

李仁杰, 宋 涛, 高 婕, 等. 基于改进 YOLOv5 的自然环境下番茄患病叶片检测模型[J]. 江苏农业学报, 2024, 40(6): 1028-1037.

doi: 10.3969/j.issn.1000-4440.2024.06.009

基于改进 YOLOv5 的自然环境下番茄患病叶片检测模型

李仁杰, 宋 涛, 高 婕, 李 东, 高 鹏, 李炳鑫, 杨 坡

(河北工业大学电子信息工程学院, 天津 300000)

摘要: 针对自然环境下番茄叶片存在的复杂背景和密集遮挡情况, 提出一种改进的 YOLOv5 模型, 用于实时检测自然环境下番茄叶片的病害。首先, 使用 RepVGG 模块代替 YOLOv5 中主干网络的卷积层, 改善主干网络的特征提取能力, 减少模型的内存占用, 加速模型的推理速度; 其次, 在颈部的 C3 模块中引入注意力机制模块 CBAM, 提高模型在复杂背景下对番茄患病叶片的检测精度以及对遮挡目标的识别率; 最后, 引用新的损失函数 SIoU, 加速模型的收敛速度并降低模型的损失值。研究表明, 相比于原 YOLOv5 模型, 改进模型的平均精度提升 3.0 个百分点, 平均精度高达 98.9%, 说明改进模型在自然环境下对番茄患病叶片的检测更具优势。

关键词: 病害检测; 深度学习; YOLOv5; 注意力机制

中图分类号: TP391; S641.2

文献标识码: A

文章编号: 1000-4440(2024)06-1028-10

Tomato diseased leaf detection model based on improved YOLOv5 in natural environment

LI Renjie, SONG Tao, GAO Jie, LI Dong, GAO Peng, LI Bingxin, YANG Po

(School of Electronic and Information Engineering, Hebei University of Technology, Tianjin 300000, China)

Abstract: Aiming at the complex background and dense occlusion of tomato leaves in natural environment, an improved YOLOv5 model was proposed for real-time detection of tomato leaf diseases in natural environment. Firstly, the RepVGG module was used to replace the convolutional layer of the backbone network in YOLOv5, which improved the feature extraction capability of the backbone network, reduced the memory occupation of the model and accelerated the reasoning speed of the model. Secondly, the attention mechanism module CBAM was introduced into C3 module in neck part to improve the detection accuracy of tomato diseased leaves and the recognition rate of shielded targets in the complex background. Finally, a new loss function SIoU was introduced to accelerate the convergence speed of the model and reduce the loss value of the model. The research results showed that compared with the original YOLOv5 model, the average precision of the improved model increased by three percentage points, and the average accuracy was as high as 98.9%, indicating that the improved model was more advantageous in the detection of tomato diseased leaves in natural environment.

Key words: disease detection; deep learning; YOLOv5; attention mechanism

收稿日期: 2023-04-24

基金项目: 河北省重点研发计划项目(22370701D)

作者简介: 李仁杰(1997-), 男, 安徽芜湖人, 硕士研究生, 研究方向为计算智能与无线网络。(E-mail) 2270464695@qq.com

通讯作者: 宋 涛, (E-mail) songtao@hebut.edu.cn

番茄营养丰富、口感独特、种植面积大, 是重要的经济作物之一^[1]。番茄种植过程中常见的病害有早疫病、晚疫病、斑点病等。不同的病害会导致番茄叶片生理形态发生不同的变化, 如形变、变色、卷曲、腐烂等, 检测番茄叶片形态变化是识别番茄病害

的关键。在实际生产过程中,生产者往往采用人工识别的方法去识别番茄叶片病害种类,这种识别方法效率低且效果不佳。因此,如果能够实现对番茄病害的实时检测,就能够及时指导生产者进行植物保护,从而达到提高番茄产量和质量的生产目标。

针对作物病害检测问题,传统机器学习识别方法由于具有计算效率高、鲁棒性强等优势而广受欢迎。但是传统机器学习识别方法需要具有丰富经验和专业知识的人去设计特征提取过程,且特征提取过程需要进行大量的试验和调整,会花费大量时间和精力^[2]。

相较于传统机器学习识别方法,深度学习方法可以使用更深、更复杂的模型来自动地学习特征,从而避免了手动设计特征的繁琐过程^[3],因此,深度学习方法在农业生产中得到了广泛的运用。Ozgun 等^[4]通过改变 Faster R-CNN 模型的参数实现对甜菜叶斑病的自动检测,获得了 95.48% 的总体分类正确率。Waheed 等^[5]提出一种用于检测玉米早期病害的卷积神经网络模型,该模型的识别准确率达到 98.06%。Liu 等^[6]将 Inception 结构、深度可分离卷积和稠密连通结构相结合,构建了一个高效的葡萄叶片害虫检测网络,该模型的准确率可达 97.22%。

虽然深度学习方法在农业病害检测方面取得了不错的成果,学者们构建的作物病害检测模型有着较高的识别率,但是大多数检测模型的研究都是基于简单的实验室环境进行的^[7],而自然场景下番茄叶片的

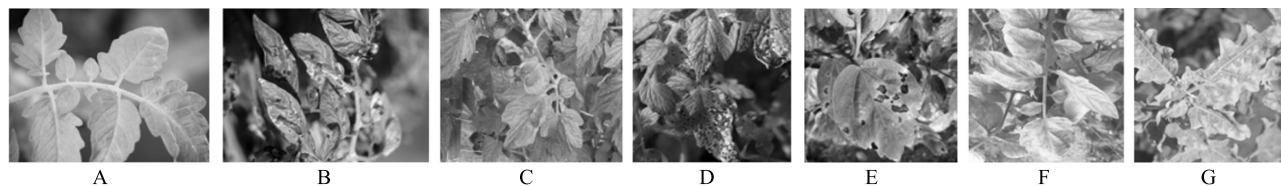
检测过程中普遍存在背景复杂、叶片被遮挡等情况,复杂的背景信息会引入大量的噪声信息,而叶片之间的遮挡会使目标区域的部分信息丢失,这大大增加了误检率和漏检率,从而降低了算法的检测精度^[8]。因此,本研究拟提出一种改进的 YOLOv5 模型,使用 RepVGG 模块代替主干网络中的 CBS 结构,以提升模型的检测精度,并加速模型的推理;颈部(Neck)中的 C3 模块引入注意力机制模块 CBAM 并命名为 C3CBAM,以使模型在训练时加强对目标区域的关注,提升对遮挡目标的识别率和对复杂背景的区分能力,有效减少漏检和误检的情况;使用损失函数 SIoU 代替 YOLOv5 模型原本的损失函数 CIoU,以加速模型的收敛并降低损失值,以期在自然环境下番茄患病叶片的实时、精准检测奠定基础。

1 材料与方法

1.1 数据集

本研究使用的数据集来自于公开的 PlantDoc^[9] 数据集,数据集包括 7 类番茄叶片,即健康的番茄叶片、患早疫病的番茄叶片、患晚疫病的番茄叶片、患叶斑病的番茄叶片、患叶菌斑病的番茄叶片、患叶黄病的番茄叶片、患花叶病的番茄叶片,一共 1 178 张。7 类番茄叶片的图像如图 1 所示。

本研究使用的番茄患病叶片、健康叶片数据集均采用 LabelImg 标签标注工具进行手动注释,绘制出叶片边界框和叶片患病的种类,以数据集中部分番茄叶片图片为例,LabelImg 工具标注的结果见图 2。



A~G 依次为健康的番茄叶片、患早疫病的番茄叶片、患晚疫病的番茄叶片、患叶斑病的番茄叶片、患叶菌斑病的番茄叶片、患叶黄病的番茄叶片、患花叶病的番茄叶片。

图 1 番茄叶片数据集部分样本

Fig.1 Partial samples of tomato leaf data set

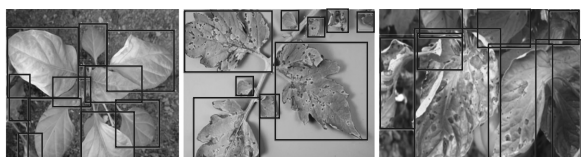
1.2 数据预处理

为了防止由于训练数据太少而导致的过度拟合和不收敛问题,对数据集样本进行数据增强,数据增强的方法包括高斯模糊、水平翻转、随机旋转、色彩抖动 4 种方法。扩充之后的数据集共有 3 130 张图片。扩充后的数据集按 8 : 1 : 1 的比例划分为训练

集、测试集和验证集。本研究原始数据集中番茄叶片的种类和数量见表 1。

1.3 试验环境

试验设备为台式计算机,选择 Windows10 作为操作系统,搭载 PyTorch 深度学习框架,编程语言选择 Python 3.8,CUDA 版本为 11.3。硬件部分:处理



A:健康的番茄叶片;B:患早疫病的番茄叶片;C:患叶斑病的番茄叶片。

图2 使用 LabelImg 标注数据集的结果

Fig.2 Annotated data set results using LabelImg

器为 Intel® Core™ i7-11700KF,主频为 3.60 GHz,内存为 32 G,显卡为 NVIDIA GeForce RTX3090。模型网络的训练轮次设为 200 个周期,批量大小设置为 64,优化器默认采用随机梯度下降 (Stochastic gradient descent,SGD),动量参数取值为 0.937,初始学习率默认设为 0.01。

表1 番茄病害叶片数据集

Table 1 Tomato diseased leaf data set

类别	数量(张)
健康的番茄叶片	110
患早疫病的番茄叶片	158
患晚疫病的番茄叶片	202
患叶斑病的番茄叶片	278
患叶菌斑病的番茄叶片	202
患叶黄病的番茄叶片	140
患花叶病的番茄叶片	88

1.4 YOLOv5 模型结构

YOLOv5 模型具有推理速度快、精度高、体积小等优点,本研究选取 YOLOv5 模型作为基线网络,在此基础上对模型进行改良。YOLOv5 模型结构见图 3,模型主要分为输入层 (Input)、主干网络 (Backbone)、颈部 (Neck) 和头部 (Head) 4 个部分。

在输入模块对输入的图像进行预处理。利用马赛克 (Mosaic) 数据增强方法增加数据集的多样性;自适应锚框运用 K 均值聚类算法计算出最佳初始锚框值;用自适应图片缩放功能调整图像的大小,使得图像能够被高效处理,同时保持较好的目标检测性能。

主干网络由焦点 (Focus) 切片、CBS、CSP 和空间金字塔池化 (SPP) 4 个模块构成,输入图片采用 Focus 进行切片操作,然后经过 CBS 和 CSP 模块完成特征提取,最后利用空间金字塔池化模块对输入

特征图进行池化操作,实现图片自适应尺寸的输出。

颈部采用了特征金字塔网络 (FPN) 结构加路径聚合网络 (PANet) 结构,FPN 的作用是将上层的强语义信息融入底层,PANet 的作用则是自底向上传递特征图的定位信息^[10]。

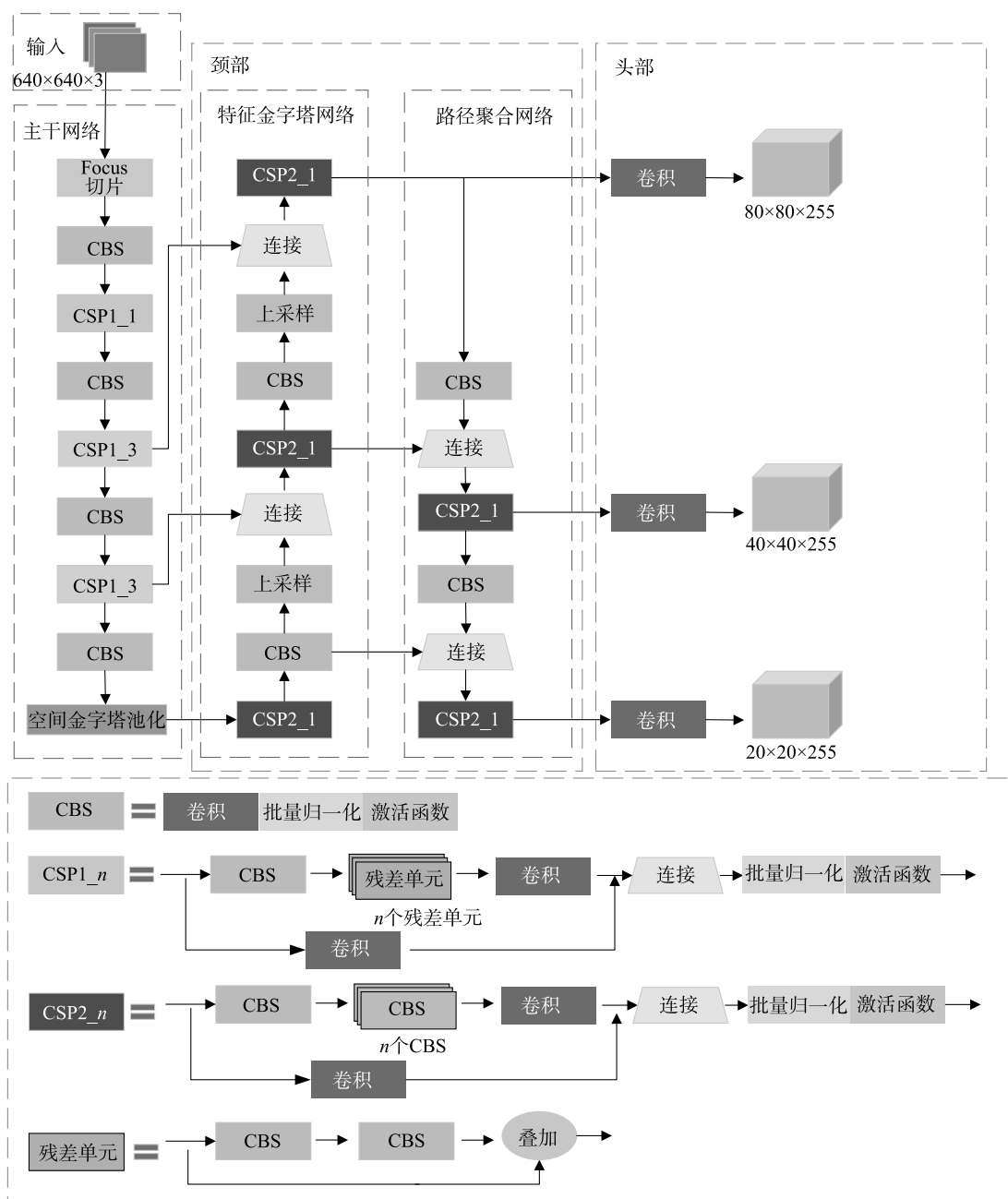
头部输出 20×20 、 40×40 、 80×80 共 3 个不同大小的特征图,每个特征图对应不同大小的预测框,每个预测框都包含置信度、分类概率和边界框位置信息,采用非极大值抑制算法 (NMS) 保留符合阈值的预测框。

1.5 YOLOv5 模型改进方法

1.5.1 引入 RepVGG 模块 RepVGG 模块^[11] 利用结构重参数化方法设计了一种可融合的网络结构,在训练阶段,RepVGG 模块采用由 3×3 卷积、 1×1 卷积和恒等映射模块组成的多分支网络结构,如图 4A 所示,多分支结构可以使模型在训练时融合不同尺度的特征图信息,加强网络的特征提取能力,但是多分支结构会增加硬件内存消耗,这是因为每路分支计算都会占用硬件的内存空间,直到所有分支计算结束融合后才释放内存^[12]。因此在推理阶段,RepVGG 模块通过结构重参数化的方式将多分支的卷积算子融合到主分支结构中,合并成卷积核为 3×3 的单分支网络,如图 4B 所示,单分支结构可以有效减少内存开销,提升模型推理速度。

本研究引入 RepVGG 模块来对 YOLOv5 网络结构进行优化设计,将主干网络中原有的 CBS 模块替换为 RepVGG 模块,为了兼顾模型的复杂度和性能,主干网络中的 C3 模块的数量缩减为原来的 $1/3$ 。融合 RepVGG 模块的主干网络 (图 5),RepVGG 模块能使网络在进行番茄患病叶片检测时兼具多分支模型与单路模型的优点。在训练阶段,多分支结构可以通过融合不同大小的卷积核来获得不同大小的感受野,强化了网络的特征表达能力,从而改善了网络模型对患病叶片的检测效果;在推理阶段,RepVGG 模块通过结构重参数化方式将多分支结构解耦为单路结构,提升了推理速度,以满足模型实时检测的需求。

1.5.2 CBAM 注意力机制模块 CBAM^[13] 模块融合了通道注意力模块 (CAM) 和空间注意力模块 (SAM)。在通道注意力模块中,输入尺寸为长度 (C) \times 高度 (H) \times 宽度 (W) 的特征图会经过上下路并行的平均池化层和全局池化层,输出大小为 $C \times 1 \times 1$ 的特征图,然后通过共享网络多层感知机 (MLP) 将



Focus: 焦点; CBS: 卷积、批量归一化、激活函数组成的模块; CSP1_n: CBS、卷积、残差单元、连接、批量归一化、激活函数组成的模块; CSP2_n: CBS、卷积、连接、批量归一化、激活函数组成的模块。

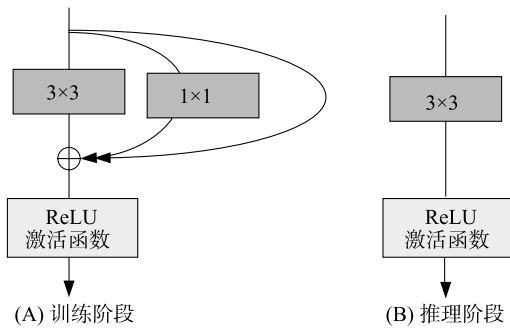
图3 YOLOv5 模型的结构

Fig.3 Structure of YOLOv5 model

特征图大小缩小至 $C/r \times 1 \times 1$, 其中 r 为缩减比, 将这 2 个输出结果进行逐元素相加, 再经过激活函数得到大小为 $C \times 1 \times 1$ 的特征图, 将最后得到的特征图与原特征图相乘, 得到大小为 $C \times H \times W$ 的通道注意力特征图传入空间注意力模块。

在空间注意力模块中, 输入特征图依次经过全

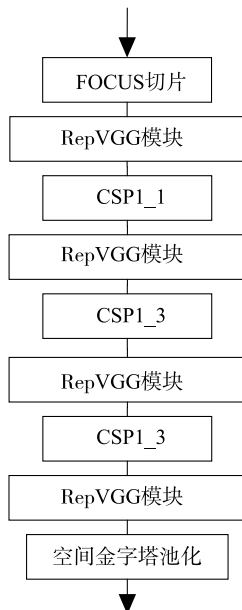
局池化层、平均池化层, 得到大小为 $1 \times H \times W$ 的 2 个特征图, 将得到的特征图进行通道上的相加, 再采用 7×7 的卷积核进行卷积操作, 最后通过激活函数得到 $1 \times H \times W$ 大小的特征图, 此时得到的 $1 \times H \times W$ 大小的特征图反映了特征图上每个区域的权重, 数值越大表明重要性越高。将该反映权重的特征图与原特



ReLU:修正线性单元。

图4 RepVGG 模块结构

Fig.4 Structure of RepVGG module



Focus:焦点;CSP1_n:CBS、卷积、残差单元、连接、批量归一化、激活函数组成的模块,结构见图3。

图5 融合 RepVGG 模块的主干网络结构

Fig.5 Backbone network structure of the fusion RepVGG module

征图进行相乘,使得输出的特征图维度变为原来的 $C \times H \times W$,就能通过不同区域的加权值大小来区分特征图中区域的重要程度高低。CBAM 模块的结构如图 6 所示。

自然环境下采集到的番茄叶片图像背景复杂且目标对象多,患病叶片之间存在较严重的遮挡情况,检测时可能会造成对遮挡叶片的漏检和误检。因此,在 C3 模块上引入了 CBAM 注意力机制模块,目的是希望检测模型能够对图片的重要位置和重要内容进行关注,区分复杂背景和需要关注的目标对象区域,提高对遮挡目标的识别率进而改善模型的检

测性能。本研究在 YOLOv5 颈部的 C3 模块中加入了 CBAM 注意力机制,命名为 C3CBAM 模块,颈部原始的 C3 模块结构如图 7A 所示,改进后的 C3CBAM 模块结构如图 7B 所示。

1.5.3 损失函数设计 YOLOv5 模型的损失函数为边界回归损失值(L_{box})、目标置信度损失值(L_{obj})和分类损失值(L_{cls})的加权和,损失值($LOSS$)的计算方法如公式(1)所示:

$$Loss = \lambda_1 L_{\text{box}} + \lambda_2 L_{\text{obj}} + \lambda_3 L_{\text{cls}} \quad (1)$$

式中, $\lambda_1 \sim \lambda_3$ 为权重参数。

YOLOv5 的边界回归损失值(L_{box})采用 CIoU 损失函数表示^[14],计算方法如公式(2)所示:

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

式中, IoU ^[15]表示预测框与真实框交集和并集的比值; $\rho^2(b, b^{gt})$ 表示预测框与真实框中心点的欧氏距离; c 表示预测框与真实框最小外接矩形的对角线长度; α 为权重参数,计算方法如公式(3)所示; v 表示纵横比的一致程度, v 的计算方法如公式(4)所示。

$$\alpha = \frac{v}{(1 - IoU) + v} \quad (3)$$

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

式中, w 、 h 分别表示预测框的宽度、高度; w^{gt} 、 h^{gt} 分别表示真实框的宽度、高度。

CIoU 损失函数引入了预测框和真实框的宽高比,但是忽略了预测框和真实框之间的方向性,这样网络在训练过程中会发生预测框随意匹配的问题。因此本研究采用 SIoU^[16]函数代替 CIoU 函数,SIoU 函数通过引入预测框和真实框之间的向量角度信息,重新定义了惩罚标准,具体包括以下几个部分:

第一,角度损失(Λ),计算方法如公式(5)所示:

$$\Lambda = 1 - 2 \times \sin^2 \left[\arcsin \left(\frac{c_h}{\sigma} \right) - \frac{\pi}{4} \right] \quad (5)$$

式中, c_h 表示真实框和预测框中心点的高度差;

σ 表示真实框和预测框中心点的距离; $\arcsin(\frac{c_h}{\sigma})$ 即表示真实框和预测框中心点坐标形成的线段与水平线的夹角。角度损失计算方法如图 8 所示。

第二,距离损失(Δ),计算方法如公式(6)所示:

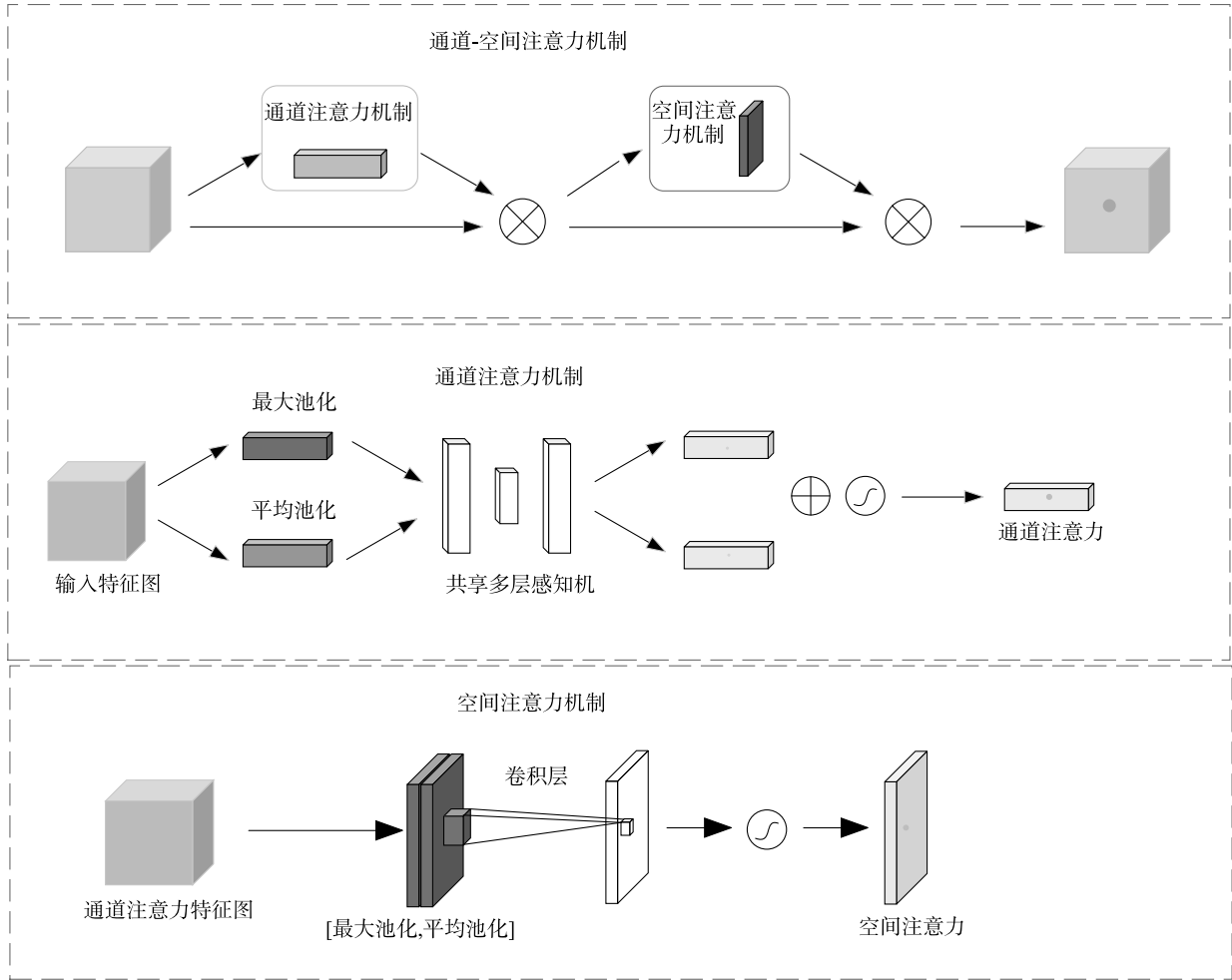


图6 CBAM 注意力机制模块结构

Fig.6 Structure diagram of CBAM attention mechanism module

$$\Delta = 2 - e^{-\gamma \rho_x} - e^{-\gamma \rho_y} \quad (6)$$

其中:

$$\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_w} \right)^2, \rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_h} \right)^2, \gamma = 2 - \Delta \quad (7)$$

式中, Δ 表示角度损失值; $(b_{c_x}^{gt}, b_{c_y}^{gt})$ 、 (b_{c_x}, b_{c_y}) 分别表示真实框、预测框的中心点位置坐标; c_w 、 c_h 分别表示包含真实框、预测框的最小外接矩形的宽度和高度。距离损失计算方法如图9所示。

第三, 形状损失 (Ω) 的计算方法如公式(8)所示:

$$\Omega = (1 - e^{-w_w})^\theta + (1 - e^{-w_h})^\theta \quad (8)$$

其中:

$$w_w = \frac{|w - w^{gt}|}{(w, w^{gt})_{\max}}, w_h = \frac{|h - h^{gt}|}{(h, h^{gt})_{\max}} \quad (9)$$

式中, w 、 h 分别表示预测框的宽度、高度, w^{gt} 、

h^{gt} 分别表示真实框的宽度、高度; $(w, w^{gt})_{\max}$ 表示取 w 、 w^{gt} 2 个变量的最大值; $(h, h^{gt})_{\max}$ 表示取 h 、 h^{gt} 2 个变量的最大值; 参数 θ 的大小与模型对形状损失的关注度高低呈正相关, θ 值过大会影响预测框的移动, 本研究通过遗传算法计算得到最佳的 θ 值接近 4, 因此设置 θ 的范围为 $[2, 6]$ 。

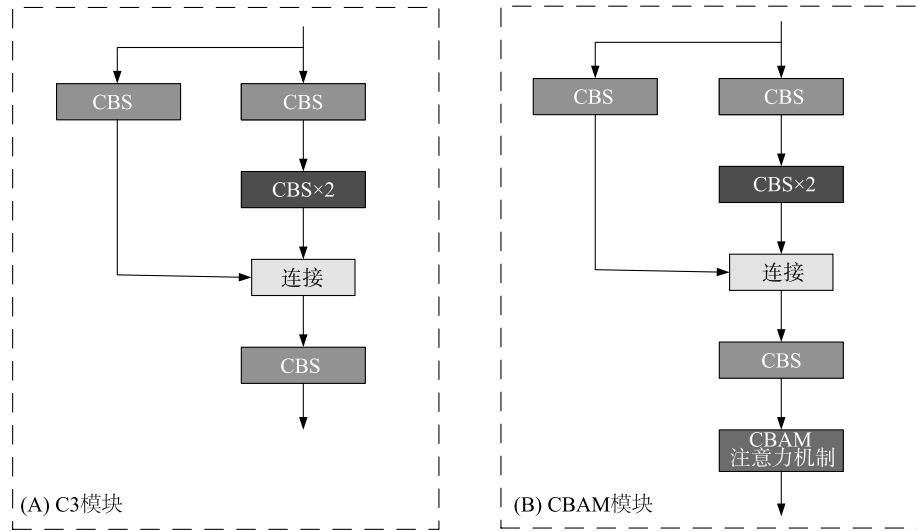
最后得到 SIoU 函数, 与现有的损失函数相比, SIoU 函数通过关注预测框和真实框之间的方向性, 加速了网络的收敛并降低了网络损失函数值 (L_{SIoU})。

$$L_{\text{SIoU}} = 1 - \text{IoU} + \frac{\Delta + \Omega}{2} \quad (10)$$

式中, IoU 表示预测框与真实框交集和并集的比值大小; Δ 表示距离损失; Ω 表示形状损失。

1.6 评价指标

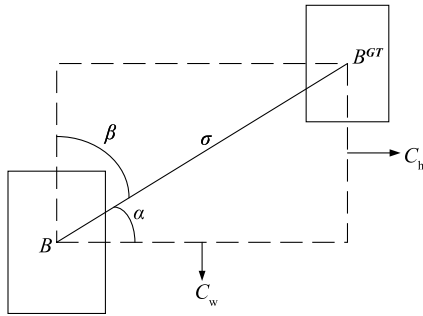
采用精度 (*Precision*)、召回率 (*Recall*)、平均精



CBS:卷积、批量归一化、激活函数组成的模块,结构见图3;CBAM:空间-通道注意力机制。

图7 C3模块和C3CBAM模块的结构

Fig.7 Network structural diagram of C3 and C3CBAM modules



B :预测框中心坐标; B^{GT} :真实框中心坐标; σ :预测框和真实框中心点的距离; α :预测框和真实框中心坐标形成的线段与水平线的夹角; β :预测框和真实框中心坐标形成的线段与垂直线的夹角; C_w :预测框和真实框中心点水平距离; C_h :预测框和真实框中心点垂直距离。

图8 角度损失计算示意

Fig.8 Schematic diagram of angle loss calculation

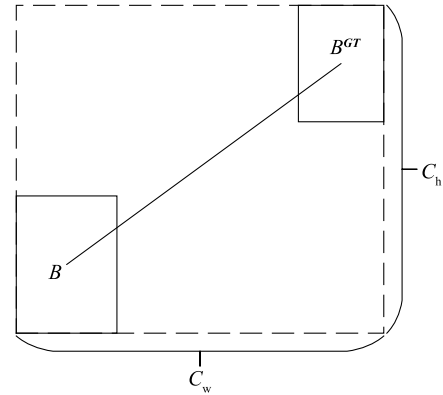
度(mAP)3个指标衡量改进的YOLOv5模型的预测效果,计算公式如下:

$$Precision = \frac{TP}{TP+FP} \quad (11)$$

$$Recall = \frac{TP}{TP+FN} \quad (12)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (13)$$

式中, TP 表示实际为正样本且被预测为正样本



B :预测框中心坐标; B^{GT} :真实框中心坐标; C_w :预测框和真实框最小外接矩形宽度; C_h :预测框和真实框最小外接矩形高度。

图9 距离损失计算示意

Fig.9 Schematic diagram of distance loss calculation

的数量; FP 表示实际为负样本但被预测为正样本的数量; FN 表示实际为正样本但被预测为负样本的样本数量; AP_i 表示第*i*个样本的平均准确度; N 表示总的类别数。

2 结果与分析

2.1 对比试验

为了验证CBAM注意力机制的有效性,选取CA^[17]、SE^[18]、ECA^[19]等注意力机制进行对比试验,不同注意力机制模块插入的位置都相同,并且保证网络其他结构不变,损失函数采用的是YOLOv5原

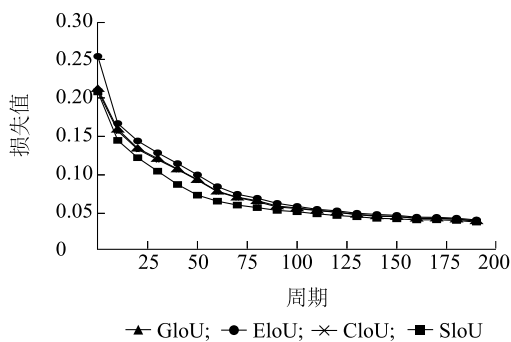
本的 CIoU 损失函数。表 2 显示,C3CBAM 注意力机制模块对模型整体性能的提升效果最佳,精度提升了 0.2 个百分点,召回率提升了 1.9 个百分点,平均精度相较于原模型提高了 1.4 个百分点。

表 2 不同注意力机制对比试验结果

Table 2 Comparative experimental results of different attention mechanisms

模型	精度 (%)	召回率 (%)	平均精度 (%)
YOLOv5	93.5	93.1	95.9
YOLOv5+C3SE	93.5	94.5	97.0
YOLOv5+C3CBAM	93.7	95.0	97.3
YOLOv5+C3CA	93.8	94.0	96.6
YOLOv5+C3ECA	93.7	93.3	96.6

本研究使用了当前目标检测中常用的几种损失函数进行对比试验,在 YOLOv5 网络的基础之上,分别采用 GIoU^[20]、CIoU、EIoU^[21]和 SIoU 作为边界框回归损失函数,以此来验证 SIoU 损失函数对改进模型的有效性,将每轮网络训练中得到的损失函数值绘制成损失函数曲线图。图 10 显示,模型在 100 个周期之后逐渐收敛,其中,EIoU 的收敛效果最差,GIoU 和 CIoU 的收敛效果相当,SIoU 不但收敛速度最快,而且模型训练的损失值也小于其他损失函数。



GIoU、EIoU、CIoU、SIoU 表示不同的损失函数。

图 10 损失函数曲线对比

Fig.10 Comparison of loss function curves

为了进一步验证 SIoU 对模型的有效性,对加入不同损失函数的模型训练完成之后,分析其损失值和 mAP 。表 3 显示,相较于 GIoU、EIoU 和 CIoU, YOLOv5 模型采用 SIoU 之后的整体效果最好,其损失值为 0.039 09, mAP 为 96.8%,相较于以 CIoU 为

损失函数的原模型,损失值下降了 0.001 58, mAP 上升了 0.9 个百分点。由此可见,SIoU 相较于其他损失函数,对模型性能的改善效果最优,得到的损失值最低。

表 3 不同损失函数对比试验结果

Table 3 Comparative experimental results of different loss functions

模型	损失值	平均精度 (%)
YOLOv5+GIoU	0.040 77	95.8
YOLOv5+CIoU	0.040 67	95.9
YOLOv5+EIoU	0.041 88	95.6
YOLOv5+SIoU	0.039 09	96.8

2.2 消融试验

本研究对 YOLOv5 模型一共提出了 3 项改进措施,为了进一步论证改进模块的合理性与有效性,采用消融试验进行分析论证。表 4 显示,在主干网络中引入了 RepVGG 模块代替卷积层进行特征提取,多分支结构加强了模型对番茄患病叶片的特征提取能力,提升了模型的性能,其中精度提升 1.4 个百分点,召回率提升 0.4 个百分点,平均精度提升 0.7 个百分点。对模型 Neck 部分增加 CBAM 注意力机制模块后,模型精度提升 0.2 个百分点,召回率提升 1.9 个百分点,平均精度提升 1.4 个百分点,这是因为模型在引入 CBAM 注意力机制模块后,CBAM 注意力机制强化了对番茄患病叶片图像中背景和目标的区分能力,使模型更好地注意到患病叶片的细节,从而提升了复杂环境和目标被遮挡情况下的识别率,降低了误检率和漏检率。模型使用 SIoU 损失函数后,精度提升 1.2 个百分点,召回率提升 0.3 个百分点,平均精度提升 0.9 个百分点,相较于原模型,采用 SIoU 损失函数后模型关注了真实框与预测框之间的方向性,从而提升了模型的收敛速度,降低了模型的损失值。完成 3 项改进后,模型的精度上升 2.9 个百分点,召回率上升 4.2 个百分点,平均精度上升 3.0 个百分点。综上所述,本研究提出的 3 项改进措施都能有效改善模型的训练效果,进一步验证了改进模型的有效性。

2.3 不同模型的对比试验

为了验证改进模型的有效性,将本研究构建模型与 SSD 模型^[22]、Faster R-CNN 模型^[23]、YOLOv3

模型^[24]、YOLOv5 模型进行对比试验。表 5 显示,虽然 Faster R-CNN 模型的平均精度达到了 96.5%,比 YOLOv5 模型高 0.6 个百分点,但是模型的参数量为 137.1 M,浮点计算量为 370.2,要远大于 YOLOv5 模型。SSD 模型和 YOLOv3 模型的平均精度比 YOLOv5 模型低,且参数量和浮点计算量比 YOLOv5 模型高。与 SSD、Faster R-CNN 模型相比,改进模型整体更具优势,与原 YOLOv5 模型相比,虽然改进模型检测单张图片的时间增加了 0.002 s,但是平均精度提升了 3.0 个百分点,且模型的参数量为 5.5 M,浮点计算量为 15.8,较改进前都有所降低。

表 4 消融试验结果

Table 4 Results of ablation tests

改进策略	精度 (%)	召回率 (%)	平均精度 (%)
YOLOv5	93.5	93.1	95.9
YOLOv5+RepVGG	94.9	93.5	96.6
YOLOv5+CBAM	93.7	95.0	97.3
YOLOv5+SIOU	94.7	93.4	96.8
YOLOv5+RepVGG+CBAM+SIOU	96.4	97.3	98.9

表 5 不同模型对比试验结果

Table 5 Comparative experimental results of different models

模型	平均精度 (%)	参数量 (M)	浮点计算量	单张图片检测用时 (s)
SSD	90.5	26.3	62.7	0.054
Faster R-CNN	96.5	137.1	370.2	1.039
YOLOv3	94.8	61.5	155.4	0.017
YOLOv5	95.9	7.1	16.5	0.009
本研究构建模型	98.9	5.5	15.8	0.011

为了更加直观地反映本研究改进模型相对于原 YOLOv5 模型在检测性能方面的提升效果,以及对复杂环境下密集遮挡的番茄患病叶片识别率的提高,使用测试集的样本图片进行测试分析,测试结果(图 11)表明,改进模型可以检测出 YOLOv5 模型漏检的部分遮挡叶片,并且改进模型的平均精度要高于 YOLOv5 模型。

3 讨论

在实时检测番茄患病叶片的过程中,为了克服

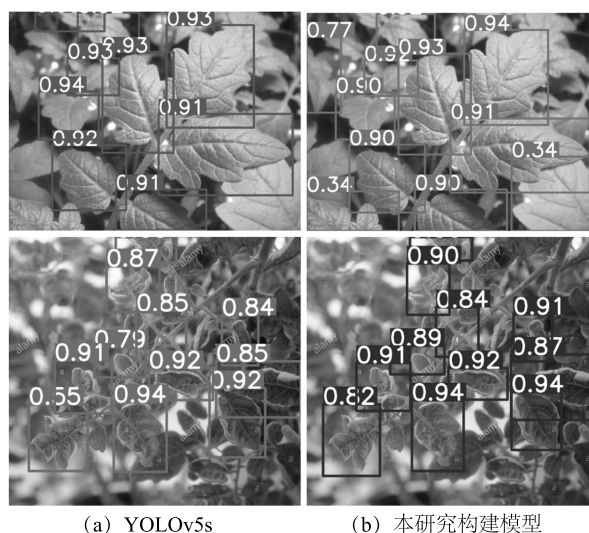


图 11 测试结果对比

Fig.11 Comparison of test results

自然环境下复杂背景和叶片遮挡对模型精度的影响,本研究提出了一种改进的 YOLOv5 检测模型,改进模型对 YOLOv5 的网络做出了合理的调整。首先,使用 RepVGG 模块代替 YOLOv5 中主干网络的卷积层,RepVGG 模块加强了主干网络对番茄患病叶片的特征提取能力,改善了模型在复杂背景下对番茄患病叶片的检测效果;其次,在颈部引入了 CBAM 注意力机制,CBAM 注意力机制的加入使得网络获得了图像的全局上下文信息,更加容易区分复杂背景和需要关注的目标对象区域,从而进一步提升了网络模型在复杂背景下对被遮挡的番茄患病叶片的检测效果;最后,本研究采用更加合理的损失函数 SIOU,SIOU 通过关注真实框和预测框的之间的方向性,使得算法在训练过程中收敛得更快,且损失函数值更低,从而改善了模型的推理精度。研究表明,改进的 YOLOv5 模型的平均精度达到 98.9%,模型单张图片检测用时为 0.011 s,满足实时检测的要求,解决了自然环境下番茄病害叶片的检测问题。

参考文献:

- [1] 周蓉,赵统敏,赵丽萍,等. 高光泽番茄育种及果实表面光泽度[J]. 江苏农业学报,2018,34(6):1437-1440.
- [2] LI Y, NIE J, CHAO X. Do we really need deep CNN for plant diseases identification[J]. Computers and Electronics in Agriculture, 2020,178:105803.
- [3] 周惠汝,吴波明. 深度学习在作物病害图像识别方面应用的研究

- 究进展[J]. 中国农业科技导报,2021,23(5):61-68.
- [4] OZGUVEN M M, ADEM K. Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms [J]. *Physica A:Statistical Mechanics and Its Applications*,2019, 535:122537.
- [5] WAHEED A, GOYAL M, GUPTA D, et al. An optimized dense convolutional neural network model for disease recognition and classification in corn leaf[J]. *Computers and Electronics in Agriculture*,2020,175:105456.
- [6] LIU B, DING Z, TIAN L, et al. Grape leaf disease identification using improved deep convolutional neural networks[J]. *Frontiers in Plant Science*,2020,11:1082.
- [7] 邵明月,张建华,冯全,等. 深度学习在植物叶部病害检测与识别的研究进展[J]. *智慧农业(中英文)*,2022,4(1):29-46.
- [8] 张志远,罗铭毅,郭树欣,等. 基于改进 YOLOv5 的自然环境下樱桃果实识别方法[J]. *农业机械学报*,2022,53(增刊1):232-240.
- [9] SINGH D, JAIN N, JAIN P, et al. PlantDoc: a dataset for visual plant disease detection[M]. New York:ACM,2020:249-253.
- [10] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]. Salt Lake City:IEEE,2018.
- [11] DING X, ZHANG X, MA N, et al. RepVGG:making vgg-style convnets great again[C]. Nashville:IEEE,2021.
- [12] 孙龙清,王泊宁,王嘉煜,等. 基于 G-RepVGG 和鱼类运动行为的水质监测方法[J]. *农业机械学报*,2022,53(增刊2):210-218.
- [13] WOO S, PARK J, LEE J Y, et al. CBAM:convolutional block attention module[C]. Berlin:Springer,2018.
- [14] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression [C]. Menlo Park: AAAI,2020.
- [15] YU J, JIANG Y, WANG Z, et al. Unitbox:an advanced object detection network[C]. New York:ACM,2016.
- [16] GEVORGYAN Z. SIoU loss:more powerful learning for bounding box regression[J]. *ArXiv Preprint ArXiv*,2022. DOI:10.48550/arXiv.2205.12740.
- [17] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]. Nashville:IEEE,2021.
- [18] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. Piscataway:IEEE,2018.
- [19] WANG Q, WU B, ZHU P, et al. ECA-Net:efficient channel attention for deep convolutional neural networks[C]. Seattle:IEEE, 2020.
- [20] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression [C]. Long Beach:IEEE,2019.
- [21] ZHANG Y, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. *Neurocomputing*, 2022,506:146-157.
- [22] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]. Berlin:Springer,2016.
- [23] REN S, HE K, 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.
- [24] REDMON J, FARHADI A. YOLOv3: an incremental improvement [J]. *ArXiv Preprint ArXiv*, 2018. DOI: 10.48550/arXiv.1804.02767.

(责任编辑:王 妮)