doi:10.6041/j.issn.1000-1298.2024.S2.025

基于EDH-YOLO的轻量型温室番茄检测方法

毕泽洋[1]　杨立伟[1]　吕树盛[2]　宫艳晶[2]　张俊宁[3]　赵礼豪[3]

（1. 中国农业大学信息与电气工程学院，北京 100083；2. 中国农业机械化科学研究院集团有限公司，北京 100083；
3. 北京信息科技大学机电工程学院，北京 100192）

摘要：针对番茄采摘机器人识别算法包含复杂的网络结构和庞大的参数体量，严重限制检测模型的响应速度问题，本文提出一种改进的轻量级YOLO v5（EDH-YOLO）算法。为了能够在保持较高识别精度的同时大幅降低计算复杂度和模型内存占用量，引入EfficientNet-B0的轻量级网络作为YOLO v5算法的骨干网络；为了在训练过程中更好地定位目标物体的同时提高检测算法精度，引入DIoU损失函数；为了降低模型计算复杂度和提高模型表达能力，引入一种轻量化的Hardswish激活函数。实验结果显示，EDH-YOLO模型在识别效果损失较小的情况下，精确率、召回率和平均精度均值分别为95.9%、93.1%和96.8%，模型内存占用量仅为7.3 MB，检测速度为53.2 f/s，对比YOLO v5原模型内存占用量降低55.3%，EDH-YOLO模型检测速度提升60.0%。对比Faster R-CNN、YOLO v7和YOLO v8，EDH-YOLO模型在不同光照和遮挡等情况下具有较高鲁棒性。同时，将EDH-YOLO模型通过模型转换部署到安卓（Android）平台中，优化模型推理过程，满足温室复杂环境下番茄目标果实实时识别需求，可为设施环境下基于移动边缘计算的机器人目标识别及自动采收作业提供技术支持。

关键词：温室番茄；深度学习；目标识别；YOLO v5；轻量化网络

中图分类号：TP391.4　　**文献标识码：**A　　**文章编号：**1000-1298(2024)S2-0246-09

Lightweight Greenhouse Tomato Detection Method Based on EDH-YOLO

BI Zeyang[1]　YANG Liwei[1]　LÜ Shusheng[2]　GONG Yanjing[2]　ZHANG Junning[3]　ZHAO Lihao[3]

(1. *College of Information and Electrical Engineering, China Agricultural University, Beijing* 100083, *China*
2. *Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing* 100083, *China*
3. *School of Mechatronics Engineering, Beijing Information Science and Technology University, Beijing* 100192, *China*)

Abstract: Considering the issue that the tomato-picking robot's recognition algorithm has a complex network structure and a large number of parameters, which severely limit the detection model's response speed, an improved lightweight YOLO v5 (EDH-YOLO) algorithm was proposed. To significantly reduce computational complexity and model size while maintaining high recognition accuracy, the lightweight EfficientNet-B0 network was introduced as the backbone of the YOLO v5 algorithm. To better locate target objects during training and improve detection accuracy, the DIoU loss function was introduced. To reduce the model's computational complexity and enhance its expressive ability, the lightweight Hardswish activation function was introduced. Experimental results indicated that the EDH-YOLO model achieved accuracy, recall, and average precision of 95.9%, 93.1%, and 96.8%, respectively, with minimal loss in recognition performance. The size of model was only 7.3 MB, and the detection speed reached 53.2 f/s. Compared with the original YOLO v5 model, the model size was reduced by 55.3%, and the detection speed of the EDH-YOLO model was increased by 60.0%. Compared with Faster R-CNN, YOLO v7, and YOLO v8, the EDH-YOLO model demonstrated higher robustness under various lighting and occlusion conditions. Additionally, the EDH-YOLO model was deployed on the Android platform through model conversion to optimize the inference process, meet real-time recognition requirements for tomato fruits in complex greenhouse environments, and provide technical support for robot target recognition and automatic harvesting operations based on mobile edge computing in facility environments.

Key words: greenhouse tomato; deep learning; object recognition; YOLO v5; lightweight network

收稿日期：2024-07-18　修回日期：2024-09-22

基金项目：北京高校重点研究培育项目(2021YJPY201)

作者简介：毕泽洋(1997—)，男，博士生，主要从事深度学习和机器视觉研究，E-mail: bizeyang@cau.edu.cn

通信作者：杨立伟(1981—)，女，副教授，博士生导师，主要从事智慧农业研究，E-mail: yangliwei@cau.edu.cn

0　引言

番茄是世界三大贸易蔬菜之一，我国是世界最大的番茄生产国[1]。利用机器人进行番茄采摘作业可以提高采摘效率、降低劳动成本，有利于提高果农的经济效益[2-3]。识别是实现采摘的前提，提高识别精确率可以减少非目标番茄的误采，从而显著提升采摘效率。在实际生产中果实目标之间存在的严重遮挡、光照不均等情况增加了其识别难度，因此需要进一步探索如何提高模型精度。

目前，国内外相关文献针对番茄检测算法研究主要分为传统目标检测算法与深度学习目标检测算法。基于传统图像处理方法的番茄果实识别研究，普遍从颜色、纹理和边缘等特征进行识别[4-7]，在一定程度上实现番茄的识别，但识别速度较慢，识别1幅图像至少需5 s，达不到实时需求，同时传统番茄检测算法对番茄检测的结果很容易受到光照、遮挡等影响导致番茄识别率低，很难满足实际工作需要。然而，随着计算机硬件的不断升级，基于深度学习的果实目标检测算法取得了巨大的进展，也应用到温室番茄的识别当中[8-10]。目前，目标检测算法主要分为两类：第1类为两阶段算法（two-stage）：先计算候选区域（region proposal）然后进行CNN分类的R-CNN系列网络，如R-CNN、Fast R-CNN[11]、Faster R-CNN[12]和Mask R-CNN[13]等，此类算法精确率较高，但神经网络模型较大，计算速度慢，无法满足实时检测的场景需求[14-15]；第2类为一阶段算法（one-stage）：在检测的过程中可以直接同时输出定位和分类结果，如YOLO[16]（You only look once）和SSD[17]等，此类算法模型小，计算速度快，对于小目标识别有较好的效果，更容易部署到智能手机和嵌入式平台等边缘计算设备，但是此类模型识别的精确率低于第一类。

近年来，越来越多学者致力于研究轻量化目标检测网络，以兼顾模型的移植部署、检测速度和精度，以期为果园采摘机器人提供技术支持[3,18-20]。此外，通过对目标检测模型嵌入统计功能，可自动计数识别果实数量，可进一步提升机器人作业的自动化程度，并且深度学习技术已成功应用于芒果[21]、柑橘[22]和苹果[23]等作物的目标数量统计中。

针对温室复杂环境、番茄个体表征因素、检测模型参数量大、计算量高、内存占用量大以及移动端部署困难等问题，本文提出一种改进型的轻量化YOLO v5算法（EDH-YOLO目标检测算法），为了在训练过程中更好地定位目标物体的同时提高检测算法的精度和鲁棒性，引入DIoU损失函数；为了降低模型计算复杂度，引入一种轻量化的Hardswish激活函数。根据温室番茄场景的复杂性，拍摄不同天气、不同光照角度、不同遮挡情况和不同拍摄角度的番茄图像，确保数据集的丰富性和多样性；通过对数据样本进行预处理和图像增强，提高数据集质量的同时更加适用于温室番茄的果实目标检测。将改进后的YOLO v5网络模型部署到手机端，并集成果实数量统计功能，以期为基于移动边缘计算的设施环境机器人作业提供技术支撑。

1　材料与方法

1.1　数据样本采集

为了提高深度学习模型在复杂温室环境中对番茄果实目标识别的鲁棒性、实时性和自适应性，于北京市通州区农业机械研究所京鹏智慧温室和山东省威海市荣成市的温室番茄园进行图像采集，番茄种植方式为吊蔓式栽培。

为增加样本多样性，在农艺专家的指导下，分别于2021年12月23日中午（阴天）、2022年1月9日下午（晴天）、2022年12月31日中午（晴天）和2023年2月11日下午（阴天）采集不同成熟度（绿熟期和坚熟期）、不同拍摄角度（俯拍、仰拍、平拍）、不同光照角度（侧光、逆光、顺光）、不同遮挡方式（果实遮挡果实、枝干遮挡果实、树叶遮挡果实）以及不同天气（晴天和阴天）的番茄图像。采集的可见光图像分辨率为4 928像素×3 264像素，颜色模式为sRGB，图像保存为jpg格式。使用Nikon D7000型单反相机和手机进行拍摄，共采集2 116幅图像，筛选后选择1 532幅作为训练样本。通过将番茄的最小外接矩形框作为真实框（Ground truth），提高了真实框内数据参数的准确性。复杂情况下温室番茄图像如图1所示。

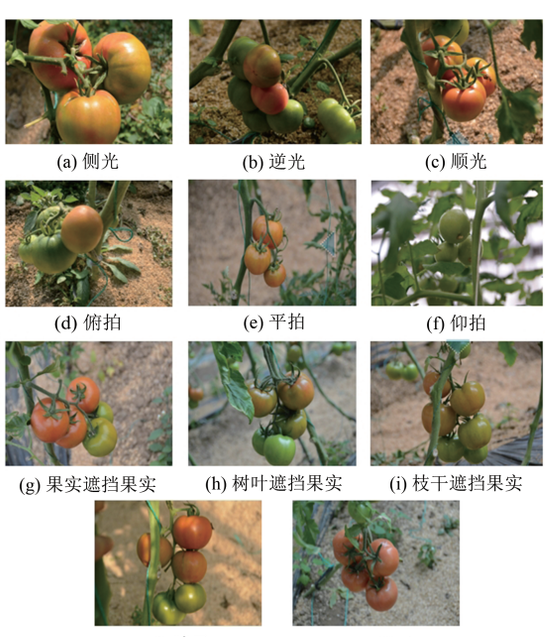

图1　复杂情况下温室番茄图像

Fig. 1　Images of greenhouse tomatoes in complex situations

1.2　数据样本预处理和图像增强

在复杂情况下的温室番茄图像中，不同天气和不同的角度下光照情况也不同，进而导致采集到的番茄果实图像颜色差异很大。此外，温室番茄果实生长的

姿态各异(不同成熟度、大小、形状和姿态等)、重叠遮挡严重等情况导致果实的形状特征难以提取。为了获取准确的数据参数,将筛选分类后的番茄图像通过标注软件LabelIMG对其数据集进行人工标注。直方图均衡化是一种增强图像对比度的方法,本文采用自适应直方图均衡化和随机调整图像饱和度(-5%~5%)的方法来减小光照强度对图像质量的影响,同时也会增加番茄与番茄、番茄与背景的对比度,提高数据集质量;通过缩放、水平/垂直翻转或旋转±90°等方法,解决番茄果实生长姿态不同的问题,改善样本集数据不均衡。此外,通过给采集到的图像添加噪声,增加数据的多样性,以提高模型对噪声的鲁棒性,为避免深度学习模型过拟合,通过将图像尺寸按一定比例缩放,降低检测过程受距离的影响。温室番茄果实图像增强如图2所示。最终扩增样本集图像至4 260幅,共计标签数为14 995个,数据集中不同成熟度的番茄果实数量如表1所示,按照标签数将图像以比例8∶1∶1将其分为训练集、测试集和验证集。

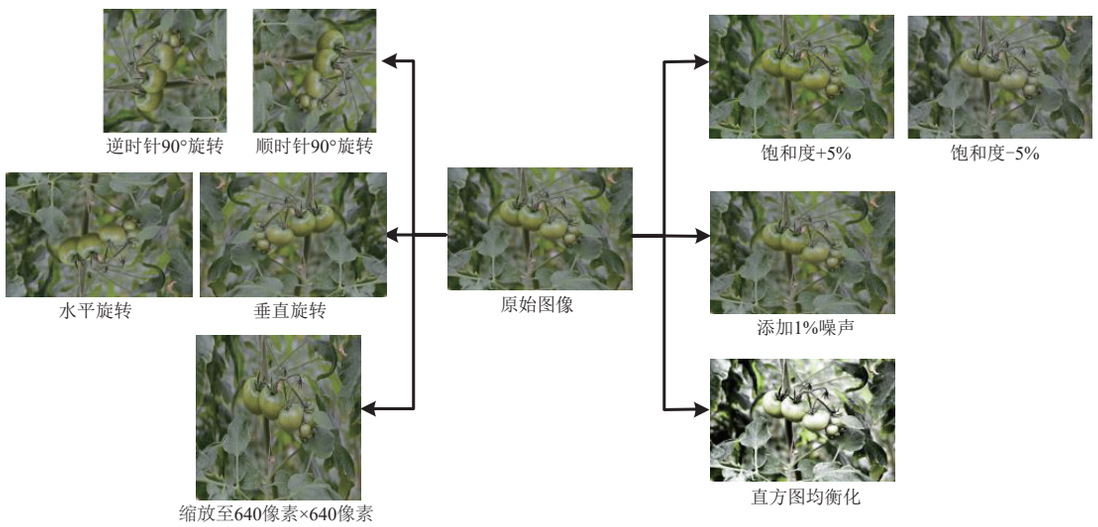

图2 温室番茄果实图像增强

Fig. 2 Greenhouse tomato fruit image enhancement

表1 不同成熟度番茄果实数量

Tab. 1 Number of tomato fruits at different maturity levels

成熟度	训练集	验证集	测试集	总计
绿色番茄	7 810	1 046	1 002	9 858
红色番茄	4 126	487	524	5 137
总计	11 936	1 533	1 526	14 995

2 理论基础

2.1 YOLO v5目标检测网络结构

YOLO v5目标检测模型由于较高的识别精度和较快的检测速度,目前已经大量运用到农业领域[24-26],随着目标检测技术的快速发展,YOLO系列也频繁更新迭代,目前已经推出了YOLO v7[27]、YOLO v8[28]等版本并进行改进,但是目前目标检测方面的文献多集中在第五代YOLO模型上,采用YOLO v5方便进行对比和改进,部署和加速也是使用YOLO v5居多,并且由于迭代较快,实质上YOLO v5之后的系列网络版本相差不大,因此,本研究使用YOLO v5模型来识别复杂环境中的温室番茄目标独特性检测,因此需在YOLO v5原模型基础上进行改进。

YOLO v5网络结构主要包括Input、Backbone、Neck和Prediction。Input部分负责图像输入,通过Mosaic数据增强、自适应锚框计算和自适应图像缩放等方法进行预处理,将图像统一缩放至标准尺寸。Backbone由Focus、CBL、CSP和SPP模块组成,主要用于特征提取。CSP模块增强了网络深度和特征提取能力,Focus模块则通过图像切片进一步提取番茄果实特征。Neck结构由FPN和PAN模块组成,增强特征学习能力,提高对语义和定位信息的处理性能,实现特征融合。在Prediction结构中,通过3个检测头分别对原始图像进行8、16、32倍下采样,生成3个不同大小的特征向量,实现温室番茄果实目标预测。

目前,YOLO v5共有5个版本,分别为YOLO v5n、YOLO v5s、YOLO v5m、YOLO v5l和YOLO v5x,不同版本具有不同的宽度和深度,适用于不同数据集。因本研究为成熟番茄和未成熟番茄两个类别,原始模型具有参数冗余,选择YOLO v5s作为基础模型,以兼顾实时检测和部署方便。因此,针对近年来表现优秀的轻量化网络SqueezeNet[29]、MobileNet系列[30-32]、EfficientNet系列[33-34]、ShuffleNet系列[35-36]、PP-LCNet[37]、GhostNet[38-39]等,进行基于YOLO v5网络特征和特征融合网络进行轻量化改进。

2.2 轻量化温室番茄目标检测方法

2.2.1 基于EfficientNet特征提取网络结构改进

EfficientNet模型采用一种称为复合缩放(Compound scaling)的扩张策略,以实现网络结构在深度、宽度和分辨率方面的均衡扩展。EfficientNet轻量化

网络共包括深度可分离卷积、反向残差结构和SENet网络[40]。采用EfficientNet系列中的基础结构EfficientNet-B0作为YOLO v5的主要特征提取网络，以达到模型轻量化优化的目的。EfficientNet-B0网络结构框架如表2所示。

表2　EfficientNet-B0网络结构框架

Tab. 2　EfficientNet-B0 network architecture framework

特征层	基本单元	输入尺寸/（像素×像素）	通道数	层数
1	Conv3×3	224×224	32	1
2	MBConv1, k3×3	112×112	16	1
3	MBConv6, k3×3	112×112	24	2
4	MBConv6, k5×5	56×56	40	2
5	MBConv6, k3×3	28×28	80	3
6	MBConv6, k5×5	14×14	112	3
7	MBConv6, k5×5	14×14	192	4
8	MBConv6, k3×3	7×7	320	1
9	Conv1×1&P&FC	7×7	1 280	1

EfficientNet的主干网络结构采用了多层反向残差结构和深度可分离卷积，网络输入经过一个标准卷积层后，逐渐加深的残差块被堆叠起来，每个残差块内部包含了SENet结构，网络的末端是全局平均池化和全连接层，用于输出最终的分类结果。以EfficientNet-B0网络为例，该网络首先会进行一个单独的3×3卷积，第一次提取输入图像的特征，该卷积层通常具有较小的卷积核和较小的步长，以保留图像的空间信息。经过初始卷积之后，图像特征会依次经过多个MBConv模块，MBConv结构如图3所示。

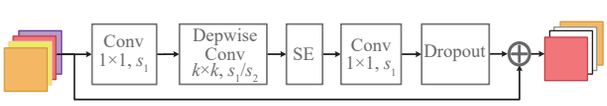

图3　MBConv结构图

Fig. 3　MBConv structure diagram

MBConv模块是EfficientNet的核心结构，它结合反向残差结构和SENet注意力机制，其主要步骤为：首先，使用1×1卷积对输入特征进行升维，提高通道数目；其次，对升维后的特征进行深度可分离卷积操作，分别执行深度卷积和逐点卷积，以提取更高层次的特征信息；然后将经过深度可分离卷积的特征传入SENet注意力模块，用于计算每个通道的权重，从而对特征进行加权处理；接着通过1×1卷积对加权后的特征进行降维操作，减少通道数目；最后将Dropout模块的输出结果和MBConv结构的输入特征进行残差连接。在经过一系列MBConv模块之后，特征会经过一个1×1卷积层进行通道数目调整，随后通过全局平均池化操作对特征进行空间压缩，得到1×1×C的特征向量。将全局平均池化后的特征向量输入到一个全连接层中，用于计算不同类别的预测概率。最后，使用Softmax激活函数将预测概率归一化，输出网络的最终结果。

2.2.2　激活函数改进

深度网络中激活函数的选择对训练动态和任务性能有显著影响[41]。为了提高模型表达能力，同时考虑模型计算速度，选择合适的激活函数尤为重要。Hardswish激活函数在MobileNetV3架构[40]中被提出，具有无上界、有下界、平滑和非单调等特点，可使神经网络层具有更丰富的表达能力，有效缓解ReLU激活函数神经元坏死的问题。本文使用Hardswish激活函数替换SiLU激活函数，原因是相较于SiLU激活函数，它用分段线性模拟代替了计算成本高的Sigmoid处理，具有数值稳定性好和计算速度快的优点，同时使模型具有丰富的表达能力。其中，Hardswish激活函数公式为

$$f(x) = x\frac{\mathrm{ReLU6}(x+3)}{6} \tag{1}$$

其中，ReLU6表示ReLU激活函数，引入ReLU6函数作为门限函数将输入值限制在[-3，3]之间，通过截断和线性缩放的操作使得Hardswish函数更具有鲁棒性，对于模型的训练和推理过程中的梯度传播更有利。

2.2.3　损失函数改进

YOLO v5原始模型中采用GIoU Loss[42]，其使用Bounding box（目标位置）的损失函数，使用二进制交叉熵和Logits损失函数计算类概率和目标得分的损失。但是，GIoU Loss存在一种问题，当预测框在目标框的内部时，GIoU损失函数的值相同，其预测框位置是不同于目标框的，损失函数收敛慢，结合温室番茄复杂的环境，不能解决相对位置关系的区分，因此这种方法无法精确地确定到效果最佳的预测框。

为解决上述问题，引入DIoU Loss[43]作为预测框偏差的偏差指标，引入YOLO v4中的DIoU损失函数（图4），其公式为

$$loss_{\mathrm{DIoU}} = 1 - DIoU \tag{2}$$

其中

$$DIoU = IoU - \frac{\rho^2(b,b^{gt})}{c^2} \tag{3}$$

式中　b、b^{gt}——预测框和真实框中心点

ρ——两个中心点之间欧氏距离

c——最小外接矩形对角线距离

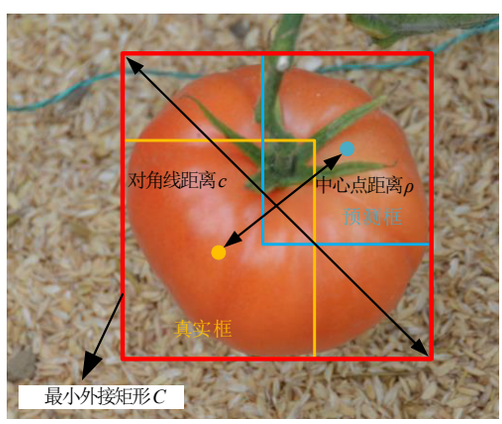

图4　DIoU损失函数示意图

Fig. 4　Schematic of DIoU loss function

IoU 损失函数计算公式为

$$IoU = \frac{\left|B \bigcap B^{gt}\right|}{\left|B \bigcup B^{gt}\right|} \tag{4}$$

式中 B——预测框 B^{gt}——真实框

2.3 轻量化模型改进

由于 EfficientNet 轻量化网络能够有效减少模型参数量和计算量，减小生成模型内存占用量，同时还不会造成太多的精度下降，因此将 YOLO v5s 的 Backbone 层中的主干网络替换为 EfficientNet 的主干网络，同时保留原本 YOLO v5s 中的 SPPF 模块，目的是能够在减少参数量和计算量的同时，对不同特征层的特征进行融合，提高模型对于特征的提取能力，从而提高模型训练精度；为提高模型表达能力并考虑计算速度，引入轻量化的 Hardswish 激活函数；通过最小化 DIoU 损失函数，直接缩短预测框与真实框的距离，提高收敛速度，从而在训练过程中更好地定位目标物体，并提高检测精度。改进轻量化 YOLO v5s 目标检测模型结构如图 5 所示，其中浅蓝色虚线框中的内容为创新点。

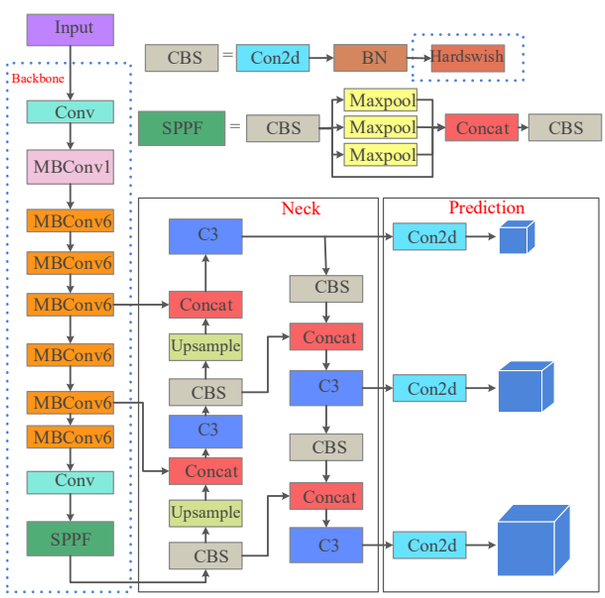

图 5 改进 YOLO v5 网络结构

Fig. 5 Improved YOLO v5 network structure diagram

3 实验结果与分析

3.1 实验平台

使用 Pytorch 深度学习框架在 Windows 10 系统中搭建网络模型并进行训练，系统运行内存大小为 128 GB，处理器型号为 Intel（R）Xeon（R）CPU E5-2620 v4 @ 2.10 GHz，显卡为 2 张 NVIDIA TITAN V，大小为 24 GB。

3.2 评价指标

采用精确率（Precision，P）、召回率（Recall，R）和平均精度均值（Mean average precision，mAP）作为番茄目标识别的评价指标，IoU 阈值为 0.5 计算得到 mAP。

3.3 改进后损失函数

EDH-YOLO 目标检测算法损失函数结果如图 6 所示，前 150 轮迭代周期中网络快速拟合，损失值明显减小，后 50 轮迭代周期损失值基本平稳，表明模型训练效果不再随着迭代次数的增加而上升，表明该 EDH-YOLO 网络模型损失函数收敛。

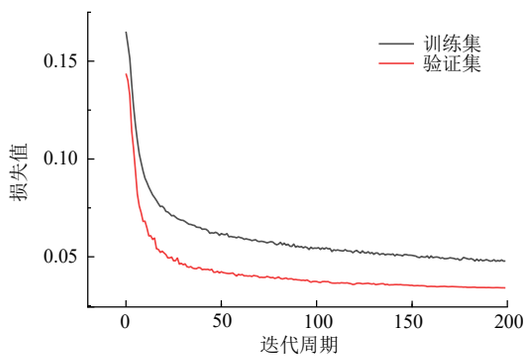

图 6 损失值曲线

Fig. 6 Loss curves

3.4 不同骨干网络检测结果对比

为进一步评估 EfficientNet-B0 的性能，选择 MobileNetV3、ShuffleNetV2、GhostNet、PP-LCNet、EfficientNet 和 EfficientNetV2 等主流网络作为 YOLO v5 算法的主干特征提取网络，并在自建的温室番茄数据集上比较这些网络的性能。选用训练后的 R、P、模型内存占用量、平均精度均值、检测速度以及浮点运算量作为评价指标，验证算法的性能，实验结果如表 3 所示。

从表 3 可以看出，YOLO v5s 网络模型准确率、召回率和平均精度均值分别为 97.6%、97.3% 和 98.7%，温室番茄果实识别效果最优，但模型内存占用量和浮点运算量最大，因此导致检测速度相较其他轻量化模型最慢。YOLO v5s-ShuffleNetV2 网络模型内存占用量最小，但其精确率、召回率以及平均精度均值明显过低，因此也非最优模型。YOLO v5s-EfficientNetV2 网络模型的性能明显优于其他轻量化主干网络，但其召回率过低，可能导致遗漏更多特征信息，因此不适用。YOLO v5s-MobileNetV3-small 网络模型也因其过低的召回率导致模型无法检测出所有的正例样本，从而无法做出正确的决策或预测。

综合表 3 分析，YOLO v5s-EfficientNet-B0 网络模型的精确率、召回率和平均精度均值分别为 95.6%、92.1% 和 96.5%。相比原模型，该模型内存占用量减少 49.0%，检测速度提升 60.0%，在保持较高识别精确率的同时兼顾精度和速度。因此，YOLO v5s-EfficientNet-B0 网络模型检测效率更高，更适合部署在移动端嵌入式设备中。

3.5 消融实验

虽然 EfficientNet-B0 网络模型在速度和精度上表现不错，但与原始模型相比，其识别效果有所下降。为弥补这种下降，从而提高其性能和精度。对 YOLO v5s 展开消融实验，消融实验结果如表 4 所示。

表3　使用不同轻量化主干网络的检测结果

Tab. 3　Detection results using different lightweight backbone networks

网络模型	*P*/%	*R*/%	平均精度均值/%	模型内存占用量/MB	浮点运算量	检测速度/(f·s^{-1})
YOLO v5s	97.6	97.3	98.7	14.3	1.58×10^{10}	33.4
YOLO v5s-MobileNetV3-small	92.2	88.8	94.4	3.0	2.50×10^{9}	62.1
YOLO v5s-MobileNetV3-large	96.1	91.2	96.2	10.2	1.02×10^{10}	35.5
YOLO v5s-ShuffleNetV2	92.6	91.6	94.5	1.9	1.80×10^{9}	79.4
YOLO v5s-GhostNet	95.3	92.5	96.5	9.9	7.60×10^{9}	37.4
YOLO v5s-PP-LCNet	95.6	90.5	95.5	6.1	5.80×10^{9}	57.2
YOLO v5s-EfficientNet-B0	95.6	92.1	96.5	7.3	7.10×10^{9}	53.3
YOLO v5s- EfficientNetV2	96.3	88.5	94.8	6.3	4.60×10^{9}	56.1

表4　消融实验

Tab. 4　Ablation experiment　%

网络模型	*P*	*R*	平均精度均值
YOLO v5s+①	95.6	92.1	96.5
YOLO v5s+①+②	95.9	92.3	96.7
YOLO v5s+①+③	96.0	90.4	96.2
YOLO v5s+①+②+③	95.9	93.1	96.8

注：①为EfficientNet-B0，②为DIoU损失函数，③Hardswish激活函数。

从表4可以看出，YOLO v5s+①+③网络模型精确率最高，为96%，但是其召回率最低，意味模型有漏检情况，会导致模型在识别真实目标时的表现不佳，从而影响最终的系统性能。YOLO v5s+①+②网络模型和YOLO v5s+①+②+③网络模型相较于YOLO v5s+①均有着更好的识别效果。但是相较于YOLO v5s+①+②网络模型，YOLO v5s+①+②+③网络模型的召回率提升0.8个百分点，在复杂环境中的温室番茄果实目标的识别效果更好。因此将YOLO v5s+①+②+③网络模型命名为EDH-YOLO网络模型。

3.6　不同算法对比实验

为了进一步验证改进后的EDH-YOLO网络模型算法性能，分别与Faster R-CNN目标检测算法以及YOLO系列算法（YOLO v7和YOLO v8）等目标检测算法进行对比实验，结果如表5所示。

表5　不同算法对比实验

Tab. 5　Comparative experiments of different algorithms

网络模型	输入尺寸/(像素×像素)	*R*/%	*P*/%	平均精度均值/%	模型内存占用量/MB	检测速度/(f·s^{-1})	浮点运算量
Faster R-CNN	600×600	87.7	98.5	98.7	108.2	11.5	1.37×10^{11}
YOLO v7	224×224	88.1	95.9	93.4	74.8	30.5	1.03×10^{11}
YOLO v8	224×224	80.0	93.4	85.1	22.5	38.4	2.84×10^{10}
EDH-YOLO	224×224	93.1	95.9	96.8	7.3	53.2	7.10×10^{9}

由表5可知，Faster R-CNN目标检测算法在精确率和平均精度均值表现优异，分别达到98.5%和98.7%。然而，其召回率只有87.7%，并且由于模型内存占用量较大和浮点运算量较高，检测速度相对较慢。YOLO v8结果最不理想，分析认为它仍处于初步发展阶段，有待进一步改进。YOLO v7召回率较低，仅为88.1%。相较而言，EDH-YOLO目标检测算法具有最高的召回率，为93.1%，且其平均精度均值比YOLO v7高出3.4个百分点。此外，EDH-YOLO目标检测算法模型内存占用量仅为7.3 MB，远低于其他3种模型，其检测速度为53.2 f/s，为4种算法最快。因此，在保持高精度的前提下，EDH-YOLO目标检测算法能够满足在温室复杂情况下番茄果实目标实时识别的要求，且整体性能优于其他算法。

对单果、遮挡情况（果实遮挡果实、枝干遮挡果实以及树叶遮挡果实）和不同光照情况（顺光、逆光以及侧光）3种情况下模型检测效果对比如图7所示，其中黄色矩形框表示漏检目标，深蓝色矩形框表示误检目标，矩形框带有“ripe_tomatoes”和“unripe_tomatoes”标签的分别表示成熟番茄的预测结果和未成熟番茄的预测结果。对于单果的检测中，4种算法均能预测果实目标，其中EDH-YOLO和Faster R-CNN模型检测精确率较高，分别为100%和96%，但YOLO v7和YOLO v8检测精确率低，分别为75%和30%；对于不同遮挡情况下，YOLO v7、YOLO v8和Faster R-CNN均存在漏检情况，且YOLO v8漏检情况严重，此外，YOLO v7存在将绿色叶片识别为绿色未成熟番茄果实误检情况；在不同光照情况下，YOLO v7、YOLO v8和Faster R-CNN虽然漏检情况不严重，但在顺光情况下存在相同的漏检情况，并且YOLO v7侧光情况下存在将绿色未成熟番茄果实识别成红色成熟番茄果实误检情况。因此，为了避免在番茄采摘过程中出现误检和漏检而造成机械臂的错误操作，同时为提高采摘机器人的整体采摘效率，EffcientNet-DHYOLO目标检测算法更加适用于番茄采摘任务。

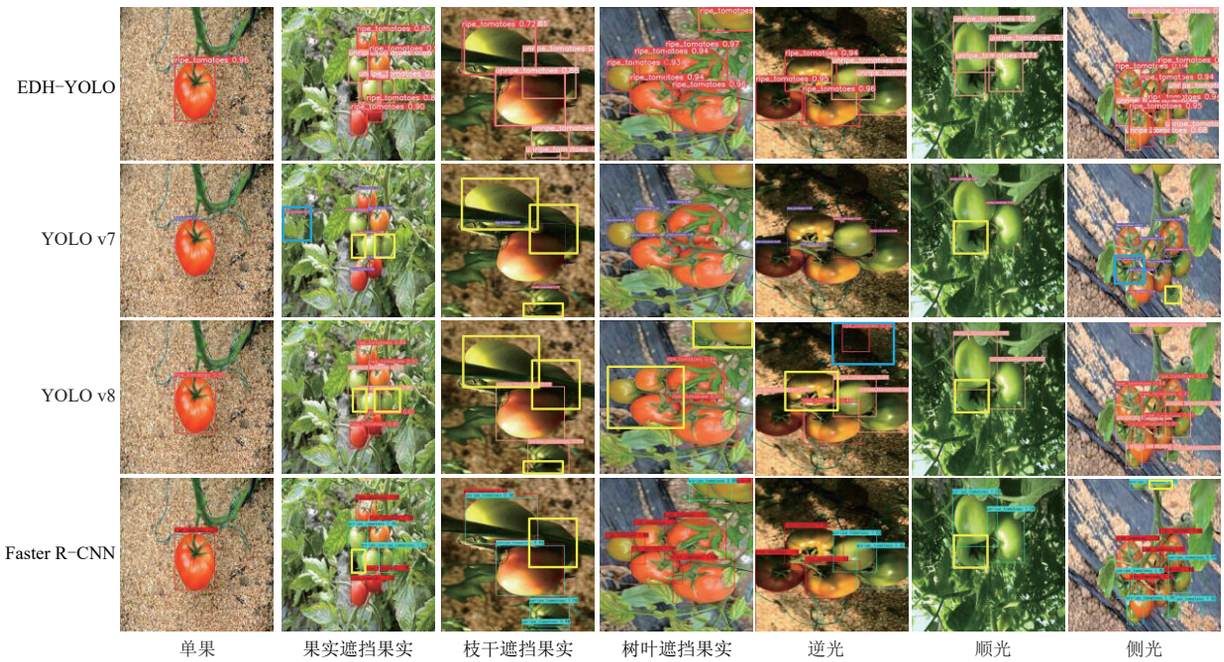

图7 不同算法对比实验效果

Fig. 7 Comparison of experiment results of different algorithms

3.7 轻量化模型部署

本文软件开发环境为：①Android Studio Dolphin 2021.3.1 on Windows 10，设定手机尺寸为6.0英寸，分辨率为1 080像素×2 340像素，安卓系统为11.0 google APIs x86（API级别为30）。②PyCharm Community Edition 2020.3.3 x64：将训练好的深度学习模型转换成ptl格式的集成开发环境（IDE）。③Pytorch框架：使用该深度学习框架对温室番茄果实目标检测模型进行训练、验证并部署到智能手机上。温室番茄果实实时识别软件的基本功能包括：①实现用户登录信息管理。②温室番茄果实目标检测模型的加载。③实现“打开相册”、“相机拍照”和“实时检测”等模块的加载。④检测结果可视化以及统计检测目标数目温室番茄果实实时识别软件整体框架如图8所示。

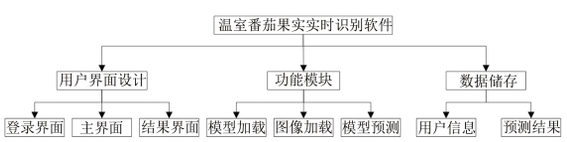

图8 温室番茄果实实时识别软件整体框架图

Fig. 8 Overall framework of real-time identification software for greenhouse tomato fruits

将上述所提出的EDH-YOLO目标检测算法，部署到Android平台的温室番茄果实实时识别软件。该软件可通过打开相册检测、拍照检测和实时检测等功能，如图9所示，利用图像或视频数据进行模型检测，从而获取温室番茄果实成熟度、位置以及数量等详细信息。



（a）相册识别效果 （b）相机拍照识别效果 （c）实时识别果实

图9 轻量化模型部署效果

Fig. 9 Lightweight model deployment effects

4 结论

（1）提出了一种基于改进YOLO v5的轻量级温室番茄果实目标检测算法（EDH-YOLO）。通过引入轻量级网络EfficientNet-B0、DIoU损失函数和轻量化的Hardswish激活函数，实现了在保持较高识别精度的同时降低计算复杂度和模型内存占用量。实验结果表明，相较于YOLO v5算法，EDH-YOLO算法减少模型内存占用量55.3%，提高模型检测速度60.0%。

（2）设计并开发了一款基于Android的温室番茄果实实时识别软件。该软件具备用户登录信息管理、相册检测、拍照检测、实时检测和统计目标数量等功能模块。软件操作简便，对用户设备环境要求较低，有较高实用性，并能满足温室复杂环境下番茄果实目标（如不同成熟度、大小、形状和姿态等）实时检测的需求。

参 考 文 献

[1] 孙永珍，贺靖，魏芳，等. “十三五”我国番茄产业发展及其国际竞争力评价[J]. 中国瓜菜，2023，36(1)：112－116.
SUN Yongzhen，HE Jing，WEI Fang，et al. Evaluation on the development and international competitiveness of China’s tomato industry during the 13th Five-Year Plan period[J]. China Cucurbits and Vegetables，2023，36(1)：112－116.(in Chinese)

[2] 王海楠，弋景刚，张秀花. 番茄采摘机器人识别与定位技术研究进展[J]. 中国农机化学报，2020，41(5)：188－196.
WANG Hainan，YI Jinggang，ZHANG Xiuhua. Research progress on recognition and localization technology of tomato picking robot[J]. Journal of Chinese Agricultural Mechanization，2020，41(5)：188－196.(in Chinese)

[3] 宋怀波，尚钰莹，何东健. 果实目标深度学习识别技术研究进展[J]. 农业机械学报，2023，54(1)：1－19.
SONG Huaibo，SHANG Yuying，HE Dongjian. Review on deep learning technology for fruit target recognition[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(1)：1－19.(in Chinese)

[4] 王丽丽. 番茄采摘机器人关键技术研究[D]. 北京：北京工业大学，2017.
WANG Lili. Research on key technology of tomato harvesting robot[D]. Beijing：Beijing University of Technology，2017.(in Chinese)

[5] 梁喜凤，章艳. 串番茄采摘点的识别方法[J]. 中国农机化学报，2016，37(11)：131－134，149.
LIANG Xifeng，ZHANG Yan. Recognition method of picking point for tomato cluster[J]. Journal of Chinese Agricultural Mechanization，2016，37(11)：131－134，149.(in Chinese)

[6] 冯青春，程伟，杨庆华，等. 基于线结构光视觉的番茄重叠果实识别定位方法研究[J]. 中国农业大学学报，2015，20(4)：100－106.
FENG Qingchun，CHENG Wei，YANG Qinghua，et al. Identification and localization of overlapping tomatoes based on linear structured light vision system[J]. Journal of China Agricultural University，2015，20(4)：100－106.(in Chinese)

[7] 马翠花，张学平，李育涛，等. 基于显著性检测与改进Hough变换方法识别未成熟番茄[J]. 农业工程学报，2016，32(14)：219－226.
MA Cuihua，ZHANG Xueping，LI Yutao，et al. Identification of immature tomatoes base on salient region detection and improved Hough transform method[J]. Transactions of the CSAE，2016，32(14)：219－226.(in Chinese)

[8] 周云成，许童羽，郑伟，等. 基于深度卷积神经网络的番茄主要器官分类识别方法[J]. 农业工程学报，2017，33(15)：219－226.
ZHOU Yuncheng，XU Tongyu，ZHENG Wei，et al. Classification and recognition approaches of tomato main organs based on DCNN [J]. Transactions of the CSAE，2017，33(15)：219－226.(in Chinese)

[9] LIU Z，WU J，FU L，et al. Improved kiwifruit detection using pre-trained VGG16 with RGB and NIR information fusion[J]. IEEE Access，2020，8：2327－2336.

[10] 王文杰，贡亮，汪韬，等. 基于多源图像融合的自然环境下番茄果实识别[J]. 农业机械学报，2021，52(9)：156－164.
WANG Wenjie，GONG Liang，WANG Tao，et al. Tomato fruit recognition based on multi-source fusion image segmentation algorithm in open environment[J]. Transactions of the Chinese Society for Agricultural Machinery，2021，52(9)：156－164.(in Chinese)

[11] LI J，LIANG X，SHEN S，et al. Scale-aware Fast R-CNN for pedestrian detection[J]. IEEE Transactions on Multimedia，2017，20(4)：985－996.

[12] 孙龙清，孙希蓓，吴雨寒，等. 基于DRN-Faster R-CNN的复杂背景多目标鱼体检测模型[J]. 农业机械学报，2021，52(增刊)：245－251，315.
SUN Longqing，SUN Xibei，WU Yuhan，et al. Multi-target fish detection model based on DRN-Faster R-CNN in complex background[J]. Transactions of the Chinese Society for Agricultural Machinery，2021，52(Supp.)：245－251，315.(in Chinese)

[13] HE K，GKIOXARI G，DOLLÁR P，et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision，2017：2961－2969.

[14] 刘芳，刘玉坤，林森，等. 基于改进型YOLO的复杂环境下番茄果实快速识别方法[J]. 农业机械学报，2020，51(6)：236－244.
LIU Fang，LIU Yukun，LIN Sen，et al. Fast recognition method for tomatoes under complex environments based on improved YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery，2020，51(6)：236－244.(in Chinese)

[15] 穆龙涛，高宗斌，崔永杰，等. 基于改进AlexNet的广域复杂环境下遮挡猕猴桃目标识别[J]. 农业机械学报，2019，50(10)：24－34.
MU Longtao，GAO Zongbin，CUI Yongjie，et al. Kiwifruit detection of far-view and occluded fruit based on improved AlexNet[J]. Transactions of the Chinese Society for Agricultural Machinery，2019，50(10)：24－34.(in Chinese)

[16] REDMON J，DIVVALA S，GIRSHICK R，et al. You only look once：unified，real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition，2016：779－788.

[17] KOLEKAR A，DALAL V. Barcode detection and classification using SSD (single shot multibox detector) deep learning algorithm[C]// Proceedings of the 3rd International Conference on Advances in Science & Technology (ICAST)，2020.

[18] 张伏，陈自均，鲍若飞，等. 基于改进型YOLOv4-LITE轻量级神经网络的密集圣女果识别[J]. 农业工程学报，2021，37(16)：270－278.
ZHANG Fu，CHEN Zijun，BAO Ruofei，et al. Recognition of dense cherry tomatoes based on improved YOLOv4-LITE lightweight neural network[J]. Transactions of the CSAE，2021，37(16)：270－278.(in Chinese)

[19] 王政，许兴时，华志新，等. 融合YOLO v5n与通道剪枝算法的轻量化奶牛发情行为识别[J]. 农业工程学报，2022，38(23)：130－140.
WANG Zheng，XU Xingshi，HUA Zhixin，et al. Lightweight recognition for the oestrus behavior of dairy cows combining YOLO v5n and channel pruning[J]. Transactions of the CSAE，2022，38(23)：130－140.(in Chinese)

[20] 李茂，肖洋轶，宗望远，等. 基于改进YOLOv8模型的轻量化板栗果实识别方法[J]. 农业工程学报，2024，40(1)：201－209.

LI Mao, XIAO Yangyi, ZONG Wangyuan, et al. Detecting chestnuts using improved lightweight YOLOv8[J]. Transactions of the CSAE, 2024, 40(1): 201 - 209. (in Chinese)

[21] WANG Z, WALSH K, KOIRALA A. Mango fruit load estimation using a video based MangoYOLO—Kalman filter—hungarian algorithm method[J]. Sensors, 2019, 19(12): 2742.

[22] ZHUANG J, LUO S, HOU C, et al. Detection of orchard citrus fruits using a monocular machine vision-based method for automatic fruit picking applications[J]. Computers and Electronics in Agriculture, 2018, 152: 64 - 73.

[23] HÄNI N, ROY P, ISLER V. Apple counting using convolutional neural networks[C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018: 2559 - 2565.

[24] 闫彬,樊攀,王美茸,等. 基于改进YOLO v5m的采摘机器人苹果采摘方式实时识别[J]. 农业机械学报, 2022, 53(9): 28 - 38,59.
YAN Bin, FAN Pan, WANG Meirong, et al. Real-time apple picking pattern recognition for picking robot based on improved YOLOv5m [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022,53(9): 28 - 38,59. (in Chinese)

[25] 范天浩,顾寄南,王文波,等. 基于改进YOLO v5s的轻量化金银花识别方法[J]. 农业工程学报, 2023, 39(11): 192 - 200.
FAN Tianhao, GU Ji'nan, WANG Wenbo, et al. Lightweight honeysuckle recognition method based on improved YOLO v5s[J]. Transactions of the CSAE, 2023, 39(11): 192 - 200. (in Chinese)

[26] 杨佳昊,左昊轩,黄祺成,等. 基于YOLO v5s的作物叶片病害检测模型轻量化方法[J]. 农业机械学报, 2023, 54(增刊1): 222 - 229.
YANG Jiahao, ZUO Haoxuan, HUANG Qicheng, et al. Lightweight method for crop leaf disease detection model based on YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp.l): 222 - 229. (in Chinese)

[27] 唐鲁婷,黄洪琼. 基于YOLOv7的轻量化水下目标检测算法[J]. 电光与控制, 2024, 31(9):92 - 97.
TANG Luting, HUANG Hongqiong. Lightweight underwater target detection algorithm based on YOLOv7[J]. Electronics Optics & Control, 2024, 31(9):92 - 97.. (in Chinese)

[28] 程换新,矫立浩,骆晓玲,等. 改进YOLOv8的遥感图像检测算法[J]. 无线电工程, 2024, 54(5):1155 - 1161.
CHENG Huanxin, JIAO Lihao, LUO Xiaoling, et al. Improved YOLOv8 for remote sensing image detection[J]. Radio Engineering, 2024, 54(5):1155 - 1161. (in Chinese)

[29] ILESANMI S, JONHBOSCO A, AHIARA W, et al. An ensemble statistical evaluation of medical image embedding with SqueezeNet neural network[C]//2022 5th Information Technology for Education and Development (ITED). IEEE, 2022: 1 - 6.

[30] JIA L, WANG Y, ZANG Y, et al. MobileNetV3 with CBAM for bamboo stick counting[J]. IEEE Access, 2022, 10: 53963 - 53971.

[31] WU Y, XIANG W. Design of intelligent sorting trash can system based on MobileNet model[C]//2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC). IEEE, 2022: 521 - 525.

[32] 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 Xiaofei, WANG Linbo, LIU Jingyan, 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. (in Chinese)

[33] WASKITA A, AMDA J, SIHONO D, et al. EfficientNetV2 based for MRI brain tumor image classification[C]//2023 International Conference on Computer, Control, Informatics and its Applications (IC3INA). IEEE, 2023: 171 - 176.

[34] AGGARWAL S, SAHOO A, BANSAL C, et al. Image classification using deep learning: a comparative study of VGG-16, InceptionV3 and EfficientNet B7 models[C]//2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE, 2023: 1728 - 1732.

[35] REZAEE K, MOUSAVIRAD S, KHOSRAVI M, et al. An autonomous UAV-assisted distance-aware crowd sensing platform using deep ShuffleNet transfer learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(7): 9404 - 9413.

[36] DONG J, YUAN J, LI L, et al. An efficient semantic segmentation method using pyramid ShuffleNet V2 with vortex pooling[C]//2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI). IEEE, 2019: 1214 - 1220.

[37] 胡阵,马宗军,黄传宝,等. PSOS-YOLO v5s:一种轻量级玉米雄穗快速检测算法[J]. 无线电工程, 2024,54(6):1446 - 1453.
HU Zhen, MA Zongjun, HUANG Chuanbao, et al. PSOS-YOLO v5s: a lightweight maize tassel rapid detection algorithm[J]. Radio Engineering, 2024,54(6):1446 - 1453. (in Chinese)

[38] WANG Z, LI T. A lightweight CNN model based on GhostNet[J]. Computational Intelligence and Neuroscience, 2022, 2022(1): 8396550.

[39] 顾文娟,魏金,阴艳超,等. 基于改进 DeepLabv3+的番茄图像多类别分割方法[J]. 农业机械学报, 2023,54(12): 261 - 271.
GU Wenjuan, WEI Jin, YIN Yanchao, et al. Multi-category segmentation method of tomato image based on improved DeepLabv3+[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(12): 261 - 271. (in Chinese)

[40] RAO B, RAO B. An effective WBC segmentation and classification using MobilenetV3 - ShufflenetV2 based deep learning framework[J]. IEEE Access, 2023, 11: 27739 - 27748.

[41] LI Y, GENG T, STEIN S, et al. GAAF: searching activation functions for binary neural networks through genetic algorithm[J]. Tsinghua Science and Technology, 2023, 28(1): 207 - 220.

[42] ZHANG Q, LIU Y, ZHANG Y, et al. Improved YOLOv3 integrating SENet and optimized GIoU loss for occluded pedestrian detection [J]. Sensors, 2023, 23(22): 9089.

[43] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993 - 13000.