88.60	68.50	84.80	81.20	5.8	4.70
YOLOv5s	90.40	72.00	86.50	82.40	16.5	16.00
YOLOv8n	86.70	67.50	85.30	78.10	8.7	5.60
YOLOv8s	89.80	73.10	84.60	83.70	28.6	20.00
YOLO11n	90.56	70.35	90.11	82.74	6.6	5.23
本研究模型 Improved model	93.21	75.64	92.59	86.98	5.6	4.76

综上所述,相比于主流算法模型,本模型参数量仅略高于YOLOv5n,而精度取得2.5%以上的提升,在性能和计算量方面取得良好的平衡,适合在木材端面检测设备上部署。

3.6 可视化分析

使用梯度加权类激活映射(Gradient-weighted class activation mapping, Grad-CAM)生成模型的热力图,Grad-CAM通过梯度加权类激活映射,突出显示模

3.5 主流算法对比试验

为探究本研究算法的检测性能,选取一些当前主流的目标检测算法Faster R-CNN、YOLOv3-tiny、YOLOv5n、YOLOv5s、YOLOv8n、YOLOv8s等进行对比试验,试验结果见表4。相比于其他主流目标检测方法,本模型具备更高的检测性能,mAP@0.5、Precision和Recall分别达到93.21%、92.59%和86.98%。在mAP@0.5上,本研究模型相比于YOLOv3-tiny提升7.91%、YOLOv5n提升4.61%、YOLOv5s提升2.81%、YOLOv8n提升6.51%、YOLOv8s提升3.41%。在准确率上,相比其他模型至少提升5.69%。在召回率上,本研究模型相比其他模型,至少提升3.28%。在权重文件大小上,本研究算法相比于YOLOv3-tiny减小75.21%、YOLOv5s减少70.25%、YOLOv8n减少15%、YOLOv8s减少76.2%,仅比YOLOv5n增加1.28%。在浮点计算数上,本研究模型相比于YOLOv3-tiny减小60.8%,相比YOLOv5s减少66.1%,相比YOLOv8n减少35.6%,相比YOLOv8s减少80.4%,相比YOLOv5n减小3.4%。

型做预测时关注的图像区域,其原理是利用目标类别的梯度信息反向传播至特定卷积层,加权融合特征图,突出模型决策依据的区域。试验结果如图12所示。由图12可知,YOLO11n对小目标的关注度较低,且对背景的识别度不足。对比而言,本研究模型关注点集中于木材截面,对背景的抑制较好,且对小目标的关注度更高,而且模型注意力集中于木材截面的中心,这可以使预测边界框更为准确,增强模型的检测性能。

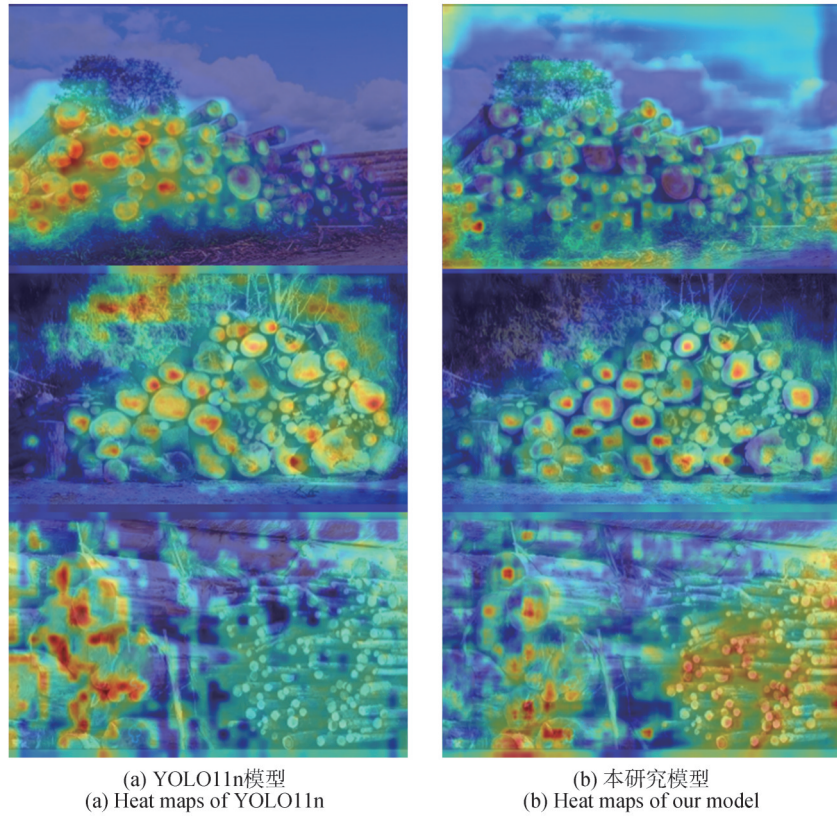


图 12 模型热力图可视化对比

Fig. 12 Visual comparison of heat maps

YOLO11n 模型和改进后的模型对木材端面的检测效果对比如图 13 所示。由图 13 可知,改进后的模型目标检测的置信度提升明显,对微小目标的检测更为灵敏;在密集目标、低分辨率和复杂背景等场景下,改

进的模型都有更好的表现,YOLO11n 存在多处漏检的情况,而改进的模型成功检测出目标。综上所述,改进后的模型对木材端面具有良好的识别效果。

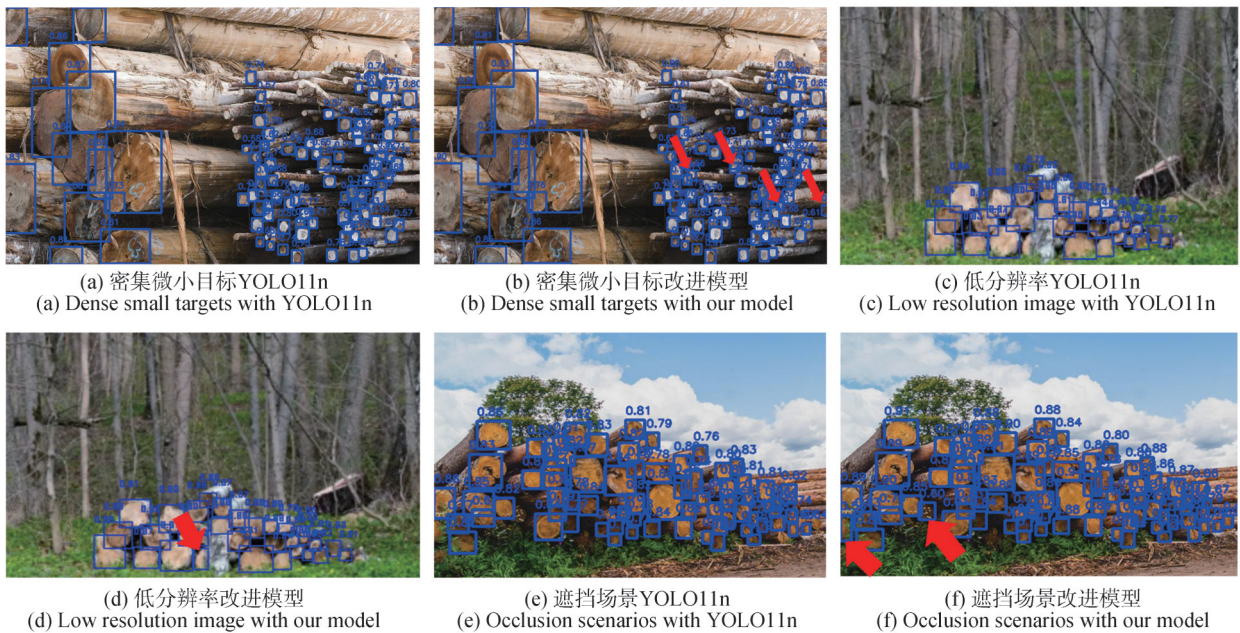


图 13 不同场景检测效果结果对比

Fig. 13 Comparison of detection results in different scenarios

4 结 论

针对木材端面识别背景复杂、小目标分布密集、目标形状尺度多样、光线和拍摄角度差异大等问题,本研究提出基于YOLO11的改进模型,通过多维度技术创新实现检测精度提升。模型改进方案包含以下核心优化策略。1)轻量化主干网络重构,采用轻量级PP-LCNet替换YOLO11主干网络,其基于深度可分离卷积策略,降低模型复杂度,适合移动端部署,通过大核卷积与SE模块扩大感受野,提升大目标检测精度;2)添加自适应特征增强机制,在颈部网络嵌入无参注意力SimAM模块,此模块基于能量函数构建三维注意力权重,能够抑制背景干扰,进一步提升模型性能;3)引入NWD损失函数,将边界框建模为高斯分布,解决IoU对小目标位置敏感问题,大幅提升小目标检测精度。

对比基准模型,本模型具有更高的端面识别精度和更低的计算量,mAP@0.5提升2.65%,mAP@0.95提升5.29%,浮点计算数下降15.15%。研究表明,本研究模型在原木木材材积检测领域有着较好的应用价值,并为移动端检测装备的部署和应用提供参考。未来的工作是将本改进算法集成到资源有限嵌入式设备中,实现木材端面检测设备的软硬件一体设计,并关注其实际使用情况。

参 考 文 献

- [1] 赵亚凤,任洪娥. 遗传算法和同态滤波在原木端面图像处理中的应用[J]. 东北林业大学学报, 2014, 42(2): 129-132.
ZHAO Y F, REN H E. Genetic algorithm and homomorphic filter in image processing of log surface[J]. Journal of Northeast Forestry University, 2014, 42(2): 129-132.
- [2] 唐浩,王克俭,李晓焯,等. 基于色差聚类的原木图像端面检测与统计[J]. 计量学报, 2020, 41(6): 682-688.
TANG H, WANG K J, LI X Y, et al. Logs end detection and statistics by color difference clustering[J]. Acta Metrologica Sinica, 2020, 41(6): 682-688.
- [3] 陈广华,张强,陈梅倩,等. 双目视觉的原木径级快速检测算法[J]. 北京交通大学学报, 2018, 42(2): 22-30.
CHEN G H, ZHANG Q, CHEN M Q, et al. Rapid detection algorithms for log diameter classes based on binocular vision[J]. Journal of Beijing Jiaotong University, 2018, 42(2): 22-30.
- [4] 林耀海,景林,王长纓,等. 基于圆弧的原木轮廓的识别与验证[J]. 福建农林大学学报(自然科学版), 2016, 45(6): 649-654.
LIN Y H, JING L, WANG C Y, et al. Outline identification and verification of cross section of log based on arc edges[J]. Journal of Fujian Agriculture and Forestry University (Natural Science Edition), 2016, 45(6): 649-654.
- [5] 林耀海,杨泽灿,张泽均. 结合图像和图形特征的原木轮廓识别[J]. 福建农林大学学报(自然科学版), 2020, 49(3): 412-417.
LIN Y H, YANG Z C, ZHANG Z J. Outline extraction of logs cross section base-upon both image and graphics features[J]. Journal of Fujian Agriculture and Forestry University (Natural Science Edition), 2020, 49(3): 412-417.
- [6] 余平平,林耀海,赖云锋,等. 融合BiFPN和YOLOv5s的密集型原木端面检测方法[J]. 林业工程学报, 2023, 8(1): 126-134.
YU P P, LIN Y H, LAI Y F, et al. Dense log end face detection method using the hybrid of BiFPN and YOLOv5s[J]. Journal of Forestry Engineering, 2023, 8(1): 126-134.
- [7] 胡笑天,王克俭,王超,等. 一种基于改进SSD的原木端面识别方法[J]. 林业工程学报, 2023, 8(1): 141-149.
HU X T, WANG K J, WANG C, et al. Development of log end face recognition method based on improved SSD[J]. Journal of Forestry Engineering, 2023, 8(1): 141-149.
- [8] 李佳雨,刘晋浩. 基于实例分割模型的原木检尺径方法[J]. 北京林业大学学报, 2023, 45(3): 153-159.
LI J Y, LIU J H. A method of log diameter measurement based on instance segmentation model[J]. Journal of Beijing Forestry University, 2023, 45(3): 153-159.
- [9] 曾小山,张小波. 基于YOLO-RW模型的机器视觉原木端面识别定位[J]. 森林工程, 2023, 39(5): 144-153.
ZENG X S, ZHANG X B. A log end face recognition and positioning model based on YOLO-RW[J]. Forest Engineering, 2023, 39(5): 144-153.
- [10] 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). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 779-788.
- [11] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv: 1804. 02767, 2018.
- [12] JOSEPH N, JACOB S. YOLOv5 is here: State-of-the-art object detection at 140 FPS[EB/OL]. [2020-06-10].
- [13] TERVEN J, CORDOVA-ESPARZA D. A comprehensive 17 review of Yolo: From Yolov1 to Yolov8 and beyond[J]. arXiv preprint arXiv: 2304. 00501, 2023.
- [14] CUI C, GAO T, WEI S, et al. PP-LCNet: A lightweight CPU convolutional neural network [J]. arXiv: 2109. 15099, 2021.
- [15] CHOLLET F. Xception: Deep learning with depthwise separable convolutions [C]//2017 IEEE Conference on

- Computer Vision and Pattern Recognition (CVPR). July 21–26, 2017, Honolulu, HI, USA. IEEE, 2017: 1800–1807.
- [16] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27–November 2, 2019, Seoul, Korea. IEEE, 2019: 1643–1655.
- [17] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [M]//Computer Vision – ECCV 2018. Cham: Springer International Publishing, 2018: 3–19.
- [18] YANG L, ZHANG R Y, LI L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks [C]//International Conference on Machine Learning. PMLR, 2021: 11863–11874.
- [19] ZHOU X, JIANG L, GUAN X J, et al. Infrared small target detection algorithm with complex background based on YOLO-NWD [C]//2022 4th International Conference on Image Processing and Machine Vision. March 25–27, 2022. Hong Kong, China. ACM, 2022: 6–12.
- [20] WANG J W, XU C, YANG W, et al. A normalized Gaussian Wasserstein distance for tiny object detection [J]. arXiv: 2110. 13389, 2021.
- [21] 董刚, 谢维成, 黄小龙, 等. 深度学习小目标检测算法综述 [J]. 计算机工程与应用, 2023, 59(11): 16–27.
- DONG G, XIE W C, HUANG X L, et al. Review of small object detection algorithms based on deep learning [J]. Computer Engineering and Applications, 2023, 59(11): 16–27.
- [22] OUYANG D L, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning [C]//ICASSP 2023 – 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). June 4–10, 2023, Rhodes Island, Greece. IEEE, 2023: 1–5.
- [23] AZAD R, NIGGEMEIER L, HÜTTEMANN 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 (WACV). January 3–8, 2024, Waikoloa, HI, USA. IEEE, 2024: 1287–1297.