

DOI:10.19651/j.cnki.emt.2417703

# 基于 RT-DETR 的复杂果园环境下青橘检测方法<sup>\*</sup>

秦建华<sup>1,2</sup> 陈振伦<sup>1,2</sup> 万保雄<sup>3</sup> 陆泰良<sup>3</sup> 雷军乐<sup>1,2</sup>

(1. 桂林理工大学广西高校先进制造与自动化技术重点实验室 桂林 541006; 2. 桂林理工大学机械与控制工程学院 桂林 541004; 3. 广西特色作物研究院广西桂北特色经济作物种质创新与利用重点实验室 桂林 541004)

**摘要:** 青橘智能收获依赖快速精准的检测技术。针对青橘尺寸多样、果园环境复杂及果实与背景相似度高导致的检测精度不足和漏检问题,本研究提出了一种轻量且高精度的青橘检测模型(RT-GCTR)。该模型采用大感受野小波卷积模块(WCLRF\_Block)增强多尺寸目标特征提取,结合多尺度多头自注意力机制(MSMHSA)构建多尺度融合模块(MSMH-AIFI),自适应聚合多尺度特征,并引入 SPDCov 与 CSP-OmniKernel 模块设计 SCOK-CCFF 特征金字塔,提升小目标检测精度。实验结果表明,RT-GCTR 在训练数据集 1 和测试数据集 2 上的 AP50 分别为 92.0% 和 92.2%,优于其他先进模型。与 RT-DETR-r18 相比,RT-GCTR 的参数量和浮点运算量分别减少 26.7% 和 25.4%,在 NVIDIA Jetson Orin NX 上检测速度达 10.3 fps。本研究在降低复杂度的同时,提升了精度和实时性,满足边缘设备需求。

**关键词:** 青橘; Transformer; 深度学习; 目标检测

**中图分类号:** TP242; S225; TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.2; 210.7

## Green citrus detection method in complex orchard environments based on RT-DETR

Qin Jianhua<sup>1,2</sup> Chen Zhenlun<sup>1,2</sup> Wan Baoxiong<sup>3</sup> Lu Tailiang<sup>3</sup> Lei Junle<sup>1,2</sup>

(1. Key Laboratory of Advanced Manufacturing and Automation Technology in Guangxi Universities, Guilin University of Technology, Guilin 541006, China; 2. College of Mechanical and Control Engineering, Guilin University of Technology, Guilin 541004, China; 3. Guangxi Key Laboratory of Germplasm Innovation and Utilization of Specialty Commercial Crops in North Guangxi, Guangxi Academy of Specialty Crops, Guilin 541004, China)

**Abstract:** The intelligent harvesting of green citrus relies on fast and accurate detection technology. To address the issues of insufficient detection accuracy and missed detection caused by the diverse sizes of green citrus, complex orchard environments, and high similarity between fruits and backgrounds, this study proposes a lightweight and high-precision green citrus detection model (RT-GCTR). The model employs a large receptive field wavelet convolution module (WCLRF\_Block) to enhance multi-scale target feature extraction. It integrates a multi-scale multi-head self-attention mechanism (MSMHSA) to construct a multi-scale fusion module (MSMH-AIFI) for adaptive feature aggregation. Additionally, it introduces SPDCov and CSP-OmniKernel modules to design the SCOK-CCFF feature pyramid, improving small target detection accuracy. Experimental results show that RT-GCTR achieves AP50 scores of 92.0% and 92.2% on training dataset 1 and test dataset 2, respectively, outperforming other advanced models. Compared to RT-DETR-r18, it reduces parameters and computations by 26.7% and 25.4%, respectively, and achieves a detection speed of 10.3 fps on the NVIDIA Jetson Orin NX. This study improves accuracy and real-time performance while reducing complexity, making it suitable for edge device applications.

**Keywords:** green citrus; Transformer; deep learning; object detection

## 0 引言

世界柑橘组织数据显示,从 2021 年夏季到 2022 年冬

季,全球柑橘产量达 1.58 亿吨,中国贡献了 4 460 万吨,占全球产量的 28%<sup>[1]</sup>。青橘因其良好的储存稳定性和丰富的营养成分,被广泛应用于生物医学、皮肤美白、饮料及精

收稿日期:2024-12-22

\* 基金项目:国家自然科学基金(52465026)、国家现代农业产业技术体系广西创新团队(nycytxgxcxtd-2023-13-02)、广西自然科学基金(2021GXNSFAA220091)项目资助

油提取等多个领域<sup>[2-3]</sup>。然而,青橘采摘具有严格的时效要求,需在短时间内采摘背景色相似、分布密集的果实。目前,青橘采摘仍主要依赖人工,存在劳动强度大、效率低、成本高等问题,且随着劳动力的减少,迫切需要具备自主控制能力的机器人提升采摘效率<sup>[4]</sup>。因此,开发快速、精准的青橘检测方法,适应复杂果园环境,成为推动自动化青橘采摘机器人发展的关键。

传统的青橘检测方法主要依靠机器视觉技术,通过手工提取的特征,如颜色、形状、纹理和光谱信息等进行识别。例如, Lin 等<sup>[5]</sup>提出了一种使用基于颜色和轮廓特征的支持向量机(support vector machine, SVM)识别柑橘和番茄果实的方法。Zhao 等<sup>[6]</sup>采用绝对变换差值总和(sum of absolute transformed differences, SATD)和支持向量机(SVM)进行绿色柑橘检测,达到了 83% 的准确率。尽管这些方法在特定条件下能够有效检测柑橘果实,但其手工特征的表达能力较弱,导致在复杂环境中的鲁棒性和准确性较低。

相较于传统方法,深度学习技术,尤其是卷积神经网络(convolutional neural network, CNN),通过多层次的特征提取和抽象学习,显著提升了物体检测的精度和速度,尤其在农业领域的应用逐渐增多<sup>[7-8]</sup>。Zheng 等<sup>[9]</sup>提出的 YOLO-BP 网络,通过裁剪(cross stage partial darknet53, CSPDarknet5)骨干并设计双向特征金字塔网络(Bi-PANet),在绿色柑橘检测中取得了 91.55% 的平均精度,但检测速度仅为 18 fps。宋中山等<sup>[10]</sup>将密集连接机制(densely connected convolutional networks, DenseNet)引入 YOLOv3,提升了交并比,但计算开销较大,不利于边缘设备的部署。Zhang 等<sup>[11]</sup>在 YOLOv8 集成了 Refined-EfficientNetV2 主干网络和(convolutional block attention module, CBAM)模块,提升了绿色柑橘特征的提取能力,但适应非结构化果园的复杂场景仍存在挑战。吕强等<sup>[12]</sup>针对 YOLOv5s,引入全局注意力机制(global attention mechanism, GAM),提升了对果实与背景的区别能力,但在极端光照条件下鲁棒性仍显不足。这类方法将目标定位与分类任务合并,特征提取时易忽略细节,特别是对小目标与背景相似的目标效果较差。此外, YOLO 过于依赖局部特征,受限于卷积神经网络较小的感受野,不利于捕获尺寸多样化的目标特征,难以在果实遮挡和光照变化等复杂场景中建模上下文信息。

近年来,基于 Transformer<sup>[13]</sup>的深度学习网络已成为物体检测领域的重要研究方向。Transformer 通过自注意力机制在捕捉长程依赖关系和建模复杂模式方面展现了显著优势<sup>[14]</sup>,尤其在农业领域,Transformer 被广泛应用于作物病害、杂草检测、果实识别等多个场景。例如, Márquez 等<sup>[15]</sup>提出的 Transformer 架构用于杂草的场地特定检测,虽提升了模型对环境变化的适应性,但在复杂背景和多样化杂草种类的场景中,仍存在准确性下降的问题。Ji 等<sup>[16]</sup>

利用可变形注意力机制与 Transformer 架构相结合,提高了青苹果检测的准确率,但在果实遮挡或目标尺寸变化较大的情况下,检测准确性仍受到影响。Li 等<sup>[17]</sup>提出的(real-time detection transformer, RT-DETR)模型有效应对了复杂农场环境中的鸡只检测,并展示了 Transformer 在多尺度检测中的潜力,但该模型引入多个模块(如空深变变压器模块和上下文变变压器),显著增加了计算复杂度,限制了其实时性和部署能力。

尽管 Transformer 技术在农业物体检测中展现出巨大潜力,但现有方法普遍面临计算复杂度高、环境适应性差、对尺寸多样化目标检测和复杂背景区分能力不足等局限性,这些问题在大规模和动态农业环境中尤为突出。因此,开发更加高效、适应复杂环境的 Transformer 检测模型,成为推动农业智能化发展的关键。

在青橘检测中,由于目标尺寸差异大、细小目标多且分布密集,同时果实与叶片、枝条颜色相似,再加上果园复杂环境的干扰,现有方法的鲁棒性和准确性受到严重影响。为了解决这些问题,本研究基于 RT-DETR<sup>[18]</sup>算法,提出了一种轻量的青橘检测网络(real-time green citrus detection transformer, RT-GCTR),旨在保证高精度的同时,减少计算量,提高模型的适应性和实时性。与现有方法相比,本研究的方法在处理多尺度、小目标及背景干扰方面表现出更优的性能,为自动化青橘采摘机器人提供了更加精准的视觉技术支持。本研究的主要贡献如下:

1) 设计(wavelet convolutional large receptive field block, WCLRFB)模块,结合大感受野和小波卷积,通过较小的计算开销有效扩展了感受野。该模块利用小波变换分解图像特征,增强了对青橘图像细节和全局信息的捕捉能力,提升了对不同尺寸目标的特征提取能力。

2) 设计(multi-scale multi-head adaptive feature integration, MSMH-AIFI)多尺度融合模块,引入多尺度自注意力机制,进一步强化了小目标的检测能力。通过融合不同尺度的特征,MSMH-AIFI 模块不仅提升了对细微特征的捕捉能力,还加强了目标与背景的区别,使得模型在复杂环境下的鲁棒性得到增强,并减少了噪声干扰对目标识别的影响。

3) 设计(spatial-channel optimized knowledge and cross-level feature fusion, SCOK-CCFF)特征金字塔结构,优化了浅层与深层特征的融合效率。特征金字塔结构在低计算成本下有效融合了多尺度信息,提升了小目标检测精度,并增强了目标与背景之间的对比度,从而使青橘在复杂环境中的识别更加精准。

4) 在青橘训练数据集 1 和测试数据集 2 上的实验结果表明,RT-GCTR 模型显著提升了实时检测性能,并有效解决了青橘检测中的误检、错位和漏检问题。该模型不仅提高了检测精度,还展示了出色的泛化能力,验证了其在实际应用中的有效性和鲁棒性。

# 1 RT-GCTR 网络架构

作为首个实时端到端目标检测器,RT-DETR 由骨干网络、混合编码器、Transformer 解码器和辅助预测头组成。骨干网络通过 CNN 提取多尺度特征,为后续处理提供低级、中级和高级特征表示;混合编码器由基于注意力的(adaptive feature integration, AIFI)模块和基于 CNN 的 CCFM 模块构成, AIFI 模块增强高级特征的层次性和丰富性,(cross-level cross-feature fusion module, CCFM)模块促进不同级别特征的融合与交互;处理后的特征通过 Transformer 解码器生成目标预测,辅助预测头进一步优化检测性能。RT-DETR 通过 IOU 感知查询改进目标查询初始化,并可调整解码器层来优化实时性。

为应对青橘检测中的复杂环境和模型规模限制,本研究在 (real-time detection transformer with resnet-18 backbone, RT-DETR-r18)基础上进行了改进,提出了轻量化增强模型 RT-GCTR。该模型的结构设计如图 1 所示。首先, WCLRf\_Block 模块通过引入不同感受野的特征提取层,有效增强了模型在青橘检测中的多尺度特征提取能力。在青橘与背景相似物体(如叶片、枝条)的复杂场景中, WCLRf\_Block 通过扩大感受野,提升了对青橘目标的识别能力。这不仅为后续模块提供了更加丰富和精确的特征信息,还有效减少了背景干扰,增强了目标与背景之间的区分度。在此基础上, MSMH-AIFI 模块利用多尺度多头自

注意力机制进一步强化了 WCLRf\_Block 模块提取的特征表达。MSMH-AIFI 的多尺度特性确保了模型能够在青橘目标尺寸差异较大的情况下,捕捉到不同尺度目标的特征,并通过增强上下文信息的学习,提高目标与背景的区分能力。这样,MSMH-AIFI 为模型提供了对复杂背景环境中细节信息的增强表示,使得模型能够更好地适应不同尺度、不同背景和不同光照条件下的青橘目标。最后, SCOK-CCFF 特征金字塔结构通过高效融合浅层与深层特征,优化了 WCLRf\_Block 提取的多尺度信息和 MSMH-AIFI 聚合的细节特征。SCOK-CCFF 不仅提高了多尺度信息的利用效率,还使得模型在多尺度目标检测中表现更加稳健,尤其对于小目标的检测精度有显著提升。在青橘目标与背景相似物体(如枝条、叶片)对比度较低的复杂环境下, SCOK-CCFF 能够有效增强目标与背景之间的区分度,从而使得模型能够在复杂背景下精确检测青橘目标,提高检测的鲁棒性。

综上所述, WCLRf\_Block 架构提供了增强的多尺度特征, MSMH-AIFI 模块通过多尺度自注意力机制强化了特征表达,并且 SCOK-CCFF 特征金字塔结构则通过优化特征的融合效率进一步提升了模型的鲁棒性和对小目标的识别能力。这些模块的协同作用确保了 RT-GCTR 模型能够在复杂环境下更加准确地识别和定位青橘,特别是在处理小目标和背景干扰方面,提升了模型的整体检测精度与鲁棒性。以下各节将详细介绍增强模型的具体设计细节。

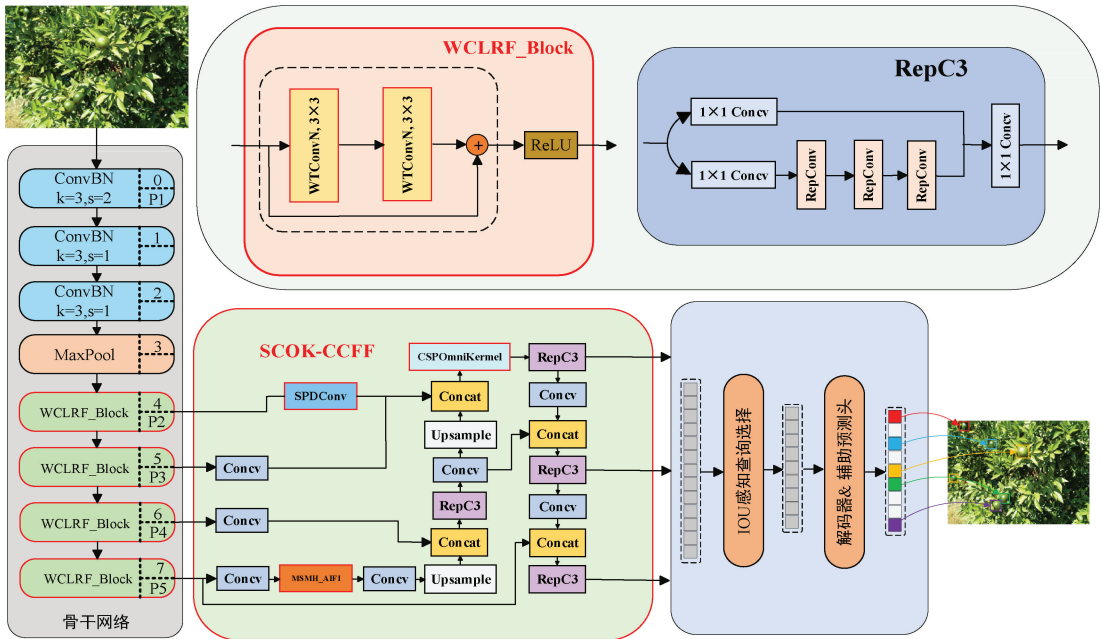


图 1 RT-GCTR 模型网络架构

Fig. 1 Network structure of RT-GCTR model

## 1.1 设计 WCLRf\_Block 模块

RT-DETR 的主干网络基于 (residual network, ResNet)架构,采用四层 (basic residual block, BasicBlock)

结构以平衡计算效率与性能。然而,在青橘检测任务中, BasicBlock 存在局限性:较小的感受野限制了全局上下文信息的捕捉,难以同时处理大目标的全局信息和小目标的

局部细节;缺乏多尺度特征融合,导致在复杂背景下检测精度下降;此外,较高的参数量增加了计算负担,影响模型部署效率。

为解决这些问题,本研究提出 WCLRF\_Block 模块,借鉴(wavelet transform convolution, WTConv)设计思想<sup>[19]</sup>。首先,对输入图像进行多级小波分解,低频子带捕捉全局特征,高频子带捕捉局部细节信息。然后,在每个频率子带上应用小卷积核进行卷积操作,有效扩展感受野,提升对不同尺度青橘目标的适应能力。最后,逆小波变换(inverse wavelet transform, IWT)将各子带的卷积结果重构融合,保留多尺度特征,从而提高检测精度,同时通过低分辨率子带和小卷积核减少计算开销和参数量。WTConv 具体过程由图 2 给出。

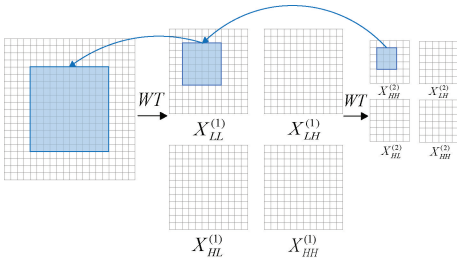


图 2 WTConv 的工作原理  
Fig. 2 WTConv's working principle

首先,输入图像  $X_{LL}^{(i-1)}$  通过小波变换(WT)进行分解为低频分量  $X_{LL}^{(i)}$  和 高频分量  $X_H^{(i)}$ , 过程如下:

$$X_{LL}^{(i)}, X_H^{(i)} = WT(X_{LL}^{(i-1)}) \quad (1)$$

其中,低频分量  $X_{LL}$  表示表示图像中的全局信息(整

体结构),高频分量  $X_H$  表示表示图像中的细节信息(边缘和纹理)。

接着,对分解后的低频分量  $X_{LL}^{(i)}$  和 高频分量  $X_H^{(i)}$  使用小尺寸卷积核  $W_i$  分别进行独立卷积操作,过程如式(2)所示。

$$Y_{LL}^{(i)}, Y_H^{(i)} = Conv(W_i, (X_{LL}^{(i)}, X_H^{(i)})) \quad (2)$$

最后,将不同频率子带上的卷积结果通过逆小波变换(IWT)重建输出。IWT 是线性操作,即  $IWT(X + Y) = IWT(X) + IWT(Y)$ , 低频卷积结果与下一层累积结果相加,再与高频卷积结果一起通过 IWT 重建当前层的输出。操作过程如下:

$$Z^{(i)} = IWT(Y_{LL}^{(i)} + Z^{(i+1)}, Y_H^{(i)}) \quad (3)$$

### 1.2 设计 MSMH-AIFI 模块

AIFI 模块作为 RT-DETR 的核心组件,通过自适应注意力机制加权融合多尺度特征,提升目标检测的精度和稳健性。然而,在应对绿色柑橘的复杂环境(如光照变化、天气影响、遮挡)和多尺度目标时,AIFI 模块中的多头自注意力面临挑战:光照和天气变化导致的视觉特征差异使得多头自注意力难以适应,尤其是目标与背景颜色相似时,可能导致信息丢失。此外,传统自注意力过于依赖局部特征,难以处理遮挡目标,影响目标完整性。在多尺度目标处理中,传统机制难以动态调整不同尺度目标的权重,小尺度目标易被大尺度目标淹没,且尺度间信息交互不足,降低检测精度。为此,本研究在 AIFI 模块中引入多尺度多头自注意力机制(multi-scale multi-head self-attention, MSMHSA)<sup>[20]</sup>,设计了 MSMH-AIFI 模块,其结构如图 3 所示。MSMH-AIFI 可有效捕捉尺度间的交互信息,进而提升其对多尺度目标的处理能力和复杂环境的适应性。

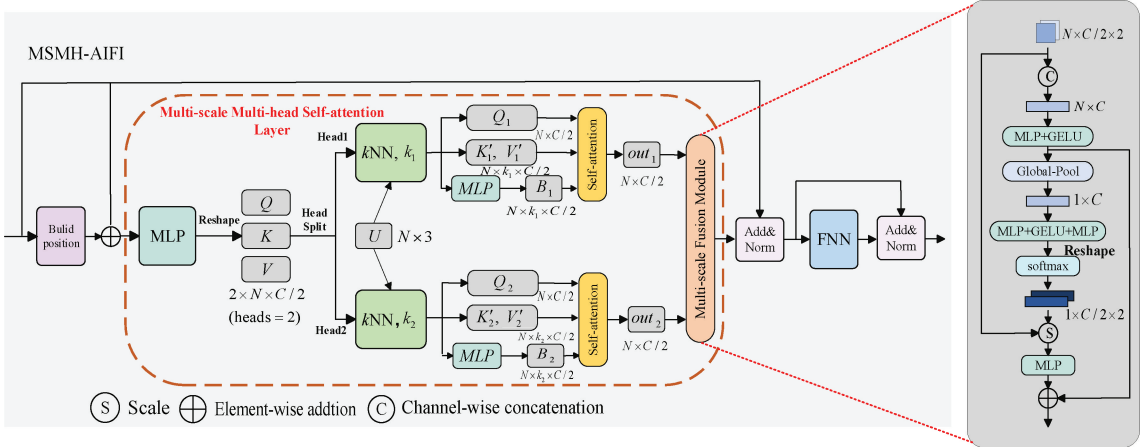


图 3 MSMH-AIFI 网络架构  
Fig. 3 Network structure of MSMH-AIFI

MSMHSA 的核心是通过在多个尺度上应用多头自注意力机制,动态捕捉不同尺度特征之间的关系,进一步加强尺度间信息交互。首先,输入数据(青橘图像)经过不同尺度的特征提取。每个尺度的特征表示通过卷积或池化操作得到,生成多个尺度的特征图。

$$X_s = ConvPool(X, k_s), s = 1, 2, \dots, S \quad (4)$$

其中,  $X_s \in \mathbb{R}^{N \times D_s}$  是第  $s$  种尺度的特征表示,  $N$  是数据的长度,  $D_s$  是第  $s$  种尺度的特征维度,  $k_s$  是不同尺度的卷积核大小。

其次,对于每个  $s$  尺度使用多头自注意力机制来计算

注意力得分并加权求和。该过程如下:

1) 生成查询  $Q$ 、键  $K$  和值  $V$  向量。

$$Q_s = X_s W_Q, K_s = X_s W_K, V_s = X_s W_V \quad (5)$$

其中,  $W_Q, W_K, W_V$  是线性变换矩阵, 分别将输入特征映射到查询、键和值的空间。

2) 计算注意力得分矩阵  $A_s^{(h)}$ 。

$$A_s^{(h)} = \text{softmax} \left( \frac{Q_s^{(h)} (K_s^{(h)})^T}{\sqrt{d_k}} \right) \quad (6)$$

其中,  $Q_s^{(h)}$  和  $K_s^{(h)}$  是第  $h$  个头的查询和查询和键向量,  $d_k$  是键向量的维度。

3) 加权求和得到每个头的输出  $O_s^{(h)}$  并将多个头的输出拼接和线性变换。

$$O_s^{(h)} = A_s^{(h)} V_s^{(h)} \quad (7)$$

$$O_s = \text{concat}(O_s^{(1)}, O_s^{(2)}, \dots, O_s^{(H)}) W_O \quad (8)$$

其中,  $H$  是头的数量,  $W_O$  是线性变换矩阵, 最终输出  $O_s$  是第  $s$  种尺度的多头自注意力输出。

最后, 融合来自不同尺度的多头自注意力输出。

$$O_{avg} = \sum_{s=1}^S \omega_s O_s \quad (9)$$

其中,  $\omega_s$  是每个尺度的权重。

### 1.3 设计 SCOK-CCFF 特征金字塔

RT-DETR 使用 CCFF 特征金字塔结构进行跨尺度特征融合, 先通过卷积层将输出特征通道映射至 256, 再通过自顶向下和自底向上的双向配对融合实现多尺度特征整合。然而, CCFF 在青橘小目标检测中存在明显不足。由于青橘尺寸较小, 传统金字塔结构依赖高层特征 (P3、P4、P5) 进行检测, 但下采样导致细节信息丢失, 无法充分表达小目标特征, 容易漏检或误检。此外, 青橘与背景颜色相似, 高层语义信息难以区分目标与背景, 光照变化和遮挡进一步削弱目标边缘与细节, 降低检测性能。虽然加入 P2 层可补充小目标信息, 但会显著增加算量和后处理复杂度, 难以满足实时检测要求。

为解决上述问题, 本研究提出了 SCOK-CCFF 特征金字塔结构, 旨在优化小目标特征表达、提高目标与背景的可区分能力、降低计算负担并提升多尺度特征融合效率。SCOK-CCFF 的结构如图 4 所示, 包含两个核心模块: 空间金字塔空洞卷积 (spatial pyramid dilated convolution, SPDCConv) 和跨阶段部分全核卷积 (cross stage partial omnkernel, CSP-OmniKernel)。

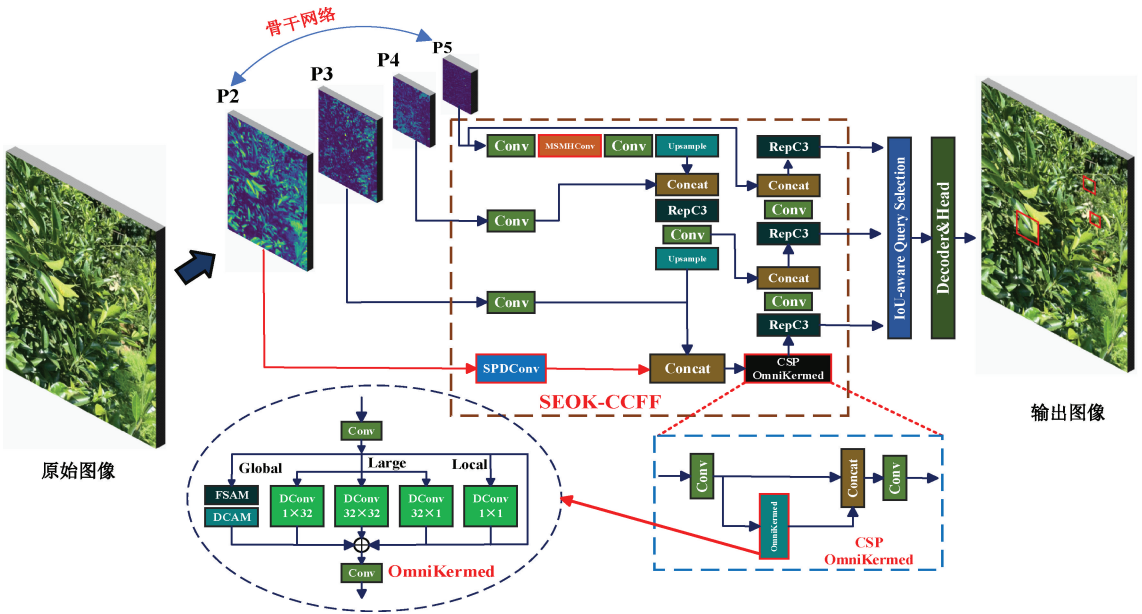


图 4 SCOK-CCFF 的网络结构

Fig. 4 Network structure of SCOK-CCFF

在 P2 特征层后引入 SPDCConv 模块<sup>[21]</sup>, 以解决小目标细节表达不足的问题。P2 层分辨率高, 包含丰富的小目标细节信息, 但直接用于检测会导致计算量激增。SPDCConv 通过轻量级卷积操作高效处理 P2 特征, 去噪并提取关键小目标信息, 随后上采样传递至 P3 层。在 P3 层, 优化后的 P2 特征与 P3 特征通过 Concat 融合, 有效弥补了 P3 层在小目标表达方面的不足, 同时避免了直接引入 P2 所带来的计算负担。针对 P3 与 P2 的融合, 传统拼接或卷积操

作无法充分建模不同尺度特征间的关系, 导致信息冗余和表达低效。因此, 融合后的特征上引入 CSP-OmniKernel 模块进行深度整合。CSP-OmniKernel 模块借鉴 CSP (cross stage partial) 设计思想, 将输入特征分为两部分: 一部分直接传递, 另一部分进入 OmniKernel 模块<sup>[22]</sup> 进行深度处理。OmniKernel 通过全局、大分支和局部分支的多层次设计优化特征: 全局分支利用 (feature spatial attention module, FSAM) 和 (dense channel attention

module,DCAM)捕获长程上下文信息,优化目标与背景的分;大分支通过  $32 \times 32$  卷积提取中大尺度结构,保持语义信息完整性;局部分支通过  $1 \times 1$  和  $32 \times 1$  卷积聚焦小目标边缘细节与形态特征,实现多尺度特征表达与交互。

SPDCConv 与 CSP-OmniKernel 模块的协同设计有效解决了青橘检测中的细节丢失、目标与背景区分困难和计算负担问题,显著提升了复杂场景下小目标检测的精度与实时性能。

## 2 实验与结果分析

### 2.1 实验平台和参数

所有模型训练和测试均在同一台计算机上进行,配置如下:Windows 11(64 位)、Intel Core i7-13620 H 处理器、NVIDIA GeForce RTX 4050 6 GB 显卡、Anaconda 3.5、CUDA 12.2、Pytorch 2.0.0、Python 3.8.19。在训练过程中,输入图像大小为  $640 \times 640$ ,训练轮次设为 200,批量大小为 4,初始学习率为 0.01,最终调整为 0.0001,采用 AdamW 优化算法以避免局部最优。为确保实验一致性,本研究未使用预训练权重。

### 2.2 实验数据集采集

本研究使用两个青橘图像数据集进行实验:数据集 1 用于模型训练和测试,数据集 2 用于模型泛化性验证。数据集 1 采集于中国广西钦州市浦北县的非结构化露天柑橘果园(北纬  $22^{\circ}2'7''$ ,东经  $109^{\circ}5'6''$ ),拍摄设备为 iPhone 13(256 GB)内置高分辨率数码相机( $f/1.8$ ,曝光时间  $1/26$  s,焦距 26 mm),拍摄距离约为  $40 \sim 150$  cm。通过多角度、全方位的单帧拍摄方式,充分模拟采摘机器人的作业视角,共采集了 1 000 幅青橘原始图像(分辨率为  $2\,532 \times 1\,170$ ,JPG 格式),涵盖各种天气(阴天、晴天、雨天)、光照(逆光、顺光)和拍摄角度(水平视图、俯视图、仰视图)。部分样本如图 5 所示。

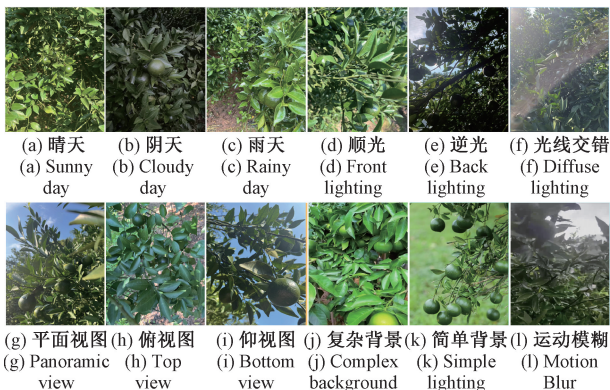


图 5 数据集 1 部分图像示例

Fig. 5 Sample images from dataset 1

数据集 2 由两部分组成,第 1 部分来自百度飞桨 AI Studio 的柑橘数据集,经过清洗后获得 100 幅青橘图像;第 2 部分通过 Google 浏览器爬取自不同地区和环境条件下

的青橘图像,共 300 幅。整个数据集共计 400 幅图像,涵盖了多样的柑橘种类、拍摄设备、气候条件和环境变化,旨在测试模型的鲁棒性和跨区域适应能力。图 6 展示了数据集 2 的部分样本图像。

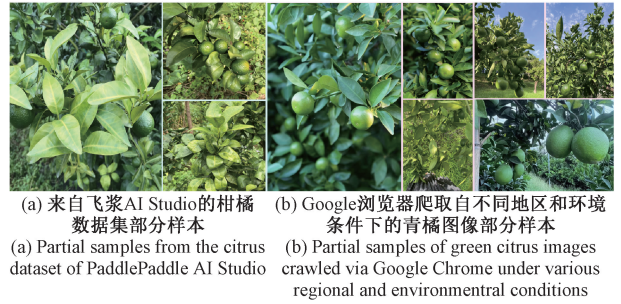


图 6 数据集 2 部分图像示例

Fig. 6 Sample images from dataset 2

### 2.3 数据集预处理与特点分析

使用 LabelImg 工具,按照 VOC 格式对数据集 1 和数据集 2 的青橘进行标注,生成 XML 文件,标签为“citrus”。标注遵循完全封闭目标于边界框内的原则,并通过两名研究人员的多轮交叉核验确保标注准确,避免主观性错误。

为增强算法鲁棒性和模型泛化能力,对数据集 1 和数据集 2 的青橘图像进行了数据增强,操作包括添加高斯噪声、平移、亮度调节、裁剪和旋转,以模拟作业环境中的噪声干扰、相机移动、光照变化、目标遮挡和相机旋转。通过这些方法随机组合应用,数据集 1 扩增至 3 000 幅,数据集 2 扩增至 1 200 幅。数据集 1 按 8:2 比例划分为训练集和验证集,样本分配如表 1 所示。

表 1 青橘数据集样本分布

Table 1 Sample distribution of the green mandarin dataset

数据集	图像数量	目标总数	训练集	训练集目标数量	验证集	验证集目标数量
1	3 000	24 750	2 400	20 803	600	3 947
2	1 200	9 235	—	—	—	—

为明确网络优化方向,本研究分析了青橘数据集的特性。如图 7 所示,青橘目标尺寸差异较大。根据 MS COCO 数据集定义,小于  $32 \text{ pixel} \times 32 \text{ pixel}$  的目标为小目标<sup>[23]</sup>,数据集 1 中包含大量微小目标,给精确的目标框回归带来困难,影响高精度定位。因此,优化应聚焦于提升对小目标和尺寸差异大的目标检测精度,尤其在复杂环境下。

### 2.4 评估指标

本研究采用精度(precision,P)、召回率(recall,R)、均值平均精度(mean average precision,mAP)作为模型检测青橘准确性的评价指标,公式如式(10)~(12)所示。

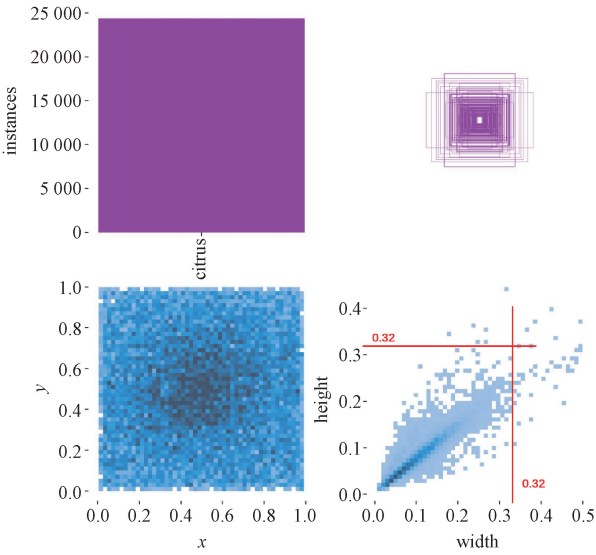


图7 数据集1特点

Fig. 7 Characteristics of dataset 1

$$P = \frac{TP}{TP + FP} \times 100\% \quad (10)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (11)$$

$$AP = \int_0^1 P(R) dR \quad (12)$$

其中, TP(真阳性)表示被准确识别为正例的青橘数量; FP(假阳性)表示错误预测为正例的青橘数量; FN(假

阴性)表示错误预测为负例的青橘数量; AP为 P、R和坐标轴围成的面积, 评估每个类别的预测精度, AP用于评估所有类别的平均精度。其中, AP50(average precision at IoU threshold 0.50)表示在IoU(交并比)阈值为0.50时的平均精度, AP50-95(average precision over IoU thresholds from 0.5 to 0.95)表示在IoU阈值从0.5~0.95范围内的平均精度。算量(GFLOPs)、参数量(Params)用于衡量网络模型复杂程度, 如式(13)~(14)。

$$GFLOPs = W \times H \times K \times K \times C_{in} \times C_{out} \quad (13)$$

$$Parameters = C_{in} \times C_{out} \times K \times K \quad (14)$$

式中:  $W$ 、 $H$ 分别表示特征图的宽和高;  $C_{in}$ 和 $C_{out}$ 分别为输入和输出特征图的通道数,  $K$ 为卷积核大小。

每秒帧数(frames per second, FPS)则评价模型检测速度, 公式如下:

$$FPS = \frac{1000}{t_{preprocess} + t_{inference} + t_{NMS}} \quad (15)$$

其中,  $t_{preprocess}$ 、 $t_{inference}$ 、 $t_{NMS}$ 分别表示模型进行预测时需要的预处理、推理和NMS时间。

## 2.5 消融试验

对原始RT-DETR-r18模型进行了3项改进: A:用WCLRF\_Block替代BasicBlock; B:用MSMH-AIFI替换原始AIFI; C:设计SCOK-CCFF特征金字塔进行多尺度特征融合。每个改进模块单独集成到原始模型中, 以验证其效果。相应结果如表2所示, 其中,  $\checkmark$ 表示使用对应增强模块,  $\times$ 表示未使用, 最佳指标以粗体标示。

表2 不同模型消融实验结果比较

Table 2 Comparison of ablation experiment results for different models

实验序号	A	B	C	P/%	R/%	AP50/%	FPS	Params/M	Flops/G
1	$\times$	$\times$	$\times$	92.9	80.7	89.5	70.9	20.2	58.6
2	$\checkmark$	$\times$	$\times$	93.3	84.1	91.6	78.8	<b>14.2</b>	<b>43.7</b>
3	$\times$	$\checkmark$	$\times$	92.7	82.3	90.6	73.7	20.1	58.5
4	$\times$	$\times$	$\checkmark$	92.8	<b>84.5</b>	91.9	68.2	20.8	66.9
5	$\checkmark$	$\checkmark$	$\checkmark$	<b>93.9</b>	84.3	<b>92.0</b>	<b>81.4</b>	14.8	52.2

从表2可知, 使用WCLRF\_Block替换BasicBlock后, 模型性能显著提升, P、R和AP50分别提高0.4%、3.7%和2.1%, 检测速度提升7.9 fps, 参数量和浮点数运算量分别减少29.7%和25.4%。这表明WCLRF\_Block的设计有效扩展了特征提取感受野, 增强了模型对多尺度目标特征的感知能力。在AIFI中引入MSMHSA后, R提高1.6%, AP50提高1.1%, 参数量和计算量略有下降, 多尺度特征信息的加入细化了特征表达, 有效抑制了复杂环境中的干扰。加入SCOK-CCFF后, 参数量和浮点运算量增加至20.8 M和66.9 G, 这可能是由于浅层高分辨特征信息的引入增加了模型的学习难度。但也实现了最大性能提升, R和AP50分别提高3.8%和2.4%, 验证了其对于细

小青橘目标检测的有效性。相较于RT-DETR-r18, RT-GCTR模型的P、R和AP50分别提高1.0%、3.6%和2.5%, 参数量和浮点运算量分别减少26.7%和25.4%, 检测速度达到了81.4 fps, 在精度、速度和计算开销三者之间实现了良好平衡。

为更直观评估改进模块的有效性, 采用Grad-CAM方法<sup>[24]</sup>生成热图分析模型关注区域。热图结果如图8所示, 图中的红色部分越深, 表示关注度越高。具体来说, 基准模型在检测青橘时, 焦点较为混乱, 部分区域出现错误检测。加入WCLRF\_Block后, 感受野扩展使模型更准确地关注重要特征, 但对细小目标的关注仍不足。结合MSMH-AIFI后, 多尺度特征的交互信息的加

入有效关注了小目标区域。进一步引入 SCOK-CCFF 后,模型在细小目标和复杂背景下的关注度显著提高,整体检测性能得到提升。这些结果验证了改进模型的有效性。

有效性。

为验证改进模型的有效性和可行性,图 9 可视化了不同模型在 AP50-95 和验证集损失 (Val Loss) 上的对比。

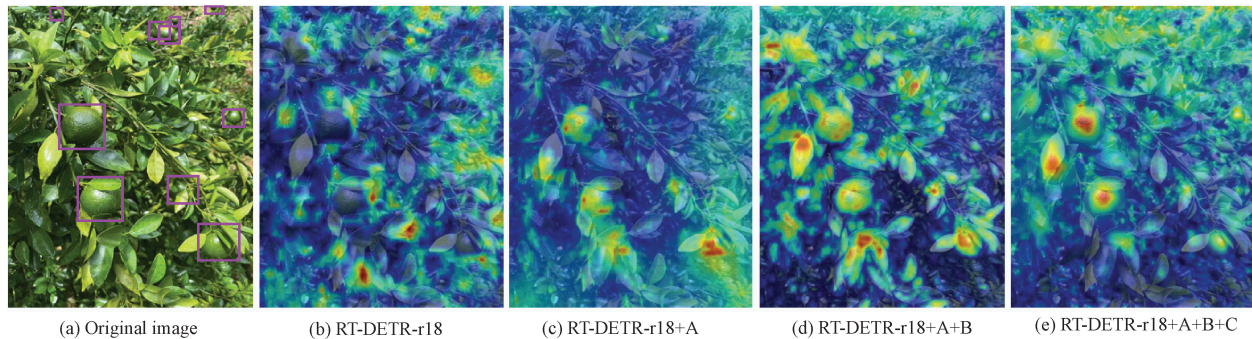
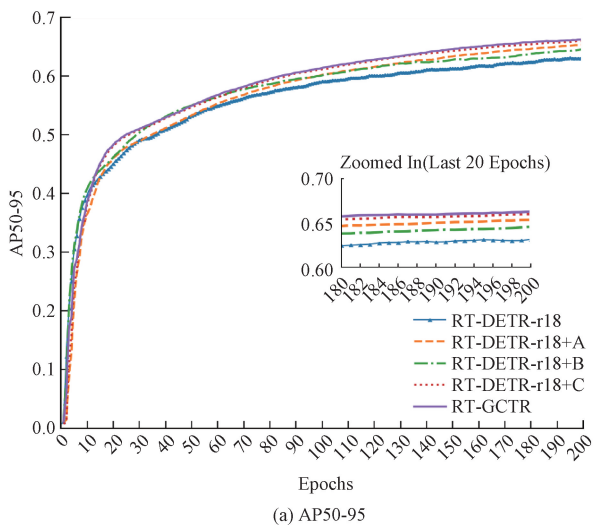
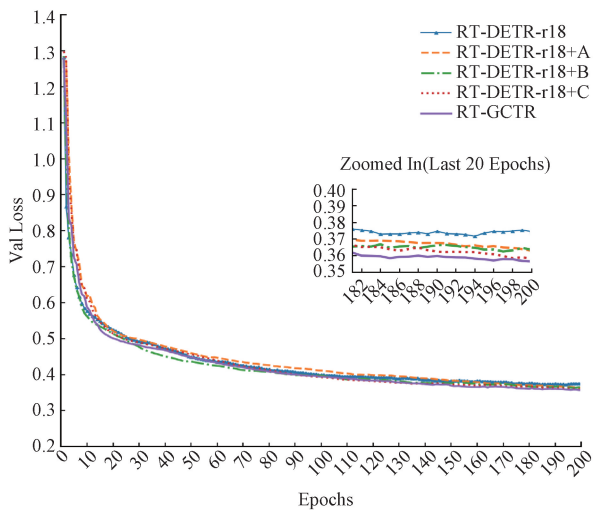


图 8 Grad-CAM 热图

Fig. 8 Grad-CAM heatmap



(a) AP50-95



(b) Val Loss

图 9 不同模型细节对比

Fig. 9 Detailed comparison of different models

如图 9(a) 所示,RT-DETR-r18 + B 和 RT-GCTR 的 AP50-95 值在初期快速增长并逐渐收敛,且始终高于 RT-DETR-r18,表明 SCOK-CCFF 特征金字塔在小目标检测上具有显著优势,适合高精度任务。图 9(b) 展示了验证集的损失曲线,RT-GCTR 以更低的损失值和更快的收敛速度稳定在 0.3564,证明其在青橘检测中的稳定性和泛化能力,能有效避免过拟合或欠拟合。

2.6 对比试验

为证明改进的 RT-GCTR 模型的优异检测性能,本研究选择了 9 种现阶段先进目标检测算法在数据集 1 上进行对比实验。这其中包括先进的轻量化网络 (YOLOv3-tiny、YOLOv5、YOLOv6、YOLOv8 和 YOLOv11) 和基于 Transformer 架构的网络模型 (RT-DETR-r18、RT-DETR-r18 + P2 (直接加入 P2 特征层的模型)、RT-DETR-r18 + SCOK-CCFF 和 RT-DETR-r34)。实验结果如表 3 所示,其中使用粗体表示最佳性能指标。

表 3 结果表明,RT-GCTR 的 P、R、AP50、参数量和计算量分别为 93.9%、84.3%、92.0%、14.8 M 和 52.2 GFlops。对比先进轻量级 YOLO 系列模型,RT-GCTR 的 AP50 比 YOLOv3-tiny、YOLOv5n、YOLOv6n、YOLOv8n、YOLOv10n、YOLOv11n 分别高出 8.3%、3.0%、9.1%、1.7% 和 3.9%。这是由于 YOLO 算法仅依靠 CNN 进行特征提取,对全局上下文信息的建模能力有限,而 RT-GCTR 引入了更为细化的浅层特征。相较于 RT-DETR-r18、RT-DETR-r18 + P2 (直接加入 P2 特征层的模型)、RT-DETR-r18 + SCOK-CCFF 和 RT-DETR-r34 等 Transformer 架构的模型,RT-GCTR 在 AP50、参数量、计算量和检测速度上表现最佳,AP50 提高了 0.1%~2.5%,参数量和计算量分别降低了 21.7%~52.9% 和 10.9%~42.4%,检测速度提升了 15.5%~58.7%。与 RT-DETR-r18 + P2 相比,RT-DETR-r18 + SCOK-CCFF 在降低

18.1%计算量的基础上 AP50 提升了 1.1%，这更进一步验证了 SCOK-CCFF 设计的有效性。综上，RT-GCTR 在综合性能上优于其他先进目标检测模型。

表 3 不同模型在数据集 1 的检测结果对比

Table 3 Comparison of detection results for different models on Dataset 1

模型	P/%	R/%	AP50/%	FPS	Params/M	Flops/G
YOLOv3-tiny	90.8	75.8	83.7	148.2	9.5	14.3
YOLOv5	92.5	80.3	89.0	145.3	7.81	18.7
YOLOv6	90.5	72.6	82.9	<b>153.6</b>	<b>4.2</b>	<b>11.5</b>
YOLOv8	93.9	81.9	90.3	139.5	9.83	23.3
YOLOv11	91.9	78.9	88.1	150.1	9.41	21.3
RT-DETR-r18	92.9	80.7	89.5	70.9	20.2	58.6
RT-DETR-r18+P2	91.4	79.6	88.8	52.4	18.9	81.7
RT-DETR-r18+SCOK-CCFF	92.8	<b>84.5</b>	91.9	68.2	20.8	66.9
RT-DETR-r34	92.2	84.0	91.3	51.3	31.4	90.6
RT-GCTR	<b>93.9</b>	84.3	<b>92.0</b>	81.4	14.8	52.2

图 10 和 11 可视化了不同检测模型在复杂环境中对青橘的检测结果,其中,漏检目标以淡蓝色三角形标记,误检目标以暗红色三角形标记。图 10 展示了一个包含不同尺寸柑橘且受光线干扰的复杂场景, YOLO 系列算法存在较多漏检,其中 YOLOv3-tiny、YOLOv5、YOLOv6、YOLOv8 和 YOLOv11 分别遗漏 3、4、5、4 和 5 个目标,这可能源于其特征提取层在低对比度特征提取上的局限性。

Transformer 架构检测算法对小目标特征更敏感,但难以区分与背景相似的果实,导致错检问题,RT-DETR-r34 误检目标最多,达 6 个。相比之下,RT-GCTR 在该场景下实现了完美检测。图 11 进一步分析了果实与叶片相似度高、枝叶遮挡显著的场景,结果显示,RT-GCTR 准确检测到所有柑橘,而其他模型均存在至少一个检测错误。综上,RT-GCTR 展现了卓越的检测性能和复杂环境适应能力。



图 10 包含不同尺寸目标且受光线干扰的复杂场景

Fig. 10 In scenarios with varying sizes of mandarins and light interference



图 11 果实与叶片相似、枝叶遮挡显著的场景

Fig. 11 In scenarios with high similarity between fruits and leaves and significant foliage occlusion

### 3 讨 论

#### 3.1 模型泛化性试验

为验证 RT-GCTR 的泛化能力和鲁棒性,本研究在青橘数据集 2 上对 RT-GCTR 与多种先进目标检测模型进行了推理验证,实验结果如表 4 所示。

表 4 不同模型在数据集 2 推理结果对比

Table 4 Comparison of inference results of different models on dataset 2

模型	P/%	R/%	AP50/%
YOLOv3-tiny	89.3	73.3	83.9
YOLOv5	91.8	78.3	88.5
YOLOv6	89.6	73.7	84.6
YOLOv8	91.1	80.1	89.6
YOLOv11	91.9	79.8	89.3
RT-DETR-r18	91.6	82.1	90.7
RT-DETR-r18+P2	92.4	80.7	90.0
RT-DETR-r18+SCOK-CCFF	92.8	83.0	91.2
RT-DETR-r34	93.4	83.3	91.4
RT-GCTR	<b>93.6</b>	<b>85.2</b>	<b>92.2</b>

从表 4 可以看出,RT-GCTR 在该数据集上的表现优于所有对比模型,其精度 ( $P = 93.6\%$ )、召回率 ( $R =$

$85.2\%$ )和 AP50 ( $92.2\%$ )均达到了最高值。相较之下, YOLO 系列和 RT-DETR 系列算法在召回率和 AP50 上有所逊色,这表明现有模型对跨区域环境的适应能力存在一定局限。尤其在复杂场景中,现有方法难以有效区分目标与背景,从而影响检测精度。

RT-GCTR 的优异性能得益于其在多尺度特征提取、目标与背景区分以及跨尺度信息融合方面的优化。通过引入 WCLRF\_Block、MSMH-AIFI 和 SCOK-CCFF 模块, RT-GCTR 能够有效减少误检和漏检率,并在复杂背景和多变环境条件下维持较高的检测精度。这些模块的联合优化使得 RT-GCTR 在训练集(数据集 1)和测试集(数据集 2)上的性能均表现优异,体现了其卓越的泛化能力和鲁棒性。

#### 3.2 模型部署

为评估 RT-GCTR 在智能系统设备边缘计算中的性能,表 5 展示了 RT-DETR-r18 和 RT-GCTR 在 NVIDIA Jetson Orin NX 上部署的实验结果。表中使用粗体表示最佳性能指标。图 12 可视化了部署过程和预测结果。具体

表 5 改进前后模型在 NX 上的实验结果

Table 5 Experimental results of the model on NX before and after improvement

模型	P/%	R/%	AP50/%	FPS
RT-DETR-r18	92.6	80.8	89.4	8.6
RT-GCTR	<b>93.4</b>	<b>84.1</b>	<b>91.9</b>	<b>10.3</b>



图 12 RT-GCTR 在 NX 上的部署过程和预测结果

Fig. 12 Deployment process and prediction results of RT-GCTR on NX

来说,相较于 RT-DETR-r18,RT-GCTR 在检测速度上提升了 19.8%,达到了 10.3 fps,满足实时性要求;同时,P、R 和 AP50 分别提升了 0.4%、3.6%和 2.4%,表明该模型在复杂果园环境中能够更准确地检测柑橘目标。因此,RT-GCTR 在 Jetson Orin NX 平台上具备更优异的检测性能和实时性,特别适合部署在计算能力有限且对实时性要求较高的边缘设备中。

### 3.3 未来工作

本研究中,RT-GCTR 展示了较强的检测性能,但在某些情境下仍存在漏检和误检问题,主要与目标尺度差异、背景干扰、光照变化、遮挡、数据不平衡和区域差异等因素有关。首先,由于青橘的尺度差异较大,尤其在近距离或远距离拍摄时,模型可能漏检小目标或远距离目标,而尺度较大的目标可能导致误检。其次,复杂背景(如枝条、叶片等物体与青橘颜色和纹理相似)会导致误检,这在非结构化果园或杂乱环境中尤为突出。光照变化(如强烈逆光或阴影)和部分遮挡(如树枝或叶片遮挡目标)也会影响模型识别,导致漏检现象。最后,尽管实验数据集涵盖了多样的环境变化,数据分布不均仍可能影响模型表现,特别是在某些区域或特定环境条件下,模型性能可能有所下降。

针对上述问题,未来的研究将着重于以下几个方面进行改进和优化:

1) 多尺度训练与增强:针对目标尺度变化,未来将引入更多多尺度数据和数据增强方法,通过生成不同尺度的目标和背景,提升模型在不同尺度目标下的表现。

2) 背景抑制与增强区分能力:为减少背景干扰,将进一步改进模型的背景抑制能力。例如,通过结合注意力机制,增强模型在复杂背景下对目标与背景的区分能力。

3) 光照与遮挡适应性增强:未来将加入光照变化与遮挡模拟的训练数据,采用合成数据或数据增强方法,提升模型在光照变化和遮挡情况下的鲁棒性。

4) 跨区域适应性研究:扩展数据集,增加来自不同地区、环境的图像数据,特别是不同种类的果园和气候条件的数据,以提升模型在不同区域的适应性和泛化能力。

## 4 结 论

本研究开发了一种轻型青柑检测网络 RT-GCTR,用于平衡检测精度与成本,为青柑橘智能采摘提供视觉支持。首先,设计了大感受野小波卷积模块(WCLRF\_Block)优化骨干网络,通过小波变换提升图像细节与全局信息捕捉能力,增强多尺度特征提取效果。其次,融合尺度多头自注意力机制(MSMHSA)与重构 AIFI 模块,有效捕捉尺度间交互信息,提升多尺度目标学习能力与复杂环境适应性。最后,开发了 SCOK-CCFF 特征金字塔,通过 SPD-Conv 去噪并提取小目标信息,结合 CSP-OmniKernel 深度整合特征,提出了 SCOK-CCFF 特征金字塔,显著提升小目标检测精度与背景区分能力。在数据集 1 的实验中,RT-GCTR 实现了 93.9%的准确率、81.4%的召回率、92.0%的 AP50 和 81.4 的 FPS,参数量为 14.8 M,浮点运算量为 52.2 G。在 NVIDIA Jetson Orin NX 上,AP50 与 FPS 分别为 91.7%与 10.3 帧/s,满足实时采摘需求。在数据集 2 上也实现了最佳检测性能。该研究为青橘数字化采摘提供了有效视觉参考,并可应用于其他目标检测任务。

### 参考文献

- [1] HOU CH J, ZHANG X D, TANG Y, et al. Detection and localization of citrus fruit based on improved you only look once v5s and binocular vision in the orchard[J]. *Frontiers in Plant Science*, 2022, 13: 972445.
- [2] OKU H, KITAGAWA F, KATO Y, et al. Anti-allergic effects of the subcritical water extract powder of Citrus unshiu in mouse and Guinea pig models[J]. *Journal of Medicinal Food*, 2021, 24(5): 533-540.
- [3] MITRA S, LAMI M S, UDDIN T M, et al. Prospective multifunctional roles and pharmacological potential of dietary flavonoid narirutin[J]. *Biomedicine & Pharmacotherapy*, 2022, 150: 112932.
- [4] TAO Y J, TAO H F, ZHUANG ZH H, et al. Quantized iterative learning control of communication-constrained systems with encoding and decoding mechanism [J]. *Transactions of the Institute of Measurement and Control*, 2024, 46(10): 1943-1954.
- [5] LIN G CH, TANG Y CH, ZOU X J, et al. Fruit detection in natural environment using partial shape matching and probabilistic Hough transform [J]. *Precision Agriculture*, 2020, 21: 160-177.
- [6] ZHAO CH Y, LEE W S, HE D J. Immature green citrus detection based on colour feature and sum of absolute transformed difference(SATD) using colour images in the citrus grove [J]. *Computers and Electronics in Agriculture*, 2016, 124: 243-253.
- [7] 金婷婷,房建东,赵于东. 注意力特征融合的番茄叶部

- 早期病斑诊断算法[J]. 电子测量技术, 2024, 47(4): 156-164.
- JIN T T, FANG J D, ZHAO Y D. Early lesion diagnosis algorithm for tomato leaves based on attention feature fusion[J]. *Electronic Measurement Technology*, 2024, 47(4): 156-164.
- [8] 陈晓, 夏颖. 基于改进 MobileViT 网络的番茄叶片病害识别[J]. 电子测量技术, 2023, 46(14): 188-196.
- CHEN X, XIA Y. Tomato leaf disease recognition based on improved MobileViT network[J]. *Electronic Measurement Technology*, 2023, 46(14): 188-196.
- [9] ZHENG ZH H, XIONG J T, LIN H, et al. A method of green citrus detection in natural environments using a deep convolutional neural network[J]. *Frontiers in Plant Science*, 2021, 12: 705737.
- [10] 宋中山, 刘越, 郑禄, 等. 基于改进 YOLOV3 的自然环境下绿色柑橘的识别算法[J]. 中国农机化学报, 2021, 42(11): 159-165.
- SONG ZH SH, LIU Y, ZHENG L, et al. Recognition algorithm for green citrus in natural environment based on improved YOLOV3[J]. *Journal of Chinese Agricultural Mechanization*, 2021, 42(11): 159-165.
- [11] ZHANG Y F, LI L, CHUN CH P, et al. Data-driven Bayesian Gaussian mixture optimized anchor box model for accurate and efficient detection of green citrus[J]. *Computers and Electronics in Agriculture*, 2024, 225: 109366.
- [12] 吕强, 林刚, 蒋杰, 等. 基于改进 YOLOv5s 模型的自然场景中绿色柑橘果实检测[J]. 农业工程学报, 2024, 40(18): 147-154.
- LYU Q, LIN G, JIANG J, et al. Green citrus fruit detection in natural scenes based on improved YOLOv5s model [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2024, 40(18): 147-154.
- [13] HAN K, AN X, WU E H, et al. Transformer in transformer [J]. *Advances in Neural Information Processing Systems*, 2021, 34: 15908-15919.
- [14] 刘铁, 段勇. 融合 CNN 和 Transformer 的机器人室内场景识别[J]. 电子测量与仪器学, 2023, 37(5): 223-229.
- LIU T, DUAN Y. Indoor scene recognition of robots by integrating CNN and Transformer[J]. *Journal of Electronic Measurement and Instrumentation*, 2023, 37(5): 223-229.
- [15] MARQUEZ G F, BARBA M A D, FRANCO M E L, et al. Enhancing site-specific weed detection using deep learning transformer architectures [J]. *Crop Protection*, 2025, 190: 107075.
- [16] JI W, ZHAI K L, XU B, et al. Green apple detection method based on multi-dimensional feature extraction network model and Transformer module. [J]. *Journal of Food Protection*, 2024, 88(1): 100397.
- [17] LI X X, CAI M R, TAN X J, et al. An efficient transformer network for detecting multi-scale chicken in complex free-range farming environments via improved RT-DETR[J]. *Computers and Electronics in Agriculture*, 2024, 224: 109160.
- [18] ZHAO Y A, LYU W Y, XU SH L, et al. Detrs beat yolos on real-time object detection [C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 16965-16974.
- [19] FINDER S E, AMOYAL R, TREISTER E, et al. Wavelet convolutions for large receptive fields [C]. *European Conference on Computer Vision*. Springer, Cham, 2025: 363-380.
- [20] WU H L, HUANG P, ZHANG M, et al. CMTFNet: CNN and multiscale transformer fusion network for remote sensing image semantic segmentation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1-12.
- [21] YANG Z M, WU Q L, ZHANG F, et al. A new semantic segmentation method for remote sensing images integrating coordinate attention and SPD-Conv [J]. *Symmetry*, 2023, 15(5): 1037.
- [22] CUI Y N, REN W Q, KNOLL A. Omni-kernel network for image restoration [C]. *AAAI Conference on Artificial Intelligence*, 2024, 38(2): 1426-1434.
- [23] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: Common objects in context [C]. *Computer Vision-ECCV 2014: 13th European Conference*, Springer International Publishing, 2014: 740-755.
- [24] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization [C]. *IEEE International Conference on Computer Vision*, 2017: 618-626.

## 作者简介

秦建华, 博士后, 教授, 硕士生导师, 主要研究方向为机器人、信息检测及处理。

E-mail: qinh2@sina.com

陈振伦, 硕士研究生, 主要研究方向为机器人、计算机视觉。

E-mail: 2845381059@qq.com

万保雄(通信作者), 硕士, 副研究员, 主要研究方向为农业推广。

E-mail: wan77118@163.com

陆泰良, 博士, 助理研究员, 主要研究方向为桃种植资源创新利用和桃栽培技术研究。

E-mail: 595315654@qq.com

雷军乐, 博士, 教授, 硕士生导师, 主要研究方向为农业机械机理。

E-mail: 40911409@qq.com