

改进的YOLOv7茶叶病害识别模型

李治林, 李遇鑫, 谢本亮*

贵州大学大数据与信息工程学院, 贵州 贵阳

收稿日期: 2023年9月27日; 录用日期: 2023年11月22日; 发布日期: 2023年11月29日

摘要

针对茶叶病害图像背景复杂、目标小、易漏检等问题, 提出一种改进YOLOv7的茶叶病害识别模型。该模型首先引入混合注意力模块ACmix加强对小目标的敏感度, 解决茶叶病害目标小, 易漏检的问题。其次, 采用C3模块替换Neck部分的ELAN-W模块以提高网络性能。最后使用Alpha-IoU损失函数优化原YOLOv7模型中的CIoU损失函数, 提升模型对检测目标的定位能力。实验结果表明, 改进后模型的平均检测精度mAP达到93.3%, 比YOLOv7模型提高了1.8%, 在FPS增加的同时模型参数量降低了3.5 M。该研究内容可以为茶园病害的智能化监控设备提供支持。

关键词

目标检测, 茶叶病害, YOLOv7, ACmix, 损失函数

Improved Tea Disease Recognition Model for YOLOv7

Zhilin Li, Yuxin Li, Benliang Xie*

College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

Received: Sep. 27th, 2023; accepted: Nov. 22nd, 2023; published: Nov. 29th, 2023

Abstract

An improved YOLOv7 model is proposed to address the challenges of complex background, small target size, and easy omission in tea disease image recognition. Firstly, the model introduces the Hybrid Attention module ACmix to enhance sensitivity towards small targets and solve the problem of easily missing tea disease targets. Secondly, the C3 module is used to replace the ELAN-W module in the Neck part to improve network performance. Lastly, the Alpha-IoU loss function is used to optimize the original YOLOv7 model's CIoU loss function, enhancing the model's ability to

*通讯作者。

locate detection targets. Experimental results show that the improved model achieves an average precision of 93.3%, a 1.8% improvement over the YOLOv7 model, while reducing model parameters by 3.5 M, without decreasing the FPS. This research can provide support for intelligent monitoring devices for tea garden diseases.

Keywords

Target Detection, Tea Disease, YOLOv7, ACmix, Loss Function

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

茶叶是我国的主要经济作物之一，在农村产业结构调整，增加农民收入，加速农民致富奔小康中起到积极的作用[1]。然而，茶树在种植过程中容易受到细菌、真菌和病毒等的入侵而感染病害，这些病害会大大降低茶叶的产量和质量。若不能及时发现这些病害并加以控制，将会导致茶叶种植者和整个行业的巨大经济损失[2]。传统上，对于茶叶病害的诊断主要是依赖于有经验的生产者的视觉观察[3]。然而，这种方法主观性强、耗时长，容易出错。此外，这种方法需要对于茶叶病害有专业的知识，不具有普适性。因此，寻找一种快速、准确地识别对茶叶病害进行诊断具有重要的现实意义。

随着计算机技术的发展，众多研究人员使用机器学习和图像处理方法来鉴定农作物疾病[4] [5]。余秀丽等[6]对小麦叶部病害图像去噪及分割病斑，然后提取病斑特征，最后使用 SVM 算法对小麦叶部病害图像进行分类识别。张云龙等[7]使用分割算法对病斑区域进行分割，并计算病斑的颜色特征和差直方图，结合 SVM 分类器，对 3 种常见苹果叶部病害图像识别率达 96.52%。林彬彬等[8]在分割病斑区域基础上，使用多种方法提取病斑的颜色特征，进行特征选择后使用 K 近邻算法进行茶叶病害的智能诊断识别。然而早期机器学习方法需要人为进行特征提取和选择，较为复杂且识别效果差，难以推广和应用。

近年来，深度学习的快速发展，卷积神经网络由于其出色的特征提取能力，因此越来越多的学者使用卷积神经网络对农作物病害进行识别[9] [10] [11]。贾璐等[12]提出了一种 MANet 葡萄病害识别模型，对 4 种常见葡萄叶片图像的识别准确率达 87.93%。杨康等[13]基于 SSD 模型架构，使用 MobileNet 替换 SSD 模型的原始主干网络，并构建了 RFB 模块实现多尺度检测，对 4 类三七病害的检测精度达 82.0%。李就好等[14]对 Faster-RCNN 进行改进，增加建议框的尺寸个数，并在特征提取网络 ResNet-50 的基础下融入了 FPN 网络，改进网络的识别准确率达到 86.39%。孙道宗等[15]对 YOLOv4 模型的主干网络进行改进，并在模型中嵌入 CBAM 注意力机制，对 3 类病害的识别精度达 93.85%。杨文姬等[16]设计一种基于改进 Yolov5 的深度学习方法，用于检测苹果和番茄病害，改进后的 Yolov5 病害检测算法 mAP 达到 95.7%。李伟豪等[17]提出了一种 Yolov7-TSA 模型对茶叶病害检测和分类，对 7 类茶叶病害和健康茶叶的识别准确率达 94.2%。然而，尽管上述研究都取得了较好的识别结果，但仍存在一些问题。如模型参数量大，难以部署在茶园边缘设备上，识别准确率不佳等。为解决这些问题，本文提出了一种改进 YOLOv7 的茶叶病害识别模型。

2. 数据集描述

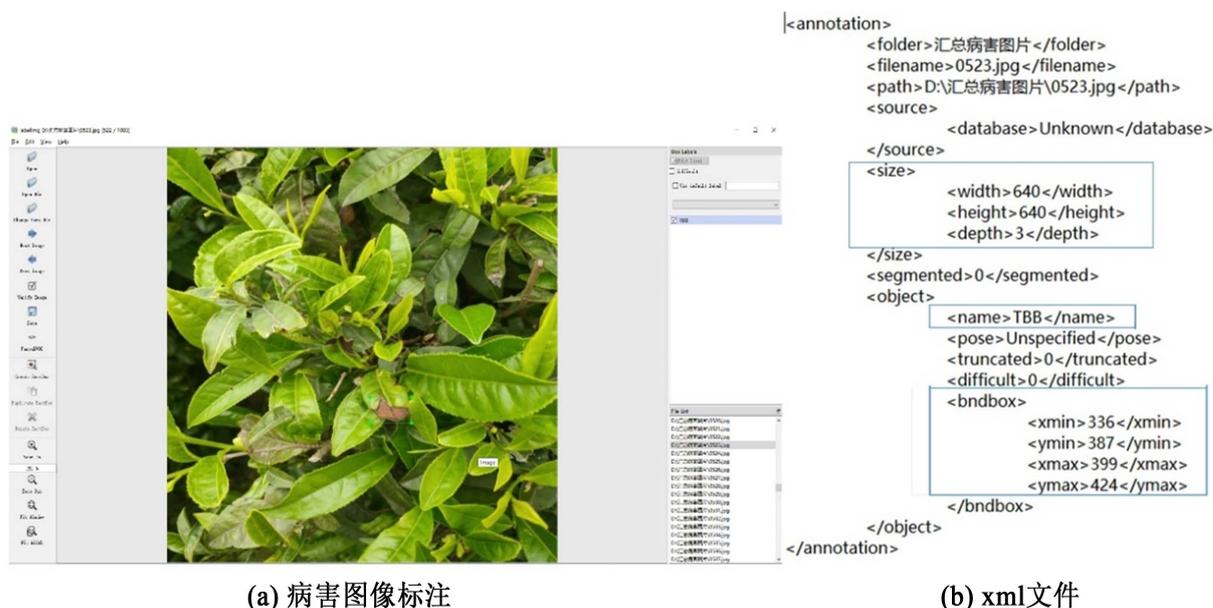
2.1. 数据采集

本文实验数据分别采集自贵州省贵阳市花溪区久安古茶园和羊艾茶园，拍摄时间分别为 2023 年 4 月

1日至7日时间段,此时正处于茶叶病害的高发期,便于病害图像的采集。拍摄时,手持拍摄设备(OPPO Reno Ace 智能手机)距茶叶病害叶片10~20 cm处进行拍摄,保存格式为JPG。在不同光照条件和不同成像背景情况下,采集到4类茶常见茶叶病害图像共1083张,其中,茶芽枯病(TBB)335张,茶轮斑病(TGB)312张,茶藻斑病(TALS)221张,茶叶枯病(TLB)215张。

2.2. 数据集预处理

训练深度学习模型时,训练样本的多样性能增强模型的学习能力,提高模型的泛化能力。由于茶树大多生长在崎岖山区,数据采集较为困难,采集到的样本有限,因此考虑采用数据增强方法来扩充样本。我们使用随机旋转、随机翻转(水平和垂直两个方向)、高斯模糊、亮度调整和添加高斯噪声等五种数据增强方法对部分茶叶病害样本进行增强,增强结束后共得到4333张茶叶病害图像。对增强后的茶叶病害图像,使用LabelImg标注工具对进行标注,标注后生成与每张病害图像种类信息、病害位置信息相对应的xml文件,与茶叶病害图像组成最终的数据集。图1显示了图像标注示例和标注后生成的对应xml文件。最后,按照7:3将最终数据集划分为训练集和验证集,其中训练集图像数量为3033张,验证集图像数量为1300张。



xml文件中,第一个框为病害图片信息,大小为 $640 \times 640 \times 3$,第二个框为病害种类名称,第三个框为标注框的位置信息。

Figure 1. Example of disease image annotation with generated xml file

图1. 病害图像标注示例与生成xml文件

3. 改进的YOLOv7茶叶病害检测模型设计

3.1. YOLOv7 简述

YOLOv7模型由Chien-Yao Wang等人[18]于2022年7月提出,较YOLO系列的其它模型相比,其具有更快的检测速度、更高的识别精度、易于部署等优点。YOLOv7的检测思路与YOLOv5相似,其结构如图2所示。从结构图可以看到它由四个部分组成,即输入模块(Input)、主干网络(Backbone)、特征融合网络(Neck)和预测网络(Head)。

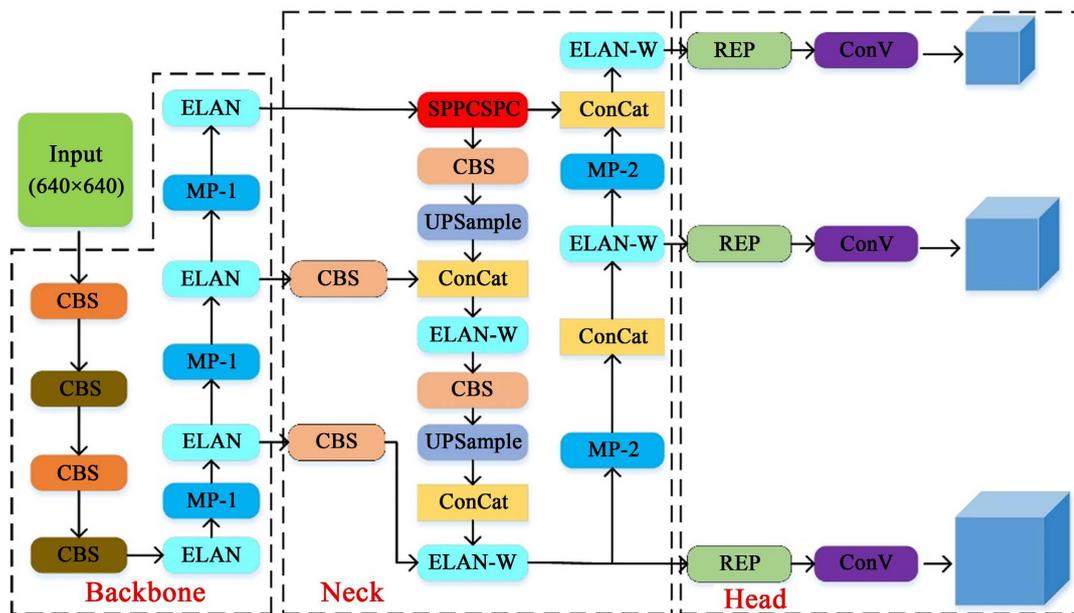


Figure 2. YOLOv7 network architecture diagram
图 2. YOLOv7 网络结构图

其中，输入模块主要是将输入图像尺寸统一缩放到 640×640 大小，并对输入图像进行数据增强操作等预处理。主干网络由多个基础卷积模块(CBS)、扩展高效聚合网络(ELAN)和最大池化卷积模块(MP)组成，主要负责对输入图像进行特征提取，得到三个不同尺度的特征图。特征融合网络由路径聚合特征金字塔网络(PAFPN)组成，通过引入自上而下的路径，使得底层信息更容易向高层传递，从而实现不同层次特征的高效整合。预测网络由重参数化卷积(RepConv)对从特征融合网络输入的三个特征图通道数进行调整，最后分别通过 1×1 卷积预测目标的置信度、类别和位置。

3.2. 改进的 YOLOV7 模型

3.2.1. 注意力机制

在注意力机制中，卷积注意力模块多注重输入与输出间的关系，而自注意力模块则着重于输入与输入之间内部关联。因此，人们通常认为它们是独立的两个注意力模块。但潘旭然等发现卷积注意力模块和自注意力模块都严重依赖 1×1 卷积操作，基于这一发现提出了一种混合注意力机制 ACmix 模块[19]，其结构如图 3 所示。该模块结合了卷积操作和自注意力模块的优点，具有更强的特征提取能力，更小的计算开销。

具体来说，ACmix 模块原理如下：第一阶段，先通过 3 个 1×1 卷积将 $H \times W \times C$ 的输入特征进行投影，并分别重构为 N 部分，得到 $3N$ 个尺寸为 $H \times W \times C/N$ 的中间特征。第二阶段，对于卷积路径，中间特征通过 $3N \times k^2 N$ 的全连接层后，对生成的特征进行移位和聚合以及卷积处理，得到 $H \times W \times C$ 的特征；对于自注意力路径，将中间特征聚合到 N 组中，每组包含三个特征，每个特征来自不同的 1×1 卷积输出，对应的 3 个尺寸为 $H \times W \times C/N$ 的特征图分别作为查询、键和值，并遵循传统的多头自注意力模型，通过移位、聚合、卷积处理得到 $H \times W \times C$ 的特征。最后，将两条路径的输出进行相加操作，强度由两个可学习的标量控制，如式(1)所示：

$$F_{out} = \alpha F_{att} + \beta F_{conv} \quad (1)$$

式中， F_{out} 表示路径的最终输出； F_{att} 表示自注意力分支的输出； F_{conv} 表示卷积注意力分支的输出； α 和 β 的值均为 1。

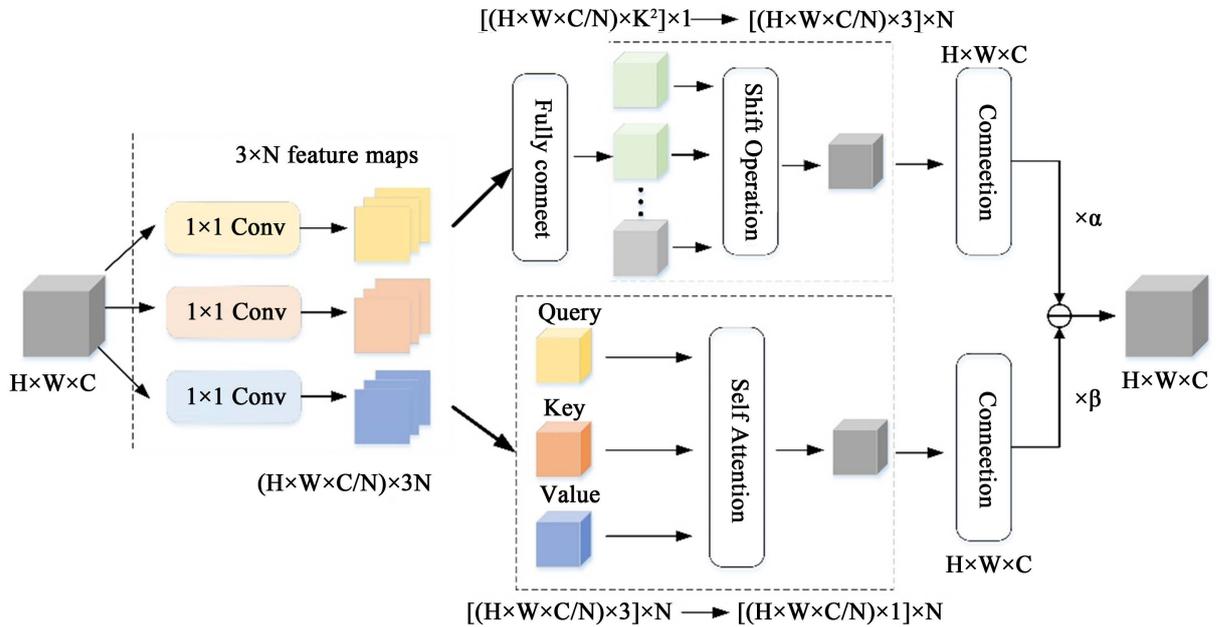


Figure 3. ACmix attention module
图 3. ACmix 注意力模块

3.2.2. C3 模块

YOLOv7 网络中包含了大量的 E-ELAN 模块，该模块在特征提取方面效果不佳，此外 E-ELAN 结构中包含了较多卷积模块和残差连接，会使网络的计算量增加，推理速度降低。C3 模块为 YOLOv5 [20] 中重要特征提取模块，它由三个标准卷积和多个瓶颈层组成，结构图如图 4 所示。其原理是一个分支经过 1×1 卷积降低特征通道数，另一个分支经过 1×1 卷积和 N 个 Bottleneck 先降低通道数再进一步提取和增强特征信息，最后两个分支拼接在一起并通过 1×1 卷积来融合它们。

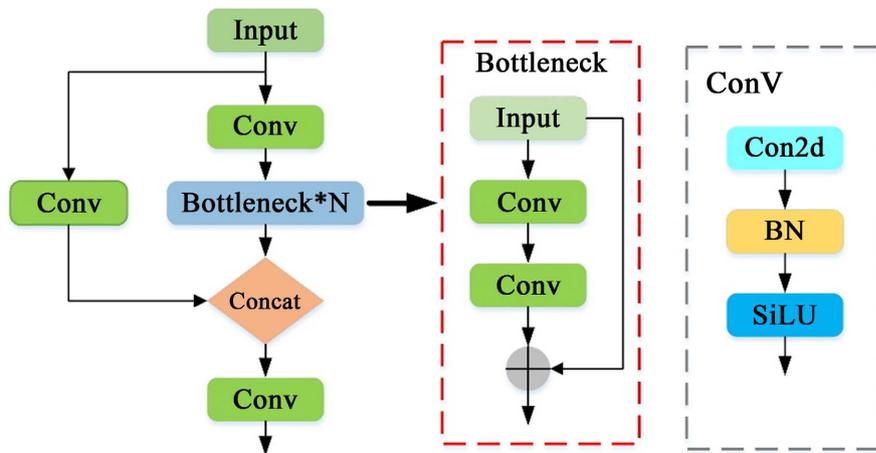


Figure 4. C3 module structure
图 4. C3 模块结构

C3 结构采用了多尺度特征融合和跨通道信息传递，增强了模型的特征表达能力。因此，为解决使用 E-ELAN 模块带来的不利影响，本文拟将特征融合网络部分 E-ELAN 替换为 C3 模块，以进一步增强 YOLOv7 网络的特征提取能提和融合能力，同时提高模型推理速度。

3.2.3. 改进损失函数

目标检测是通过边界框回归来定位图像中的物体，因此良好的边界框损失函数对于准确定位目标极其重要。在早期的目标检测任务中，常采用 IoU 作为定位损失函数。但 IoU 损失存在两个问题：一是预测框与真实框不相交时，IoU 损失会造成梯度消失问题，导致收敛速度变慢，检测不准确。二是预测框和真实框相交时无法反映重叠程度。为了解决这两个问题，后来研究人员又提出了几种改进 IoU 损失，如 GIoU、DIoU 和 CIoU 等等。在 YOLOv7 模型中，便采用了 CIoU 损失函数进行边界框回归。CIoU 损失函数考虑了三个几何因素，包括归一化中心点距离的最小化以及重叠区域和长宽比的一致性。但当预测框与真实框的长宽比相同且共中心点时 CIoU 会退化为 IOU，CIoU 损失函数将得不到稳定表达。

基于上述原因，本文采用 Alpha-IoU 损失函数[21]改进原网络中的 CIoU 损失函数。Alpha-IoU 损失函数通过控制功率参数 α ，提高了边界框回归的精度，同时考虑了距离、角度及重叠面积这三个关键因素，从而加快了模型收敛速度。改进损失函数如式(2)所示：

$$L_{\alpha-CIoU} = 1 - IoU^\alpha + \frac{\rho^{2\alpha}(b, b^{gt})}{c^{2\alpha}} + (\beta v)^\alpha$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (2)$$

$$\beta = \frac{v}{1 - IoU + v}$$

式中， $\rho^{2\alpha}$ 代表 2 个中心点之间的欧氏距离， (b, b^{gt}) 表示预测框和真实框的中心点坐标， c 表示能够同时包含预测框和真实框的最小外接矩形的对角线距离， $\frac{w^{gt}}{h^{gt}}$ 表示真实框的宽高比， $\frac{w}{h}$ 表示预测框的宽高比。本实验中 α 设为 3，更有利于提升模型的精度。

4. 实验结果与分析

4.1. 实验环境

本实验使用 python 语言并基于 PyTorch 深度学习框架进行开发，硬件及软件配置如表 1 所示。

Table 1. Configuration of the experimental environment

表 1. 实验环境配置表

类别	相关配置
CPU	Intel(R) Core(TM) i5-11400F
GPU	NVIDIA GeForce RTX 3060
内存	16 GB
硬盘	1TB
操作系统	Windows 10 64 位

实验中输入图像尺寸为 640×640 大小，训练迭代次数为 500 轮，批量大小为 16，优化器为 SGD 优化器，学习率设为 0.001。

4.2. 评价指标

为了评估改进模型对茶叶病害识别性能，采用查准率(Precision, P)、查全率(Recall, R)、平均准确率

(mAP)、检测速度(FPS)、模型参数量(Params)作为评价指标。各评价指标定义公式如下:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (3)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (4)$$

$$mAP = \frac{1}{N} \int_0^1 Precision(Recall) d(Recall) \quad (5)$$

$$FPS = \frac{frameNum}{elapsedTime} \quad (6)$$

式中, TP 为检测出正确的正样本数; FP 为检测出错误的正样本数; FN 为检测出错误的负样本数; N 为检测的类别数; frameNum 为检测图片的总数; elapsedTime 为耗费的总时间。

4.3. 消融实验

为了验证我们所提出改进点对 YOLOv7 模型的影响, 本实验部分在自制茶叶病害数据集上进行了四组消融实验, 实验结果见表 2 所示。

Table 2. Ablation experiments

表 2. 消融实验

基础模型	ACmix	C3	Alpha-IoU	mAP@0.5 (%)	FPS (f/s)	Params (M)
	-	-	-	91.5	65	37.2
	√	-	-	92.6	63	38.5
YOLOv7	-	√	-	91.3	82	32.8
	-	-	√	92.1	64	37.2
	√	√	√	93.3	79	33.7

从表 2 可知, 加入 ACmix 注意力模块后, mAP@0.5 提升了 1.1 个百分点, 模型参数量略有增加, 而 FPS 略有下降, 但是变化幅度都不大, 这说明加入 ACmix 注意力在不增加参数量的同时显著提高了网络的特征提取能力; 用 C3 模块替换 ELAN-W 之后, mAP@0.5 和模型参数量都略微有所下降, 其中 mAP@0.5 下降 0.2 个百分点, FPS 提升为 82 帧, 而模型参数量下降了 4.4 M, 这表明加入 C3 模块后显著减少了模型的计算量和参数量; 使用 Alpha-IoU 优化原损失函数后, 在没有 FPS 和模型参数量增加的情况下, mAP@0.5 增加了 0.6 个百分点。最后, 加入所有改进点之后, mAP@0.5 较基准模型提升了 1.8 个百分点, FPS 提升为 79 帧, 模型参数量下降了 3.5 M。消融实验结果表明, 加入三个改进点后的模型效果最佳, 兼具精度、速度和模型大小。

为了更直观地展示改进后模型的性能, 我们选取不同病害图片进行展示, 对比改进前后模型的检测效果, 如图 5 所示。

从图 5 可以看出, 对于第一行图像中目标, 原始 YOLOv7 可以较好的检测出茶芽枯病, 但发生了两处误检。而改进后的 YOLOv7 消除了误检, 对茶芽枯病的置信度更高, 识别更准确。对于第二行图像中目标, 虽然 YOLOv7 能准确检测出茶藻斑病和茶轮斑病两个目标, 但置信度不高, 而改进后模型对两个目标的置信度更高。对于第三行图像中目标, 原始 YOLOv7 能准确识别出茶轮斑病, 但是置信度过低, 而改进后的 YOLOv7 提高了对茶轮斑病的置信度。对于第四行图像中茶叶枯病, 原始 YOLOv7 的目标框发生了重叠, 而改进后的 YOLOv7 消除了这个情况并提升了目标框的置信度。

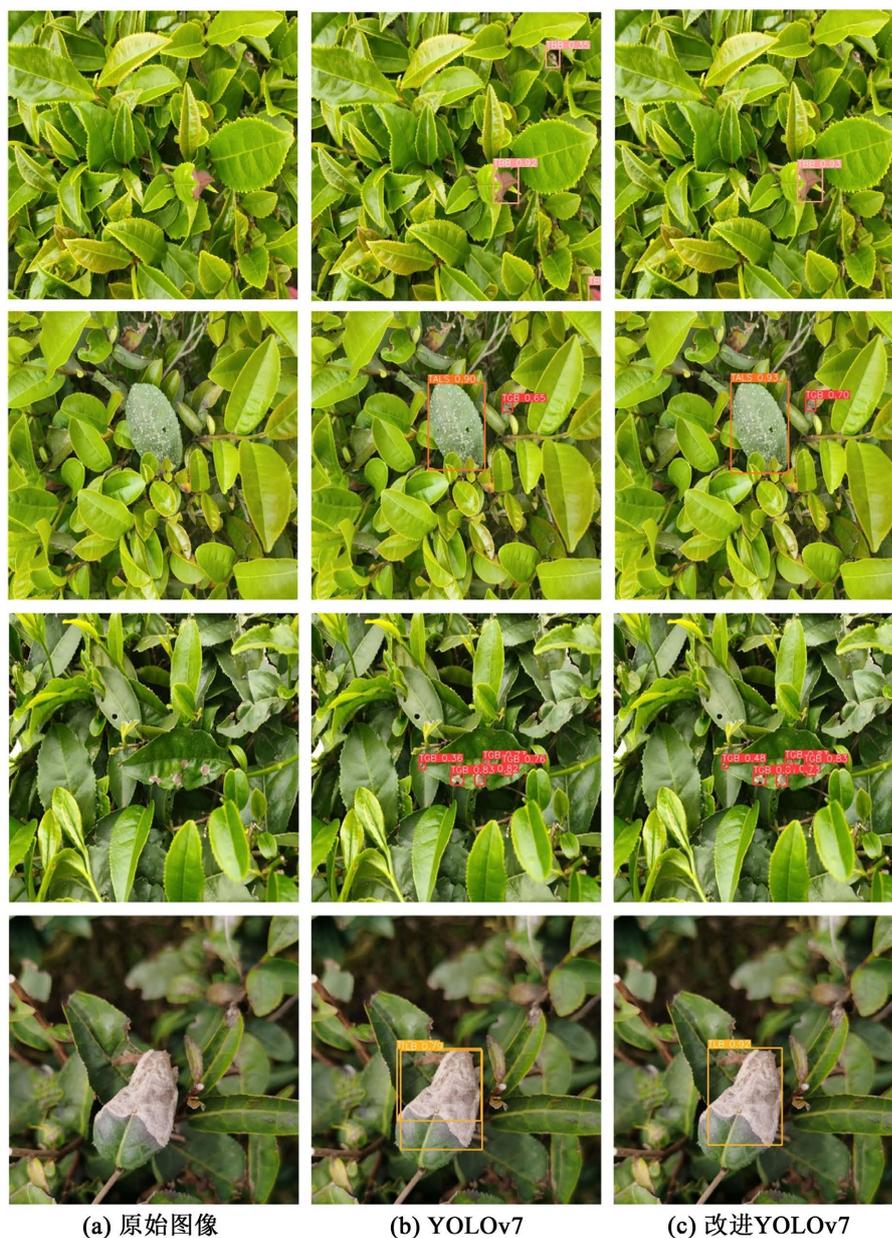


Figure 5. Comparison of recognition effect of YOLOv7 before and after improvement
图 5. 改进前后 YOLOv7 的识别效果对比

4.4. 不同模型性能对比

为了验证本文所提出的改进模型的有效性,将改进后 YOLOv7 模型与当前主流检测模型在自制茶叶病害数据集上进行对比实验,实验结果如表 3 所示。

通过表 3 结果可以看出,本文模型各方面都优于基准模型。此外,在多个指标上与先进方法相比,均具有明显优势。其中, P、R 和 mAP@0.5 均是最优的,分别为 94.1%、92.3%和 93.3%。在检测速度方面, YOLOV7-Tiny 检测速度最快,本文改进模型为 79 f/s,高于 YOLOv7 14 f/s。在参数量方面, YOLOV7-Tiny 参数量最低,仅为 6.1 M,而本文改进模型排第三名,为 33.7 M。综合准确率、检测速度和模型大小等评价指标考虑,本文改进模型综合性能更具有优势。

Table 3. Comparison of the improved model with other detection models
表 3. 改进模型与其他检测模型的对比

模型	P (%)	R (%)	mAP@0.5 (%)	FPS (f/s)	Params (M)
YOLOv4	84.3	85.1	83.9	34	243.4
YOLOv5s	89.5	86.4	87.3	85	7.2
YOLOV7-Tiny	90.2	89.5	88.4	110	6.1
YOLOv7	91.6	89.7	91.5	65	37.2
Ours	94.1	92.3	93.3	79	33.7

5. 结论

本文提出了一种改进的 YOLOv7 茶叶病害识别模型，其 mAP@50 和 FPS 分别高于原始 YOLOv7 模型 1.8 个百分点和 14 帧，参数量减少了 3.5 M，改进模型较原始模型性能有较大提升。此外，将改进后模型与当前主流目标检测网络 YOLOv4、YOLOv5s 以及 YOLOV7-Tiny 相比，其精确率、召回率和 mAP@0.5 为最佳，参数量大小和 FPS 分别为第三名。综合各项评价指标来看，本文改进模型满足茶园环境中茶叶病害识别的需要，可以将模型部署到茶园终端，对茶叶发病情况进行监控，为茶产业的智能化发展提供了研究基础，具有重要的理论价值和实际意义。

基金项目

贵州省基础研究项目(No. ZK [2023]060)。

参考文献

- [1] 师念, 汪悦. 因茶兴业因茶致富[N]. 陕西日报, 2021-06-25(005).
- [2] Bao, W., Fan, T., Hu, G., et al. (2022) Detection and Identification of Tea Leaf Diseases Based on AX-RetinaNet. *Scientific Reports*, **12**, 2183. <https://doi.org/10.1038/s41598-022-06181-z>
- [3] Lee, S.H., Lin, S.R. and Chen, S.F. (2020) Identification of Tea Foliar Diseases and Pest Damage under Practical Field Conditions Using a Convolutional Neural Network. *Plant Pathology*, **69**, 1731-1739. <https://doi.org/10.1111/ppa.13251>
- [4] 魏丽冉, 岳峻, 李振波, 等. 基于核函数支持向量机的植物叶部病害多分类检测方法[J]. 农业机械学报, 2017, 48(S1): 166-171.
- [5] 杨昕薇, 谭峰. 基于贝叶斯分类器的水稻病害识别处理的研究[J]. 黑龙江八一农垦大学学报, 2012, 24(3): 64-67, 79.
- [6] 余秀丽, 徐超, 王丹丹, 等. 基于 SVM 的小麦叶部病害识别方法研究[J]. 农机化研究, 2014, 36(11): 151-155, 159.
- [7] 张云龙, 袁浩, 张晴晴, 等. 基于颜色特征和差直方图的苹果叶部病害识别方法[J]. 江苏农业科学, 2017, 45(14): 171-174.
- [8] 林彬彬, 邱新法, 何永健, 等. 茶树病害智能诊断识别算法研究[J]. 江苏农业科学, 2019, 47(6): 85-91.
- [9] 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.
- [10] 毛锐, 张宇晨, 王泽玺, 等. 利用改进 Faster-RCNN 识别小麦条锈病和黄矮病[J]. 农业工程学报, 2022, 38(17): 176-185.
- [11] 储鑫, 李祥, 罗斌, 等. 基于改进 YOLOv4 算法的番茄叶部病害识别方法[J]. 江苏农业学报, 2023, 39(5): 1199-1208.
- [12] 贾璐, 叶中华. 基于注意力机制和特征融合的葡萄病害识别模型[J]. 农业机械学报, 2023, 54(7): 223-233.
- [13] 杨康, 熊凯, 周平, 等. 基于改进 SSD 算法的轻量化三七病害检测[J]. 华南农业大学学报, 2023, 44(3): 447-455.

- [14] 李就好, 林乐坚, 田凯, 等. 改进 Faster R-CNN 的田间苦瓜叶部病害检测[J]. 农业工程学报, 2020, 36(12): 179-185.
- [15] 孙道宗, 刘欢, 刘锦源, 等. 基于改进 YOLOv4 模型的茶叶病害识别[J]. 西北农林科技大学学报(自然科学版), 2023, 51(9): 145-154.
- [16] 杨文姬, 胡文超, 赵应丁, 等. 基于改进 Yolov5 植物病害检测算法研究[J]. 中国农机化学报, 2023, 44(1): 108-115.
- [17] 李伟豪, 詹炜, 周婉, 等. 轻量型 Yolov7-TSA 网络在茶叶病害检测识别中的研究与应用[J]. 河南农业科学, 2023, 52(5): 162-169.
- [18] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M. (2023) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, 17-24 June 2023, 7464-7475. <https://doi.org/10.1109/CVPR52729.2023.00721>
- [19] Pan, X., Ge, C., Lu, R., *et al.* (2022) On the Integration of Self-Attention and Convolution. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, 18-24 June 2022, 805-815. <https://doi.org/10.1109/CVPR52688.2022.00089>
- [20] Zhu, X., Lyu, S., Wang, X., *et al.* (2021) TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-Captured Scenarios. *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshop*, Montreal, 11-17 October 2021, 2778-2788. <https://doi.org/10.1109/ICCVW54120.2021.00312>
- [21] He, J., Erfani, S., Ma, X., *et al.* (2021) α -IoU: A Family of Power Intersection over Union Losses for Bounding Box Regression. *Advances in Neural Information Processing Systems*, **34**, 20230-20242.