doi:10.6041/j.issn.1000-1298.2023.10.022

基于改进 YOLO v7 的轻量化樱桃番茄成熟度检测方法

苗荣慧[1,2]　李志伟[1]　武锦龙[1,2]

(1.山西农业大学信息科学与工程学院,太谷 030801;2.山西农业大学农业工程学院,太谷 030801)

摘要:为在自然环境下自动准确地检测樱桃番茄果实的成熟度,实现樱桃番茄果实自动化采摘,根据成熟期樱桃番茄果实表型特征的变化以及国家标准 GH/T 1193—2021 制定了 5 级樱桃番茄果实成熟度级别(绿熟期、转色期、初熟期、中熟期和完熟期),并针对樱桃番茄相邻成熟度特征差异不明显以及果实之间相互遮挡问题,提出一种改进的轻量化 YOLO v7 模型的樱桃番茄果实成熟度检测方法。该方法将 MobileNetV3 引入 YOLO v7 模型中作为骨干特征提取网络,以减少网络的参数量,同时在特征融合网络中加入全局注意力机制(Global attention mechanism,GAM)模块以提高网络的特征表达能力。试验结果表明,改进的 YOLO v7 模型在测试集下的精确率、召回率和平均精度均值分别为 98.6%、98.1% 和 98.2%,单幅图像平均检测时间为 82 ms,模型内存占用量为 66.5 MB。对比 Faster R-CNN、YOLO v3、YOLO v5s 和 YOLO v7 模型,平均精度均值分别提升 18.7、0.2、0.3、0.1 个百分点,模型内存占用量也最少。研究表明改进的 YOLO v7 模型能够为樱桃番茄果实的自动化采摘提供技术支撑。

关键词:樱桃番茄;成熟度检测;YOLO v7;轻量化;MobileNetV3;全局注意力机制

中图分类号:S252　　**文献标识码**:A　　**文章编号**:1000-1298(2023)10-0225-09　　**OSID**:

Lightweight Maturity Detection of Cherry Tomato Based on Improved YOLO v7

MIAO Ronghui[1,2]　LI Zhiwei[1]　WU Jinlong[1,2]

(1. *College of Information Science and Engineering*, *Shanxi Agricultural University*, *Taigu* 030801, *China*

2. *College of Agricultural Engineering*, *Shanxi Agricultural University*, *Taigu* 030801, *China*)

Abstract: Automatic and accurate detection of cherry tomato maturity in natural environment is the foundation for achieving automatic cherry tomato picking. According to the changes in phenotypic characteristics of cherry tomato during its mature period and the national standard GH/T 1193—2021, and regarding the lack of significant differences in adjacent maturity characteristics of cherry tomatoes and the problem of mutual occlusion between fruits, a lightweight maturity detection method of cherry tomato with five levels, including green, turning, pink, lightred and red was proposed based on improved YOLO v7 model. In this model, MobileNetV3 was introduced into the original YOLO v7 model as backbone for feature extraction to reduce the parameters of the original model; global attention mechanism (GAM) module was added to the feature fusion network to improve the feature expression ability of the model. The experimental results showed that the precision, recall and mean average precision of the improved model were 98.6%, 98.1% and 98.2%, respectively, the average detection time of a single image was 82 ms, and the memory occupied by the model was 66.5 MB. Compared with Faster R-CNN, YOLO v3, YOLO v5s and YOLO v7 models, the mean average precision (mAP) was improved by 18.7, 0.2, 0.3 and 0.1 percentage points, respectively. The average accuracy of the improved YOLO v7 model was also improved, and memory usage of the model was optimal. The results showed that the improved YOLO v7 model can provide effective exploration for automated cherry tomato fruit picking.

Key words: cherry tomato; maturity detection; YOLO v7; lightweight; MobileNetV3; global attention mechanism

收稿日期:2023-07-05　修回日期:2023-07-25

基金项目:山西省基础研究计划项目(202203021212414、202203021212428)和山西农业大学青年科技创新基金项目(J142102257)

作者简介:苗荣慧(1990—),女,讲师,博士生,主要从事图像分析与机器视觉研究,E-mail:ronghui092@163.com

通信作者:李志伟(1969—),男,教授,博士生导师,主要从事精细农业装备及计算机控制研究,E-mail:lizhiweitong@163.com

0 引言

视觉识别系统作为果蔬采摘机器人系统的组成部分,在果蔬目标识别定位、自动采摘和果蔬估产等方面具有至关重要的作用[1]。樱桃番茄因其营养价值高、风味独特而被广泛种植[2]。由于樱桃番茄果实高度密集、果实小,且叶片遮挡严重、果蔬高度不一,采收成为最费时耗力的工作。目前,樱桃番茄的采摘主要依靠人工,成本高、效率低且不能保证采摘的果实处于最佳采摘时期。对果实成熟期的判定直接决定果实的运输和贮藏方式,对果实的价格及风味也会造成极大的影响[3]。因此,研究具有高识别率的成熟度自动检测系统对于确定不同成熟度果实的分布区域,实现樱桃番茄的自动化采摘具有重要意义。

近年来,国内外学者尝试采用机器视觉及光谱分析方法实现苹果[4]、香蕉[5]、芒果[6]以及柿子[7]等果实成熟度的检测。赵娟等[8]针对富士苹果的3种成熟度建立了极限学习机和支持向量回归分类模型,实验结果表明可见/近红外光谱结合成熟度评价指数可实现苹果成熟度分类。周文静等[9]提出一种基于改进圆形 Hough 变换的田间红提葡萄果穗成熟度判别方法,该研究成果可为果农适宜期收获葡萄及自动化采摘提供重要指导。薛建新等[10]利用高光谱成像技术对4种不同成熟阶段的沙金杏成熟度进行了判别研究。然而,以上方法大多采用人工方式进行特征提取,算法的特征表达能力有限,泛化能力也不高,难以胜任自然条件下果实成熟度的检测工作。

深度学习技术在目标检测领域有巨大的优势,由于其对目标高维特征的高度提取,使复杂情况下不同成熟度的果实识别成为可能。陈锋军等[11]提出一种改进的 EfficientDet 的油橄榄果实成熟度检测方法,改进的模型不仅具有良好的检测精度,而且模型占用内存较小,能够胜任自然条件下推广应用的实际需求。龙洁花等[12]提出改进 Mask R-CNN 的温室环境下不同成熟度番茄果实的分割方法,该方法对绿熟期、半熟期、成熟期番茄果实分割的平均精度均值都较高,可为番茄采摘机器人精准作业提供依据。TIAN 等[13]提出了一种改进的 YOLO v3 模型,用于在不同光照条件、复杂背景以及遮挡等情况下检测不同成熟度的苹果。熊俊涛等[14]提出一种轻量化 YOLO v5-Lite 模型对自然环境下木瓜成熟度进行检测,该方法对不同拍摄距离、不同遮挡情况以及不同光照情况下的果实均能实现准确识别。可见卷积神经网络在果实成熟度检测研究中优势显著,其中 YOLO v7 模型具有检测精度高、速度快、支持高分辨率图像、支持多种类型目标等优点,已被应用到行人检测[15]、工业产品缺陷检测[16]等领域。目前,在每秒传输 5~160 帧图像范围内,YOLO v7 模型的速度和精度都超过了所有已知的目标检测模型。因此,选取 YOLO v7 模型为基础模型进行樱桃番茄成熟度的检测。

目前,基于卷积神经网络的水果成熟度检测方法,存在如下问题:大多数检测精度高的方法,模型计算复杂度高,检测速度慢;而模型计算复杂度低,检测速度快的方法,检测精度却又不高。对于平衡好检测速度、检测精度和模型计算复杂度的樱桃番茄检测方法,当前的研究还较少。樱桃番茄果实具有如下特性:每个花序结果少则一到两个,多则十几个,果实之间往往存在遮挡问题;樱桃番茄果实在生长发育过程中易受光照、环境等因素的影响,同一花序所结果实成熟度不一致且相邻成熟度特征差异不明显。这些都会导致 YOLO v7 模型对相邻成熟度及遮挡果实的检测能力较差。针对以上问题,本文提出一种改进的 YOLO v7 轻量化樱桃番茄成熟度检测模型。首先,在数据采集时对具有不同成熟度的部分图像进行跟踪采集,对标注数据进行修订以提高模型对相邻成熟度果实的检测能力。然后,将 YOLO v7 模型的骨干特征提取网络替换为轻量级的网络 MobileNetV3,大幅降低模型的参数量和计算量。在特征融合模块引入 GAM 注意力模块来加强网络的特征表达能力,降低因模型参数量少而造成检测精度下降的影响,提高相邻成熟度果实和被遮挡果实的检测精度。最后,通过试验验证改进的 YOLO v7 模型对自然环境下樱桃番茄果实成熟度的检测效果以及确定成熟果实的分布区域,为樱桃番茄果实的自动化采摘提供参考依据。

1 材料与方法

1.1 图像采集

试验研究区域选取山西省晋中市太谷区范村镇象谷村番茄种植基地,该基地主要栽培品种为红美玉、黑珍珠、迷彩、红珍珠、黄珍珠等。本文以产量最高的红珍珠为研究对象,于 2023 年 3—5 月使用 Oppo Reno5 型手机采集温室条件下的不同植株及不同成熟度的樱桃番茄图像。采集时间包括上午、中午和下午,以获取不同光照条件下的图像,共获取 3 468 像素 ×4 624 像素的樱桃番茄图像 1 560 幅,其中,部分图像采用跟踪拍摄方式对樱桃番茄成熟过程进行采集。采集图像类型包括单目标图像、多目标图像、顺光图像、逆光图像、有枝叶遮挡图像、无枝

叶遮挡图像和重果图像等，如图 1 所示。



(a) 单目标+无遮挡　(b) 多目标+重果+有枝叶遮挡　(c) 顺光+有枝叶遮挡　(d) 逆光+重果

图 1　樱桃番茄样本图像

Fig. 1　Sample images of cherry tomato

1.2　果实成熟度等级划分

国家标准 GH/T 1193—2021[17]将番茄成熟度等级根据果皮和果肉颜色变化划分为 6 级，即未熟期、绿熟期、变色期、红熟前期、红熟中期和红熟后期。而本文研究的重点是无损检测挂枝状态下樱桃番茄的成熟度，与直接采摘果实通过剖面获取果肉颜色判断成熟度相比，挂枝状态无法获取果肉颜色变化；与普通番茄相比，樱桃番茄因其果实尺寸小、生长发育周期短，导致相邻成熟度差异较小。因此，本文根据樱桃番茄实际生长发育情况及果皮颜色变化将 6 级成熟度（表 1）合并为 5 级成熟度。其中，6 级成熟度的第 1 级为未熟期，该时期果实尚未充分生长发育定形，催熟困难不宜采摘，将第 1 级删除；6 级成熟度的 2 ~ 6 级与 5 级成熟度的 1 ~ 4 级成熟度完全一致，都是根据果皮颜色的变化进行划分。因此，将 6 级成熟度的 2 ~ 6 级根据果皮颜色定义为 5 级，分别为绿熟期、转色期、初熟期、中熟期和完熟期，如表 2 所示。从表 2 可知，不同成熟度的番茄会直接影响其加工、贮藏及运输方式，采收成熟度因用途的不同而有所差别，适时采摘具有重要意义。

表 1　番茄 6 级成熟度描述

Tab. 1　Maturity description of tomato at six levels

级别	成熟度	描述	采摘情况
1	未熟期	果实及种子尚未充分生长发育定形，果皮绿色，无光泽	催熟困难，不宜采摘贮藏
2	绿熟期	果实定形，果面有光泽，由绿色变为白绿色，种子已长大，周围呈胶状	可人工催熟，采摘贮藏或远距离运输
3	变色期	由绿熟到红熟的过渡期，果脐周围开始出现黄色或淡红色晕斑，果实着红面不到 10%	可采摘贮藏，短距离运输
4	红熟前期	一至三成红熟，果实着红面积占比 10% ~30%	可采摘，隔 2 日上市鲜销
5	红熟中期	四至六成红熟，果实着红面积占比 30% ~60%	可采摘，隔 1 日上市鲜销
6	红熟后期	七至十成红熟，果实着红面积占比 60% ~100%	可采摘，当日上市鲜销

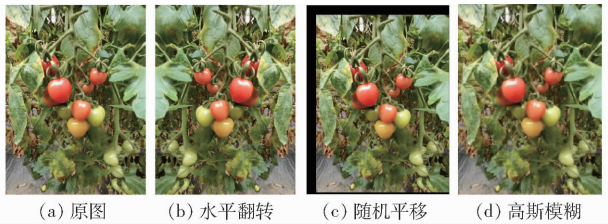
表 2　樱桃番茄 5 级成熟度描述

Tab. 2　Maturity description of cherry tomato at five levels

级别	成熟度	RGB 图像	描述	采摘情况
1	绿熟期		果实大小定形，果面为绿色	可人工催熟，采摘贮藏或远距离运输
2	转色期		果脐周围开始出现黄色或淡红色晕斑，果实着红面积占比为 10% ~30%	可采摘贮藏，短距离运输
3	初熟期		果实着红面积占比 30% ~60%	可采摘，隔 2 日上市鲜销
4	中熟期		果实着红面积占比 60% ~90%	可采摘，隔 1 日上市鲜销
5	完熟期		果实表面 90% 以上为红色	可采摘，当日上市鲜销

1.3　数据集构建

由于樱桃番茄数据集的数据样本量不足以让模型在训练时达到收敛状态，为了改善网络训练模型效果和提高模型泛化能力，采用数据增强方法增加样本数量，防止网络因训练样本不足导致过拟合问题。本研究采用高斯模糊、水平翻转、垂直翻转、非等比例缩放、随机平移、透视变换和随机裁剪等数据增强方式随机组合，对数据集进行扩充，部分数据增强样本如图 2 所示。通过数据增强共获得 8 745 幅样本图像，按照比例 7∶1∶2将数据集划分为训练集（6 121 幅）、验证集（875 幅）、测试集（1 749 幅）。



(a) 原图　(b) 水平翻转　(c) 随机平移　(d) 高斯模糊

图 2　樱桃番茄数据增强图像

Fig. 2　Data enhanced images of cherry tomato

本研究训练的模型为 YOLO 系列格式的数据集，采用 LabelImg 软件对图像中不同成熟度樱桃番茄果实进行标注，格式为 .txt 文档，标注规则为：①对图像中果实相互遮挡但不影响人工判断成熟度等级的果实分别进行标注。②对果实遮挡严重，人工难以辨别成熟度等级的果实不予标注。③对于未长大的绿色樱桃番茄不予标注。由于樱桃番茄同一

花序所结果实成熟度不一致，使得有的图像涵盖多个成熟度级别。数据集中不同成熟度樱桃番茄果实数量如表 3 所示。

表 3 数据集中不同成熟度樱桃番茄标注数量

Tab. 3 Annotation number of cherry tomato with different maturities

成熟度	训练集	验证集	测试集	总计
绿熟期	7 125	768	1 215	9 108
转色期	6 828	528	869	8 225
初熟期	4 254	617	998	5 869
中熟期	8 795	1 288	4 387	14 470
完熟期	9 123	1 376	3 454	13 953
总计	36 125	4 577	10 923	51 625

1.4 樱桃番茄成熟度检测模型

樱桃番茄在生长发育过程中易受光照和环境因素的影响，同一花序所结果实成熟度存在差异，本研究构建的数据集中多数图像包含多个成熟度级别，且相邻成熟度特征差异不明显，同时存在枝叶遮挡及果实遮挡问题，导致准确检测果实的成熟度较困难。现有基于深度学习的卷积神经网络模型虽然检测精度高，但模型计算复杂度高、检测速度慢。因此，为了平衡好检测速度、检测精度和模型计算复杂度，同时，有效解决影响相邻成熟度和被遮挡果实的检测问题，本研究对 YOLO v7 模型进行了改进。

首先，为了达到模型在有较高检测精度和检测速度的同时，模型参数尽可能少的目的，将 YOLO v7 的骨干特征提取网络替换为轻量级 MobileNetV3 网络。为了提高相邻成熟度的检测能力，在特征融合网络中加入 GAM 注意力模块以提高网络的表达能力，改进后的 YOLO v7 轻量化樱桃番茄模型结构如图 3 所示。

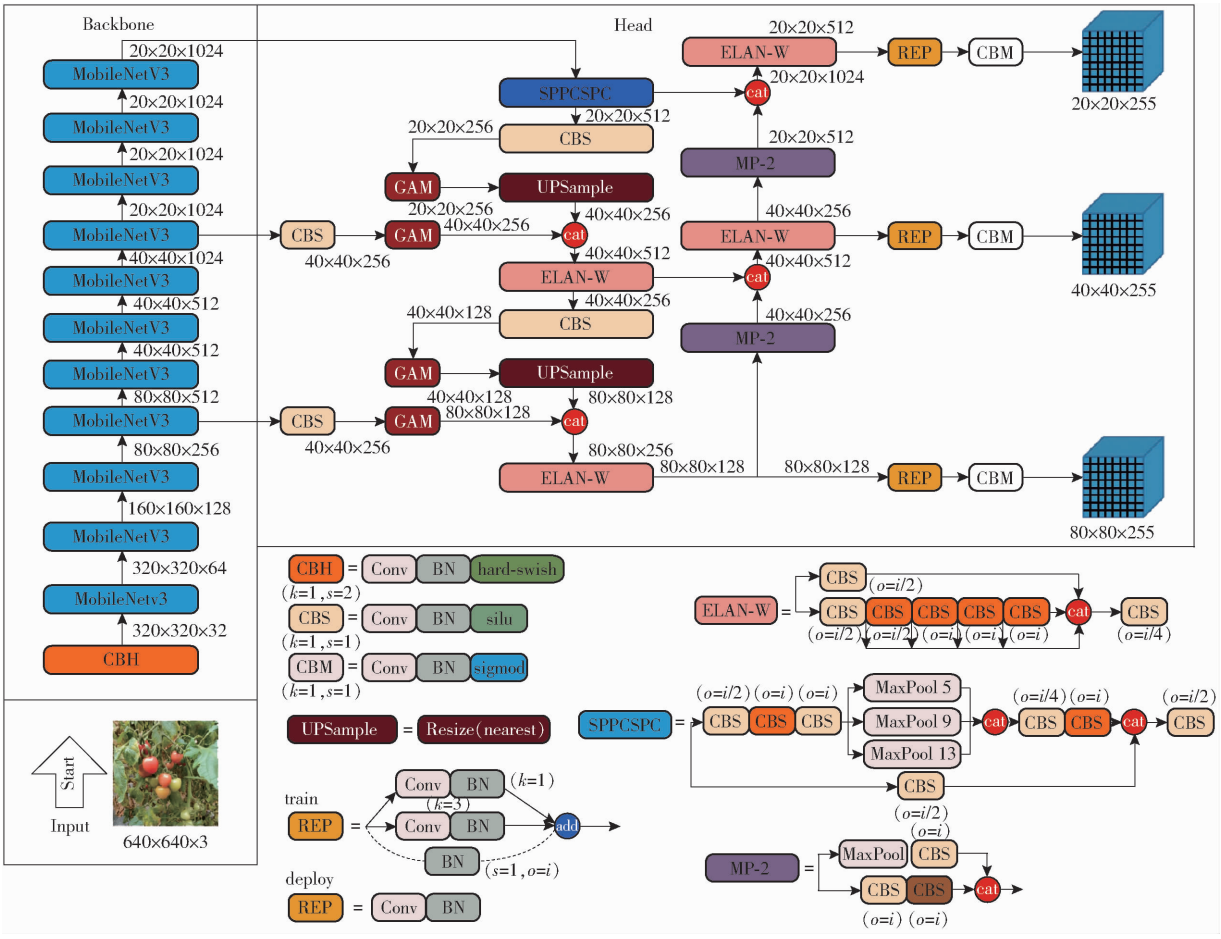


图 3 改进的 YOLO v7 轻量化网络模型

Fig. 3 Lightweight network model of improved YOLO v7

1.4.1 MobileNetV3 模型

MobileNet 网络是由 Google 团队提出的一种轻量级神经网络模型，使用深度可分离卷积(Depthwise separable convolution)和反向残差结构的思想构建模型[18]。与传统卷积神经网络相比，MobileNet 在准确率小幅度降低的前提下大大减少了模型参数量与运算量[19]。MobileNetV3 综合了 MobileNetV2 和 MobileNetV1 的特点，在 MobileNetV2 网络的基础上加入了 SE(Squeeze and excitation networks)模块[20]，其网络结构如图 4 和表 4 所示。

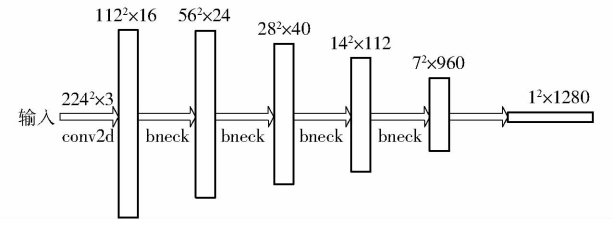


图 4　MobileNetV3 网络结构

Fig. 4　Network structure of MobileNetV3

表 4　MobileNetV3 网络参数

Tab. 4　Network parameters of MobileNetV3

输入	操作	升维通道数	输出通道数	SE	NL	步长
$224^2\times3$	conv2d		16	×	HS	2
$112^2\times16$	bneck,3×3	16	16	√	RE	1
$112^2\times16$	bneck,3×3	64	24	×	RE	2
$54^2\times24$	bneck,3×3	72	24	×	RE	1
$54^2\times24$	bneck,5×5	72	40	√	RE	2
$28^2\times40$	bneck,5×5	120	40	√	RE	1
$28^2\times40$	bneck,5×5	120	40	√	RE	1
$28^2\times40$	bneck,3×3	240	80	×	HS	2
$14^2\times80$	bneck,3×3	200	80	×	HS	1
$14^2\times80$	bneck,3×3	184	80	×	HS	1
$14^2\times80$	bneck,3×3	184	80	×	HS	1
$14^2\times80$	bneck,3×3	480	112	√	HS	1
$14^2\times112$	bneck,3×3	672	112	√	HS	1
$14^2\times112$	bneck,5×5	672	160	√	HS	2
$7^2\times160$	bneck,5×5	960	160	√	HS	1
$7^2\times160$	bneck,5×5	960	160	√	HS	1
$7^2\times160$	conv2d,1×1		960	×	HS	1
$7^2\times960$	pool,7×7			×		1
$1^2\times960$	conv2d,1×1		1 280	×	HS	1
$1^2\times1\,280$	conv2d,1×1		n	×		1

注：“√”表示本层引入了 SE 模块，“×”表示本层未引入 SE 模块，n 表示分类类别数。

表 4 中，第 1 列为每个特征层输入向量的尺寸；第 2 列为每个特征层所进行的操作；第 3 和第 4 列分别代表了 Bottleneck 结构中逆残差结构升维后的通道数和 Bottleneck 输出的通道数；SE 代表本层是否引入注意力机制 SE 模块；NL 代表激活函数的类型，其中 HS 代表 hard-swish，RE 代表 ReLU；最后一列代表 Bottleneck 结构中卷积的步长。conv2d 表示普通卷积，bneck（The inverted residual with linear bottleneck block）表示由线性瓶颈的逆残差结构组成的特征提取块，pool 表示池化层。MobileNetV3 网络操作过程：首先利用 1×1 的卷积进行升维，再引入 MobileNetV2 线性瓶颈的逆残差结构；然后进行 3×3 深度可分离卷积，降低网络的计算量；再通过注意力模型 SE－Net，使网络关注更加有用的通道信息来调整每个通道的权重；最后，使用 hard-swish 激活函数代替 swish 函数，减少运算量以提高模型性能。该网络能够在保证模型轻量化的基础上，提高模型精度及运行速度，因此，本文使用 MobileNetV3 网络作为 YOLO v7 模型的骨干网络实现特征提取。

线性瓶颈的逆残差结构[21]是一种可以有效减小计算量并且防止细节特征消失的特征提取结构，图 5 所示是具有瓶颈结构的深度可分离卷积模块。特征首先会在卷积块中通过 1×1 的卷积进行维度扩张，再使用 3×3 的卷积进行深度可分离卷积，同时模块在深度可分离卷积的同时融入了轻量级注意力机制将特征图进行全局平均池化（Global average pooling，GAP），并经过 ReLU 和 hard-swish 激活函数后和特征图相乘，调整特征图的每个通道的权重。（图 5 中 Dwise 3×3 表示 3×3 卷积核的深度可分离卷积，BN 表示 Batch Normalization，GAP 表示全局平均池化，Conv 1×1 表示 1×1 的卷积）。

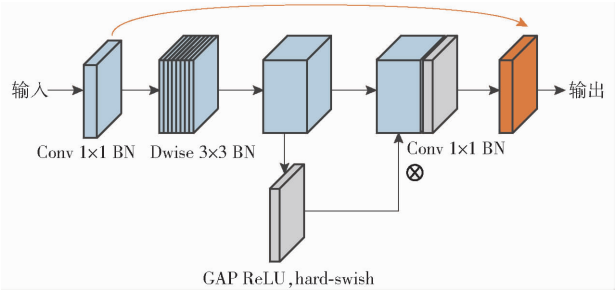


图 5　线性瓶颈的逆残差结构块

Fig. 5　Inverted residual with linear bottleneck block

1.4.2　加入 GAM 注意力模块的特征融合网络

卷积块注意力模块（Convolutional block attention module，CBAM）是一种结合了空间注意力模块和通道注意力的卷积注意力模块[22]，通过学习可以自动获取每个特征空间和特征通道的重要程度，并根据重要程度分配不同的权重来加强重要特征的提取，抑制其他无关的信息。但由于信息减少及维度分离，该机制只能利用有限的感受野实现视觉表示，缺乏空间和通道的交互作用。GAM 是一种可以减少信息缩减并放大全局维度交互特征的注意力机制[23]，它采用 CBAM 中的顺序通道-空间注意机制，并对子模块进行了重新设计，其模型结构如图 6 所示。

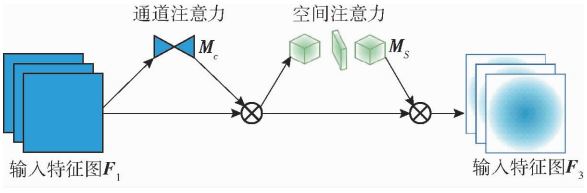


图 6　GAM 注意力模块

Fig. 6　GAM attention module

GAM 能够稳定地提高不同架构和深度的卷积神

经网络的性能,具有良好的数据拓展能力和鲁棒性。

$$\boldsymbol{F}_2 = \boldsymbol{M}_c(\boldsymbol{F}_1) \otimes \boldsymbol{F}_1 \tag{1}$$

$$\boldsymbol{F}_3 = \boldsymbol{M}_s(\boldsymbol{F}_2) \otimes \boldsymbol{F}_2 \tag{2}$$

其中:$\boldsymbol{F}_1 \in \mathbf{R}^{C\times H\times W}$为输入特征图;$\boldsymbol{M}_c$ 和 $\boldsymbol{M}_s$ 分别为通道和空间注意特征图;$\otimes$为矩阵点乘。

通道注意力子模块采用 3D 组合保留跨三维度信息,然后采用多层感知机(Multilayer perceptron, MLP)让特征向量在不同维度间充分交互,使得模型能够获取更多非线性特征和组合特征信息,如图 7 所示,图中 C、H、W 分别为特征图的通道数、高和宽。其中 MLP 代表多层感知机(全连接层);Sigmoid 为激活函数。

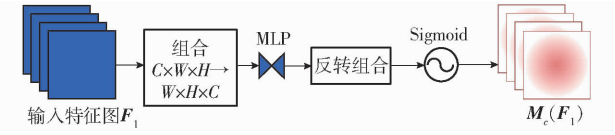


图 7 GAM 通道注意力子模块

Fig. 7 Channel attention submodule of GAM

为了更多关注空间信息,在空间注意力子模块中采用 2 个卷积层对空间信息进行融合。由于最大池化减少了信息量且对模型具有消极影响,该模块中删除了池化过程来更好地保留特征图,如图 8 所示。

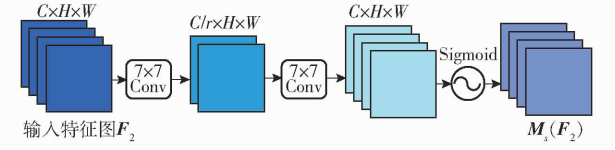


图 8 GAM 空间注意力子模块

Fig. 8 Spatial attention submodule of GAM

1.5 试验环境配置与网络参数设置

试验所用操作系统为 Windows 10,CPU 型号为 Intel(R) Core(TM) i7 - 12700F @ 2.10 GHz,GPU 型号为 NVIDIA GeForce RTX 3080,运行内存为 32 GB,1 TB 机械硬盘,编程语言为 Python 3.9,使用的深度学习框架为 Pytorch 1.13.0,GPU 加速库为 CUDA 11.7、CUDNN 8.4.1。为缩短网络训练的时间,采用冻结训练的方法,冻结骨干网络训练时初始学习率设为 0.001,批处理量设为 16,动量设置为 0.9;解冻之后的网络训练学习率设为 0.000 1,批大小设为 8,动量设置为 0.9,迭代次数为 300 次。

1.6 模型评价指标

对樱桃番茄果实成熟度的检测过程需要考虑检测精度和速度。在模型检测精度方面,选用精确率(Precision, P)、召回率(Recall, R)和平均精度(Average precision, AP)作为评价指标;在模型检测性能方面,选取平均精度均值(mAP)、平均检测时间(Average detection time, ADT)和模型内存占用量作为评价指标。

2 实验结果与分析

2.1 改进 YOLO v7 模型试验

为验证改进 YOLO v7 模型的性能,对测试集中的 1 749 幅樱桃番茄图像进行测试评价,表 5 为本文算法在不同成熟度级别的检测结果。从表 5 中可知,本文算法的平均精度均值可以达到 98.2%,精确率为 98.6%,召回率为 98.1%,所有成熟度级别检测结果均在 96% 以上。部分检测示例如图 9 所示,从图 9 中可以看出,本文算法能够准确检测出不同成熟度的樱桃番茄,对于图 9a 中掉落在地上的小目标以及距离较远的小目标本文识别结果也较好。由于果实成熟时间不一致,图像包含多个成熟度级别,对于刚进入初熟期的果实与进入中熟期的果实特征差异较小,极易发生误判,改进的算法可以提取果实表型形状的细粒度特征实现不同成熟度的准确检测。图 9 中樱桃番茄果实可能受枝叶以及其他果实的遮挡,改进的 YOLO v7 模型可进一步融合遮挡果实的位置信息以及语义信息,提高对遮挡果实成熟度的检测能力。综上所述,改进的 YOLO v7 模型能够准确地检测果实的成熟度,对小目标、多目标、枝叶遮挡、重果、光照影响等都有较好的检测效果。

表 5 本文算法不同成熟度级别检测结果

Tab. 5 Detection results with different maturities of proposed algorithm %

成熟度级别	精确率	召回率	平均精度均值
绿熟期	99.2	99.6	99.2
转色期	99.8	96.4	99.0
初熟期	96.1	98.0	96.2
中熟期	98.9	97.0	97.9
完熟期	99.0	99.6	98.8
平均值	98.6	98.1	98.2

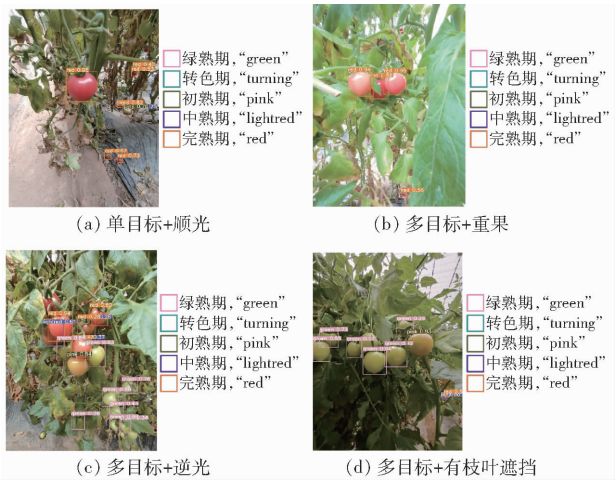


图 9 改进后的 YOLO v7 模型检测结果

Fig. 9 Detection results of improved YOLO v7 model

2.2 改进 YOLO v7 模型消融试验

为了进一步验证改进 YOLO v7 模型的性能,本

研究设置了消融试验验证 4 组网络的性能，分别为在 YOLO v7 特征融合中加入 GAM 模块、替换 YOLO v7 骨干网络为 MobileNetV3、YOLO v7 和本文算法。从定量的角度分析 4 组网络的性能，对测试集中的 1 749 幅樱桃番茄图像进行客观评价，评价指标包括模型检测精度、平均检测时间等，评价对比结果如表 6 所示。从表 6 可知，改进后模型平均精度均高于其他 3 组模型。

表 6　改进的 YOLO v7 模型消融试验结果

Tab. 6　Ablation test results of improved YOLO v7 model

模型	AP/%					mAP/%	ADT/ms
	绿熟期	转色期	初熟期	中熟期	完熟期		
YOLO v7 + GAM	53.4	78.5	89.6	75.4	68.9	71.8	137
YOLO v7 + MobileNetV3	58.8	90.6	64.2	93.3	64.3	74.2	245
YOLO v7	99.2	1.00	97.0	96.8	98.7	98.1	93
改进的 YOLO v7	99.2	99.0	96.2	97.9	98.8	98.2	82

2.3　不同目标检测网络对比结果分析

为定性评价改进 YOLO v7 模型的检测结果，将改进的模型与 Faster R－CNN[24]、YOLO v3[25]模型、YOLO v5s[26]模型以及原始 YOLO v7 模型对测试集中的樱桃番茄图像进行算法对比，算法对比结果如图 10 所示。

从图 10b 可知，该算法在单果目标图像中能够识别小目标，但部分结果存在重复识别情况；在多目标图像中，存在重复识别情况，算法将图像中部分树叶误识别为果实，对重果的识别效果也差，并且存在漏检情况。从图 10c 可知，该算法在单果目标图像中能够较好地识别小目标、远距离目标；在多目标图像中，重果识别效果较差，部分图中将树叶识别为果实，存在误识别，也存在漏检情况。从图 10d 可知，该算法在单果目标图像中能够较好地识别小目标；在多目标图像中，针对重果、枝叶遮挡、多目标等情况识别效果都较好。从图 10e 可知，该算法在单果目标图像中识别小目标情况较差；在多目标图像中，绿色多目标和重果识别效果较差，部分图中将树叶识别为果实，存在误识别，也存在漏检情况。从图 10f 可知，该算法在单果目标图像中能够较好地识别小目标、远距离目标；在多目标图像中，针对重果、枝叶遮挡、多目标等情况识别效果都较好。

为定量比较改进模型的性能，将改进模型与 Faster R－CNN、YOLO v3 模型、YOLO v5s 模型以及原始 YOLO v7 模型在测试图像上进行性能比较，如



(a) 原图　(b) Faster R-CNN　(c) YOLO v3　(d) YOLO v5s　(e) YOLO v7　(f) 改进的YOLO v7

图 10　不同目标检测网络对比结果

Fig. 10　Comparison results of different detection networks

表 7 所示。从表 7 中可知,与其他模型相比,改进的 YOLO v7 樱桃番茄成熟度检测模型在精确率、召回率、平均精度指标都是最优的,分别为 98.6%、98.1%、98.2%,其平均精度均值 mAP 与 Faster R-CNN、YOLO v3 模型、YOLO v5s 模型以及原始 YOLO v7 模型相比分别提升 18.7、0.2、0.3、0.1 个百分点。本文模型在原始 YOLO v7 模型的基础上进行了轻量化改进,改进之后的模型内存占用量也是最少的。

表 7 不同检测模型在测试集中的性能结果

Tab.7 Performance of different detection models in test sets

模型	精确率/%	召回率/%	mAP/%	ADT/ms	模型内存占用量/MB
Faster R-CNN	74.2	92.8	79.5	356	91.19
YOLO v3	96.0	96.5	98.0	89	89.15
YOLO v5s	96.9	96.0	97.9	98	90.15
YOLO v7	98.3	97.8	98.1	93	122.20
改进的 YOLO v7	98.6	98.1	98.2	82	66.50

3 结论

(1)提出了一种改进的轻量化 YOLO v7 网络模型用于樱桃番茄成熟度检测。该网络将 YOLO v7 原始模型中骨干网络替换为轻量型 MobileNetV3 结构并在特征融合网络中引入 GAM 注意力机制,不仅能够充分并有效利用全局维度的交互特征,还使得网络能够平衡运算速度和模型复杂度的问题。

(2)为了验证改进 YOLO v7 模型的性能,本文设置了 4 组网络进行消融试验验证并对试验结果进行了定量分析,4 组网络分别为在 YOLO v7 特征融合中加入 GAM 模块、替换 YOLO v7 骨干网络为 MobileNetV3、YOLO v7 和本文算法,试验结果表明改进后的模型在模型检测精度和平均检测时间方面均优于其他 3 组模型。

(3)在相同的试验条件下,通过与 Faster R-CNN、YOLO v3、YOLO v5s、YOLO v7 模型对比,改进的模型在樱桃番茄数据集上取得了较优的效果,平均精度均值 mAP 分别提升 18.7、0.2、0.3、0.1 个百分点。试验充分证明了本文提出的轻量化模型不仅在评价指标上有所提高,同时在视觉上也取得了较好的效果。通过本研究可以将成熟度检测应用于农产品智能采摘中,为其他农产品智能采摘机器人的视觉识别、目标定位以及分级等后续工作提供依据。

参 考 文 献

[1] 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)

[2] 王凡,彭彦昆,汤修映,等. 樱桃番茄可溶性固形物含量的可见/近红外透射光谱无损检测[J]. 中国食品学报,2018,18(10):235-240.
WANG Fan, PENG Yankun, TANG Xiuying, et al. Near infrared nondestructive testing of soluble solids content of cherry tomato[J]. Journal of Chinese Institute of Food Science and Technology, 2018, 18(10): 235-240. (in Chinese)

[3] 邵园园,王永贤,玄冠涛,等. 基于高光谱成像的肥城桃品质可视化分析与成熟度检测[J]. 农业机械学报,2020,51(8):344-350.
SHAO Yuanyuan, WANG Yongxian, XUAN Guantao, et al. Visual detection of SSC and firmness and maturity prediction for Feicheng peach by using hyperspectral imaging[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 344-350. (in Chinese)

[4] MULYANI E D S, SUSANTO J P. Classification of maturity level of fuji apple fruit with fuzzy logic method[C]// International Conference on Cyber and IT Service Management (CITSM). IEEE, 2017: 1-4.

[5] MAZEN F, NASHAT A A. Ripeness classification of bananas using an artificial neural network[J]. Arabian Journal for Science and Engineering. Section A, Sciences, 2019, 44(8): 6901-6910.

[6] MIM F S, GALIB S M, HASAN M F, et al. Automatic detection of mango ripening stages: an application of information technology to botany[J]. Scientia Horticulturae, 2018, 237: 156-163.

[7] MOHAMMADI V, KHEIRALIPOUR K, GHASEMI V M. Detecting maturity of persimmon fruit based on image processing technique[J]. Scientia Horticulturae, 2015, 184: 123-128.

[8] 赵娟,全朋坤,马敏娟,等. 富士苹果采收成熟度光谱无损预测模型对比分析[J]. 农业机械学报,2018,49(12):347-354.
ZHAO Juan, QUAN Pengkun, MA Minjuan, et al. Comparative analysis of harvest maturity model for fuji apple based on visible/near spectral nondestructive detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(12): 347-354. (in Chinese)

[9] 周文静,查志华,吴杰. 改进圆形 Hough 变换的田间红提葡萄果穗成熟度判别[J]. 农业工程学报,2020,36(9):205-213.
ZHOU Wenjing, ZHA Zhihua, WU Jie. Maturity discrimination of "red globe" grape cluster in grapery by improved circle Hough transform[J]. Transactions of the CSAE, 2020, 36(9): 205-213. (in Chinese)

[10] 薛建新,张淑娟,张晶晶. 基于高光谱成像技术的沙金杏成熟度判别[J]. 农业工程学报,2015,31(11):300-307.

XUE Jianxin, ZHANG Shujuan, ZHANG Jingjing. Ripeness classification of Shajin apricot using hyperspectral imaging technique[J]. Transactions of the CSAE, 2015, 31(11): 300 - 307. (in Chinese)

[11] 陈锋军,张新伟,朱学岩,等. 基于改进 EfficientDet 的油橄榄果实成熟度检测[J]. 农业工程学报,2022,38(13):158 - 166.
CHEN Fengjun, ZHANG Xinwei, ZHU Xueyan, et al. Detection of the olive fruit maturity based on improved EfficientDet [J]. Transactions of the CSAE, 2022, 38(13): 158 - 166. (in Chinese)

[12] 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)

[13] TIAN Y, YANG G, WANG Z, et al. Apple detection during different growth stages in orchards using the improved YOLO - V3 model[J]. Computers and Electronics in Agriculture, 2019, 157: 417 - 426.

[14] 熊俊涛,韩咏林,王潇, 等. 基于 YOLO v5 - Lite 的自然环境木瓜成熟度检测方法[J]. 农业机械学报,2023,54(6):243 - 252.
XIONG Juntao,HAN Yonglin,WANG Xiao, et al. Method of maturity detection for papaya fruits in natural environment based on YOLO v5 - Lite[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(6):243 - 252. (in Chinese)

[15] 赵元龙,单玉刚,袁杰. 改进 YOLOv7 与 DeepSORT 的佩戴口罩行人跟踪[J]. 计算机工程与应用,2023,59(6):221 - 230.
ZHAO Yuanlong, SHAN Yugang, YUAN Jie. Wearing mask pedestrian tracking based on improved YOLOv7 and DeepSORT [J]. Computer Engineering and Applications, 2023, 59(6): 221 - 230. (in Chinese)

[16] 倪昌双,李林,罗文婷,等. 改进 YOLOv7 的沥青路面病害检测[J]. 计算机工程与应用,2023,59(13):305 - 316.
NI Changshuang, LI Lin, LUO Wenting, et al. Disease detection of asphalt pavement based on improved YOLOv7[J]. Computer Engineering and Applications, 2023, 59(13):305 - 316. (in Chinese)

[17] GH/T 1193—2021 番茄[S]. 2021.

[18] 易子娟,贾渊. 基于改进 MobileNetV3 的色环电阻识别[J]. 计算机系统应用,2023,32(4):361 - 367.
YI Zijuan, JIA Yuan. Recognition of color-ring resistors based on improved MobileNetV3 [J]. Computer Systems & Applications, 2023, 32(4): 361 - 367. (in Chinese)

[19] 彭心睿,潘晴,田妮莉. WCF - MobileNetV3:轻量型新冠肺炎 CXR 图像识别网络[J]. 计算机工程与应用,2023,59(14):224 - 231.
PENG Xinrui, PAN Qing, TIAN Nili. WCF - MobileNetV3: lightweight COVID - 19 CXR image recognition network[J]. Computer Engineering and Applications, 2023, 59(14):224 - 231. (in Chinese)

[20] 刘雪,沈长盈,吕学泽,等. 基于改进 MobileNetV3 - Large 的鸡蛋新鲜度识别模型[J]. 农业工程学报,2022,38(17):196 - 204.
LIU Xue, SHEN Changying, LÜ Xueze, et al. Recognizing egg freshness using an improved MobileNetV3 - Large[J]. Transactions of the CSAE, 2022, 38(17): 196 - 204. (in Chinese)

[21] 张林,谢刚,谢新林,等. 融合 MobileNetv3 与 Transformer 的钢板缺陷实时检测算法[J/OL]. 计算机集成制造系统:1 - 14 [2023 - 06 - 20]. http://kns.cnki.net/kcms/detail/11.5946.TP.20230104.1212.012.html.
ZHANG Lin, XIE Gang, XIE Xin Lin, et al. Real-time detection algorithm of steel plate defects integrating MobileNetv3 and Transformer[J/OL]. Computer Integrated Manufacturing Systems: 1 - 14[2023 - 06 - 20]. http://kns.cnki.net/kcms/detail/11.5946.TP.20230104.1212.012.html. (in Chinese)

[22] 王美华,吴振鑫,周祖光. 基于注意力改进 CBAM 的农作物病虫害细粒度识别研究[J]. 农业机械学报,2021,52(4):239 - 247.
WANG Meihua, WU Zhenxin, ZHOU Zuguang. Fine-grained identification research of crop pests and diseases based on improved CBAM via attention[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4):239 - 247. (in Chinese)

[23] WU T, KU T, ZHANG H. Research for image caption based on global attention mechanism[C]//Second Target Recognition and Artificial Intelligence Summit Forum,2020.

[24] 朱红春,李旭,孟炀,等. 基于 Faster R - CNN 网络的茶叶嫩芽检测[J]. 农业机械学报,2022,53(5):217 - 224.
ZHU Hongchun, LI Xu, MENG Yang, et al. Tea bud detection based on Faster R - CNN network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5):217 - 224. (in Chinese)

[25] 武星,齐泽宇,王龙军,等. 基于轻量化 YOLOv3 卷积神经网络的苹果检测方法[J]. 农业机械学报,2020,51(8):17 - 25.
WU Xing, QI Zeyu, WANG Longjun, et al. Apple detection method based on light-YOLOv3 convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8):17 - 25. (in Chinese)

[26] 李健源,柳春娜,卢晓春,等. 基于改进 YOLOv5s 和 TensorRT 部署的鱼道过鱼监测[J]. 农业机械学报,2022,53(12):314 - 322.
LI Jianyuan, LIU Chunna, LU Xiaochun, et al. Fish passage monitoring based on improved YOLOv5s and TensorRT deployment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(12):314 - 322. (in Chinese)