

基于改进 YOLOv7-tiny 的多光谱苹果表层缺陷检测

化春键^{1,2*}, 孙明春^{1,2}, 蒋毅^{1,2}, 俞建峰^{1,2}, 陈莹³¹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;³江南大学物联网工程学院, 江苏 无锡 214122

摘要 针对苹果表层存在多种缺陷类型、对不同缺陷的检测方法不同的问题,提出一种基于改进 YOLOv7-tiny 的缺陷检测模型,结合相机采集的 RGB+NIR 多光谱图像对苹果表层多种缺陷进行了检测和分类。首先,为了提取更多有效的特征信息,提高对缺陷的定位能力,在主干网络中使用坐标注意力(CA)机制聚合坐标信息,同时在主干网络后添加上下文转换器(CoT)模块以增加全局感受野;其次,为了增强高效聚合网络的特征融合能力,将其与加权双向特征金字塔结合,调整结构中各分支的占比;最后,为了解决难易样本不均衡的问题,将损失函数更换为 Focal-EIoU 损失。改进后网络的平均精度均值(mAP)@0.5 提升了 1.2 个百分点,达到 93.2%,识别速度为 89.3 frame/s。研究结果表明,本文研究内容为苹果表层的缺陷检测提供了更加高效的方法,同时为苹果的分级提供了更加精确的依据。

关键词 缺陷检测; 苹果表层; 多光谱图像; 深度学习; YOLOv7-tiny; 注意力机制

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP231895

Multispectral Apple Surface Defect Detection Based on Improved YOLOv7-tiny

Hua Chunjian^{1,2*}, Sun Mingchun^{1,2}, Jiang Yi^{1,2}, Yu Jianfeng^{1,2}, Chen Ying³¹School of Mechanical Engineering, Jiangnan University, Wuxi 214122, Jiangsu, China;²Jiangsu Key Laboratory of Advanced Food Manufacturing Equipment & Technology, Wuxi 214122, Jiangsu, China;³School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, Jiangsu, China

Abstract A defect detection model based on improved YOLOv7-tiny is proposed herein to address the problem of different detection methods for different defects on apple surface. Combined with RGB+NIR multispectral images collected by a camera, various defects on the apple surface are detected and classified. First, to extract more effective feature information and improve the ability to locate defects, coordinate attention (CA) is used to aggregate coordinate information in the backbone network, and a contextual transformer (CoT) module is added behind the backbone network to increase the global receptive field. Second, it is combined with the weighted bidirectional feature pyramid to adjust the proportion of each branch in the structure to enhance the feature fusion ability of efficient layer aggregation networks. Finally, the loss function is replaced by Focal-EIoU loss to solve the problem of unbalanced samples. The mean average precision (mAP) @0.5 of the improved network increases by 1.2 percentage points to 93.2%, and the recognition speed is 89.3 frames/s. The research content of this paper provides a more efficient method for apple surface defect detection and a more accurate basis for apple grading.

Key words defect detection; apple surface; multi-spectral image; deep learning; YOLOv7-tiny; attention mechanism

1 引言

苹果具有营养丰富、产出高、效益高等优点^[1]。我

国是苹果第一生产大国,2020/2021 产季我国苹果产量达 4406.6 万吨,占世界产量的 54.7%。虽然我国苹果产量位列前茅,但出口产品的竞争力较低^[2],原因之

收稿日期: 2023-08-10; 修回日期: 2023-09-19; 录用日期: 2023-10-09; 网络首发日期: 2023-11-07

基金项目: 国家自然科学基金(62173160)

通信作者: *277795559@qq.com

一是我国在苹果品质检测上做的工作不够,导致果品品质较差。而在苹果质检中表层缺陷的检测尤为重要,它直接影响到苹果的市场销售价格和消费者的购买欲望^[3]。苹果在生长的过程中会产生一些显性的缺陷,而在采摘和运输过程中容易产生一些肉眼难以发现的隐性缺陷。因此,通过融合多源信息,实现对苹果表层多种缺陷的无损检测已成为该领域的研究热点。

机器视觉是苹果无损检测的重要技术,在水果的分级中有较为广泛的应用。程磊^[4]使用改进粒子群算法对苹果表面是否存在缺陷进行检测,漏检率为 4.5%。Zhang 等^[5]利用模糊 C 均值算法和非线性规划遗传算法对苹果缺陷进行检测,准确率达 98%。上述基于机器视觉的研究往往难以检测苹果表层的隐性缺陷^[6],因此许多学者选择使用光谱图像,尤其是近红外光谱图像对苹果表层缺陷进行检测。Huang 等^[7]选定三个波段的图像对苹果的碰伤进行检测,静态测试和在线测试中的准确率分别达 91.5% 和 74.6%。邵志明等^[8]利用近红外相机成像和图像阈值分割方法对苹果表面早期损伤进行检测,准确率达到 90%。

近年来,深度学习在果蔬外部品质检测领域得到了广泛的应用。da Costa 等^[9]使用 ResNet50 模型对番茄外部缺陷进行检测,平均精度达到 94.6%。Fan 等^[10]使用改进 YOLOv4 网络和近红外相机实现了对苹果缺陷的实时检测,平均准确率达到 93.9%,但未对缺陷进行分类。Tang 等^[11]利用多波段图像和改进的 U-Net 进行苹果表面缺陷检测,平均交并比(mIoU)和平均 F1 分数(mF1)分别达到了 91% 和 95%。范晓飞等^[12]提出了一种基于四通多光谱图像和改进 YOLOv4 模型的检测方法,用于对玉米外观品质的快速识别和定位,平均精度均值(mAP)达到 98.02%。

现有研究的苹果表层缺陷检测类别较少,且少有研究针对苹果表层缺陷进行一次性综合评定和检测。因此,本文结合可见光和近红外图像的优点,融合多源信息,使用多光谱图像对苹果的表层缺陷进行检测,建

立了一个可见光-近红外(RGB-NIR)苹果表层缺陷数据集,基于改进 YOLOv7-tiny 建立了一个多光谱图像苹果表层缺陷检测模型。通过在网络中引入坐标注意力(CA)机制和上下文转换器(CoT)模块,提高网络对缺陷的特征提取能力和定位能力;通过改进的 Bi-LAN 结构,提高模型的特征融合能力;通过替换损失函数,改善难易样本不均衡的问题。

2 数据集与改进算法

2.1 数据集的建立

所使用的待检测苹果为随机挑选的红富士(Red Fuji)和奶油富士(Cream Fuji)两个品种的缺陷果,产地为山东烟台栖霞。根据中国农业部苹果等级规格(NY/T 1793—2009)^[13]和实验对象实际情况,将待检测缺陷类别分为显性缺陷、轻微碰压伤、刺伤、果锈 4 种,并将果梗和花萼这两种易混淆的目标列入检测对象中。

对待检测苹果使用不同质量的钢球进行坠落实验,模拟不同程度的碰压伤,所得损伤表面无表皮破裂,略有凹陷,部分仅凭肉眼难以发现;使用不同的金属尖锐物体对苹果表面模拟穿刺或划伤。选用 JAI 公司的 AD-080 GE 型 2CCD 多光谱棱镜面阵相机对处理 5 min 左右的苹果进行图像的采集。文献[6]表明在波长 850~960 nm 之间随着波长的增加,显性缺陷尺寸逐渐变小,隐性缺陷逐渐明显,故本文选用波长为 850 nm 和 940 nm 的两个光源,使用相机分别进行图像采集,所得 RGB 图像和两个近红外波段下的图像如图 1 所示。近红外 850 nm 光源下的图像显性缺陷更明显,面积也更大,而 940 nm 光源下的隐性缺陷更加明显,更容易辨别。为了更好地平衡可见光和近红外下两种图像的缺陷,本文使用 940 nm 光源作为近红外图像采集时的光源。

对采集到的 1180 对 RGB-NIR 图片使用 Labelimg 工具进行标注,按照 16:4:5 将数据集随机划分成训练

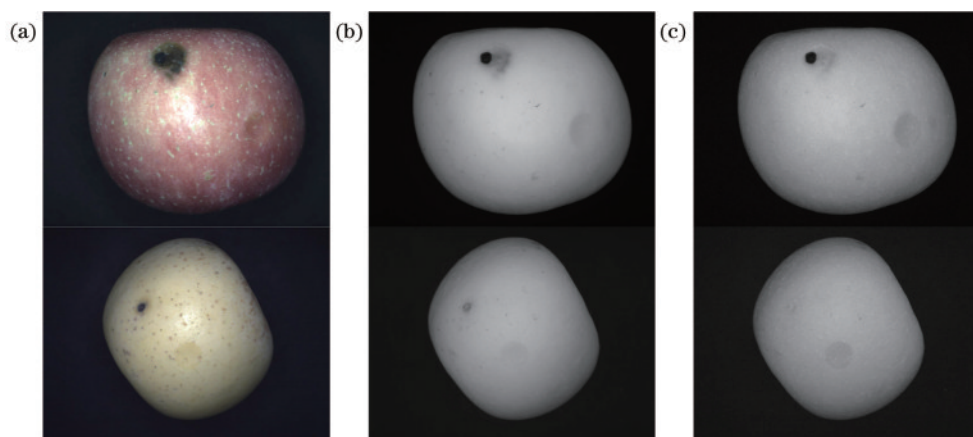


图 1 光谱图像对比。(a)RGB;(b)近红外(850 nm);(c)近红外(940 nm)
Fig. 1 Spectral image comparison. (a) RGB; (b) NIR (850 nm); (c) NIR (940 nm)

集、验证集、测试集。为了进一步提升训练模型的鲁棒性,对原始数据集进行扩充,对划分后的原始图片施加随机的运动模糊和垂直翻转,以模拟流水线上实时检测的情况,得到 2360 对 RGB-NIR 图像。各类别数目

分别为:果梗(stem)720个,花萼(calyx)594个,显性缺陷(defect)1738个,轻微碰压伤(bruise)1186个,刺伤(stab)902个,果锈(rust)490个。苹果上各种缺陷及果梗、花萼的 RGB 和 NIR 局部图像如图 2 所示。

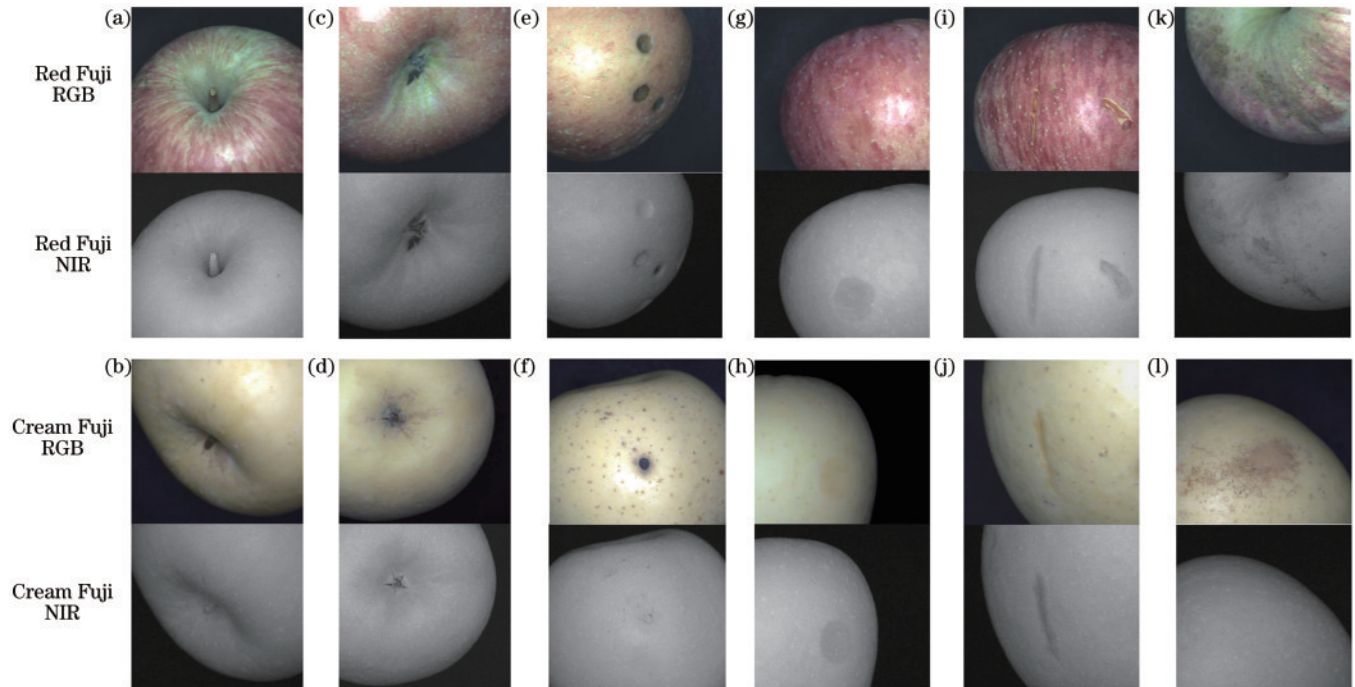


图 2 各类别图像。(a)(b)果梗;(c)(d)花萼;(e)(f)显性缺陷;(g)(h)轻微碰压伤;(i)(j)刺伤;(k)(l)果锈

Fig. 2 Images of each category. (a) (b) Stem; (c) (d) calyx; (e) (f) defect; (g) (h) bruise; (i) (j) stab; (k) (l) rust

2.2 YOLOv7-tiny 网络

YOLOv7 算法模型由 Wang 等^[14]在 2022 年提出,其中的 YOLOv7-tiny 版本由 YOLOv7 模型精简而来,在保证检测精度的基础上对模型进行了压缩,使其拥有更少的参数量和更快的检测速度,适用于本文快速检测的需求,同时较小的模型也适合于部署在检测流水线上,故选择使用 YOLOv7-tiny 作为基础网络进行改进。

YOLOv7-tiny 主要包括特征提取网络(Backbone)、特征融合网络(Neck)、特征检测输出层(Head)。在输入端使用 mosaic 技术对输入图片进行随机拼接,丰富图片背景,提高网络的鲁棒性。Backbone 主要由高效聚合网络(ELAN)组成,通过控制最短、最长梯度路径,使更深的模块可以更加有效地进行学习并得到更好的收敛。在 Neck 部分采用路径聚合网络(PANet)结构,对浅层特征与深层特征进行融合,达到增强特征的作用。在 Head 部分使用标准卷积进行通道数调整,获取最终的检测结果。

2.3 改进 YOLOv7-tiny 目标检测模型构建

改进后的 YOLOv7-tiny 模型结构如图 3 所示。成对的 RGB 图像和 NIR 图像经通道合并(merge channels)后得到 4 通道图像,在经过图像大小重置(reshape)后得到 $640 \times 640 \times 4$ 大小的图像。Backbone 用于提取 4 通道图片中的特征,生成 3 个不同尺度的特

征图 P3、P4 和 P5;将得到的特征图输入 PANet 中,沿自顶向下和自底向上两条路径进行多尺度特征融合,对高层语义信息和低层定位信息进行双向传递,增强主干网络的表征能力;最后将融合后的特征分别输入到 YOLO Head 中进行预测框的回归和目标类别的判断。

2.4 改进方法

为了提高对缺陷的定位能力,在主干特征提取网络的 P3 和 P4 特征图后添加 CA 机制^[15],其结构如图 4 所示。CA 机制通过在图像的宽度上和高度上进行位置编码,获得全局的位置信息,可以更好地对有效特征进行提取,减少背景无关特征的影响,获取更多苹果缺陷的信息。

CA 机制首先使用水平和垂直方向的平均池化对特征图进行编码操作,随后对其通道进行级联,级联后的特征图经过卷积和激活操作后分成两个张量,再分别进行卷积和激活操作,得到宽度方向和高度方向的两个坐标注意力图;最终的输出是原始分支和这两个注意力图的相乘。CA 机制同时考虑了位置关系和通道关系,捕获了跨通道的信息,还包含方向与位置信息,使得模型能够更加准确地定位到缺陷的区域。

为了进一步获取来自全局感受野的特征,提高对果锈等大面积不规则缺陷的检测能力,将 CoT 模块插入空间金字塔跨阶段部分通道(SPPCSPC)模块后。

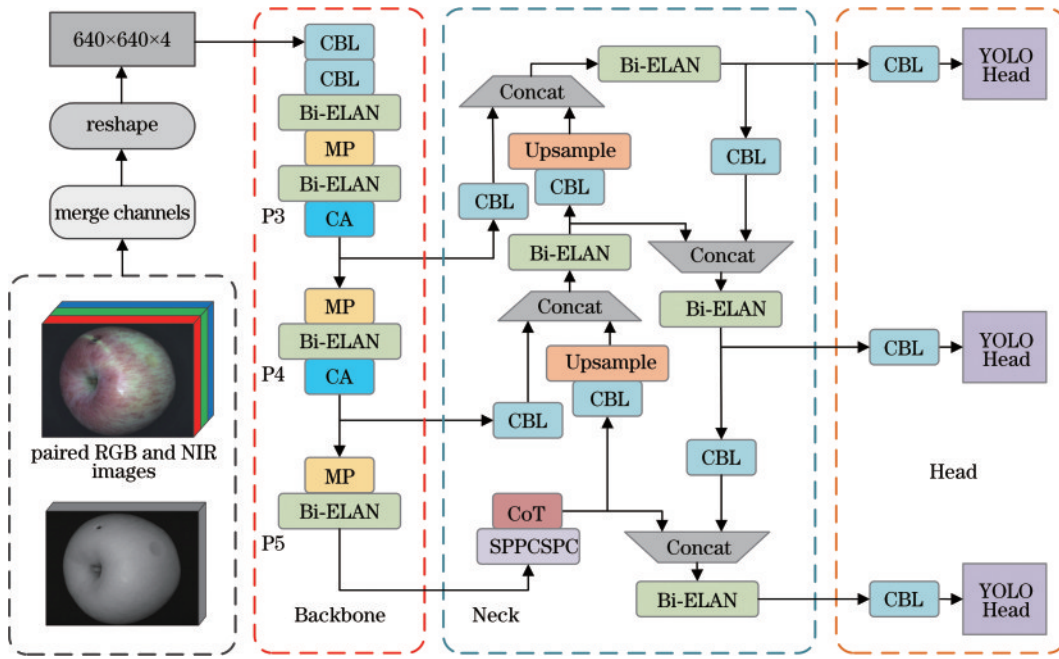


图 3 改进 YOLOv7-tiny 网络结构
Fig. 3 Improved YOLOv7-tiny network structure

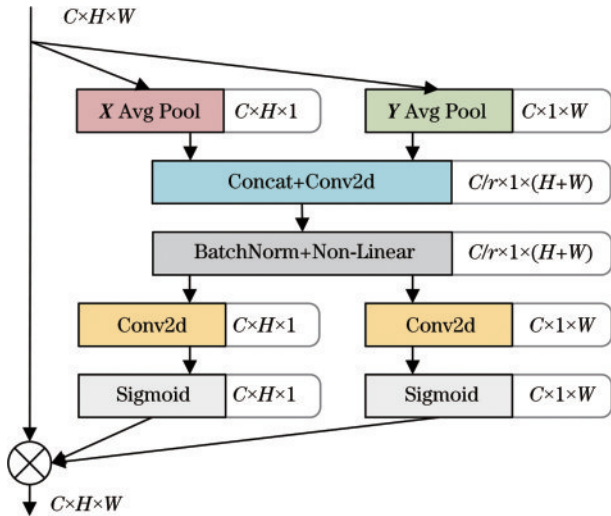


图 4 CA 模块结构
Fig. 4 CA module structure

基于 Transformer 的全局交互机制,有效感受野可以迅速扩大,通过自注意力机制可以建模输入数据之间的全局依赖关系^[16],获取更加精确的目标位置信息。CoT 模块结构如图 5 所示,对于输入大小为 $H \times W \times C$ 的特征矩阵 X ,对其进行特征映射操作,得到键 K 、查询 Q 、值 V ,公式为

$$\begin{cases} K = X \\ Q = X \\ V = XW_v \end{cases}, \quad (1)$$

式中: W_v 是将 X 转换为 V 的特征映射矩阵。然后,使用大小为 3 的卷积核对 K 进行卷积操作,得到反映局部邻近位置静态上下文信息的特征映射 K_1 ,使用 K_1 和 Q 得到注意力矩阵 A ,其计算公式为

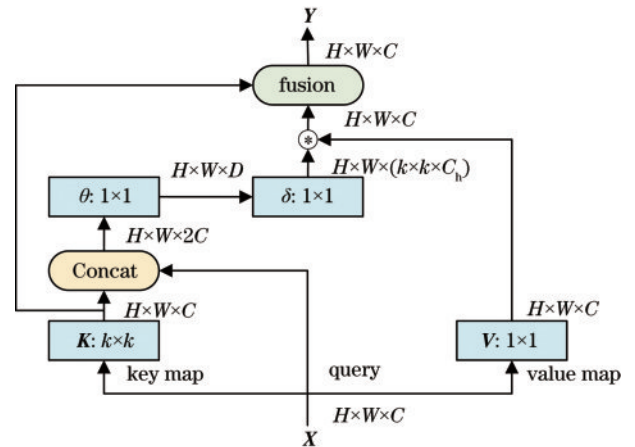


图 5 CoT 模块结构
Fig. 5 CoT module structure

$$A = [K_1, Q]W_\theta W_\delta, \quad (2)$$

式中: $[\cdot]$ 代表级联操作; W_θ 和 W_δ 为两个 1×1 的卷积操作。基于所得到的注意力矩阵 A 和通过特征映射得到的 V ,计算用于捕捉动态上下文特征交互的特征映射 K_2 ,计算公式为

$$K_2 = V \circledast A, \quad (3)$$

式中: \circledast 表示局部矩阵乘法运算。最终的输出 Y 为静态上下文表示 K_1 和动态上下文表示 K_2 的融合。CoT 模块是一种可以充分发掘邻近上下文信息和全局信息的自注意力学习模块,在 SPPCSPC 后加入该模块,在保证增加较少参数数量的同时对全局特征进行提取,同时获取一定的上下文信息,使特征提取能力得到加强。

YOLOv7-tiny 将 YOLOv7 中的扩展高效聚合网络 (E-ELAN) 精简为 ELAN,以满足降低模型大小、提

高推理速度的需求。虽然这种结构可以在保证获取最短梯度路径的同时提高融合效率,但是仅对网络上层之间进行通道的拼接,特征信息难以充分融合^[17],因此借鉴加权双向特征金字塔网络(Bi-FPN)^[18]的思想,对 ELAN 模块中的 4 个分支模块进行加权级联的操作,自动分配权重,提高模块的特征提取能力和对有用特征的识别能力,得到的新模块命名为 Bi-ELAN。原始 ELAN 和 Bi-ELAN 的结构如图 6 所示,Bi-ELAN 将原始 ELAN 中的 Concat 级联操作替换为了 Bi-Concat 加权级联。

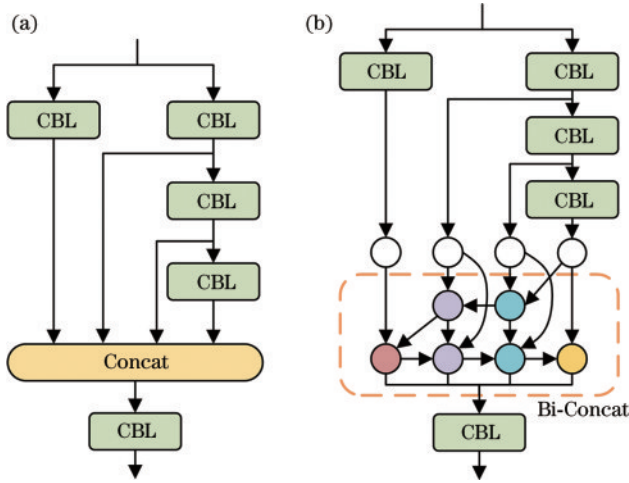


图 6 ELAN 和 Bi-ELAN 结构。(a)ELAN;(b)Bi-ELAN
Fig. 6 ELAN and Bi-ELAN structures. (a) ELAN;
(b) Bi-ELAN

YOLOv7-tiny 默认使用完全交并比(CIoU)作为边界框回归的损失函数,CIoU 考虑了重叠面积、中心距离和纵横比,但纵横比不能反映实际的边长差异,因此将边界框回归的损失函数更改为可以反映实际边长差异的高效交并比(EIoU)^[19]。EIoU 损失函数的计算公式为

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - R_{IoU} + \frac{\rho^2(b, b^{gt})}{(\omega^c)^2 + (h^c)^2} + \frac{\rho^2(\omega, \omega^{gt})}{(\omega^c)^2} + \frac{\rho^2(h, h^{gt})}{(h^c)^2}, \quad (4)$$

式中: L_{IoU} 为重叠面积损失; L_{dis} 为距离损失; L_{asp} 为宽高损失; b, ω 和 h 分别代表预测框的中心点坐标、宽和高; b^{gt}, ω^{gt} 和 h^{gt} 分别代表真实框的中心点坐标、宽和高; ω^c 和 h^c 是覆盖预测框和真实框的最小外接矩形的宽度和高度; $\rho(\cdot)$ 表示欧氏距离; R_{IoU} 为预测框和真实框的交并比。此外,边界框回归还存在着高质量锚框和低质量锚框不均衡的问题,Focal-EIoU 损失使用 Focal 损失解决该问题,计算公式为

$$L_{Focal-EIoU} = R_{IoU}^\gamma L_{EIoU}, \quad (5)$$

式中: γ 是控制异常值抑制程度的参数。总的来说,Focal-EIoU 损失将锚框的宽和高分开计算,同时使用 Focal 损失调节难易样本,有助于提升收敛的速度和回

归的精度,同时可以保障样本的整体质量。

3 实验内容

3.1 实验环境与参数配置

所提算法的构建与改进所使用的深度学习框架 PyTorch 版本为 1.7.1,CUDA 版本为 11.0,操作系统版本为 Ubuntu 20.04.4 LTS, GPU 为 NVIDIA GeForce RTX 3090。实验使用随机梯度下降法(SGD)作为优化器对神经网络进行优化,加快训练过程。初始学习率设置为 0.01,SGD 动量参数统一设为 0.937,权值衰减参数设为 0.0005。

3.2 评价指标

实验采用 4 种指标进行模型性能的评价,分别是反映所有类别的识别平均准确率的 mAP、模型每秒可以检测图片数量的处理速度、反映模型参数的内存占用量的参数量(Params)、反映模型的计算量的浮点运算数(FLOPs)。mAP 的计算公式为

$$P_{mAP} = \frac{1}{k} \sum_{i=1}^k P_{AP_i}, \quad (6)$$

式中: k 为待检测类别数; P_{AP} 为每类的平均识别精度。目标置信度阈值 IoU 为 0.5 时的 mAP 记作 mAP@0.5。

3.3 不同图像模型对比实验

针对两种图像设计不同的模型,模型分为仅使用 3 通道 RGB 图像的 YOLOv7-tiny(RGB)模型、仅使用单通道 NIR 图像的 YOLOv7-tiny(NIR)模型、使用 4 通道图像的 YOLOv7-tiny(RGB-NIR)模型以及改进后的 YOLOv7-tiny(RGB-NIR)模型。不同模型各类别 AP 以及 mAP@0.5 对比如表 1 所示。从实验结果可以看出:当仅使用 RGB 图像进行缺陷检测时,模型对 bruise 的平均精度较低,为 89.4%,比仅使用 NIR 图像的模型低 9.0 百分点;而仅使用 NIR 图像进行检测时,模型对 defect 和 rust 的平均精度较低,分别为 71.5% 和 15.4%,分别较仅使用 RGB 图像的模型低 21.3 百分点和 53.1 百分点;YOLOv7-tiny(RGB-NIR)模型对 defect、bruise、stab 的平均检测精度提升,对 stem、calyx、rust 的平均检测精度有所下降,原因是两种图像之间产生了一定的影响;改进后的算法消除了一定的负面影响,同时有效提高了检测精度,除对 bruise 的平均检测精度保持与原网络相近之外对其他类别的平均检测精度都得到了提升,相比 YOLOv7-tiny(RGB-NIR),对 stem 的平均检测精度提高 2.1 百分点,对 calyx 的平均检测精度提高 0.5 百分点,对 defect 的平均检测精度提高 1.2 百分点,对 stab 的平均检测精度提高 0.2 百分点,对 rust 的平均检测精度提高 3.4 百分点,对所有类别的平均检测精度提升 1.2 百分点。总的来说,改进后的算法可以较好地权衡 RGB 和 NIR 两种图像中的缺陷类型,达到最优的检测效果。

4个模型的精确率-召回率(P-R)曲线对比如图7所示,可以看到仅使用NIR图像的P-R曲线在最下面,由下到上的模型依次是YOLOv7-tiny(NIR)、YOLOv7-tiny

(RGB)、YOLOv7-tiny(RGB-NIR)、improved YOLOv7-tiny(RGB-NIR)。改进算法的P-R曲线将其他曲线包裹,证明改进后的算法的平均精度达到了最优。

表1 不同模型的平均精度对比

Table 1 Comparison of average accuracy of different models

Model	AP / %						mAP@0.5 / %
	stem	calyx	defect	bruise	stab	rust	
YOLOv7-tiny (RGB)	97.1	99.8	92.8	89.4	94.1	68.5	90.3
YOLOv7-tiny (NIR)	97.3	96.7	71.5	98.4	95.1	15.4	79.1
YOLOv7-tiny (RGB-NIR)	96.0	99.2	93.3	98.5	97.2	67.7	92.0
Improved YOLOv7-tiny (RGB-NIR)	98.1	99.7	94.5	98.4	97.4	71.1	93.2

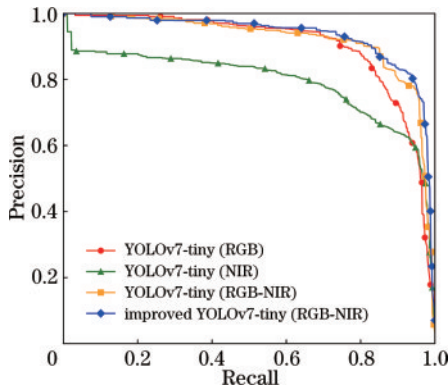


图7 不同模型的P-R曲线对比

Fig. 7 Comparison of P-R curves of different models

使用改进后的网络模型对测试集的图片进行检测实验,部分检测结果如图8所示,可以看到包括显性缺

陷、轻微碰压伤、刺伤、果锈在内的多种缺陷以及果梗和花萼都得到了有效的检测识别,同时在一个苹果上存在的多种缺陷或果梗和花萼都得到了很好的区分,证明了改进网络的有效性。

为进一步验证改进网络的优势,对改进前后的部分检测结果进行对比,如图9所示。可以看到:图9(a)和图9(b)中的果锈未被原始网络模型识别,而图9(f)和图9(g)中的果锈成功地被改进后的模型框出;图9(c)下方红色果皮被原始模型误识别为果锈,右上角的刺伤由于接近边缘也未被识别,图9(h)中果皮颜色成功地被改进后的模型正确区分,且边缘处的刺伤得到精确识别;图9(d)中苹果右下角和图9(e)中苹果右侧的显性缺陷未被识别,图9(i)和图9(j)中显性缺陷成功地被改进后的模型识别。改进后的网络的特征提取能力得到了加强,对各类别的辨别能力得到了提升。

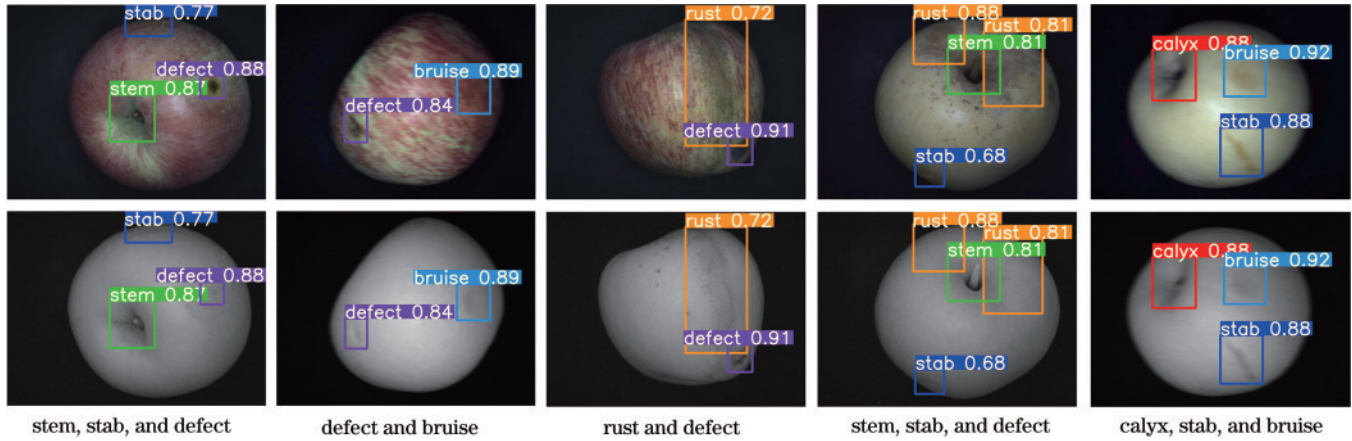


图8 改进算法的检测结果

Fig. 8 Detection results of the improved algorithm

3.4 消融实验

为了进一步验证各个模块的有效性,对改进网络的各个模块进行消融实验,各模块使用情况和实验结果如表2所示,使用mAP@0.5、Params、FLOPs作为主要评价指标。从表2可以看出:原始的YOLOv7-tiny在RGB-NIR数据集上的mAP@0.5为92.0%,模型参数量为 6.03×10^6 ,计算量为 13.28×10^9 ;在主干网络中间的两个特征图后添加CA机制后,mAP@0.5

提升了0.4个百分点,说明CA机制能够较好地提取坐标信息和通道信息,提高模型的特征表达能力;在SPPCSPC后加入CoT模块后,mAP@0.5提升0.9个百分点,说明CoT模块获取到了一定的全局信息,提高了对缺陷特征的捕捉能力;使用Bi-ELAN后,mAP@0.5提升了0.6个百分点,说明改进特征融合方式对提升网络性能有一定的帮助;使用Focal-EIoU后,mAP@0.5提升了0.7个百分点,说明Focal-EIoU对缓解低质

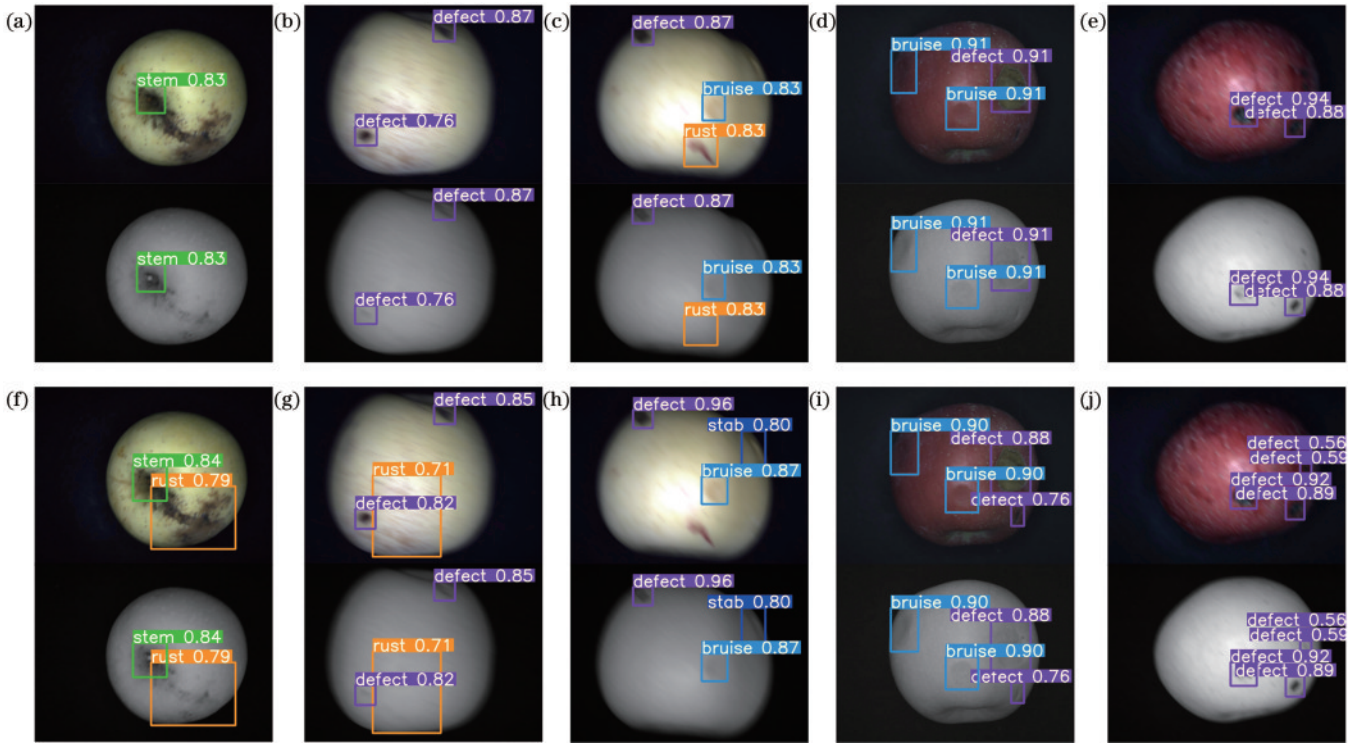


图 9 模型改进前后的检测结果对比。(a)~(e)原始 YOLOv7-tiny 的检测效果;(f)~(j)改进 YOLOv7-tiny 的检测效果
 Fig. 9 Comparison of detection results before and after model improvement. (a)~(e) Original YOLOv7-tiny' detection results; (f)~(j) improved YOLOv7-tiny' detection results

表 2 消融实验结果

Table 2 Results of ablation experiment

Experiment No.	CA	CoT	Bi-ELAN	Focal-EIoU	mAP@0.5 / %	Params / 10 ⁶	FLOPs / 10 ⁹
1					92.0	6.03	13.28
2	✓				92.4	6.04	13.32
3		✓			92.9	6.61	13.74
4			✓		92.6	6.03	13.28
5				✓	92.7	6.03	13.28
6	✓		✓	✓	92.9	6.04	13.32
7		✓	✓	✓	93.1	6.61	13.74
8	✓	✓	✓	✓	93.2	6.62	13.78

量样本造成的损失振荡、平衡难易样本有较好的效果。实验 6 和实验 7 进一步验证了 CoT 和 CA 的有效性,分别去掉这两个模块后,mAP@0.5 降低了 0.3 个百分点和 0.1 百分点。改进后模型的参数量为 6.62×10^6 ,较原算法增加了 0.59×10^6 ,计算量为 13.78×10^9 ,较原算法增加了 0.5×10^9 ,mAP@0.5 从 92.0% 提升到 93.2%。总的来说,所提改进算法在增加了少量参数数量和计算量的情况下,提高了整体的 mAP@0.5。

3.5 不同目标检测算法性能对比

为进一步对比模型性能,在相同条件下,分别使用 Faster-RCNN^[20]、Swin-Transformer^[21](基于 Mask-RCNN)、YOLOv3^[22]、YOLOv5s、YOLOv7-tiny 及所提算法在 RGB-NIR 数据集上进行实验,输入都为 4 通道

图像,加载预训练模型并将训练 epoch 设置为 200,使用测试集对 6 种算法进行性能评估,其平均精度、参数量、计算量、检测速度等指标如表 3 所示。根据表 3 可知:所提算法的 mAP@0.5 最高,两阶段的目标检测算法 Faster-RCNN 的 mAP@0.5 最低;基于 Mask-RCNN 的 Swin-Transformer 虽然相较于 Faster-RCNN 在 mAP 上提升了 1.4 个百分点,但参数量和计算量的增加导致检测速度降低了 0.2 frame/s;单阶段的 YOLO 系列算法在检测速度上有较大的优势,YOLOv3 和 YOLOv5s 的 mAP@0.5 较为相近,但 YOLOv3 庞大的参数量和计算量导致检测速度比 YOLOv5s 低了 42.9 frame/s;YOLOv5s 和 YOLOv7-tiny 在参数量和运算量上接近,YOLOv7-tiny 的 mAP@0.5 比 YOLOv5s 的高 1.5 百分点,检测速度比 YOLOv5s 高 4.5 frame/s; Improved

表 3 不同模型的性能对比

Table 3 Performance comparison of different models

Model	mAP@0.5 /%	Params /10 ⁶	FLOPs /10 ⁹	Speed /(frame·s ⁻¹)
Faster-RCNN	83.7	41.15	91.35	6.0
Swin-Transformer	85.1	44.77	92.04	5.8
YOLOv3	89.5	61.55	155.58	61.3
YOLOv5s	90.5	7.04	16.22	104.2
YOLOv7-tiny (baseline)	92.0	6.03	13.28	108.7
Improved YOLOv7-tiny	93.2	6.62	13.78	89.3

YOLOv7-tiny 的 mAP@0.5 比 YOLOv5s 高 2.7 个百分点,比 YOLOv7-tiny 高 1.2 个百分点,检测速度与改进前相比有所下降,但仍达 89.3 frame/s。

4 结 论

在 YOLOv7-tiny 的基础上提出了一种多光谱图像的苹果表层缺陷检测算法,可以用于检测显性缺陷、碰压伤、刺伤、果锈等苹果表层缺陷,同时可以识别苹果的果梗和花萼。实验结果表明:在网络结构中引入可以选择性突出关键位置信息的 CA 机制以及可以获得全局感受野的 CoT 模块后,可提高对缺陷的定位能力和检测精度;将网络中的 ELAN 结构改进为 Bi-ELAN 结构,可提高网络的特征融合能力;使用 Focal-EIoU 替换原损失函数,可改善难易样本不均衡的问题。与改进前 4 通道输入的 YOLOv7-tiny 相比,所提改进模型在增加较少参数数量和运算量的前提下将检测精度提升了 1.2 个百分点,能有效识别 4 种缺陷以及果梗和花萼,同时保持较高的检测速度。未来的工作将进一步扩充数据集,同时精简网络,使模型可以部署到苹果检测及包装流水线上,用于苹果的分级检测。

参 考 文 献

- [1] 2021 年度中国苹果产业发展报告(总篇)精简版[J]. 中国果菜, 2023, 43(1): 1-8.
China apple industry development report in 2021 (general), condensed version[J]. China Fruit & Vegetable, 2023, 43(1): 1-8.
- [2] 牛桂草, 宋卓展, 刘畅, 等. 中国苹果贸易竞争力评价与分析[J]. 河北农业科学, 2022, 26(3): 97-100.
Niu G C, Song Z Z, Liu C, et al. Evaluation and analysis of apple trade competitiveness in China[J]. Journal of Hebei Agricultural Sciences, 2022, 26(3): 97-100.
- [3] 马佳佳, 王克强. 水果品质光学无损检测技术研究进展[J]. 食品工业科技, 2021, 42(23): 427-437.
Ma J J, Wang K Q. Research progress of optical nondestructive testing technology for fruit quality[J]. Science and Technology of Food Industry, 2021, 42(23): 427-437.
- [4] 程磊. 基于改进粒子群算法的苹果表面缺陷检测[J]. 食品与机械, 2018, 34(3): 141-145.
Cheng L. Apple surface defect detection research based on improved particle swarm optimization algorithm[J].

Food & Machinery, 2018, 34(3): 141-145.

- [5] Zhang W Z, Hu J, Zhou G X, et al. Detection of apple defects based on the FCM-NPGA and a multivariate image analysis[J]. IEEE Access, 2020, 8: 38833-38845.
- [6] 庞琦. 基于深度学习和光谱成像技术的苹果表皮显/隐性缺陷检测研究[D]. 上海: 上海海洋大学, 2022.
Pang Q. Research on detection of apple skin explicit/implicit defects based on deep learning and spectral imaging technology[D]. Shanghai: Shanghai Ocean University, 2022.
- [7] Huang W Q, Li J B, Wang Q Y, et al. Development of a multispectral imaging system for online detection of bruises on apples[J]. Journal of Food Engineering, 2015, 146: 62-71.
- [8] 邵志明, 王怀彬, 董志城, 等. 基于近红外相机成像和阈值分割的苹果早期损伤检测[J]. 农业机械学报, 2021, 52(S1): 134-139.
Shao Z M, Wang H B, Dong Z C, et al. Early bruises detection method of apple surface based on near infrared camera imaging technology and image threshold segmentation method[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(S1): 134-139.
- [9] da Costa A Z, Figueroa H E H, Fracarolli J A. Computer vision based detection of external defects on tomatoes using deep learning[J]. Biosystems Engineering, 2020, 190: 131-144.
- [10] Fan S X, Liang X T, Huang W Q, et al. Real-time defects detection for apple sorting using NIR cameras with pruning-based YOLOV4 network[J]. Computers and Electronics in Agriculture, 2022, 193: 106715.
- [11] Tang Y, Bai H Y, Sun L J, et al. Multi-band-image based detection of apple surface defect using machine vision and deep learning[J]. Horticulturae, 2022, 8(7): 666.
- [12] 范晓飞, 王林柏, 刘景艳, 等. 基于改进 YOLO v4 的玉米种子外观品质检测方法[J]. 农业机械学报, 2022, 53(7): 226-233.
Fan X F, Wang L B, Liu J Y, et al. Corn seed appearance quality estimation based on improved YOLO v4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 226-233.
- [13] 中华人民共和国农业部. 苹果等级规格: NY/T 1793—2009[S]. 北京: 中国农业出版社, 2010.
Ministry of Agriculture of the People's Republic of China. Apple grade specifications: NY/T 1793—2009 [S]. Beijing: China Agriculture Press, 2010.
- [14] Wang C Y, Bochkovskiy A, Liao H Y 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), June 17-24, 2023, Vancouver, BC, Canada. New York: IEEE Press, 2023: 7464-7475.
- [15] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13708-13717.
- [16] 李建, 杜建强, 朱彦陈, 等. 基于 Transformer 的目标检测算法综述[J]. 计算机工程与应用, 2023, 59(10): 48-64.
Li J, Du J Q, Zhu Y C, et al. Survey of Transformer-based object detection algorithms[J]. Computer Engineering and Applications, 2023, 59(10): 48-64.
- [17] 齐向明, 董旭. 改进 YOLOv7-tiny 的钢材表面缺陷检测算法[J]. 计算机工程与应用, 2023, 59(12): 176-183.
Qi X M, Dong X. Improved YOLOv7-tiny algorithm for steel surface defect detection[J]. Computer Engineering and Applications, 2023, 59(12): 176-183.
- [18] Tan M X, Pang R M, Le Q V. EfficientDet: scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 10778-10787.
- [19] Zhang Y F, Ren W Q, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146-157.
- [20] 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.
- [21] Liu Z, Lin Y T, Cao Y, et al. Swin Transformer: hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10-17, 2021, Montreal, QC, Canada. New York: IEEE Press, 2022: 9992-10002.
- [22] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08) [2023-06-05]. <https://arxiv.org/abs/1804.02767>.