doi:10.6041/j.issn.1000-1298.2026.02.019

基于颜色特征量化和改进 YOLO v8 的番茄成熟度分级检测方法

张领先[1]　周　沁[1]　姚天雨[1]　裴鑫达[1]　赵立群[2]　满　杰[2]　钱　井[2]
(1.中国农业大学信息与电气工程学院,北京 100083;2.北京市农业技术推广站,北京 100029)

摘要:番茄的成熟度与其品质密切相关,是生产中采摘和分拣等环节的重要依据。针对作物成熟度分级检测系统功能简单,人工升级系统成本较大的问题,本文以番茄为例,采集并构建自然场景下番茄图像数据集,设计以番茄果实成熟度分级算法为基础的番茄图像半自动标注算法对采集后的数据进行标注,在 YOLO v8 模型基础上,将 FPN 结构替换为 BiFPN 结构实现更高效的多尺度特征融合,利用 SE 注意力机制对空间和通道进行融合特征提取,引入 Focal SIoU 损失函数对预测框与真实框之间的角度差异进行度量,构建基于颜色特征量化和改进 YOLO v8 的番茄成熟度分级检测模型 YOLO v8_BFS,识别番茄生长过程的 5 个不同成熟度。试验结果表明,本文模型较好地解决了自然复杂场景下番茄成熟度分级检测的错漏检问题,在模型浮点运算量(FLOPs)、参数量(Params)和内存占用量有少量增加的条件下,本文模型的平均精度均值为 94.10%,相较原模型 YOLO v8 提高 3.0 个百分点。通过与 Faster R-CNN-Resnet50、YOLO v5、YOLO v7-tiny、YOLO v8、YOLO v10 和 YOLO 11 目标检测模型对比,本文在检测精度具有显著优势,为番茄成熟度的检测提供了一种可靠的方法。

关键词:番茄成熟度;自然场景;颜色特征量化;YOLO v8
中图分类号:TP391　　**文献标识码**:A　　**文章编号**:1000-1298(2026)02-0193-10　　**OSID**:

Tomato Ripeness Grading Detection Method Based on Color Feature Quantification and Improved YOLO v8

ZHANG Lingxian[1]　ZHOU Qin[1]　YAO Tianyu[1]　PEI Xinda[1]　ZHAO Liqun[2]　MAN Jie[2]　QIAN Jing[2]
(1. *College of Information and Electrical Engineering*, *China Agricultural University*, *Beijing* 100083, *China*
2. *Beijing Agricultural Technology Extension Station*, *Beijing* 100029, *China*)

Abstract: The ripeness of tomatoes is closely related to their quality, and it serves as a crucial basis for key production processes such as harvesting and sorting. To address the issues of simple functionality in crop ripeness grading and detection systems, and high costs associated with manual system upgrades, taking tomatoes as an example. It collected and constructed a tomato image dataset under natural scenarios, and a semi-automatic tomato image annotation algorithm was designed based on the tomato fruit ripeness grading algorithm to annotate the collected data. Building on the YOLO v8 model, the FPN structure was replaced with the BiFPN structure to achieve more efficient multi-scale feature fusion. It utilized the SE attention mechanism for fused feature extraction across spatial and channel dimensions, and introduced the Focal SIoU loss function to measure the angular difference between the predicted bounding box and the ground truth box. This results in the development of the tomato ripeness grading and detection model YOLO v8_BFS which was based on color feature quantization and the improved YOLO v8, and can identify five different ripeness stages during tomato growth. Experimental results showed that the proposed model effectively solved the problems of false detection and missed detection in tomato ripeness grading and detection under complex natural scenarios. While there was a slight increase in model computational complexity (FLOPs), parameter count (Params), and memory storage size, the detection accuracy of the proposed model reached 94.10%, which was 3.0 percentage points higher than that of the original YOLO v8 model. Compared with target detection models such as Faster R-CNN-Resnet50, YOLO v5, YOLO v7-tiny, YOLO v8, YOLO v10, and YOLO 11, the proposed model

收稿日期:2024-10-26　修回日期:2024-12-03
基金项目:国家自然科学基金项目(62176261)
作者简介:张领先(1970—),男,教授,博士生导师,主要从事农业信息化研究,E-mail:zhanglx@cau.edu.cn

demonstrated significant advantages in detection accuracy, providing a reliable method for tomato ripeness detection.

Key words: tomato ripeness; natural scene; color feature quantification; YOLO v8

0 引言

番茄是三大世界性贸易蔬菜之一,具有较高的营养价值和经济价值[1]。番茄果实的成熟度很大程度上影响着番茄果实的品质与口感[2],传统粗放型的生产经营方式存在着机械化、智能化较为落后的问题,生产作业中的各个环节依然高度依赖人工[3-7],一定程度上阻碍着果蔬产业未来的高质量发展。此外,传统的图像处理技术难以处理自然复杂场景下的识别任务,仅局限于室内简单场景,且存在着步骤复杂、精确率低等问题。

自然场景下存在着枝叶等遮挡物,多果实之间也可能存在遮挡情况,此外光照等外部条件也会随天气而变化,从而对果实目标的识别、分级带来了较大困难。文献[8]提出了自然场景下成熟番茄的分割方法,基于 YUV 颜色模型,确定红熟番茄的颜色阈值范围,使用双阈值分割算法对成熟番茄进行识别、分割。文献[9]提出基于改进 Hough 圆变换算法的成熟番茄果实识别方法,首先对番茄图像进行下采样,通过颜色特征进行背景与果实的分割后,依次采用 Sobel 算子、改进 Hough 圆变换算法进行边缘检测和番茄果实识别,该算法对于简单背景的果实可以准确识别,但处理茎叶遮挡、果实重叠等复杂情况时,准确率有较大幅度下降,综合准确率为 86.3%。近几年,深度学习技术快速发展,广泛应用于多个领域,并且在农业领域上也已经显示出了巨大的优势和潜力,使得自然复杂场景下快速、精准地识别果蔬成为可能。文献[10]基于深度卷积神经网络构建了柑橘目标识别模型,通过深度卷积神经网络提取柑橘在自然环境下的视觉特征、深层池化结构提取高层语义特征,采用非极大值抑制算法进行预测,并基于迁移学习,最终识别平均准确率为 86.6%。文献[11]提出了一种夜间自然环境下的番茄果实检测方法,通过 CIoU 损失函数和自适应锚定框对 YOLO v5 网络模型进行改进,实现了对夜间自然环境下番茄绿色果实红色果实的快速识别,综合平均识别精度为 96.2%。文献[12]基于 YOLO v5 首次开展白芦笋的检测,通过 Light C3、DW Conv 等结构对 YOLO v5 进行改进,实现了快速、精准的白芦笋检测,并将模型部署在嵌入式设备上为收割机器人提供定位功能。

本文基于图像处理和深度学习技术开展自然场景下番茄成熟度分级检测研究,构建基于 YOLO v8 的番茄成熟度分级检测模型。通过将 FPN[13]结构替换为 BiFPN[14]结构,利用 SE(Squeeze-and-excitation)注意力机制[15]对空间和通道进行融合特征提取,引入 Focal SIoU 损失函数对预测框与真实框之间的角度差异进行度量,改进模型为 YOLO v8_BSF,通过检测结果可视化对比,以验证 YOLO v8_BSF 模型解决复杂背景的能力。

1 材料与方法

1.1 番茄图像采集与预处理

口感型番茄是近年来为满足消费者高品质需求而出现的新型番茄品种,具备皮薄、肉厚、多汁、风味浓郁等特点,深受市场欢迎,且价格和经济效益都高于菜用型和加工型番茄[16]。所用番茄图像数据于 2023 年 5 月上旬分别采集于北京市朝阳区的连栋温室大棚(440 幅)和北京市密云区的日光温室大棚(439 幅),所用采集设备为小米 13 Pro 型手机。采集时考虑了自然条件下多种复杂场景,包括不同光照强度、不同果实数量、不同遮挡情况等,共采集 879 幅图像,总计 2 818 枚番茄果实,其中绿熟期 1 192 枚,变色期 408 枚,红熟前期 528 枚,红熟中期 210 枚,红熟后期 480 枚。为进一步加强数据集的丰富度,通过对单幅图像进行几何变换和色彩变换,例如旋转、缩放、平移、亮度变换、对比度变换等,增加数据集的多样性。此外,采用 Mosaic 对多幅图像进行组合变换,更大程度地减少原图像和增强图像的相似性[17]。

1.2 颜色特征量化的番茄成熟度分级标准

选取中华人民共和国供销合作行业标准 GH/T 1193—2021[18]作为分级参考依据。该标准将番茄果实成熟度划分为 6 个等级:未熟期、绿熟期、变色期、红熟前期、红熟中期、红熟后期,番茄成熟度划分依据如表 1 所示,其中未熟期番茄果实尚未发育定形、不宜采摘贮藏,不在本研究范围内,本文主要围绕其余 5 个等级进行研究。

随着番茄果实成熟度提高,果实内叶绿素减少至消失,而以番茄红素为代表的类胡萝卜素含量逐渐增加,从而显示出番茄果实表皮及果肉由绿至红的成熟过程[19-20]。从供销合作行业标准也可以看出,对于各成熟度等级之间的划分也主要围绕着

表 1　番茄不同成熟度果实形态特征

Tab. 1　Morphological characteristics of tomatoes at different maturity stages

成熟度	形态特征
未熟期	果实及种子尚未充分生长发育定形
绿熟期	果实定形，果面有光泽，由绿色变为白绿色
变色期	果脐周围开始出现淡红或黄色晕斑，果实着红面不到 10%
红熟前期	果实着红面 10% ~30%
红熟中期	果实着红面 40% ~60%
红熟后期	果实着红面 70% ~100%

果实着红面这一颜色特征。因此，将基于颜色特征量化，探索适用性更强的番茄果实成熟度分级标准。

在 HSV 颜色模型中，H(Hue)通道即色调，表示人感官对颜色的感受；S(Saturation)通道即饱和度，表示颜色深浅程度；V(Value)通道即亮度，表示色彩明亮程度。HSV 颜色模型相对于 RGB 颜色模型能够更直观地表达颜色明暗程度、鲜艳程度以及色调，更容易排除复杂自然场景下不同光照强度对番茄果实颜色特征的影响[22-25]，图 1 为同一番茄原图像不同亮度在 HSV 颜色模型下的效果，在番茄原图像基础上，增加和减少亮度的参数值模拟现实场景中自然环境和采集设备不同所带来的亮度差异，由图 1 可以看出，同一番茄图像在不同亮度下，V 通道差别最为明显，在 S 通道上略有差别，而在 H 通道上几乎没有差别，即像素的整体亮度由 V 通道集中确定，H 通道受亮度影响很小。

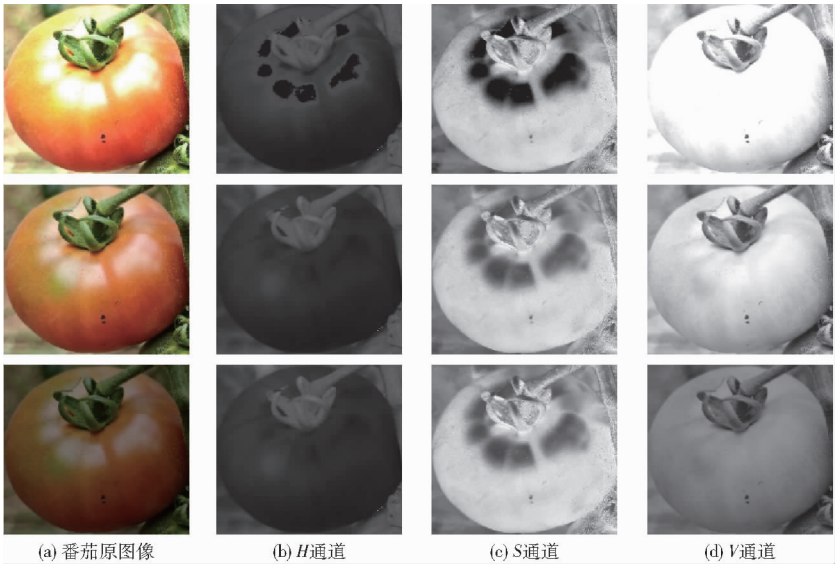

图 1　同一番茄图像不同亮度在 HSV 颜色模型下的效果

Fig. 1　Effect of the same tomato image across various brightness levels within HSV color

图 2 为不同成熟度番茄果实图像和其颜色分布直方图(HSV－H)。从图 2 可以看出，随着番茄果实成熟度增加，颜色分布直方图波峰也随之左移。其中图 2 第 2 行为全区间[0°，180°]图示，第 3 行为[0°，45°]区间图示，显然，番茄果实图像 H 通道分量与其成熟度有着较强的相关性和规律性。因为 H 通道受亮度的影响很小，图 2 颜色分布直方图 H 通道分量值左移的主要原因是番茄成熟度的增加，而非亮度影响，因此，以此为依据制定量化分级标准是可行的。基于供销合作行业标准，并结合番茄果实图像与 H 通道颜色分布直方图的对比分析工作，得到具体的番茄果实成熟度划分标准如表 2 所示。

1.3　颜色特征量化的番茄图像半自动标注算法

人工标注的过程存在着许多问题，例如设备和环境等客观因素影响标注、经验和注意力集中度等主观因素影响标注等，本文提出番茄图像的半自动标注算法以解决上述问题。具体流程主要分为 3 部分：①获取自然场景原始图像中番茄果实坐标信息。②将果实图像从原始图像中提取出来，并对果实图像进行颜色特征量化与自动分级处理。③提供番茄果实图像、自动分级结果、各成熟度的像素面积占比等信息辅助人工进行研判和校正，形成最后的番茄果实坐标和成熟度信息。

基于番茄成熟度分级标准所设计的番茄成熟度自动分级算法步骤为：

(1)对于遮挡少、背景简单、颜色均匀的番茄图像，判断标准为该图像中除绿熟期外的某一成熟度等级颜色范围所占的像素面积超过 45%，并进一步以图像中该成熟度等级颜色范围的像素面积为基础，如果后一个成熟度等级颜色范围的像素面积超过此等级的 20%，则分级为后一个等级，否则分级为该等级。

(2)由于绿熟期存在和枝叶等颜色相近的因素

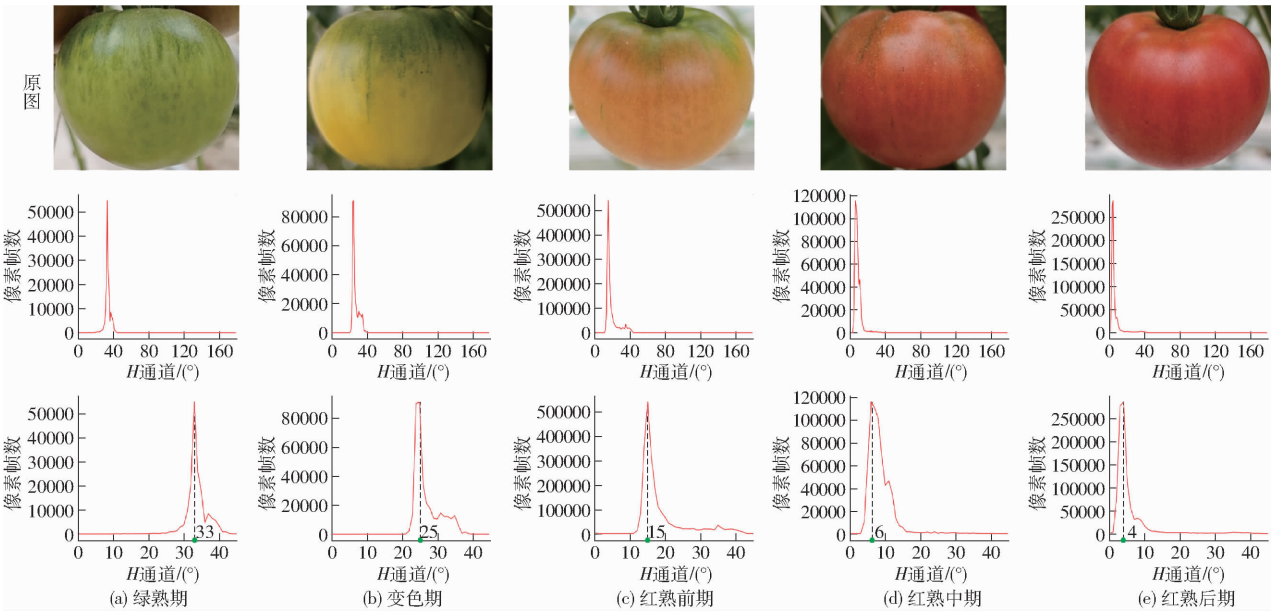

图 2 不同成熟度的番茄果实图像和其颜色分布直方图(HSV－H)

Fig. 2 Images of tomatoes at different ripeness stages and their color distribution histograms (HSV－H)

表 2 颜色特征量化的番茄果实成熟度划分标准

Tab. 2 Tomato fruit maturity classification standard based on color feature quantification

H 通道分量集中分布范围/(°)	成熟度
[0,4]	红熟后期
[5,9]	红熟中期
[10,18]	红熟前期
[19,27]	变色期
[28,45]	绿熟期

影响,该成熟度颜色范围内的像素面积会偏大,对绿熟期单独处理,将像素面积占比阈值设为 90%,并且后 1 个成熟等级也就是变色期的颜色范围像素面积占绿熟期的比例阈值设为 5%,其余同步骤(1)。

(3)针对剩余颜色不均匀或是颜色分布于成熟度等级界限附近的番茄图像,将相邻 2 个成熟度等级的颜色面积合并计算(除绿熟期外共合并为 3 组,变色期与红熟前期、红熟前期与红熟中期、红熟中期与红熟后期),如果任意 1 组颜色面积占比超过 50%,则进行组内二次判断,如果组内后 1 个等级颜色面积超过前 1 个等级的 20%,则分级为后 1 个等级,否则分级为前 1 个等级。

(4)同样地,对绿熟期单独处理。将绿熟期和变色期合并为 1 组,当该组像素面积相加占比大于 95%,并且变色期颜色范围的像素面积占比大于 10%,分级为变色期。

(5)降低步骤(3)的组合像素面积相加阈值为 30%,其余同步骤(3)。

(6)对剩余颜色不均匀、遮挡等情况复杂的图像,虽还可以继续探索适当的分级,但已经很难通过传统图像处理方法解决。并且前面算法步骤如存在着分级错误的情况,最终还是需要通过人工对量化自动分级的结果进行研判和校正。故本研究不再进一步分级处理,按照像素面积占比最大的成熟度等级进行直接分级。

基于以上 6 个步骤的番茄成熟度自动分级算法,得到分级正确示例如图 3(图 3a ~ 3c 列从上至下分别是番茄果实原图像、颜色分布直方图(HSV－*H*)和各成熟度等级颜色区间比重饼图)所示。如果分级错误,还需要进行人工矫正。

1.4 试验环境与评价指标

试验所用软硬件环境:CPU 为 AMD Ryzen 9 5900HX,GPU 为 NVIDIA GeForce RTX 3080 Laptop,深度学习框架为 PyTorch 1.12.1,GPU 加速库为 CUDA 11.8。训练过程中,初始学习率为 0.01,采用余弦退火策略对学习率进行调整,随着训练批次的增加逐渐降低学习率。批次大小设置为 16(Faster R－CNN 模型受硬件配置限制,设置为 8),训练迭代次数为 300。使用随机梯度下降对模型参数进行更新,动量因子设置为 0.97。通过非极大值抑制(NMS)方式消除冗余的预测框,NMS 阈值为 0.5。

本文在自然场景下进行番茄成熟度分级检测,未来可能作为智能移动终端视觉系统进行应用,在评价模型时,不仅要考虑到模型检测精度,也应综合考虑模型运行消耗的资源,从而尽可能降低设备运行的配置要求。因此,采取评价指标包括交并比(IoU)设定为 0.5 的平均精度均值(mean Average Precision, mAP)、浮点运算量(FLOPs)、参数量(Params)和模型内存占用量。

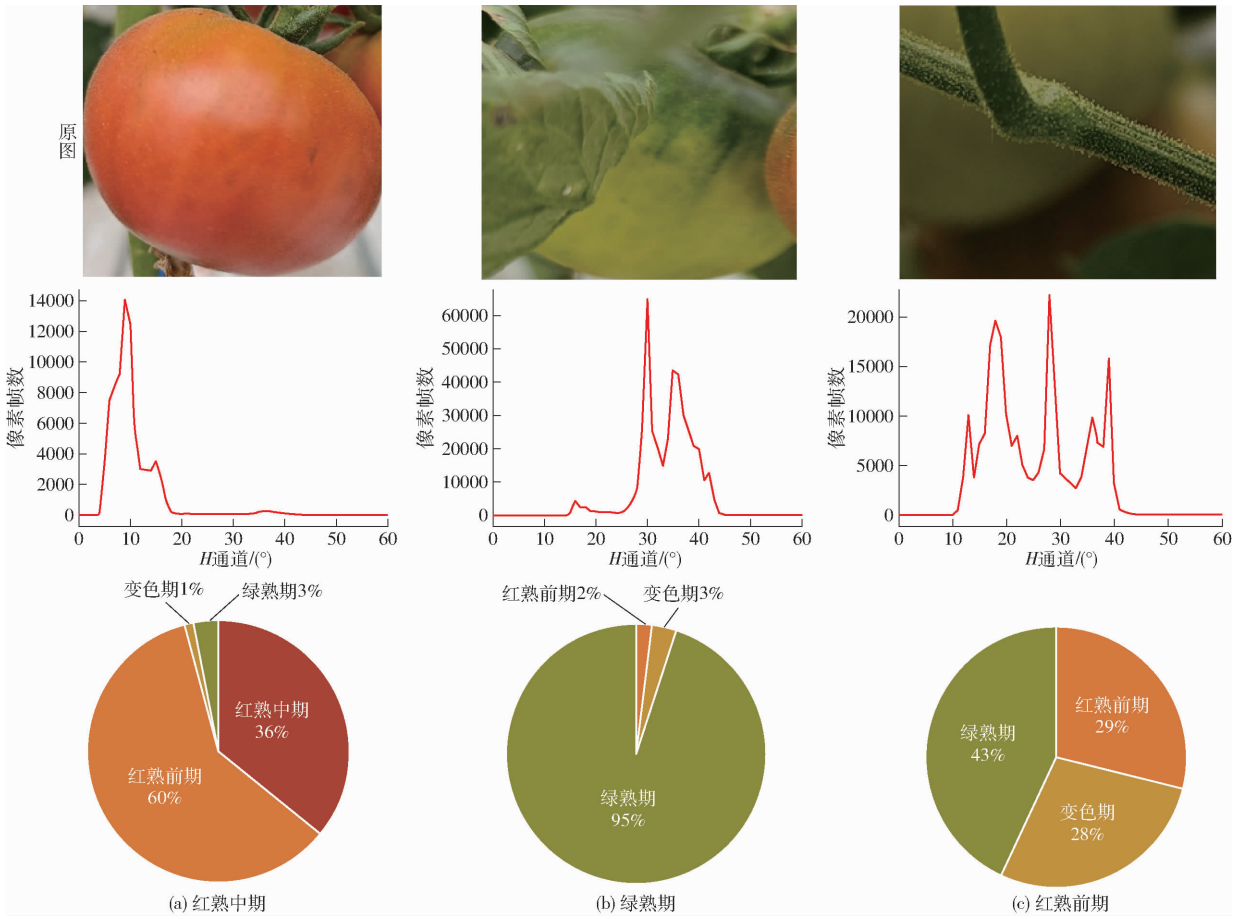

图 3　分级正确示例

Fig. 3　Example of correct grading

2　YOLO v8_BFS 模型构建

YOLO v8 作为单阶段目标检测算法，基于 YOLO v5 进行了多方面改进，提升了检测速度和精度，达到 SOTA 效果。其网络结构分为 Backbone、Neck 和 Head 3 部分，主要改进包括用 C2f 结构替代 YOLO v5 中的 C3 结构，增加了跳层连接和切分操作，以增强特征提取能力；引入解耦头（Decoupled - Head），将目标定位回归和类别分类任务解耦，分别计算回归损失和分类损失，提高模型的收敛速度和检测精度；以及通过 Task - Aligned Assigner 实现无锚框，动态调整正负样本分配，消除对锚框的依赖，从而提升检测准确性和灵活性，增强了模型对不同形状物体的拟合能力。

以 YOLO v8 为基准模型，通过将 FPN 结构替换为 BiFPN 结构，实现更高效的多尺度特征融合，提高模型的特征表达能力和融合效率；利用 SE 注意力机制对空间和通道进行融合特征提取，加强模型对关键特征信息的提取能力；引入 Focal SIoU 损失函数对预测框与真实框之间的角度差异进行度量，通过以上改进措施，降低复杂背景对检测效果的影响，提高模型的特征提取能力，从而更好地解决番茄成熟度分级检测准确率偏低的问题。改进后的模型称为 YOLO v8_BSF，如图 4 所示。

2.1　将 FPN 结构替换为 BiFPN 结构

YOLO v8 使用网络 FPN 很好地解决了浅层特征与深层特征的信息流动问题，但各尺度的特征图不仅存在着语义信息和特征信息不同的特点，还存在着各自更适合于检测不同尺度目标的特点。在 FPN 中，对于不同尺度的特征图进行特征融合时，只是简单相加。为了解决这一问题，BiFPN 提出一种简单而高效的加权双向特征金字塔网络，通过引入可学习的变量控制各尺度特征图的权重，从而使各尺度特征图对于融合后特征的贡献比例更加科学。此外 BiFPN 还移除了只有单边输入的结点，同时加入一些额外的残差连接通路，也即跨尺度的跳跃连接，减少了计算量，同时加强了各尺度之间的特征信息流通，促进更多特征的高效融合。

2.2　添加 SE 注意力机制

SE 注意力机制通过自动学习的方式，对输入图像的每个通道权重进行动态调整，基于权重矩阵赋予图像中不同区域以不同权重，即提升重点关注的特征通道权重，降低无关特征通道的权重，从而使模型有重点地侧重某些通道，提高整体的处理效率。

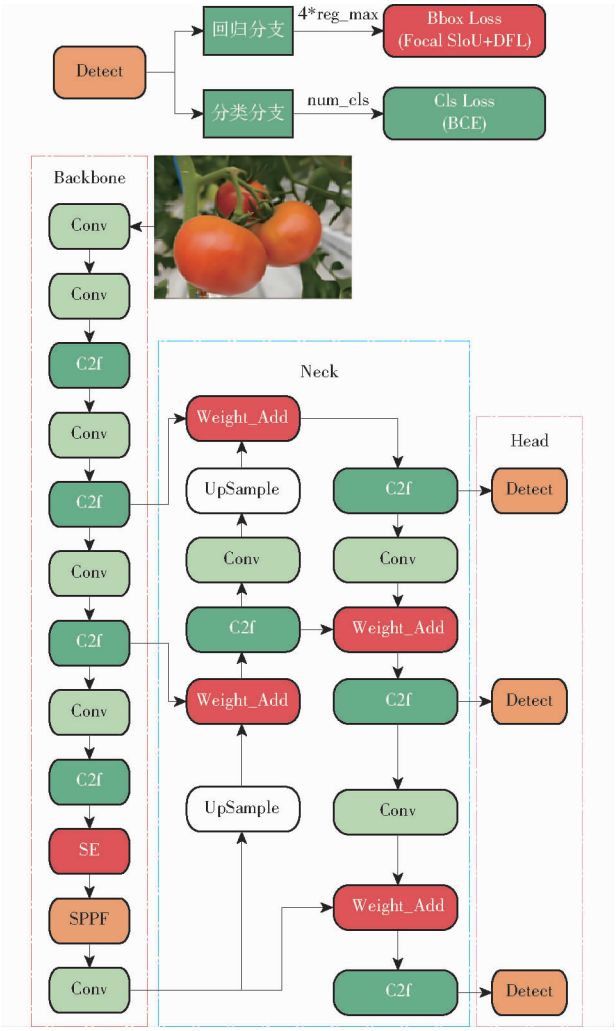

图 4 改进模型 YOLO v8_BSF 模型结构

Fig. 4 Structure of improved model YOLO v8_BSF

SE 注意力机制模块如图 5 所示,左侧输入的长方体颜色统一,表示每个通道权重相同,而在经过 SE 注意力机制处理(即图 5 的 $F_{sq}(\cdot)$ 和 $F_{ex}(\cdot,W)$ 操作)后,右侧输出的长方体颜色各异,表示每个通道被赋予了不同的权重。通过 SE 注意力机制,可以使模型更专注于输入特征图的重要数据,从而提升模型整体效率和性能。

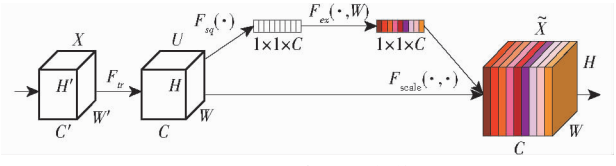

图 5 SE 注意力机制模块

Fig. 5 SE attention mechanism block

2.3 引入 Focal SIoU 损失函数

YOLO v8 对分类和回归任务解耦后,将 CIoU[26] 和 DFL[27] 组合作为边框回归的损失函数以提高模型的检测精度。但 CIoU 损失函数并没有考虑目标框的方向信息,会导致模型在某些场景下表现不佳。而 SIoU 损失函数[28-29] 针对这一问题,在 CIoU 损失的基础上,进一步考虑了真实框和预测框之间的向量角度,在函数中添加了角度惩罚指标,可以更准确地衡量预测框和真实框之间的相似性,有效地对预测框回归的自由度进行控制,提高模型的收敛速度和检测精度。SIoU 损失函数由 4 个部分组成,包括角度损失、距离损失、形状损失和 IoU 损失,计算公式为

$$L_{SIoU}=1-SIoU=1-IoU+\frac{\Delta+\Omega+\Lambda}{3} \tag{1}$$

其中角度损失 Λ 为

$$\Lambda=1-2\sin^2\left(\arcsin x-\frac{\pi}{4}\right) \tag{2}$$

其中

$$x=\frac{C_h}{\sigma}=\sin a$$

$$\sigma=\sqrt{(b_{c_x}^{gt}-b_{c_x})^2+(b_{c_y}^{gt}-b_{c_y})^2}$$

$$C_h=\max(b_{c_y}^{gt},b_{c_y})-\min(b_{c_y}^{gt},b_{c_y})$$

其中距离损失 Δ 为

$$\Delta=\sum_{t=x,y}(1-\mathrm{e}^{-\gamma\rho_t}) \tag{3}$$

其中

$$\rho_x=\left(\frac{b_{c_x}^{gt}-b_{c_x}}{c_w}\right)^2\quad\rho_y=\left(\frac{b_{c_y}^{gt}-b_{c_y}}{c_h}\right)^2$$

$$\gamma=2-\Lambda$$

其中形状损失 Ω 为

$$\Omega=\sum_{t=w,h}(1-\mathrm{e}^{-\omega_t})^{\theta} \tag{4}$$

其中

$$\omega_w=\frac{|w-w^{gt}|}{\max(w,w^{gt})}\quad\omega_h=\frac{|h-h^{gt}|}{\max(h,h^{gt})}$$

其中 IoU 损失为

$$IoU=\frac{|B\cap B^{GT}|}{|B\cup B^{GT}|} \tag{5}$$

式中 θ——对形状损失的注意程度控制参数,取 4

此外,目标检测边界框回归任务中,也存在训练样本不平衡的问题。因此,本文参考 Focal EIoU 损失函数[30],将 Focal L1 损失函数与 SIoU 损失函数结合,得到 Focal SIoU 损失函数计算公式为

$$L_{Focal_{SIoU}}=IoU^{\gamma}L_{SIoU} \tag{6}$$

式中 γ——控制异常值抑制程度的参数,取 0.5

采用 Focal SIoU 损失函数代替 YOLO v8 的 CIoU 损失函数,使模型在更好地衡量回归损失的同时,也能更集中于高质量实例。

3 试验结果与分析

3.1 消融试验

在试验过程中,以 YOLO v8n 作为基准模型,以自建数据集为试验数据,分别在数据不增强的模型和数据增强的模型下进行消融试验,试验结果如表 3。可以发现,数据增强模型在模型内存占用量

表 3　数据增强策略消融试验结果

Tab. 3　Ablation results of data augmentation improvement strategies

模型	数据增强	mAP@0.5/%	浮点运算量	参数量	模型内存占用量/MB
YOLO v8	不增强	89.70	8.20×10^9	3.01×10^6	5.94
	增强	91.10			5.95

几乎不变的前提下，平均精度均值提高 1.4 个百分点，故在后续试验中采用数据增强模型。

表 4 为不同模型消融试验结果。由表 4 可知，将 FPN 结构替换为 BiFPN 结构后，检测精度也有了明显提高，mAP@0.5 上升到 92.65%，浮点运算量、参数量和内存占用量都只有略微增加，充分证明了

表 4　模型改进策略消融试验结果

Tab. 4　Ablation results of model improvement strategies

模型	BiFPN	SE	Focal SIoU Loss	mAP@0.5/%	浮点运算量	参数量	模型内存占用量/MB
YOLO v8	×	×	×	91.10	8.20×10^9	3.01×10^6	5.95
YOLO v8_B	√	×	×	92.65	8.30×10^9	3.03×10^6	5.99
YOLO v8_BS	√	√	×	93.01	8.30×10^9	3.04×10^6	6.01
YOLO v8_ BSF	√	√	√	94.10	8.30×10^9	3.04×10^6	6.01

注：√表示采用此模块，×表示不采用此模块。

相较于 FPN 结构，BiFPN 结构能够更有效地促进数据信息流动，提高模型的性能；在模型的 Neck 层继续加入 SE 注意力模块后，在模型浮点运算量、参数量和内存占用量几乎没有变动的情况下，mAP@0.5 增加 0.36 个百分点，证明 SE 注意力模块能够提升模型对于关键特征的提取能力，增加模型的检测精度；同样，在模型参数量、模型浮点运算量和内存占用量不变的情况下，Focal SIoU 损失函数也进一步提高了检测精度，证明该函数能够更好地衡量 BBR 任务中的预测损失。YOLO v8_BSF 整体检测精度有了显著提高，mAP@0.5 达到 94.10%，相较于原模型提升 3.0 个百分点。

如图 6 所示，虽基准模型 YOLO v8 的损失函数下降和检测精度上升都更加快速，但随着迭代次数增长，明显出现了过拟合现象，而改进模型 YOLO v8_BSF 在 100 次迭代后，检测精度仍在稳步提升，模型学习能力和避免过拟合的性质都相对更好。

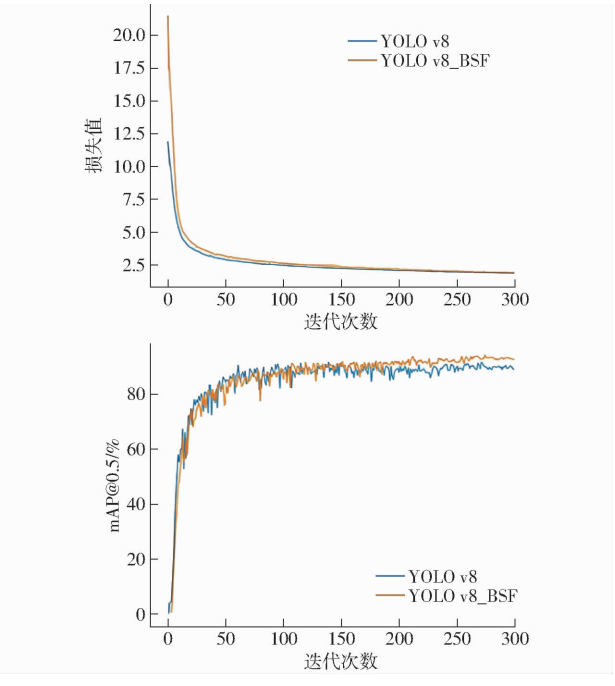

图 6　模型训练过程损失曲线和检测精度曲线

Fig. 6　Loss curves and detection accuracy curves during model training

图 7 为 YOLO v8 和 YOLO v8_BSF 处理后的可视化检测结果。第 1 列是背景简单的番茄图像，YOLO v8 和改进模型均能较好地实现对果实及其成熟度的检测；第 2 列是遮挡情况略微复杂的番茄图像，YOLO v8 对于被大幅度遮挡的果实出现了误检情况，而 YOLO v8_BSF 仍实现了对图中番茄果实的正确检测；第 3 列是果实多、距离各异，但几乎无遮挡的番茄图像，YOLO v8 和 YOLO v8_BSF 均正确实现了对果实及其成熟度的检测；第 4 列是背景复杂、番茄果实多、遮挡情况严重的图像，YOLO v8 将逆光的红熟中期番茄果实错检测为了红熟后期，而 YOLO v8_BSF 实现了正确检测。分析可得，自然场景下番茄成熟度分级检测难点主要集中在背景复杂、遮挡严重等情况。YOLO v8_BSF 有效地解决了这一问题，能够较好地实现高精度的番茄成熟度分级检测。

3.2　不同模型分级试验

为了更好地评估模型 YOLO v8_BSF 的性能，验证模型在番茄成熟度分级检测中的可行性，选取主流的目标检测模型进行模型性能与分级横向对比试验，试验结果如表 5 和图 8 所示。

由表 5 可知，YOLO 系列模型的轻量级版本相对于两阶段检测模型 Faster R－CNN 来说，在模型浮点运算量（FLOPs）、参数量（Params）和内存占用量上均有显著优势，以 YOLO v8 为例，模型浮点运算量、参数量和内存占用量仅为 Faster R－CNN－Resnet50 的 0.87%、10.59% 和 1.89%。在 YOLO 系列模型中，YOLO v7－tiny 的检测精度、浮点运算



图 7 可视化检测结果

Fig. 7 Visualization of detection results

表 5 不同模型性能试验结果

Tab. 5 Experiment results of different models

模型	mAP@0.5/%	浮点运算量	参数量	模型内存占用量/MB
Faster R-CNN-Resnet50	89.91	9.40×10^{11}	2.84×10^{6}	315.00
YOLO v5	90.53	4.20×10^{9}	1.70×10^{6}	3.69
YOLO v7-tiny	90.52	1.32×10^{10}	6.02×10^{6}	11.70
YOLO v8	91.10	8.20×10^{9}	3.01×10^{6}	5.95
YOLO v10	89.80	8.40×10^{9}	2.71×10^{6}	5.51
YOLO 11	92.50	6.40×10^{9}	2.59×10^{6}	5.24
YOLO v8_BSF	94.10	8.30×10^{9}	3.04×10^{6}	6.01

量、参数量和内存占用量等均劣于 YOLO v5 模型，YOLO v5 模型尽管浮点运算量、参数量和内存占用量均最小，但检测精度却不高。YOLO v10 的检测精度低于 YOLO 系列其他 5 个轻量版本模型，并且在浮点运算量、参数量和内存占用量上也没有显著优势，而改进 YOLO v8_BSF 模型的检测精度优于 YOLO 系列的其他 5 个轻量版本模型，平均精度均值为 94.1%，同时浮点运算量、参数量和内存占用量仅为 8.30×10^{9}、3.04×10^{6} 和 6.01 MB，能够满足

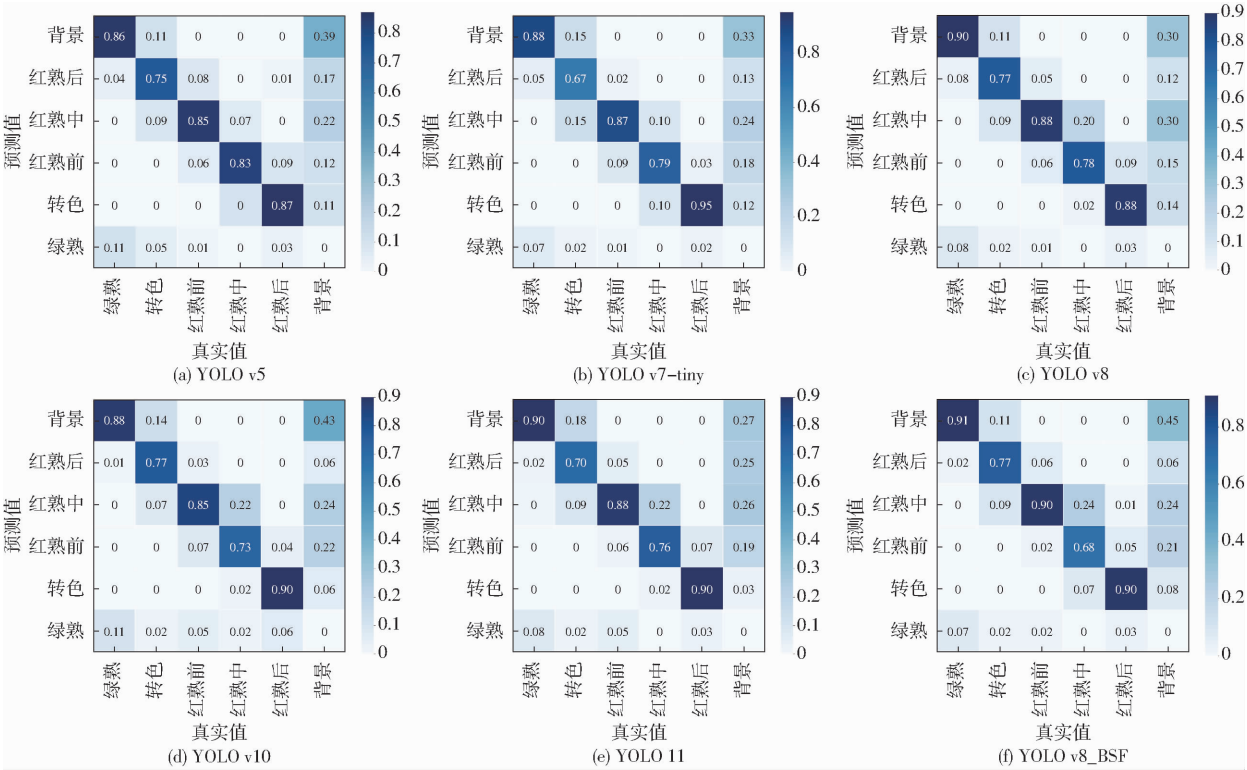

图 8 不同模型混淆矩阵

Fig. 8 Confusion matrix of different models

本研究的高精度和轻量化需求。

图 8 为不同模型分级结果的混淆矩阵，矩阵中区域颜色越深，表示概率越大，矩阵对角线为预测正确区域。从图 8 可知，对各成熟度果实误检，都是将果实检测为相近成熟度；绿熟期和红熟后期颜色特征鲜明、只有一个相近成熟度，从而检测准确率相对更高；绿熟期与叶片颜色相近，易被漏检，以上均符合番茄成熟度分级检测的特点。

3.3　泛化能力测试

实际生产中番茄的种植场景多样、品种多样，为了进一步测试番茄成熟度分级检测模型的泛化能力和鲁棒性，采用公开数据集 laboro_tomato_big 进行试验。YOLO v8_BSF 模型平均精度均值 mAP@0.5 为 88.30%，可视化检测结果如图 9 所示。

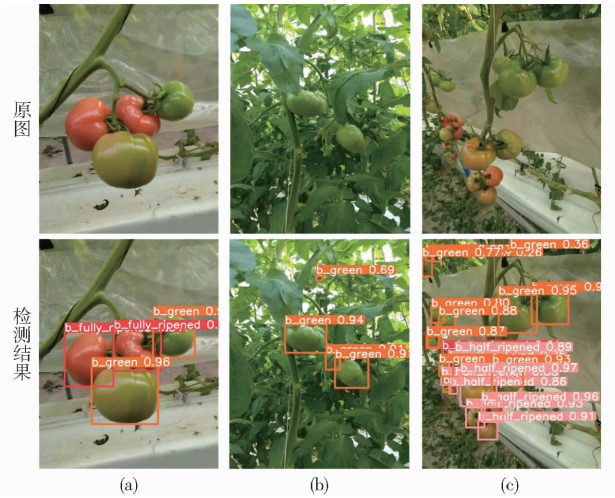

图 9　公开数据集检测结果可视化

Fig. 9　Visualization of detection results on public dataset

相较于自建数据集，YOLO v8_BSF 在公开数据集 laboro_tomato_big 的检测精度有 5.8 个百分点的差距。主要原因有：①公开数据集采集的番茄品种纹理特征不明显，从而未成熟绿色番茄与番茄茎叶较为相近，为识别带来困难（图 9b）。②公开数据集背景较为复杂、远距离小果实较多的图像占据比例较大，为检测带来困难。③由于对图像分辨率进行压缩时会造成信息损失，不利于小目标检测，因此进一步降低了整体检测精度（图 9c）。

3.4　番茄成熟度分级检测系统

为了将本研究实际应用到农业种植场景中，帮助番茄种植人员通过简单便捷的操作方式进行番茄果实检测、获取其成熟度以及相关信息，番茄成熟度分级检测系统使用 YOLO v8_BFS 模型训练得到的 best.pt 权重文件对番茄成熟度进行检测。为贴近实际生产需求，提高系统的实用性，系统包含番茄成熟度检测、检测历史和反馈历史功能。用户在检测页面（图 10a）可选择决定从相册中上传或者直接拍摄上传进行成熟度检测，检测完成自动跳转检测结果页面（图 10b）；在检测结果页面点击下方的结果反馈按钮即可跳转结果反馈页面（图 10c），结果反馈页面简要展示了检测结果，主要是提供对本次检测结果的反馈功能，并且展示本次检测结果的历史反馈记录；检测历史（图 10d）界面，以日期倒序展示用户的检测历史，点击某一条检测记录后，则会跳转到该条记录的检测结果页面。

系统编程环境为 Python 3.8.13、MySQL 5.7、微信开发者工具。该系统为番茄生产提供了一个自动化、智能化的成熟度分级检测工具，提升了检测效率，为未来农业智能化发展提供了有力支持。

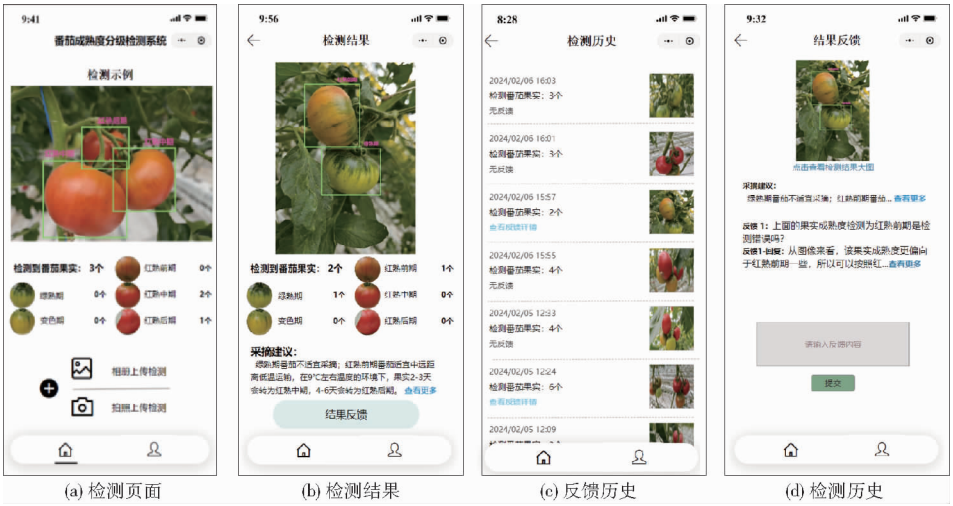

图 10　系统操作界面

Fig. 10　System operation interface

4　结论

（1）提出了一种基于颜色特征的番茄成熟度

细粒度分级标准,并开发了以番茄果实成熟度分级算法为基础的番茄图像半自动标注算法。研究选取了 5 个成熟度等级,分析了颜色特征与成熟度之间的关系,并基于 HSV 颜色模型量化成熟度分级标准。该半自动标注分级算法能够有效消除人工标注的不足。

(2)提出了一种基于改进 YOLO v8 的高精度番茄成熟度分级检测模型 YOLO v8_BSF:引入 BiFPN 结构和 SE 注意力机制,在复杂自然场景下提高了检测稳定性和精度。试验结果表明,mAP@0.5 达到 94.10%,YOLO v8_BSF 在复杂背景条件下表现优异,且与主流模型相比综合性能更佳。

(3)开发了番茄成熟度分级检测系统,能满足移动场景的实际应用需求。通过合理的系统架构设计,系统功能测试结果显示,该系统能够实现便携的番茄成熟度检测。

参 考 文 献

[1] NEMATI A, SENDI J J, FATHIPOUR Y. Combined effects of gibberellin and vermiwash on the life history and antioxidant system of phthorimaea absoluta (Meyrick) in tomato plants[J]. Scientific Reports, 2025, 15(1):4427.

[2] LIU Y, WEI C, YOON S C, et al. Development of multimodal fusion technology for tomato maturity assessment[J]. Sensors, 2024, 24(8):2467.

[3] 李峰. 蔬菜机械化收获技术及其发展需求分析[J]. 农机使用与维修, 2024(3):78 - 80.
LI Feng. Analysis of mechanized vegetable harvesting technology and its development needs[J]. Agricultural Machinery Use and Maintenance, 2024(3):78 - 80. (in Chinese)

[4] 杨洲, 李君略, 刘清化. 水果生产机械化发展动态[J]. 现代农业装备, 2017(5):13 - 18.
YANG Zhou, LI Junlue, LIU Qinghua. Development trends in mechanization of fruit production[J]. Modern Agricultural Equipment, 2017(5):13 - 18. (in Chinese)

[5] 宋卫堂. 日光温室蔬菜生产全程机械化的一种解决方案[J]. 中国农机化学报, 2018, 39(9): 26 - 29.
SONG Weitang. A solution for full mechanization of vegetable production in solar greenhouses[J]. Journal of Chinese Agricultural Mechanization, 2018, 39(9):26 - 29. (in Chinese)

[6] 安树杰, 陈怀勐,祝保英,等. 连栋温室番茄生产关键环节机械化栽培技术应用[J]. 蔬菜, 2021(增刊 1):67 - 69.
AN Shujie, CHEN Huaimeng, ZHU Baoying, et al. Application of key mechanized cultivation techniques in greenhouse tomato production[J]. Vegetables, 2021(Supp. 1):67 - 69. (in Chinese)

[7] 曾祥伟, 张天柱, 何小凡, 等. 设施番茄产量影响因素分析[J]. 新农业,2020(1):47 - 49.
ZENG Xiangwei, ZHANG Tianzhu, HE Xiaofan, et al. Analysis of factors affecting tomato yields in facilities[J]. New Agriculture, 2020(1):47 - 49. (in Chinese)

[8] 林伟明, 胡云堂. 基于 YUV 颜色模型的番茄收获机器人图像分割方法[J]. 农业机械学报, 2012, 43(12): 176 - 180.
LIN Weiming, HU Yuntang. Tomato harvesting robot image segmentation method based on YUV color model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(12):176 - 180. (in Chinese)

[9] 冯俊惠, 李志伟, 戎有丽, 等. 基于改进 Hough 圆变换算法的成熟番茄果实识别[J]. 中国农机化学报, 2021, 42(4): 190 - 196.
FENG Junhui, LI Zhiwei, RONG Youli, et al. Recognition of ripe tomato fruit based on improved Hough circle transform algorithm[J]. Journal of Chinese Agricultural Mechanization, 2021, 42(4):190 - 196. (in Chinese)

[10] 毕松, 高峰, 陈俊文, 等. 基于深度卷积神经网络的柑橘目标识别方法[J]. 农业机械学报, 2019, 50(5): 181 - 186.
BI Song, GAO Feng, CHEN Junwen, et al. Citrus target recognition method based on deep convolutional neural networks[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5):181 - 186. (in Chinese)

[11] 何斌, 张亦博, 龚健林, 等. 基于改进 YOLO v5 的夜间温室番茄果实快速识别[J]. 农业机械学报, 2022, 53(5): 201 - 208.
HE Bin, ZHANG Yibo, GONG Jianlin, et al. Fast recognition of greenhouse tomatoes at night based on improved YOLO v5 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5):201 - 208. (in Chinese)

[12] ZHANG P, LIU X, YUAN J, et al. YOLO v5 - spear: a robust and real-time spear tips locator by improving image augmentation and lightweight network for selective harvesting robot of white asparagus[J]. Biosystems Engineering, 2022, 218: 43 - 61.

[13] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117 - 2125.

[14] TAN M, PANG R, LE Q V. Efficientdet: scalable and efficient object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781 - 10790.

[15] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018: 7132 - 7141.

[16] 杨明宇,曾剑波,李明博,等. 密云区设施番茄产业发展现状,问题与对策[J]. 蔬菜, 2023(11):52 - 55.
YANG Mingyu, ZENG Jianbo, LI Mingbo, et al. Current situation, problems and countermeasures of facility tomato industry in Miyun District[J]. Vegetables, 2023(11):52 - 55. (in Chinese)

[17] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.

[18] 中华人民共和国供销合作总社. 番茄:GH/T 1193—2021[S]. 北京:中华全国供销合作总社,2021:1 - 6.

(下转第 224 页)

2211.12905, 2022.
[27] TAN M. Efficientnet: rethinking model scaling for convolutional neural networks[C]// Proceedings of the 36th International Conference on Machine Learning. Long Beach, USA: PMLR, 2019: 6105 - 6114.
[28] 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 (ICCV). Seoul, South Korea: IEEE, 2019: 3007 - 3016.
[29] LI C, ZHOU A, YAO A. Omni-dimensional dynamic convolution[J]. arXiv Preprint, arXiv:2209.07947, 2022.
[30] QUAN Y, ZHANG D, ZHANG L, et al. Centralized feature pyramid for object detection[J]. IEEE Transactions on Image Processing, 2023, 32: 4341 - 4354.
[31] LI J, WEN Y, HE L. Scconv: spatial and channel reconstruction convolution for feature redundancy[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE, 2023: 6153 - 6162.
[32] LI Y, HOU Q, ZHENG Z, et al. Large selective kernel network for remote sensing object detection[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE, 2023: 16794 - 16805.
[33] YANG L, ZHANG R Y, LI L, et al. Simam: a simple, parameter-free attention module for convolutional neural networks[C]// Proceedings of the 38th International Conference on Machine Learning, PMLR, 2021: 11863 - 11874.
[34] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning[C]// ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2023: 1 - 5.
[35] CAO Y, XU J, LIN S, et al. Gcnet: non-local networks meet squeeze-excitation networks and beyond[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops. Seoul, South Korea: IEEE, 2019: 1971 - 1980.
[36] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: convolutional triplet attention module[C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Online Virtual Conference, 2021: 3139 - 3148.
[37] HU J, SUN G, LI X, et al. Gather-excite: exploiting feature context in convolutional neural networks[C]// Proceedings of the 32nd International Conference on Neural Information Processing Systems (NeurIPS). Montreal, Canada, 2018: 9401 - 9411.

(上接第 202 页)

[19] 杨鸣娟, 张坤生. 番茄品种、成熟度和热处理对番茄中番茄红素含量的影响[J]. 食品研究与开发, 2003(3): 99 - 101.
YANG Mingjuan, ZhANG Kunsheng. Effects of tomato variety, maturity and heat treatment on lycopene content in tomatoes [J]. Food Research and Development, 2003(3):99 - 101. (in Chinese)
[20] 庞淑艳. 关于番茄果实颜色的主要研究进展[J]. 现代农业, 2015(2):119 - 120.
PANG Shuyan. Major research progress on tomato fruit color[J]. Modern Agriculture, 2015(2):119 - 120. (in Chinese)
[21] 陈文静. 番茄果实成熟过程中番茄红素含量变化及关键基因的表达分析[D]. 石家庄:河北科技大学, 2020.
CHEN Wenjing. Changes in lycopene content and key gene expression during tomato fruit ripening[D]. Shijiazhuang:Hebei University of Science and Technology, 2020. (in Chinese)
[22] 袁奋杰, 周晓, 丁军, 等. 基于 FPGA 的 RGB 和 HSV 色空间转换算法实现[J]. 电子器件, 2010, 33(4): 493 - 497.
YUAN Fenjie, ZHOU Xiao, DING Jun, et al. Implementation of RGB and HSV color space conversion algorithm based on FPGA[J]. Electronic Devices, 2010, 33(4):493 - 497. (in Chinese)
[23] 马玲, 张晓辉. HSV 颜色空间的饱和度与明度关系模型[J]. 计算机辅助设计与图形学学报, 2014, 26(8):1272 - 1278.
MA Ling, ZHANG Xiaohui. Relationship model of saturation and brightness in HSV color space[J]. Journal of Computer-Aided Design and Computer Graphics, 2014, 26(8):1272 - 1278. (in Chinese)
[24] 陶华伟, 赵力, 奚吉, 等. 基于颜色及纹理特征的果蔬种类识别方法[J]. 农业工程学报, 2014, 30(16): 305 - 311.
TAO Huawei, ZHAO Li, XI Ji, et al. Fruit and vegetable classification method based on color and texture features[J]. Transactions of the CSAE, 2014, 30(16):305 - 311. (in Chinese)
[25] 李天华, 孙萌, 丁小明, 等. 基于 YOLO v4 + HSV 的成熟期番茄识别方法[J]. 农业工程学报, 2021, 37(21): 183 - 190.
LI Tianhua, SUN Meng, DING Xiaoming, et al. Tomato maturity recognition method based on YOLO v4 + HSV[J]. Transactions of the CSAE, 2021, 37(21):183 - 190. (in Chinese)
[26] ZHENG Z, WANG P, LIU W, et al. Distance - IoU loss: faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993 - 13000.
[27] LI X, WANG W, WU L, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection[J]. Advances in Neural Information Processing Systems, 2020, 33: 21002 - 21012.
[28] GEVORGYAN Z. SIoU loss: more powerful learning for bounding box regression[J]. arXiv preprint arXiv:2205.12740, 2022.
[29] 王晓明,陈智宇,董文涛,等. 基于 YOLO v7x 的接触网吊弦缺陷检测方法[J]. 华东交通大学学报, 2024,41(3):65 - 73.
WANG Xiaoming, CHEN Zhiyu, DONG Wentao, et al. Catenary dropper defect detection method based on YOLO v7x[J]. Journal of East China Jiaotong University, 2024,41(3):65 - 73. (in Chinese)
[30] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IoU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146 - 157.