

利用改进 YOLOv5s 模型检测番茄果实成熟度及外观品质

孙宇朝^{1,2}, 李守豪^{1,2}, 夏秀波^{1,3,*}, 杨 珩^{2,*}, 李民赞², 张焕春^{1,3}

(¹ 山东省烟台市农业科学研究院, 山东烟台 265500; ² 中国农业大学智慧农业系统集成研究教育部重点实验室, 北京 100083; ³ 烟台市智慧农业研究中心, 山东烟台 265500)

摘要: 以粉果番茄为试验材料, 基于深度学习方法开展了番茄果实成熟度和外观品质的检测研究。

试验中共采集番茄图片数据 2 036 张, 通过处理扩增至 5 316 张, 然后将数据进行标注和文件转换, 构建了试验用数据集; 通过在 YOLOv5s 模型中加入 CA 注意力机制、替换 Stem block 结构、结合识别需求优化检测层尺度、替换 K-means++聚类算法来实现 SC-YOLOv5s 识别精度提升, 提高模型的特征表达能力; 通过在 SC-YOLOv5s 模型中加入 Fire module 结构进行轻量化卷积、降低 Bottleneck 模块的参数量来实现 SC-YOLOv5s-lite 轻量化设计, 提升模型在硬件上的检测速度; 将 SC-YOLOv5s-lite 模型在训练集上进行训练优化、消融试验和性能对比, 结果表明, SC-YOLOv5s-lite 模型内存大小为 7.73 M, 其准确率为 89.04%, 召回率 83.35%, 平均精度 91.34%, 检测时间为 143 ms, 相比于 YOLOv5s, 模型参数量降低了 45.57%, 模型大小压缩了 44.86%, 平均精度提升 3.98%, 检测时间减少 20.99%, 优势明显, 更适合于硬件上部署。

关键词: 番茄; 成熟度; 外观品质; 检测; 深度学习; 计算机视觉

中图分类号: S 641.2

文献标志码: A

文章编号: 0513-353X (2024) 02-0396-15

Detecting Tomato Fruit Ripeness and Appearance Quality Based on Improved YOLOv5s

SUN Yuchao^{1,2}, LI Shouhao^{1,2}, XIA Xiubo^{1,3,*}, YANG Wei^{2,*}, LI Minzan², and ZHANG Huanchun^{1,3}

(¹Yantai Academy of Agricultural Sciences in Shandong Province, Yantai, Shandong 265500, China; ²Key Laboratory of Smart Agriculture Systems, Ministry of Education, China Agricultural University, Beijing 100083, China; ³Yantai Smart Agriculture Research Center, Yantai, Shandong 265500, China)

Abstract: The study was conducted on the detection of tomato fruit maturity and appearance quality based on deep learning methods using pink tomato as the experimental material. Two thousand and thirty-six tomato image data were collected and amplified to 5 316 through preprocessing. Then, the data was annotated and converted into files to construct an experimental dataset. The experiment improves the accuracy of SC-YOLOv5s by adding CA attention mechanism, replacing the Stem block structure, optimizing the detection layer scale based on recognition requirements, and replacing the K-means++ clustering algorithm to improve the model's feature expression ability. By adding a fire module structure to

收稿日期: 2023-10-30; **修回日期:** 2023-12-29

基金项目: 山东省现代优势产业集群 + 人工智能试点示范单位项目(鲁工信工联〔2020〕89号); 烟台市设施番茄育种攻关团队项目(烟农[2023]174号)

* 通信作者 Author for correspondence (E-mail: xiuboxia@163.com, cauyw@cau.edu.cn)

SC-YOLOv5s for lightweight convolution and reducing the parameter count of the Bottleneck module, the SC-YOLOv5s-lite lightweight design is achieved, improving the detection speed of the model on hardware; Train and optimize the SC-YOLOv5s-lite model on the training set. The results showed that the memory usage of the SC-YOLOv5s-lite model was 7.73 M, with an accuracy rate of 89.04%, a recall rate of 83.35%, an average accuracy of 91.34%, and a detection time of 143 ms. Compared to YOLOv5s, the model parameter quantity is reduced by 54.57%, model size is compressed by 44.86%, with an average accuracy improvement of 3.98%, and the detection time is reduced by 20.99%. It has obvious advantages and is more suitable for hardware deployment.

Keywords: tomato; maturity; appearance quality; detection; deep learning; computer vision

2021 年中国番茄种植面积约为 111.3 万 hm², 产量约为 6 609 万 t, 占全球 1/3 以上。但目前生产中番茄人工采摘效率低、成本高, 成为制约产业发展的重要因素(刘铮和魏传成, 2021; 黄波, 2022)。利用人工智能技术实现某些生产环节智能化、无人化, 是提高生产效率和质量的重要措施(杜名扬和张天柱, 2022; 冯帆, 2022; 罗锡文 等, 2022)。

在生产中, 要实现产品的智能化采摘, 首先需要准确地识别果实, 尤其是成熟的果实。在人工智能方面, 深度学习中的卷积神经网络具有更快的处理速度和更高的准确率(张淮峰 等, 2006), 已成为农产品识别的重要研究方向。基于深度学习的计算机视觉技术已被广泛应用于农作物的产量预测, 产品分级、品质管控和病虫害诊断等(周升和 等, 2023)。张磊等(2022)使用 YOLOv4 主干网络与 VGG19 卷积层相结合的改进模型, 经迁移学习后对番茄果实进行了识别检测, 改进模型的平均精度(mean Average Precision, mAP)达到了 89.04%; 熊俊涛等(2020)利用 YOLOv5-lite 对自然环境下树上木瓜的成熟度进行了识别, 模型的平均精度相比 YOLOv5s 提高了 1.1%, 检测时间达到了 7 ms, 而模型大小仅为 11.3 M; 孙丰刚等(2022)通过引入 CBMA 和 BIFPN 后, 得到了改进的 GHTR2-YOLOv5s 模型, 经过迁移学习后对苹果果实病虫害识别的平均精度提升了 0.7%; 黄彤镔等(2022)在自然环境下利用基于 YOLOv5 的改进模型对柑橘进行了识别, 其精度值达到 91.3%。但是模型的识别场景单一, 对复杂背景中重叠遮挡柑橘的识别困难, 存在漏检现象。前人研究主要集中在成熟果实的识别上, 而同时能实现产品外观品质检测与果实成熟度分级的较少, 并且检测结果受光照和遮挡等的影响较大。

本试验在温室自然条件下采集了涵盖不同天气条件、遮挡等多种因素的番茄果实图像, 利用改进的深度学习方法进行模型训练, 来实现对不同成熟度果实的准确识别与外观品质分级, 以期为后续机械化、智能采摘和产量预估提供理论和技术支持。

1 材料与方法

1.1 数据集采集

粉果番茄图像数据分别于 2022 年 7 月 1 日和 8 月 15 日自烟台市农业科学研究院温室采集。采集设备为 iPhone 8plus 手机智能终端。采集的图片涵盖了晴天、阴天、早上和傍晚等天气条件, 以及果实叠加度($\leq 70\%$)、枝叶遮挡度($\leq 70\%$)、背景环境、成熟度等多种因素, 同时还包括单一因素和多类因素混搭等类型, 共采集到 2 036 张番茄图片, 图像分辨率为 $3\,000 \times 4\,000$, 格式为 jpg。

1.2 番茄成熟度和外观品质分级

参照 T/TFHT C007-2019 行业标准, 对果实成熟度和外观品质进行分级(表 1)。图 1 是对应类别番茄果实的分级图像数据展示。

表 1 番茄果实不同成熟度与外观品质分级标准
Table 1 Classification standards for different maturity and appearance quality of tomato fruits

成熟度分类 Maturity classification	外观品质分级 Appearance quality grading	主要特点 Main features
成熟果 Ripe fruit	优等 Superior	商品果, 果实周正, 无瑕疵。着色率 $\geq 80\%$ Commodity fruit, and the fruit is round and flawless. Coloring rate $\geq 80\%$
	二等 Second-class	非商品果, 但可食用, 有瑕疵, 裂纹, 轻微日灼。着色率 $\geq 70\%$ Non commercial fruit, but edible, with defects, cracks, and slight sunburn. Coloring rate $\geq 70\%$
	劣等 Inferior	非商品果, 有明显的腐烂, 挤压伤, 裂果, 表面有虫。着色率 $\geq 70\%$ Non commercial fruits, with obvious decay, crushing damage, cracking, and insect infestation on the surface. Coloring rate $\geq 70\%$
半熟果 Semi-ripe fruit	不分等级 Not graded	白粉色, 或者着色率 $< 70\%$ White and pink, or coloring rate $< 70\%$
未熟果 Unripe fruit	不分等级 Not graded	白果、绿果 White and green fruits



图 1 番茄果实不同成熟度及外观品质分级样本

a: 优等成熟果; b: 二等成熟果; c: 劣等成熟果; d: 半熟果; e、f: 未熟果。

Fig. 1 Tomato samples with different maturity and appearance quality grading of tomato fruits

a: Ripe superior fruit; b: Ripe second-class fruit; c: Ripe inferior fruit; d: Semi-ripe fruit; e, f: Unripe fruit.

1.3 数据扩增与标注

通过网络模型过拟合(张会丰等, 2022)或欠拟合, 提高模型的泛化能力。首先对番茄图像进行预处理(查文文等, 2023), 通过采用随机图像缩放和旋转, 翻转, 随机剪裁或补零, 镜像平移(朱德利等, 2023), 对比度变化, 亮度变化, 添加不同范围和大小的高斯噪声(陈乔松等, 2023)等预处理方法, 可以增加训练样本的多样性, 提升算法模型的泛化能力。经过上述数据增强后, 将原有的 2 036 张图片扩增至 5 316 张, 其中存在光照明暗度 1 237 张, 果实叠加 2 781 张, 枝叶遮挡 2 355 张, 多因素混搭 3 876 张。

按照制定的分类标准, 采用 Labelme 图像标注工具对扩增后图片中的番茄依次进行边界框标注(阳茜汕等, 2023), 具体见图 2。标注完毕的番茄图片信息保存在一个与图片同名的.json 文件中。将标注之后的.json 数据转换成.txt 格式文件(吴故等, 2023)。将数据集按照 8:1:1 的比例划分为训练集、测试集和验证集(王亮和张超, 2023), 图片数量依次为 4 253、532、531 张。

1.4 模型训练环境搭建与改进

1.4.1 模型训练环境搭建

采用以 Python 语言为主 Pytorch 的深度学习框架, 具体硬件信息见表 2。考虑到需要将模型部署到嵌入式设备中, 选用 YOLOv5 系列中最轻型的 YOLOv5s 作为基准模型进行改进, 改进方向为精度提升和轻量化设计(马金林等, 2022)。

1.4.2 模型精度提升

改进方法主要有加入 CA 注意力机制(李岩舟等, 2023)、替换 Stem block 结构、检测层尺度优化、替换 K-means++聚类算法。实现这 4 种改进策略的番茄识别算法称为 SC-YOLOv5s。具体过程如下:

(1) CA 注意力机制: 在 YOLOv5s 网络模型中加入 CA 注意力机制。如图 3 所示, CA 结构包含 Coordinate 协调信息嵌入和 Coordinate Attention 协调注意力信息生成两个核心步骤。将 CA 注意力机制加入到 YOLOv5s 模型 C3 模块之后加强了 YOLOv5s 模型对重要特征信息的关注和提取能力。

(2) Stem block 结构: 选择去掉原 YOLOv5s 网络结构中的 Focus 层和第一个 Conv 层, 使用 Stem block 结构(Wang et al., 2022) 作为 YOLOv5s 网络输入的第一层, 来有效提升网络的特征提取能力。Stem block 网络结构如图 4 所示, Stem block 结构从多个层次中提取特征丰富了输入图像的特征层, 因此有效地提升了 YOLOv5s 网络的特征表达能力。同时从图 4 中还可以看出, Stem block 结构中每个卷积层的通道数相对较少, 最终输出的通道数仅为 32, 所以 Stem block 模块对 YOLOv5s 的参数进行了有效缩减。

(3) 检测层尺度优化: 在制作数据集时采集的番茄图片的分辨率为 $3\ 000 \times 4\ 000$, 采集的番茄为粉果, 每张图片中番茄的尺寸均处于 40×40 分辨率以上, 成熟的果实一般是 30×30 分辨率。由于试验检测番茄果实的成熟度及其外观品质, 重点是对已成熟番茄进行检测, 即使模型漏检了一些番茄幼果(小个的未成熟番茄), 不影响检测和采摘效果。因此去掉了小型番茄这类小目标检测层的网络结构, 来减少模型的参数量。

(4) K-means++聚类 anchor box: Means++是一种改进的 K-means 聚类算法, 其通过一种启发式的方式选择初始聚类中心点, 有效避免局部最优解, 更好地进行初始化聚类中心, 从而提高聚类结果的准确性和稳定性。因而采用 K-means++聚类 anchor box 替代了 YOLOv5 中 K-means 聚类算法。

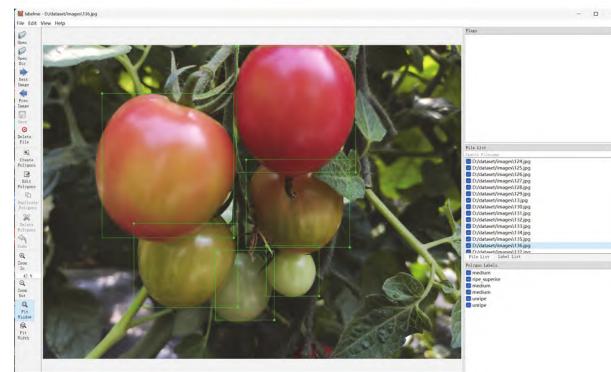


图 2 Labelme 图像标注面板
Fig. 2 Labelme image annotation panel

表 2 试验平台硬件及软件信息

Table 2 Hardware and software information of the test platform

配置名称	Configuration name	详细信息	Detailed information
CPU		Intel (R) Core i7-12700H	
GPU		NVIDIA GEFORCE RTX 3060	
显存	Video storage	16GB	
操作系统	Operating system	Windows11	
深度学习框架		Pytorch1.9.0	
Deep learning framework			
开发语言	Development language	Python3.9	
开发 IDE	Develop IDE	Pycharm	
虚拟环境	Virtual environment	Anaconda	

具体流程图见图 5。

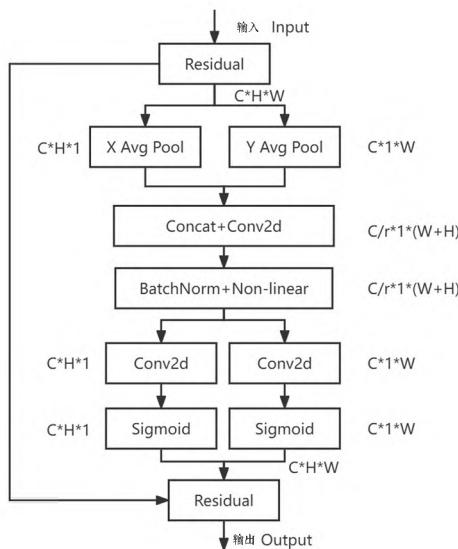


图 3 CA 结构图

Fig. 3 CA structure diagram

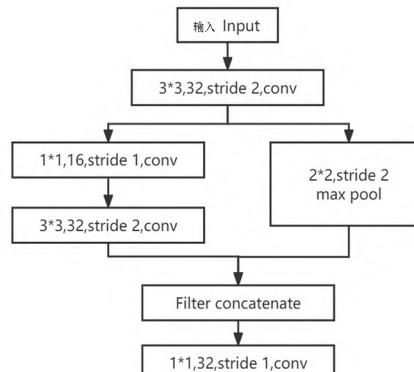


图 4 Stem block 结构图

Fig. 4 Structure diagram of Stem block

1.4.3 模型轻量化设计

番茄成熟度及外观品质检测在要求检测精度的同时，对实时性也有很高要求，所以对 SC-YOLOv5s 算法进行模型通道轻量化设计，新的测算法称为 SC-YOLOv5s-lite。

(1) 添加 Fire module 结构：在保证检测精度的同时，为了实现模型轻量化和快速检测能力，采用了 Fire module 结构代替 SC-YOLOv5s 网络中 3×3 、步长为 1 的卷积层，来提高网络的计算速度，并保证检测的精度。具体流程见图 6 和图 7。

(2) 残差模块深度适配：对 YOLOv5s 中 C3 模块的 Bottleneck（高彦钊和陶常勇，2023）残差模块进行改进，改进后的 Bottleneck 结构如图 8 所示，其中 3×3 卷积的通道数量已经减少了一半。虽然提高了模型的性能，但是由于更深的 Bottleneck 卷积层带来了模型参数量增长和多次叠加无关背景的问题，为了降低模型大小要继续进行 Bottleneck 数量的重新适配。YOLOv5 主干网络存在 4 个 C3 模块，4 个 C3 中的

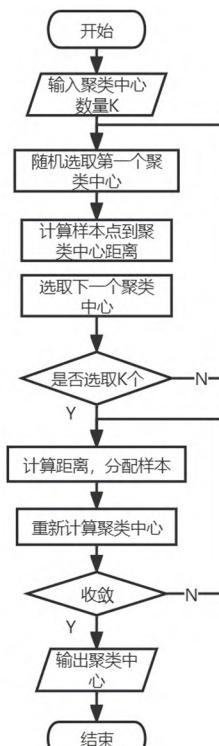


图 5 K-means++聚类算法流程图

Y: 是; N: 否.

Fig. 5 Flowchart of K-means++ clustering algorithm
 Y: Yes; N: No.

Bottleneck 模块数依次为 1、3、3、1 个, 经过试验发现, 在对番茄数据集进行检测的时候, C3 中的 Bottleneck 模块存在部分冗余, 删减这些冗余的 Bottleneck 模块并不影响检测模型的速度和精度, 对比结果显示将 Bottleneck 的数量降低为 1、2、2、1 个的时候可以兼顾检测速度和精度。

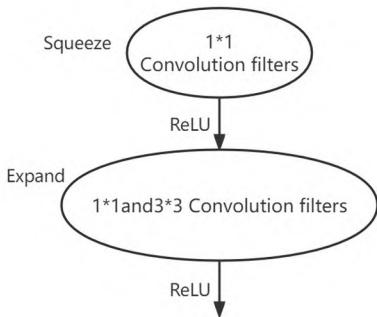


图 6 Fire module 结构图
Fig. 6 Fire module structure diagram

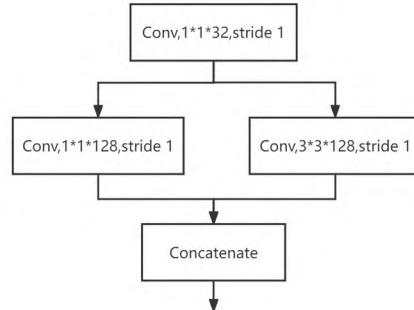


图 7 Fire module 详细连接图
Fig. 7 Fire module detailed connection diagram

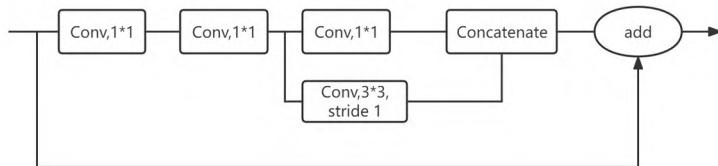


图 8 改进后 Bottleneck 结构图
Fig. 8 Structure diagram of improved Bottleneck

1.5 模型的评价指标

对于部署到嵌入式设备的深度学习模型, 必须要考虑到模型的大小, 识别精度和检测速度等(贺杰 等, 2022)。用精确率(AP)、召回率(Recall)、平均精度(mAP)、 F_1 [$F_1 = (2 \times \text{精确率} \times \text{召回率}) / (\text{精确率} + \text{召回率})$]、检测时间(t), 模型大小和模型参数量等作为模型的评价指标(王海群 等, 2023)。

2 结果与分析

2.1 模型消融试验

2.1.1 SC-YOLOv5s 消融试验

从表 3 中可以看出, YOLOv5s 模型加入 CA 注意力机制后模型的平均精度提高了 1.15%, 表明 CA 注意力机制确实使模型更加关注有效特征信息, 但模型的参数数量增加了 11 590 个, 检测时间增加了 4 ms, 表明加入 CA 注意力机制并不会显著增加计算时间; 在模型网络输入的第一层中引入 Stem block 结构后, 模型的平均精度提升了 1.35%, 但检测时间又增加了 8 ms, 说明 Stem block 结构可以对模型的特征层进行丰富, 提高特征表达能力; 结合数据集番茄对象特征对模型检测层尺度进行优化, 去掉 80×80 小目标检测层后, 模型大小降低了 24.84%, 模型参数量降低了 25.51%, 平

均精度提升了 0.13%，说明针对检测层进行尺度优化对本数据集是有效的，可以明显降低模型大小；将 YOLOv5s 中的 K-means 聚类算法替换成 K-means++聚类后，发现模型的平均精度又提升了 0.76%，说明用 K-means++聚类算法优化初始聚类中心能提高模型的识别效果。综上可知，在递进使用了 4 种改进方法后的 SC-YOLOv5s 模型的平均精度比 YOLOv5s 提高了 3.42%，其模型大小比 YOLOv5s 低 24.25%，其检测时间仅增加 7 ms，由此可知：SC-YOLOv5s 在保证模型的检测速度不受影响的同时，降低了模型大小并提高了模型的平均精度值。

表 3 SC-YOLOv5s 消融试验
Table 3 SC-YOLOv5s ablation test

模型 Model	精度提升策略 Precision improvement strategy				参数量 Parameter quantity	模型大小/M Model memory	平均精度/% mAP	检测时间/ms Detection time
	CA 注意力 机制	Stem Block 结构	优化检测层尺度 Optimize detection layer scale	替换 K-means++聚类 Replace K-means++ clustering				
YOLOv5s	—	—	—	—	7 103 997	14.02	81.19	181
YOLOv5s	✓	—	—	—	7 115 587	14.13	82.12	185
YOLOv5s	✓	✓	—	—	7 098 934	14.17	83.23	193
YOLOv5s	✓	✓	✓	—	5 287 521	10.65	83.34	188
SC-YOLOv5s	✓	✓	✓	✓	5 287 521	10.62	83.97	188

2.1.2 SC-YOLOv5s-lite 消融试验

为验证 SC-YOLOv5s 轻量化设计的有效性，进行 SC-YOLOv5s-lite 的消融试验，以 $3\ 000 \times 4\ 000$ 分辨率的图片作为模型输入，进行检测时间和平均精度测试，结果详见表 4。替换为 Fire module 结构后的模型相比于 SC-YOLOv5s，模型的平均精度下降了 0.76%，模型大小下降了 26.18%，参数量下降了 26.21%，检测时间减少了 21.28%；继续减少 Bottleneck 模块数量后的模型相比于 SC-YOLOv5s，其平均精度提高了 0.54%，模型大小下降了 27.21%，参数量减少了 26.88%；检测时间减少了 23.94%。综合对比 YOLOv5s 可知，SC-YOLOv5s-lite 模型的参数量降低了 45.57%，模型大小降低了 44.86%，平均精度提高了 3.98%，检测时间减少了 20.99%，说明改进后的模型参数量更少，检测速度更快，精度更高，更加有利于模型工程化应用，部署到端侧识别时模型会有更优的性能。

表 4 SC-YOLOv5s-lite 消融试验
Table 4 SC-YOLOv5s lite ablation test

模型 Model	轻量化策略 Lightweight strategy			参数量 Parameter quantity	模型大小/M Model memory	平均精度/% mAP	检测时间/ms Detection time
	Fire module 结构替换 3×3 卷积层	减少 Bottleneck 模块数量	Fire module structure replace 3×3 convolutional layers				
YOLOv5s	—	—	—	7 103 997	14.02	81.19	181
SC-YOLOv5s	—	—	—	5 287 521	10.62	83.97	188
SC-YOLOv5s	✓	—	—	3 901 521	7.84	83.33	148
SC-YOLOv5s-lite	✓	✓	✓	3 866 032	7.73	84.42	143

2.2 识别模型的性能比较

YOLOv5n 和 YOLOv5-lite 都是 YOLOv5 的轻量级改进版本，是适合移植到嵌入式设备的小型模型，为测试改进的模型相较于其他轻量化模型的检测性能，试验特选取了 YOLOv5s、YOLOv5n、YOLOv5-lite 等 3 种轻量化目标检测模型进行横向比较。

由表 5 可知, 在结合番茄果实识别实际需求优化后的 SC-YOLOv5s-lite, 其模型大小为 7.73 M, 仅大于 YOLOv5n 模型, 但其平均精度比 YOLOv5n 高 6.44%, 检测时间比 YOLOv5n 长 9.16%, 在检测速度和模型大小上稍劣于 YOLOv5n, 但平均精度高于 YOLOv5n。对比 YOLOv5s 和 YOLOv5s-lite 发现, 改进模型在模型大小、平均精度、检测时间上都要优于它们。从总体上看, 改进的模型相对于主流的目标检测算法如 YOLOv5s 来说, 在模型大小、检测准确率和检测速度三方面达到了一个平衡, 因此更加适合应用于嵌入式设备上。

表 5 不同网络模型性能参数对比
Table 5 Comparison of performance parameters of different network models

模型 Model	参数量 Parameter quantity	模型大小/M Model memory	平均精度/% mAP	检测时间/ms Detection time
YOLOv5n	2 754 122	3.87	79.31	131
YOLOv5s	7 103 997	14.12	81.23	181
YOLOv5-lite	4 368 521	10.98	80.86	162
SC-YOLOv5s-lite	3 866 032	7.73	84.42	143

2.3 SC-YOLOv5s-lite 模型训练测试

2.3.1 SC-YOLOv5s-lite 量化结果

优化后的 SC-YOLOv5s-lite 模型通过训练使其在统一的测试集上进行测试。从表 6 可以看出, SC-YOLOv5s-lite 模型对优等成熟果、劣等成熟果、未熟果 3 个类别的识别效果优秀, 精确率达到 92% 以上, F_1 值均达到 90% 以上; 对二等成熟果这一类别的检测中, 精确率和召回率都偏低, 误检较多, F_1 值也略低, 分析原因可能是二等成熟果的检测特征不明显, 而且在标注期间对二等成熟果的标注标准不稳定, 导致容易被误检为劣等成熟果和半熟果。SC-YOLOv5s-lite 模型整体的准确率达到了 89.04%, 召回率 83.35%, F_1 值为 86.98%, 把 iou 阈值设为 0.5 时, 平均精度达到了 91.34%, 不同 iou 阈值 (0.5 ~ 0.95, 步长 0.05) 上的平均精度为 84.42%。

表 6 SC-YOLOv5s-lite 算法下不同成熟度和外观品质果实量化结果
Table 6 Quantitative results of fruits with different maturity and appearance quality using the SC-YOLOv5s lite algorithm

成熟度类别 Maturity classification	外观品质分级 Appearance quality grading	召回率/%	精确率/%	F_1 值/%	准确率/%	平均精度/% mAP
		Recall	AP	F_1 -score	Accuracy	mAP @ 0.5
成熟果 Ripe fruit	优等果 Superior fruit	89.52	94.88	92.12		
	二等果 Second-class fruit	64.39	85.72	73.53		
	劣等果 Inferior fruit	93.13	95.07	94.09		
半熟果 Semi-ripe fruit	—	81.48	88.93	85.04		
	未熟果 Unripe fruit	—	88.24	92.08	90.12	
整体 Total	—	83.35	—	86.98	89.04	91.34
						84.42

2.3.2 SC-YOLOv5s-lite 识别可视化分析

为了测试 SC-YOLOv5s-lite 模型对数据集 5 种类别的实际识别情况, 把测试集中 532 张图像放到模型中进行检验。从测试效果看 (图 9), 每张图片中番茄的成熟度都能得到正确的检测, 成熟番茄的外观品质也能得到有效区分。

针对枝叶遮挡, 光照明暗度, 果实叠加, 多类混搭这几种复杂场景下 SC-YOLOv5s-lite 模型的识别能力进行验证。结果 (图 10) 表明, 在枝叶遮挡和果实叠加情况下, SC-YOLOv5s-lite 模型识别准确率有所降低, 但是在不同光照明暗度和多种因素混搭情况下的准确率不受影响, 表明在数据

集中增加不同明暗度和多类别混搭的数据有利于保证模型的在这种场景下识别的稳定性。



图 9 番茄果实成熟度和外观品质检测效果图

a ~ c: 成熟优等果; d ~ f: 成熟二等果; g ~ i: 成熟劣等果; j ~ l: 半熟果; m ~ o: 未熟果。

Fig. 9 Effect diagram of tomato fruit detection with different maturity and appearance quality

a ~ c: Mature and high-quality fruit; d ~ f: Mature and second class fruit; g ~ i: Mature and inferior fruit;

j ~ l: Semi-ripe fruit; m ~ o: Unripe fruit.



图 10 复杂环境下 SC-YOLOv5s-lite 预测验效果图

a、b: 枝叶遮挡环境; c、d: 不同光照明暗度; e、f: 果实叠加环境; g、h: 多类因素混搭环境。

Fig. 10 SC-YOLOv5s lite prediction validation effect in complex environments

a, b: The environment obstructed by branches and leaves; c, d: The darkness of different lighting conditions;
e, f: The overlapping environment of fruits; g, h: A mixed environment of multiple factors.

2.3.3 SC-YOLOv5s-lite 田间检测情况

为了验证 SC-YOLOv5s-lite 模型检测番茄成熟度及外观品质的有效性, 在烟台农业科学研究院温室进行现场试验。图 11 分别为暗光、顺光、单目标、多目标场景下的检测结果。暗光场景(图

11, a) 下劣等成熟果被准确检测，远距离的两个未熟果也被识别到。多目标场景（图 11, b）下 4 个未熟果被成功检测，被遮挡的一个半熟果也被检测到。顺光场景（图 11, c）下两个半熟果被检测到，两个未熟果也被检测到，两个绿色枝叶被误检为半熟果。单目标场景（图 11, d）下准确识别到图中一个优等成熟果。在使用树莓派（微型电脑）进行实际测试中发现识别的平均 FPS 在 3~4 帧之间，画面较为流畅。可知经过改进的 SC-YOLOv5s-lite 提高了番茄成熟度及外观品质检测识别的速度。将不同场景下识别结果与真实类别做对比，发现番茄成熟度及外观品质检测整体识别率达到了 89.04%，在番茄温室大棚识别时适应性良好。

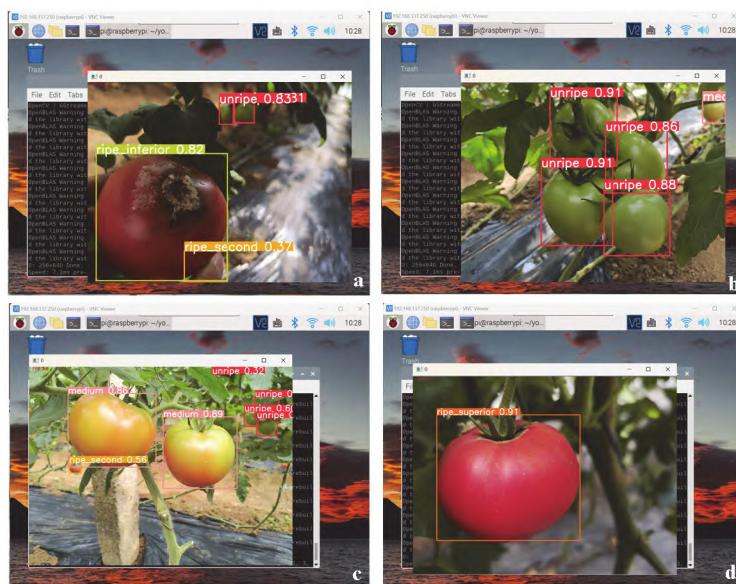


图 11 田间番茄不同场景实时检测效果图

Fig. 11 Real time detection effect diagram of tomatoes in different scenes in the field

2.3.4 SC-YOLOv5s-lite 混淆矩阵和热力图分析

SC-YOLOv5s-lite 模型训练后的混淆矩阵和热力图见图 12 和图 13。从图 12 可以看出，实际为未熟果的样例中有 92% 预测正确，有 1% 被误检为半熟果，漏检率 7%；实际为半熟果的样例中有 88% 预测正确，有 2% 被误检为未熟果，有 8% 被误检为二等成熟果，有 2% 被漏检；实际为优等成熟果的样例中有 94% 预测正确，有 1% 被误检为半熟果，有 4% 被误检为二等成熟果，有 1% 被误检为劣等成熟果；实际为二等成熟果的样例中有 85% 预测正确，有 5% 被误检为半熟果，有 4% 被误检为优等成熟果，有 3% 被误检为劣等成熟果；实际为劣等成熟果的样例中有 95% 预测正确，有 1% 被误检为优等成熟果，有 4% 被误检为二等成熟果。试验发现半熟果和二等成熟果误检率较多，原因可能是因为半熟果表面存在绿色区域和红色区域，与未熟果、二等成熟果的表面特征存在一定的相似性，容易被误检。二等成熟果的标准因为难以制定和严格执行，受到人为主观因素影响较大，所以在数据人工标注阶段对二等成熟果的标注标准波动较大，导致其他类别的果实也易被识别为二等成熟果。未熟果因为表面特征单一且较为明显，所以误检率较低；优等成熟果表面光滑红润，容易识别，所以误检率较低；劣等成熟果表面多存在偏暗区域，与表面红色区域形成对比，特征也较为明显，所以识别率较高。同时通过对真实背景分析后，可以发现有 47% 的背景被预测为了未熟果，因为背景中番茄枝叶呈现绿色，与未熟番茄果实的绿色完全一致，所以背景中的枝叶极易被虚检成未熟果。

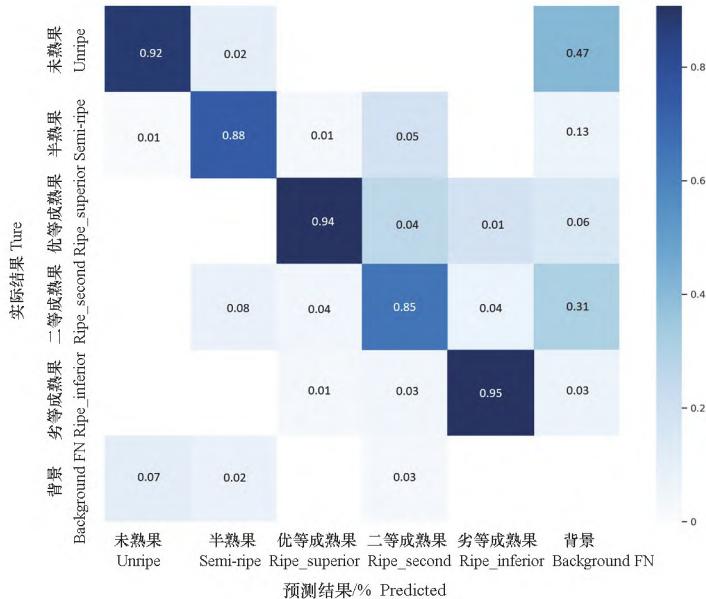


图 12 SC-YOLOv5s-lite 模型的混淆矩阵
Fig. 12 Confusion matrix of the SC-YOLOv5s-lite model

由图 13, a 中可以看出, 未熟果标签的数量最多, 二等成熟果标签的数量最少, 所以增加二等成熟果标签的数据量可以进一步提升这一类别的精确率。通过热力图 (图 13, b ~ d) 信息, 可以发

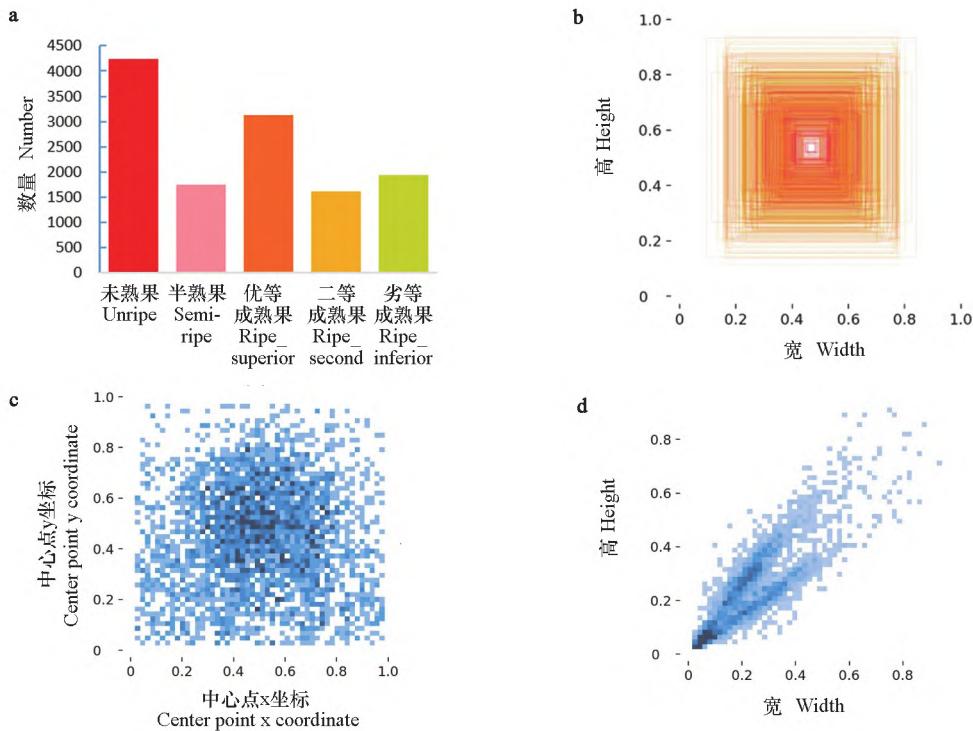


图 13 SC-YOLOv5s-lite 模型的热力图

a: 不同类别的标签数; b: 目标框在图片中的相对位置; c: 目标框中心点坐标在图片中占比; d: 目标框长和宽在图片中占比。

Fig. 13 Thermal diagram of the SC-YOLOv5s-lite model

a: The number of labels in different categories; b: The relative position of the target box in the image; c: The proportion of the center point coordinates of the target box in the image; d: The proportion of the length and width of the target box in the image.

现模型识别的番茄目标整体在图片中的相对位置，目标框中心点集中在整幅图片的中心位置，长和宽占比整幅图的长和宽为 22%~40%之间。

3 讨论

目前目标检测方法主要分为传统的机器学习算法和基于深度学习的算法。传统机器学习算法严重依赖于手工特征设计、难度高、鲁棒性差 (Kang, 2022)，一般只适用于某一特定果实的识别，而实际识别中会遇到果蔬种类多、品种杂、果实的形状各异、背景复杂等问题，无法做到识别场景的通用性，也无法对果实特征进行准确的提取；基于深度学习的果实识别方法可以有效地解决传统方法对果实识别的局限性。利用卷积神经网络可以全面自主学习果实图像特征 (宋玲 等, 2023)，只需要采集大量的果实图像对卷积神经网络进行训练，使网络自主学习图像中的果实特征，避免了复杂繁重的人工特征设计过程，使模型训练具有了更好的操作性。目前番茄果实目标检测研究中仍存在一些问题。如目前番茄成熟度检测主要针对成熟果实 (刘芳 等, 2020; 岳有军 等 2021; 张磊 等, 2022)，缺乏同步的外观品质分级检测，后续还需要对果实外观品质进行再次分级；而且番茄植株枝叶比较茂盛，并且是一穗多果，实际检测时容易受到果实重叠、枝叶遮挡等因素的干扰，影响识别准确率和识别效果 (冯俊惠 等, 2021; 郗厚印 等, 2021)；深度学习检测模型的参数量和运算量都很大，将模型部署到性能受限的嵌入式设备上会存在检测速度过慢的问题 (祖林禄, 2023)。

本试验基于深度学习方法在番茄成熟度检测研究的基础上增加了对成熟果实外观品质分级的检测。采集了番茄图片数据 2 036 张，通过对随机图像缩放和旋转，翻转，随机剪裁或补零，镜像平移，对比度变化，亮度变化，添加不同范围和大小的高斯噪声等处理，将图片数据扩增到 5 316 张，然后将数据进行标注和文件转换，构建涵盖了光照明暗度、果实叠加度、枝叶遮挡度、背景环境、成熟度等多种因素，以及多类因素混搭的试验用数据集；采用一阶段目标检测算法对番茄成熟度及外观品质进行检测。针对 YOLOv5s 模型部署在树莓派上存在检测精度较低和检测速度过慢的问题，试验通过在 YOLOv5s 中加入 CA 注意力机制、替换 Stem block 结构、结合识别需求优化检测层尺度、替换 K-means++聚类算法来实现 SC-YOLOv5s 精度提升，提高模型的特征表达能力；通过在 SC-YOLOv5s 中加入 Fire module 结构进行轻量化卷积、降低 Bottleneck 模块的参数量来实现 SC-YOLOv5s-lite 轻量化设计，提升模型在树莓派上的检测速度；最后将 SC-YOLOv5s-lite 模型在训练集上进行训练优化、消融实验和性能对比，以精确率、召回率、平均精度作为评价指标，结果表明：SC-YOLOv5s-lite 模型内存占用为 7.73 M，其准确率为 89.04%，召回率 83.35%，平均精度为 91.34%，检测时间为 143 ms，相比于 YOLOv5s，模型参数量降低了 45.57%，模型大小压缩了 44.86%，平均精度提升 3.98%，检测时间减少 20.99%，优势明显，更适合于硬件上部署。

References

- Chen Qiao-song, Chen Peng-chang, Li Pei, Zhang Ya-ling, Deng Xin, Sun Kai-wei, Wang Jin. 2023. Multi-scale feature selection and fusion for object detection. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 35 (2): 227 – 234. (in Chinese)
陈乔松, 陈鹏昌, 李 佩, 张亚玲, 邓 欣, 孙开伟, 王 进. 2023. 基于多尺度特征选择与融合的目标检测方法. 重庆邮电大学学报 (自然科学版), 35 (2): 227 – 234.
- Du Ming-yang, Zhang Tian-zhu. 2022. Beautiful unmanned farms in the digital landscape of agriculture. Vegetables, (9): 1 – 9. (in Chinese)
杜名扬, 张天柱. 2022. 农业数字化风景下的靓点无人农场. 蔬菜, (9): 1 – 9.
- Feng Fan. 2022. The application of internet of things technology in modern agriculture under the background of smart agriculture. The Farmers

- Consultant, (21): 25 – 27. (in Chinese)
- 冯 帆. 2022. 智慧农业背景下物联网技术在现代农业中的应用. 农家参谋, (21): 25 – 27.
- Feng Jun-hui, Li Zhi-wei, Rong You-li, Sun Zhi-li. 2021. Identification of mature tomatoes based on an algorithm of modified circular hough transform. Journal of Chinese Agricultural Mechanization, 42 (4): 190 – 196. (in Chinese)
- 冯俊惠, 李志伟, 戎有丽, 孙志立. 2021. 基于改进 Hough 圆变换算法的成熟番茄果实识别. 中国农机化学报, 42 (4): 190 – 196.
- Gao Yan-zhao, Tao Chang-yong. 2023. Hardware-accelerated consistent computing structure for signal processing and deep learning. Journal of National University of Defense Technology, 45 (2): 112 – 120. (in Chinese)
- 高彦钊, 陶常勇. 2023. 信号处理与深度学习硬件加速的一致性计算结构. 国防科技大学学报, 45 (2): 112 – 120.
- He Jie, Ma Ting-juan, Zhang Yong-liang, Zhang Ya-juan. 2022. Construction of evaluation model and index system of blended learning in higher vocational education aiming at deep learning. Vocational and Technical Education, 43 (26): 64 – 70. (in Chinese)
- 贺 杰, 马婷娟, 张永良, 张亚娟. 2022. 指向深度学习的高职混合式学习评价模型及指标体系构建. 职业技术教育, 43 (26): 64 – 70.
- Huang Bo. 2022. Countermeasures and analysis of talent loss in the background of rural revitalization. Practical Rural Technology, (4): 46 – 48. (in Chinese)
- 黄 波. 2022. 乡村振兴背景下的人才流失问题对策与分析. 农村实用技术, (4): 46 – 48.
- Huang Tong-bin, Huang He-qing, Li Zhen, Lü Shi-lei, Xue Xiu-yun, Dai Qiu-fang, Wen Wei. 2022. Citrus fruit recognition method based on the improved model of YOLOv5. Journal of Huazhong Agricultural University, 41 (4): 170 – 177. (in Chinese)
- 黄彤镔, 黄河清, 李 震, 吕石磊, 薛秀云, 代秋芳, 温 威. 2022. 基于 YOLOv5 改进模型的柑橘果实识别方法. 华中农业大学学报, 41 (4): 170 – 177.
- Kang S. 2022. Research on intelligent video detection of small targets based on deep learning intelligent algorithm. Computational Intelligence and Neuroscience, 16(5): 67 – 69.
- Li Yan-zhou, He Yan-zhou, Qin Feng, Qian Wan-qiang, Wu Mei, Qiao Xi. 2023. Identification of spartina alterniflora based on convolutional neural network. Journal of Chinese Agricultural Mechanization, 44 (4): 159 – 166. (in Chinese)
- 李岩舟, 何艳洲, 覃 锋, 钱万强, 吴 媚, 乔 曦. 2023. 基于卷积神经网络的互花米草识别研究. 中国农机化学报, 44 (4): 159 – 166.
- Liu Fang, Liu Yu-kun, Lin Sen, Guo Wen-zhong, Xu Fan, Zhang Bai. 2020. Fast recognition method for tomatoes under complex environments based on improved YOLO. Transactions of the Chinese Society for Agricultural Machinery, 51 (6): 229 – 237. (in Chinese)
- 刘 芳, 刘玉坤, 林 森, 郭文忠, 徐 凡, 张 白. 2020. 基于改进型 YOLO 的复杂环境下番茄果实快速识别方法. 农业机械学报, 51 (6): 229 – 237.
- Liu Zheng, Wei Chuan-cheng. 2021. Key points, difficulties and countermeasures for promoting the comprehensive revitalization of rural areas. Economic Review Journal, (10): 122 – 128. (in Chinese)
- 刘 铮, 魏传成. 2021. 推进乡村全面振兴的重点、难点及对策. 经济纵横, (10): 122 – 128.
- Luo Xi-wen, Liao Juan, Zang Ying, Ou Ying-gang, Wang Pei. 2022. Developing from mechanized to smart agricultural production in China. Strategic Study of CAE, 24 (1): 46 – 54. (in Chinese)
- 罗锡文, 廖 娟, 臧 英, 区颖刚, 汪 浩. 2022. 我国农业生产的发展方向: 从机械化到智慧化. 中国工程科学, 24 (1): 46 – 54.
- Ma Jin-lin, Zhang Yu, Ma Zi-ping, Mao Kai-ji. 2022. Research progress of lightweight neural network convolution design. Journal of Frontiers of Computer Science and Technology, 16 (3): 512 – 528. (in Chinese)
- 马金林, 张 裕, 马自萍, 毛凯绩. 2022. 轻量化神经网络卷积设计研究进展. 计算机科学与探索, 16 (3): 512 – 528.
- Song Ling, Cao Mian, Hu Xiao-chun, Jia Pei-yuan, Chen Yan, Chen Ning-jiang. 2023. Detection of cassava leaf diseases under complicated background based on YOLOX. Transactions of the Chinese Society for Agricultural Machinery, 54 (3): 301 – 307. (in Chinese)
- 宋 玲, 曹 勉, 胡小春, 贾沛沅, 陈 燕, 陈宁江. 2023. 基于 YOLOX 的复杂背景下木薯叶病害检测方法. 农业机械学报, 54 (3): 301 – 307.
- Sun Feng-gang, Wang Yun-lu, Lan Peng, Zhang Xu-dong, Chen Xiu-de, Wang Zhi-jun. 2022. Identification of apple fruit diseases using improved YOLOv5s and transfer learning. Transactions of the Chinese Society of Agricultural Engineering, 38 (11): 171 – 179. (in Chinese)
- 孙丰刚, 王云露, 兰 鹏, 张旭东, 陈修德, 王志军. 2022. 基于改进 YOLOv5s 和迁移学习的苹果果实病害识别方法. 农业工程学报, 38 (11): 171 – 179.
- Wang Hai-qun, Zhang Cheng-jun, Zhang Yi. 2023. A lightweight flame detection algorithm based on YOLOv4. Journal of Shandong University of

- Science and Technology(Natural Science), 42 (1): 91 – 99. (in Chinese)
- 王海群, 张成君, 张 怡. 2023. 基于 YOLOv4 的轻量级火焰检测算法. 山东科技大学学报(自然科学版), 42 (1): 91 – 99.
- Wang J J, Feng C, Wang L Y, Li G L, He B. 2022. Detection of weak and small targets in forward-looking sonar image using multi-branch shuttle neural network. IEEE Sensors Journal, 22 (7): 6772 – 6783.
- Wang Liang, Zhang Chao. 2023. A lightweight pedestrian detection method based on YOLOv5. Industrial Control Computer, 36 (4): 84 – 86, 89. (in Chinese)
- 王 亮, 张 超. 2023. 一种基于 YOLOv5 的轻量型行人检测方法. 工业控制计算机, 36 (4): 84 – 86, 89.
- Wu Di, Li Ming-hui, Ma Wen-kai, Li Rui-tong, Li Yan. 2023. Surface defect detection of steel based on improved Yolov5. Journal of Shaanxi University of Science & Technology, 41 (2): 162 – 169. (in Chinese)
- 吴 敌, 李明辉, 马文凯, 李睿童, 李 艳. 2023. 基于改进 YOLOv5 的钢材表面缺陷检测. 陕西科技大学学报, 41 (2): 162 – 169.
- Xi Hou-yin, Zhang Dong, Zhou Tao, Yang Yun-xiao. 2021. Research on the method of picking up overlapping tomato fruits by picking robot. Journal of Agricultural Mechanization Research, 43 (12): 17 – 23, 50. (in Chinese)
- 郗厚印, 张 栋, 周 涛, 杨云霄. 2021. 采摘机器人识别抓取重叠番茄果实的方法研究. 农机化研究, 43 (12): 17 – 23, 50.
- Xiong Jun-tao, Zheng Zhen-hui, Liang Jia-en, Zhong Zhuo, Liu Bo-lin, Sun Bao-xia. 2020. Citrus detection method in night environment based on improved YOLO v3 network. Transactions of the Chinese Society for Agricultural Machinery, 51 (4): 199 – 206. (in Chinese)
- 熊俊涛, 郑镇辉, 梁嘉恩, 钟 灼, 刘柏林, 孙宝霞. 2020. 基于改进 YOLO v3 网络的夜间环境柑橘识别方法. 农业机械学报, 51 (4): 199 – 206.
- Yang Qian-shan, Zhou De-bing, Tu Wan-ying, Liu Du-yang. 2023. Concrete crack detection research based on offline data enhancement and YOLOv5. Shanxi Architecture, 49 (8): 63 – 66. (in Chinese)
- 阳茜汕, 卓德兵, 涂婉莹, 刘杜洋. 2023. 基于离线数据增强与 YOLOv5 的混凝土裂缝检测研究. 山西建筑, 49 (8): 63 – 66.
- Yue You-jun, Sun Bi-yu, Wang Hong-jun, Zhao Hui. 2021. Object detection of tomato fruit based on cascade RCNN. Science Technology and Engineering, 21 (6): 2387 – 2391. (in Chinese)
- 岳有军, 孙碧玉, 王红君, 赵 辉. 2021. 基于级联卷积神经网络的番茄果实目标检测. 科学技术与工程, 21 (6): 2387 – 2391.
- Zha Wen-wen, Pan Wei-hao, Chen Cheng-peng, Chen Wei, Gu Li-chuan, Jiao Jun. 2023. Study on pose recognition and tracking of herd-raised pigs based on edge computing and improved YOLOv5. Journal of Northeast Agricultural University, 54 (3): 83 – 96. (in Chinese)
- 查文文, 潘伟豪, 陈成鹏, 陈 威, 姜丽川, 焦 俊. 2023. 基于边缘计算与改进 YOLOv5 的群养生猪姿态识别及跟踪研究. 东北农业大学学报, 54 (3): 83 – 96.
- Zhang Huai-feng, He Xiang-jian, Wu Qiang. 2006. Generic object detection: a survey. Journal of Yunnan Minzu University(Natural Sciences Edition), 15 (4): 261 – 267. (in Chinese)
- 张淮峰, 何祥健, 吴 强. 2006. 通用目标检测算法研究进展与评述. 云南民族大学学报(自然科学版), 15 (4): 261 – 267.
- Zhang Hui-feng, Xing Yu-tong, Zhang Miao, Li Wen-xu, Li Cong. 2022. Comparison of epiphyseal localization accuracy by different deep learning models. Chinese Journal of Forensic Medicine, 37 (3): 267 – 271. (in Chinese)
- 张会丰, 邢雨彤, 张 淦, 李文旭, 李 聪. 2022. 不同深度学习模型的骨骼定位准确率比较. 中国法医学杂志, 37 (3): 267 – 271.
- Zhang Lei, Liu Qi-fang, Nie Hong-mei, Wang Chen, Niu Fan. 2022. Tomato fruit detection based on improved YOLOV4 network model. Journal of Chinese Agricultural Mechanization, 43 (12): 162 – 169. (in Chinese)
- 张 磊, 刘琪芳, 聂红玫, 王 晨, 牛 帆. 2022. 基于改进 YOLOV4 网络模型的番茄果实检测. 中国农机化学报, 43 (12): 162 – 169.
- Zhou Shen-he, Shao Wu-chang, Wang Qiang. 2023. Overview of small object detection algorithms based on deep learning. Digital Technology & Application, 41 (2): 156 – 158. (in Chinese)
- 周升和, 邵武长, 王 强. 2023. 基于深度学习的小目标检测算法综述 数字技术与应用, 41 (2): 156 – 158.
- Zhu De-li, Wen Rui, Xiong Jun-yi. 2023. Lightweight corn silk detection network incorporating with coordinate attention mechanism. Transactions of the Chinese Society of Agricultural Engineering, 39 (3): 145 – 153. (in Chinese)
- 朱德利, 文 瑞, 熊俊逸. 2023. 融合坐标注意力机制的轻量级玉米花丝检测. 农业工程学报, 39 (3): 145 – 153.
- Zu Linlu. 2023. Research on the data-driven model for predicting tomato growth in solar greenhouse[Ph. D. Dissertation]. Tai'an: Shandong Agricultural University. (in Chinese)
- 祖林禄. 2023. 数据驱动的日光温室番茄果实生长预测模型研究[博士论文]. 泰安: 山东农业大学.