

基于改进 YOLO v8n 网络的番茄成熟度实时检测算法

任晶秋¹ 万恩晗¹ 单蜜¹ 张光华¹ 卢为党²

(1. 东北石油大学电气信息工程学院, 大庆 163318; 2. 浙江工业大学信息工程学院, 杭州 310014)

摘要: 为应对番茄采摘面临的果农老龄化、劳动力短缺和人工成本上涨等挑战,解决在复杂果园环境下番茄采摘机器人视觉系统成熟度检测精度低和实例分割不准确等问题,本文提出一种基于改进 YOLO v8n 网络的番茄成熟度实时检测算法。首先,通过在 YOLO v8n 网络中引入通道嵌入位置注意力模块和改进大核卷积块注意力模块,能够在浅层网络保留番茄目标位置信息,建立目标区域之间的长距离依赖关系,从而增加 YOLO v8n 网络对显著番茄特征的关注。然后,在 LaboroTomato 数据集上进行了对比实验,改进 YOLO v8n 相较于原 YOLO v8n 网络,检测和分割的 mAP@50 和 mAP@50-95 分别提高 0.4、1.4 个百分点和 0.3、1.2 个百分点。最后,实现了改进 YOLO v8n 网络在低成本、低算力和低功耗的 Jetson Nano 平台上的轻量化部署,模型内存占用量由满溢减少到 2.4 GB,推理速度加倍。该研究可为番茄采摘机器人在复杂场景下实时、准确检测番茄成熟度提供技术支撑。

关键词: 番茄成熟度; 目标检测; YOLO v8n; 注意力模块; 轻量化; Jetson Nano

中图分类号: TP391.4; S24

文献标识码: A

文章编号: 1000-1298(2025)03-0374-09

OSID:



Improved YOLO v8n Network for Real-time Detection of Tomato Maturity

REN Jingqiu¹ WAN Enhao¹ SHAN Mi¹ ZHANG Guanghua¹ LU Weidang²

(1. School of Electrical and Information Engineering, Northeast Petroleum University, Daqing 163318, China

2. College of Information Engineering, Zhejiang University of Technology, Hangzhou 310014, China)

Abstract: To address the numerous challenges faced in tomato harvesting, such as the aging of farmers, labor shortages, and rising labor costs, and resolve issues related to the low maturity detection accuracy and inaccurate instance segmentation of tomato harvesting robots in complex orchard environments, an improved YOLO v8 network-based real-time tomato maturity detection algorithm was proposed. Firstly, by introducing the channel embedded positional attention module and an improved large kernel convolutional block attention module into the YOLO v8n network, the algorithm can effectively retain the positional information of tomato targets in the shallow network layers and establish long-range dependencies between target regions, thereby significantly increasing the attention of the YOLO v8n network to critical tomato features. Then a series of comprehensive and rigorous comparative experiments were conducted on the LaboroTomato dataset, demonstrating that the improved YOLO v8n network achieved 0.4 percentage points, 1.4 percentage points, and 0.3 percentage points, 1.2 percentage points improvements in detection and segmentation mAP@50 and mAP@50-95, respectively, compared with that of the original YOLO v8n network. Finally, the improved YOLO v8n network was lightweight deployed on the low-cost, low-computation, and low-power Jetson Nano platform, successfully reducing memory usage from overflow to 2.4 GB and doubling the inference speed. The research result can provide robust technical support for the real-time and accurate detection of tomato maturity by tomato harvesting robots in complex scenarios, significantly enhancing the overall efficiency and effectiveness of automated tomato harvesting operations.

Key words: Tomato maturity; target detection; YOLO v8n; attention module; lightweight; Jetson Nano

收稿日期: 2024-02-08 修回日期: 2024-05-30

基金项目: 国家自然科学基金项目(62271447)

作者简介: 任晶秋(1979—),女,副教授,主要从事信息与信号处理研究,E-mail: renjingqiunepu@126.com

通信作者: 张光华(1979—),男,副教授,博士,主要从事基于新一代通信体制的信号处理研究,E-mail: dqzgh@nepu.edu.cn

0 引言

我国番茄种植面积和产量长期位居全球首位,番茄产业已成为农民增收和出口创汇的重要途径^[1-2]。然而,番茄采摘作为最费工费时费力的生产环节,面临果农老龄化、劳动力短缺和人工成本上涨等诸多挑战^[3]。采摘机器人的出现可有效应对这些挑战,并且提高农产品生产效率和品质^[4]。机器视觉系统是采摘机器人核心组成部分,负责图像采集、果实检测与识别以及三维重建与定位等任务,以引导末端执行机构接近果实并完成抓取和采摘动作^[5]。机器视觉系统能够在复杂环境干扰下快速而精确地识别番茄成熟度,这对于番茄采摘机器人能否成功完成预期采摘任务至关重要。

基于深度学习的番茄成熟度检测方法可提取图像高维特征,受环境干扰程度低,近年来在成熟度检测领域逐渐得到应用和推广^[6-16]。当前研究工作中集中在改进现有目标检测模型,用于番茄成熟度检测。这类模型通常表现出优异性能,并且可以进行个性化定制,相比重新设计专用于番茄成熟度检测模型,这种方法成本更低、效率更高。当前,对这类模型的改进主要集中在3个方面:①数据方面,构建样本丰富的数据集,采用数据增强方法提高模型鲁棒性。②网络结构方面,增强主干网络对番茄特征的提取能力,增加颈部网络多尺度特征融合,在头部结构添加小目标检测头。③轻量化方面,采用轻型主干网络,通过剪枝和量化去除冗余结构,引入轻量化卷积降低参数量。

目前,虽然在番茄成熟度检测方面已积累了许多优秀案例,但在面对枝叶遮挡、果实重叠和光线干扰等复杂环境,以及对相邻成熟度特征相似的番茄进行检测时,番茄成熟度检测精度仍有较大提升空间。一些研究虽取得了较高精度,但是以牺牲检测速度和增加参数量为代价。另一些研究虽然平衡了

精度、速度和轻量化的矛盾,但尚未进行相应的实时检测应用。针对上述问题,本文提出一种用于番茄成熟度检测的改进 YOLO v8n 网络。首先,将通道嵌入位置注意力模块(Channel embedded positional attention module, CePAM)用于在浅层网络中保留有用细节特征的位置信息。然后,改进卷积块注意力模块(Convolutional block attention module, CBAM)用于建立番茄区域之间的长距离依赖关系。其次,在 YOLO v8n 的主干网络引入这两种注意力模块,使其更加关注复杂环境中相邻成熟度特征差异小的番茄,从而提高检测精度。同时,这两种注意力模块都避免增加 YOLO v8n 网络参数量和计算量。以期实现改进 YOLO v8n 网络在低成本、低算力和低功耗的嵌入式 AI 平台 Jetson Nano 上的轻量化部署,使其能够通过摄像头实时检测番茄成熟度,为番茄采摘机器人自动化采摘作业提供技术支撑。

1 数据与方法

1.1 数据集

LaboroTomato 是专为目标检测和实例分割任务而设计的番茄图像数据集,包含了不同成熟阶段的番茄图像^[17]。这些番茄图像均在果园环境中通过两个不同分辨率的独立相机拍摄,分辨率分别为 3 024 像素×4 032 像素和 3 120 像素×4 160 像素。数据集中的番茄图像按大小分为普通番茄和樱桃番茄。结合现行国标 GH/T 1193—2021^[18]和文献[19]对番茄成熟度等级定义,根据两种番茄实际生长状况和生产需求,划分出3类番茄成熟度:完全成熟、半成熟和绿色,因此番茄成熟度标签共6种。成熟度划分标准是完全成熟果实表面着红面积在90%以上,半成熟果实表面着红面积在30%~90%之间,绿色果实表面着红面积在0~30%之间。该数据集共包含804幅图像,按比例7:2:1划分训练集、验证集和测试集。图1为训练集中部分番茄图像。

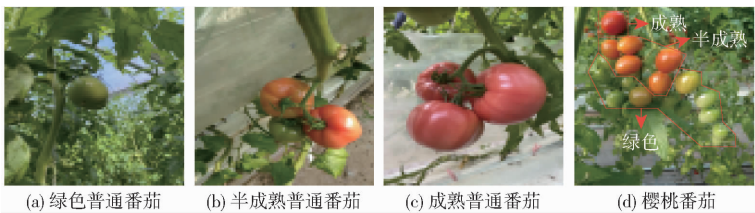


图1 不同成熟阶段普通番茄和樱桃番茄

Fig. 1 Large tomatoes and cherry tomatoes at different stages of maturity

图像尺寸增大往往伴随着训练和推理时间延长,并且导致网络感受野减小,影响中高层特征提取,进而造成检测性能下降^[20-21]。因此,将训练集

和验证集图像同时缩放并拉伸至 YOLO v8 要求的 640 像素×640 像素。这两种操作在一定程度上降低了图像清晰度,改变了番茄原始比例,但为算法提

供了规范化输入,有助于减少处理不同尺寸图像所需的计算资源,并且提升了处理速度。

鉴于番茄数据集的样本量不足以使模型在训练过程中充分收敛,采用水平翻转和垂直翻转的数据增强方法^[19],如图2所示。通过这种方法,将训练集由563幅图像扩充到1282幅图像,从而增加样本多样性,提高卷积神经网络泛化能力和鲁棒性。



图2 数据增强方法示例图

Fig.2 Data augmentation methods

1.2 YOLO v8 网络

YOLO v8 网络具有推理速度快、模型尺寸小以及支持多种硬件平台的特点^[22],将其作为番茄成熟度检测模型的基础算法,能够满足实时性和轻量化的需求。虽然在一般场景下的粗粒度检测任务中,YOLO v8 已表现出较高精度,但在处理相邻番茄成熟度特征差异较小的细粒度检测任务时,现有的主干网络不足以在枝叶遮挡、果实重叠等复杂背景干扰的情况下区分相似的特征。因此,以参数量最少的 YOLO v8n 网络作为改造对象,通过引入注意力机制,在保持网络轻量化并且训练数据较少的情况下,从复杂背景的干扰中对番茄特征给予更多关注,以提高实时检测效率和精度。

1.3 通道嵌入位置注意力模块 CePAM

首先对输入特征 X 沿水平维度 H 和垂直维度 W 分别做全局最大池化和全局平均池化,第 c 通道在高度 h 和宽度 w 处最大池化输出可分别表示为

$$Y_c^{\max}(h) = \max_{0 \leq j \leq W} X_c(h, j) \quad (1)$$

$$Y_c^{\max}(w) = \max_{0 \leq i \leq H} X_c(i, w) \quad (2)$$

同理,第 c 通道在高度 h 和宽度 w 处的平均池化输出可分别表示为

$$Y_c^{\text{avg}}(h) = \frac{1}{W} \sum_{0 \leq j \leq W} X_c(h, j) \quad (3)$$

$$Y_c^{\text{avg}}(w) = \frac{1}{H} \sum_{0 \leq i \leq H} X_c(i, w) \quad (4)$$

式中 X ——输入的中间特征图, $X \in \mathbf{R}^{C \times H \times W}$

$(H, 1)$ 、 $(1, W)$ ——池化空间范围

$Y(h)$ ——沿水平方向池化输出特征图,

$$Y(h) \in \mathbf{R}^{C \times H \times 1}$$

$Y(w)$ ——沿垂直方向池化输出特征图,

$$Y(w) \in \mathbf{R}^{C \times 1 \times W}$$

对池化操作得到的4个空间上下文描述符进行两两交叉融合。由于在训练阶段,采用了水平翻转和垂直翻转的数据增强操作,因此交叉融合考虑图像翻转后维度的变化,融合后描述符可分别表示为

$$Y_c^h = [Y_c^{\max}(h), Y_c^{\text{avg}}(w)] \quad (5)$$

$$Y_c^w = [Y_c^{\max}(w), Y_c^{\text{avg}}(h)] \quad (6)$$

式中 $[\cdot, \cdot]$ ——沿空间维度的级联运算符

将两个中间特征向量 $Y^h \in \mathbf{R}^{C \times 2H \times 1}$ 和 $Y^w \in \mathbf{R}^{C \times 1 \times 2W}$ 发送到共享网络,共享网络由两个共享参数的 1×1 卷积变换 W_1 、 W_2 ,以及 ReLU 激活函数和 Sigmoid 激活函数组成。为了减少模型复杂度开销,以适当的缩减率 r (Reduction ratios, $r = 16$) 减少 W_1 的通道数。两个中间特征向量经过共享网络之后,通过逐元素相乘的方式合并特征向量,即

$$Y = \sigma(W_2 \delta(W_1 Y^h)) \times \sigma(W_2 \delta(W_1 Y^w)) \quad (7)$$

式中 $\times$ ——逐元素相乘符号

σ ——Sigmoid 函数

W_1 、 W_2 ——两个共享参数的 1×1 卷积, $W_1 \in \mathbf{R}^{C \times (C/r)}$, $W_2 \in \mathbf{R}^{(C/r) \times C}$

δ ——ReLU 激活函数

最后再对融合后特征图进行 3×3 卷积操作,将水平维度和垂直维度降低到和输入特征图 X 相同大小。之后通过乘法操作将注意力广播,最后 CePAM 注意力模块输出 Y 可表示为

$$Y_c(i, j) = X_c(i, j) \cdot \sigma(f^{3 \times 3}(Y_c(i, j))) \quad (8)$$

式中 $f^{3 \times 3}$ ——大小为 3×3 、步长为1的卷积核

CePAM 注意力模块整体结构如图3所示,全局平均池化和全局最大池化用来聚合输入特征图的特征,最大池化突出前景,保留主要特征,平均池化突出背景,保留背景信息。在上述池化操作中并不是直接从整个空间维度聚合特征图的空间信息,而是分别沿水平方向和垂直方向将图像特征聚合到两个空间维度 H 和 W 上,得到4个方向可知的空间上下文描述符,每个上下文描述符都沿同一空间维度捕获目标区域间的长距离依赖关系,在另一个空间维度上建立目标位置信息,可以更准确地定位并识别目标区域^[23]。与池化整幅特征图相比,通过目标位置信息编码通道关系和长距离依赖关系,可以避免在较远位置之间建立大量不必要连接,从而减少计算开销^[24]。

1.4 改进 CBAM 注意力模块中的空间注意力

CBAM 是一种简单而高效的前馈卷积神经网络注意力模块,通过学习不同特征通道和特征空间的重要性权重生成通道注意力图和空间注意力图,然后将两个注意力图逐次与输入的中间特征图相乘,自适应提取重要特征。由于 CBAM 注意力机制的

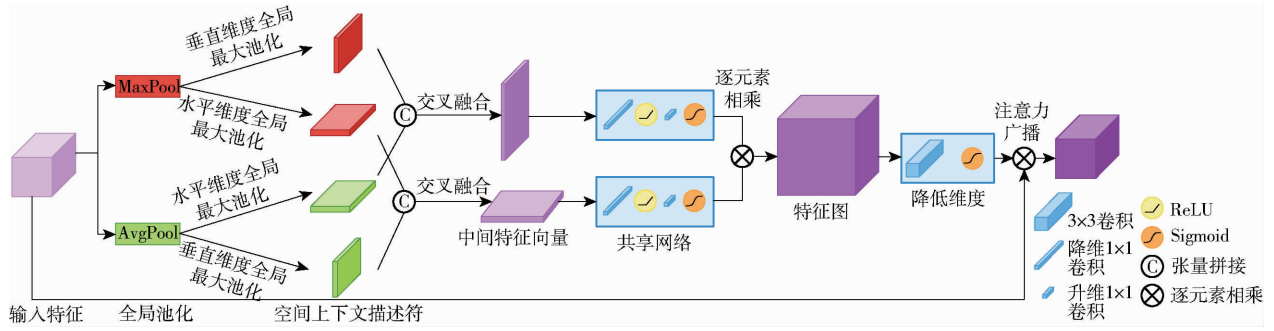


图 3 通道嵌入位置注意力模块结构图

Fig. 3 Channel embedded positional attention module structure diagram

参数量少,计算开销可忽略不计,非常适合嵌入到轻量型目标检测网络中。

大尺度卷积核具有更大的感受野,能够聚合远距离上下文信息,小尺度卷积核可以有效提取局部细节特征^[25]。然而,CBAM 的空间注意力模块采用单一尺度的 7×7 卷积核获取远距离空间信息的方式并不完全有效。如果继续增大卷积核尺寸容易导致难以兼顾局部特征和计算量成倍增加的问题。受 Visual Attention Network 中大核注意力启发,将分解后大核卷积替换 7×7 卷积^[26]。具体来说,将 15×15 的大核卷积分解成 5×5 深度卷积、扩张率为 3 的 5×5 深度扩张卷积和 1×1 点卷积的组合,之后重新引入 Softmax 激活函数,将每个空间位置分配的值转化为对应空间特征重要性权重,与通道注意力输出相适应。大核卷积具体分解过程如下:

假设大核卷积核数为 K ,则深度卷积核数

$K_{DW-Conv}$ 为

$$K_{DW-Conv} = \frac{K}{d} \quad (9)$$

深度扩张卷积核数 $K_{DW-D-Conv}$ 为

$$K_{DW-D-Conv} = 2d - 1 \quad (10)$$

式中 d ——深度扩张卷积的扩张率

改进后空间注意力模块输出可表示为

$$Y_s(X) = \sigma(f_{Conv}^{1 \times 1}(f_{DW-D-Conv}^{5 \times 5}(f_{DW-Conv}^{5 \times 5} \cdot ([AvgPool(X); MaxPool(X)])))) \quad (11)$$

式中 $f_{Conv}^{1 \times 1}$ —— 1×1 点卷积

$f_{DW-D-Conv}^{5 \times 5}$ ——扩张率为 3 的 5×5 深度扩张卷积

$f_{DW-Conv}^{5 \times 5}$ —— 5×5 深度卷积

将改进空间注意力模块融入到 CBAM 中形成大核卷积块注意力模块 (Large kernel convolutional block attention module, LKCBAM), 其结构如图 4 所示。

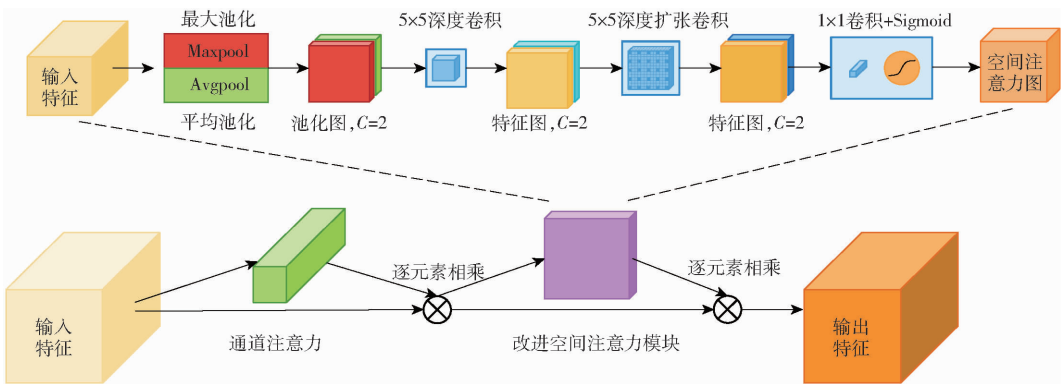


图 4 LKCBAM 注意力模块结构图

Fig. 4 LKCBAM attention module structure diagram

分别实验不同的扩张率和不同的深度卷积核、深度扩张卷积核,对网络带来的影响,发现大核数为 13 时, 5×5 的深度卷积和扩张率为 3 的 5×5 深度扩张卷积可以带来显著的性能提升,推测可能的原因:①使用相对较小的 5×5 深度卷积可以有效地捕获局部上下文信息,并且防止引入过多目标区域之外的冗余信息。②扩张率为 3 的 5×5 深度扩张

卷积的实际感受野为 13×13 ,与 5×5 深度卷积叠加后,感受野达到 17×17 ,比 7×7 卷积更能捕获长距离依赖关系。③深度卷积和深度扩张卷积的组合能够同时提取局部和全局的空间信息,同时避免了使用大卷积核带来的计算成本过高的问题。

1.5 改进 YOLO v8 网络

通过在特征提取网络引入注意力模块,提出一

种基于改进 YOLO v8n 网络的番茄成熟度检测模型。为了应对枝叶遮挡、果实重叠等复杂环境,改进 YOLO v8n 网络执行两种检测任务:①通过实例分

割将果实的边界精确标记出来,以应用于自动化采摘。②通过目标检测检测番茄成熟度。改进后网络结构如图 5 所示。

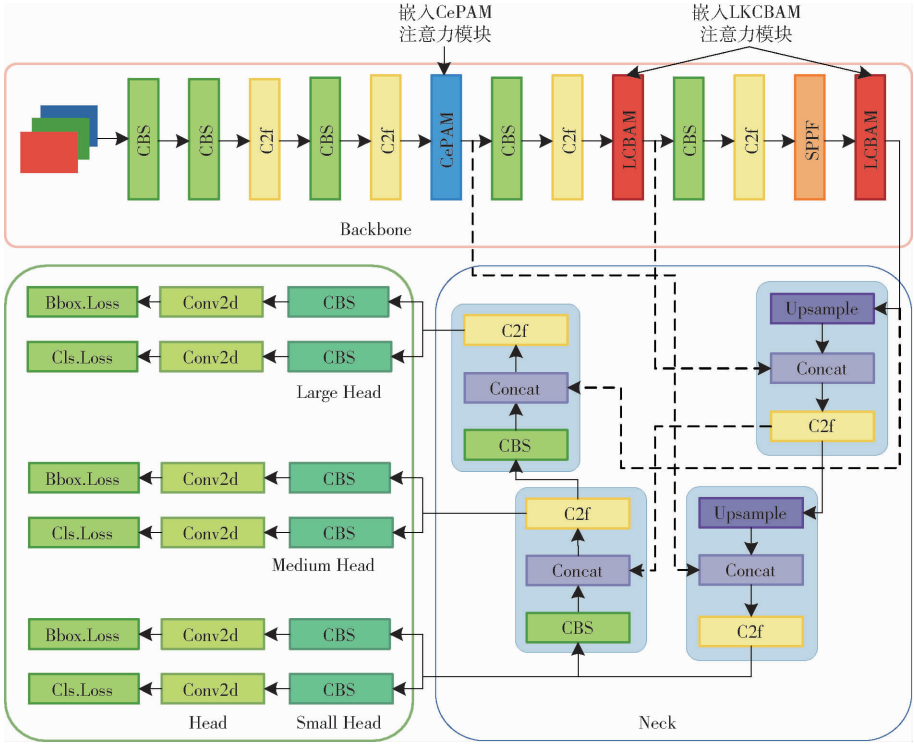


图 5 改进 YOLO v8n 网络结构图

Fig. 5 Improved YOLO v8n network structure diagram

主干网络完成特征提取后,P3、P4 和 P5 特征层参与特征融合,因此分别在这 3 个特征层最后引入 CePAM 和 LCBAM 注意力模块。P3 层作为浅层网络感受野比较小,一般包含更多的位置和细节信息,在此阶段最后加入 CePAM 注意力模块,处理不同通道的细节特征权重分配关系,通过对每个通道进行加权,增加对有用细节特征通道的关注。同时,两个空间维度上的池化操作保留了细节特征的位置信息,有助于在 P3 层特征融合后的小目标检测头对樱桃番茄精确定位与识别。

P4 层作为中层网络,在此阶段的最后引入改进后的 LCBAM 注意力模块,通道注意力强化普通番茄目标的特征,抑制不重要的果园环境背景特征。改进后的空间注意力得益于较大的感受野,可以更加关注普通番茄目标的空间位置。

P5 层作为深层网络,感受野比较大,包含更多的语义和全局信息,在此阶段最后继续引入 LCBAM 模块。改进后空间注意力可以有效地捕获局部信息和两个目标区域的长距离依赖关系,能够强化果园环境之间的相关性和番茄之间的相关性,弱化果园环境和番茄之间的相关性,从而明确区分前景和背景。

2 实验设置

实验硬件配置为 GeForce RTX 3090 24G GPU 显卡、Intel(R) Xeon(R) Platinum 8358P CPU 处理器和 80 GB RAM,软件配置为 Ubuntu 20.4 Linux 操作系统、Python 3.8.10 和 CUDA 12.0,YOLO v8n 以及改进型基于 Pytorch 深度学习框架运行。

模型训练参数为:学习率 0.01、训练轮次 300、SGD 动量 0.937、衰减率 0.000 5、批处理 32、随机种子 1 234、进程数 0。

3 评价指标

为了验证模型性能,采用精确率(Precision)、召回率(Recall)、平均精度均值(mAP)、参数量(Parameters)、浮点运算次数(FLOPs)和推理速度等评价指标^[27]。

4 结果与分析

改进 YOLO v8n 网络被训练 300 轮次,训练损失和验证损失下降过程如图 6 和图 7 所示。在前 50 个轮次,模型损失快速减小,网络迅速拟合;在 150~300 个轮次,模型损失趋于平稳并且收敛,没有出现欠拟合和过拟合现象。

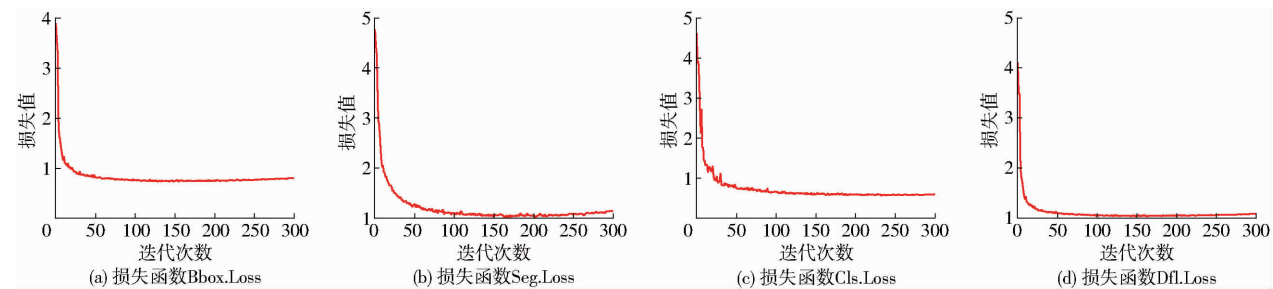


图 6 训练损失下降过程

Fig. 6 Training loss reduction process

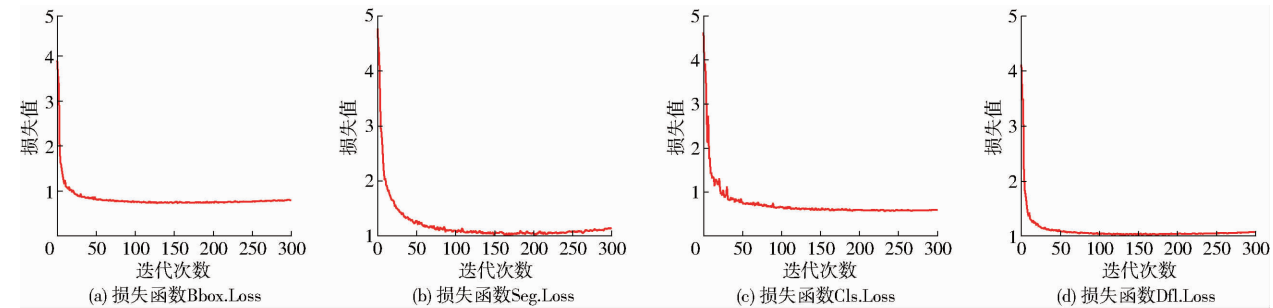


图 7 验证损失下降过程

Fig. 7 Validation loss reduction process

为验证改进 YOLO v8n 网络的有效性,将 YOLO v5n、YOLO v7 和 YOLO v8n 网络与改进 YOLO v8n 网络进行评估对比,并对改进 YOLO v8n 网络进行消融实验。所有算法的训练过程均在相同配置环境下进行,并且均未使用预训练权重。需要指出的是,由于本文无法满足 YOLO v7 网络训练实例分割任务所需的算力条件,因此不将其纳入实例分割性能对比的考虑范围。

由表 1 可知,尽管改进 YOLO v8n 网络在参数量和计算量方面相较于 YOLO v5n 网络有所增加,但除目标检测任务的召回率和实例分割任务的精确率之外,其他性能指标均得到有效提升,其中两个 mAP@ 50 – 95 提升最为显著,分别达到 4、3.6 个百

分点。虽然 YOLO v7 网络在各性能指标上相较于改进 YOLO v8n 网络提高约 2 个百分点,但 YOLO v7 网络深度的增加带来了高计算量和高参数量,不适宜将其部署在资源受限的嵌入式环境中。改进 YOLO v8n 相较于原始版本,在不增加额外计算负担的前提下,在目标检测任务上,精确率、召回率、mAP@ 50 和 mAP@ 50 – 95 分别提高 – 1.5、2.6、0.4、1.4 个百分点,在实例分割任务上,精确率、召回率、mAP@ 50 和 mAP@ 50 – 95 分别提高 – 1.6、2.9、0.3、1.2 个百分点。另外,在推理速度方面,改进 YOLO v8n 与原 YOLO v8n 网络相当,优于 YOLO v7 和 YOLO v5n 网络。

消融实验结果如表 2 所示,尽管改进 YOLO v8n

表 1 4 种算法性能指标对比结果

Tab. 1 Comparison of performance metrics of four algorithms

模型	目标检测任务/%				实例分割任务/%				模型尺寸		推理 时间/ms
	精确率	召回率	mAP@ 50	mAP@ 50 – 95	精确率	召回率	mAP@ 50	mAP@ 50 – 95	参数量	浮点运算次数	
YOLO v5n	79.8	79.5	84.4	67.9		78.5	84.2	65.6	1.90×10^6	6.80×10^9	1.7
YOLO v7	81.8	81.7	87.1	74.1					3.72×10^7	1.05×10^{11}	14.3
YOLO v8n	81.6	76.2	84.9	70.5	81.4	75.8	84.6	68.0	3.30×10^6	1.20×10^{10}	1.2
改进 YOLO v8n	80.1	78.8	85.3	71.9	79.8	78.7	84.9	69.2	3.30×10^6	1.25×10^{10}	1.1

网络结合 CePAM 模块使得精确率有所降低,但召回率分别提高 2.5、2.9 个百分点,并且 mAP@ 50 均提高 0.3 个百分点。CePAM 模块通过在中间特征图中应用水平和垂直池化来嵌入特征位置信息,有效解决了模型对被遮挡番茄检测不足的问题,从而在一定程度上提高了召回率。引入 LKCBAM 模块后,

改进 YOLO v8n 网络在 mAP@ 50 上略有降低,但其他指标均获得不同程度提升。具体来说,精确率分别提高 0.4、0.8 个百分点,召回率分别提高 1.1、1.5 个百分点。原因在于,深度卷积与深度扩张卷积的组合有助于提取番茄的细节特征,并通过扩大感受野建立番茄之间的长距离依赖关系,从而同时

表 2 CePAM 和 CBAM 消融实验对比结果

Tab.2 Comparison of CePAM and CBAM ablation experiments

模型	目标检测任务				实例分割任务			
	精确率	召回率	mAP@ 50	mAP@ 50 - 95	精确率	召回率	mAP@ 50	mAP@ 50 - 95
YOLO v8n	81. 6	76. 2	84. 9	70. 5	81. 4	75. 8	84. 6	68. 0
YOLO v8n + CePAM	80. 0	78. 7	85. 2	70. 5	79. 7	78. 7	84. 9	68. 6
YOLO v8n + LKCBAM	82. 0	77. 3	84. 8	71. 7	82. 2	77. 3	84. 2	69. 1
YOLO v8n + CePAM + LKCBAM	80. 1	78. 8	85. 3	71. 9	79. 8	78. 7	84. 9	69. 2

增强了精确率和召回率。最后,通过同时使用 CePAM 和 LKCBAM 模块,均衡了精确率和召回率之间的矛盾关系,除精确率外,其他性能指标相较于上述 3 组模型都获得了提高。

为验证改进 YOLO v8n 网络实际检测效果,分别使用 YOLO v5n、YOLO v8n、改进 YOLO v8n 网络对测试集图像进行检测,部分检测和分割结果如图 8 所示。在普通番茄成熟度检测中,YOLO v5n 和

YOLO v8n 网络都出现了两个重叠番茄边界分割模糊和缺失的情况,相比之下,改进 YOLO v8n 网络实例分割准确,边界区分明显,检测置信度高。在樱桃番茄成熟度检测中,YOLO v5n 和 YOLO v8n 网络都出现了不同程度的误检问题,尤其是在枝叶遮挡、背景杂乱情况下,将地面误检成了番茄,原因在于两者都未有效建立环境区域和目标区域之间的长距离依赖关系。

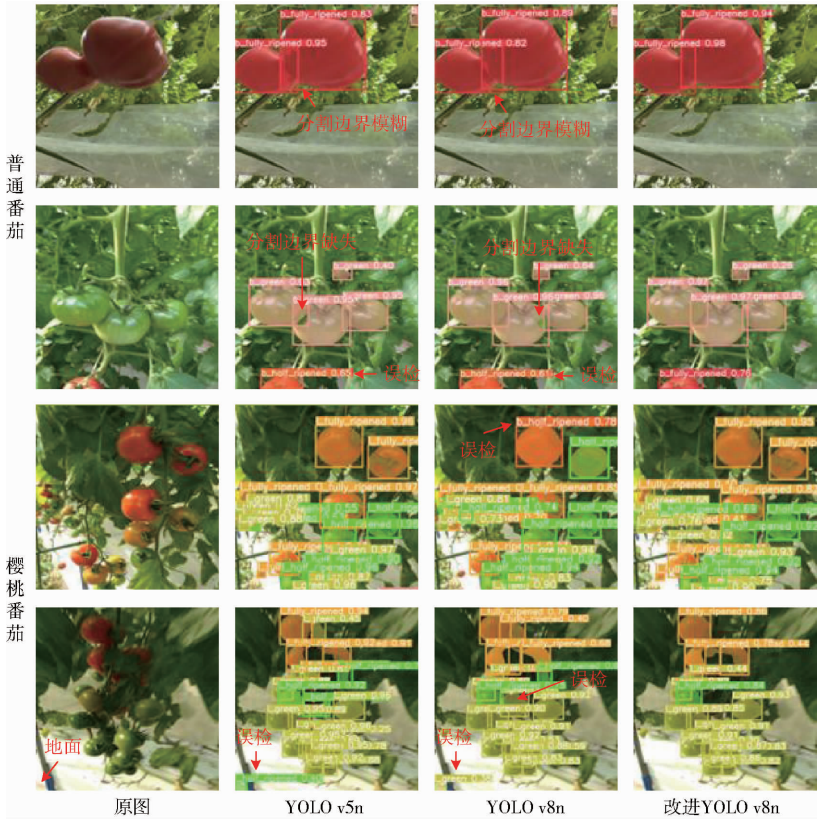


图 8 3 种算法实际检测和分割效果对比

Fig.8 Comparison of actual detection and segmentation effects of three algorithms

Grad - CAM 是一种解释卷积神经网络模型的可视化工具,使用该工具能够以热力图的形式清晰展示神经网络对被预测图像的关注区域^[28]。原 YOLO v8n 和改进 YOLO v8n 网络 P3 和 P4 两个浅层特征层和小目标检测层热力掩膜如图 9 和图 10 所示。

对比 P3 层和 P4 层热力图可以看出,P3 层 CePAM 注意力模块可以在浅层网络保留更多的樱

桃番茄特征,使得小目标检测层能够更加准确地识别樱桃番茄。P4 层 LKCBAM 模块建立距离较远的两个目标区域的长距离依赖关系,使得番茄像素区域之间联系紧密,番茄与背景区域之间区分明显。从整体热力图对比来看,改进 YOLO v8n 网络不但更加关注普通番茄目标,而且对于樱桃番茄这种小目标检测能力也有较大的提升,对于复杂的环境背景并没有给予过多关注。

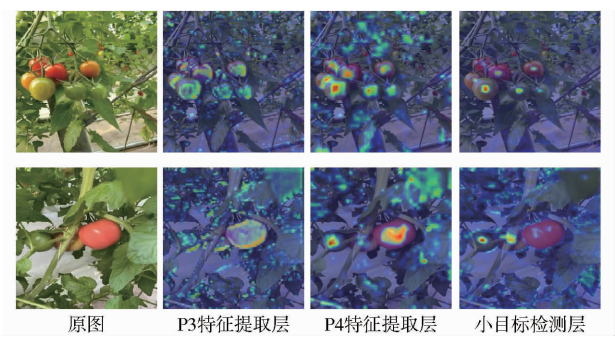


图 9 YOLO v8n 部分网络层热力分布图

Fig. 9 Heat map of some network layers of YOLO v8n

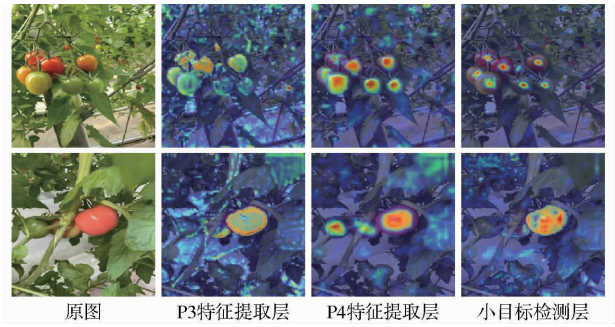


图 10 改进 YOLO v8n 部分网络层热力分布图

Fig. 10 Heat map of some network layers of improved YOLO v8n

5 轻量化部署

以 Jetson Nano 为核心硬件,设计一款能够满足番茄成熟度实时检测需求的农用嵌入式 AI 平台,如图 11 所示。该平台具有低成本、低功耗和便携式的特点。受限于 Jetson Nano 的计算能力,直接运行改进 YOLO v8n 网络进行摄像头实时检测时,会出现 4 GB 内存满溢导致操作系统不能运行的现象。TensorRT 是 NVIDIA 公司推出的一款高性能神经网络推理引擎,与 NVIDIA 的 GPU 相结合,使得神经网络可以提供最大推理吞吐量和低延迟速度。TensorRT 可在保证检测精度的情况下实现改进 YOLO v8n 网络推理加速,基于 TensorRT 的模型部署流程如图 12 所示。

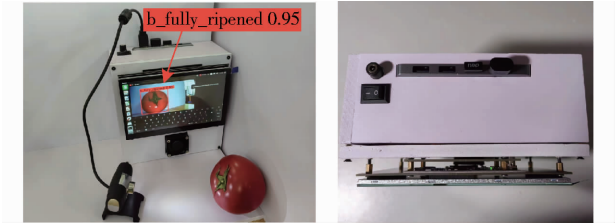


图 11 运行番茄成熟度实时检测模型的农用嵌入式 AI 平台

Fig. 11 Agricultural embedded AI platform running real-time tomato maturity detection model

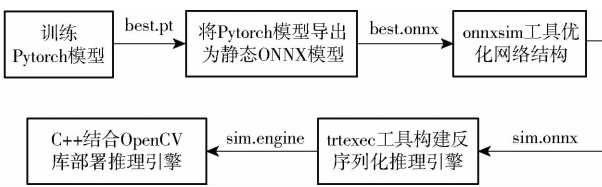


图 12 改进 YOLO v8n 在 Jetson Nano 上的部署流程图
Fig. 12 Improved YOLO v8n deployment process on Jetson Nano

为验证改进 YOLO v8n 网络在嵌入式平台上的实时运行效果,选择相邻成熟度特征差异小的 5 个完全成熟、4 个半成熟和 4 个未成熟樱桃番茄,使用改进 YOLO v8n 网络对 13 个番茄进行成熟度检测,检测结果如图 13 所示。实验结果表明,改进 YOLO v8n 网络将所有番茄成熟度全部准确检测出来,检测置信度在 0.7 左右,最高置信度达 0.86,模型内存占用量 2.4 GB,实时检测速度为 10 f/s。并且改进 YOLO v8n 网络能够长时间稳定运行,可以满足采摘机器人等农用设备对番茄成熟度实时检测的需求。

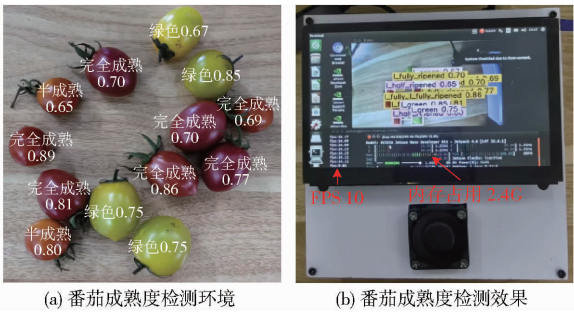


图 13 改进 YOLO v8n 在 Jetson Nano 上的摄像头实时检测效果

Fig. 13 Improved YOLO v8n real-time camera detection effect on Jetson Nano

6 结束语

提出了 CePAM 注意力模块,改进了 CBAM 模块中的空间注意力,通过将两种模块嵌入到 YOLO v8n 网络中,提出一种带有注意力机制的改进 YOLO v8n 网络用于复杂环境下的番茄成熟度检测与实例分割。3 种验证实验结果表明,改进 YOLO v8n 网络在不增加额外计算负担的同时提高了检测与分割精度。检测效果对比实验和热力图实验结果表明,使用两种注意力模块可以使模型更加关注番茄目标区域,即使在枝叶遮挡和果实重叠的情况下,也能够准确识别并分割番茄目标。同时,在基于 Jetson Nano 的嵌入式平台上进行了轻量化部署,整套低成本、低功耗的视觉装置为果园番茄采摘设备的视觉系统开发提供了新的思路。

参 考 文 献

- [1] TIEMAN D, ZHU G, RESENDE Jr M F R, et al. A chemical genetic roadmap to improved tomato flavor[J]. *Science*, 2017, 355(6323): 391–394.
- [2] Food and Agriculture Organization of the United Nations. FAO Statistical Database[EB/OL]. (2023-12-27). <https://www.fao.org/faostat/zh/#/data/QCL>. (in Chinese)
- [3] 刘成良, 贡亮, 苑进, 等. 农业机器人关键技术研究现状与发展趋势[J]. *农业机械学报*, 2022, 53(7): 1–22, 55.
LIU Chengliang, GONG Liang, YUAN Jin, et al. Current status and development trends of agricultural robots[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(7): 1–22, 55. (in Chinese)
- [4] 刘继展. 温室采摘机器人技术研究进展分析[J]. *农业机械学报*, 2017, 48(12): 1–18.
LIU Jizhan. Research progress analysis of robotic harvesting technologies in greenhouse[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2017, 48(12): 1–18. (in Chinese)
- [5] 苟国旻, 闫建伟, 张富贵, 等. 水果采摘机器人视觉系统与机械手研究进展[J]. *计算机工程与应用*, 2023, 59(9): 13–26.
GOU Yuanmin, YAN Jianwei, ZHANG Fugui, et al. Research progress on vision system and manipulator of fruit picking robot[J]. *Computer Engineering and Applications*, 2023, 59(9): 13–26. (in Chinese)
- [6] MU Y, CHEN T S, NINOMIYA S, et al. Intact detection of highly occluded immature tomatoes on plants using deep learning techniques[J]. *Sensors*, 2020, 20(10): 2984.
- [7] LIU G, NOUAZE J C, TOUKO M P L, et al. YOLO-tomato: a robust algorithm for tomato detection based on YOLOv3[J]. *Sensors*, 2020, 20(7): 2145.
- [8] 龙洁花, 赵春江, 林森, 等. 改进 Mask R-CNN 的温室环境下不同成熟度番茄果实分割方法[J]. *农业工程学报*, 2021, 37(18): 100–108.
LONG Jiehua, ZHAO Chunjiang, LIN Sen, et al. Segmentation method of the tomato fruits with different maturities under greenhouse environment based on improved Mask R-CNN[J]. *Transactions of the CSAE*, 2021, 37(18): 100–108. (in Chinese)
- [9] WANG Z, LING Y, WANG X, et al. An improved Faster R-CNN model for multi-object tomato maturity detection in complex scenarios[J]. *Ecological Informatics*, 2022, 72: 101886.
- [10] KIM J Y, PYO H R, JANG I, et al. Tomato harvesting robotic system based on deep-ToMaToS: deep learning network using transformation loss for 6D pose estimation of maturity classified tomatoes with side-stem[J]. *Computers and Electronics in Agriculture*, 2022, 201: 107300.
- [11] SU F, ZHAO Y, WANG G, et al. Tomato maturity classification based on SE-YOLOv3-MobileNetV1 network under nature greenhouse environment[J]. *Agronomy*, 2022, 12(7): 1638.
- [12] LI R, JI Z, HU S, et al. Tomato maturity recognition model based on improved YOLO v5 in greenhouse[J]. *Agronomy*, 2023, 13(2): 603.
- [13] WANG X, WU Z, JIA M, et al. Lightweight SM-YOLOv5 tomato fruit detection algorithm for plant factory[J]. *Sensors*, 2023, 23(6): 3336.
- [14] GUO J, YANG Y, LIN X, et al. Revolutionizing agriculture: real-time ripe tomato detection with the enhanced tomato-YOLO v7 system[J]. *IEEE Access*, 2023, 11: 133086–133098.
- [15] 刘建航, 何鉴恒, 陈海华, 等. 基于改进 YOLO v4 和 ICNet 的番茄串检测模型[J]. *农业机械学报*, 2023, 54(10): 216–224, 254.
LIU Jianhang, HE Jianheng, CHEN Haihua, et al. Development of detection model for tomato clusters based on improved YOLO v4 and ICNet[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(10): 216–224, 254. (in Chinese)
- [16] 孙宇朝, 李守豪, 夏秀波, 等. 利用改进 YOLO v5s 模型检测番茄果实成熟度及外观品质[J]. *园艺学报*, 2024, 51(2): 396–410.
SUN Yuchao, LI Shouhao, XIA Xiubo, et al. Using the improved YOLO v5s model to detect tomato fruit maturity and appearance quality[J]. *Acta Horticulturae Sinica*, 2024, 51(2): 396–410. (in Chinese)
- [17] ZHANG W, LIU Y, CHEN K, et al. Lightweight fruit-detection algorithm for edge computing applications[J]. *Frontiers in Plant Science*, 2021, 12: 740936.
- [18] GH/T 1193—2021 番茄[S]. 2021.
- [19] 苗荣慧, 李志伟, 武锦龙. 基于改进 YOLO v7 的轻量化樱桃番茄成熟度检测方法[J]. *农业机械学报*, 2023, 54(10): 225–233.
MIAO Ronghui, LI Zhiwei, WU Jinlong. Lightweight maturity detection of cherry tomato based on improved YOLO v7[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(10): 225–233. (in Chinese)
- [20] LUKE J, JOSEPH R, BALAJI M. Impact of image size on accuracy and generalization of convolutional neural networks[J]. *Int. J. Res. Anal. Rev. (IJRAR)*, 2019, 6(1): 70–80.
- [21] SAPONARA S, ELHANASHI A. Impact of image resizing on deep learning detectors for training time and model performance [C]//International Conference on Applications in Electronics Pervading Industry, 2021: 10–17.

based on correlation-matter element analysis[J]. *Journal of Cleaner Production*, 2023, 408: 137231.

[13] ZUO Y J, GUO Y D, SONG C C, et al. Study on soil water and heat transport characteristic responses to land use change in Sanjiang Plain[J]. *Sustainability*, 2019, 11(1): 157.

[14] YANG Q F, ZHANG P Y, MA Z P, et al. Agricultural economic resilience in the context of international food price fluctuation—an empirical analysis on the main grain-producing areas in Northeast China[J]. *Sustainability*, 2022, 14(21): 14102.

[15] 王立权. 兴凯湖灌区开展用水户参与灌溉管理的实践与探讨[J]. *现代化农业*, 2005(9): 4-5.

[16] 吴普特, 孙世坤, 王玉宝, 等. 作物生产水足迹量化方法与评价研究[J]. *水利学报*, 2017, 48(6): 651-660, 669.
WU Pute, SUN Shikun, WANG Yubao, et al. Research on the quantification methods for water footprint of crop production [J]. *Journal of Hydraulic Engineering*, 2017, 48(6): 651-660, 669. (in Chinese)

[17] SUN S K, WU P T, WANG Y B, et al. The impacts of interannual climate variability and agricultural inputs on water footprint of crop production in an irrigation district of China[J]. *Science of the Total Environment*, 2013, 444: 498-507.

[18] CAO X C, SHU R, REN J, et al. Variation and driving mechanism analysis of water footprint efficiency in crop cultivation in China[J]. *Science of the Total Environment*, 2020, 725: 138537.

[19] XU Y N, LI J N, LI H, et al. The water-food-energy nexus evaluation and optimization of cropping system in the North China Plain: a case of county scale[J]. *Journal of Cleaner Production*, 2024, 443: 141200.

[20] KONG Y, HE W J, ZHANG Z F, et al. Spatial-temporal variation and driving factors decomposition of agricultural grey water footprint in China[J]. *Journal of Environmental Management*, 2022, 318: 115604.

[21] WANG C Y, CHEN J, GU L, et al. A pathway analysis method for quantifying the contributions of precipitation and potential evapotranspiration anomalies to soil moisture drought[J]. *Journal of Hydrology*, 2023, 621: 129570.

[22] 李铁男. 黑龙江省水资源承载力评价[R]. 哈尔滨: 黑龙江省水利科学研究院, 2017.

[23] 王征. 黑龙江省锦西灌区水资源优化配置与农业节水潜力分析[D]. 哈尔滨: 东北农业大学, 2018.
WANG Zheng. Optimal allocation of water resources and analysis of agricultural water-saving potential in Jinxi Irrigation District, Heilongjiang Province[D]. Harbin: Northeast Agricultural University, 2018. (in Chinese)

[24] 付强, 刘巍, 吕纯波, 等. 灌区灌溉水利用效率分析及用水优化管理[M]. 北京: 科学出版社, 2020.

[25] 郭相平, 高爽, 吴梦洋, 等. 中国农作物水足迹时空分布与影响因素分析[J]. *农业机械学报*, 2018, 49(5): 295-302.
GUO Xiangping, GAO Shuang, WU Mengyang, et al. Analysis of temporal-spatial distribution and influencing factors of water footprint in crop production system of China [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018, 49(5): 295-302. (in Chinese)

[26] HOU R J, QI Z Y, LI T X, et al. Mechanism of snowmelt infiltration coupled with salt transport in soil amended with carbon-based materials in seasonally frozen areas[J]. *Geoderma*, 2022, 420: 115882.

[27] LIU J L, LIU Q Q, LI J W, et al. Response of soil moisture and water use efficiency to straw mulching amount and mulching period in black soil zone of northeast China[J]. *Canadian Journal of Soil Science*, 2023, 103(4): 634-641.

[28] GRONLE A, LUX G, BÖHM H, et al. Effect of ploughing depth and mechanical soil loading on soil physical properties, weed infestation, yield performance and grain quality in sole and intercroops of pea and oat in organic farming[J]. *Soil & Tillage Research*, 2015, 148: 59-73.

[29] HE Y B, LIU D M, YAO Y M, et al. Spatializing growth suitability for spring soybean cultivation in Northeast China[J]. *Journal of Applied Meteorology and Climatology*, 2013, 52(4): 773-783.

(上接第 382 页)

[22] 罗斌,李家超,周亚男,等. 基于改进 YOLO v8 的轻量化稻瘟病孢子检测方法[J]. *农业机械学报*, 2024, 55(11): 32-38.
LUO Bin, LI Jiachao, ZHOU Ya'nan, et al. Lightweight rice blast spores detection method based on improved YOLO v8[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(11): 32-38. (in Chinese)

[23] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 13713-13722.

[24] HOU Q, ZHANG L, CHENG M M, et al. Strip pooling: rethinking spatial pooling for scene parsing[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 4003-4012.

[25] DING X, ZHANG X, HAN J, et al. Scaling up your kernels to 31x31: revisiting large kernel design in CNNs[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 11963-11975.

[26] GUO M H, LU C Z, LIU Z N, et al. Visual attention network[J]. *Computational Visual Media*, 2023, 9(4): 1-20.

[27] PADILLA R, NETTO S L, DA SILVA E A B. A survey on performance metrics for object-detection algorithms[C] // *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE, 2020: 237-242.

[28] SELVARAJU R R, DAS A, VEDANTAM R, et al. Grad-CAM: why did you say that? [J]. *arXiv Preprint*, arXiv:1611.07450, 2016.