doi:10.6041/j.issn.1000-1298.2025.10.049

基于改进 YOLO v10n 的百香果成熟果实检测方法

祝诗平　张　越　唐茂杰　邹佳岐　刘寅峰
(西南大学工程技术学院, 重庆 400716)

摘要: 在自然环境下识别定位成熟百香果是实现其自动化采摘的首要任务。然而,在自然环境下,成熟百香果采摘面临光照变换、树枝遮挡和果实重叠等问题,需要一个高效的视觉系统作为依托,实现成熟百香果自动化采摘。为此,本文提出一种基于改进 YOLO v10n 的百香果成熟果实检测模型(RC-YOLO v10n)。在自然环境下采集百香果图像,并对图像中百香果果实进行标注(未成熟、近成熟、成熟时期果实)。以单阶段检测模型 YOLO v10n 为基础,在其 Neck 端引入 RFAConv 模块替换第 9 层 C2f 模块与第 12 层 C2fCIB 模块,以增强模型特征表达能力和特征融合效果;并使用 CARAFE 上采样算子替换原模型中的线性插值上采样模块,以丰富上采样语义信息。试验结果表明,相比原始模型,本文模型精确率提升 3 个百分点,召回率提升 1.6 个百分点,平均精度均值提升 4.6 个百分点,模型内存占用量为 13.72 MB,单幅图像检测时间为 0.22 s。本文模型性能优于 Faster R-CNN 以及 YOLO 系列目标检测算法。将改进前后模型部署到 Jetson Nano 上进行测试,结果表明,改进模型检测效果相比原模型有明显提升,单幅图像检测时间为 1.78 s,改进模型具有较好的应用价值。结合果园百香果实际采收情况给出了 2 种采摘策略:采摘近成熟与成熟百香果,采摘准确率为 89.25%;只采摘成熟百香果,采摘准确率为 92.08%。本文可为成熟百香果识别和自动化采摘提供参考。

关键词: 百香果; 成熟果实; YOLO v10n; RFAConv; CARAFE
中图分类号: TP391.4; S24　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)10-0549-09　　**OSID**:

Ripe Fruit Detection Model of Passion Fruit Based on Improved YOLO v10n

ZHU Shiping　ZHANG Yue　TANG Maojie　ZOU Jiaqi　LIU Yinfeng
(*College of Engineering and Technology*, *Southwest University*, *Chongqing* 400716, *China*)

Abstract: Identifying and locating mature passion fruits in a natural environment is the primary task for achieving their automated picking. However, in natural environments, the picking of ripe passion fruits faces significant challenges such as changes in light, tree branch obstructions, and overlapping fruits, which require an efficient visual system as a support to achieve automated picking of ripe passion fruits. To this end, a passion fruit mature fruit detection model RC-YOLO v10n was proposed based on the improved YOLO v10n. The passion fruit images were collected in a natural environment and the passion fruit fruits in the images were labeled as unripe, nearly ripe, and ripe fruits. Based on the single-stage detection model YOLO v10n, the RFAConv module was introduced at its Neck end to replace the C2f module of the 9th layer and the C2fCIB module of the 12th layer, in order to enhance the feature expression ability and feature fusion effect of the model. And the CARAFE upsampling operator was used to replace the linear interpolation upsampling module in the original model to enrich the semantic information of upsampling. The test results showed that, compared with the original model, the precision of the model was increased by 3 percentage points, the recall rate was increased by 1.6 percentage points, the mean average precision was increased by 4.6 percentage points, the memory usage of the model was 13.72 MB, and the detection time of a single image was 0.22 s. The performance of the model was superior to that of Faster R-CNN and the YOLO series of object detection algorithms. The models before and after the improvement were deployed on Jetson Nano for testing. The results showed that the detection effect of the improved model was significantly improved compared with the original model. The detection time of a single image was 1.78 s. The improved model had good application value. Based on

收稿日期: 2024-06-04　修回日期: 2024-08-27
基金项目: 教育部中国高校产学研创新基金——德州专项(2021DZ005)
作者简介: 祝诗平(1969—),男,教授,博士,主要从事智能检测与控制研究,E-mail: zspswu@126.com

the actual harvesting situation of passion fruits in the orchard, two picking strategies were provided: the first one was to pick near-ripe and ripe passion fruits, with a picking accuracy rate of 89.25%. Secondly, only ripe passion fruits were picked, with a picking accuracy rate of 92.08%. The research result can provide a reference for the research on the identification and automated picking of mature passion fruits.

Key words: passion fruit; ripe fruit; YOLO v10n; RFAConv; CARAFE

0 引言

百香果具有极高的食用价值、药用价值以及经济价值[1]。目前对百香果果实采摘主要依靠人工采摘,这种方式使得其产收效率低、成本高[2]。因此,实现百香果自动化采摘具有重要的实际意义与经济效益。

百香果成熟果实检测是实现自动化采摘的首要任务[3],然而传统方法不适用于复杂环境下百香果果实的检测[4]。随着深度学习[5]在目标检测领域的不断发展,果实的检测识别取得了显著进展,并在自然环境中的果实实时检测取得了较好效果[6]。以 Faster R-CNN[7]、Mask R-CNN[8]为代表的二阶段算法在运算过程中会首先确定目标候选框再通过卷积神经网络完成候选框的分类,这类算法虽然准确度高,但检测速度较慢,不适用于复杂自然环境下果实的实时检测[1]。TU 等[9]基于改进 Faster R-CNN 模型检测百香果,并将 RGB 与深度结合,有效提高了百香果果实与其成熟度的检测性能。YU 等[10]利用 Mask R-CNN 对复杂生长环境下的草莓进行检测,虽然识别精度上取得了较好效果,但检测的实时性能有待提高。李翠明等[11]基于 Faster R-CNN 对苹果进行检测,将 ResNet50 作为主干网络,结合高效通道注意力机制与多尺度融合特征金字塔,实现了复杂场景下苹果的识别与定位。

以 SSD[12]、YOLO 系列算法为代表的单阶段检测算法是对目标直接进行输入端到输出端的回归与分类,这种算法在自然环境中的果实实时检测取得了较好效果[13-15]。综上,基于深度学习的果实检测已受到广泛关注,可为百香果成熟果实检测提供有效参考。然而,在自然环境下,成熟百香果的采摘面临光照变换、树枝遮挡和果实重叠等问题,因此需要一个高效的视觉系统作为依托,实现成熟百香果的自动化采摘。

本文以百香果为识别对象,对检测速度快且检测精度高的 YOLO v10n 模型[16]进行改进。首先将 Neck 端第 9 层 C2f 模块与第 12 层的 C2fCIB 模块替换为感受野注意力卷积(Receptive-field attention convolution,RFAConv)模块,增强模型的学习表达能力;再使用内容感知特征重组(Content-aware reassembly of features,CARAFE)上采样模块,有效提高网络特征提取能力,以有效提高模型的检测精度。通过试验验证 RFAConv-CARAFE-YOLO v10n 模型(RC-YOLO v10n)性能,以实现成熟百香果高效识别,为百香果采摘机器人识别与定位提供参考。

1 材料与方法

1.1 图像采集

采集的百香果图像来自重庆市北碚区某果蔬种植合作社和生态园,于 2023 年 10 月—2024 年 1 月使用手机分别采集露天与温室条件下不同植株及不同成熟时期的百香果图像。为了获得不同光照下的图像,采集时间段为 08:00、13:00、18:00 前后。在图像采集过程中,为提高数据多样性,有效提高模型泛化能力,采集图像类型包括单目标、多目标、顺光、逆光、有无枝叶遮挡、有无果实重叠等。对采集图像进行挑选,去除部分过于模糊以及无法使用的图像,共获取百香果图像 2 090 幅。

1.2 数据集构建

为使检测模型在训练中获取稳定的泛化能力,对百香果图像进行数据预处理,并构建百香果试验数据集。首先将获取的 2 090 幅图像按比例 8∶1∶1[17]随机划分为训练集、验证集、测试集,以进行后续的模型训练和试验。划分后训练集包含 1 674 幅图像,验证集包含 208 幅图像,测试集包含 208 幅图像。为方便后续开展采摘研究,将百香果分为 3 个类别,分别是未成熟、近成熟和成熟。为模拟实际应用场景,提升网络检测能力,分别对训练集、验证集、测试集随机选取含有成熟时期百香果图像进行数据增强。数据增强方式有 2 种:①使图像在(-15°,15°)范围中随机旋转。②模拟不同光照条件随机调整图像饱和度。数据增强后图像如图 1 所示。扩充后训练集包含 2 096 幅图像,验证集、测试集各包含 262 幅图像。



(a) 原图 (b) 随机角度翻转 (c) 调整饱和度

图 1 数据增强后的数据集图像示例

Fig. 1 Example of dataset images after data augmentation

采用 LabelImg 软件对百香果进行标注，标注具体图像如图 2 所示。百香果 3 个类别（未成熟、近成熟、成熟）对应的标注标签分别为“immature”、“near-mature”、“mature”。



图 2　图像类别标注信息

Fig. 2　Image category annotation information

标注后生成 XML 文件，文件存储信息包括：目标类别、标注框坐标位置信息。通过脚本代码将 XML 文件转换成试验网络能够识别的 TXT 文件。最终数据集图像数量与百香果果实标注数量如表 1 所示。

表 1　百香果数据集数量

Tab. 1　Details of passion fruit dataset

数据集	图像数量/幅	标注果实数量/个		
		未成熟	近成熟	成熟
训练集	2 096	2 561	832	741
验证集	262	308	98	91
测试集	262	271	104	88
总计	2 620	3 140	1 034	920

2　百香果检测模型构建

2.1　YOLO v10n 模型

YOLO v10 是性能优越的一阶段模型，适用于输入端到输出端的实时检测。YOLO v10 根据模型深度与宽度不同可分为 5 个版本，其中 YOLO v10n 结构最简洁且检测时间最短[16]，因此本文选用 YOLO v10n，在追求检测精度的同时最大程度地减少模型的冗余度。YOLO v10n 主要由输入端、主干网络、颈部、头部和输出端组成。

相较于 YOLO v8，YOLO v10 设计了部分创新模块，旨在提升模型全局建模能力的同时减少计算开销。首先是空间通道解耦下采样（Spatial-channel decoupling downsampling，SCDown），其结构如图 3 所示。SCDown 由两个卷积操作组成，第 1 个卷积用于调制通道维度，实现通道转换；第 2 个深度卷积实现 2 倍下采样，最大限度地提高了下采样时的信息保留，提升网络性能。



图 3　SCDown 结构

Fig. 3　Structure diagram of SCDown

其次将 C2f 结构中 Bottleneck 模块替换为紧凑型倒置模块，成为 C2fCIB 模块。CIB 模块主要由深度卷积与逐点卷积构成，结构如图 4c 所示。深度卷积对输入特征图的每个通道分别进行卷积操作，然后利用点卷积对深度卷积生成的多个特征图进行融合，减少参数量。这种紧凑架构设计，可有效减少网络冗余。

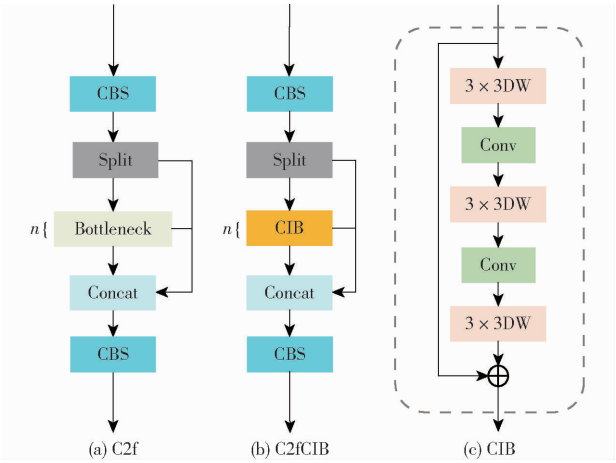


图 4　C2f 与 C2fCIB 结构对比

Fig. 4　Structure comparison between C2f and C2fCIB

此外，部分自注意模块（Partial self-attention，PSA）在 1×1 卷积后将特征图按通道均匀地划分为两部分，仅将其中一部分输入到由多头自注意模块（Multi-head self-attention module，MHSA）和前馈网络（Feed-forward network，FFN）组成的 N_{PSA} 模块中。然后将两个部分特征通过 1×1 卷积融合，从而实现效率和精度的平衡。PSA 结构如图 5 所示。

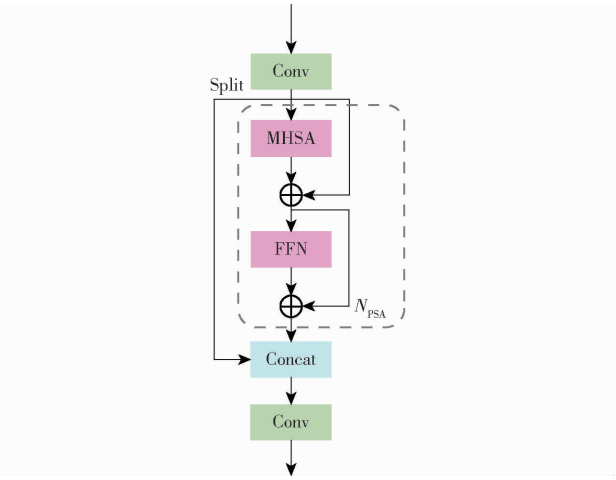


图 5　PSA 结构

Fig. 5　Structure diagram of PSA

YOLO v10 检测头采用一对一和一对多结合的双检测头，结构如图 6 所示。在训练阶段 2 个部分都参与训练，使得模型能够得到一对多检测头的监督；在推理阶段，只使用一对一检测头，可避免非极大值抑制运算的后处理，提高推理速度。

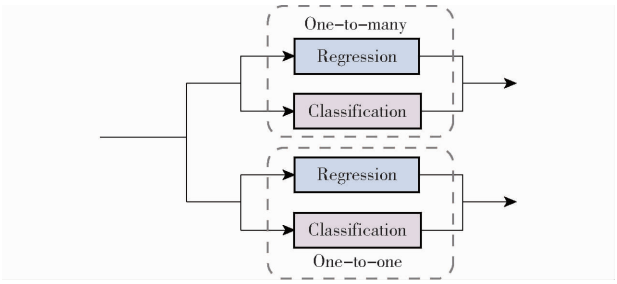


图 6 YOLO v10 检测头结构

Fig. 6 Structure diagram of YOLO v10 detection header

2.2 改进 YOLO v10n 模型

为适应复杂环境背景下成熟百香果识别与检测，以 YOLO v10n 为基础模型，对比在网络不同位置改进 RFAConv 模块，将 Neck 端第 9 层的 C2f 模块与第 12 层 C2fCIB 模块替换为 RFAConv 模块，有效解决了卷积核参数共享的问题，提高了卷积神经网络性能；在 Neck 端将原模型中的插值上采样模块替换成 CARAFE 上采样模块，提升特征融合能力，从而提高识别效率和识别精度。改进模型通过引入少量参数，得到最大的网络性能优化。改进模型（RC－YOLO v10n）网络结构如图 7 所示，图中 C 表示通道数，H 和 W 分别表示特征图高度和宽度，C_m 表示降维后特征层通道数，σ 表示上采样倍数，

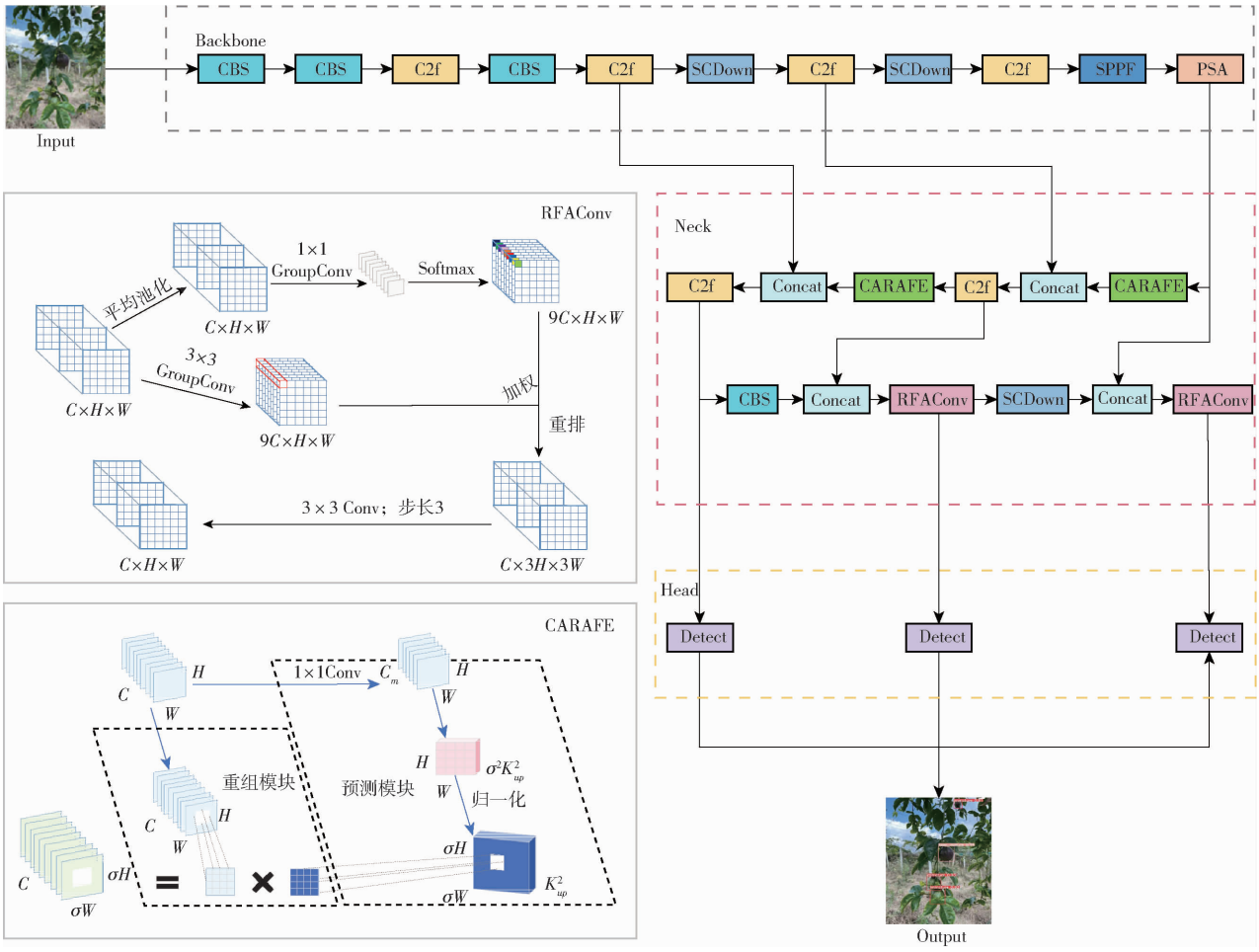


图 7 RC－YOLO v10n 模型结构

Fig. 7 Structure diagram of RC－YOLO v10n model

K_{up} 表示预测上采样核尺寸。

2.2.1 感受野注意力卷积

为解决当前已有的空间注意力机制（如卷积注意力（Convolutional block attention module，CBAM）[18] 和协调注意力（Coordinate attention，CA）[19]）在卷积核参数共享上受限的问题，ZHANG 等[20] 提出了一种新的注意机制——感受野注意力（Receptive-field attention，RFA），RFA 注重感受野滑块中不同的特征，还对感受野中映射的空间特征进行了优先级排序，有效解决了卷积核参数共享的问题。而后通过将感受野注意力机制与标准卷积运算结合设计了一种 RFAConv，可以代替标准卷积来提高卷积神经网络性能，RFAConv 结构如图 7 所示。

当百香果图像输入时，为了减少参数量，利用平均池化汇聚每个感受野特征的全局信息，对百香果特征信息进行压缩；使用 1×1 组卷积将原始特征映射到新特征中，然后使用 Softmax 来强化特征信息；而后将增强后特征与直接使用 3×3 组卷积映射的新特征加权后进行重排调整，调整后得到的特征图

为不重叠的感受野空间特征，其汇总了每个感受野滑块的特征信息，这完全弥补了现有 CA 和 CBAM 注意机制的不足。调整后，根据特征高度和宽度采用相应卷积操作来提取特征信息。

2.2.2　CARAFE 上采样算子

上采样常用于将低分辨率输入映射到高分辨率输出，增加了特征图尺寸，提高了网络的特征提取能力[21]。本研究使用 CARAFE 上采样算子[22]替换原模型中的插值上采样模块，跟以往的上采样不同的是，为更好适应不同特征之间的空间差异，CARAFE 可以根据输入的特征进行自适应完成上采样操作。CARAFE 主要由 2 个部分组成：第 1 部分是核预测模块，首先采用 1 × 1 的卷积核对输入的特征图像进行降维处理，有效减少了后续的计算成本。降维后特征层通道数为[23] $C_m = \sigma^2 K_{up}^2$。压缩后的特征图输入内容编码器后，编码器负责进行内容编码，并设置 $K_{encoder} \times K_{encoder}$（编码器卷积核尺寸）卷积层预测上采样核，改变 $K_{encoder}$ 可改变编码器的感受野，能更有效利用特征信息。最后每个重组核在空间上用 Softmax 函数进行归一化处理得到预测的上采样核。第 2 部分是重组模块，将核预测模块得到的上采样核输入重组模块，与输入的特征图重新组合，得到的结果即为上采样的结果，最后得到的特征图语义信息会比原始特征图更丰富。CARAFE 上采样模块具体结构如图 7 所示。

2.3　模型训练

2.3.1　超参数与环境配置

所有模型均使用相同的超参数进行训练，其中迭代次数为 300；初始学习率设置为 0.001；动量因子为 0.937；批大小为 16；优化器为 Adam；损失函数为 CIoU。改进模型和参与对比试验的模型均在相同环境下进行试验。所用操作系统为 Windows 10，CPU 型号为 Intel Core i5 - 12400F，GPU 型号为 NVIDIA GeForce RTX 3060，运行内存为12 GB，编程语言为 Python 3.9，深度学习框架为 Pytorch 2.0.1。

2.3.2　评价指标

为衡量成熟百香果果实检测模型性能，采用精确率（Precision）、召回率（Recall）、平均精度均值（mean Average Precision，mAP）、单幅图像检测时间（Detection time）、参数量、浮点运算量和模型内存占用量等指标对模型进行评估，其中，平均精度均值采用 IoU 为 0.5 时的结果，即 mAP@0.5。

3　结果与分析

3.1　改进 YOLO v10n 模型试验

3.1.1　改进前后模型对比

改进前后模型对不同时期百香果检测性能对比结果如表 2 所示。由表 2 可知，改进模型精确率、召回率、平均精度均值均有提升，相比原模型精确率提升 3 个百分点，召回率提升 1.6 个百分点。相比原模型未成熟百香果平均精度均值提升 3.6 个百分点，达到 95.60%；近成熟百香果检测平均精度均值大幅提升 7.3 个百分点；而成熟百香果检测平均精度均值提升 14.1 个百分点，达 94.10%。因此 RC - YOLO v10n 可对不同成熟时期百香果实现有效准确识别。RC - YOLO v10n 模型对比原模型，虽引入了少量参数，但改进模型性能得到较大提升，可为后续应用研究提供一定理论支持。

表 2　改进前后模型性能对比

Tab.2　Comparison of model performance before and after improvement

模型	精确率/%	召回率/%	mAP@0.5/%			参数量	浮点运算量	模型内存占用量/MB
			未成熟	近成熟	成熟			
YOLO v10n	81.80	83.90	92.00	79.60	80.00	2.70×10^6	8.20×10^9	10.33
RC - YOLO v10n	84.80	85.50	95.60	86.90	94.10	3.59×10^6	9.60×10^9	13.72

3.1.2　改进前后可视化分析

为直观对比改进前后模型检测能力，将 YOLO v10n 模型和改进模型在测试集上的可视化结果进行比较，结果如图 8 所示。检测的目标置信度设置为 0.5、IoU 设置为 0.6、输入图像尺寸为 640 像素 × 640 像素。分别测试多目标、顺光、逆光、枝叶遮挡、果实重叠 5 种实时检测中常见的复杂环境情况，通过对比改进前后漏检、误检与置信度情况判断模型检测性能是否有效提升。由图 8 可以看出，YOLO v10n 模型在多目标情况下存在 1 个漏检（蓝色圆框为漏检标注），可能是由于模型对小目标的特征提取不足；在顺光、枝叶遮挡 2 种情况下不存在漏检与误检（黄色圆框为误检标注）；在逆光情况下，原模型存在 1 个误检，可能是由于网络模型对目标的特征提取能力较弱，且受复杂背景干扰较大；在果实重叠情况下，原模型存在 1 个漏检，改进模型优化了这一问题，但对远距离的小目标检测能力有待进一步加强。综上，RC - YOLO v10n 模型增强了网络特征提取能力，在 5 种复杂环境情况除了较远目标几乎不存在漏检与误检。此外，相比原模型，RC -YOLO

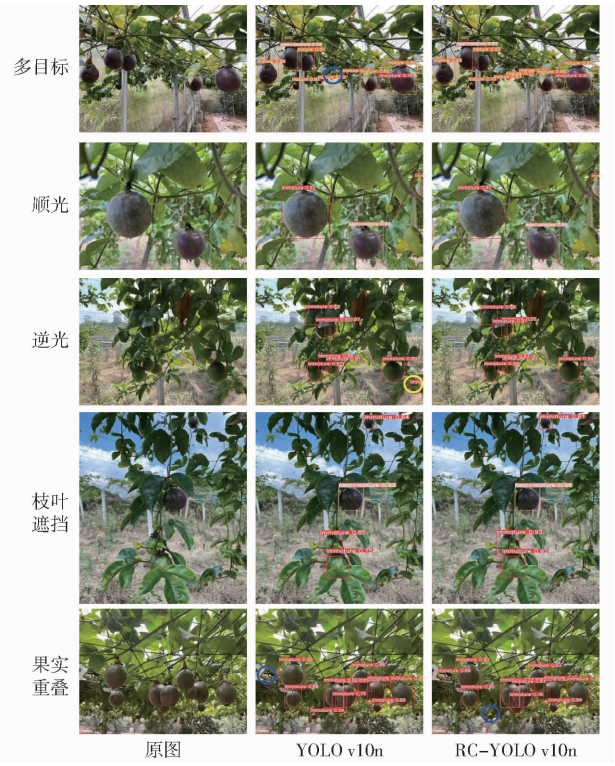


图 8 改进前后检测效果对比

Fig. 8 Comparison of detection results before and after improvement

v10n 模型置信度也有明显提升,为成熟百香果实时检测提供了有效方法。

3.2 改进 YOLO v10n 模型消融试验

3.2.1 RFAConv 模块不同改进位置结果对比

为探究 RFAConv 模块在模型不同位置改进性能,对比了在不同层改进 RFAConv 模块的结果。由表 3 可知,RFAConv 模块在 4 个改进位置的性能相比原模型均有提升,平均精度均值不小于 88.00%。除了位置 3、4 的改进方法,其他改进方法精确率较原模型分别提高 1.5、2.2 个百分点;除了位置 3 的改进方法,其他改进方法召回率较原模型分别提高 0.6、2.9、0.5 个百分点;4 种改进方法的平均精度均值较原模型分别提高 2.9、3.2、1.2、4.2 个百分点,虽位置 4 的改进方法平均精度均值最优,但其参数量与模型内存占用量最大,不适合后续移动端应用。综上,在模型中引入 RFAConv 模块,有助于提升模型性能。其中,位置 2 的改进方法在 4 种改进方法中的综合表现最好,相比原模型虽增加了少量参数,但网络性能得到了有效提升。

表 3 不同改进位置结果对比

Tab. 3 Comparison of results for different improvement positions

位置	精确率/%	召回率/%	mAP@0.5/%	参数量	模型内存占用量/MB
1	83.30	84.50	90.50	3.02×10^6	11.54
2	84.00	86.80	90.80	3.46×10^6	13.21
3	78.70	83.80	88.80	3.90×10^6	14.85
4	81.50	84.40	91.80	4.22×10^6	16.11
YOLO v10n	81.80	83.90	87.60	2.70×10^6	10.33

注:位置 1 为 Backbone 中所有 C2f 模块替换为 RFAConv;2 为 Neck 中第 9 层 C2f 模块与第 12 层 C2fCIB 模块替换为 RFAConv;3 为 Neck 中所有 C2f 与 C2fCIB 模块替换为 RFAConv;4 为 Backbone 与 Neck 中所有 C2f 与 C2fCIB 模块替换为 RFAConv。

为更好地理解模型预测过程,利用梯度加权类激活映射(Gradient-weighted class activation mapping,Grad-CAM)[24]生成自然环境下检测热力图,比较不同改进位置的检测效果。Grad-CAM 通过计算特征图的梯度来生成可视化热力图,可用于分析模型在不同区域的关注度。由图 9a 可以看出,改进位置 1 虽能对百香果特征进行提取,但特征提取能力较弱。图 9c 相比图 9a 对百香果特征的提取能力有进一步加强,但易受背景干扰,可能是由于将部分背景错误识别为了特征。图 9d 相比图 9c 对背景的错误识别有所改善,且特征提取能力有明显增强。由图 9b 可以看出,将本文所采用的 RFAConv 改进方法可从复杂环境中有效提取百香果的特征,从而有效提高检测准确性。

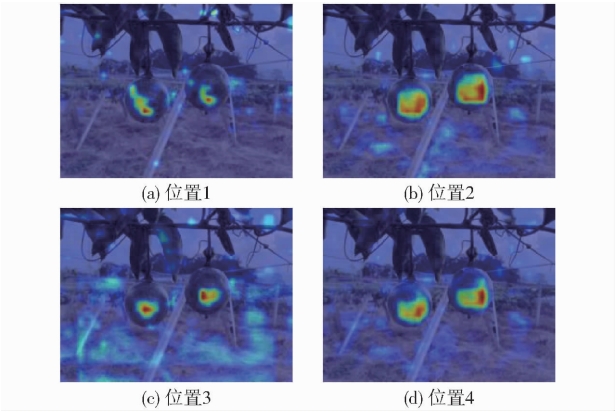


图 9 不同改进位置热力图

Fig. 9 Heat maps of different improvement locations

3.2.2 改进模型消融试验

为进一步验证改进 YOLO v10n 模型性能,设置了消融试验对 4 组网络性能进行验证,分别为 RFAConv 模块改进、CARAFE 上采样算子替换、YOLO v10n 模型以及 RC-YOLO v10n 模型。由表 4 可知,原模型对百香果的识别精确率为 81.80%,召回率为 83.90%,平均精度均值为 87.60%;在添加 RFAConv 模块后,精确率提升 2.2 个百分点,召回率提升 2.9 个百分点,平均精度均值提升 3.2 个百分点;在添加 CARAFE 上采样算子后,虽召回率有所下降,精确率提升 0.4 个百分点,平均精度均值提升 3.3 个百分点。当同时添加 RFAConv 模块和 CARAFE 上采样算子后(即 RC-

表 4　消融试验结果

Tab. 4　Ablation test results　%

模型	精确率	召回率	mAP@0.5
YOLO v10n	81.80	83.90	87.60
YOLO v10n + RFAConv	84.00	86.80	90.80
YOLO v10n + CARAFE	82.20	83.00	90.90
RC - YOLO v10n	84.80	85.50	92.20

YOLO v10n)，模型综合性能最优，各项指标均有小幅提升，精确率提升 3 个百分点，召回率提升 1.6 个百分点，平均精度均值提升 4.6 个百分点。通过消融试验结果可知，改进模型可进一步提升百香果检测性能。

3.3　不同模型结果对比

为客观评价改进模型检测效果，使用常见经典模型 Faster R - CNN[7] 以及 YOLO 系列目标检测算法 YOLO v5n[25]、YOLO v7 - tiny[26]、YOLO v8n[27]、YOLO v8s[27]、YOLO v9s[28]、YOLO v10n 与 RC - YOLO v10n 进行对比试验。同时为了探究模型改进方法的性能，将改进方法应用在 YOLO v5n、YOLO v8n 两种模型内存占用量较小的模型上。以上模型统一在本文构建的百香果数据集上进行训练与测试，同一环境下对比试验结果如表 5 所示。

表 5　不同模型试验结果对比

Tab. 5　Comparison of test results of different models

模型	精确率/%	召回率/%	mAP@0.5/%	检测时间/s	模型内存占用量/MB
Faster R - CNN	61.10	70.40	63.00	1.03	107.88
YOLO v5n	78.60	84.10	86.40	0.11	6.74
YOLO v7 - tiny	81.80	83.30	87.80	0.26	22.89
YOLO v8s	80.10	80.10	86.80	0.22	42.48
YOLO v8n	80.90	82.00	88.60	0.29	11.49
YOLO v9s	84.30	83.10	90.20	0.40	37.17
YOLO v10n	81.80	83.90	87.60	0.20	10.33
RC - YOLO v5n	80.70	81.90	89.80	0.12	8.46
RC - YOLO v8n	83.80	84.30	92.10	0.30	14.10
RC - YOLO v10n	84.80	85.50	92.20	0.22	13.72

由表 5 可知，Faster R - CNN 模型精确率、召回率、平均精度均值均比其他一阶段模型低，模型内存占用量最大的同时所需检测时间也最多。相比二阶段网络，以 YOLO 为代表的一阶段网络更适用于自然环境下百香果的实时检测。在 YOLO 系列模型中，相比 YOLO v5n、YOLO v10n，RC - YOLO v10n 模型内存占用量与检测时间有所增加，但其精确率、召回率、平均精度均值分别比 YOLO v5n 模型高 6.2、1.4、5.8 个百分点；比 YOLO v10n 模型高 3、1.6、4.6 个百分点。与 YOLO v7 - tiny 模型相比，RC - YOLO v10n 模型精确率、召回率、平均精度均值分别高 3、2.2、4.4 个百分点，且模型内存占用量仅为 YOLO v7 - tiny 模型的 60%。与 YOLO v8n 模型相比，RC - YOLO v10n 模型精确率、召回率、平均精度均值分别比 YOLO v8n 模型高 3.9、3.5、3.6 个百分点，且模型更加轻量化，所需检测时间更少。RC - YOLO v10n 模型的精确率、召回率、平均精度均值较 YOLO v8s 模型提升 4.7、5.4、5.4 个百分点；较 YOLO v9s 模型提升 0.5、2.4、2 个百分点。RC - YOLO v10n 模型内存占用量为 YOLO v8s 模型的 1/3；较 YOLO v9s 模型小 23.45 MB。综上，可表明本文所提模型综合性能更佳，可作为自然环境下成熟百香果果实检测的有效方法。

改进 RC - YOLO v5n、RC - YOLO v8n、RC - YOLO v10n 相较各自的原模型，检测时间与模型内存占用量虽有少量增加，但模型精确率、召回率、平均精度均值均有明显提升。因此，可认为提出的改进方法有效提升了模型性能。

3.4　改进前后模型部署测试

为进一步验证改进模型在移动端应用的有效性，将改进前后模型部署到嵌入式平台进行测试，部署测试如图 10 所示。试验平台采用的操作系统为 Ubuntu 18.04，环境配置为 JetPack 4.6、Cuda 10.2、Cudnn 8.2、TensorRT 8.0；编程语言为 C + +；深度学习框架为 Pytorch 1.12.1；嵌入式设备为 Jetson Nano B01，是一款由英伟达公司研发的体积小巧、功能强大的人工智能嵌入式开发板。其 CPU 为四核 ARM，A57@1.43 GHz、GPU 为 128 核 Maxwell，内存为 4 GB。

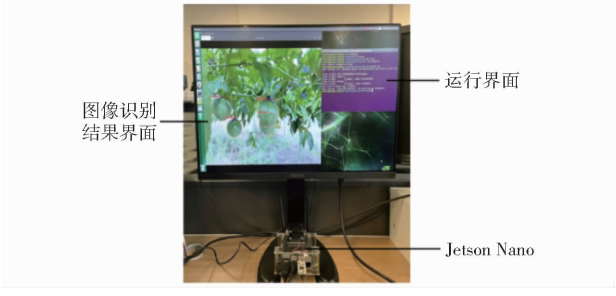


图 10　部署测试

Fig. 10　Deployment test diagram

在深度学习模型实际部署中，传统的直接部署方法往往受限于框架兼容性和运行性能瓶颈。为减少传统部署中对特定框架的依赖，采取了一种高效的模型部署策略。首先通过 PyTorch 框架训练 YOLO v10n 与 RC - YOLO v10n 模型，提炼出最优模型权重；为确保模型在不同硬件平台上的通用性和稳定性，采用开放神经网络交换（Open neural

network exchange,ONNX)格式,以静态批处理(Static batch)的方式导出模型权重,其优势在于明确了推理时的所有维度,确保了推理一致性,并适用于广泛的实际应用场景。随后运用高性能深度学习推理优化器——TensorRT,这种优化器可利用层融合、内存优化和精度校准等优化技术,极大地提升了模型性能。TensorRT 通过内置的 libnvonnxparser. so 库,能够直接解析 ONNX 文件,从而精确构建出优化后的模型进行测试。通过这种部署方法只需训练得到的最优模型权重即可,不依赖框架推理,可有效减少模型运行时间,可实现深度学习模型快速、高效部署。

通过对比改进前后模型在 Jetson Nano 平台上部署测试情况,可进一步验证改进模型性能的明显提升。如表 6 所示,尽管改进模型在模型内存占用量上略有增加,但其检测时间相较原模型仅多 0.05 s,对实时检测效果影响不大。此外,从图 11 可以看出,原模型在百香果检测过程中存在明显漏检现象(蓝色圆框为漏检),改进模型不仅不存在漏检情况,而且检测置信度也有所提升。综合考虑检测速度与检测效果可知,改进模型可满足实际应用中对检测准确性和高效性的要求,具有较好的应用价值。

表 6 改进前后模型部署测试对比

Tab. 6 Comparison of model deployment tests before and after improvement

模型	输入图像尺寸/(像素×像素)	模型内存占用量/MB	检测时间/s
YOLO v10n	640×640	10.33	1.73
RC-YOLO v10n	640×640	13.72	1.78



图 11 改进前后模型部署测试

Fig. 11 Model deployment test diagrams before and after improvement

3.5 采摘准确率预测

根据果园百香果实际采收情况,近成熟时期百香果根据距离可储藏运输外地销售,成熟时期百香果可就近上市鲜销。采摘策略可分为 2 种:①采摘近成熟与成熟百香果,此策略下只要识别为近成熟百香果或成熟百香果就会被采摘。②只采摘成熟百香果。RC-YOLO v10n 在测试集上的推理结果如表 7 所示,根据推理结果对两种采摘率准确率进行计算(准确率为识别正确数量占识别总数百分比),计算可得 2 种采摘策略准确率分别为 89.25%、92.08%。

表 7 测试集推理结果

Tab. 7 Test set detection results

果实类别	识别正确数	识别错误数	总计
未成熟	257	33	290
近成熟	73	12	85
成熟	93	8	101

4 结论

(1)采集自然环境中各种复杂情况下图像,将图像中百香果分为未成熟、近成熟和成熟 3 个类别。在 YOLO v10n 基础上引入 RFAConv 模块;并使用 CARAFE 上采样算子替换原模型中的插值上采样模块。改进模型增强了对百香果果实的特征提取能力,从而有效提高了模型的检测性能。

(2)为验证 RC-YOLO v10n 模型性能设置了 4 组网络定量进行消融试验,结果表明改进模型在引入少量参数后,检测性能得到有效提升。通过与 Faster R-CNN、YOLO 系列模型对比,改进模型在百香果数据集上取得较好效果,证明本文模型在识别准确度上有明显提升,对百香果的检测识别可以为成熟百香果智能采摘的识别与定位提供一定理论依据。

(3)采取一种高效部署策略,将改进前后模型部署至 Jetson Nano。部署测试结果表明,改进模型可满足实际应用中对检测准确性和高效性的要求,具有较好的应用价值。最后结合果园百香果实际采收情况,给出 2 种采摘策略:采摘近成熟与成熟百香果,采摘准确率为 89.25%;只采摘成熟百香果,采摘准确率为 92.08%。采摘策略可为果园百香果实际采收提供参考。

参 考 文 献

[1] 张展榜. 基于嵌入式平台的百香果识别和定位技术研究[D]. 福州:福建农林大学, 2022.
ZHANG Zhanbang. Research on identification and positioning technology of passion fruit based on embedded platform[D]. Fuzhou: Fujian Agriculture and Forestry University, 2022. (in Chinese)

[2] 张日红, 欧炬基, 丁力行, 等. 基于视觉定位的串并混联型百香果采摘机器人研究[J]. 中国农机化学报, 2023, 44(10): 209-216, 244.
ZHANG Rihong, OU Juji, DING Lixing, et al. Research on series parallel hybrid passion fruit picking robot based on vision

positioning[J]. Journal of Chinese Agricultural Mechanization, 2023, 44(10): 209 - 216, 244. (in Chinese)
[3] 唐熔钗. 基于深度学习的百香果成熟度实时检测算法研究[D]. 桂林：桂林电子科技大学, 2022.
TANG Rongchai. Research on real-time detection algorithm of passion fruit maturity via deep learning[D]. Guilin: Guilin University of Electronic Science and Technology, 2022. (in Chinese)
[4] 罗志聪, 李鹏博, 宋飞宇, 等. 嵌入式设备的轻量化百香果检测模型[J]. 农业机械学报, 2022, 53(11): 262 - 269, 322.
LUO Zhicong, LI Pengbo, SONG Feiyu, et al. Lightweight passion fruit detection model based on embedded devices[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(11): 262 - 269, 322. (in Chinese)
[5] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436 - 444.
[6] 宋怀波, 尚钰莹, 何东健. 果实目标深度学习识别技术研究进展[J]. 农业机械学报, 2023, 54(1): 1 - 19.
SONG Huaibo, SHANG Yuying, HE Dongjian. Review on deep learning technology for fruit target recognition[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 1 - 19. (in Chinese)
[7] REN S, HE K, GIRSHICK R, et al. Faster R - CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137 - 1149.
[8] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R - CNN[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961 - 2969.
[9] TU S, XUE Y, ZHENG C, et al. Detection of passion fruits and maturity classification using Red - Green - Blue Depth images [J]. Biosystems Engineering, 2018, 175: 156 - 167.
[10] YU Y, ZHANG K, YANG L, et al. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask - RCNN[J]. Computers and Electronics in Agriculture, 2019, 163: 104846.
[11] 李翠明, 杨柯, 申涛, 等. 基于改进 Faster R - CNN 的苹果采摘视觉定位与检测方法[J]. 农业机械学报, 2024, 55(1): 47 - 54.
LI Cuiming, YANG Ke, SHEN Tao, et al. Vision detection method for picking robots based on improved Faster R - CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 47 - 54. (in Chinese)
[12] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multiBox detector[J]. arXiv Preprint, arXiv:1512.02325, 2015.
[13] 王昱, 姚兴智, 李斌, 等. 基于改进 YOLO v7 - tiny 的甜椒畸形果识别算法[J]. 农业机械学报, 2023, 54(11): 236 - 246.
WANG Yu, YAO Xingzhi, LI Bin, et al. Malformed sweet pepper fruit identification algorithm based on improved YOLO v7 - tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 236 - 246. (in Chinese)
[14] BAI Y, YU J, YANG S, et al. An improved YOLO algorithm for detecting flowers and fruits on strawberry seedlings[J]. Biosystems Engineering, 2024, 237: 1 - 12.
[15] 罗志聪, 何陈涛, 陈登捷, 等. 基于轻量化 YOLO v8s - GD 的自然环境下百香果快速检测模型[J]. 农业机械学报, 2024, 55(8): 291 - 300.
LUO Zhicong, HE Chentao, CHEN Dengjie, et al. Passion fruit rapid detection model based on lightweight YOLO v8s - GD [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(8): 291 - 300. (in Chinese)
[16] WANG A, CHEN H, LIU L, et al. YOLO v10: real-time end-to-end object detection[J]. arXiv Preprint, arXiv:2405.14458, 2024.
[17] 赵辉, 黄镖, 王红君, 等. 基于改进 YOLO v7 的农田复杂环境下害虫识别算法研究[J]. 农业机械学报, 2023, 54(10): 246 - 254.
ZHAO Hui, HUANG Biao, WANG Hongjun, et al. Pest identification method in complex farmland environment based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 246 - 254. (in Chinese)
[18] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[J]. arXiv Preprint, arXiv:1807.06521, 2018.
[19] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713 - 13722.
[20] ZHANG X, LIU C, YANG D, et al. RFAConv: innovating spatial attention and standard convolutional operation[J]. arXiv Preprint, arXiv:2304.03198, 2023.
[21] 张旭, 陈慈发, 董方敏. 基于改进 YOLOv7 的 PCB 缺陷检测算法[J]. 计算机工程, 2024, 50(12): 318 - 328.
ZHANG Xu, CHEN Cifa, DONG Fangmin. PCB defect detection algorithm based on improved YOLO v7[J]. Computer Engineering, 2024, 50(12): 318 - 328. (in Chinese)
[22] WANG J, CHEN K, XU R, et al. Carafe: content-aware reassembly of features[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 3007 - 3016.
[23] 范天浩, 顾寄南, 王文波, 等. 基于改进 YOLO v5s 的轻量化金银花识别方法[J]. 农业工程学报, 2023, 39(11): 192 - 200.
FAN Tianhao, GU Ji'nan, WANG Wenbo, et al. Lightweight honeysuckle recognition method based on improved YOLO v5s [J]. Transactions of the CSAE, 2023, 39(11): 192 - 200. (in Chinese)

(下转第 595 页)

[18] 刘成,岳训. 基于 YOLO v5 + DeepSORT 算法的羊群游走同步群体决策行为研究[J]. 农业机械学报,2024,55(6):229 - 236.
LIU Cheng, YUE Xun. Experiment of synchronized group decision-making behavior under herding walk model based on YOLO v5 + DeepSORT algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(6): 229 - 236. (in Chinese)

[19] 黄成龙,张忠福,华向东,等. 基于改进 Faster R - CNN 和 Deep Sort 的棉铃跟踪计数[J]. 农业机械学报,2023,54(6):205 - 213.
HUANG Chenglong, ZHANG Zhongfu, HUA Xiangdong, et al. Cotton boll tracking and counting based on improved Faster R - CNN and Deep Sort[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 205 - 213. (in Chinese)

[20] CAO J, WENG X, PANG J, et al. Observation-Centric SORT: rethinking sort for robust multi-object tracking[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 9686 - 9696.

[21] AHARON N, ORFAIG R, BOBROVSKY B Z. BoT - SORT: robust associations multi-pedestrian tracking[EB/OL]. (2022 - 06 - 29)[2024 - 04 - 28]. https://arxiv.org/abs/2206.14651.

[22] WANG C, YEH I, LIAO H. YOLOv9: learning what you want to learn using programmable gradient information[EB/OL]. (2024 - 02 - 21)[2024 - 04 - 28]. https://arxiv.org/abs/2402.13616.

[23] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 8759 - 8768.

[24] ZHANG Y, SUN P, JIANG Y, et al. ByteTrack: multi-object tracking by associating every detection box[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2022: 1 - 21.

[25] 王雒,李飚,傅瑞罡. 基于改进 ByteTrack 算法的红外地面多目标跟踪方法[J]. 计算机科学,2023,50(9):176 - 183.
WANG Luo, LI Biao, FU Ruigang. Infrared ground multi-object tracking method based on improved ByteTrack algorithm[J]. Computer Science, 2023, 50(9): 176 - 183. (in Chinese)

[26] PADILLA R, NETTO S L, DA SILVA E A B. A survey on performance metrics for object-detection algorithms[C]// 2020 International Conference on Systems, Signals and Image Processing (IWSSIP). IEEE, 2020: 237 - 242.

[27] 韩锟,彭晶莹. 基于改进 YOLOX 与多级数据关联的行人多目标跟踪算法研究[J]. 铁道科学与工程学报,2024, 21(1): 94 - 105.
HAN Kun, PENG Jingying. Pedestrian multi-object tracking algorithm based on improved YOLOX and multi-level data association[J]. Journal of Railway Science and Engineering, 2024, 21(1): 94 - 105. (in Chinese)

[28] 梁秀英,贾学镇,何磊,等. 基于 YOLO v8n-seg 和改进 Strongsort 的多目标小鼠跟踪方法[J]. 农业机械学报,2024, 55(2):295 - 305,345.
LIANG Xiuying, JIA Xuezhen, HE Lei, et al. Multi-target mice tracking based on YOLO v8n-seg and improved Strongsort[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 295 - 305, 345. (in Chinese)

(上接第 557 页)

[24] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad - CAM: visual explanations from deep networks via gradient-based localization[J]. arXiv Preprint, arXiv: 1610.02391, 2016.

[25] ZHU X K, LYU S C, WANG X, et al. TPH - YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2778 - 2788.

[26] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLO v7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. arXiv Preprint, arXiv: 2207.02696, 2022.

[27] LOU H T, DUAN X H, GUO J M, et al. DC - YOLO v8: small-size object detection algorithm based on camera sensor[J]. Electronics, 2023, 12(10): 2323.

[28] WANG C Y, YEH I H, LIAO H Y M. YOLO v9: learning what you want to learn using programmable gradient information[J]. arXiv Preprint, arXiv: 2402.13616, 2024.