



朱琦,周德强,盛卫锋,等. 基于 DSCS-YOLO 的苹果表面缺陷检测方法[J]. 南京农业大学学报,2024,47(3):592-601.

ZHU Qi,ZHOU Deqiang,SHENG Weifeng,et al. Apple surface defect detection method based on DSCS-YOLO[J]. Journal of Nanjing Agricultural University,2024,47(3):592-601.

## 基于 DSCS-YOLO 的苹果表面缺陷检测方法

朱琦<sup>1,2</sup>,周德强<sup>1,2\*</sup>,盛卫锋<sup>1</sup>,左文娟<sup>1</sup>,朱家豪<sup>1,2</sup>

(1.江南大学机械工程学院,江苏 无锡 214122;2.江苏省食品先进制造装备技术重点实验室,江苏 无锡 214122)

**摘要:**[目的]针对苹果无损检测过程中表面缺陷检测精度低的问题,提出一种基于 DSCS-YOLO 的苹果表面缺陷检测方法。[方法]首先为提高网络对表面缺陷细节特征的提取能力,设计一种基于 Dense 模块以及 SE 模块的深浅特征选择模块 DSCS(deep and shallow feature selection module),采用 DSCS 替换 Backbone 中的 C3 模块,在保留表面缺陷浅层信息的基础上强化对重要特征的学习,并起到削弱冗余特征的作用;针对 Backbone 与 Neck 部分输出信息过多导致的参数耦合问题,利用解耦头原理对 Head 层部分进行分层预测。其次采用 ELU 激活函数改进原有解耦头,简化末端结构,使网络训练更加容易。最后针对表面缺陷标注困难的问题,采用 Wise-IoU 损失函数代替 CIoU 损失函数,为不同质量的标注提供非线性增益,实现网络的动态聚焦学习。[结果]DSCS-YOLO 提高了对小目标的检测能力,在苹果表面缺陷测试集上平均精度均值达到 90.9%,相较于 YOLOv3-tiny、YOLOv5s、YOLOX-s 以及 SSD 分别提高了 4.5%、1.9%、6.3%、16.3%,检测效果最优。同时模型参数量为 9.54 M,推理时间仅为 2.8 ms,检测速度快,满足实际应用需求。[结论]改进后的 DSCS-YOLO 提高了 YOLOv5s 算法的精度,实现了苹果表面缺陷的精准识别。

**关键词:**苹果;表面缺陷;检测;特征融合;动态聚焦;损失函数;YOLOv5s;注意力机制

中图分类号:TP391.4;S238

文献标志码:A

文章编号:1000-2030(2024)03-0592-10

## Apple surface defect detection method based on DSCS-YOLO

ZHU Qi<sup>1,2</sup>,ZHOU Deqiang<sup>1,2\*</sup>,SHENG Weifeng<sup>1</sup>,ZUO Wenjuan<sup>1</sup>,ZHU Jiahao<sup>1,2</sup>

(1.College of Mechanical Engineering,Jiangnan University,Wuxi 214122,China;

2.Jiangsu Key Laboratory of Advanced Food Manufacturing Equipment & Technology,Wuxi 214122,China)

**Abstract:**[Objectives]Aiming at the low detection accuracy of apple surface defects in the process of non-destructive testing,an apple surface defect detection method based on DSCS-YOLO was proposed. [Methods]Firstly,in order to improve the network's ability to extract detailed features of surface defects,a deep and shallow feature selection module DSCS based on Dense and SE modules was designed. The C3 module in Backbone was replaced by DSCS,which strengthened the learning of important features and weakened redundant features on the basis of retaining superficial information of surface defects. In order to solve the problem of parameter coupling caused by excessive output information of Backbone and Neck,the Decoupled Head principle was used to make hierarchical prediction of Head layer. At the same time,the ELU activation function was used to improve the original Decoupled Head,simplify the terminal structure,and make the network training easier. Finally,to solve the difficulty of labeling surface defects,Wise-IoU loss function was used to replace CIoU loss function,which provided nonlinear gain for labeling of different qualities and realized dynamic focused learning of the network. [Results]DSCS-YOLO improved the detection of small targets,and the mean average precision of the apple surface defect test set reached 90.9%,which was 4.5%,1.9%,6.3% and 16.3% higher than that of YOLOv3-tiny,YOLOv5s,YOLOX-s and SSD,respectively,showing the best detection effect. At the same time,the number of model parameters was 9.54 M,and the inference time was only 2.8 ms. The detection speed was fast,which could meet the needs of practical applications. [Conclusions]The improved DSCS-YOLO improved the accuracy of YOLOv5s algorithm and realized the accurate identification of apple surface defects.

**Keywords:**apple;surface defect;detection;feature fusion;dynamic focusing;loss function;YOLOv5s;attention mechanism

苹果是我国重要农作物,其产量与种植面积均为世界第一<sup>[1]</sup>。但由于我国检测与分级等苹果后处理技术方面落后,导致苹果的质量参差不齐,无法为企业和农民带来更高的销售利润。表面缺陷检测是苹果

收稿日期:2023-07-31

基金项目:无锡市科技局资助项目(G20222014)

\*通信作者:周德强,副教授,研究方向为机器视觉与无损检测,E-mail:zhoudeqiang@jiangnan.edu.cn。

质量检测中重要的一环,但目前检测方式仍以人工目测为主,成本高且效率低下,难以适用于大规模生产。因此,实现苹果表面缺陷的自动化、高精度检测对保证苹果质量、提高企业和农民的收益具有重要的实际意义。

近年来,随着科学技术的发展,深度学习逐渐应用于农产品检测中。薛勇等<sup>[2]</sup>提出使用 GoogleNet 模型并采用迁移学习的方式进行训练,实现了苹果表面缺陷检测。王江晴等<sup>[3]</sup>针对现有植物病害识别模型过于复杂的问题,采用 Ghost 模块对 VGG16 进行改进,并引入 Ranger 优化器辅助训练,最终在 PlantVillage 数据集上准确率达到 99.37%。张倩如等<sup>[4]</sup>采用 GhostNet 与迁移学习结合的方式,同时采用 Dropout 层提高网络分类的准确率,实现了小麦秸秆表面显微图像的完整性分类。王大庆等<sup>[5]</sup>基于深度迁移学习理论提出一种两阶段迁移学习的 EfficientNet 模型,两阶段迁移学习的 EfficientNet 有效解决了训练不充分和域迁移等问题,解决了识别复杂背景玉米叶部病害过程中出现的域迁移和小样本问题。以上均为传统深度学习在农业方面的应用,相较于机器学习和传统图像处理算法,其无需针对检测对象设计复杂的特征提取器,且能通过增加数据的方式提高模型的检测精度与鲁棒性,在农业检测领域具有一定的优势。但传统的深度学习主要对图像整体进行判断,却忽视了图像的细节特征,仅能识别单一、简单的对象。而苹果表面缺陷复杂多变,图像细节的丢失会导致精度受到严重影响。

YOLO<sup>[6]</sup>等目标检测网络的提出,使网络能针对标注的图像细节特征进行学习,提高了对复杂目标的检测精度。Liang 等<sup>[7]</sup>首先通过 BiSeNet V2 分割出图像中的苹果缺陷潜在区域,缩小检测范围,再采用剪枝后的 YOLOv4 对其检测,该方法具有较高的识别精度。陶兆胜等<sup>[8]</sup>提出一种基于 YOLOv5 网络模型改进的番茄叶片病害检测模型 YOLOv5s-TLD,模型平均精度均值和召回率较原 YOLOv5s 模型分别提高 1.9 和 2.5 个百分点。徐会杰等<sup>[9]</sup>在 YOLOv3 的 Head 部分增加高分辨率的检测层,提高对小目标的检测精度,再通过 K-means++ 算法计算先验框参数,实现了玉米叶片病虫害的检测。Gai 等<sup>[10]</sup>通过对 YOLOv4 引入 DenseNet 结构,并更改先验框以适应樱桃形状,实现了对不同成熟度樱桃的检测。赵文博等<sup>[11]</sup>采用 Ghost 模块降低 YOLOv5s 模型复杂度,并通过引入 BIFPN、EIoU 以及 Focal loss 损失函数提高了对甘蔗茎节的识别精度。以上研究表明,YOLO 已经被逐渐应用于各种复杂对象的检测之中,其相对于传统的深度学习算法展现出更广阔的应用前景。并且 YOLO 检测算法具有较高的灵活性,研究人员能针对被检测对象的特点灵活改进模型结构或训练策略,提高对检测目标的识别能力,为实现苹果表面缺陷检测提供了解决方案。

综上所述,传统深度学习网络主要是对图像整体特征进行判断,仅能检测简单、单一的对象,而 YOLO 系列模型通过对图像细节进行标注学习,提高了对缺陷等局部特征的检测精度。因此,结合 YOLOv5s 高效、高鲁棒性的优势,同时针对苹果表面缺陷的特点,对网络结构与 IoU 损失函数进行改进,本文提出了一种基于 DSCS-YOLO 的苹果表面缺陷检测方法。该方法提高了 YOLOv5s 模型的检测精度,为苹果无损检测提供了技术支持。

## 1 材料与方 法

### 1.1 数据集的采集与制作

试验采用的苹果种类为红富士苹果,通过设备采集苹果图像样本共 968 张。采集设备为海康 MV-CA013-A0UC 工业相机,分辨率为 1 280×1 024,最大帧率为 201 fps。苹果表面缺陷检测实际上是缺陷的类别检测,通过对图像的局部特征进行缺陷存在性判断,将其划分为有缺陷和无缺陷 2 类。为提高网络对缺陷特征的学习能力,又根据缺陷形态差异将其分为机械损伤和病害损伤 2 种,并使用 Labelimg1.8.6 标注软件将缺陷标注为 YOLO 格式。采用传统数据增强方法,通过 Imgaug0.4.0 对样本进行镜像、高斯模糊、缩放、亮度变化等方式随机融合处理。最后数据集扩充为 5 808 张图像样本,存在苹果表面缺陷 4 104 处,其中机械损伤缺陷 1 602 处,病害损伤缺陷 2 502 处。将数据增强完成后的数据集按照 8-U1-U1 的概率随机抽样,将其分为训练集、验证集、测试集,分别用于模型的训练、选择与测试。数据增强融合处理流程如图 1 所示。

### 1.2 YOLOv5s 检测原理

目前存在的目标检测网络主要分为以 YOLO、SSD<sup>[12]</sup> 为主的单阶段检测网络和以 Faster RCNN<sup>[13]</sup> 系列为主的双阶段检测网络 2 种。YOLOv5s 是 YOLO 系列中经典的目标检测架构,其针对训练的每个阶段均

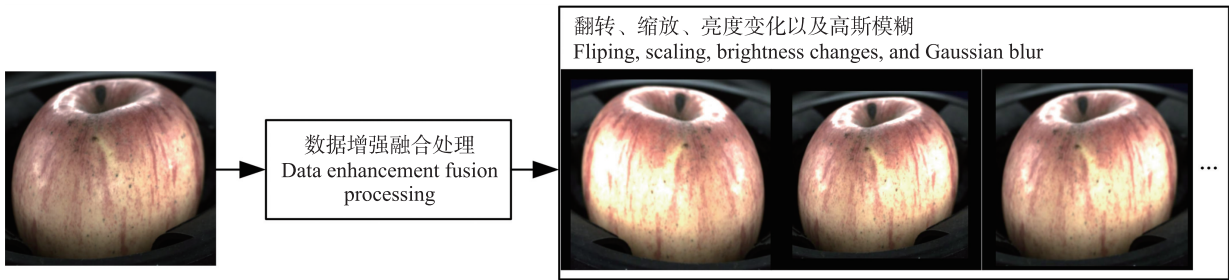


图 1 数据增强融合处理

Fig. 1 Data enhancement fusion processing

提出了高效的处理方法。在预训练阶段,采用 Mosaic 数据增强<sup>[14]</sup>提高了样本多样性。训练阶段,其主要由 Backbone、Neck、Head 三个部分共同完成任务。Backbone 通过 CBS 模块对特征矩阵的大小和通道数进行调整,采用 C3 模块对特征进行提取与变换,SPPF 模块依据特征金字塔原理<sup>[15]</sup>对多尺度特征进行融合。Neck 部分主要对 Backbone 输出特征进行上采样和拼接,实现信息的融合。最后 Head 部分在 Neck 输出特征基础上利用卷积层将其映射成目标框的置信度、坐标信息以及分类信息完成最终预测。整个训练流程采用 CIoU 损失函数约束模型的学习方向,完成目标检测任务。YOLOv5s 整体结构图如图 2 所示。

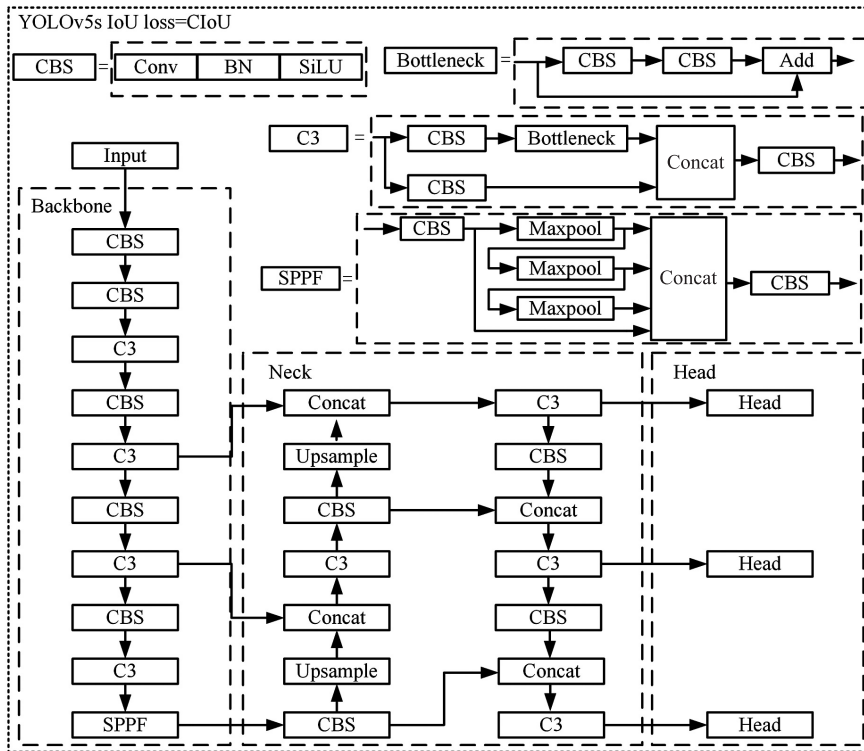


图 2 YOLOv5s 结构图

Fig. 2 YOLOv5s structure diagram

### 1.3 YOLOv5s 改进方法

1) 采用 DSCS 模块改进 Backbone 网络结构。苹果表面缺陷多为中小目标,需要依赖图像细节特征,但原始 Backbone 采用 C3 与 CBS 处理时,并未进行浅层细节特征的保留,导致缺陷检测的精度下降。且 Backbone 内部缺乏信息交流,训练相对更加困难。因此,在 SE 模块<sup>[16]</sup>以及 Dense 模块<sup>[17]</sup>的设计思想上,提出一种深浅特征选择结构 DSCS 保存图像中的细节特征并加强信息的交流与融合。其中 Dense 模块、SE 模块以及 DSCS 模块原理图分别如图 3、图 4、图 5 所示。

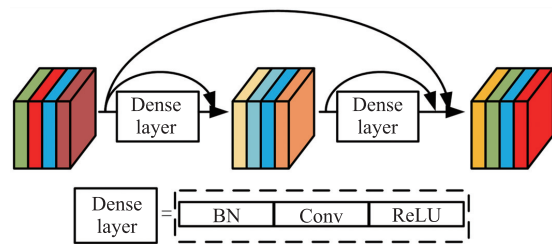


图 3 Dense 模块原理图

Fig. 3 Dense module schematic diagram

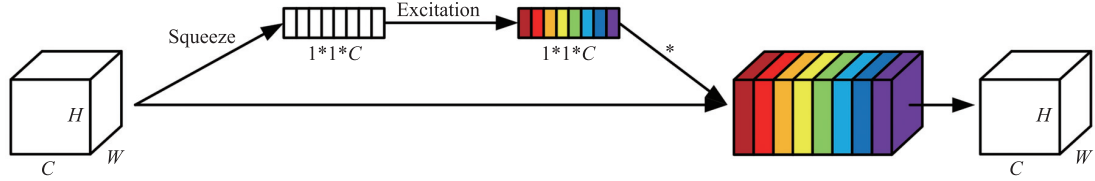


图 4 SE 模块原理图

Fig. 4 SE module schematic diagram

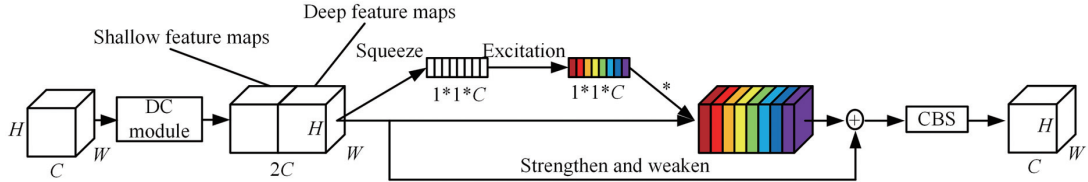


图 5 DSCS 模块原理图

Fig. 5 DSCS module schematic diagram

Dense Layer 由 BN 层、Conv 层和 ReLU 激活函数构成, Dense 模块输出包含所有 Dense Layer 层输出, 并加强了网络间的信息流动。DSCS 中的 DC 模块(Dense CBS)将 Dense 模块中 Dense Layer 层替换为 CBS 层, 在保留 Dense 模块优势的同时也增强了 Backbone 的特征提取能力。设 DC 模块由  $n$  个 DC Layer 层组成,  $y_0$  为输入, 则第  $t$  个 DC Layer 的输出  $Y_t$  如公式(1)所示, DC 模块输出  $C_{out}$  计算方法如公式(2)所示。进而采用 SE 模块对 DC 模块输出特征进行融合与筛选, 最后将处理完成后信息与 DC 模块输出进行 add 操作, 实现了对重要权重的强化并削弱冗余权重的作用。且为保证通道数前后一致, 需要在 DSCS 尾部加入  $1 \times 1$  CBS 模块进行降维处理。则 DSCS 输出  $S_{out}$  计算方法如公式(3)所示。

$$Y_t = \text{CBS}(\text{Concat}(Y_0, Y_1, \dots, Y_{t-1})) \tag{1}$$

$$C_{out} = \text{Concat}(Y_0, Y_1, \dots, Y_n) \tag{2}$$

$$S_{out} = \text{CBS}(\text{SE}(C_{out}) + C_{out}) \tag{3}$$

2) 采用 ELU 激活函数的改进解耦头 EDH(ELU Decoupled Head) 替换原 Head 部分<sup>[18]</sup>。EDH 主要从网络结构上以及激活函数上对原有的 Head 进行改进。当 Head 层输入信息量过于丰富时, 容易出现参数耦合现象。而且在加入 DSCS 结构后, 前序网络的输出更加丰富, 参数耦合更加严重。此时采用 EDH 对 Head 进行改进能促进 DSCS 与 EDH 的协同作用, 使网络精度进一步提升。EDH 模块原理图如图 6 所示。

EDH 采用多层解耦卷积的方式对输出进行解耦处理, 并采用 ELU 激活函数替换原 DH(Decoupled Head)中的 SiLU 激活函数。当前序网络输出出现微小扰动时, 若 Head 部分非线性能力过大, 则会对该扰动放大, 使网络更难以训练。因此, 相较于 Backbone 强大的特征提取能力, 采用 ELU 激活函数可以简化末端复杂度, 使网络容易训练且更加稳定。

3) 采用 Wise-IoU<sup>[19]</sup> 替换 CIoU 损失函数。YOLOv5s 采用 CIoU 计算边界框回归损失, 考虑苹果表面缺陷的复杂性, 在标注时容易产生标注误差, 导致网络往错误的方向收敛。因此, 采用 Wise-IoU 损失能自适

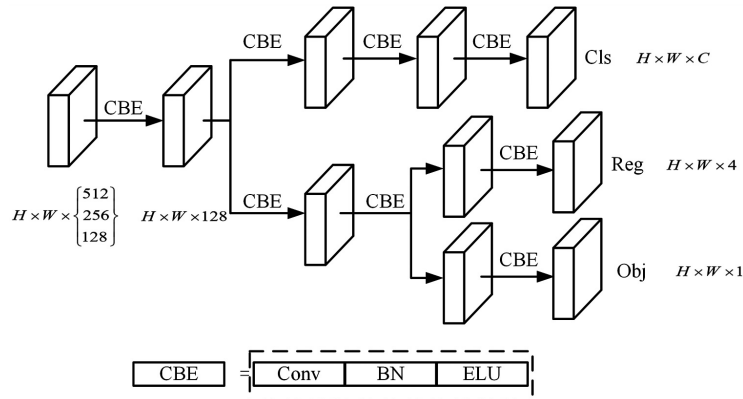


图 6 EDH 模块原理图

Fig. 6 EDH module schematic diagram

应区分标注质量的高低,防止低质量标注产生的有害梯度,实现网络的动态聚焦学习。Wise-IoU 损失值  $L_{WIoU}$  计算方式如式(4)~(7)所示。其中  $\delta$  与  $\alpha$  为引入超参数,  $\overline{L_{IoU}}$  为预测框动态交并比损失均值,  $L_{IoU}^*$  为当前预测框交并比损失值,  $\beta$  为当前预测框的离群度,  $R_{WIoU}$  为预测框与标注的几何差异,  $w_g, h_g$  为最小包围框的宽和高,  $r$  为该框分配所得增益。

$$L_{WIoU} = rR_{WIoU}L_{IoU}^* \tag{4}$$

$$R_{WIoU} = \exp\left(\frac{(x-x_{gt})^2 + (y-y_{gt})^2}{(w_g^2 + h_g^2)}\right) \tag{5}$$

$$r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{6}$$

$$\beta = \frac{L_{IoU}^*}{\overline{L_{IoU}}} \tag{7}$$

由式(4)、(5)可知,  $R_{WIoU}$  与  $L_{IoU}^*$  不受参数影响,当  $r > 1$  时,为标注框损失分配更多的权重,实现了网络的聚焦功能。由式(6)可知,增益  $r$  取决于超参数  $\delta, \alpha$  以及框的离群度  $\beta$ ,设置超参数为不同离群度的框提供非线性增益。由式(7)可知,Wise-IoU 首先计算预测框的动态交并比损失均值  $\overline{L_{IoU}}$ ,并将其作为网络理应得到的损失值。同时计算每个框  $L_{IoU}^*$  与  $\overline{L_{IoU}}$  之比作为离群度  $\beta$ ,  $\beta$  用于评价标注框合理性,当框损失接近  $\overline{L_{IoU}}$  时则说明该框为合理标注框,应该为其分配更多权重。 $\alpha$  则用于控制  $r$  的非线性强度,若固定  $\beta$ ,  $\beta$  的值则表示分配正向增益的离群度范围大小。为直观表示  $\delta$  对  $r$  的影响,现令  $\alpha = 2.2$ ,取不同  $\delta$  值绘制增益  $r$  与离群度  $\beta$  的曲线,由图7可知:当  $\alpha$  固定时,  $\delta$  越大,所得曲线截得  $r=1$  的部分越长,说明为更大范围标注框提供正向增益。综上所述,改进后网络结构图如图8所示。

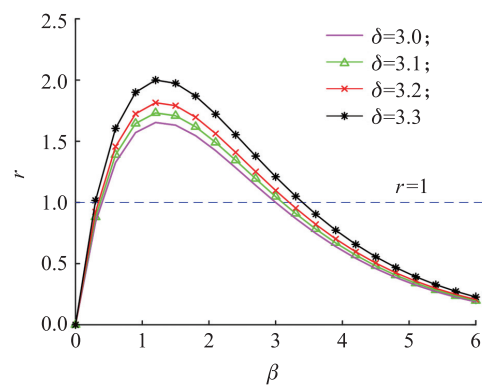


图7 取不同  $\delta$  值时增益  $r$  与离群度  $\beta$  关系的曲线  
Fig.7 Curves of the relationship between gain  $r$  and outlier  $\beta$  at different values of  $\delta$

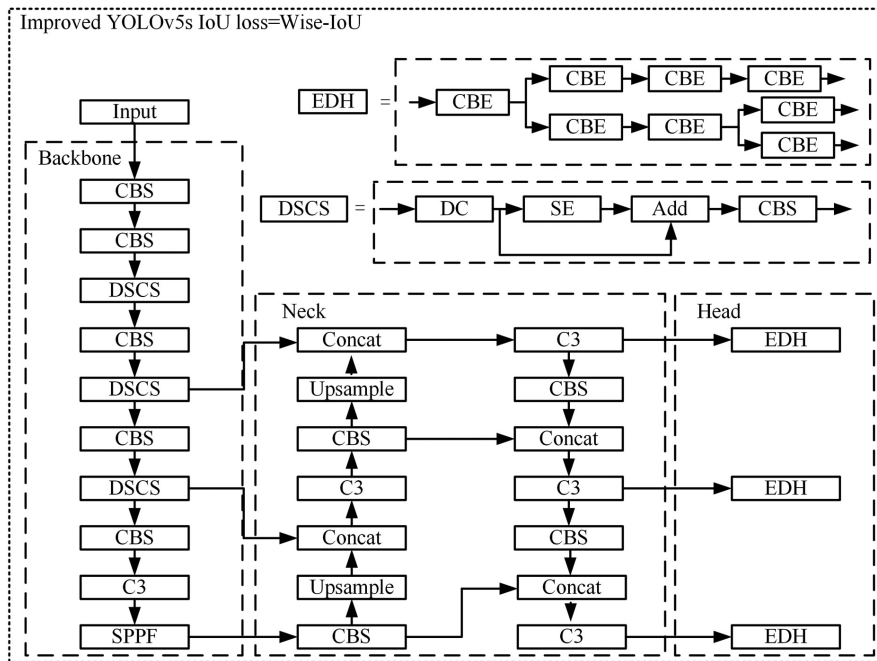


图8 改进后的网络结构图

Fig.8 Improved network structure diagram

1.4 试验环境与评价指标

试验采用 Windows 11 操作系统台式机,处理器型号为 Intel i9-13900K 3.00GHz,机带 RAM 大小为 64 G,

显卡型号为 NVIDIA GeForce RTX 4080。采用 Python3.9.13、Pytorch1.13.1、CUDA11.4、Cudnn8.8.0 深度学习开发包。选择 BatchSize 大小为 32,学习率为 0.01,采用初始学习率系数为 1、最终学习率系数为 0.01 的线性衰减策略进行训练,动量参数为 0.937,权重衰减参数为 0.000 5,迭代 epoch 数为 400。

选择精确率 (precision,  $P$ )、召回率 (recall,  $R$ ) 以及平均精度均值 (mean average precision,  $mAP_{0.5,0.95}$ ) 指标评价模型精度。其中  $P$  表示预测结果为正例的样本中实际正样本所占的比例,  $R$  表示预测结果为正例且实际为正样本占有所有实际正样本的比例, 平均精度  $A$  为  $P$ - $R$  曲线与坐标轴围成的区域面积, 而  $mAP$  为所有检测类别的平均精度  $A$  的平均值,  $mAP_{0.5,0.95}$  表示在不同  $IoU$  阈值 (0.5~0.95) 上的平均  $mAP$ 。  $P$ 、 $R$ 、 $mAP$  计算公式如式 (8)~(11) 所示, 其中  $T_p$  为预测为正例实际为正样本的数量,  $F_p$  为预测为正例实际为负样本数量,  $F_N$  为预测为负例实际上为正样本数量,  $A$  为平均精度。同时选择模型参数量、计算量 GFLOPs 评价模型复杂度, 选择模型对单张图片的推理时间  $t$  评价模型运算速度。

$$P = T_p / (T_p + F_p) \times 100\% \quad (8)$$

$$R = T_p / (T_p + F_N) \times 100\% \quad (9)$$

$$A = \int_0^1 PRdR \quad (10)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N A_i \quad (11)$$

## 2 结果与分析

### 2.1 DSCS 改进试验

在 DSCS 网络结构方面, DSCS 模块由 DC (Dense CBS) 模块和 SE 模块组成。为研究 DSCS 网络结构上的作用, 采用 Dense 模块、SE 模块以及 DSCS 模块试验进行分析。试验结果如表 1 所示: 仅采用 Dense 模块替换 C3 模块会导致  $mAP$  较 YOLOv5s 下降 2%。对模型结构分析可知, Dense 模块虽然保留了 Backbone 中的浅层信息, 但并未进行融合和选择, 且模块中的 ReLU 激活函数会导致神经元失活从而阻碍网络训练。而直接采用 SE 模块时, 由于 C3 模块未保存浅层信息, 阻碍了网络信息流动, SE 不能充分发挥通道信息处理优势, 导致  $mAP$  下降 0.8%。而采用 DSCS 模块对 Backbone 改进, 实现了深层信息以及浅层信息的融合, 增强了重要信息的学习, 综合效果更强, 其  $mAP$  相较于 YOLOv5s 提高 1.2%。

表 1 DSCS 网络结构对比试验

Table 1 DSCS network structure comparison experiments

模型 Model	参数量/M Parameters	计算量 GFLOPs	精确率/% Precision	召回率/% Recall	平均精度均值/% mAP
YOLOv5s	7.04	16.0	98.0	98.3	89.0
YOLOv5s+Dense	7.60	21.1	97.9	98.8	87.0
YOLOv5s+SE	7.05	16.0	97.8	98.8	88.2
YOLOv5s+DSCS	7.67	20.9	98.2	99.0	90.2

在 DSCS 激活函数方面, DSCS 模块由 DC (Dense CBS) 模块和 SE 模块组成, 而其中 DC 模块与 Dense 模块的主要区别在于激活函数的差异。为验证 SiLU 替换 ReLU 对于 DSCS 的效果, 分别采用 DSCS (ReLU) 以及 DSCS 进行试验。如表 2 所示: 采用 DSCS (ReLU) 模块替换 C3 模块, 该方法  $mAP$  较 YOLOv5s 提升 0.8%, 证明了 DSCS (DC+SE) 结构的合理性。但在 Backbone 阶段采用更强大激活函数会有更好效果, SiLU 激活函数非线性能力强于 ReLU 激活函数, 并且解决了 ReLU 存在的死区问题。因此, 采用 DSCS 替换 C3 模块会使  $mAP$  较 YOLOv5s 提升 1.2%, 较 DSCS (ReLU) 提升更高。

表 2 DSCS 激活函数对比试验

Table 2 DSCS activation function comparison experiments

模型 Model	参数量/M Parameters	计算量 GFLOPs	精确率/% Precision	召回率/% Recall	平均精度均值/% mAP
YOLOv5s	7.04	16.0	98.0	98.3	89.0
YOLOv5s+DSCS (ReLU)	7.67	20.9	97.9	98.9	89.8
YOLOv5s+DSCS	7.67	20.9	98.2	99.0	90.2

### 2.2 解耦头改进试验

针对 Head 部分, 设计引入解耦头对比试验。如表 3 所示: 采用 EDH 对 YOLOv5s+DSCS 结构进行改进使  $mAP$  提高 0.6%, 而采用 EDH 对 YOLOv5s 结构进行改进并未得到提升。对模型结构分析可知, YOLOv5s 网络末端不包含大量浅层信息, 参数耦合现象相对较弱, 采用 EDH 进行解耦处理无法得到较好

效果。但采用 DSCS 模块替换主干网络中的 C3 模块时,为后续网络结构提供了大量浅层信息,在网络训练末端特征信息更加丰富,原始 Head 层参数无法满足庞大的信息量,参数耦合现象严重。因此表 3 验证了 EDH 与 DSCS 模块之间存在的协同作用,并使网络 mAP 得到显著提升。

表 3 解耦头对比试验

Table 3 The Decoupled Head comparison experiments

序列 Order number	模型 Model	参数量/M Parameters	计算量 GFLOPs	精确率/% Precision	召回率/% Recall	平均精度均值/% mAP
1	YOLOv5s	7.04	16.0	98.0	98.3	89.0
2	1+EDH	8.91	26.2	98.2	98.6	89.0
3	YOLOv5s+DSCS	7.67	20.9	98.2	99.0	90.2
4	3+EDH	9.54	31.2	98.3	98.9	90.8

由于 EDH 与原解耦头 Decoupled Head (DH) 区别在于激活函数不同,且根据上述表 3 结果,YOLOv5s+DSCS 末端参数耦合现象较 YOLOv5s 强,因此选用 YOLOv5s+DSCS 作为 Baseline 进行不同激活函数试验。试验结果如表 4 所示:由于原 DH 使用的 SiLU 激活函数非线性过强导致末端训练困难,导致网络 mAP 较 Baseline 下降 2.2%。而采用 ReLU 激活函数的 Decoupled head (RDH) 虽然能优化网络训练,但由于其死区的存在也导致 mAP 下降 1.5%。由于 ELU 函数较 SiLU 简单,且函数处处可微,不会存在神经元失活现象,适用于网络末端训练,并使 mAP 提高 0.6%。

表 4 解耦头激活函数对比试验

Table 4 Comparison experiments of decoupled head's activation function

序列 Order number	模型 Model	参数量/M Parameters	计算量 GFLOPs	精确率/% Precision	召回率/% Recall	平均精度均值/% mAP
1	YOLOv5s+DSCS (Baseline)	7.67	20.9	98.2	99.0	90.2
2	1+DH	9.54	31.2	98.6	99.1	88.0
3	1+RDH	9.54	31.2	99.0	99.1	88.7
4	1+EDH	9.54	31.2	98.3	98.9	90.8

### 2.3 Wise-IoU 损失改进试验

依据表 5,设计 Wise-IoU 消融试验验证 Wise-IoU 对每个改进模块的影响。Wise-IoU 通过参数  $\delta$  与  $\alpha$  为不同质量的标注框添加非线性增益,但  $\delta$  与  $\alpha$  参数依赖于网络结构。试验获得  $\delta=3.0$  与  $\alpha=2.2$  时,试验序列 8 达到最优效果,mAP 较未添加 Wise-IoU (序列 7) 提高 0.1%。同时采用相同的  $\delta$  与  $\alpha$  进行试验,序列 2、4 的 mAP 较序列 1、3 分别提升 0.4%、1%。但如序列 6 mAP 反而下降 0.1%。分析网络结构以及 Wise-IoU 参数影响可知,由于 YOLOv5s+DSCS 相对于其他网络而言,其浅层信息量较大且没有解耦头进行处理,末端参数较少导致输出波动较大,合理标注框的离群度范围也会随之增大,此时设置  $\delta=3.0$  无法为所有合理标注框提供正向增益从而导致精度有所下降。由图 7 可知,应该扩大  $\delta$  值才能使 YOLOv5s+DSCS 精度得到提升。

表 5 Wise-IoU 消融试验

Table 5 Wise-IoU ablation experiments

序列 Order number	模型 Model	参数量/M Parameters	计算量 GFLOPs	精确率/% Precision	召回率/% Recall	平均精度均值/% mAP
1	YOLOv5s	7.04	16.0	98.0	98.3	89.0
2	1+Wise-IoU	7.04	16.0	98.0	98.4	89.4
3	YOLOv5s+EDH	8.91	26.2	98.2	98.6	89.0
4	3+Wise-IoU	8.91	26.2	98.4	98.9	90.0
5	YOLOv5s+DSCS	7.67	20.9	98.2	99.0	90.2
6	5+Wise-IoU	7.67	20.9	97.9	98.3	90.1
7	YOLOv5s+DSCS+EDH	9.54	31.2	98.9	99.2	90.8
8	7+Wise-IoU	9.54	31.2	98.5	99.1	90.9

根据表 5 结果,为验证 Wise-IoU 中  $\delta$  参数对 YOLOv5s+DSCS 结构的影响,设计不同  $\delta$  值的 Wise-IoU 对比试验。通过分析 Wise-IoU 计算方法可知,对于网络末端输出波动较大的网络, $\delta$  的最优值应大于网络末端输出波动较小的网络。因此固定  $\alpha=2.2$ ,逐步调整  $\delta$  大小寻找最优值。从表 6 可见:当  $\delta=3.2$  时,mAP 得到提升,较未加入 Wise-IoU 时提高 0.2%,验证了 Wise-IoU 对于 YOLOv5s+DSCS 同样具有适用性。

表6 Wise-IoU 参数对于 YOLOv5s+DSCS 影响试验

Table 6 Experiments for effects of Wise-IoU parameters on YOLOv5s+DSCS

序列 Order number	模型 Model	参数量/M Parameters	计算量 GFLOPs	精确率/% Precision	召回率/% Recall	平均精度均值/% <i>mAP</i>
1	YOLOv5s+DSCS( Baseline)	7.67	20.9	98.2	99.0	90.2
2	1+Wise-IoU( $\delta=3.0$ )	7.67	20.9	97.9	98.3	90.1
3	1+Wise-IoU( $\delta=3.1$ )	7.67	20.9	98.2	98.8	89.2
4	1+Wise-IoU( $\delta=3.2$ )	7.67	20.9	98.7	98.1	90.4
5	1+Wise-IoU( $\delta=3.3$ )	7.67	20.9	98.4	98.2	89.7

为对比 Wise-IoU 较其他损失函数的作用,在 YOLOv5s+DSCS+EDH 基础上分别添加 SIoU、DIoU、AlphaIoU、Wise-IoU 进行对比试验。如表 7 所示:添加 Wise-IoU 的 *mAP* 提升 0.1%,DIoU 未得到提升,SIoU 以及 AlphaIoU 反而下降 0.9%和 0.2%。试验证明,Wise-IoU 实现对不同质量的框提供不同的非线性增益,从一定程度上缓解了标注误差产生的影响,对于复杂多变的苹果缺陷数据集具有一定的适用性。

表7 不同损失函数对比试验

Table 7 Comparison experiments of different loss functions

序列 Order number	模型 Model	参数量/M Parameters	计算量 GFLOPs	精确率/% Precision	召回率/% Recall	平均精度均值/% <i>mAP</i>
1	YOLOv5s+DSCS+EDH	9.54	31.2	98.9	99.2	90.8
2	1+SIoU	9.54	31.2	98.6	98.7	89.9
3	1+DIoU	9.54	31.2	98.4	99.0	90.8
4	1+AlphaIoU	9.54	31.2	97.3	97.8	90.6
5	1+Wise-IoU	9.54	31.2	98.5	99.1	90.9

## 2.4 YOLOv5s 消融试验

为直观表明改进模块的作用,进行了一系列消融试验。从表 8 可见:首先针对 YOLOv5s 网络结构丢失浅层信息的问题,采用 DSCS 替换 Backbone 中 C3 模块,使 *mAP* 提高 1.2%。但添加 DSCS 模块为 Head 部分引入的信息过于丰富,从而加重了末端的参数耦合现象,因此采用使用 ELU 激活函数的解耦头 EDH 对末端参数进行解耦,进一步使 *mAP* 提高 0.6%。最后采用 Wise-IoU 实现网络的动态聚焦学习,缓解了标注误差造成的影响,最终苹果表面缺陷检测 *mAP* 达到 90.9%,较 YOLOv5s 提升 1.9%。

表8 消融试验

Table 8 Ablation experiments

模型 Model	参数量/M Parameters	计算量 GFLOPs	精确率/% Precision	召回率/% Recall	平均精度均值/% <i>mAP</i>
YOLOv5s	7.04	16.0	98.0	98.3	89.0
YOLOv5s+DSCS	7.67	20.9	98.2	99.0	90.2
YOLOv5s+DSCS+EDH	9.54	31.2	98.9	99.2	90.8
YOLOv5s+DSCS+EDH+Wise-IoU (DSCS-YOLO)	9.54	31.2	98.5	99.1	90.9

## 2.5 与不同主流检测模型进行对比

采用同一苹果表面缺陷数据集对不同单阶段网络模型进行训练,选择验证集上最优结果对测试集进行测试,试验结果如表 9 所示。DSCS-YOLO 对苹果缺陷检测的 *mAP* 达到 90.9%,相较于 YOLOv3-tiny、YOLOv5s、YOLOX-s 以及 SSD 分别提高了 4.5%、1.9%、6.3%、16.3%。其中 SSD 精度最低且参数量以及运算量也偏大,而改进的 DSCS-YOLO 仅在增加少量参数和计算量情况下,得到较高的 *mAP*。DSCS-YOLO 单张图片推理时间为 2.8 ms,较原 YOLOv5s 增加 1.1 ms,但总体上仍然能够满足实际应用需求。

表9 不同模型对比试验

Table 9 Comparison experiments with different models

模型 Model	参数量/M Parameters	计算量 GFLOPs	平均精度均值/% <i>mAP</i>	模型推理时间/ms Inference time
YOLOv3-tiny	8.69	13.0	86.4	1.5
YOLOv5s	7.04	16.0	89.0	1.7
YOLOX-s	8.94	26.8	84.6	5.3
SSD	25.42	88.5	74.6	20.8
DSCS-YOLO	9.54	31.2	90.9	2.8

分别采用 YOLOv5s 模型、改进后的 DSCS-YOLO 模型对苹果缺陷数据集进行检测,其结果如图 9 所示。

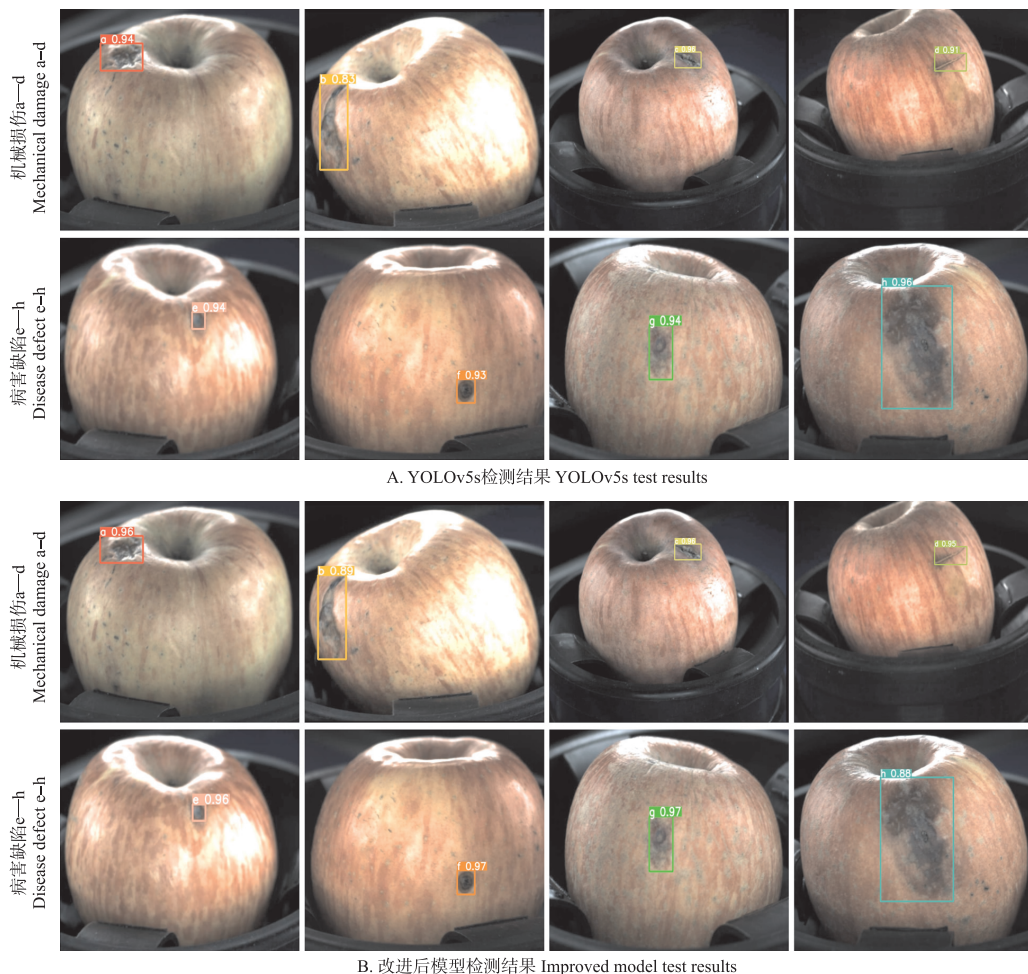


图 9 苹果表面缺陷检测结果对比

Fig. 9 Comparison of detection results of apple surface defects

YOLOv5s 模型与 DSCS-YOLO 模型对苹果缺陷检测均具有良好的识别效果,但改进后的 DSCS-YOLO 由于引入了 DSCS 模块,为网络提供了更多的细节信息,使得网络模型对中小缺陷(缺陷 a—g)的辨识能力强于原 YOLOv5s 模型。对于较大缺陷(缺陷 h)而言,其需要更大的感受野,因此辨识能力有所下降。且根据表 9 结果, DSCS-YOLO 模型在测试集上  $mAP$  较 YOLOv5s 提高 1.9%,说明改进后模型整体检测效果强于 YOLOv5s 模型,能适用于苹果表面缺陷检测。从图 10 的 DSCS-YOLO 与 YOLOv5s 在验证集上的  $mAP$  曲线可知, DSCS-YOLO 训练时在验证集上平均精度均值也较 YOLOv5s 有明显提升。

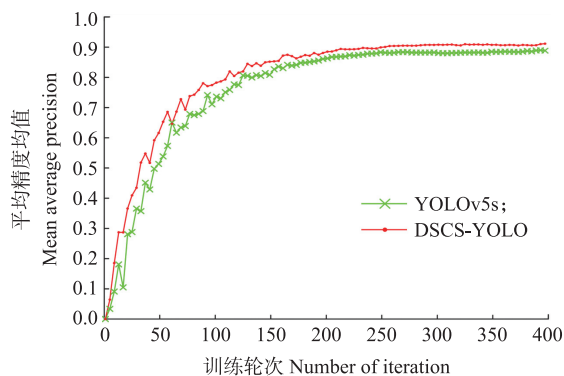


图 10 不同模型训练  $mAP$  曲线

Fig. 10 Training  $mAP$  curve of different models

### 3 结论

本文针对无损检测中苹果表面缺陷检测精度低的问题,结合 YOLO 系列网络的优势,对原有 YOLOv5s 网络结构改进。首先设计了一种深浅特征选择模块 DSCS,保存缺陷中细节信息的同时实现信息的融合与选择,但前序网络信息过于丰富会加剧 Head 部分的参数耦合。在此基础上通过 EDH 对不同信息分组预测,实现对参数解耦的同时显现出与 DSCS 的协同作用,最后采用 WIoU 损失函数自适应区分标注框质量,实现网络动态聚焦学习,提高了缺陷检测的精度。在测试集上试验结果表明,本文所提出的方法对苹果表

面缺陷检测的平均精度均值达 90.9%,较 YOLOv5s 提升 1.9%,实现了苹果表面缺陷的精准识别。

本研究虽取得一定成果,但仍存在一定不足:1)改进后模型精度虽然提高 1.9%,但改进后的模型计算量提升了近一倍,单张图片推理时间增加 1.1 ms。因此,可以从模型轻量化方面对模型进一步完善,减少模型运算量并提高网络运行速度,便于后续进行模型的部署与应用。2)目前苹果种类以及缺陷种类数据集可以进行补充,以便进一步完善模型,应用于更多种类的缺陷检测。

#### 参考文献 References:

- [1] 余梅,文艺,毛敏. 我国苹果对外贸易格局及发展趋势[J]. 中国果树,2022(7):100-104.  
Yu M, Wen Y, Mao M. The pattern and development trend of China's apple foreign trade[J]. China Fruits, 2022(7):100-104(in Chinese).
- [2] 薛勇,王立扬,张瑜,等. 基于 GoogLeNet 深度迁移学习的苹果缺陷检测方法[J]. 农业机械学报,2020,51(7):30-35.  
Xue Y, Wang L Y, Zhang Y, et al. Defect detection method of apples based on GoogLeNet deep transfer learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7):30-35(in Chinese with English abstract).
- [3] 王江晴,冀星,莫海芳,等. 基于轻量化 VGG 的植物病虫害识别[J]. 中国农机化学报,2022,43(4):25-31.  
Wang J Q, Ji X, Mo H F, et al. Plant disease detection based on lightweight VGG[J]. Journal of Chinese Agricultural Mechanization, 2022, 43(4):25-31(in Chinese with English abstract).
- [4] 张倩如,王云飞,吕帅朝,等. 基于改进 GhostNet 的小麦秸秆表皮结构完整性分类方法[J]. 南京农业大学学报,2022,45(4):788-798. DOI:10.7685/jnau.202108015.  
Zhang Q R, Wang Y F, Lü S C, et al. Integrity classification of wheat straw epidermis based on improved GhostNet[J]. Journal of Nanjing Agricultural University, 2022, 45(4):788-798(in Chinese with English abstract).
- [5] 王大庆,禄琳,于兴龙,等. 基于深度迁移学习的 EfficientNet 玉米叶部病害识别[J]. 东北农业大学学报,2023,54(5):66-76.  
Wang D Q, Lu L, Yu X L, et al. Maize leaf diseases identification using EfficientNet based on deep-transfer learning[J]. Journal of Northeast Agricultural University, 2023, 54(5):66-76(in Chinese with English abstract).
- [6] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016:779-788.
- [7] Liang X T, Jia X Y, Huang W Q, et al. Real-time grading of defect apples using semantic segmentation combination with a pruned YOLO V4 network[J]. Foods, 2022, 11(19):3150.
- [8] 陶兆胜,石鑫宇,王勇,等. 基于改进 YOLOv5s 的番茄叶片病害检测方法[J]. 沈阳农业大学学报,2023,54(6):712-721.  
Tao Z S, Shi X Y, Wang Y, et al. Tomato leaf diseases detection method based on improved YOLOv5s Journal of Shenyang Agricultural University, 2023, 54(6):712-721(in Chinese with English abstract).
- [9] 徐会杰,黄仪龙,刘曼. 基于改进 YOLOv3 模型的玉米叶片病虫害检测与识别研究[J]. 南京农业大学学报,2022,45(6):1276-1285. DOI:10.7685/jnau.202110039.  
Xu H J, Huang Y L, Liu M. Research on pest detection and identification of corn leaf based on improved YOLOv3 model[J]. Journal of Nanjing Agricultural University, 2022, 45(6):1276-1285(in Chinese with English abstract).
- [10] Gai R L, Chen N, Yuan H. A detection algorithm for cherry fruits based on the improved YOLO-v4 model[J]. Neural Computing and Applications, 2023, 35(19):13895-13906.
- [11] 赵文博,周德强,邓于然,等. 基于改进 YOLOv5 的甘蔗茎节识别方法[J]. 华中农业大学学报,2023,42(1):268-276.  
Zhao W B, Zhou D Q, Deng G R, et al. Sugarcane stem node recognition method based on improved YOLOv5[J]. Journal of Huazhong Agricultural University, 2023, 42(1):268-276(in Chinese with English abstract).
- [12] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[EB/OL]. 2015; arXiv: 1512.02325. <https://arxiv.org/abs/1512.02325>.
- [13] Ren S Q, He K M, 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, 2017, 39(6):1137-1149.
- [14] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. 2020; arXiv:2004.10934. <https://arxiv.org/abs/2004.10934>.
- [15] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9):1904-1916.
- [16] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[EB/OL]. 2017; arXiv:1709.01507. <https://arxiv.org/abs/1709.01507>.
- [17] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017:2261-2269.
- [18] 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>.
- [19] Tong Z J, Chen Y H, Xu Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. 2023; arXiv:2301.10051. <https://arxiv.org/abs/2301.10051>.