

doi:10.12068/j.issn.1005-3026.2025.20249020

## 基于改进YOLOv8的脱水蔬菜异物检测方法

胡博<sup>1</sup>, 熊华德<sup>2</sup>, 刘尧<sup>3</sup>, 张勇军<sup>1</sup>

(1. 北京科技大学 工程技术研究院, 北京 102206; 2. 北京科技大学 设计研究院有限公司, 北京 102206;

3. 中国科学院 理化技术研究所, 北京 100190)

**摘要:** 针对脱水蔬菜生产过程中人工质检工作量大、检测效率低、工人质检标准不一致等问题, 提出了一种基于改进YOLOv8的异物检测方法YOLOv8n-BCS, 以辅助工人提高质检效率并减轻劳动强度。YOLOv8n-BCS模型在主干网络中引入ShuffleNetV2和BoTNet(bottleneck transformer network), 在颈部网络结构融入SimAM(simple attention module)注意力机制和轻量化上采样算子CARAFE(content-aware reassembly of features), 同时采用SIoU(similarity intersection over union)函数计算回归损失。使用NVIDIA GeForce RTX 3080服务器进行训练测试, 实验结果表明: YOLOv8n-BCS模型精确率 $P$ 为96.8%, 召回率 $R$ 为94.7%, 调和平均值 $F1$ 为95.7%, 平均精度均值(mAP)为97.1%, 帧率为231 f/s, 模型体积为6.1 MB。相比对照模型, YOLOv8n-BCS模型容量减小, 检测速度和平均精度均值提升, 可为脱水蔬菜智能检测分拣系统优化提供技术参考。

**关键词:** 脱水蔬菜异物; YOLOv8; 目标检测; 图像识别; 轻量化

中图分类号: TP 393.4

文献标志码: A

文章编号: 1005-3026(2025)11-0019-12

## Foreign Object Detection Method for Dehydrated Vegetables Based on Improved YOLOv8

HU Bo<sup>1</sup>, XIONG Hua-de<sup>2</sup>, LIU Yao<sup>3</sup>, ZHANG Yong-jun<sup>1</sup>

(1. Institute of Engineering Technology, University of Science and Technology Beijing, Beijing 102206, China;

2. Design and Research Institute Co., Ltd., University of Science and Technology Beijing, Beijing 102206, China;

3. Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, Beijing 100190, China.

Corresponding author: ZHANG Yong-jun, E-mail: zhangyj@ustb.edu.cn)

**Abstract:** Problems such as heavy workload, low manual detection efficiency, and inconsistent quality inspection standards of workers during the production process of dehydrated vegetables exist. To address these issues, a foreign object detection method based on improved YOLOv8, namely YOLOv8n-BCS, was proposed. This method could assist workers in improving quality inspection efficiency and reducing labor intensity. The YOLOv8n-BCS model introduced ShuffleNetV2 and BoTNet (bottleneck transformer network) into the backbone network and incorporated the simple attention module (SimAM) attention mechanism and lightweight upsampling operator content-aware reassembly of features (CARAFE) into the neck structure. The similarity intersection over union (SIoU) function was also used to calculate regression loss. By using an NVIDIA GeForce RTX 3080 server for training and testing, the experimental results show that the YOLOv8n-BCS model has an accuracy  $P$  of 96.8%, a recall  $R$  of 94.7%, a harmonic mean  $F1$  of 95.7%, a mean average accuracy (mAP) of 97.1%, a frame rate of 231 f/s, and a model volume of 6.1 MB. Compared with the control model, the YOLOv8n-BCS model has reduced capacity, as well as improved detection speed and average accuracy, providing a technical reference for optimizing intelligent detection and sorting systems for dehydrated vegetables.

**Key words:** foreign object in dehydrated vegetables; YOLOv8; object detection; image recognition; lightweight

收稿日期: 2024-04-10

作者简介: 胡博(1994—),男,湖北襄阳人,北京科技大学博士研究生;张勇军(1973—),男,山东东营人,北京科技大学教授,博士生导师。

脱水蔬菜具有体积小、质量轻、易于运输和长时间储存等优点,广泛应用于航空航天、航海、军事等领域<sup>[1-2]</sup>。脱水蔬菜不合格品可能会在原材料运输、处理和烘干等过程中产生<sup>[3]</sup>,外源性异物可能在脱水蔬菜生产的各个环节混入产品中,造成食品安全问题,因此生产企业需要对产品中的杂质异物进行极其严格的检查<sup>[4]</sup>。

研究人员对于农产品及其二次加工品的质检过程自动化已有不少探索,主要涉及产品的颜色、大小、形状、纹理与损伤等<sup>[5-6]</sup>。当前农产品的质检主要为无损检测<sup>[7]</sup>,利用光谱、机器视觉、高光谱成像等技术检测其品质<sup>[8-10]</sup>,如色选机、重金属探测仪等设备。王红军等<sup>[11]</sup>针对马铃薯质量和形状的分级建立了预测模型;Blanc<sup>[12]</sup>设计的质量评级系统对葡萄等串状水果的质量评级能力有显著提升;叶晋涛等<sup>[13]</sup>提出了基于图像特征与SVM神经网络的哈密瓜成熟度分级方法,该方法对八九成熟哈密瓜识别的精确率达到了97.22%;魏康丽等<sup>[14]</sup>通过特征提取的方法建立了苹果脆片外部品质等级评价标准;Baigvand等<sup>[15]</sup>设计了一套基于机器视觉的分拣系统,通过机器视觉技术将无花果干分类为5个定性等级,并用气动单元进行不同等级无花果干的分拣。

目前脱水蔬菜生产仍需要人工质检,生产过程存在劳动强度大、工人质检标准不一致、质检效率低等问题。针对以上问题,本文作者前期开发了一套基于机器视觉与随动跟踪的脱水蔬菜智能识别分拣系统,采用YOLOv3模型在一定程度上提升了效率。但是目前YOLOv3性能已经落后,为进一步提升模型的识别速度和精度,本文基于YOLOv8n,在模型中引入ShuffleNetV2网络、BoTNet网络,轻量化上采样算子CARAFE、SimAM注意力机制和SIoU损失函数,以提升检测速度和平均精度均值(mAP),实现模型轻量化,为脱水蔬菜智能检测分拣系统优化提供技术参考。

## 1 网络模型

### 1.1 YOLOv8目标检测网络

YOLOv8是YOLO(you only look once)系列中的最新算法,相比于广泛应用的YOLOv3和YOLOv5网络,YOLOv8具有检测精度高、内存占用低和检测速度快等优点。在确保精度的同时,YOLOv8进一步实现了模型的轻量化,其网络模型的权重文件显著减小<sup>[16]</sup>。目前YOLOv8已在工业安

全场景<sup>[17-18]</sup>、虫害监测<sup>[19-21]</sup>和麦穗检测<sup>[22-24]</sup>等领域得到改进和应用。

YOLO系列算法可分为主干(Backbone)、颈部(Neck)和头部(Head)3部分结构<sup>[16]</sup>。YOLOv8网络模型体系结构包含YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l和YOLOv8x共5种模型,不同模型差异在于特征提取的深度和宽度,其中YOLOv8n检测速度最快。YOLOv8参考YOLOv7中的E-ELAN(efficient expandable layer aggregation network)结构<sup>[25]</sup>,将YOLOv5中的C3模块替换为C2f(cross stage partial networks with feature fusion)模块,使用更多分支跨层连接,在增强梯度流的同时保持模型的性能和轻量化。颈部网络依旧沿用了路径聚合网络(path aggregation network, PAN)结构,避免了卷积神经网络大型框架主干网络优化中出现重复的梯度信息,在不损失精度的情况下减少了模型尺寸。对比YOLOv5结构可知,YOLOv8删除了PAN-FPN(PAN-feature pyramid network)上采样阶段中的卷积结构。在头部网络,YOLOv8参考YOLOX将分类和检测头分离<sup>[26]</sup>,可将类别和位置特征解耦合以提升模型效率;并把Anchor-Base替换成Anchor-Free无锚框检测头,可减少预测数量,避免定位精度低和样本不均衡的情况。YOLOv8采用了Task Aligned Assigner分配策略<sup>[27]</sup>,损失计算涵盖分类和回归两个分支:分类分支使用VFL(varifocal loss)损失函数进行训练;回归分支使用DFL(distribution focal loss)和CIoU(corrected intersection over union)损失函数,以提升对边界框预测的精准性。YOLOv8模型结构如图1所示。

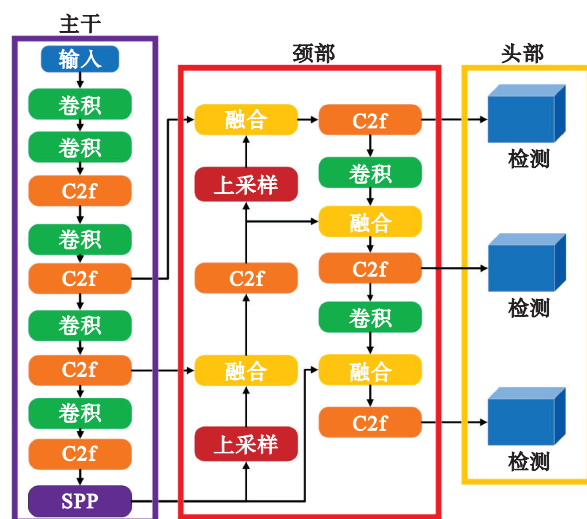


图1 YOLOv8网络结构

Fig. 1 Structure of YOLOv8 network

### 1.2 改进的 YOLOv8 目标检测网络

YOLOv8 在复杂背景下的小规格目标检测精度存在不足,且模型参数量较多,因此需对 YOLOv8 模型进一步优化.这样既能减少模型的参数量,又能提高检测精度,从而提升模型的推理速度.

#### 1.2.1 基于 YOLOv8 的网络结构改进

优化后的网络模型由 3 个部分构成:主干、颈部和头部.其中,主干网络负责对异物图像特征进行提取,颈部网络负责进行特征融合,而头部网络负责进行预测回归.为方便下文表述,将改进模型命名为 YOLOv8n-BCS,网络结构如图 2 所示.

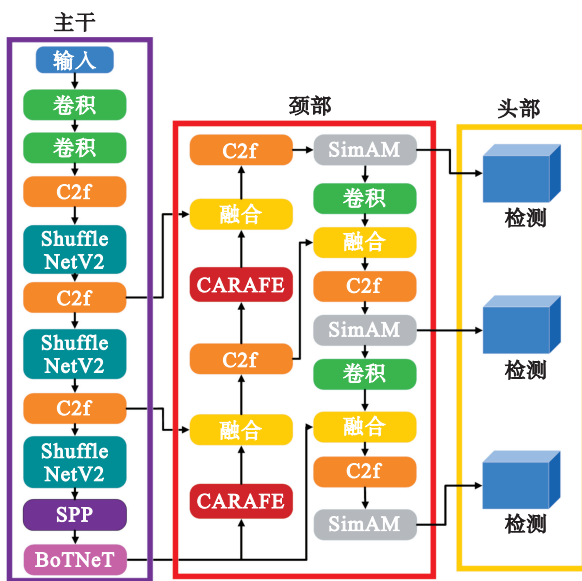


图 2 YOLOv8n-BCS 网络结构

Fig. 2 Structure of YOLOv8n-BCS network

YOLOv8n-BCS 在主干特征提取网络中先引入主干网络 ShuffleNetV2 和 BoTNet,以简化主干结构并增大模型全局感受野;同时在颈部网络中引入简单有效的轻量化上采样算子 CARAFE 和 SimAM 注意力机制,从而加快上采样速度且不添加参数直接推断特征图的三维权重,提高模型的灵活性、收敛能力以及卷积神经网络的表达能力;最后将输出端的头部损失函数从 CIoU 改为 SIoU,以提升模型的收敛能力并减小计算回归损失.

#### 1.2.2 ShuffleNetV2 网络结构

ShuffleNetV2 是一种轻量化网络结构,由 ShuffleNetV1 发展而来,引入主干网络可降低特征层的参数量.ShuffleNetV1 使用大量类瓶颈结构,使网络的输入和输出通道数不同,导致训练运算量增大,训练速度降低;ShuffleNetV1 包含很

多逐点组卷积,会增大设备硬件压力;网络中的部分模块采用多分支结构,即模块中包含多个小算子堆叠不同的卷积层和池化层,导致网络结构碎片化,降低运算效率和检测速度;ShuffleNetV1 使用 ReLU (rectified linear unit) 和 shortcut 结构,所需内存较大,限制了网络速度.

ShuffleNetV2 结构如图 3 所示,V2 在 V1 版本的基础上移除 ResNet 残差单元,并引入了通道分离(channel split)模块,将特征图通道维度分成 2 个支路<sup>[28]</sup>.左侧支路保持原有映射,右侧支路输入和输出通道数深度相同,含有 3 个连续卷积,将 2 个支路输出结果进行合并操作,并对合并结果进行通道混洗(channel shuffle)确保特征信息融合.ShuffleNetV2 在未损失检测精度的前提下,降低了运算量,同时提升了模型运算速度.

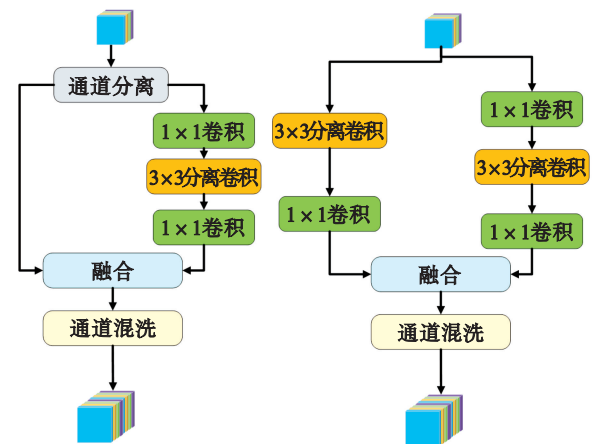


图 3 ShuffleNetV2 网络结构

Fig. 3 Structure of ShuffleNetV2 network

#### 1.2.3 BoTNet 网络

现场收集的高分辨率异物照片中,部分包含复杂的蔬菜背景,难以辨别异物信息.为增强 YOLOv8 模型网络的特征提取能力与全局信息处理能力,本文在主干网络引入了 BoTNet 架构.

BoTNet 是一种 Transformer 和 CNN (convolutional neural network) 相结合的骨干网络,其网络结构是在残差的最后 3 个瓶颈层中使用全局多头自注意力 (multi-head self-attention, MHSA) 替换 3x3 空间卷积<sup>[29]</sup>.全局多头自注意力可以收集并集中处理特征图中包含的目标信息,提升对特征的辨别能力.全局多头自注意力由多个单头自注意力块组成,其结构如图 4 所示<sup>[21]</sup>.

单头自注意力计算公式为<sup>[21]</sup>

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Soft max} \left( \frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} \right) \mathbf{V}. \quad (1)$$

式中： $Q, K, V$ 分别为 query, key 和 value 矩阵，可通过输入矩阵的线性变换获得； $QK^T$ 为用于计算  $Q$  与  $K$  的相似度矩阵； $d_k$ 为比例缩放因子，用于调节  $QK^T$  避免数值过大；Softmax 为与  $V$  矩阵映射的归一化函数，用于计算通道相关性。

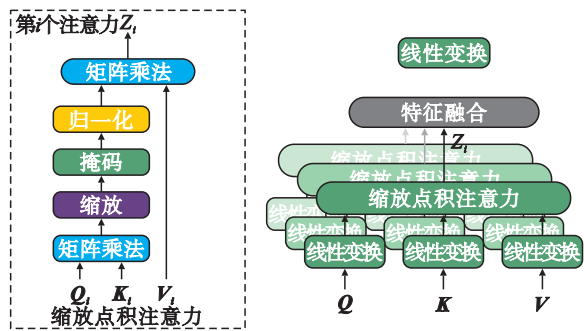


图 4 MHA 结构图

Fig. 4 Structure diagram of MHA

单头注意力块因特征子空间受限导致建模能力有限，因此采用多头自注意力机制将输入线性投影到多个特征子空间中，并通过多个相互独立的注意力层分别处理，输出向量由一层全连接层获得最终结果。

在 YOLOv8 主干网络中引入 BoTNet，可将自注意力机制与卷积结合，增强模型获取全局信息的能力并缓解卷积信息丢失的情况。BoTNet 扩展了模型专注于不同位置的能力，既可用卷积网络从高分辨率图像中获取低分辨率特征图，也可用自注意力机制来处理卷积特征图包含的信息。

BoTNet 采用  $1 \times 1$  卷积，相比  $3 \times 3$  卷积有效减少了网络参数量，在提升检测性能的同时保证了网络模型的轻量化。

### 1.2.4 上采样算子 CARAFE

在 YOLOv8 中，上采样常用的是 Upsample 算子，可以使用线性插值 (linear)、最近邻插值 (nearest) 等方法。这些方法主要依靠最邻近像素值，无法精确提取图像中微小的信息变化，导致图像特征丢失。进行放大操作时，图像会变得模糊，此时系统的感受野较小。为提高模型的全局感受野，本文引入轻量级上采样算子 CARAFE，其易于集成且运行速度快，能在一个大感受野内聚合上下文信息。相对于最近邻和双线性等上采样算子，CARAFE 只引入很少的参数量并取得了显著提升。CARAFE<sup>[30]</sup> 利用输入特征图来预测上采样核，每个上采样核的位置不同，然后根据预测的上采样核进行特征重组。

CARAFE 具体流程如图 5 所示，由上采样核预测模块和特征重组模块两部分组成。通道压缩器对特征图进行  $1 \times 1$  的卷积压缩，特征图尺寸变为  $H \times W \times C_m$ ，再由内容编码器和上采样核预测器进行特征重组计算，得到上采样核尺寸为  $\sigma H \times \sigma W \times k_{up}^2$ 。上采样核通过 Softmax 函数进行归一化，使得卷积核权重和为 1。特征重组模块分别将每个位置对应映射回输入特征图，取出中心  $k_{up} \times k_{up}$  区域并与预测出该点的上采样核做点积运算，得到尺寸为  $\sigma H \times \sigma W \times C$  的输出特征图。

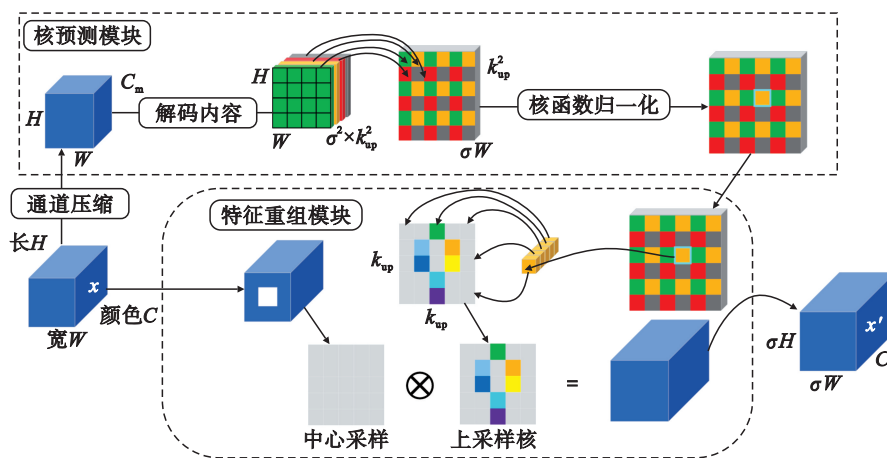


图 5 CARAFE 结构图

Fig. 5 Structure diagram of CARAFE

### 1.2.5 SimAM 注意力机制

注意力机制实质上是一种分配机制，能有效增强神经网络的提取能力，可根据对象的重要程度重新分配权重，在数理统计、语音识别和翻译

等领域有广泛应用<sup>[31-32]</sup>。目前主流注意力机制主要采用增加特征通道和特征空间来提高模型的性能，但通过深化网络结构会加大模型的运算量，若超过网络承载运算量会使模型检测效果受

到影响<sup>[26]</sup>.

在现场收集的异物图像中,图片的分辨率为 2 448 像素×2 048 像素.图像中异物占比相对较少,而无效信息较多,经过多次卷积迭代会产生大量冗余信息,最终淹没目标特征信息,致使模型精度降低<sup>[24]</sup>,为此引入了 SimAM 注意力机制.

SimAM 注意力机制是一个概念简单但有效的注意力模块.相比于现有的通道注意力和空间注意力机制,SimAM 直接推断特征图中三维的注意力权重而且不新增参数<sup>[33]</sup>,如图 6 所示. SimAM 基于神经科学理论,从  $H, W, C$  3 个维度通过优化能量函数来查找每个神经元,并实现将代码控制在 10 行以内. SimAM 的另一优点是大部分运算符是基于定义的能量函数解选择的,避免了过多的结构调整,在改善卷积网络的表征能力上具有灵活性和有效性.

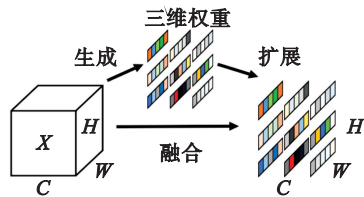


图 6 SimAM 结构图<sup>[33]</sup>

Fig. 6 Structure diagram of SimAM<sup>[33]</sup>

### 1.2.6 SIoU 损失函数

CIoU 损失函数未考虑到真实框和预测框的角度偏差,主要依赖于目标框回归指标的聚合. CIoU 中预测框宽、高相对于长宽比的梯度是一对相反数,因此宽和高不能同时增大或减小,存在收敛速度慢的情况. YOLOv8n 模型中引入 SIoU 损失函数可避免这一问题,SIoU 考虑到真实框和

预测框之间的角度,重新校正角度惩罚程度,可使预测框快速漂移到附近轴<sup>[34-35]</sup>,只需要回归一个坐标,有效减少自由度数量.

SIoU 损失函数计算方法如式(2)所示.

$$L_{SIoU} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (2)$$

式中:IoU 为交并比; $\Delta$  为距离损失,其计算方式如式(3)所示; $\Omega$  为形状损失,其计算方式如式(4)所示.

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

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

$$\gamma = 2 - \Delta \quad (5)$$

$$\Delta = 1 - 2 \times \sin^2 \left( \arcsin(x) - \frac{\pi}{4} \right) \quad (6)$$

其中: $\Delta$  为角度损失,计算方式如式(6)所示; $\theta$  为对形状损失的关注程度; $\omega_t$  为预测框与真实框的差值比; $\gamma$  为调节因子; $\rho_t$  为距离损失权重; $x$  和  $y$  为坐标方向; $w$  和  $h$  代表宽高; $t$  为预测框的数量.

## 2 改进模型实验结果与分析

### 2.1 图像采集与数据集搭建

脱水蔬菜产品种类繁多且不同品种蔬菜差异很大,图像识别模型会受到不同背景影响,导致检测效果有明显差异.为此本文以葱为研究对象进行异物图像采集,脱水葱产品规格为 5~15 mm 厚的葱粒.蔬菜异物图像采集地点位于江苏某脱水蔬菜生产厂,使用工业相机共采集 9 361 张图像,分辨率为 2 448 像素×2 048 像素,其中包含彩条、杂草、头发、塑料片等外源性异物,如图 7 所示.

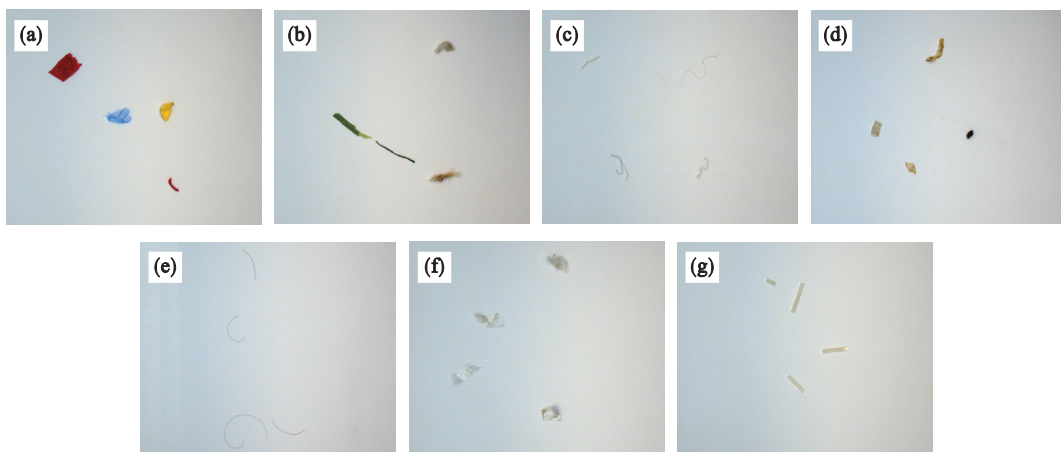


图 7 异物图像

Fig. 7 Foreign object images

(a)一彩条;(b)一杂草;(c)一棉线;(d)一带泥蔬菜;(e)一头发;(f)一块状塑料;(g)一条状塑料.

在模型训练前,使用标注工具 LabelImg 对上述 9 361 张蔬菜异物图像进行人工标注,采用 PASCAL VOC 标注格式.蔬菜异物图像的训练集、测试集和验证集按照 8:1:1 的比例进行划分,数据集构建方法如表 1 所示.

### 2.2 模型评价指标

目标检测模型性能评估的关键指标分别为精确率  $P$ 、召回率  $R$ 、平均精度均值 mAP、损失值  $L_{loss}$ 、每秒模型处理帧数 FPS 与模型大小(weights).精确率  $P$  表示在所有被识别标记的目标中正确目标所占比率,召回率  $R$  表示在所有异物中正确识别出来的比率<sup>[16]</sup>.精确率  $P$  和召回率  $R$  是一对矛盾的度量标准,相互此消彼长.mAP 是精确率与召回率曲线与坐标轴所围面积的平均值,是评价目标检测模型性能的关键指标,mAP 越大代表模型识别效果越好<sup>[36-38]</sup>.综合分析上述指标可以证明模型识别性能.

### 2.3 实验环境与参数设置

本文训练测试的平台硬件配置环境:CPU 采用 Intel(R) Core(TM) i9-13900KF 处理器(主频 3.00 GHz),GPU 为 NVIDIA GeForce RTX 3080 显卡,内存 32 GB,主板为华硕 PRIME Z790-P.操

作系统为 Windows 11,软件平台使用 PyCharm,模型训练环境基于 PyTorch 1.11.0 框架,Python 版本为 3.8,CUDA 版本为 11.3,cuDNN 版本为 7.6.训练参数设置为 300 个 epoch,batch size 设置为 32,输入的图片分辨率为 640 像素×640 像素,其他参数设置不变.

表 1 蔬菜异物数据集

Table 1 Foreign object dataset of vegetables				
异物类别	训练集	验证集	测试集	总计
彩条	1 481	186	186	1 853
头发	1 444	181	181	1 806
杂草	1 680	210	210	2 100
塑料片	1 040	130	130	1 300
带泥蔬菜	401	50	50	501
棉线	1 441	180	180	1 801
总计	7 487	937	937	9 361

### 2.4 训练与检测结果

从图 8 中可以看出,损失函数整体波动较小,在前 30 个 epoch 训练中收敛速度快,损失值达到了 0.039;经过 300 个 epoch 训练后,损失值收敛到 0.025.最终结果的精确率  $P$  为 96.8%,召回率  $R$  为 94.7%,mAP 为 97.1%.

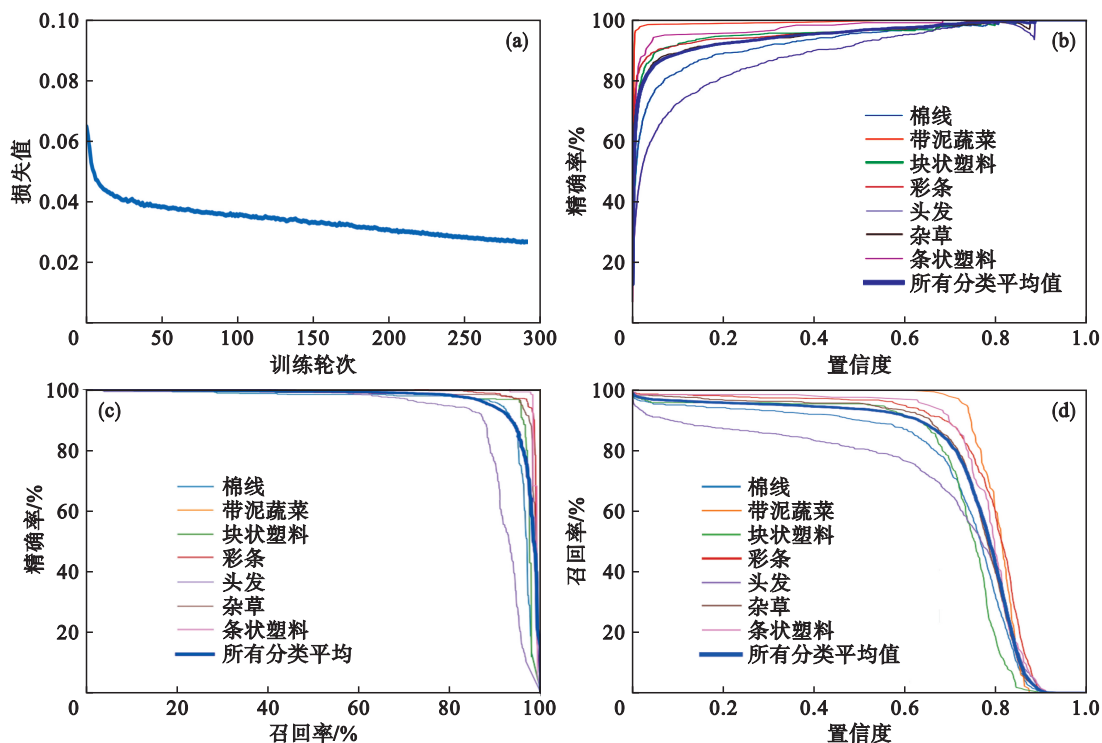


图 8 模型训练过程

Fig. 8 Model training process

(a)—损失值;(b)—精确率;(c)—精确率与召回率曲线;(d)—召回率.

使用训练好的改进模型 YOLOv8n-BCS 对测试集中的 937 张异物图像进行识别,部分异物

识别结果如图 9 所示(图中数字代表置信度).对于局部遮挡的异物图像和多目标的异物图像,

YOLOv8n-BCS 模型能够准确检测异物信息.

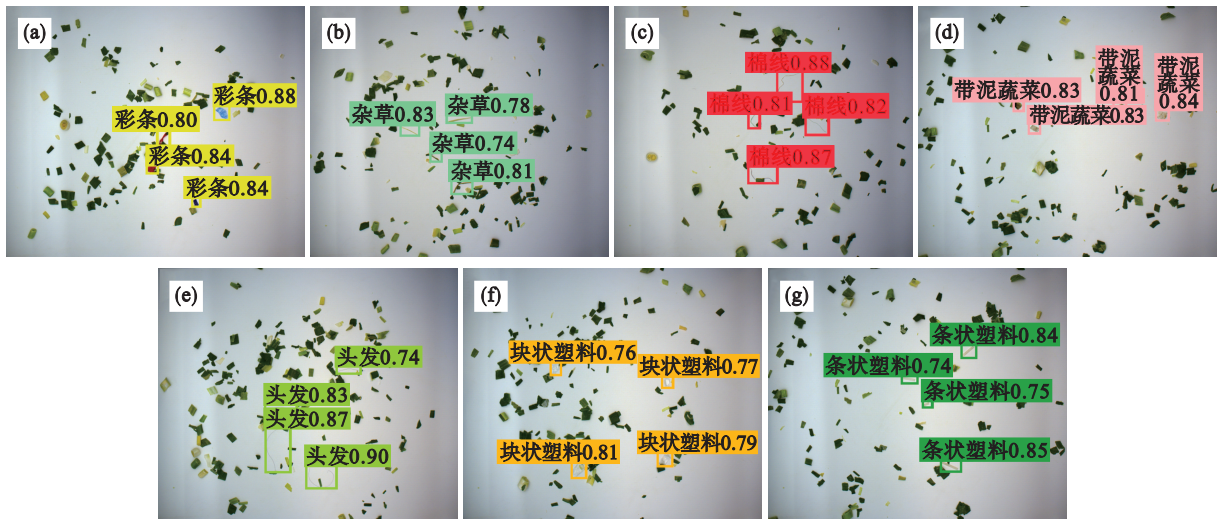


图 9 异物图像识别结果

Fig. 9 Image recognition results of foreign objects

(a)一彩条; (b)一杂草; (c)一棉线; (d)一带泥蔬菜; (e)一头发; (f)一块状塑料; (g)一条状塑料.

### 2.5 脱水蔬菜异物数据集测试结果分析

为验证 YOLOv8n-BCS 模型的性能, 选取 YOLOv3<sup>[39]</sup>、YOLOv5s 作为对照组进行比较. 其中, YOLOv8n-BCS 与 YOLOv3, YOLOv5s, YOLOv8n 模型的 mAP 训练曲线变化如图 10 所示. YOLOv8n-BCS 与其他模型检测精度和性能的实验结果如表 2 所示.

从图 10 中可以看出, 相比 YOLOv3 模型, YOLOv5s, YOLOv8n 和 YOLOv8n-BCS 得益于轻量化网络架构, 3 个模型在收敛速度上有明显优势. 经过 150 轮次迭代训练后, YOLOv3, YOLOv5s 模型与 YOLOv8n-BCS 和 YOLOv8n 模型都趋于收敛, 但模型最终平均精度均值仍存在差距. 由表 2 可知, 在检测脱水蔬菜异物方面, YOLOv8n-BCS 相比 YOLOv3 和 YOLOv5s, 精确率  $P$  分别提高 3.7% 和 0.8%, 召回率  $R$  分别提高 4.0% 和 0.8%, 调和平均值  $F1$  分别提升 3.8%, 0.8%, 平均精度均值

分别提高 3.8% 和 0.8%. YOLOv8n-BCS 与 YOLOv8n 相比, 精确率提升 0.7%, 召回率  $R$  提升 0.5%, 调和平均值  $F1$  提升 0.6%, 平均精度均值提高 0.6%, FPS 性能提升约 16.4%, 模型存储空间略有降低. 综合各项指标看, YOLOv8n-BCS 模型性能均优于 YOLOv3, YOLOv5s 和 YOLOv8n.

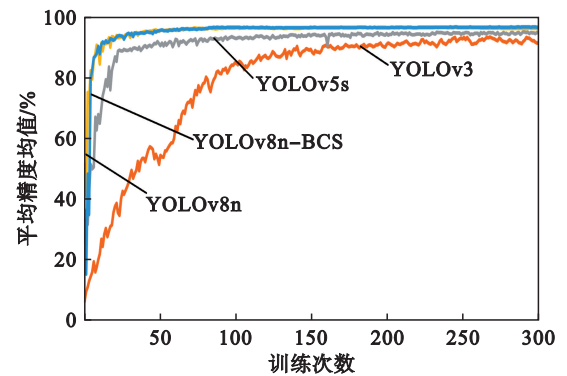


图 10 mAP 曲线对比

Fig. 10 Comparison of mAP curves

表 2 不同模型性能对比

Table 2 Performance comparison of different models

模型参数	YOLOv3	YOLOv5s	YOLOv8n	YOLOv8n-BCS
$P/\%$	93.1	96.0	96.1	96.8
$R/\%$	90.7	93.9	94.2	94.7
$F1/\%$	91.9	94.9	95.1	95.7
mAP/%	92.3	96.3	96.5	97.1
FPS/( $f \cdot s^{-1}$ )	143.483	224.116	198.847	231.476
模型大小/MB	235	18.6	6.2	6.1

## 2.6 消融实验与分析

为验证所提出的 YOLOv8n-BCS 模型的有效性,在保持相同平台和参数设置的情况下,在

测试集上对模型进行消融实验,使用精确率  $P$ 、召回率  $R$ 、平均精度均值 mAP 和调和平均值  $F1$  作为评价指标,实验结果如表 3 所示。

表 3 模型消融实验  
Table 3 Model's ablation experiment

序号	基础模型	ShuffleNetV2+ BoTNET	CARAFE	SimAM	SIoU	$P/\%$	$R/\%$	$F1/\%$	mAP/%	FPS $f \cdot s^{-1}$
模型 1	YOLOv8n	×	×	×	×	96.1	94.2	95.1	96.5	198.847
模型 2	YOLOv8n	√	×	×	√	96.3	94.2	95.2	96.7	210.452
模型 3	YOLOv8n	×	×	√	√	96.1	94.0	95.0	96.5	202.926
模型 4	YOLOv8n	×	√	×	√	96.1	94.3	95.2	96.4	205.424
模型 5	YOLOv8n	×	√	√	√	96.3	94.5	95.4	96.6	208.867
模型 6	YOLOv8n	√	√	×	√	96.7	93.5	95.1	96.6	205.723
模型 7	YOLOv8n	√	×	√	√	96.4	94.1	95.2	96.7	212.428
模型 8	YOLOv8n	√	√	√	√	96.8	94.7	95.7	97.1	231.476

从表 3 可知,模型 1 是对照组 YOLOv8n 模型. BoTNet 与 SIoU 经过验证可以提升模型检测效果<sup>[24]</sup>,本文不单独对 BoTNet 和 SIoU 进行实验. ShuffleNetV2 和 BoTNet 结构都位于主干网络中,因此将两模块一起进行实验,模型 2 至模型 8 均引入 SIoU. 与对照组相比,模型 2 引入 ShuffleNetV2, BoTNet 和 SIoU 后,其精确率  $P$  提升 0.2%、召回率  $R$  保持不变、调和平均值  $F1$  提升 0.1%,平均精度均值提高 0.2%,FPS 提升约 5.8%. 这证明模型引入 ShuffleNetV2 和 BoTNet 能有效获取目标特征信息,并提升模型对异物的检测效果. 与对照组相比,模型 3 引入 SimAM 和 SIoU 后,其精确率  $P$  和平均精度均值 mAP 保持不变,召回率  $R$  与调和平均值  $F1$  分别降低了 0.2% 和 0.1%,FPS 提升约 2.1%. 实验证明引入 SimAM 能提升模型网络识别速度,但单独使用会稍微降低异物的检测性能. 与对照组相比,引入 CARAFE 和 SIoU 的模型 4 精确率  $P$  保持不变、召回率  $R$  和调和平均值  $F1$  提高 0.1%,平均精度均值 mAP 降低 0.1%,FPS 提升约 3.3%. 对比模型 1~模型 4 可知,模型 2 的性能呈小幅提升,模型 3 和模型 4 部分性能出现了下降, SimAM, CARAFE, ShuffleNetV2+BoTNet 和 SIoU 结合均可以提升模型的检测速度,其中 ShuffleNetV2+BoTNet 模块的性能提升最为显著.

与对照组相比,引入 CARAFE, SimAM 和 SIoU 的模型 5 精确率  $P$  提高 0.2%,召回率  $R$  提高 0.3%,调和平均值  $F1$  提高 0.3%,平均精度均值 mAP 提高 0.1%,FPS 提升约 5.0%. 与对照组相比,

模型 6 引入 CARAFE、ShuffleNetV2、BoTNet 和 SIoU,其精确率  $P$  提升 0.6%,召回率  $R$  降低 0.7%,调和平均值  $F1$  保持不变,平均精度均值 mAP 提高 0.1%,FPS 提升约 3.5%. 与对照组相比,模型 7 引入 SimAM, ShuffleNetV2, BoTNet 和 SIoU,其精确率  $P$  提升 0.3%、召回率  $R$  降低 0.1%、调和平均值  $F1$  提高 0.1%,平均精度均值 mAP 提高 0.2%,FPS 提升约 6.8%. 与对照组相比,模型 8 引入 ShuffleNetV2, BoTNet, CARAFE, SimAM 和 SIoU,其精确率  $P$  提高 0.7%、召回率  $R$  提高 0.5%、调和平均值  $F1$  提高 0.6%,平均精度均值 mAP 提高 0.6%,FPS 提升约 16.4%. 对比模型 5~模型 8 可知,相比模型 2~模型 4 单独应用 ShuffleNetV2+BoTNet, CARAFE, SimAM 结构,相互组合可以进一步提升检测速度和精度. ShuffleNetV2+BoTNet 与 CARAFE 结合时,精确率提升明显,但召回率稍有降低. YOLOv8n-BCS(即模型 8)同时引入所有结构,所有性能指标均为最优.

为进一步测试改进模型效果,从现场采集照片中选择两张复杂背景多目标图片,使用 YOLOv8n-BCS 与性能较好的模型 2 和模型 7 进行对比,验证结果如图 11 所示.

由图 11 可知,对于含有复杂背景的小样本异物图像,YOLOv3 存在漏检和定位不准确的现象; YOLOv5s 比 YOLOv3 稍好,但对小目标检测效果较差; YOLOv8n 与模型 2、模型 7、YOLOv8n-BCS 模型均能识别图像中的异物且定位准确. 与其他模型相比, YOLOv8n-BCS 模型在异物检测效果上有所提高.

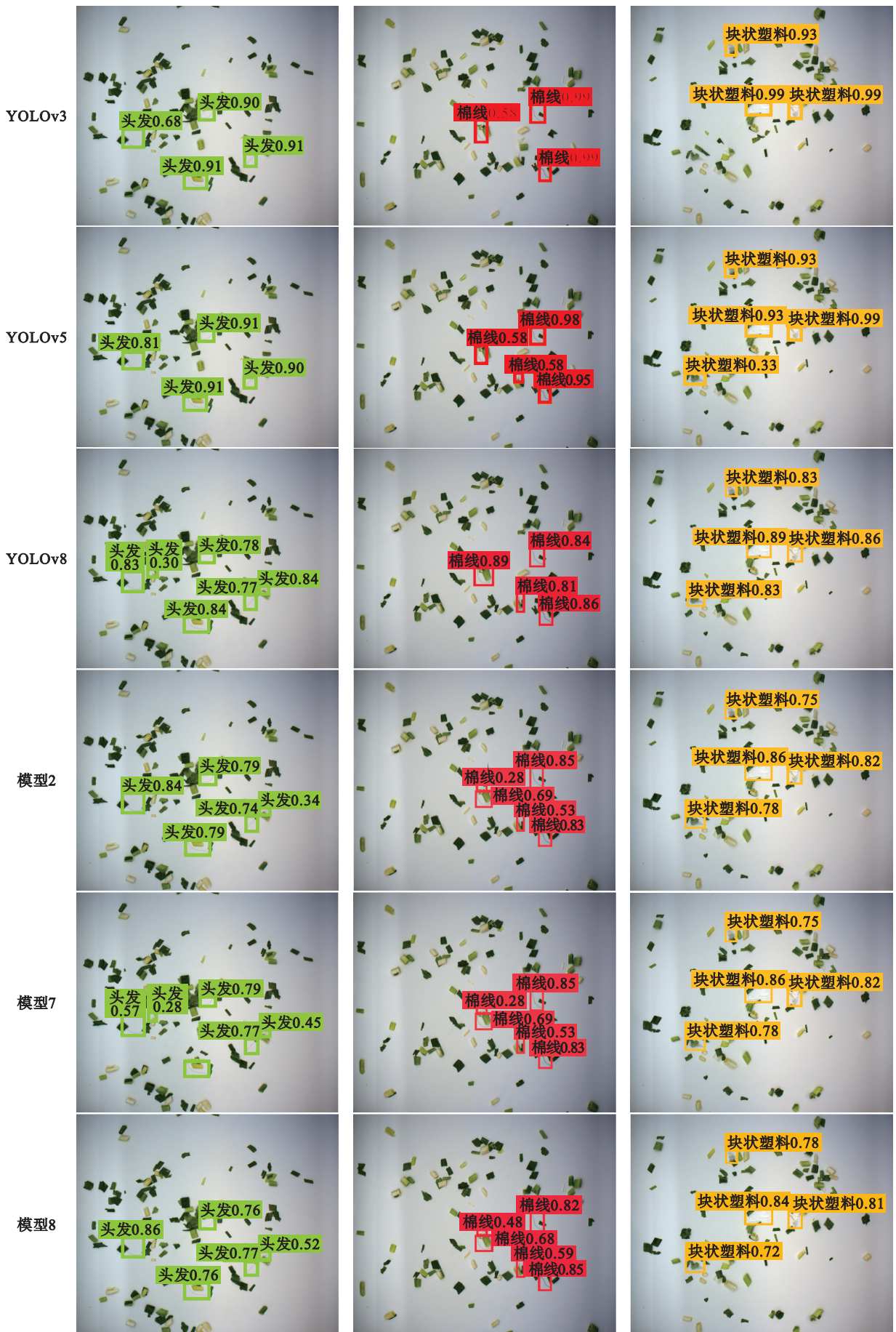


图 11 不同模型检测效果对比

Fig. 11 Comparison of detection effects of different models

### 3 现场应用

本文所开发的脱水蔬菜智能检测分拣系统已投入使用,其系统硬件由皮带输送机、CMOS 相机、照明系统和 Delta 机器人等组成,系统实物如图 12 所示.

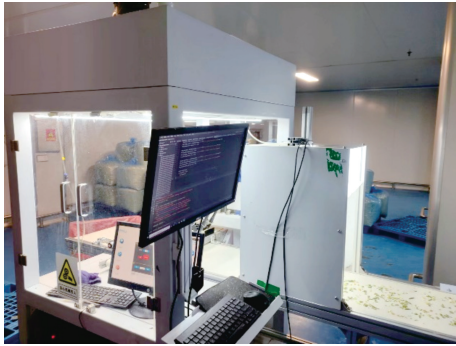


图 12 脱水蔬菜异物分拣系统  
Fig. 12 Foreign object sorting systems for dehydrated vegetables

检测分拣系统使用时,Delta 机器人采用世界坐标系,CMOS 相机采用相机坐标系,CMOS 相机的工作区域使用像素坐标系.通过相机视角和安装高度,像素坐标系 $(u, v)$ 坐标可以转化为相机坐标系坐标 $(X_c, Y_c, Z_c)$ .结合分拣设备和相机的相对位置,相机坐标系坐标可计算得到世界坐标系 $(X_w, Y_w, Z_w)$ 的坐标.最终实现像素坐标系与世界坐标系的相互转换,具体原理如图 13 所示.检测分拣系统中,机器人与视觉系统采用 eye-to-hand 结构,视觉部分捕获异物目标后,目标经皮带传送至 Delta 机器人进行分拣.当多个目标同时出现时,系统进行优先级排序,Delta 机器人会根据优先级依次分拣异物<sup>[40]</sup>.

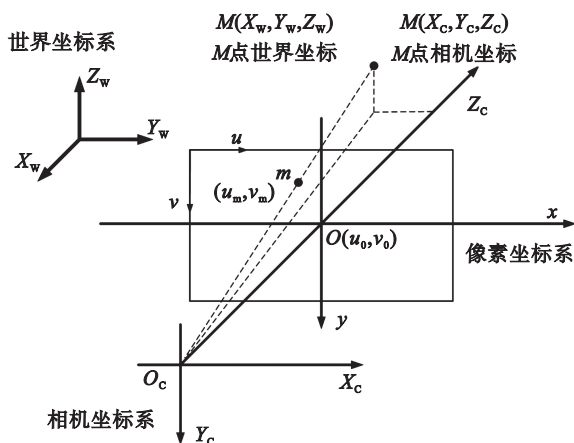


图 13 相机成像模型

Fig. 13 Imaging model of camera

脱水蔬菜智能检测分拣系统可以辅助工人分拣异物以提高生产效率,帮助企业降本增效.脱水蔬菜生产过程需要经历 3 道人工分拣工序,脱水蔬菜智能检测分拣系统可替代前两道人工分拣.该系统 1 h 可以分拣 8 kg 脱水蔬菜,相较于工人平均分拣速度提高了 4 倍,现场分拣效率提高约 50%,目前已应用于脱水葱、玉米和豌豆的生产中.

### 4 结 论

1) 改进模型 YOLOv8n-BCS 引入 ShuffleNetV2 网络,BoTNet 网络,上采样算子 CARAFE 和 SimAM 注意力机制,可提升小目标检测与特征提取能力;使用 SIoU 损失函数可提高收敛性能.

2) 相比于 YOLOv3, YOLOv5s, YOLOv8n 模型, YOLOv8n-BCS 模型在检测速度和特征提取能力方面均有更好表现.对脱水蔬菜异物进行检测时, YOLOv8n-BCS 模型的检测精确率  $P$  为 96.8%, 召回率  $R$  为 94.7%, 调和平均值  $F1$  为 95.7%, 平均精度均值 mAP 为 97.1%, 模型存储空间为 6.1MB.

3) 结合实际应用,设计并开发了脱水蔬菜智能检测分拣系统,该系统可帮助脱水蔬菜企业降本增效,辅助工人分拣葱、玉米和豌豆中的异物,使分拣效率提升约 50%.

#### 参考文献:

- [1] 单杨. 中国果蔬加工产业现状及发展战略思考[J]. 中国食品学报, 2010, 10(1): 1-9.  
(Shan Yang. Current situation and development strategic consideration of the fruits & vegetables processing industry in China [J]. *Journal of Chinese Institute of Food Science and Technology*, 2010, 10(1): 1-9.)
- [2] 陈丰, 陈晓波, 王瑞鸿. 脱水蔬菜加工技术及其研究进展[J]. 农产品加工, 2013(9): 6-7.  
(Chen Feng, Chen Xiao-bo, Wang Rui-hong. Processing technology and research progress of dehydrated vegetables [J]. *Farm Products Processing*, 2013(9): 6-7.)
- [3] 孙小静, 刘军, 邹宇晓, 等. 脱水蔬菜加工过程中品质变化的研究进展[J]. 食品工业科技, 2014, 35(20): 388-392.  
(Sun Xiao-jing, Liu Jun, Zou Yu-xiao, et al. Research progress in the quality change of dehydrated vegetables in processing [J]. *Science and Technology of Food Industry*, 2014, 35(20): 388-392.)
- [4] Bhargava A, Bansal A, Goyal V. Machine learning-based detection and sorting of multiple vegetables and fruits [J]. *Food Analytical Methods*, 2022, 15(1): 228-242.
- [5] Kawamura S, Natsuga M, Takekura K, et al. Development of an automatic rice-quality inspection system [J]. *Computers and Electronics in Agriculture*, 2003, 40(1/3):

- 115–126.
- [6] Zhao Y S, Gong L, Huang Y X, et al. Robust tomato recognition for robotic harvesting using feature images fusion[J]. *Sensors*, 2016, 16(2): 173.
- [7] Hayashi S, Shigematsu K, Yamamoto S, et al. Evaluation of a strawberry-harvesting robot in a field test [J]. *Biosystems Engineering*, 2010, 105(2): 160–171.
- [8] Pandey V K, Srivastava S, Dash K K, et al. Machine learning algorithms and fundamentals as emerging safety tools in preservation of fruits and vegetables: a review[J]. *Processes*, 2023, 11(6): 1720.
- [9] Ji W, Qian Z J, Xu B, et al. A nighttime image enhancement method based on Retinex and guided filter for object recognition of apple harvesting robot [J]. *International Journal of Advanced Robotic Systems*, 2018, 15(1): 1–12.
- [10] Agarwal M, Gupta S K, Biswas K K. Development of efficient CNN model for tomato crop disease identification [J]. *Sustainable Computing: Informatics and Systems*, 2020, 28: 100407.
- [11] 王红军, 熊俊涛, 黎邹邹, 等. 基于机器视觉图像特征参数的马铃薯质量和形状分级方法[J]. 农业工程学报, 2016, 32(8): 272–277.  
(Wang Hong-jun, Xiong Jun-tao, Li Zou-zou, et al. Potato grading method of weight and shape based on imaging characteristics parameters in machine vision system [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2016, 32(8): 272–277.)
- [12] Blanc P. Unit for sorting and packaging products capable of being hung on a hooking member for the purpose of their conveyance, such as bunches of fruits, in particular table grapes or truss tomatoes: US6957940B2[P]. 2005–10–25.
- [13] 叶晋涛, 王运祥, 杨杰, 等. 哈密瓜颜色特征提取及成熟度分级的研究[J]. 石河子大学学报(自然科学版), 2016, 34(1): 106–111.  
(Ye Jin-tao, Wang Yun-xiang, Yang Jie, et al. Research on maturity and classification of Hami melon based on color space feature extraction [J]. *Journal of Shihezi University (Natural Science)*, 2016, 34(1): 106–111.)
- [14] 魏康丽, 王振杰, 孙柯, 等. 基于计算机视觉对苹果脆片外观品质分级[J]. 南京农业大学学报, 2017, 40(3): 547–555.  
(Wei Kang-li, Wang Zhen-jie, Sun Ke, et al. External quality classification of apple chips based on computer vision [J]. *Journal of Nanjing Agricultural University*, 2017, 40(3): 547–555.)
- [15] Baigvand M, Banakar A, Minaei S, et al. Machine vision system for grading of dried figs [J]. *Computers and Electronics in Agriculture*, 2015, 119: 158–165.
- [16] 马超伟, 张浩, 马新明, 等. 基于改进 YOLOv8 的轻量化小麦病害检测方法[J]. 农业工程学报, 2024, 40(5): 187–195.  
(Ma Chao-wei, Zhang Hao, Ma Xin-ming, et al. Lightweight wheat disease detection method based on improved YOLOv8[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2024, 40(5): 187–195.)
- [17] 王春梅, 刘欢. YOLOv8–VSC: 一种轻量级的带钢表面缺陷检测算法[J]. 计算机科学与探索, 2024, 18(1): 151–160.  
(Wang Chun-mei, Liu Huan. YOLOv8–VSC: lightweight algorithm for strip surface defect detection [J]. *Journal of Frontiers of Computer Science and Technology*, 2024, 18(1): 151–160.)
- [18] 胡杰珍, 杨靖荣, 邓培昌, 等. 基于 YOLO 算法的金属表面腐蚀图像识别分析系统[J]. 材料保护, 2025, 58(9): 124–133.  
(Hu Jie-zhen, Yang Jing-rong, Deng Pei-chang, et al. Image recognition and analysis system for metal surface corrosion based on YOLO algorithm [J]. *Materials Protection*, 2025, 58(9): 124–133.)
- [19] Ye R, Shao G Q, He Y, et al. YOLOv8–RMDA: lightweight YOLOv8 network for early detection of small target diseases in tea[J]. *Sensors*, 2024, 24(9): 2896.
- [20] 王少聪, 王会强, 丁小明, 等. 基于改进 YOLO v8n 的草莓温室飞虫情监测方法[J]. 农业工程学报, 2025, 41(17): 184–193.  
(Wang Shao-cong, Wang Hui-qiang, Ding Xiao-ming, et al. Pest monitoring in strawberry greenhouses using improved YOLO v8n [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2025, 41(17): 184–193.)
- [21] 白凯, 张玉杰, 苏邓文, 等. 基于改进 YOLO v8n 的花生叶片病害检测方法[J]. 农业机械学报, 2025, 56(6): 518–526, 564.  
(Bai Kai, Zhang Yu-jie, Su Deng-wen, et al. Peanut leaf disease detection method based on improved YOLO v8n[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2025, 56(6): 518–526, 564.)
- [22] Fang C, Yang X. Lightweight YOLOv8 for wheat head detection[J]. *IEEE Access*, 2024, 12: 66214–66222.
- [23] 刘梦姝, 张春琪, 晁金阳, 等. 基于 YOLO v8n 改进的小麦病害检测系统[J]. 农业机械学报, 2024, 55(S1): 280–287, 355.  
(Liu Meng-shu, Zhang Chun-qi, Chao Jin-yang, et al. Improved wheat disease detection system based on YOLO v8n[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(S1): 280–287, 355.)
- [24] 苑迎春, 耿俊, 许楠, 等. 基于 YOLO v8–TRP 模型的小麦麦穗识别方法研究[J]. 农业机械学报, 2025, 56(11): 499–508.  
(Yuan Ying-chun, Geng Jun, Xu Nan, et al. Wheat ear recognition method using YOLOv8–TRP model [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2025, 56(11): 499–508.)
- [25] Wang C Y, Bochkovskiy A, Liao H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, 2023: 7464–7475.
- [26] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021 [EB/OL]. 2021: arXiv: 2107.08430. <https://arxiv.org/abs/2107.08430>.
- [27] 鲍禹辰, 徐增波, 田丙强. 基于 YOLOv8 改进的服装疵点检测算法[J]. 东华大学学报(自然科学版), 2024, 50(4): 49–56.  
(Bao Yu-chen, Xu Zeng-bo, Tian Bing-qiang. Improved garment defect detection algorithm based on YOLOv8 [J]. *Journal of Donghua University (Natural Science)*, 2024, 50(4): 49–56.)
- [28] Ma N N, Zhang X Y, Zheng H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design [C]//Computer Vision–ECCV 2018. Cham: Springer, 2018: 122–138.
- [29] Srinivas A, Lin T Y, Parmar N, et al. Bottleneck transformers for visual recognition [EB/OL]. (2021–01–21) [2024–03–30]. <https://arxiv.org/abs/2101.11605>
- [30] Wang J Q, Chen K, Xu R, et al. CARAFE: content-aware ReAssembly of FEatures [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, 2020: 3007–3016.