

郑宇达,陈仁凡,杨长才,等.基于改进YOLOv5s模型的柑橘病虫害识别方法[J].华中农业大学学报,2024,43(2):134-143.  
DOI:10.13300/j.cnki.hnlkxb.2024.02.016

## 基于改进YOLOv5s模型的柑橘病虫害识别方法

郑宇达<sup>1</sup>,陈仁凡<sup>1</sup>,杨长才<sup>2</sup>,邹腾跃<sup>1</sup>

1. 福建农林大学机电工程学院,福州 350002; 2. 福建农林大学计算机与信息学院,福州 350002

**摘要** 针对现有检测模型不能满足在自然环境中准确识别多种类柑橘病虫害的问题,提出一种基于改进YOLOv5s模型的常见柑橘病虫害检测方法。改进模型引入ConvNeXtV2模型,构建一个CXV2模块替换YOLOv5s的C3模块,增强提取特征的多样性;添加了动态检测头DYHEAD,提高模型对不同空间尺度、不同任务目标的处理能力;采用CARAFE上采样模块,提高特征提取效率。结果显示,改进后的YOLOv5s-CDC的召回率和平均精度均值分别为81.6%、87.3%,比原模型分别提高了4.9、3.4个百分点。与其他YOLO系列模型在多个场景下的检测对比,具有更高的准确率和较强的鲁棒性。结果表明,该方法可用于自然复杂环境下的柑橘病虫害的检测。

**关键词** 深度学习; 病虫害; YOLOv5s; 目标检测

**中图分类号** TP391.4 **文献标识码** A **文章编号** 1000-2421(2024)02-0134-10

柑橘是我国最为重要的农产品之一,种植面积与产量位居世界首位<sup>[1]</sup>,给果农带来良好的经济效益。由于柑橘的生长周期长,受气候、环境等因素的影响,极易发生病虫害,若不能准确判断柑橘病虫害种类,及时采取防治措施,对症下药,则会造成柑橘产量下降,甚至传染果园其他区域,造成严重的经济损失。在复杂的果园环境中依靠人工诊断的方式不仅效率低下且具有一定的主观性,同时大面积种植仅依靠人力诊断也是远远不够的。

随着计算机视觉技术和人工智能的兴起,图像处理技术得到大力发展,由于低成本,易于实现,这些技术被广泛应用到农业的病虫害检测中。传统的农业图像处理主要通过支持向量机<sup>[2-4]</sup>、K近邻<sup>[5]</sup>等机器学习算法,但是传统的机器学习技术需要依靠人工完成特征选择提取,泛化能力弱,且受环境限制较大,并不适合运用在实际果园环境中。

深度学习是机器学习的一个特定领域,与机器学习相比,深度学习能自动提取特征信息,减少特征提取工程的负担以及人工提取容易出现错误,适合处理大规模数据集,并提供更好的预测和泛化能力。深度学习中常见的目标检测模型有Faster R-CNN<sup>[6]</sup>、SSD<sup>[7]</sup>、YOLO系列<sup>[8-9]</sup>等,目前已被大量应

用到农业病虫害检测研究中。Zhao等<sup>[10]</sup>提出了一种新的Faster R-CNN架构检测7种不同的草莓病害,取得了较好的检测精度。Sun等<sup>[11]</sup>构建了一个轻量级卷积神经网络MEAN-SSD,降低模型大小,并在移动设备上成功检测5种苹果叶片病害。Zhang等<sup>[12]</sup>用YOLOv4和EfficientNet模型结合的方法识别了柑橘果实的病害。Li等<sup>[13]</sup>基于YOLOv5n检测黄瓜病害,通过引入协调注意力机制和Transformer结构提高了模型的检测精度,并且具备较好的鲁棒性。孙丰刚等<sup>[14]</sup>用改进的YOLOv5s实现了快速检测苹果果实病害。

目前针对柑橘病虫害的研究大多倾向于分类任务<sup>[15-17]</sup>,对象类别较少,所用数据集在实验室环境下基于简单背景进行拍摄<sup>[18-19]</sup>。由于柑橘病虫害类间特征相似、自然环境复杂、枝叶遮挡等特点使得现有模型很容易出现漏检、误检等问题,不能满足自然环境中对于多种类、多目标病虫害检测。针对以上问题,本研究基于YOLOv5s提出了一个能在复杂的果园环境下识别柑橘病虫害的目标检测方法,旨在辅助果农检测病虫害,以及为果园施药机器人等智能产品提供算法依据。

收稿日期: 2023-08-24

基金项目: 福建省自然科学基金项目(2019J01402)

郑宇达, E-mail: 403553702@qq.com

通信作者: 邹腾跃, E-mail: zouty@fafu.edu.cn

1 材料与方法

1.1 数据采集

柑橘病虫害数据集采集自福建省泉州市永春县柑橘种植园,使用华为 Mate 40 Pro 手机作为采集设备,摄像头分辨率为 4 096 像素×3 072 像素,ISO 值为 50,在不同时间段(07:00—18:00)的自然光照条件下,多角度进行拍摄。为了构建符合自然复杂环境的数据集,手机距离目标 15~35 cm,拍摄的图像中除了清晰的病虫害叶片,也包含其他叶片等周围

背景,叶片密度为果园自然状态,未进行人工干预。因此数据集包含了不同光照强度、不同目标大小、目标密集、叶片遮挡等情景,增加了数据的多样性。每个图像通过检查明确采集的柑橘病虫害类型。在去除质量较差和非典型病虫害图像后,共保留复杂背景下 6 种柑橘病虫害(黄龙病、潜叶蛾、树脂病、蚜虫、锈壁虱、蚧壳虫)叶片图像 1 649 张作为本研究的数据(图 1)。在每个病虫害类别的数据集中,大部分图像包含多个该类病虫害的目标,部分图像同时存在其他类型的病虫害,以满足多目标、多种类的检测需求。



图 1 病害类型

Fig.1 Disease types

为提高模型的训练效率,节省计算资源与训练时间,对原图像进行无损压缩,压缩后图像分辨率大小为 2 048 像素×1 536 像素。使用 LabelImg 图像标注工具对数据集中每个受害叶片进行标注,标注结果以 XML 格式进行保存,以适应不同算法的格式要求。将原图像以 9:1 的比例随机分割为训练集与验证集。由于存在样本数目不平均的问题,为增强网

络的泛化性,防止图像数量不足导致过拟合,本研究使用翻转、比例缩放、局部裁剪、改变对比度、添加高斯噪声、伽马噪声等方式进行数据增强,平衡各类病害的数量,同时去除了不符合条件的增强图像,最终得到数据图像 7 091 张,其中训练集 6 313 张,验证集 778 张。数据集及各类病害叶片标注数量如表 1 所示。

表 1 柑橘病虫害数据集信息

Table 1 Information of citrus pests and diseases data set

项目 Item	蚜虫 Aphid	蚧壳虫 Coccid	黄龙病 Yellow shoot disease	树脂病 Melanose	潜叶蛾 Leaf miner	锈壁虱 Rust mite
原始图像数 Original images	285	295	120	215	439	295
数据增强后 Data enhancement	1 140	1 180	1 115	1 159	1 317	1 180
最终标注量 Labels	4 326	5 146	4 986	4 564	5 576	4 698

1.2 YOLOv5s-CDC 检测模型改进

YOLO 系列模型中,YOLOv5 有着检测精度高、速度快、模型较小便于部署等优点,常被运用于各种环境目标检测场景。按照模型大小 YOLOv5 分为 5

个版本,通过自建数据集上的试验对比分析,本研究选择了 YOLOv5s 作为基线模型。YOLOv5s 模型主要分为输入层(input)、主干网络(backbone)、颈部网络(neck)、检测头(head)4 个部分。输入层对检测

图像进行预操作处理,将图像传入模型。主干网络用于提取输入图像中的特征,使用CSPDarknet53架构,有效提高特征传递的效率。颈部网络即特征融合层,使用特征金字塔网络FPN(feature pyramid network)<sup>[20]</sup>和路径聚合网络PAN(path aggregation network)<sup>[21]</sup>将不同尺度的特征图进行融合,提高网络对不同特征层的检测能力。检测头用来映射提取的特征,输出预测目标的位置,类别和置信度信息。

为提高在自然复杂环境中的柑橘病虫害检测性能,本研究提出一种改进模型YOLOv5s-CDC。设计了1个CXV2模块替换原模型中的C3模块、用动态检测头DYHEAD替换原有检测头,并将颈部网络中上采样方式换成CARAFE上采样算子,在不明显增加模型计算量的基础上提高对柑橘病虫害的检测精度。YOLOv5s-CDC的网络结构如图2所示。

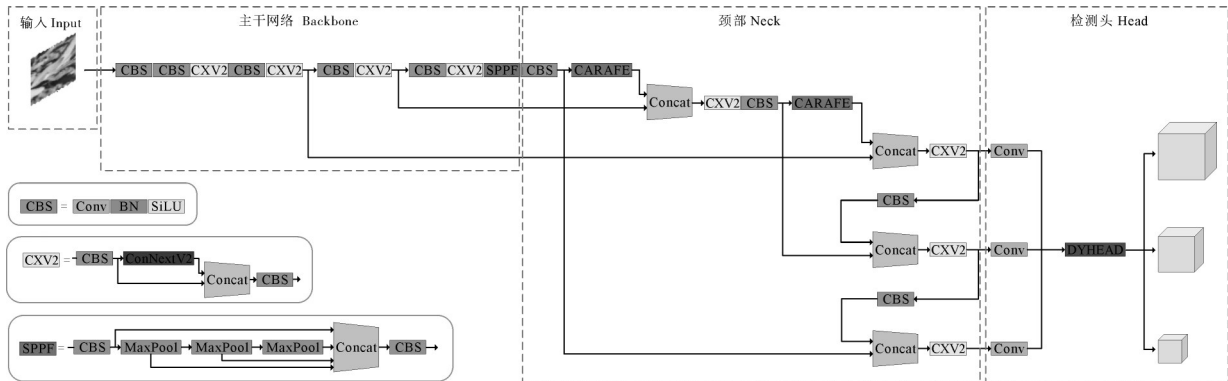


图2 YOLOv5s-CDC网络结构

Fig.2 Diagram of the YOLOv5s-CDC network

1)CXV2模块。为了提高模型的特征提取能力,获得更具分辨力的通道特征,增强模型的泛化能力,本研究引入ConvNeXtV2模型<sup>[22]</sup>,设计了1个CXV2模块用来替换YOLOv5s中的C3模块。ConvNeXtV2采用了全卷积掩码自编码器FCMAE机制,其将掩码自编码器MAE<sup>[23]</sup>机制与自监督学习技术相结合,提高模型的训练速度与精度。ConvNeXtV2模块的结构如图3所示,在ConvNeXt<sup>[24]</sup>的原有架构中添加了全局响应归一化(GRN)层,可以解决在训练过程中出现的特征崩溃问题。结合深度卷积操作来增强通道之间的特征竞争,提高通道对比度与选择性,有利于提高复杂背景图像中提取特征的多样性。

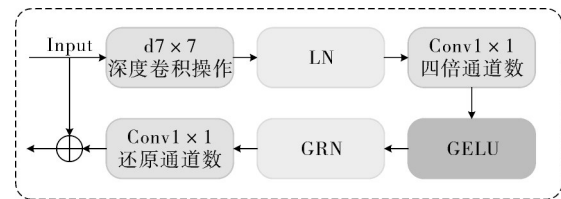


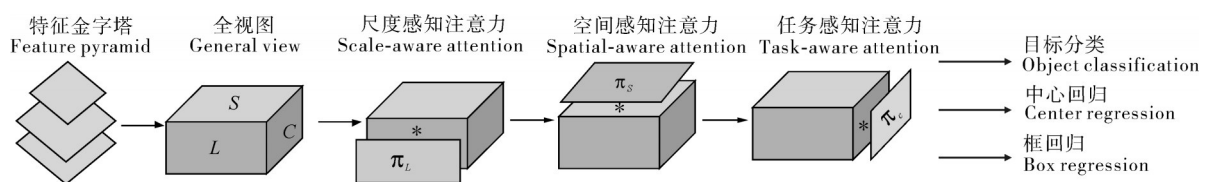
图3 ConvNeXtV2模块结构

Fig.3 Structure diagram of ConvNeXtV2

有分支进行拼接,并通过 $1 \times 1$ 标准卷积块操作恢复通道数,输出增强多样性的特征,提高模型特征表达能力。

CXV2模块中使用了类残差的结构,输入特征先经过1个 $1 \times 1$ 标准卷积块CBS,进行特征提取,减少通道数量,之后传递到2个不同的分支,其中1个分支特征被传递到ConvNeXtV2模块中,通过模块处理增强特征提取,另一分支特征保留不变,最后将所

2)DYHEAD动态检测头。柑橘不同病虫害具有病理特征相近的特点,且在自然环境中,易与背景相互混淆,病害叶片的尺度大小与不同空间位置也为特征提取与目标检测带来挑战。针对以上问题,本研究添加了动态检测头(dynamic head)<sup>[25]</sup>。动态检测头框架如图4所示,将原本的特征金字塔网络的特征进行尺度重调,重塑为具有相同尺度的三维张



L:代表特征层级 Feature hierarchy;S:代表特征图的高宽之积 Product of width and height;C:代表通道数 Number of channels.

图4 动态检测头框架

Fig.4 Framework of dynamic head



量  $F \in R^{L \times S \times C}$ , 并在每个独立维度采用不同注意力机制。

为避免全连接层带来的高计算量, 动态检测头采取3个序列注意力的方式, 其自注意力方程如公式(1)所示, 每个注意力模块仅对应一个维度。

$$W(F) = \pi_c(\pi_s(\pi_l(F) \cdot F) \cdot F) \cdot F \quad (1)$$

其中  $\pi_l$  为尺度感知注意力函数, 增加不同尺度特征感知能力;  $\pi_s$  为空间感知注意力函数, 增强模型的空间位置感知能力;  $\pi_c$  为任务感知注意力函数, 增强模型对不同任务目标的感知力。3个注意力机制模块如图5所示。在YOLOv5s主干网络后添加动态头, 旨在增强模型对不同目标尺度和特征层级的适应能力, 提升模型的检测精度和泛化能力。

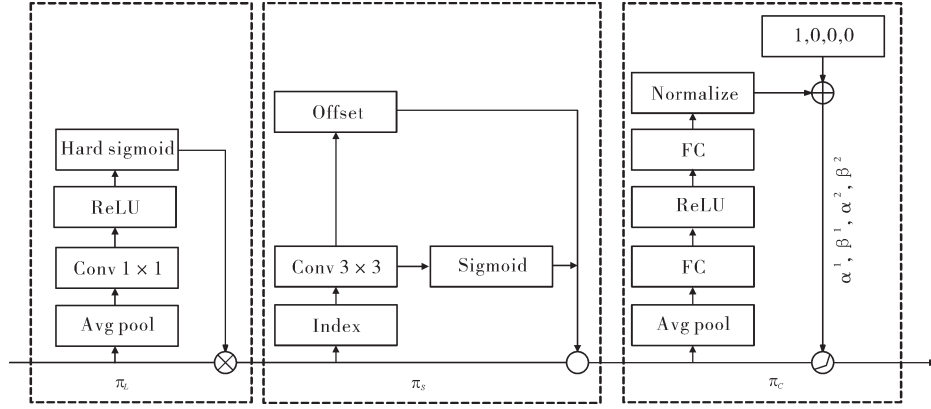
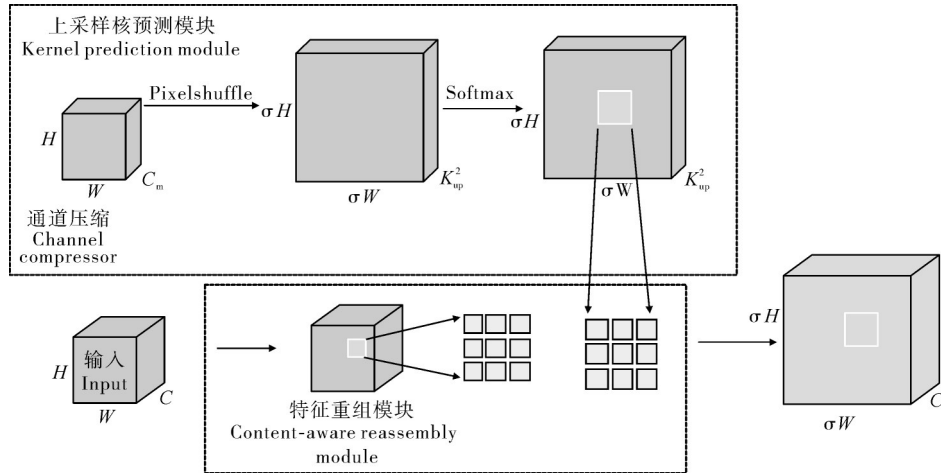


图5 动态检测头注意力机制模块

Fig.5 Attention mechanism module of dynamic head

3) CARAFE上采样模块。在YOLOv5s模型中, 特征金字塔是颈部特征提取与融合的重要组成部分, 模型需要通过上采样来将图像深层语义信息与浅层进行匹配与融合处理, YOLOv5s原模型中使用最近邻上采样的方法, 该方法只是通过像素点空

间位置决定上采样核, 从而导致感知域很小, 不能充分利用到特征图的语义信息。本研究采用CARAFE<sup>[26]</sup>上采样模块替换原模型的上采样算子。CARAFE构成如图6所示, 主要分为上采样核预测模块与特征重组模块2个部分。



$H$ : 表示特征层的高 Height of the feature layer;  $W$ : 表示特征层的宽 Width of the feature layer;  $\sigma$ : 表示上采样倍数 (通常为2) Upsampling multiple, usually set 2;  $C$ : 表示特征通道数 Number of feature channels;  $K_{up}$ : 表示预测的上采样核大小 Predicted upsampling kernel size.

图6 CARAFE模块结构

Fig.6 Structure diagram of CARAFE

在核预测模块中, 首先用  $1 \times 1$  的卷积核压缩通道数, 将原有通道数  $C$  降为  $C_m$ , 减少后续计算量, 提升效率。之后用 pixelshuffle<sup>[27]</sup> 的方式, 将特征图的高、宽、通道数重组, 对压缩通道数后的特征图用尺

寸为  $K_{encoder} \times K_{encoder}$  的卷积层预测上采样核, 预测得到的上采样核形状为  $\sigma H \times \sigma W \times K_{up}^2$ , 将该上采样核用 softmax 进行归一化处理。将特征图像传入特征重组模块, 用上采样核与特征图的每一层特征做

乘积,输出上采样工作的结果。CARAFE在不过多增加计算量的基础上,拥有较大的感知域,能根据不同特征生成不同上采样核,提高模型的全局特征提取能力。

### 1.3 实验环境配置

实验平台参数如表2所示,使用SGD优化器,训练轮次设置为300个Epoch,输入图像大小为640像素×640像素,学习率0.01,批次大小设置为16。为了防止过拟合,提高模型泛化能力,将mosaic数据增强设置为YOLOv5默认最佳值1,不使用mix-up。为了确保实验结果的可靠性与准确性,所有实验均不使用预训练权重进行训练。

表2 实验平台参数

Table 2 Experimental platform parameters

项目 Item	类型 Type
操作系统 Operating system	Windows 10
中央处理器 Central processing unit, CPU	12th Gen Intel(R) Core(TM) i5-12400F
图形处理器 Graphics processing unit, GPU	NVIDIA GeForce RTX 3060
开发环境 Development environment	Pycharm
语言 Language	Python3.7
加速环境 Accelerated environment	CUDA11.0
框架 Framework	torch1.7.1, torchvision0.8.2

### 1.4 评价指标

为了验证自然环境下模型对柑橘病虫害的检测性能,本研究采用深度学习中最常见的评价指标验证检测模型:精确率(precision,  $P$ )、召回率(recall,  $R$ )、平均精度均值(mean average presion,  $P_{mA}$ )、浮点运算量(floating-point operations, FLOPs)、参数量(parameters)。

精确率是所有检测为受害结果中真正受害数量所占比例。召回率是所检测到受害数量在实际受害总量中所占的比例。平均精度 $P_A$ 是基于精确度和召回率进行计算的, $P$ - $R$ 曲线下的面积就是 $P_A$ 的值。 $P_{mA}$ 是所有类别 $P_A$ 的平均值,它是目标检测的主要评估指标,能反映模型的整体性能,它的值越高则模型性能越好, $P_{mA}@0.5$ 指IoU设为0.5时的 $P_{mA}$ ;  $P_{mA}@.5:.95$ 表示在不同IoU阈值(0.5到0.95)上的平均 $P_{mA}$ 。FLOPs用于测量模型运行时间,它的值越低,则代表模型较小,模型执行计算量和所需时间也就越少。计算公式如下:

$$P = \frac{T_p}{T_p + F_p} \quad (2)$$

$$R = \frac{T_p}{T_p + F_N} \quad (3)$$

$$P_A = \int_0^1 P(R) dR \quad (4)$$

$$P_{mA} = \frac{1}{n} \sum_{i=1}^n P_A \quad (5)$$

其中, $T_p$ 代表真正例, $F_p$ 代表假正例, $F_N$ 代表假负例, $n$ 代表总类别数( $n=6$ )。

## 2 结果与分析

### 2.1 基线模型选择

为了更为高效和准确地实现对自然环境中柑橘病虫害的检测,从数据集中筛选仅含潜叶蛾危害这一类特征最为明显的数据对5个版本的YOLOv5进行测试,从中挑选最适合的基线模型。测试结果如表3所示,可以看出在柑橘病虫害的检测中,复杂的模型结构并不具有更为显著的检测效果,且会增加计算负担,降低网络效率。其中YOLOv5s有最高的检测精度,且计算参数与模型复杂度也相对较小,从检测效率、训练时间、计算成本等方面考虑,选择YOLOv5s作为本研究的基线模型。

表3 基线模型测试

Table 3 Baseline model testing

模型 Model	平均精度均值/% $P_{mA}@0.5$	浮点运算量 FLOPs	参数量 Parameters
YOLOv5n	93.8	$4.1 \times 10^9$	$1.76 \times 10^6$
YOLOv5s	95.3	$15.8 \times 10^9$	$7.01 \times 10^6$
YOLOv5m	93.8	$47.9 \times 10^9$	$20.85 \times 10^6$
YOLOv5l	93.2	$107.6 \times 10^9$	$46.11 \times 10^6$
YOLOv5x	92.9	$203.8 \times 10^9$	$86.17 \times 10^6$

### 2.2 消融实验

为了验证改进网络结构各个模块的有效性,以YOLOv5s模型为基线,在自建数据集上对基线网络逐个添加模块进行消融实验。训练环境参数与上述相同。如表4所示,各个模块的添加可以有效提高检测能力。添加DYHEAD动态检测头后,虽然精确率有所下降,但召回率提升0.9百分点, $P_{mA}@0.5$ 提升1.2百分点;CXV2模块的添加,关注高维度特征,提高了特征的多样性,其中 $P_{mA}@0.5$ 提升2.5百分点,召回率提升3.4百分点,提升明显;用不同尺寸核的CARAFE进行了试验,结果表明,在 $K_{up}$ 、 $K_{encoder}$ 都为3的情况下取得最佳 $P_{mA}$ 值87.3%(表5),说明过大的感知域并不能取得更好的效果,反而会增加模型计算成本。

表 4 模型消融实验结果

Table 4 Model ablation experimental results

基线网络 YOLOv5s	动态检测头 DYHEAD	特征提取 CXV2	上采样 CARAFE	精确率/% <i>P</i>	召回率/% <i>R</i>	平均精度均值/% <i>P</i> <sub>mA@0.5</sub>	平均精度均值/% <i>P</i> <sub>mA@.5:.95</sub>	浮点运算量 FLOPs
✓				92.7	76.7	83.9	69.0	16.0×10 <sup>9</sup>
✓	✓			91.6	77.6	85.1	72.3	16.8×10 <sup>9</sup>
✓		✓		90.2	80.0	84.8	71.6	16.2×10 <sup>9</sup>
✓			✓	89.6	79.1	84.7	69.9	16.0×10 <sup>9</sup>
✓	✓	✓		93.6	80.1	86.4	74.1	17.2×10 <sup>9</sup>
✓	✓	✓	✓	92.1	81.6	87.3	75.3	17.3×10 <sup>9</sup>

表 5 CARAFE 模块不同采样核参数性能对比

Table 5 Performance comparison of different sampling kernel parameters in module CARAFE

核参数 <i>K</i> <sub>encoder</sub>	核参数 <i>K</i> <sub>up</sub>	平均精度均值/% <i>P</i> <sub>mA@0.5</sub>
1	1	85.7
1	3	85.8
3	3	87.3
3	5	85.6
5	5	85.9

为验证改进模块的优异性,本研究采用双线性插值(bilinear)、双立方插值(bicubic)的上采样方式,

高效解耦头(efficient-decoupled head)、TSCODE(task-specific context decoupling)检测头模块与CARAFE、DYHEAD模块分别做横向对比实验,结果如表6~表7所示,CARAFE模块取得了最高的*P*<sub>mA@0.5</sub>值,且没有额外增加计算量;高效解耦头与DYHEAD模块虽有同样的检测精度,但FLOPs却多了11×10<sup>9</sup>,大大增加了计算负担。

最终 YOLOv5s-CDC 模型与基线模型相比,*P*<sub>mA@0.5</sub>提高3.4百分点,*P*<sub>mA@.5:.95</sub>提高6.3百分点,召回率提高4.9百分点,FLOPs仅增加1.3×10<sup>9</sup>,在略微提高模型计算量的基础上呈现较好的检测精度。

表 6 上采样模块横向对比

Table 6 Horizontal comparison of upsampling module

基线网络 YOLOv5s	上采样 CARAFE	双线性插值 bilinear	双立方插值 bicubic	精确率/% <i>P</i>	召回率/% <i>R</i>	平均精度均值/% <i>P</i> <sub>mA@0.5</sub>	平均精度均值/% <i>P</i> <sub>mA@.5:.95</sub>	浮点运算量 FLOPs
✓				92.7	76.7	83.9	69.0	16.0×10 <sup>9</sup>
	✓			89.6	79.1	84.7	69.9	16.0×10 <sup>9</sup>
		✓		90.9	76.9	83.6	69.0	16.0×10 <sup>9</sup>
			✓	91.5	78.4	84.4	69.8	16.4×10 <sup>9</sup>

表 7 检测头模块横向对比

Table 7 Horizontal comparison of detection head module

基线网络 YOLOv5s	动态头检测头 DYHEAD	高效解耦头 Efficient	TSCODE	精确率/% <i>P</i>	召回率/% <i>R</i>	平均精度均值/% <i>P</i> <sub>mA@0.5</sub>	平均精度均值/% <i>P</i> <sub>mA@.5:.95</sub>	浮点运算量 FLOPs
✓				92.7	76.7	83.9	69.0	16.0×10 <sup>9</sup>
	✓			91.6	77.6	85.1	72.3	16.8×10 <sup>9</sup>
		✓		91.9	77.8	85.1	72.0	27.8×10 <sup>9</sup>
			✓	88.8	76.6	82.2	67.6	16.3×10 <sup>9</sup>

模型总体损失如图7所示,可以看出基线模型的损失在训练前期有较大的波动,而改进模型的损失曲线则较为平滑稳定,最终取得收敛,并且损失低于基线模型,没有出现拟合、欠拟合的情况。模型平均精度如图8所示,在训练的前期改进模型与基线模型有重合交错,在第30轮左右改进的模型便开始超越基线模型,稳定提升并收敛,最终取得比基线模型更高的检测精度。

2.3 不同模型检测结果对比

为验证改进模型性能,将改进模型与SSD、YOLOv3-tiny、YOLOv4、YOLOv7-tiny、YOLOv8s模型5种主流的目标检测算法进行了比较,结果如表8所示,改进后的算法有着最高的召回率与平均精度均值。与系列最新、性能较强的YOLOv8s相比,虽然在精确率上改进算法低了0.5百分点,但是在召回率上高了6.2百分点,*P*<sub>mA@0.5</sub>提高了2.6百分点,呈现

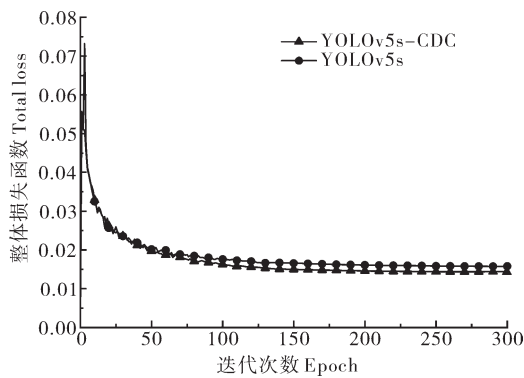


图7 模型整体损失比较

Fig.7 Comparison of total loss

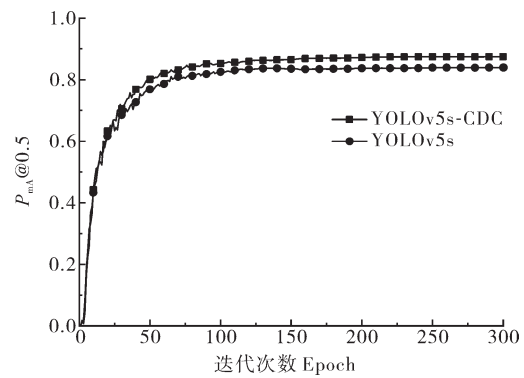


图8 模型  $P_{mA}$  比较

Fig.8 Comparison of  $P_{mA}$

较大的提升,且参数量和浮点运算量远小于YO-LOv8s。由表9可知,改进后的模型在各类病害中都取得了最高的检测精度,在总体性能上改进的算法优于YOLOv8s和其他几种算法,并且适合移动设备使用。

表8 不同模型检测性能对比

Table 8 Comparison of different detection models

模型 Model	精确率/% $P$	召回率/% $R$	平均精度均值/% $P_{mA}@0.5$	平均精度均值/% $P_{mA}@.5:.95$	浮点运算量 FLOPs	参数量 Parameters
SSD	82.4	44.5	49.6	30.3	$62.7 \times 10^9$	$26.29 \times 10^6$
YOLOv3-tiny	85.6	70.3	77.3	55.9	$12.9 \times 10^9$	$8.68 \times 10^6$
YOLOv4	87.2	79.7	84.6	73.4	$119.1 \times 10^9$	$52.5 \times 10^6$
YOLOv5s	92.7	76.7	83.9	69.0	$16.0 \times 10^9$	$7.03 \times 10^6$
YOLOv7-tiny	81.3	78.3	82.5	64.4	$13.1 \times 10^9$	$6.02 \times 10^6$
YOLOv8s	92.6	75.4	84.7	75.1	$28.4 \times 10^9$	$11.13 \times 10^6$
YOLOv5s-CDC	92.1	81.6	87.3	75.3	$17.3 \times 10^9$	$7.61 \times 10^6$

表9 不同模型中各类病虫害  $P_A$  对比

Table 9 Comparison of  $P_A$  for diseases in different models

模型 Model	黄龙病 Yellow shoot disease	潜叶蛾 Leaf miner	树脂病 Melanose	蚜虫 Aphid	锈壁虱 Rust mite	蚧壳虫 Coccid
SSD	57.2	30.1	43.1	62.6	40.1	64.5
YOLOv3-tiny	56.5	88.9	82.8	76.7	81.7	77.3
YOLOv4	64.8	96.3	89.3	81.9	88.1	87.0
YOLOv5s	60.6	97.1	87.3	83.4	88.9	86.3
YOLOv7-tiny	63.0	95.0	86.0	77.8	89.7	83.7
YOLOv8s	60.5	96.5	88.1	85.8	89.7	87.8
YOLOv5s-CDC	65.0	97.8	89.5	88.0	93.5	90.0

利用上述几种算法与改进后的YOLOv5s-CDC模型对6种柑橘病虫害进行检测,验证改进模型在互相遮盖、阴影遮挡、特征相似、密集等复杂场景中的实际检测效果。结果如图9所示,在各类型的病虫害检测中,YOLOv5s-CDC能较准确地检测出图中目标,而其余的各种算法都出现了对健康叶片及其他背景因素的误检;在检测病害特征相似的树脂病害与蚧壳虫虫害时,SSD与YOLOv7-tiny模型出现了将两者混淆误检的现象;在密集多目标场景的检测中,YOLOv5s-CDC漏检情况有明显改善,其余几种

算法都出现了较大程度的漏检。试验表明YOLOv5s-CDC在各类病虫害场景中能更好地检测出受害叶片并正确分类,取得较高的置信度,证明了改进模型具有较好的鲁棒性。

### 3 讨论

本研究将深度学习目标检测技术用于柑橘病虫害检测,对于柑橘果园防治、精准施药、保证柑橘产量有较大的现实意义。为了提高在自然果园的复杂背景下多种类、多目标的柑橘病虫害检测准确率,提



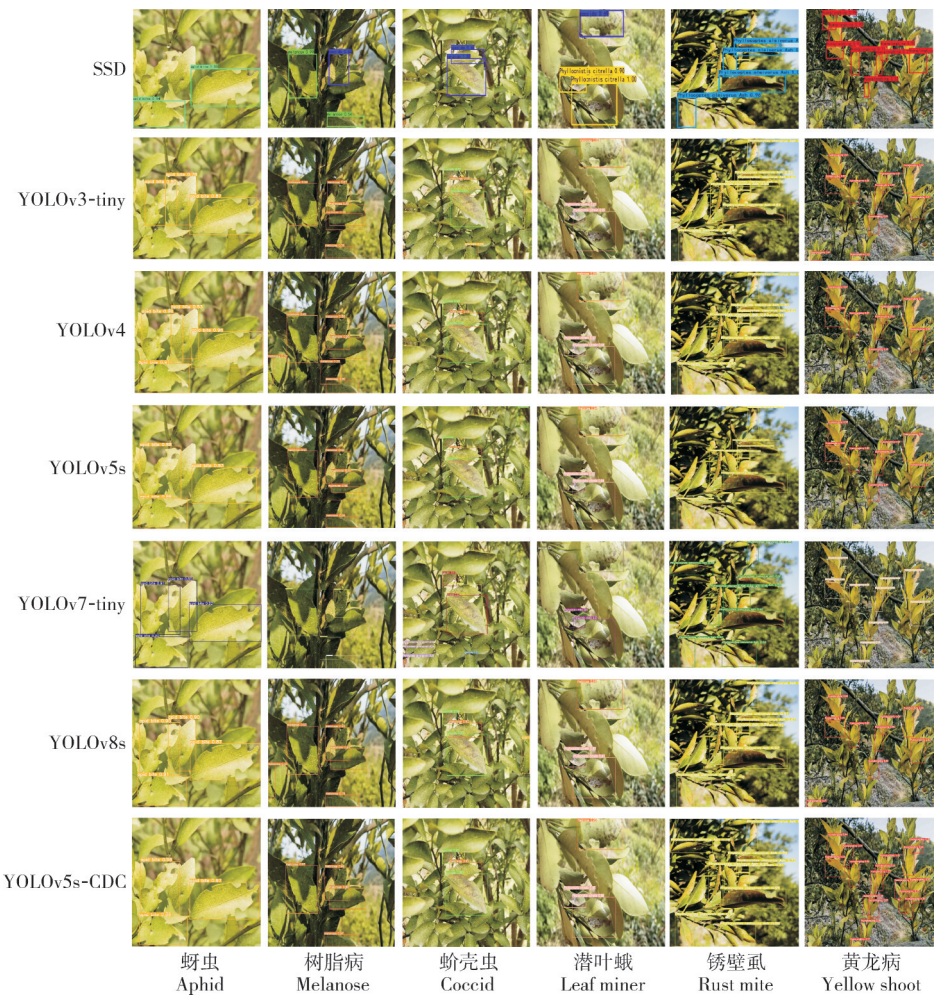


图9 不同场景下不同模型检测效果对比

Fig.9 Comparison of model detection effects in different scenarios

出了1种基于YOLOv5s的改进检测算法。该方法使用CXV2模块来替换原模型的C3模块,模块中引入ConvNeXtV2模块,增强特征提取能力;使用动态检测头代替原有检测头,增强对不同空间尺度、不同任务目标的感知力;用轻量级上采样算子CARAFE替换原有上采样方法,加强上采样工作。研究结果表明,该算法能有效提取柑橘病虫害特征,在自建数据集上的 $P_{m\lambda}@0.5$ 为87.3%,比原YOLOv5s模型提高了3.4百分点,浮点运算量为 $17.3\times10^9$ ,仅略微增加 $1.3\times10^9$ 。与其他常用算法相比,该模型具有较好的检测效果和较强的鲁棒性,证明了该模型的有效性,能满足在复杂环境下的检测需求,由于较少的参数量,便于实行快速检测与移动端部署。与宋中山等<sup>[28]</sup>的基于二值化的Faster R-CNN柑橘病虫害检测模型(4类病害,数据样本为单张叶片)相比,本研究的病虫害类型更丰富,且数据样本包含多目标与复杂的背景环境,因此模型具备更强的泛化能力,更

适用于自然田园环境下的检测。

本研究在计算机设备上进行,尚未部署到移动设备进行实地测试,所以在实际环境中可能存在一定误差。在后续的研究工作中,可以针对单类病虫害改进模型,加强对个别病虫害的检测效果,在模型的参数量和计算量上也可以进一步优化,以便于将模型部署到移动设备,实现模型的实际运用。

参考文献 References

[1] 聂扬眉. 乡村振兴下我国柑橘产业发展情况:基于价值链模型的实证[J]. 中南民族大学学报(自然科学版), 2023, 42(2): 283-288. NIE Y M. China's citrus industry under rural revitalization: empirical evidence based on the value chain model [J]. Journal of South-Central Minzu University (natural science edition), 2023, 42(2): 283-288 (in Chinese with English abstract).

[2] PUJARI J D, YAKKUNDIMATH R, BYADGI A S. SVM



- and ANN based classification of plant diseases using feature reduction technique[J]. International journal of interactive multimedia and artificial intelligence, 2016, 3(7): 6-14.
- [3] SHARIF M, KHAN M A, IQBAL Z, et al. Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection [J]. Computers and electronics in agriculture, 2018, 150: 220-234.
- [4] 杨英茹, 吴华瑞, 张燕, 等. 基于复杂环境的番茄叶部图像病虫害识别[J]. 中国农机化学报, 2021, 42(9): 177-186. YANG Y R, WU H R, ZHANG Y, et al. Tomato disease recognition using leaf image based on complex environment [J]. Journal of Chinese agricultural mechanization, 2021, 42(9): 177-186 (in Chinese with English abstract).
- [5] BALAKRISHNA K, RAO M. Tomato plant leaves disease classification using KNN and PNN[J]. International journal of computer vision and image processing, 2019, 9(1): 51-63.
- [6] 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.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C]//LEIBE B, MATAS J, SEBE N, et al. Computer vision - ECCV 2016: 14th European conference, October 11-14, 2016, Amsterdam, Switzerland; Springer, 2016: 21-37.
- [8] REDMON J, FARHADI A. YOLOv3: an incremental improvement[DB/OL]. arXiv, 2018: 1804.02767 [2023-08-24]. <https://doi.org/10.48550/arXiv.1804.02767>.
- [9] BOCHKOVSKIY A, WANG C-Y, LIAO H-Y M. Yolo v4: Optimal speed and accuracy of object detection[DB/OL]. arXiv, 2020: 2004.10934 [2023-08-24]. <https://doi.org/10.48550/arXiv.2004.10934>.
- [10] ZHAO S Y, LIU J Z, WU S. Multiple disease detection method for greenhouse-cultivated strawberry based on multiscale feature fusion Faster R-CNN[J/OL]. Computers and electronics in agriculture, 2022, 199: 107176 [2023-08-24]. <https://doi.org/10.1016/j.compag.2022.107176>.
- [11] SUN H N, XU H W, LIU B, et al. MEAN-SSD: a novel real-time detector for apple leaf diseases using improved lightweight convolutional neural networks[J/OL]. Computers and electronics in agriculture, 2021, 189: 106379 [2023-08-24]. <https://doi.org/10.1016/j.compag.2021.106379>.
- [12] ZHANG X X, XUN Y, CHEN Y H. Automated identification of citrus diseases in orchards using deep learning[J]. Biosystems engineering, 2022, 223: 249-258.
- [13] LI S, LI K, QIAO Y, et al. A multi-scale cucumber disease detection method in natural scenes based on YOLOv5[J/OL]. Computers and electronics in agriculture, 2022, 202: 107363 [2023-08-24]. <https://doi.org/10.1016/j.compag.2022.107363>.
- [14] 孙丰刚, 王云露, 兰鹏, 等. 基于改进 YOLOv5s 和迁移学习的苹果果实病害识别方法[J]. 农业工程学报, 2022, 38(11): 171-179. SUN F G, WANG Y L, LAN P, et al. Identification of apple fruit diseases using improved YOLOv5s and transfer learning [J]. Transactions of the CSAE, 2022, 38(11): 171-179 (in Chinese with English abstract).
- [15] MOHAMMAD M, AHMAD J, ALI ASGHAR N, et al. Detection of citrus black spot disease and ripeness level in orange fruit using learning-to-augment incorporated deep networks[J/OL]. Ecological informatics, 2022, 71: 101829 [2023-08-24]. <https://doi.org/10.1016/j.ecoinf.2022.101829>.
- [16] KHANRAMAKI M, ASKARI ASLI-ARDEH E, KOZEGAR E. Citrus pests classification using an ensemble of deep learning models[J/OL]. Computers and electronics in agriculture, 2021, 186: 106192 [2023-08-24]. <https://doi.org/10.1016/j.compag.2021.106192>.
- [17] SYED-AB-RAHMAN S F, HESAMIAN M H, PRASAD M. Citrus disease detection and classification using end-to-end anchor-based deep learning model [J]. Applied intelligence, 2022, 52(1): 927-938.
- [18] BARMAN U, CHOUDHURY R D, SAHU D, et al. Comparison of convolution neural networks for smartphone image based real time classification of citrus leaf disease[J/OL]. Computers and electronics in agriculture, 2020, 177: 105661 [2023-08-24]. <https://doi.org/10.1016/j.compag.2020.105661>.
- [19] LUAIBI A R, SALMAN T M, MIRY A H. Detection of citrus leaf diseases using a deep learning technique [J]. International journal of electrical and computer engineering (IJECE), 2021, 11(2): 1719-1727.
- [20] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 936-944.
- [21] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C]//2018 IEEE/CVF conference on computer vision and pattern recognition. June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 8759-8768.
- [22] WOO S, DEBNATH S, HU R H, et al. Convnext v2: co-designing and scaling convnets with masked autoencoders [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 17-24, 2023, Vancouver, BC, Canada. New York: IEEE, 2023: 16133-16142.
- [23] HE K M, CHEN X L, XIE S N, et al. Masked autoencoders are scalable vision learners [C]//2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR). June 18-24, 2022, New Orleans, LA, USA. New York: IEEE, 2022: 15979-15988.
- [24] LIU Z, MAO H Z, WU C Y, et al. A convnet for the 2020s [C]//2022 IEEE/CVF conference on computer vision and

- pattern recognition (CVPR). June 18-24, 2022, New Orleans, LA, USA. New York: IEEE, 2022: 11966-11976.
- [25] DAI X Y, CHEN Y P, XIAO B, et al. Dynamic head: unifying object detection heads with attentions [C]//2021 IEEE/CVF conference on computer vision and pattern recognition (CVPR). June 20-25, 2021, Nashville, TN, USA. New York: IEEE, 2021: 7369-7378.
- [26] 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). October 27 - November 2, 2019, Seoul, Korea (South). New York: IEEE, 2020: 3007-3016.
- [27] SHI W Z, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]//2016 IEEE conference on computer vision and pattern recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1874-1883.
- [28] 宋中山, 汪进, 郑禄, 等. 基于二值化的Faster R-CNN柑橘病虫害识别研究[J]. 中国农机化学报, 2022, 43(6): 150-158.
- SONG Z S, WANG J, ZHENG L, et al. Research on citrus pest identification based on Binary Faster R-CNN[J]. Journal of Chinese agricultural mechanization, 2022, 43(6): 150-158 (in Chinese with English abstract).

## Improved YOLOv5s based identification of pests and diseases in citrus

ZHENG Yuda<sup>1</sup>, CHEN Renfan<sup>1</sup>, YANG Changcai<sup>2</sup>, ZOU Tengyue<sup>1</sup>

1. College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou 350002, China;

2. College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China

**Abstract** Accurately identifying pests and diseases in citrus can be used to timely reduce the economic losses. A common method for detecting pests and diseases in citrus based on the improved YOLOv5s model was proposed to solve the problems that the existing models of detection cannot accurately identify multiple types of pests and diseases of citrus in the natural environment. The model was improved by introducing the ConvNeXtV2 model and constructing a CXV2 module to replace the C3 module of YOLOv5s, enhancing the diversity of extracted feature. The dynamic detection head DYHEAD was added to improve the processing ability of the model for different spatial scales and task targets. The CARAFE upsampling module was used to improve the efficiency of extracting feature. The results showed that the improved YOLOv5s-CDC had a mean recall rate and average precision of 81.6% and 87.3%, 4.9 percentage points and 3.4 percentage points higher than that of the original model, respectively. Compared with the detection with other YOLO serial models in multiple scenarios, it had higher accuracy and stronger robustness. It is indicated that this method can be used for detecting the diseases and pests of citrus in complex natural environments.

**Keywords** deep learning; pests and diseases; YOLOv5s; target detection

(责任编辑: 葛晓霞)