

基于改进 YOLACT 的果树叶墙区域实时检测方法

肖珂^{1,2} 梁聪哲¹ 夏伟光¹

(1. 河北农业大学信息科学与技术学院, 保定 071001; 2. 河北省农业大数据重点实验室, 保定 071001)

摘要: 为了解决果园因农药过量使用导致的环境污染与农药浪费问题,提出了一种基于改进 YOLACT 的果树叶墙区域(Leaf wall area, LWA)实时检测方法,用于计算深度-彩色双目相机采集视频中的叶墙区域距离及密度,为果园农药智慧喷施作业中农药喷洒剂量与喷洒距离的实时调整提供依据。首先,使用 ConvNeXt 主干网络改进了 YOLACT 模型,并引入 NAM 通道注意力机制对模型进行了优化;其次,提出了基于深度学习的果树叶墙密度检测方法;最后,通过阈值法排除深度图像中的干扰信息,简化了果树叶墙平均距离计算方法的处理流程。实验结果表明,改进 YOLACT 模型分割的 AP_{all} 为 91.6%,相较于原始模型上升 3.0 个百分点,与 YOLACT++、Mask R-CNN 和 QueryInst 模型相比分别高 2.9、1.2、4.1 个百分点;叶墙密度估计算法在叶墙顶部、中部和底部的均方根误差(Root mean square error, RMSE)分别为 1.49%、0.82%、2.20%;叶墙区域实时检测方法的处理速度可达 29.96 f/s。

关键词: 果树; 叶墙密度; YOLACT 模型; 目标检测; 实时; 平均距离

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2023)04-0276-09 OSID: 

Real-time Detection Method of Fruit Leaf Wall Area
Based on Improved YOLACT

XIAO Ke^{1,2} LIANG Congzhe¹ XIA Weiguang¹

(1. College of Information Science and Technology, Hebei Agricultural University, Baoding 071001, China
2. Hebei Key Laboratory of Agricultural Big Data, Baoding 071001, China)

Abstract: To reduce the environmental pollution and pesticide waste in orchards, a real-time method to detect the fruit tree leaf wall area (LWA) based on the improved YOLACT model was proposed to estimate average distance and density in the videos that captured by depth-color binocular camera, which can provide data for the real-time adjustment of pesticide spraying dose and spraying distance on intelligence pesticide spraying. Firstly, the YOLACT model was improved by using the ConvNeXt backbone network, and the NAM channel attention mechanism was introduced to optimize the model. Secondly, a leaf wall density estimation method based on deep learning was proposed. Finally, the average distance calculation method of LWA was proposed by excluding the interference information in the depth image through the threshold algorithm to simplify processing flow. The experimental results showed that the segmentation AP_{all} metrics of the improved YOLACT model was 91.6%, which was increased by 3.0 percentage points compared with that of the original model, and 2.9 percentage points, 1.2 percentage points and 4.1 percentage points compared with that of YOLACT++, Mask R-CNN, and QueryInst. The root mean square error (RMSE) of the leaf wall density estimation method was 1.49%, 0.82% and 2.20%. And the processing speed of the real-time LWA detection method could reach 29.96 f/s.

Key words: fruit tree; leaf wall density; YOLACT model; target detection; real-time; average distance

0 引言

农药的使用对果园中果树的病虫害防治起着关

键的作用,但同时过量的农药也会对人类的身体健康和周围的环境造成一定的危害^[1-3]。因此,如何在保证果树高质量生产的同时,降低农药的使用剂

量,是目前研究者们重点研究方向。变量喷施技术^[4-7]已然成为当下果园农药喷施研究领域的热点。变量喷施技术需要解决的一个关键性问题是如 何精准地检测到喷施区域,这依赖于精准的目标检测技术。目前,应用于果园的目标检测技术主要有超声波^[8-9]、激光雷达^[10-12]、机器视觉等方法。虽然这些方法都可以进行果园内精准目标检测,但是超声波检测法受地形影响大,激光雷达成本高昂。因此,低成本高精度的机器视觉方法被越来越多的研究者所关注。

叶墙面积^[13-16]是一个反映果园果树可施药范围的常见指标,由果园面积、果树冠层高度和果树行间距 3 个参数组成。其中冠层高度与果树行间距为计算叶墙面积的关键参数,每 10 cm 的变化都会导致叶墙面积产生 3% 到 5% 的波动^[17]。经调查研究表明,以叶墙面积作为剂量表达的协调方法可以很容易地作为农药剂量调整的调节模型,因此本研究以叶墙区域作为检测对象。

机器学习中的传统方法^[18-19],特征提取的过程繁琐单一,只能提取到目标的浅层特征,在背景颜色较为分明的果园中可以达到很好的检测效果。但是在背景颜色较为复杂的果园场景下,无法精准地检测喷施区域。基于深度学习方法^[20-21]可以提取更具区别能力的深层次特征,在背景颜色较为复杂的果园场景下仍对目标具有良好的检测效果。并且基于深度学习的方法使用的是端到端的结构,简化了数据处理流程,使得检测更加高效。有些研究者不再局限于只使用单目 RGB 相机,以彩色相机和深度相机相结合的方式对果园内的果树进行综合检测,从而获取更多、更有价值的信息。KIM 等^[22]设计了基于果园果树语义分割的智能喷洒系统,通过彩色相机和深度相机采集了 RGB 数据与深度数据,将深度图像与彩色图像相融合,并通过设置合理距离阈值的方式去除了 RGB 图像中的干扰背景,经训练后的 SegNet 网络模型准确率达到 83.79%。ZHANG 等^[23]提出了基于深度学习的苹果树分支检测方法,使用 Kinect V2 相机采集获取果园内的 RGB 图像数据与点云数据,从而获得并使用 Simple - RGB 图像和 Foreground - RGB 图像苹果树分支图像训练 SegNet 网络模型,经训练后的模型对 Simple - RGB 图像的分割准确率分别为 0.82 和 0.89,对 Foreground - RGB 图像的分割准确率分别为 0.91 和 0.92。这些方法仍具有一些缺点,它们只是使用深度数据对 RGB 数据进行前景切割,用于辅助训练深度模型,并没有提取喷施设备到喷施区域的距离信息。

为了解决上述问题,本文利用 RGB 相机与深度相机,拟研究一种基于改进 YOLACT 模型的果树叶墙区域实时检测方法。该方法使用深度学习技术对叶墙区域进行检测,通过改进后的 YOLACT 模型从 RGB 相机所采集的 RGB 彩色图像中提取叶墙区域并计算叶墙密度信息,为农药喷洒量提供控制参数。同时,通过深度相机所采集的深度图像提取果树叶墙与喷施设备的平均距离信息,为喷头调整农药喷洒距离提供控制参数。最终,构建出果树喷施区域实时检测模型。

1 材料与方法

1.1 数据获取

本研究在位于河北省保定市非结构化桃园内(38°56′53.408 4″N,115°26′0.416 4″E),采集了果树叶墙区域的 RGB 数据与深度数据。采集时间为 2021 年 9 月 30 日(当天温度为 11℃ 至 20℃,风力为西风 2 级)和 2021 年 10 月 13 日(当天温度为 15℃ 至 25℃,风力为东北风 3 级)。考虑到农药喷洒过程中可能遇到的情况,采集了多种状况下的叶墙区域图像(例如:高密度叶墙区域、低密度叶墙区域、有间隔叶墙区域)。部分图像的拍摄距离与农药喷洒距离相同。采集彩色 RGB 数据使用的设备是 SONY 公司生产的 DSC - RX0M2 型数码相机,分辨率为 1 980 像素 × 1 080 像素。采集深度数据使用的设备是 PMD 公司生产的 pico monstar 105 型深度相机,分辨率为 352 像素 × 287 像素。

1.2 深度学习数据集构建

使用 FFMPEG 工具对彩色 RGB 视频进行处理,从中截取多种状态下的叶墙区域视频帧共 400 幅。为了增强模型的鲁棒性和泛化能力,使其能够满足在不同环境下的农药喷施作业需求,使用数据增强技术扩充原始数据。使用的数据增强技术有水平与垂直翻转、亮度增强与减弱、高斯模糊。其中使用水平与垂直翻转来保证树形的多样性,使用图像亮度增强与减弱来模拟不同光照下的果园环境,使用高斯模糊来模拟处于运动模糊状态下的图像。最终将原始的 400 幅图像扩充至 1 605 幅图像,并划分为没有重叠的 3 个子集。其中 1 205 幅图像划分到训练集,200 幅划分到验证集,200 幅划分到测试集,最终构建了一个丰富的彩色 RGB 图像数据集。

1.3 叶墙区域检测方法概述

本文方法以 RGB 彩色相机与深度相机相结合的方式,对叶墙区域进行综合检测,整个方法流程分为两部分,图 1 为本文方法流程图。

第 1 部分是 RGB 数据处理:首先,由彩色相机

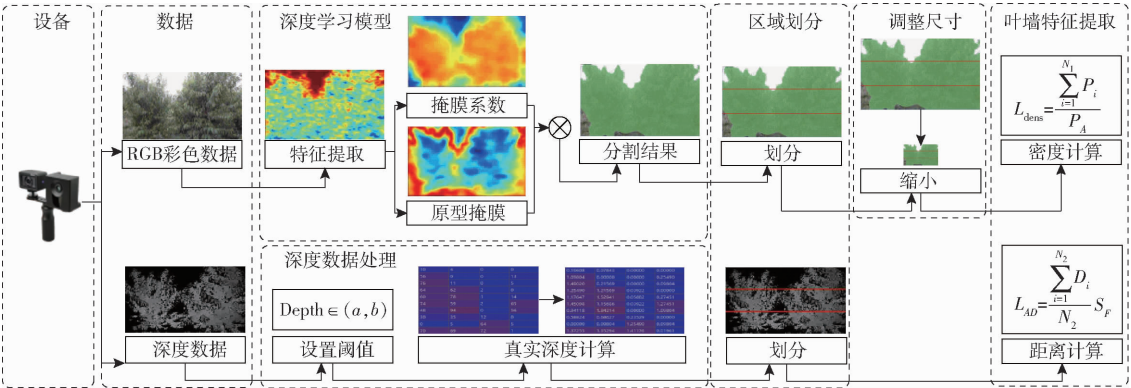


图 1 本文方法流程图

Fig.1 The proposed method flow chart

收集 RGB 数据;其次,经由 YOLACT 深度学习模型检测 RGB 图像中的叶墙区域;再次,划分喷施区域并缩小图像尺寸;最后分别计算出叶墙区域各部分的密度。第 2 部分是深度数据处理:首先,由深度相机收集深度数据;其次,根据设置的深度阈值屏蔽干扰信息并通过深度图比例因子将深度图中的像素信息转换为真实的距离信息;再次,划分喷施区域;最后,计算叶墙区域各部分到喷施设备的平均距离。

1.4 改进的 YOLACT 叶墙区域实例分割模型

YOLACT 模型是 Bolya 等^[24] 基于 One-stage 目标检测模型提出的实时实例分割模型,它将实例分割任务分为了原型掩膜 (Prototype mask) 生成与每个实例的掩膜系数 (Mask coefficients) 预测两个子任务。最后通过原型掩膜与掩膜系数的线性组合生成实例掩膜 (Instance masks)。YOLACT 模型的检测速度快,模型参数量少,可以实时对叶墙区域进行检测,但检测精度略有不足。为了提高对叶墙区域的检测精度,本文对 YOLACT 模型进行了改进。首先将模型的主干网络进行了替换,使用了特征提

取能力更强的 ConvNeXt 主干网络,然后在此基础上引入注意力机制,将 NAM 通道注意力机制添加到 ConvNeXt Block 内部与 FPN 的特征图输出部位。改进的 YOLACT 叶墙区域实例分割模型如图 2 所示,红色虚线框区域为改进部分。

1.4.1 深度特征提取主干网络 ConvNeXt

YOLACT 原始模型的默认主干网络为 ResNet,由于其特征提取能力不足,无法在复杂果园背景下对叶墙区域特征进行有效提取,最终导致模型对叶墙区域的检测精度较低。因此,模型需要特征提取能力更强、网络架构更加完善的主干网络。ConvNeXt 网络是由 Liu 等^[25] 通过借鉴以往研究者搭建网络的经验提出的纯卷积架构主干网络。在宏观设计上,ConvNeXt 与 ResNet 都是采用四阶段分布设计,但在 4 个阶段的 Block 堆叠比例上,两者有所差异,不同于 ResNet 的 [3:4:6:3],ConvNeXt 采用的是更加合理的 [1:1:3:1] 堆叠比例,让靠后的阶段拥有更多的计算量,优化了网络的结构。在卷积层方面 ConvNeXt 引入了分组卷积的思想,将卷积

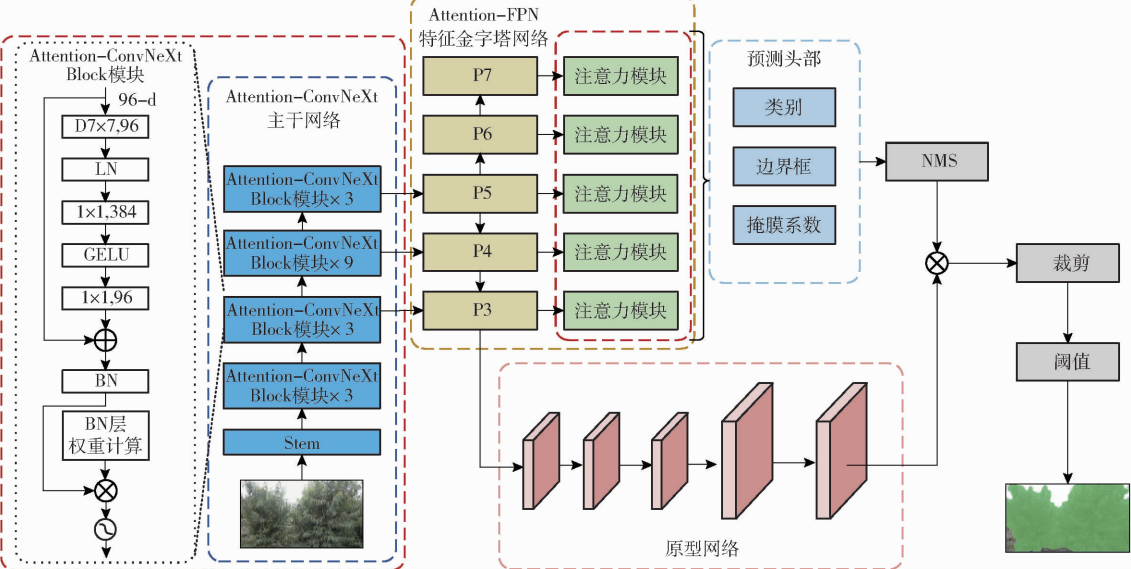


图 2 改进 YOLACT 模型

Fig.2 Improved YOLACT model

核为 3 的卷积层全部替换为分组卷积层,减少了模型的计算量,同时为了弥补模型容量的下降,将基础通道数设置为 96,这样在增加了较小计算量的同时提高了模型的精准度。在 Block 结构设计方面不同于 ResNet 的瓶颈层设计,ConvNeXt Block 采用了逆瓶颈层设计,调整了 Block 中每层计算量,在提高精准度的同时减少了模型整体的计算量。ConvNeXt 也进行了一些优化,例如将激活函数 ReLU 替换为 GELU、将 BN 替换为 LN、减少了激活函数与归一层的数量等,在计算量不变的情况下略微提升了模型的精准度。考虑到本研究检测种类的单一性,因此选择了与 RseNet-50 相对应的 ConvNeXt-T 量级的 ConvNeXt 网络作为模型的主干网络,在保证语义特征足够的前提下,降低网络的深度,提高模型对叶墙区域的检测精度。经实验表明,相对于使用其他主干网络而言,以 ConvNeXt 作为主干网络的 YOLACT 模型,具有更强的特征提取能力。

1.4.2 深度特征提取模块改进

YOLACT 模型在进行特征提取时,由于对重要特征区域的关注度不足,导致其特征提取效率较低。尤其是果园的背景颜色较为复杂,更加需要提升模型对关键特征的关注度,进而降低模型错检、漏检几率。NAM 注意力机制^[26]以批归一化 (Batch normalization, BN) 的缩放因子来表示权值的重要程度,通过对权值贡献因子的处理,可以实现有选择性抑制不重要特征的权重,达到对重要特征精准定位的效果。并且,由于是对 BN 层的缩放因子进行处理,所以避免了增加额外的参数量。因此,为了使模型提升对叶墙区域特征的关注度,降低模型的错检、漏检几率,本研究将 NAM 注意力机制的通道注意力模块与 ConvNeXt Block 模块进行融合,形成新的 Attention-ConvNeXt Block 模块(图 3),提高主干网络的特征提取效率。同时,通过在 FPN 每层卷积相对应的特征输出位置添加 NAM 通道注意力模块,将其改进为 Attention-FPN,提高其多尺度特征的提取效率。

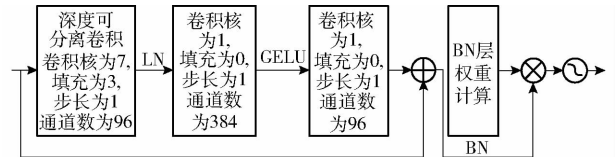


图3 Attention-ConvNeXt Block 示例图

Fig.3 Attention-ConvNeXt Block example diagram

1.5 叶墙密度估计算法

ZENG 等^[27]的研究表明冠层密度信息可以对农药精准喷施提供基础数据支持。为了实时控制农机喷洒农药的剂量,提出了一个叶墙密度估计算法。

该方法分为 3 个阶段:首先,RGB 视频帧经由改进的 YOLACT 模型进行实例分割;其次,为了降低本方法的时间成本且不影响深度模型的检测精度,本研究在密度信息进行提取之前,对改进 YOACLT 模型的检测结果进行尺寸缩小至 1/10 的处理,由原始的 1 920 像素×1 080 像素缩小到 192 像素×108 像素;最后,由于当叶墙区域出现过大大间隔时,会出现多个检测目标,同时模型也会保存多个检测结果,因此,需要对多个检测结果的掩膜区域像素数进行统计,并对叶墙密度进行计算。叶墙密度计算公式为

$$L_{dens} = \frac{\sum_{i=1}^{N_1} P_i}{P_A} \times 100\%$$

式中 L_{dens} ——叶墙密度,%
 P_i ——第 i 个检测结果中掩膜部分的像素数
 N_1 ——深度模型分割结果个数
 P_A ——所有像素数

1.6 叶墙平均距离估计算法

GU 等^[28]的研究表明喷雾距离会对农药药液的分布产生显著影响。本研究为了实时控制农机农药的喷洒距离,对 GAO 等^[29]提出的叶墙距离估计算法进行了改进,通过设置深度阈值获取有效距离信息,极大地简化了算法的处理流程。首先通过设置深度阈值过滤干扰信息,然后对叶墙到喷施设备的平均距离进行计算。叶墙平均距离计算公式为

$$L_{AD} = \frac{\sum_{i=1}^{N_2} D_i}{N_2} S_F \quad (D_i \in (a, b))$$

式中 L_{AD} ——叶墙到喷施设备的叶墙平均距离,cm
 D_i ——深度图像中像素 i 的信息
 a, b ——深度阈值
 N_2 ——深度图像中像素值位于 (a, b) 阈值内的像素数
 S_F ——深度图比例因子,取 1.96

1.7 深度学习模型的实验环境与训练参数

网络模型训练与测试的实验环境配置如表 1 所示。

表 1 网络模型训练与测试实验环境配置
Tab.1 Network model training and testing experimental environment configuration

环境属性	环境配置
CPU	Intel i7-9700
GPU	NVIDIA GeForce RTX2080 SUPER
内存/GB	32
深度学习框架	Pytorch
编程语言	Python
操作系统	Windows

本研究使用 AdamW 优化器对网络进行端到端联合方式的训练,权重衰减为 0.05;平均值的系数使用的是默认值(0.9,0.999);数据集训练次数设置为 300;批处理量设置为 2;输入网络图像设置为 224 像素×224 像素;起始学习率为 0.000 2;学习率衰减策略选择余弦退火(Cosine annealing)策略。此外,为了避免起始学习率过高而导致的梯度爆炸问题,采用了训练热身策略(Warmup):在训练的前 20 轮中,学习率从 0.000 035 缓慢上升到 0.000 2,在 20 轮后学习率才开始逐渐下降。在训练模型的过程中,它被设置为每一轮输出一次模型。

1.8 深度学习模型评价标准

因为本研究的数据集是按照 COCO 数据集格式制作而成,故选用平均精度(Average precision, AP)作为模型检测效果的评价标准。具体来说使用的是平均精度均值(Mean average precision, mAP),它是所有类的平均精度的平均,由于本研究检测类型单一,所以平均精度等价于平均精度均值。使用的平均精度标准主要为 AP₇₅与 AP_{all},AP₇₅表示交并比阈值为 0.75 时的平均测量值,AP_{all}表示在不同 IoU 阈值(0.5~0.95,步长 0.05)下的平均测量值。IoU 是模型对某一类别预测结果和真实值的交集与并集的比值,即模型产生的预测区域与原标注区域的重叠度。该值越高,模型分割效果越好,最理想的情况下是分割区域与原标注区域完全重叠,即交并比为 1。

2 实验与结果分析

2.1 深度学习模型

2.1.1 多种主干网络分割结果对比

本研究在同一数据集上对使用 ResNet-50、ResNeXt-50、ResNeSt-50、Swin Transformer-T 和 ConvNeXt-T 5 种不同主干网络的 YOLACT 模型进行了训练与测试。训练损失曲线如图 4 所示。测试集精度如表 2 所示。ConvNeXt-T 相比 ResNet-50,可以明显发现使用 ConvNeXt-T 主干网络后,模型训练阶段的损失收敛速度、模型收敛之后的损失值和测试集测试后的结果,都优于使用 ResNet-50 主干网络。ConvNeXt-T 相比 Swin Transformer-T,在 25 000 迭代次数之前的训练阶段,两者收敛速度相当,损失值相差无几,但随着训练次数的增加,ConvNeXt-T 损失曲线逐渐低于 Swin Transformer-T 的损失曲线,且经测试集测试后,ConvNeXt-T 分割与边界框的 AP_{all}与 AP₇₅指标均高于 Swin Transformer-T。虽然 ConvNeXt-T 与 ResNeXt-50、ResNeSt-50 3 个主干网络收敛后的损失值并无太大差异,但 ConvNeXt-T 在训练时的损失收敛速

度优于 ResNeXt-50 和 ResNeSt-50,并且如表 2 所示,ConvNeXt-T 在测试集上的测试结果高于其他 2 个主干网络,分割的 AP_{all}与 AP₇₅达到 90.5%和 100%,边界框的 AP_{all}与 AP₇₅达到 94.8%和 100%。ConvNeXt 的性能之所以优于其他主干网络,是因为其综合了 ResNeXt、Swin Transformer、MobileNet_v2 主干网络的优点,以纯卷积架构为基础,发挥了更强的性能。因此,本研究选择特征提取能力相对较强的 ConvNeXt-T 作为叶墙区域检测模型的主干网络。

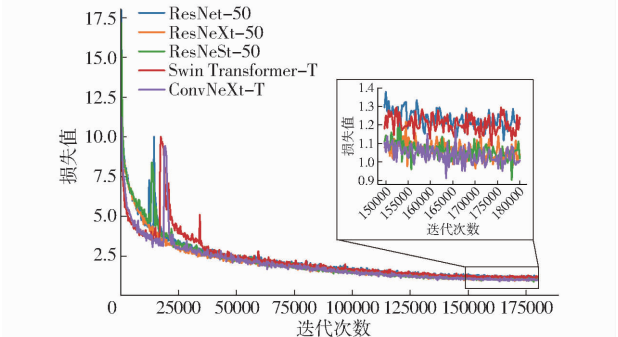


图 4 YOLACT 模型使用不同主干网络的训练损失曲线
Fig.4 Training loss curves of YOLACT model by using different backbone networks

表 2 5 种主干网络在测试集中的测试结果
Tab.2 Test results of five types of backbone networks in test set %

主干网络	分割平均精度		边界框平均精度	
	AP _{all}	AP ₇₅	AP _{all}	AP ₇₅
ResNet-50	88.6	98.0	92.8	98.9
ResNeXt-50	89.2	98.0	93.5	99.0
ResNeSt-50	89.5	99.0	91.9	97.9
Swin Transformer-T	90.0	98.0	91.1	98.9
ConvNeXt-T	90.5	100	94.8	100

2.1.2 消融实验结果与分析

消融实验可以用于验证模型的超参数或模块对整体结构产生的影响。本研究统一了模型超参数,在相同的数据集上进行了消融实验,测试集测试结果如表 3 所示,训练损失曲线如图 5 所示。可以看出,更换主干网络为 ConvNeXt-T 后,模型边界框的

表 3 消融实验测试结果
Tab.3 Ablation experiment test results %

模型	分割平均精度		边界框平均精度	
	AP _{all}	AP ₇₅	AP _{all}	AP ₇₅
YOLACT	88.6	98.0	92.8	98.9
YOLACT + ConvNeXt	90.5	99.0	94.8	100
YOLACT + NAM 通道注意力	89.6	99.0	93.4	100
YOLACT + ConvNeXt + NAM 通道注意力	91.6	100	96.3	100

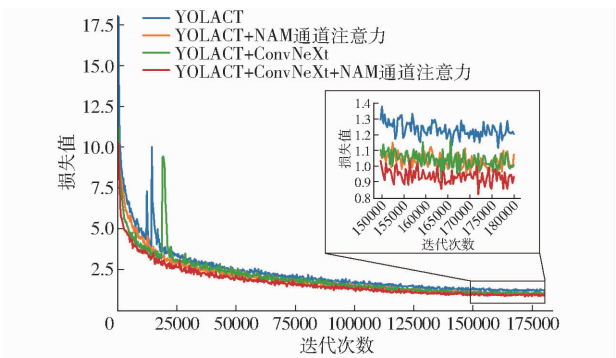


图 5 消融实验训练损失曲线

Fig. 5 Training loss curves of ablation experiment

AP_{all} 与 AP_{75} 达到 94.8% 与 100%, 与原模型相比分别上升 2.0、1.1 个百分点, 分割的 AP_{all} 与 AP_{75} 达到 90.5% 与 99.0%, 与原模型相比分别上升 1.9、1.0 个百分点。并且模型的训练损失收敛速度提升, 收敛后损失值下降约 0.2。这是由于 ConvNeXt-T 网络的特征提取能力强于 ResNet-50 网络, 这使得模型在进行后续的特征处理工作时会更加高效。添加 NAM 通道注意力机制后, 模型边界框的 AP_{all} 与 AP_{75} 达到 93.4% 与 100%, 与原模型相比分别上升 0.6、

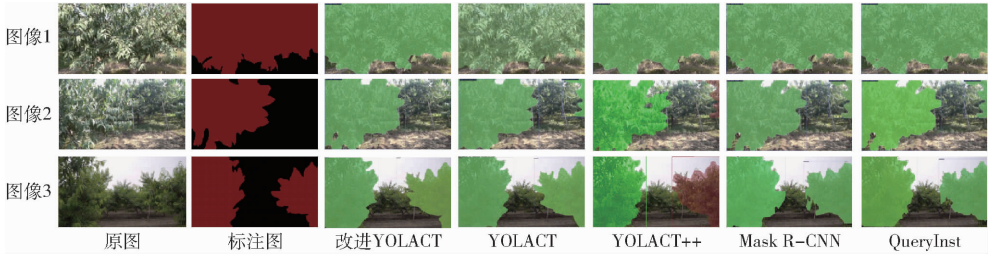


图 6 各模型分割效果

Fig. 6 Segmentation effect of each model

从图 6 中的图像 1 可以看出, 在叶墙区域是高密度且背景颜色分明的情况下, 这些算法的检测效果并没有太大差异, 叶墙区域的边缘都很好地分割了出来。然而在实际的果园环境中, 叶墙区域会出现有间隔的情况。这种情况下, 图像背景与目标区域的颜色、纹理等特征十分相近, 例如图 6 中的图像 2 与图像 3, 这使得模型对叶墙区域的检测更加困难, 特别是对边缘信息的检测。通过对模型分割结果进行分析比较, 发现在处理有间隔且背景颜色复杂的叶墙区域图像方面, 本文模型比其它模型有着更好的检测结果。如表 4 所示, 改进 YOLACT 模型分割的 AP_{all} 与 YOLACT、YOLACT++、Mask R-CNN 和 QueryInst 模型相比分别高 3.0、2.9、1.2、4.1 个百分点, 检测框的 AP_{all} 和 AP_{75} 均高于其他模型。尽管检测过程中存在一些错误检测和不确定性, 但仍然可以看出, 与其他模型相比, 本文提出的叶墙区域检测模型能够更精准地检测到叶墙区域, 边缘分割也明显优于其他方法。除了分析评估模型

1.1 个百分点, 分割的 AP_{all} 与 AP_{75} 达到 89.6% 与 99.0%, 与原模型相比都上升 1.0 个百分点。并且, 模型的训练损失收敛速度提升, 收敛后的损失值有所下降。这是由于 NAM 通道注意力机制相当于一个筛选机制, 对训练过程中发现的不重要特征进行惩罚, 减少这些特征的权重, 增加了重要特征所占的权值比重, 使模型更加关注重要特征, 故模型的性能会有所提升。在使用 ConvNeXt-T 的基础上引入注意力机制后, 模型的性能进一步提升, 最终, 改进 YOLACT 叶墙区域分割模型其边界框的 AP_{all} 与 AP_{75} 分别达到 96.3% 与 100%, 与原模型相比分别提高 3.5、1.1 个百分点, 分割的 AP_{all} 与 AP_{75} 分别达到 91.6% 与 100%, 与原模型相比分别提高 3.0、2.0 个百分点。

2.1.3 深度模型间检测实验结果分析

为了进一步分析叶墙区域分割模型的性能, 将改进 YOLACT 模型与 YOLACT、YOLACT++、Mask R-CNN^[30]、QueryInst^[31] 进行比较。其中 YOLACT 是最基础的模型。经测试集测试后的结果如图 6 所示。

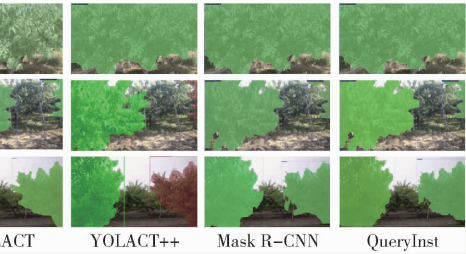


表 4 5 种分割模型在测试集中的测试结果

Tab. 4 Test results of five segmentation models in test set

模型	分割平均精度/%		边界框平均精度/%		单幅图像检测时间/ms
	AP_{all}	AP_{75}	AP_{all}	AP_{75}	
YOLACT	88.6	98.0	92.8	98.9	13
YOLACT++	88.7	96.9	89.2	97.8	13
Mask R-CNN	90.4	97.6	94.0	98.0	62
QueryInst	87.5	98.5	94.3	96.8	122
改进 YOLACT	91.6	100	96.3	100	15

的检测精度外, 还需考虑模型运行时间。从表 4 可以看出, 改进 YOLACT 模型检测 1 幅图像的时间为 15 ms, 相较于 YOLACT、YOLACT++ 这种 One-stage 模型速度有所下降, 这是由于模型的参数量有所增加, 模型计算复杂度提升所导致。但与 Mask R-CNN、QueryInst 这种 Two-stage 模型相比, 改进 YOLACT 模型的速度仍然具有很大优势, 能够进行实时目标实例分割。经过以上分析表明, 本研究所提出的模型具有较强的抗干扰能力和泛化能力。

2.2 叶墙区域检测算法检测效果分析

2.2.1 叶墙密度估算效果分析

本研究通过估算叶墙区域密度,对农机调节农药喷洒剂量提供信息支持。为了验证叶墙区域密度估算方法的有效性,本研究通过对图像中的叶墙区域进行人工标注并计算得到叶墙密度的真实测量值,之后将本方法估算的叶墙密度信息与真实测量的叶墙密度信息进行对比实验,如图 7 所示。本研究从一桃园 RGB 视频中按时间顺序截取了 30 幅视频帧样本用于本次实验。由图 7a 可以看出,与顶部和底部区域相比,叶墙中部区域的密度相对较高。这是由于这一区域的叶子往往生长得更加茂密,占据了相对较大的空间。而顶部与底部区域由于叶子

稀疏,背景包含了部分天空与地面区域,导致密度较低。图 7b~7d 分别为叶墙顶部、中部与底部区域叶墙密度估计值与测量值,可以看出各样本的估算值在测量值周围小幅度上下波动,并无太大差距。此外,本研究计算了 30 个样本测量值与估算值之间的误差,以平均误差 (Mean error, ME) 和均方根误差作为叶墙 3 部分区域的误差指标,如表 5 所示。其中 3 部分区域平均误差最大的是顶部区域,平均误差为 1.19%,最小的为中部区域,平均误差为 0.13%。叶墙顶部、中部、底部 3 部分区域的均方根误差为 1.49%、0.82% 和 2.20%,由此可见估计值与测量值的误差较小,可以证明估计值的准确性和可靠性。

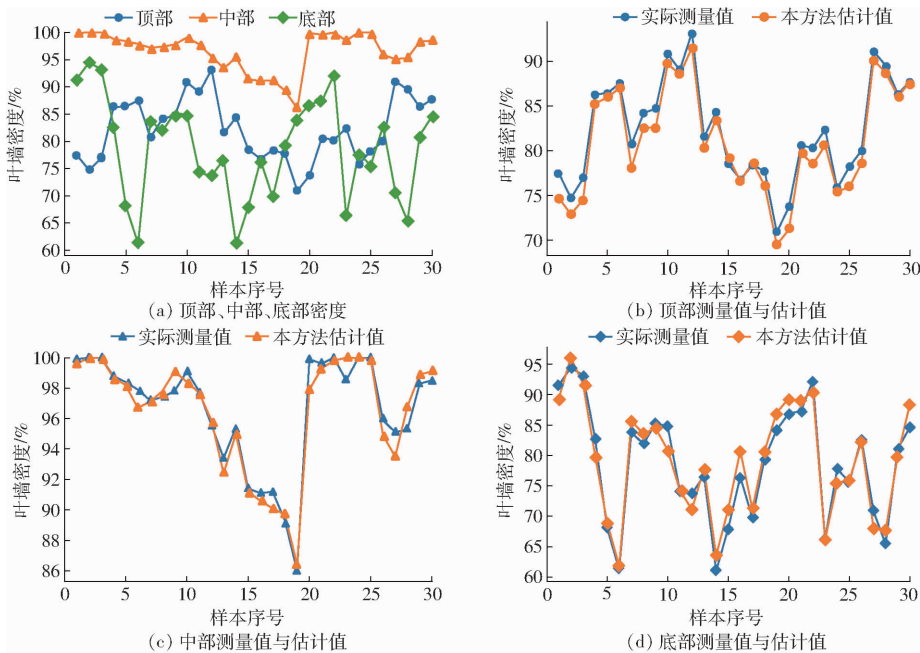


图 7 叶墙顶部、中部、底部密度
Fig. 7 Density at top, middle and bottom of leaf wall

表 5 叶墙密度误差指标				%
Tab.5 Leaf wall density error				
叶墙 区域	估计值			平均测量 值
	平均值	平均误差	均方根误差	
顶部	80.99	1.19	1.49	82.19
中部	96.49	0.13	0.82	96.62
底部	79.11	0.39	2.20	78.71

2.2.2 叶墙平均距离提取效果分析

本研究通过估算喷头到叶墙的平均距离,为农机调整喷雾距离提供参考。假设深度相机与喷头的距离为 120 cm,叶墙区域的背面边缘部分距离相机最远为 300 cm。于是,将阈值设为 [120,300] cm,用于排除干扰信息。图 8 显示了 30 个深度视频帧样本的叶墙区域平均距离变化曲线。叶墙的中部和底部部分的距离最远,而顶部部分的距离最近。本

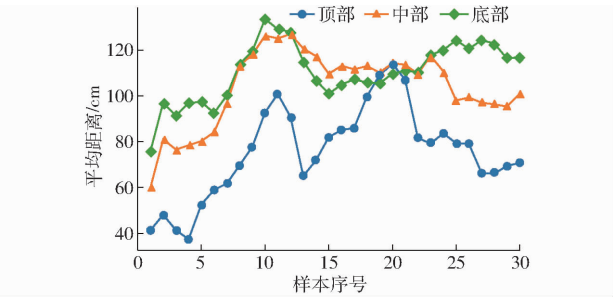


图 8 叶墙平均距离
Fig. 8 Average distance of leaf wall

研究计算了 30 个视频帧样本的平均距离,其叶墙顶部、中部与底部区域的平均距离分别为 75.5、100.5、106.9 cm。可以看出叶墙顶部区域的距离相对较近,这是由于果树顶部枝条向外较为发散所导致。

2.2.3 叶墙区域检测算法运行速度分析

通过对 30 组样本(1 幅 RGB 图像与 1 幅深度图像为 1 组)进行实验,验证本研究所提出方法的实时性。图 9 显示了这 30 组样本的处理时间。可以看出提取密度信息要比提取距离信息所耗时间更长,这是由于 RGB 图像中叶墙区域的背景复杂,需要经由深度学习技术对其进行深层次的特征提取。本研究计算了 30 组样本的平均耗时,其中提取密度信息的平均耗时为 33.37 ms,提取平均距离信息的平均耗时为 10.28 ms。本方法在串行处理的方式下处理速度为 22.90 f/s。但为了进一步缩短检测时间,本研究通过 Python 多进程技术对本方法进行优化,使密度信息与平均距离信息的提取并行进行。因此,只需对密度信息提取速度进行计算便可得出本方法整体的处理速度,经计算得本研究所提出的叶墙区域检测方法的处理速度可达 29.96 f/s,相较串行处理高 7.06 f/s,这说明了并行处理方法对节约时间成本有着显著的成效。

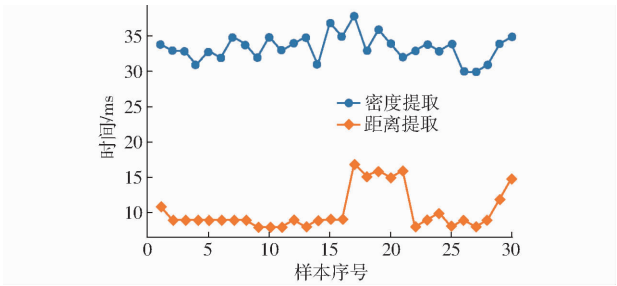


图 9 密度与距离信息提取时间

Fig.9 Density and distance information extraction time

3 结论

(1) 对 YOLACT 模型进行了改进。使用 ConvNeXt-T 主干网络替换了原始的 ResNet-50 主干网络,增强了模型的特征提取能力。并在此基础上,引入了 NAM 通道注意力机制,使模型更加注重关键特征的提取,提升了模型特征提取效率。经实验后表明,改进后的 YOLACT 模型分割的 AP_{all} 为 91.6%,与原始模型相比上升 3.0 个百分点,与 YOLACT++、Mask R-CNN 和 QueryInst 模型相比分别高 2.9、1.2、4.1 个百分点,在处理有间隔且背景颜色较为复杂的叶墙区域图像方面,本文模型比其它模型有着更好的检测效果,即该模型具有更好的普适性和鲁棒性。

(2) 构建了一个果树喷施区域实时检测模型。基于改进的 YOLACT 模型提出了叶墙密度估算方法,通过模型分割出的目标区域像素数除以当前区域总像素数得出叶墙密度,经实验表明,该方法得出的估计值与实际测量值的均方根误差分别为 1.49%、0.82%、2.20%,证明该方法具有一定的准确性和可靠性。同时,优化了叶墙平均距离估算方法,通过设置深度阈值排除干扰信息,然后计算出每个区域的平均叶墙距离。在处理速度方面,本研究使用并行处理的方式对叶墙的密度与平均距离信息进行提取,经实验表明,该方法的处理速度可达 29.96 f/s,能够实时进行叶墙区域检测。

参 考 文 献

[1] WANDKAR S V, BHATT Y C, JAIN H K, et al. Real-time variable rate spraying in orchards and vineyards: a review[J]. Journal of the Institution of Engineers: Series A, 2018, 99(2): 385-390.

[2] 翟长远,赵春江,WANG Ning,等. 果园风送喷雾精准控制方法研究进展[J]. 农业工程学报,2018,34(10):1-15. ZHAI Changyuan, ZHAO Chunjiang, WANG Ning, et al. Research progress on precision control methods of air-assisted spraying in orchards[J]. Transactions of the CSAE, 2018, 34(10): 1-15. (in Chinese)

[3] ZHANG Z Z, WANG X Y, LAI Q H, et al. Review of variable-rate sprayer applications based on real-time sensor technologies[J]. Automation in Agriculture-Securing Food Supplies for Future Generations, 2018, 8: 53-79.

[4] 魏新华,蒋杉,孙宏伟,等. PWM 间歇喷雾式变量喷施控制器设计与测试[J]. 农业机械学报, 2012, 43(12): 87-93,129. WEI Xinhua, JIANG Shan, SUN Hongwei, et al. Design and test of variable rate application controller of intermittent spray based on PWM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(12): 87-93,129. (in Chinese)

[5] TACKENBERG M, VOLKMAR C, SCHIRRMANN M, et al. Impact of sensor-controlled variable-rate fungicide application on yield, senescence and disease occurrence in winter wheat fields[J]. Pest Management Science, 2018, 74(6): 1251-1258.

[6] 莽璐,祁力钧,冀荣华,等. 温室自动变量施药系统设计[J]. 中国农业大学学报, 2009, 14(4): 114-118. MANG Lu, QI Lijun, JI Ronghua, et al. Design of variable rate spray system in greenhouse[J]. Journal of China Agricultural University, 2009, 14(4): 114-118. (in Chinese)

[7] 王润涛,刘瑶,王树文,等. 基于模糊控制的车速跟随变量喷雾系统设计与试验[J]. 农业机械学报,2022,53(6):110-117. WANG Runtao, LIU Yao, WANG Shuwen, et al. Design and experiment of speed-following variable spray system based on fuzzy control[J]. Transactions of the Chinese Society for Agricultural Machinery,2022,53(6):110-117. (in Chinese)

[8] MAGHSOUDI H, MINAEI S, GHOBADIAN B, et al. Ultrasonic sensing of pistachio canopy for low-volume precision spraying[J]. Computers and Electronics in Agriculture, 2015, 112: 149-160.

- [9] 邹伟, 王秀, 冯青春, 等. 基于超声靶标探测的果园变量喷药控制系统设计[J]. 农机化研究, 2021, 43(2): 58–63, 70. ZOU Wei, WANG Xiu, FENG Qingchun, et al. Design of variable spraying control system based on ultrasonic target detection [J]. Journal of Agricultural Mechanization Research, 2021, 43(2): 58–63, 70. (in Chinese)
- [10] 窦汉杰, 翟长远, 王秀, 等. 基于 LiDAR 的果园对靶变量喷药控制系统设计与试验[J]. 农业工程学报, 2022, 38(3): 11–21. DOU Hanjie, ZHAI Changyuan, WANG Xiu, et al. Design and experiment of the orchard target variable spraying control system based on LiDAR[J]. Transactions of the CSAE, 2022, 38(3): 11–21. (in Chinese)
- [11] NARVÁEZ F J, DEL PEDREGAL J S, PRIETO P A, et al. LiDAR and thermal images fusion for ground-based 3D characterisation of fruit trees[J]. Biosystems Engineering, 2016, 151: 479–494.
- [12] 牛润新, 张向阳, 王杰, 等. 基于激光雷达的农业机器人果园树干检测算法[J]. 农业机械学报, 2020, 51(11): 21–27. NIU Runxin, ZHANG Xiangyang, WANG Jie, et al. Orchard trunk detection algorithm for agricultural robot based on laser radar[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(11): 21–27. (in Chinese)
- [13] MOYA E G, GALLART M, LLORENS J, et al. Spray adjustments based on LWA concept in vineyard. Relationship between canopy and coverage for different application settings[C] // Aspects of Applied Biology 122, Internattional Advances in Pesticide Applications, 2014: 25–32.
- [14] GIL E, LLORENS J, LANDERS A, et al. Field validation of DOSAVIÑA, a decision support system to determine the optimal volume rate for pesticide application in vineyards[J]. European Journal of Agronomy, 2011, 35(1): 33–46.
- [15] WALKLATE P J, CROSS J V. Regulated dose adjustment of commercial orchard spraying products[J]. Crop Protection, 2013, 54: 65–73.
- [16] WALKLATE P J, CROSS J V, PERGHER G. Support system for efficient dosage of orchard and vineyard spraying products [J]. Computers and Electronics in Agriculture, 2011, 75(2): 355–362.
- [17] TOEWS R B, FRIESSLEBEN R. Dose rate expression—need for harmonization and consequences of the leaf wall area approach[J]. Erwerbs-Obstbau, 2012, 54(2): 49–53.
- [18] CHEN Y Y, HOU C J, TANG Y, et al. Citrus tree segmentation from UAV images based on monocular machine vision in a natural orchard environment[J]. Sensors, 2019, 19(24): 5558.
- [19] KESTUR R, ANGURAL A, BASHIR B, et al. Tree crown detection, delineation and counting in uav remote sensed images: a neural network based spectral-spatial method[J]. Journal of the Indian Society of Remote Sensing, 2018, 46(6): 991–1004.
- [20] 王辉, 韩娜娜, 吕程序, 等. 基于 Mask R-CNN 的单株柑橘树冠识别与分割[J]. 农业机械学报, 2021, 52(5): 169–174. WANG Hui, HAN Na'na, LÜ Chengxu, et al. Recognition and segmentation of individual citrus tree crown based on Mask R-CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 169–174. (in Chinese)
- [21] WEINSTEIN B G, MARCONI S, BOHLMAN S, et al. Individual tree-crown detection in RGB imagery using semi-supervised deep learning neural networks[J]. Remote Sensing, 2019, 11(11): 1309.
- [22] KIM J, SEOL J, LEE S, et al. An intelligent spraying system with deep learning-based semantic segmentation of fruit trees in orchards[C] // IEEE International Conference on Robotics and Automation, 2020: 3923–3929.
- [23] ZHANG J, HE L, KARKEE M, et al. Branch detection for apple trees trained in fruiting wall architecture using depth features and regions-convolutional neural network (R-CNN)[J]. Computers and Electronics in Agriculture, 2018, 155: 386–393.
- [24] BOLYA D, ZHOU C, XIAO F, et al. YOLACT: real-time instance segmentation[C] // Proceedings of the IEEE International Conference on Computer Vision, 2019: 9157–9166.
- [25] LIU Z, MAO H, WU C Y, et al. A convnet for the 2020s[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2022: 11976–11986.
- [26] LIU Y C, SHAO Z R, TENG Y Y, et al. NAM: normalization-based attention module[J/OL]. Computer Vision and Pattern Recognition[2021–11–24]. <https://arxiv.org/abs/2111.12419>.
- [27] ZENG L H, FENG J, HE L. Semantic segmentation of sparse 3D point cloud based on geometrical features for trellis-structured apple orchard[J]. Biosystems Engineering, 2020, 196: 46–55.
- [28] GU C C, LIU Z J, PAN G J, et al. Optimization of working parameters for 3MGY-200 axial air-assisted sprayer in kiwifruit orchards[J]. International Journal of Agricultural and Biological Engineering, 2020, 13(2): 81–91.
- [29] GAO G D, XIAO K, MA Y J. A leaf-wall-to-spray-device distance and leaf-wall-density-based automatic route-planning spray algorithm for vineyards[J]. Crop Protection, 2018, 111: 33–41.
- [30] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961–2969.
- [31] FANG Y X, YANG S S, WANG X G, et al. Instances as queries[C] // Proceedings of the IEEE International Conference on Computer Vision, 2021: 6910–6919.