

殷献博, 邓小玲, 兰玉彬, 等. 基于改进 YOLOX-Nano 算法的柑橘梢期长势智能识别 [J]. 华南农业大学学报, 2023, 44(1): 142-150.
YIN Xianbo, DENG Xiaoling, LAN Yubin, et al. Intelligent recognition of citrus shoot growth based on improved YOLOX-Nano algorithm[J]. Journal of South China Agricultural University, 2023, 44(1): 142-150.

基于改进 YOLOX-Nano 算法的柑橘梢期长势智能识别

殷献博[✉], 邓小玲[✉], 兰玉彬, 陈欣
(华南农业大学 电子工程学院/国家精准农业航空施药技术
国际联合研究中心, 广东 广州 510642)

摘要:【目的】采用机器视觉技术开展柑橘梢期的智能感知技术研究,以解决背景与目标颜色相似造成识别精度低的问题,实现柑橘梢期自动监测,探索算法的改进方法。【方法】根据不同卷积层提取特征的特点与不同注意力机制的作用,提出了一种基于多注意力机制改进的 YOLOX-Nano 智能识别模型,建立多元化果园数据集并进行预训练。【结果】改进的 YOLOX-Nano 算法使用果园数据集作为预训练数据集后,各类别平均精度的平均值 (Mean average precision, mAP) 达到 88.07%。与 YOLOV4-Lite 系列模型相比,本文提出的改进模型在使用较少的参数和计算量的情况下,识别精度有显著的提升,mAP 分别比 YOLOV4-MobileNetV3 和 YOLOV4-GhostNet 提升 6.58% 和 6.03%。【结论】改进后的模型在果园监测终端的轻量化部署方面更具有优势,为农情实时感知和智能监测提供了可行的数据和技术解决方案。

关键词: 柑橘; 新梢; 智能监测; 注意力机制; YOLOX-Nano; 深度学习; 智慧农业

中图分类号: S666.2; TP183

文献标志码: A

文章编号: 1001-411X(2023)01-0142-09

Intelligent recognition of citrus shoot growth based on improved YOLOX-Nano algorithm

YIN Xianbo[✉], DENG Xiaoling[✉], LAN Yubin, CHEN Xin

(College of Electronic Engineering, South China Agricultural University/National Center for International Collaboration
Research on Precision Agricultural Aviation Pesticide Spraying Technology, Guangzhou 510642, China)

Abstract: 【Objective】In order to solve the problem of low recognition accuracy due to similar color of the background and new shoots, to realize the automatic monitoring of citrus shoot stage and explore the improved method of algorithm, the machine vision technology was used to carry out the research on intelligent perceiving the growth stage of citrus shoot. 【Method】According to the characteristics of features extracted from different convolutional layers and the role of different attention mechanism, an improved YOLOX-Nano intelligent recognition model based on multi-attention mechanism was proposed, and a diversified orchard dataset for pre-training was established. 【Result】The improved YOLOX-Nano algorithm achieved the mAP (Mean average precision) of 88.07% using the orchard dataset as a pre-training dataset. Compared with the model of YOLOV4-Lite series, the improved model significantly improved the recognition accuracy with less parameters and calculation. Compared with YOLOV4-MobileNetV3 and YOLOV4-GhostNet, the mAP of improved model

收稿日期: 2021-12-24 网络首发时间: 2022-06-10 11:16:26

首发网址: <https://kns.cnki.net/kcms/detail/44.1110.S.20220609.1252.002.html>

作者简介: 殷献博, 硕士研究生, 主要从事智慧果园、机器视觉的研究, E-mail: 17320672746@163.com; 通信作者: 邓小玲, 副教授, 博士, 主要从事智慧果园、农业人工智能、无人机遥感方向的研究, E-mail: dengxl@scau.edu.cn

基金项目: 广州市重点研发计划 (202103000090); 广东省重点研发计划 (2019B020214003); 广东高校重点领域 (人工智能) 专项 (2019KZDZX1012); 高等学校学科创新引智计划 (D18019)

increased by 6.58% and 6.03% respectively. 【Conclusion】 The improved model has greater advantage for lightweight deployment at orchard monitoring terminal. The findings provide feasible data and technical solutions for agricultural real-time perception and intelligent monitoring.

Key words: Citrus; New shoot; Intelligent monitoring; Attention mechanism; YOLOX-Nano; Deep learning; Intelligent agriculture

柑橘是国内生产的主要水果^[1-2]。在传统柑橘果园管理过程中, 存在作业效率低、管理粗放、施肥施药过量以及人员作业存在风险等问题, 导致果园收益降低、农药残留污染, 不利于农业经济和绿色农业的发展。智慧农业是农业信息化发展从数字化到网络化再到智能化的高级阶段^[3]。智慧果园精准管控是实现果园智能化精准管理、增效减施、减少环境与人员危害的重要途径。近年来, 柑橘果园智能化管理的研究主要以柑橘病虫害为研究对象, 通过光谱、深度学习等技术快速获取果园病虫害信息, 实现果树病虫害级别快速定量化^[4-8]。在柑橘生产阶段的研究中, 现有的研究主要针对柑橘果实和花, 使用目标检测、语义分割等方法从模型轻量化、多尺度特征提取等多个方面改进模型, 获取果实和花的准确统计数据 and 定位信息, 为果园估产、自动采摘提供技术支持^[9-12]。这些研究主要集中于果园管理中的果实识别、产量预测、常见病虫害监测等问题, 关于柑橘新梢智能管理方面的研究鲜见报道。新梢作为柑橘植株的重要组成部分, 在柑橘生产过程中有着极其重要的作用。健壮的新梢将发育为结果母枝, 代表着下一年的产量; 而过多的夏梢会抢夺柑橘果实的营养造成果实膨胀期发育迟缓。研究表明, 新梢与柑橘潜叶蛾、木虱的发生具有较强的相关性^[13-14]。药物控梢作为主要的控梢方式, 容易因施药不当造成果实花皮、大量落果、树体黄化、树势衰退等现象^[15]。当前新梢管控多以人工巡园和种植经验为主, 通过进入果园内部查看新梢抽发量和生长阶段判断新梢管控作业节点。这种方式容易受主观因素的影响存在判断偏差, 而且果园面积较大时会出现生长阶段判断不准确的现象; 另外, 大面积果园中因不同区域的水肥或者果树长势差异导致新梢抽发情况不同, 人工巡园很难兼顾果园全部区域, 而且新梢一年抽发多次, 人工消耗巨大, 不利于大面积果园的新梢管控。因此, 使用计算机视觉技术结合农业物联网对新梢图像进行识别, 精准识别出各个新梢的当前生长阶段, 实现自动化和智能化的新梢监测, 这对于提高柑橘产量、品质以及减少劳动力消耗具有重要意义。

柑橘新梢识别难度主要在于目标颜色与植株背景颜色相近, 且随着新梢生长, 其形态、颜色逐渐接近正常枝叶。Li 等^[16]使用 RGB-D 相机结合 YOLOV3 目标检测算法检测茶叶“一尖两叶”区, 对茶叶嫩芽的检测精度达到了 93.1%。Xu 等^[17]使用 YOLOV3 识别图像中的茶蕾, 并对识别后的区域使用 DenseNet201 网络进一步分类, 剔除错误的识别结果, 可以使整体识别分类精度提高至 95.71%。袁加红等^[18]采用三基色以及组合因子对图像进行灰度化处理, 利用维纳滤波和梯度增强技术对处理的图像进行滤波去噪, 采用大津法和迭代法分割图像获取二值图像, 并基于二值化图像提取茶叶新梢的质心, 获取茶叶新梢的具体位置。Fang 等^[19]使用坐标注意力机制 (Coordinate attention, CA) 以及 Do-Conv 卷积改进的 YOLOV4-LITE 对生姜嫩芽检测, 确定新梢生长方向, 改进后的模型识别精度达到 98.73%, 计算量降低至 8.74 G。Scarlett 等^[20]提取图像中目标与其余物体的像素区分阈值, 并进行图像切分, 通过无监督的特征选择和聚类方法对葡萄新梢进行识别, 识别精度达到 86.83%, 可为葡萄早期产量估计提供支持。上述研究中, 针对新梢(嫩芽)的问题采用多种方法提高识别精度, 但均存在对数据采集设备要求高、数据预处理要求高、方法复杂度高或者需要多次处理以及处理时间长等问题, 不利于实时性要求高的农业应用和模型边缘化部署。针对新梢检测和识别的难题, 为兼顾 AI 模型的边缘化部署和实时解析, 在仅使用深度学习模型的前提下需要进一步提高模型的特征提取与区分能力。本研究在轻量化的 YOLOX-Nano 识别网络基础上, 通过使用多种注意力机制和多元化数据集提升算法对目标与背景的区分能力, 从而实现柑橘新梢智能检测和梢期的智能识别。

1 材料与方法

1.1 数据采集与预处理

本文所需柑橘新梢图像主要采集于广东省惠州市博罗县, 使用手机在距离目标 20~100 cm 的位

置进行拍摄,包括晴天、阴天、小雨等日常自然环境情况,共获得 3 000×4 000 像素图片 1 096 张。在此基础上,加入田间摄像头和网络上的新梢图像 123 张,共计 1 219 张原始图像。通过裁剪和筛选最终获得 608×608 像素的有效图像 1 681 张。根据柑橘新梢生长管控中对抽梢量、秋梢壮梢、冬梢防冻害促转绿的实际需求,以及新梢不同生长阶段的特征,将新梢分为萌芽期、生长期和展叶期,如图 1 所示。使用 LabelImg 图像标注工具对数据集进行标注,标注数据以 PASCAL VOC 的格式存储在 XML 文件中。

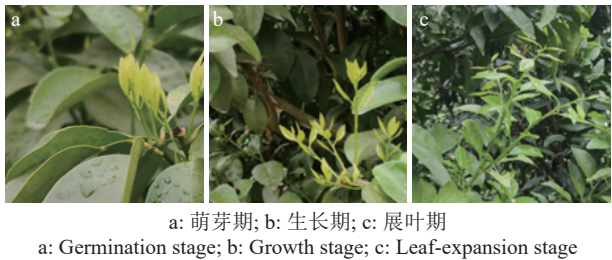


图 1 不同生长阶段的柑橘新梢
Fig. 1 Citrus shoots at different growth stages

为保证数据集的独立性,对数据集按照 8:1:1 比例进行训练集、验证集、测试集的切分,获得训练集图像 1 345 张、验证集图像 168 张、测试集图像 168 张。使用 Imgaug 数据扩充工具包对训练集进行扩充,扩充方法选用镜像、边缘检测、像素填充以及随机大小缩放等,并进行随机组合,通过筛选最终获得图像 5 440 张。

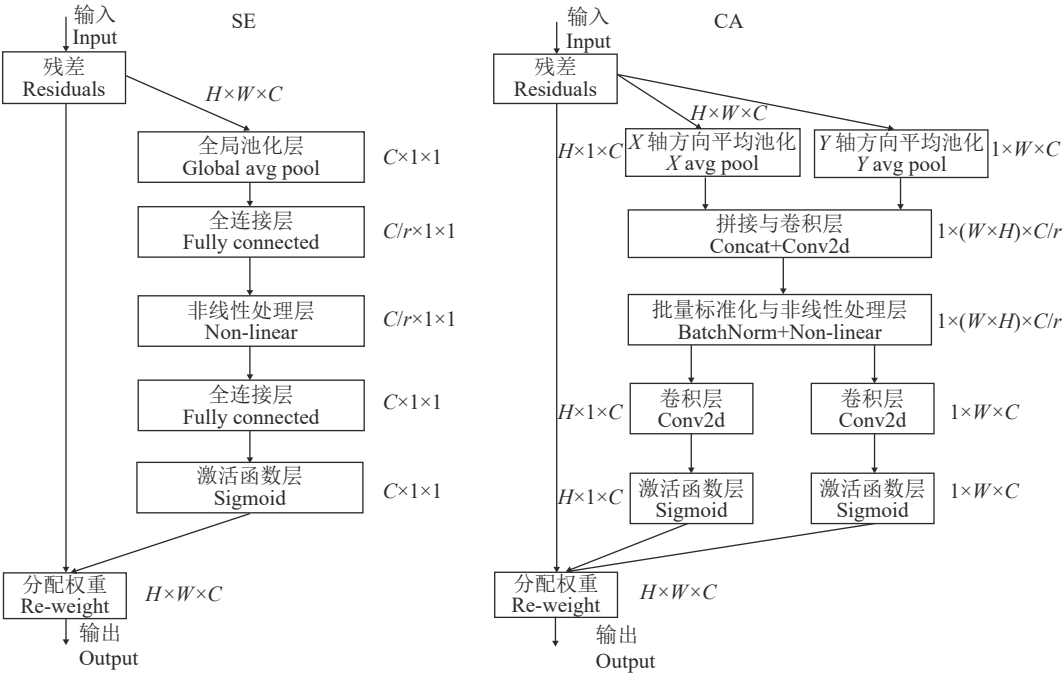
1.2 基于多注意力机制的柑橘梢期检测与识别模型

注意力机制可以使模型选择性地集中于部分信息,是模型的一种资源分配方案,可以有效地解决信息过载,同时在有效的计算资源下,使模型处理更重要的信息^[21]。标准卷积中无法实现对通道的相关性进行建模,所有通道处于平等地位,造成一些重要信息表达不够明显,注意力机制在深度学习模型中重新分配了信息的权重,通过迭代训练可以使重要的信息获得较高的权重,实现网络性能的提升。在农业环境应用中容易出现小目标、局部遮挡、背景复杂等问题,引入注意力机制可以在一定程度上改进模型性能^[22-24]。不同的注意力机制模块有着不同的作用,常见的通道注意力机制 (Squeeze-and-excitation attention, SE) 使用全局池化获得 1×1×C(C 为通道数),并使用 2 个全连接层和 1 个激活函数进行非线性处理,便于处理通道间的复杂相关性,最终获得 1×1×C 的通道权重与特征图层匹配^[25]。由于使用全局池,SE 将全局信息压缩为通道

权重,可以很好地确定不同通道间的重要性,但这种方式将不同通道的信息同等对待,忽略了通道内部特征的局域性。在较深位置,特征信息已被高度抽象到通道中,位置效果明显;但在网络初始阶段,特征较为具体、位置特征明显的图层中 SE 的作用有限。针对 SE 存在的缺陷,CA 沿特征图层的宽、高 2 个维度进行池化和卷积操作,从而获得特征编码,并在通道的 2 个方向上进行聚合。与 SE 仅简单地重新分配通道的权重相比,CA 可以实现沿一个空间方向捕获远程依赖关系、沿另一个空间方向保留精确位置信息,使模型更准确地定位并识别目标区域。Hou 等^[26]试验证明,在 ImageNet 分类任务和 COCO 目标检测任务中,CA 模型的精度比 SE 模型分别提升了 0.9% 和 0.8%,有效地提高了网络模型对信息的保存能力。SE 和 CA 的计算过程如图 2 所示。

本研究采用 YOLOX 的轻量化版本 YOLOX-Nano 作为基础网络,目标是实现对柑橘新梢的精准定位与梢期阶段的精准识别、实现边缘端智能和终端智能,使智能识别模型部署不依赖远程传输的果园环境。YOLOX-Nano 网络仅使用超参数 Depth 和 Width 控制 CSPLayer 中残差块的个数以及特征图的维度,Depth 取值 0.33、Width 取值 0.25。为进一步提高模型对柑橘新梢检测的性能,本研究对 YOLOX-Nano 网络进行了优化改良。由于轻量化后模型的通道数量减少,对通道进行权重分配使得重要的通道作用更加明显,并保证通道保存重要的目标空间和特征信息。本文在 YOLOX-Nano 网络中的 Focus 模块、SPP 模块以及 PAFPN 模块分别引入了注意力机制。由于 Focus 模块存在于网络的头部,直接面向原始图像,图像中存在丰富的空间位置信息和特征信息,因此如何保存精确的位置信息并捕获远程依赖关系对于网络特征提取十分重要。本文结合注意力机制的特点,选用 CA 对 Focus 模块进行改进,在 Focus 模块对图像进行间隔采样后添加注意力模块,将原始图像中的位置信息保留到扩充通道中,使 Focus 网络的特征保存能力进一步增强,如图 3 所示。

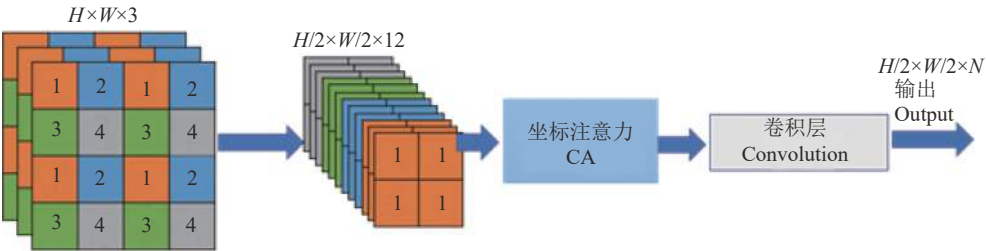
在神经网络模型中,深层的特征图层中包含更多的语义特征,位置信息已被高度抽象化为语义信息。YOLOX-Nano 在深层网络中使用 SPP 模块增大模型感受野,使用 PAFPN 模块实现特征融合,获取显著的上下文特征。由于深层网络通道主要保存抽象的语义信息,单个通道变成保存信息的独立体。因此,本研究在 SPP 模块和 PAFPN 模块使用



r 为缩放率; C 为特征图层通道数; W 、 H 为特征图层宽、高
 r is the reduction ratio; C is the number of feature map; W and H is the width and height of feature map, respectively

图 2 通道注意力机制 (SE) 与坐标注意力机制 (CA) 的计算过程

Fig. 2 Computational processes of squeeze-and-excitation attention(SE) and coordinate attention (CA)



W 、 H 为特征图层宽、高
 W and H is the width and height of feature map, respectively

图 3 改进后的 Focus 模块

Fig. 3 The Focus module after improved

SE 对拼接后的通道重新分配通道权重, 确定通道重要性。如图 4 和图 5 所示, 在 SPP 和 PAFPN 模块中加入 SE 的位置, 其中, 在 SPP 模块添加 1 处 SE, PAFPN 模块则加入了 4 处 SE。

1.3 多元化预训练数据集构建

Yosinski 等^[27] 通过试验证明深度学习模型能够学习物体的通用特征, 相比于参数随机初始化, 预训练模型参数更加合理, 在训练中能加快模型的收敛, 提高模型性能; 另一方面, 在数据集较少的情况下, 使用预训练模型对提升精度也有一定帮助。

本文的研究对象为柑橘新梢, VOC 和 COCO 数据集中的样本主要是生活场景中的物体, 这与柑橘果园场景有明显的区别, 预训练提取特征存在差异。因此, 在训练过程增加了在实验室研究基础上建立的果园场景图像数据集作为预训练数据集。自

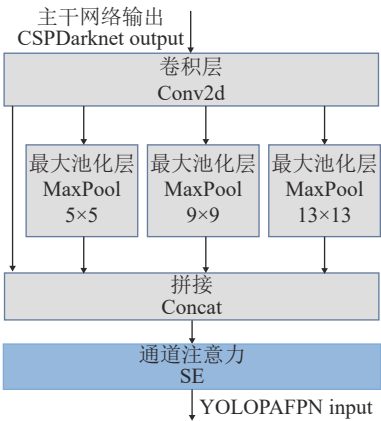


图 4 改进后的 SPP 模块

Fig. 4 The SPP module after improved

建的果园数据集图像背景与柑橘新梢十分相似, 其中, 柑橘青果、荔枝新梢与柑橘新梢具有相似特征。因此, 使用其作为预训练数据集模型进行参数

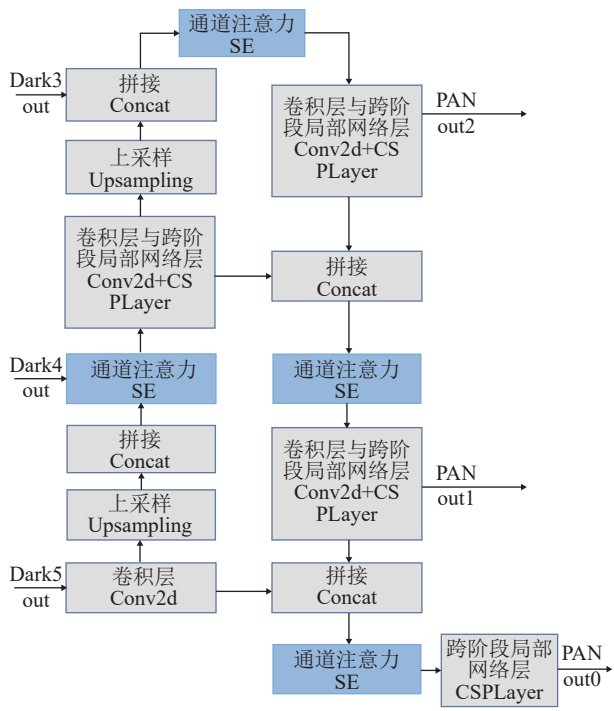


图 5 改进后 PAFPN 模块

Fig. 5 The PAFPN module after improved

初始化有利于学习目标与背景间的特征差异,使模型在特征提取过程中更明显地区分目标与背景。如图 6 所示,数据集中主要包括桃、苹果、梨、柑橘(青果)、柑橘(黄果)、荔枝新梢等目标,对应的数据集样本分别为 332、46、333、1 584、128 和 328 张。

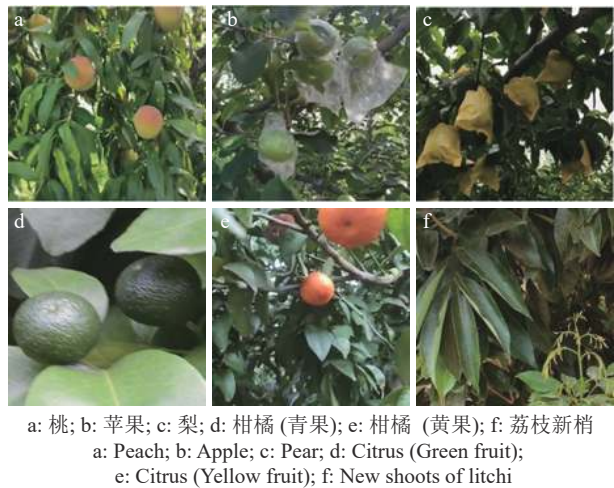


图 6 预训练数据集的果园目标

Fig. 6 Orchard goals for pretraining datasets

2 试验验证与结果

2.1 建模环境设置

本研究使用 Pytorch1.7 深度学习框架,显卡为 RTX3060 显存 12 G,底层使用 CUDA11.0 作为并行计算框架。在训练策略方面, Batch Size 为 16,除预训练数据集对比试验外,均使用 COCO 预训练数

据集进行迁移学习训练,对训练数据集迭代训练共 80 次,初始学习率为 1×10^{-3} ,学习衰减率均为 0.92。为验证模型改进的可行性以及对柑橘新梢检测和梢期识别的有效性,本研究中设计了多个对比方案:1) 当前改进模型与 YOLOV4-lite 系列以及原版 YOLOX-Nano 在模型检测精度、模型参数量、模型计算量、推理时间的差异;2) 不同注意力机制改进方案对柑橘新梢检测的精度、模型参数量与计算量的影响;3) 使用果园数据集与使用 VOC、COCO 数据集以及其他随机参数初始化方法在检测精度上的差异。

2.2 模型评价与试验对比

针对本研究中的模型目标检测与分类任务,本文选用常用的平均精度 (Average precision, AP)、各类别 AP 的平均值 (Mean average precision, mAP) 和漏检率 (Miss rate, MR) 对数值作为模型性能评价指标。同时考虑到本研究内容为多类别的识别与检测,为综合判断模型性能,针对 MR^{-2} 评价指标进行多类别求平均,获得 mMR^{-2} 。各评价指标计算如下所示:

$$P = \frac{TP}{TP + FP}, \quad (1)$$

$$R = \frac{TP}{TP + FN}, \quad (2)$$

$$AP = \int_0^1 P(R) dR, \quad (3)$$

$$mAP = \frac{1}{C} \sum_0^C AP, \quad (4)$$

式中, P 为准确率, R 为召回率, TP 为正确检测出的样本数量, FP 为误检的样本数量, FN 为漏检的样本数量, C 为类别数量。

MR^{-2} 用来量化 MR-FPPI 曲线, MR 、 $FPPI$ 为互斥指标,当模型检测阈值较低时,模型检测出的目标较多,漏检较少但误检增加;阈值升高时则误检降低,漏检增加,通过设置不同的检测阈值来获取 MR-FPPI 曲线。 MR^{-2} 计算方式较为复杂,以 FPPI 值作为横坐标,以 $\lg(MR)$ 值作为纵坐标,在 $[0.01, 1.00]$ 的范围内随机获取 9 个 FPPI 值,获取其对应的纵坐标值并求平均值。 MR^{-2} 越小表示模型性能越好,其计算过程如下所示:

$$MR = 1 - R, \quad (5)$$

$$FPPI = \frac{FP}{N}, \quad (6)$$

$$MR^{-2} = \frac{1}{9} \sum \delta(FPPI), \quad (7)$$

$$mMR^{-2} = \frac{1}{C} \sum_0^C MR^{-2}, \tag{8}$$

式中, MR 为漏检率, FPPI 为图像平均误检率, N 为图片数量, $\delta(FPPI)$ 表示 FPPI 与 $\lg(MR)$ 的映射关系, C 为类别数量。

2.3 不同注意力机制组合方案对比

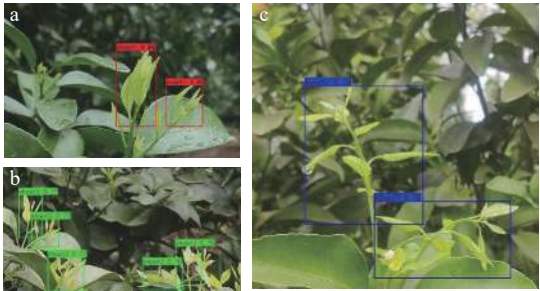
本文采用 YOLOX-Nano 作为基础对比网络, 在该网络上, 分别在 Focus、SPP 以及 PAFPN 模块使用不同的注意力机制组合进行改进, 训练了多组对比模型, 其中, Focus+CA、SPP+SE、PAFPN+SE

是本文提出的最终改进方案。模型识别结果(表 1 和图 7)表明, 在相同的训练条件下, 相比于不同的注意力组合方案, 本文提出的在 Focus 中添加 CA、在 SPP 和 PAFPN 中添加 SE 的改进方案与原始 YOLOX-Nano 相比, mAP 提高 4.39%、 mMR^{-2} 下降 0.086; 在萌芽期、生长期和展叶期均获得较好的识别效果, 而未引入注意力机制的 YOLOX-Nano 则表现出梢期识别效果的差异性, 因此, 本文提出的注意力机制引入方案同时改善了梢期阶段识别效果的平衡性。

表 1 不同改进方案检测结果对比

Table 1 Comparison of detection results using different improvement scheme								
方案 ¹⁾ Scheme	萌芽期		生长期		展叶期		mAP/%	mMR ⁻²
	Germination stage		Growth stage		Leaf-expansion stage			
	AP/%	MR ⁻²	AP/%	MR ⁻²	AP/%	MR ⁻²		
1	84.88	0.37	86.67	0.41	78.77	0.46	82.44	0.413
2	86.24	0.34	82.53	0.38	80.05	0.47	82.94	0.397
3	85.00	0.36	83.71	0.39	81.38	0.42	83.03	0.390
4	87.50	0.32	82.98	0.38	83.70	0.40	84.73	0.367
5	87.96	0.30	86.84	0.32	85.90	0.36	86.83	0.327

1) 1: YOLOX-Nano; 2: Focus+SE, SPP+SE; 3: Focus+CA, SPP+CA; 4: Focus+CA, SPP+SE; 5: Focus+CA, SPP+SE, PAFPN+SE

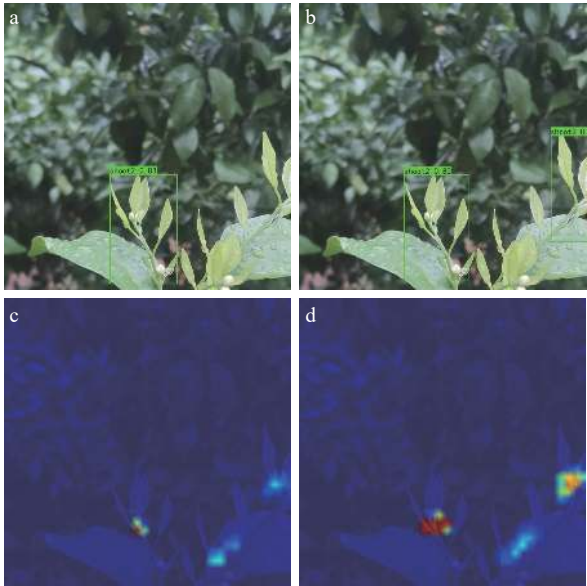


a: 萌芽期; b: 生长期; c: 展叶期
a: Germination stage; b: Growth stage; c: Leaf-expansion stage

图 7 不同生长期的检测结果

Fig. 7 Detection results at different growth stages

如图 8a、8b 所示, YOLOX-Nano 网络对边缘位置不完整的新梢存在漏检现象, 而改进后的 YOLOX-Nano 则完整地检测出图像中的新梢目标。为更好地解释模型的性能提升, 本文使用网络最终输出结果计算图像类激活图(Class activate map, CAM)。CAM 将模型的关注区域以热力图的形式可视化, 图中热力值越高表明关注程度越高。如图 8c、8d 所示, 注意力机制改进后的 YOLOX-Nano 实现了资源的重新分配, 提高了模型对特征的提取能力, 对边缘或者特征不明显的区域关注仍能有效提取特征, 使目标区域的关注度提高, 从而

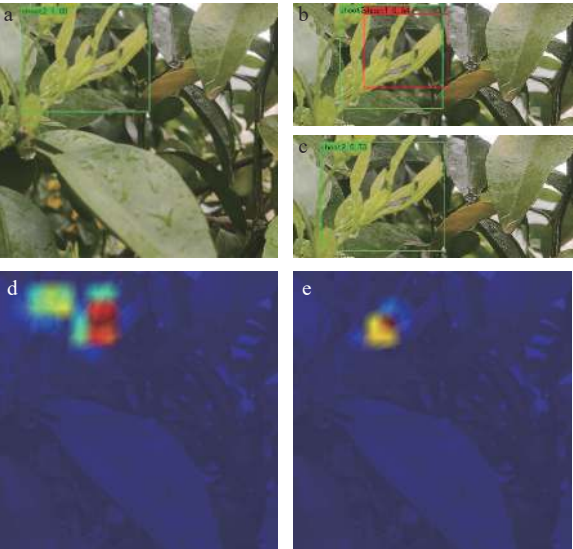


a: YOLOX-Nano 检测; b: 改进后 YOLOX-Nano 检测;
c: YOLOX-Nano 类激活图; d: 改进后 YOLOX-Nano 类激活图
a: YOLOX-Nano detection; b: Improved YOLOX-Nano detection;
c: CAM of YOLOX-Nano; d: CAM of improved YOLOX-Nano

图 8 YOLOX-Nano 和改进后 YOLOX-Nano 目标漏检及类激活图

Fig. 8 Target omission and class activate map (CAM)of Yolox-Nano and improved Yolox-Nano

避免了目标的漏检。此外, 如图 9 所示, 改进后的 YOLOX-Nano 模型关注点集中于主要目标区域, 避



a: 原始标注; b: YOLOX-Nano 检测; c: 改进后 YOLOX-Nano 检测;
d: YOLOX-Nano 类激活图; e: 改进后 YOLOX-Nano 类激活图
a: Original annotation; b: YOLOX-Nano detection; c: Improved
YOLOX-Nano detection; d: CAM of YOLOX-Nano;
e: CAM of improved YOLOX-Nano

图 9 YOLOX-Nano 和改进后 YOLOX-Nano 局部重复检测
检测结果及类激活图

Fig. 9 Local repeat detection and class activate map (CAM)
of Yolox-Nano and improved Yolox-Nano

免了因特征图中类激活点分散造成的单个目标局部重复检测的问题。

2.4 不同网络模型的综合性能对比

为验证改进后的 YOLOX-Nano 网络模型的综合性能,本研究建立与 YOLOV4 轻量化版本的对比试验,使用 Torchsummary、Torchstat 工具对网络模型进行参数量和计算量分析。表 2 数据表明,相同训练条件下针对本研究数据集,使用注意力机制改进后的 YOLOX-Nano 模型参数量仅为 YOLOV4- MobileNetV3 的 9.82%,但模型的 mAP 提升 6.58%,mMR⁻² 下降 0.033。在仅比原版 YOLOX-Nano 增加 0.01G 计算量的情况下,模型

表 2 不同网络模型性能对比

Table 2 Performance comparison of different backbone
network

网络模型 ¹⁾ Network model	参数量/MB Parameters	t/s	计算量/G FLOPs	mAP/ %	mMR ⁻²
1	43.16	0.042	7.56	80.25	0.360
2	42.01	0.052	6.98	80.83	0.373
3	4.23	0.036	1.11	82.44	0.413
4	4.24	0.037	1.12	86.83	0.327

1) 1: YOLOV4-MobileNetV3; 2: YOLOV4-GhostNet;
3: YOLOX-Nano; 4: Improved YOLOX-Nano

mAP 提升 4.39%、mMR⁻² 下降 0.086,模型综合性能更优。对于农业应用场景,更小的参数量和更少的计算量意味着更少的部署成本,因此改进后的模型对于农业应用具有明显的优势。

2.5 不同预训练数据集效果对比

为使模型初始化参数更具有合理性,本文在改进模型的基础上使用果园数据集以及 VOC、COCO 数据集进行预训练,选取最优模型作为基准模型训练柑橘新梢数据,训练超参数同“2.1”,训练 Loss 曲线如图 10 所示。Loss 曲线中 3 种不同预训练模型均具有明显的收敛趋势,但使用果园数据集预训练模型在收敛过程中更加平滑,波动较小,训练集与验证集收敛较为同步、合理。在模型识别精度方面如表 3 数据所示,使用果园数据集作为预训练数据集后,训练模型的识别精度有明显的提升,mAP 比 VOC 和 COCO 数据集分别提高了 2.76% 和 1.24%,mMR⁻² 分别下降了 0.036 和 0.020。

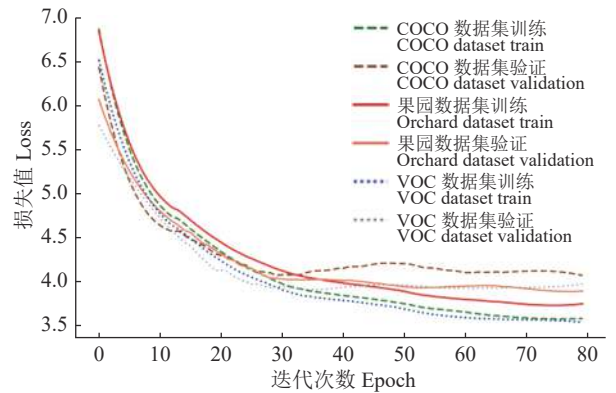


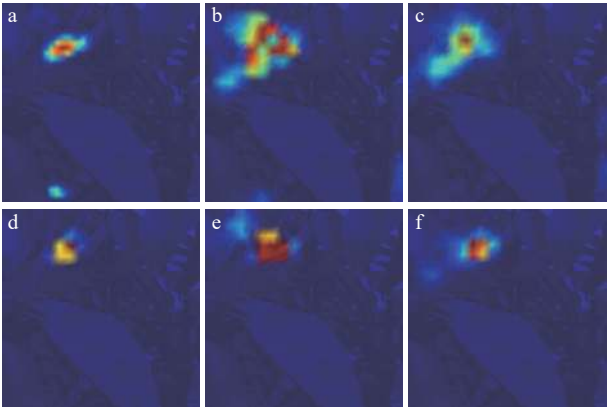
图 10 不同预训练数据集 Loss 曲线
Fig. 10 The loss curves of different pre-training datasets

为揭示不同参数初始化方式对模型性能影响的内部原因,使用“2.3”中的 CAM 可视化技术,对不同模型的最终输出进行可视化。如图 11 所示,使用正态分布初始化和 Kaiming 初始化在关注目标特征的同时,存在对图像中的其他与目标相似区域关注的现象,而正交初始化对目标关注较弱,不利于其他复杂情况下的目标检测。比较 3 种预训练数据集发现,COCO 和果园数据集预训练参数关注点更集中在新梢的主要部分,但是相比于果园数据集预训练参数,COCO 对于新梢的关注较弱;VOC 和果园数据集预训练参数均对新梢主要部分表现出较强的关注度,但 VOC 数据集预训练参数仅关注新梢外侧叶片,未将关注点集中在新梢主要区域;相比于随机初始化,使用预训练数据集后的训练参数可以有效地抑制对非目标区域的关注,且关注点更集中、与目标的主

表 3 不同参数初始化方法效果对比

Table 3 Effect comparison of different parameter initialization method

参数初始化方法 Parameter initialization method	萌芽期		生长期		展叶期		mAP/%	mMR ⁻²
	Germination stage		Growth stage		Leaf-expansion stage			
	AP/%	MR ⁻²	AP/%	MR ⁻²	AP/%	MR ⁻²		
Kaiming 初始化 Kaiming_init	41.68	0.76	46.94	0.83	36.09	0.85	41.57	0.813
正交初始化 Orthogonal_init	44.69	0.79	45.12	0.80	37.52	0.83	42.44	0.807
正态分布初始化 Normal_init	80.06	0.45	72.70	0.57	71.17	0.60	74.64	0.540
VOC数据集 VOC dataset	87.36	0.32	85.37	0.32	83.19	0.39	85.31	0.343
COCO数据集 COCO dataset	87.96	0.30	86.84	0.32	85.90	0.36	86.83	0.327
果园数据集 Orchard dataset	87.42	0.28	87.19	0.31	89.58	0.33	88.07	0.307



a: 正态分布初始化; b: Kaiming 初始化; c: 正交初始化;
d: COCO 数据集; e: VOC 数据集; f: 果园数据集
a: Normal_init; b: Kaiming_init; c: Orthogonal_init;
d: COCO dataset; e: VOC dataset; f: Orchard dataset

图 11 不同参数初始化方法的类激活图

Fig. 11 Class activate maps of different parameter initialization methods

要部分关注值更高。

根据 CAM 热力图可视化数据,以模型对目标区域的关注强弱为首要评价依据、关注区域的集中程度为次要判断依据,对模型进行评价。可以得出参数初始化 (mAP、mMR⁻²) 的优先顺序为: 果园数据集 (88.07%、0.307)、COCO 数据集 (86.83%、0.327)、VOC 数据集 (85.31%、0.343)、正态分布初始化 (74.64%、0.540)、Kaiming 初始化 (41.57%、

0.807)、正交初始化 (42.44%、0.813),与模型的识别性能基本一致。表明预训练模型的主要作用是在训练初期保证模型具有合适的初始参数,以便提取数据的准确特征、保证网络优化的正确性。试验结果证明,使用果园数据集训练的预训练模型参数对于本研究中的新梢数据集的特征提取更为合理,最终实现训练模型最优。

由于果园数据集中包含有柑橘、荔枝、苹果等多种作物,且柑橘果实与新梢同属柑橘果园场景。因此,为验证数据集中不同作物是否有利于提升模型性能,在使用全部果园数据集预训练以外,仅使用果园数据集中的柑橘果园图像部分进行预训练,对比两者对模型的性能影响。结果 (表 4) 表明,仅使用柑橘果园图像,对模型性能提升具有正向作用,但并未达到使用全部果园数据集的效果, mAP 为 87.14%,相比于使用全部果园数据集下降了 0.93%。这是因为预训练数据集中图像数量的减少造成了特征多样性降低。深度学习中,数据集的大小是影响模型性能的一个主要因素,庞大的数据集包含特征更为全面,且更有利于学习和巩固当前场景下的目标与背景特征。本研究结果表明,柑橘果园场景对模型性能提升的作用较大;增加数据集的数量,有利于丰富当前场景下果园数据特征,巩固预训练过程中模型学习到的特征。

表 4 果园内部数据集效果分析

Table 4 The effect analysis of orchard dataset

预训练数据集 Pretrain dataset	萌芽期		生长期		展叶期		mAP/%	mMR ⁻²
	Germination stage		Growth stage		Leaf-expansion stage			
	AP/%	MR ⁻²	AP/%	MR ⁻²	AP/%	MR ⁻²		
果园数据集 Orchard dataset	87.42	0.28	87.19	0.31	89.58	0.33	88.07	0.307
柑橘果园数据集 Citrus orchard dataset	89.12	0.31	85.41	0.33	86.89	0.38	87.14	0.340
柑橘果园与VOC混合数据集 Citrus orchards and VOC mixed dataset	88.36	0.33	85.98	0.33	86.61	0.34	86.98	0.330

3 结论

本文基于智慧果园对柑橘长势智能监测的应用背景,研究了柑橘新梢检测和梢期识别的轻量级神经网络,并部署于果园边缘端和终端设备实现实时感知,从而指导果园进行精准化作业决策。结论如下: 1) 采用 YOLOX-Nano 作为基础网络,使用多种注意力机制改进网络的 Focus 模块、SPP 模块以及 PAFPN 模块,能更有效地提取柑橘新梢的特征信息,在保证模型具有较少参数的基础上,提高了模型的识别分类精度,改进的模型对新梢检测的 mAP 比原模型提升了 4.39%; 2) 通过构建多元果园预训练数据集,对模型进行迁移学习,模型更精确地区分目标与背景, mAP 比 VOC、COCO 数据集模型分别提高了 2.76% 和 1.24%; 3) 通过多个模型对比,在相同训练条件下,与 YOLOV4_lite 系列网络模型相比,本文提出的基于注意力机制的 YOLOX-Nano 神经网络的改进模型在模型的参数量、计算量以及识别精度方面更具有优势,改进后的模型参数仅为 4.24 MB,降低了对部署平台的计算能力要求,为智慧果园柑橘生长精准管控提供了良好的技术支持。

参考文献:

[1] 邓秀新,束怀瑞,郝玉金,等. 果树学科百年发展回顾[J]. 农学学报, 2018, 8(1): 24-34.

[2] 刘双喜,徐春保,张宏建,等. 果园基肥施肥装备研究现状与发展分析[J]. 农业机械学报, 2020, 51(S2): 99-108.

[3] 赵春江. 智慧农业的发展现状与未来展望[J]. 华南农业大学学报, 2021, 42(6): 1-7.

[4] DENG X, ZHU Z, YANG J, et al. Detection of *Citrus Huanglongbing* based on multi-input neural network model of UAV hyperspectral remote sensing[J]. Remote-Sensing, 2020, 12(17): 2678.

[5] 戴泽翰,郑正,黄莉舒,等. 基于深度卷积神经网络的柑橘黄龙病症状识别[J]. 华南农业大学学报, 2020, 41(4): 111-119.

[6] 陆健强,林佳翰,黄仲强,等. 基于 Mixup 算法和卷积神经网络的柑橘黄龙病果实识别研究[J]. 华南农业大学学报, 2021, 42(3): 94-101.

[7] 胡嘉沛,李震,黄河清,等. 采用改进 YOLOv4-Tiny 模型的柑橘木虱识别[J]. 农业工程学报, 2021, 37(17): 197-203.

[8] 王林惠,兰玉彬,刘志壮,等. 便携式柑橘虫害实时监测系统的研制与试验[J]. 农业工程学报, 2021, 37(9): 282-288.

[9] ZHANG J, LUO S, HOU C, et al. Detection of orchard citrus fruits using a monocular machine vision-based method for automatic fruit picking applications[J]. Computers and Electronics in Agriculture, 2018, 152: 64-73.

[10] 吕石磊,卢思华,李震,等. 基于改进 YOLOv3-LITE 轻

量级神经网络的柑橘识别方法[J]. 农业工程学报, 2019, 35(17): 205-214.

[11] 邓颖,吴华瑞,朱华吉. 基于实例分割的柑橘花朵识别及花量统计[J]. 农业工程学报, 2020, 36(7): 200-207.

[12] APOLO-APOLOZ O E, MARTINEZ-GUANTER G, EGEE P, et al. Deep learning techniques for estimation of the yield and size of citrus fruits using a UAV[J]. European Journal of Agronomy, 2020, 115: 126030.

[13] 金方伦,邓江涛,敖学希,等. 柑橘潜叶蛾发生与控制柑橘夏梢的相关性及防治技术研究[J]. 湖北农业科学, 2013, 52(23): 5767-5770.

[14] 刘慧,何利庭,龚碧涯,等. 柑橘木虱在湖南发生规律的初步研究[J]. 湖南农业科学, 2019(10): 49-52.

[15] 黄永敬,李娟,陈杰忠,等. 沙糖桔控夏梢保果技术[J]. 广东农业科学, 2011, 38(14): 36-38.

[16] LI Y, HE L, JIANG M, et al. In-field tea shoot detection and 3D localization using an RGB-D camera[J]. Computers and Electronics in Agriculture, 2021, 185: 106149.

[17] XU W, ZHAO L, LI J, et al. Detection and classification of tea buds based on deep learning[J]. Computers and Electronics in Agriculture, 2022, 192: 106547.

[18] 袁加红,张中正,朱德泉,等. 名优绿茶嫩芽识别与定位方法研究[J]. 安徽农业大学学报, 2016, 43(5): 676-681.

[19] FANG L, WU Y, LI Y, et al. Ginger seeding detection and shoot orientation discrimination using an improved YOLOv4-LITE network[J]. Agronomy, 2021, 11: 2328.

[20] SCARLETT L STEVE C JULIE T, et al. A computer vision system for early stage grape yield estimation based on shoot detection[J]. Computers and Electronics in Agriculture, 2017, 137: 88-101.

[21] NIU Z, ZHONG G, YU H. A review on the attention mechanism of deep learning[J]. Neuro-computing, 2021, 452: 48-62.

[22] 李文涛,张岩,莫锦秋,等. 基于改进 YOLOv3-tiny 的田间行人与农机障碍物检测[J]. 农业机械学报, 2020, 51(S1): 1-8.

[23] YING B, XU Y, ZHANG S, et al. Weed detection in images of carrot fields based on improved YOLOv4[J]. Traitement du Signal, 2021, 38(2): 341-348.

[24] 杨蜀秦,刘杨启航,王振,等. 基于融合坐标信息的改进 YOLOv4 模型识别奶牛面部[J]. 农业工程学报, 2021, 37(15): 129-135.

[25] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[EB/OL]. [2017-09-05]. <https://arxiv.org/pdf/1709.01507.pdf>.

[26] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[EB/OL]. [2021-03-04]. <https://arxiv.org/pdf/2103.02907.pdf>.

[27] YOSINSKI J, CLUNE J, BENGIO Y, et al. How transferable are features in deep neural networks?[J]. Advances in Neural Information Processing Systems, 2014, 27: 3320-3328.