

基于改进 YOLO v5 的夜间温室番茄果实快速识别

何 斌^{1,2} 张亦博¹ 龚健林¹ 付 国^{1,2} 赵昱权¹ 吴若丁¹

(1. 西北农林科技大学水利与建筑工程学院, 陕西杨凌 712100;

2. 西北农林科技大学旱区农业水土工程教育部重点实验室, 陕西杨凌 712100)

摘要:为实现日光温室夜间环境下采摘机器人正常工作以及番茄快速识别,提出一种基于改进 YOLO v5 的夜间番茄果实的识别方法。采集夜间环境下番茄图像 2 000 幅作为训练样本,通过建立一种基于交并比的 CIOU 目标位置损失函数,对原损失函数进行改进,根据计算函数 anchor 生成自适应锚定框,确定最佳锚定框尺寸,构建改进型 YOLO v5 网络模型。试验结果表明,改进 YOLO v5 网络模型对夜间环境下番茄绿色果实识别精度、红色果实识别精度、综合平均识别精度分别为 96.2%、97.6% 和 96.8%,对比 CNN 卷积网络模型及 YOLO v5 模型,提高了被遮挡特征物与暗光下特征物的识别精度,改善了模型鲁棒性。将改进 YOLO v5 网络模型通过编译将训练结果写入安卓系统制作快速检测应用软件,验证了模型对夜间环境下番茄果实识别的可靠性与准确性,可为番茄实时检测系统的相关研究提供参考。

关键词:番茄果实;温室;识别;改进 YOLO v5;夜间

中图分类号: TP391.4

文献标识码: A

文章编号: 1000-1298 (2022)05-0201-08

OSID:



Fast Recognition of Tomato Fruit in Greenhouse at Night
Based on Improved YOLO v5

HE Bin^{1,2} ZHANG Yibo¹ GONG Jianlin¹ FU Guo^{1,2} ZHAO Yuquan¹ WU Ruoding¹

(1. College of Water Resources and Architectural Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China

2. Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semiarid Areas, Ministry of Education, Northwest A&F University, Yangling, Shaanxi 712100, China)

Abstract: In order to realize the normal operation of the picking robot and the rapid recognition of tomato in the nighttime environment of solar greenhouse, a nighttime tomato fruit detection method based on improved YOLO v5 (You only look once) was proposed. Totally 2000 tomato images in the night environment were collected as the initial training samples, and the original loss function was improved by establishing a CIOU target position loss function based on intersection and union ratio, and then an adaptive anchor frame was generated according to the anchor calculation function, the optimal anchor frame size was determined, the network structure was optimized, and an improved YOLO v5 network model was constructed, and the recognition rate of tomato fruit in night environment was improved. The experimental results showed that the average recognition accuracy of improved YOLO v5 network model for tomato green and red fruits and average recognition accuracy in night environment was 96.2% , 97.6% and 96.8% . Compared with traditional CNN convolution network model and traditional YOLO v5 model, the recognition accuracy of occluded features and features in dark light was improved and the robustness of the model was improved. The improved YOLO v5 network model compiled and wrote the training results into Android system to make a rapid detection application software, which verified the reliability and accuracy of the model for tomato fruit recognition in night environment, and provided a reference for the relevant research of tomato real-time detection system.

Key words: tomato fruit; greenhouse; recognition; improved YOLO v5; night

收稿日期: 2021-06-23 修回日期: 2021-09-19

基金项目: 陕西省科技创新引导专项(2021QFY08-01)

作者简介: 何斌(1971—),男,副教授,博士,主要从事设施农业工程研究,E-mail: ylhebin@nwfufu.edu.cn

0 引言

我国既是番茄出口大国又是消费大国,在日光温室内对番茄进行实时监控并利用采摘机器人不仅降低人工成本,同时又能提高效率,节省时间,对于番茄培育具有重要意义^[1-2]。但目前夜间环境下光线复杂,果实与叶片重叠在昏暗光线下进一步对采摘机器人与监控设备的识别造成了困难。

日光温室内番茄果实的传统识别方法主要依靠颜色空间的差异与形状对图像进行分割,随着计算机技术与机器视觉技术的发展,深度学习也应用到了番茄果实的识别方法中^[3-11]。综上所述,图像分割等处理方法对于番茄识别的精度较低,卷积神经网络(CNN)等方法实时性较差,基于深度学习的识别算法虽提高了实时性与精度,但面对夜间复杂环境下的各种影响因素考虑不全,难以满足实际要求。

随着以卷积神经网络为代表的深度学习不断地发展,目标检测已经广泛地运用到了各个领域。目标检测算法可以分为两类,第一类是基于 region proposal(候选区域)的 R-CNN^[12] 系列算法,如:R-CNN、Fast R-CNN^[13]、Faster R-CNN^[14] 等。该类算法通过两个步骤进行计算:选取候选框;对候选框进行分类或者回归。此类方法鲁棒性高,识别错误率低,但运算时间长,占用磁盘空间大,对图像信息进行重复计算,不适合进行实时检测。第二类是如 YOLO^[15]、SSD^[16] 等网络模型的 one-stage 算法^[17]。该类算法是采用不同的尺寸对图像进行遍历抽样,然后利用 CNN 提取特征后直接进行回归^[18]。该方法识别速度快,实时性强,但由于网络模型简单,识别率低于第一类方法。SSD 神经网络运行速度略低于 YOLO,检测精度略低于 Faster R-CNN,但鲁棒性较差,需要人工设置各种参数阈值,且先验框无法通过学习获得,需要手工设置,对于小目标存在低级特征卷积层数少,特征提取不充分的问题。YOLO 系列网络以 YOLO v5 为代表,由于平衡了速度快的特点,丧失了部分精度,因为没有进行区域采样,所以在小范围的信息上表现较差,具有识别物体位置精准性差、召回率低等问题。研究表明^[19],由于第二类目标检测方法实时性强,有利于提高采摘机器人及监测设备的工作效率,适用于复杂环境下实时的目标检测。

本文基于改进 YOLO v5 目标检测算法,根据夜间日光温室下复杂环境情况,采用 YOLO v5 快速精确的检测结构,融合多尺度信息模型,通过数据增强等方法提高检测精度,构建一种可在移动设备下实现夜间复杂环境的番茄本体特征识别的网络模型。

并与其他算法的对比,验证本网络模型的实时性与准确性,以期为采摘机器人与实时检测设备系统设计提供参考。

1 番茄图像数据采集

番茄图像数据采集自西北农林科技大学园艺场节能型日光温室,番茄植株采用吊蔓式栽培方法。图像采集设备为佳能 EOS-750D 型相机,采用 CMOS 传感器,APS 画幅(22.3 mm×14.9 mm),采集番茄果实、叶片、花卉、秆茎 4 类 RGB 图像共 2 000 幅用于数据训练及测试,RGB 图像分辨率为 4 000 像素×6 000 像素。由于模拟日光温室内夜间环境下采摘机器人夜间视觉系统,采集过程中在图像采集设备一侧设置对角光源^[20],如图 1 所示。为了避免由于样本数据多样性不足导致的过拟合现象,考虑夜间环境下暗光区域以及阴影区域,图像数据分为正常采样与暗光采样两种情况,如图 2 所示。图像样本中对红色成熟番茄果实与绿色未成熟果实两种情况进行区分,同时也对不同果实数量、不同枝叶遮挡程度进行区分以增加样本多样性。图 3 为夜间环境下番茄果实图像。

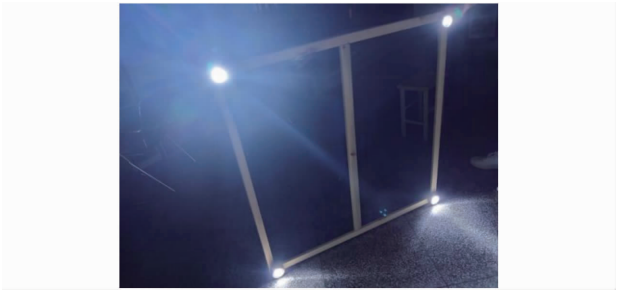
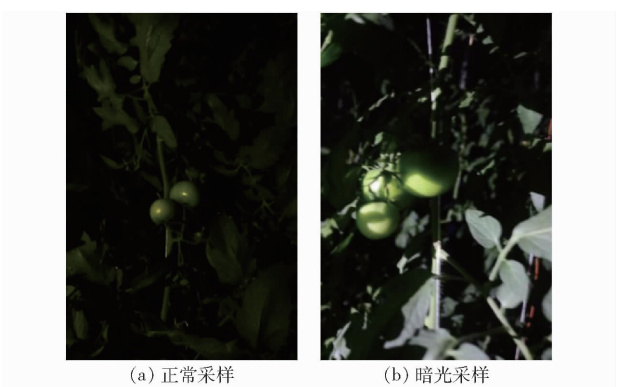


图 1 夜间照明设备图
Fig. 1 Night lighting diagram



(a) 正常采样 (b) 暗光采样

图 2 不同条件下采样实例
Fig. 2 Sample diagrams under different conditions

为了保证数据参数的准确性,在数据训练前需要人工标注数据,标注时将番茄各器官的最小外接矩形作为真实框,以此减少框内背景上的无用像素。本研究的数据训练基于 pytorch 框架,使

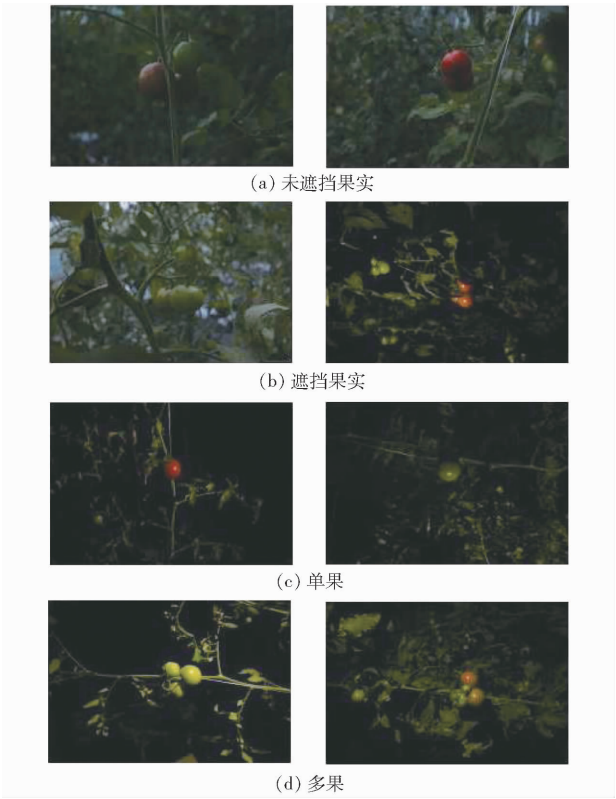


图 3 夜间环境下番茄果实

Fig. 3 Effects of night environment on tomato fruit

用线性增强技术降低出现样本不均匀的概率,即在不改变原有数据特征的情况下对原有数据进行图像处理,无需增加原始数据量,采用的数据增强手段主要有:①翻转。对图像进行 90° 翻转、水平翻转以及随机(0° ~ 180°)旋转,模拟检测过程中图像抓取的随机性。②缩放。按照一定比例缩放图像尺寸,模拟检测过程中受距离影响的图像尺寸不同。③颜色抖动。改变图像的饱和度与亮度,模拟夜间环境下不同亮度差异。④添加噪声。对数据图像添加椒盐噪声和高斯噪声,模拟拍摄过程中的噪声,同时降低高频特征防止出现过拟合现象。通过数据增强技术后得到新的数据样本量,共计 9 042 幅。其中 70% 用于数据集训练,15% 用于数据集验证,15% 用于数据集测试。

2 基于 YOLO v5 的番茄识别网络

2.1 YOLO v5 网络模型

YOLO v5 是一种单阶段目标检测算法,该算法在 YOLO v4 的基础上添加了新的改进思路,使其速度与精度都得到极大的提升。YOLO v5 网络模型主要分为输入端、Backbone 基准网络、Neck 网络、Head 输出端 4 部分。输入端:包含一个图像预处理阶段,将输入图像缩放到网络的输入尺寸,并进行归一化等操作,操作方法为 Mosaic 数据增强操作、自适应锚框计算与自适应图像缩放方法。基准网络:通常

是一些性能优异的分类器的网络,该模块用来提取一些通用的特征表示。YOLO v5 同时使用了 CSPDarknet53 结构与 Focus 结构作为基准网络。Neck 网络:使用 SPP 模块、FPN + PAN 模块位于基准网络和头网络的中间位置,利用此两个模块进一步提升特征的多样性及鲁棒性。Head 输出端:用来完成目标检测结果的输出,包含一个分类分支和一个回归分支,利用 GIOU_Loss 来代替 Smooth L1 Loss 函数,增加了相交尺度的衡量,从而进一步提升算法的检测精度。

2.2 多尺度特征提取

YOLO v5 算法中使用 Focus 与 CSPDarknet53 特征提取网络获取多尺度图像特征,相比前期版本进一步消除网格敏感性,提高对于遮挡物体特征信息的拾取,优化小目标特征信息差异,提高识别精度。该算法首先以 640 像素 × 640 像素的图像为输入进行切片操作,先变成 320 × 320 × 12 的特征图,切片过程如图 4 所示。

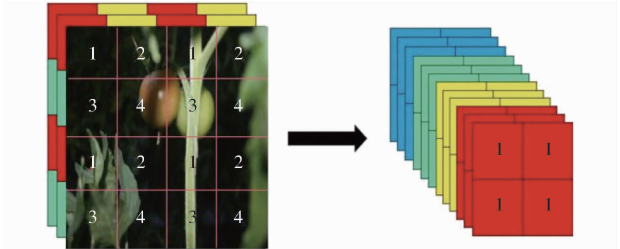


图 4 切片操作演示图

Fig. 4 Slice operation demo

再经过一次卷积操作,最终变成 320 × 320 × 32 的特征图,该操作通过增加计算量来保证图像特征信息不会丢失,将候选框宽 W 、候选框高 H 的信息集中到通道上,使得特征提取得更加充分。然后再分别进行 32、16、8 倍下采样,获得不同层次的特征图,然后通过上采样和张量拼接,将不同层次的特征图融合转化为维度相同的特征图,如图 5 所示。

其解决了其他大型卷积神经网络框架 Backbone 中网络优化的梯度信息重复问题,将梯度的变化从头到尾地集成到特征图中,因此减少了模型的参数量和每秒浮点运算次数(FLOPS),既保证了推理速度和准确率,又减小了模型尺寸。鉴于本文检测红、绿两种果实的不同目标,Darknet53 特征提取最终分别输出 13 像素 × 13 像素、26 像素 × 26 像素、52 像素 × 52 像素 3 种尺度的特征图,分别作为远近景尺度视场内各目标回归检测的依据。

2.3 自适应锚定框

在 YOLO v5 中根据目标样本边框标注信息,采用 K-means 聚类算法预先设置先验框进行回归预测,并根据特征图层次进行分配,但 K-means 聚类算

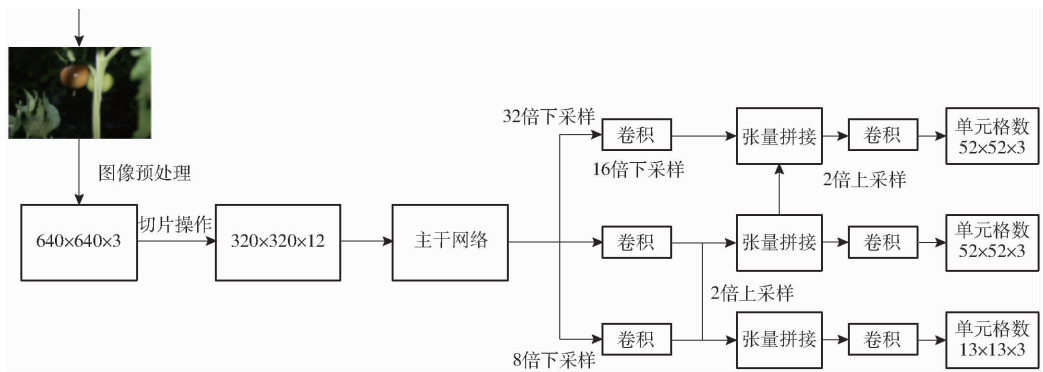


图5 Darknet53 特征流程图

Fig. 5 Darknet53 feature flowchart

法可能收敛至局部最小值,无法给出最优解,对预测框造成误差。因此对 YOLO v5 中锚定框计算方式进行改进。在原始 COCO 数据集中,针对不同特征大小的检测框尺寸如表 1 所示。通过对原始检测框的检测效果分析表明,检测框数据不协调,候选检测框长宽比达 1:8,不利于训练效果,影响真实框预测结果^[21]。因此通过 anchor 计算函数进行机器学习,迭代出每个步骤下的最优检测框尺寸,能够有效抓取小目标特征值,减小真实检测框偏离程度,并提高被遮挡物体的识别精度。重新计算检测框尺寸,设置聚类数为 9,anchor 与 bbox 比值为 8,迭代次数为 1 000 次,通过 YOLO v5 自动学习确定锚定框的最优尺寸。

表 1 原始数据集检测框尺寸分配

Tab.1 Size allocation of original data set detection frame

单元格尺寸/(像素×像素)	检测框尺寸/(像素×像素)
13×13	(116×90)、(156×198)、(373×326)
26×26	(30×61)、(62×45)、(59×119)
52×52	(10×13)、(16×30)、(33×23)

表 2 为数据集中缩放至一定水平的最佳锚定框尺寸,并用该尺寸替换原始 COCO 数据集中锚定框尺寸。

表 2 改进后检测框尺寸分配

Tab.2 Improved size allocation of detection frame

单元格尺寸/(像素×像素)	检测框尺寸/(像素×像素)
13×13	(254×69)、(314×82)、(373×95)
26×26	(174×62)、(175×43)、(209×38)
52×52	(134×38)、(135×48)、(172×35)

2.4 损失函数改进

YOLO v5 采用 GIOU_Loss 做 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,计算公式为

$$Loss_{coord} = \sum_{i=0}^S \sum_{j=0}^B l_{ij}^{obj} (1 - GIOU_{ij}) \tag{1}$$

其中
$$GIOU_{ij} = \frac{J}{U} - \frac{A - U}{A} \tag{2}$$
$$U = \hat{\omega}_i \hat{h}_i + \omega_i h_i - J \tag{3}$$

式中 $Loss_{coord}$ ——目标位置损失函数
 l_{ij}^{obj} ——单元格 i 产生的先验框 j 包含目标
 J ——边框交集面积
 U ——边框并集面积
 A ——边框最小外接矩形面积
 ω_i, h_i ——预测框高与宽
 S ——真实框与预测框最小外接矩形面积
 β ——真实框与预测框并集区域面积
 $\hat{\omega}_i, \hat{h}_i$ ——GIOU 损失函数中真实边框高与宽

当预测框和真实框重合时, GIOU 最大取 1,反之,随着二者距离的增大 GIOU 趋近于 -1,即预测框与真实框距离越远时损失值越大。但当预测框与真实框出现包含关系或出现宽和高对齐的情况时,差集为 0,则损失函数不可导,无法收敛,容易对遮挡情况下物体进行漏检。因此本文采用考虑到预测框中心点欧氏距离和重叠率参数(CIOU)的损失函数作为预测框偏差的偏差指标^[22],如图 6 所示。图中 (ω^{gt}, h^{gt}) 、 (ω, h) 分别表示预测框和真实框的高与宽, b, b^{gt} 分别表示预测框和真实框的中心点, ρ 表示两个中心点间的欧氏距离, c 表示最小外接矩形框对角线距离。则有

$$\begin{cases} CIOU_{ij} = \frac{J}{U} - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha\nu \\ \nu = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h} \right)^2 \end{cases} \tag{4}$$

式中 α ——权重函数
 ν ——真实框与预测框矩形对角线倾斜角的方差

则目标函数改进为

$$Loss_{coord} = \sum_{i=0}^{s^2} \sum_{j=0}^B l_{ij}^{obj} (1 - CIOU_{ij}) \tag{5}$$

该目标函数增加了中心点距离度量,可以直接

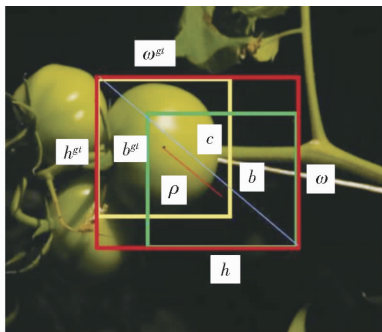


图 6 损失函数 CIOU 边框图
Fig. 6 Loss function CIOU border chart

最小化两个目标框的距离,收敛速度大于 GIOU 损失函数,且考虑到差异化情形,避免了真实框与预测框包含关系时的不收敛情况,能够有效提高物体在遮挡情况下的识别率,优化了边框之间的相互关系。

3 网络模型训练

3.1 算法运行环境与参数设置

检测网络算法在深度学习框架中运行,硬件环境为 Intel i5-9560 处理器,16GB DDR4 2 400 MHz 运行内存,显卡为 GeForce GTX 1050Ti。软件环境为 Windows 10 操作系统下 pytorch 1.8.1 深度学习框架和 CUDA 11.1 并行计算构架。

COCO 和 VOC 数据集上的训练结果初始化 YOLO v5 的网络参数,参数训练采用 SGD 优化算法,参数设置如下:图像输入尺寸为 640 像素 × 640 像素,Batchsize 为 32;最大迭代次数为 700;动量因子为 0.9;权重衰减系数为 0.000 5。采用余弦退火策略动态调整学习率,余弦退火超参数为 0.1,初始学习率为 0.001,采用 CIOU Loss 作为损失函数。

为了验证本试验方法的有效性,将改进 YOLO v5、YOLO v5 与 CNN 在数据集上进行试验对比。

3.2 训练过程与结果

采用 YOLO v5s 检测模块,100 次迭代周期中损失函数如图 7 所示,前 20 次迭代周期中损失函数值明显减小,后经历 80 次迭代损失函数值逐渐稳定至 0.03 处小幅波动,则认为该检测网络模型稳定收敛。经过非极大值抑制(NMS)处理后得到的预测框分类中,将置信度大于阈值 0.5 的预测框定义为正样本,反之为负样本。

选用检测网络性能的指标包括平均精度 M_{AP} 、绿色果与红色果检测精度 A_{PG} 与 A_{PR} 、准确率 P 、检测时间 t 、召回率 R 及交并比 (IOU)。主要衡量指标为平均精度 M_{AP} 与检测时间 t ,反映了检测网络的准确性与速度,准确率 P 为被预测为正的样本中实际为

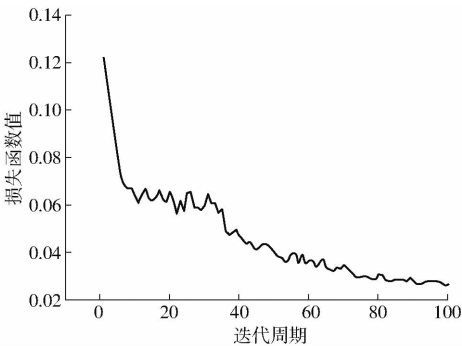


图 7 损失函数曲线
Fig. 7 Loss function curve

正的样本概率,召回率 R 为实际为正的样本中被预测为正样本的概率,交并比 IOU 为预测框与真实框的重合程度。为了对比改进后网络模型的检测效果,将改进 YOLO v5、YOLO v5 与 CNN 网络模型进行比较。CNN 网络模型选择 Faster R-CNN 网络模型,对于任意尺寸的图像,首先缩放至固定尺寸 $M \times N$,然后将图像送入网络,池化层中包含了 13 个 conv 层 + 13 个 relu 层 + 4 个 pooling 层;RPN 网络首先经过 3×3 卷积,再分别生成 positive anchors 和对应 bounding box regression 偏移量,然后计算出 proposal,而 ROI Pooling 层则利用 proposals 从 feature maps 中提取 proposals feature 送入后续全连接和 softmax 网络做 classification。Faster R-CNN 网络模型训练包含 5 个步骤:①在已经训练好的 model 上,训练 RPN 网络。②利用步骤①中训练的 RPN 网络,收集 proposals。③训练 Faster R-CNN 网络与 RPN 网络。④再次利用步骤③中训练好的 RPN 网络,收集 proposals。⑤第 2 次训练 Faster R-CNN 网络,得到结果。

在训练过程中,对各类图像占比分两组进行统计。其中第 1 组根据果实遮挡率进行划分,果实遮挡率小于等于 50% 的图像共 6 473 幅,占比 69%;果实遮挡率大于 50% 的图像共 2 929 幅,占比 31%。第 2 组根据单果、多果进行划分,单果图像共 4 012 幅,占比 43%;多果图像共 5 390 幅,占比 57%。训练集、验证集、测试集检测指标结果如表 3 所示。为还原真实场景中叶片对果实遮挡的情况,根据叶片对果实的遮挡率进行划分,遮挡率计算方式为叶片遮挡果实部分的面积与果实总面积的比值。当遮挡率小于等于 50% 时,可认为叶片对果实的遮挡无影响,当遮挡率大于 50% 时认为叶片的遮挡对识别精度造成了影响,需要额外进行说明。

3 种检测结果如表 4 所示。由表 4 可知,YOLO v5 的 M_{AP} 为 93.4%,改进 YOLO v5 模型 M_{AP} 为 96.7%,提高了 3.3 个百分点,检测单幅图像时间

表 3 训练集、验证集、测试集检测指标

Tab.3 Test index results of training set, verification set and test set

数据集	遮挡率		单果精度	多果精度	平均精度	检测时间	准确率
	小于等于 50%	大于 50%					
	$A_{PR}/\%$	$A_{PC}/\%$					
	$P_{si}/\%$	$P_{st}/\%$					
训练集	97.6	96.1	98.1	97.4	97.3	10	96.7
验证集	96.3	95.0	96.4	96.6	96.1	10	95.9
测试集	96.8	95.5	97.3	96.9	96.6	10	96.2

表 4 3 种模型各项检测指标

Tab.4 Results of all test indexes of three models

检测模型	遮挡率		单果精度	多果精度	平均精度	检测时间	准确率	IOU/%
	率小于 50%	大于 50%						
	$A_{PR}/\%$	$A_{PC}/\%$						
	$P_{si}/\%$	$P_{st}/\%$						
改进 YOLO v5	96.2	95.3	97.2	96.3	96.7	10	96.3	87.6
YOLO v5	95.7	94.3	96.9	95.1	93.4	10	93.3	85.8
CNN	97.1	95.9	98.1	96.4	97.1	67	96.0	84.1

均为 10 ms,可以看出改进 YOLO v5 模型鲁棒性优于 YOLO v5。改进 YOLO v5 模型绿色果检测精度和红色果检测精度分别较 YOLO v5 高 0.5、1.0 个百分点,可以看出改进 YOLO v5 模型因更换了损失函数,对于遮挡番茄果实预测率提高,因此模型的平均精度有所提高。同时改进 YOLO v5 模型的交并比 IOU 比 YOLO v5 模型提高了 1.8 个百分点,由此可以看出,使用计算函数 anchor 重新计算检测框尺寸并替换原有检测框尺寸能够有效提高交并比,提高检测模型的鲁棒性。

另外,虽改进 YOLO v5 模型的平均精度略低于 CNN 网络模型,但改进 YOLO v5 的准确率与召回率均略高于 CNN 模型,果实遮挡率小于等于 50%、果实遮挡率大于 50% 的差值与单果多果检测精度的差值均小于 CNN 模型,且检测时间为 CNN 模型的 6.7 倍,说明在保持准确率的情况下,改进 YOLO v5 模型的鲁棒性与实时性均高于 CNN 网络模型。又因为 CNN 模型的检测框利用滑动检测方式,因此交并比较改进 YOLO v5 低了 3.5 个百分点,说明改进 YOLO v5 对提取不同番茄果实特征信息与区分背景与目标信息的能力更为突出。

4 试验

4.1 材料与方法

为验证模型可靠性,2021 年 6 月 1 日北京时间 20:30 于西北农林科技大学北校区园艺场番茄日光

温室内进行现场试验,随机挑选 250 个区域进行数据采集,利用改进 YOLO v5 模型与 YOLO v5 模型编译的手机端应用进行番茄检测试验。

试验方法如下:①将两种网络模型的数据文件进行安卓端部署,生成手机检测 APP。②在日光温室内随机挑选区域进行数据采集,为满足随机性,图像数据包含远近、遮挡等不同形态。③对采集数据进行人工识别,对不同颜色番茄果实,遮挡与未遮挡等特殊情况进行分类。④利用两种网络模型的手机应用对采集的图像数据进行识别,并与人工识别结果进行比对,分析模型精度。

4.2 YOLO v5 安卓端部署

将 YOLO v5 模型训练后的 Best. pt 格式权重文件进行半精简转换为 onnx 格式文件,再将 onnx 文件使用腾讯 NCNN 平台编译为 bin 与 param 格式文件。利用 Netron 可视化工具对编译后的文件进行修改,去除原模型文件中的切片操作。最后使用 Android Studio 工具将修改后的模型文件写入 Android 系统生成应用软件。试验采用手机型号为 Oneplus 8t,系统版本为 Android 11.0。

4.3 结果分析

为验证实际检测效果,使用改进 YOLO v5 与 YOLO v5 两种模型训练后生成的手机应用进行现场检测,如图 8 所示。并对检测结果进行统计。以人工识别的番茄果实结果为参考,分别对两种模型的检测结果进行对比与分析评价。2 个模型分别对番茄红果、绿果的识别数量、识别总数量与人工识别数量的比值作为两个模型番茄红果、绿果的检测精度。统计结果如表 5 所示。

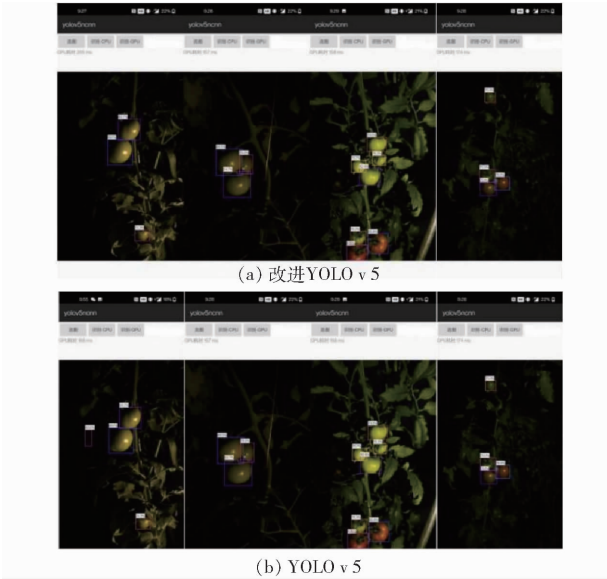


图 8 2 种模型手机应用的夜间果实检测效果

Fig.8 Night fruit detection effects of two mobile phone models

表 5 2 种模型各项检测指标

Tab. 5 Results of all test indexes of two models %			
模型	红色果实精度	绿色果实精度	总精度
改进 YOLO v5	97. 6	96. 2	96. 8
YOLO v5	96. 2	94. 4	95. 2

由表 5 可知,改进 YOLO v5 模型的番茄红果与绿果的识别精度较 YOLO v5 模型分别提高了 1. 4、1. 8 个百分点,总识别精度提高了 1. 6 个百分点,对于遮挡情况下或多果重叠情况下改进 YOLO v5 模型识别率优于 YOLO v5 模型。为了验证复杂环境下的模型检测效果,对番茄绿果、红果的单果、多果以及遮挡、未遮挡进行统计区分。选取各种情况下样本数量 50 幅进行验证,成功识别数量与样本总量比值为识别率,结果如表 6 所示。由表 6 可以看出,在遮挡率不超过 50% 的情况下,番茄红

表 6 复杂环境下番茄果实检测结果

Tab. 6 Detection results of tomato fruit in complex environment								
参数	番茄绿果				番茄红果			
	单果		多果		单果		多果	
	遮挡	未遮	遮挡	未遮	遮挡	未遮	遮挡	未遮
		挡		挡		挡		挡
样本数量/个	50	50	80	80	50	50	80	80
识别率/%	100	100	96	98	100	100	98	98

果与绿果的单果识别率均可达 100%。由于番茄绿果在夜间环境下颜色容易与叶片、秆茎等混淆,多果在重叠情况下边界条件区分较为困难,因此番茄绿果识别率及多果识别率略低于番茄红果。

5 结论

(1)提出了基于改进 YOLO v5 的识别网络模型对夜间环境下的番茄果实识别,使用计算函数 anchor 迭代出最优检测框代替原有方法,并修改原始数据中的目标检测损失函数,建立了夜间环境下番茄果实识别模型。改进 YOLO v5 的识别网络模型识别精度得到改善, M_{AP} 为 96. 7%,较 YOLO v5 的 M_{AP} (93. 4%)提高了 3. 3 个百分点。

(2)为验证模型的实际应用性,利用改进 YOLO v5 模型训练后的权重文件制作手机端应用软件进行现场检测。试验表明,改进 YOLO v5 模型检测精度较 YOLO v5 模型得到改善,针对夜间环境下绿色果实、红色果实及总果实精度分别为 96. 2%、97. 6% 和 96. 8%。

(3)改进 YOLO v5 模型对于多果重叠、遮挡等复杂情况下的番茄果实识别有显著的提升。在多果的遮挡情况下番茄绿果、红果识别精度分别达 96%、98%,相比于 YOLO v5 网络模型鲁棒性更优。

参 考 文 献

[1] 李君明,项朝阳,王孝宣,等. “十三五”我国番茄产业现状及展望[J]. 中国蔬菜, 2021(2):13 – 20.

[2] 孙永生,张逸鸣,王国政,等. 辽宁省庄河市设施番茄产业现状及发展对策[J]. 园艺与苗种, 2020,40(1):28 – 30. SUN Yongsheng,ZHANG Yiming,WANG Guozheng,et al. Present situation and development strategy of facility tomato industry in Zhuanghe City Liaoning Province[J]. Horticulture and Seedling, 2020, 40(1): 28 – 30. (in Chinese)

[3] 周云成,许彤羽,邓寒冰,等. 基于面向通道分组卷积网络的番茄主要器官实时识别[J]. 农业工程学报, 2018,34(10): 153 – 162. ZHOU Yuncheng,XU Tongyu,DENG Hanbing,et al. Real-time recognition of main organs in tomato based on channel wise group convolutional network[J]. Transactions of the CSAE, 2018,34(10): 153 – 162. (in Chinese)

[4] 周云成,许彤羽,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)

[5] 周云成,许彤羽,邓寒冰,等. 基于自监督学习的番茄植株图像深度估计方法[J]. 农业工程学报, 2019, 35(24): 173 – 182. ZHOU Yuncheng,XU Tongyu,DENG Hanbing,et al. Method for estimating the image depth of tomato plant based on self-supervised learning[J]. Transactions of the CSAE, 2019,35(24): 173 – 182. (in Chinese)

[6] 李颀,强华. 基于双目视觉与深度学习的番茄本体特征检测系统[J]. 南方农业学报, 2020, 51(1): 237 – 244. LI Qi,QIANG Hua. Tomato ontology feature detection system based on binocular vision and deep learning[J]. Journal of Southern Agriculture, 2020,51(1): 237 – 244. (in Chinese)

[7] 项荣,段鹏飞. 基于重叠边缘的夜间重叠番茄识别[J]. 华中科技大学学报(自然科学版), 2019,47(5):68 – 72. XIANG Rong,DUAN Pengfei. Recognition of overlapping tomatoes based on overlapping edges at night[J]. Journal of Huazhong University of Science and Technology, 2019,47(5):68 – 72. (in Chinese)

[8] 刘芳,刘玉坤,林森,等. 基于改进型 YOLO 的复杂环境下蕃茄果实快速识别方法[J]. 农业机械学报, 2020,51(6):229 – 237. 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): 229 – 237. (in Chinese)

[9] 项荣,张杰兰. 基于改进 PCNN 的番茄植株夜间图像分割算法[J]. 农业机械学报, 2020, 51(3):71 – 79. XIANG Rong,ZHANG Jielan. Image segmentation for tomato plants at night based on improved PCNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(3): 71 – 79. (in Chinese)

- [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] 冯俊惠,李志伟,戎有丽,等. 基于改进 Hough 圆变换算法的成熟番茄果实识别[J]. 中国农机化学报, 2021, 42(4): 191–196.
FENG Junhui, LI Zhiwei, RONG Youli, et al. Identification of mature tomatoes based on an algorithm of modified circular Hough transform[J]. Journal of Chinese Agricultural Mechanization, 2021, 42(4): 191–196. (in Chinese)
- [12] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014: 580–587.
- [13] LI J, LIANG X, SHEN S M, et al. Scale-aware Fast R – CNN for pedestrian detection [C] // IEEE Transactions on Multimedia, 2015.
- [14] AMAIA S, XAVIER G, FERRAN M, et al. Faster R – CNN features for instance search [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 394–401.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [16] WEI Liu, DRAGOMIR A, DUMITRU E, et al. SSD: single shot multibox detector[C]//ECCV, 2016: 21–37.
- [17] FRIBERG B, SENNERBY L, LINDEN B, et al. Stability measurements of one-stage Brnemark implants during healing in mandibles[J]. International Journal of Oral & Maxillofacial Surgery, 2010, 28(4):266–272.
- [18] 成伟,张文爱,冯青春,等. 基于改进 YOLOv3 的温室番茄果实识别估产方法[J]. 中国农机化学报, 2021, 42(4):177–182.
CHENG Wei, ZHANG Wen'ai, FEI Qingchun, et al. Method of greenhouse tomato fruit identification and yield estimation based on improved YOLOv3[J]. Journal of Chinese Agricultural Mechanization, 2021, 42(4): 177–182. (in Chinese)
- [19] 赵春江,王朝武,林森,等. 基于级联卷积神经网络的番茄花期识别检测方法[J]. 农业工程学报, 2020, 36(24): 143–152.
ZHAO Chunjiang, WEN Chaowu, LIN Sen, et al. Tomato florescence recognition and detection method based on cascaded neural network[J]. Transactions of the CSAE, 2020, 36(24): 143–152. (in Chinese)
- [20] 项荣,段鹏飞. 番茄采摘机器人夜间照明系统设计与试验[J]. 农业机械学报, 2016, 47(7): 8–14.
XIANG Rong, DUAN Pengfei. Design and experiment of night lighting system for tomato picking robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(7): 8–14. (in Chinese)
- [21] JU Moran, LUO Haibo, WANG Zhongbo, et al. The application of improved YOLO v3 in multi-scale target detection [C] // Applied Sciences-Basel, 2019, 9(18): 3775.
- [22] XIAO Liyun, ZHOU Peng, XU Ke, et al. Multi-directional scene text detection based on improved YOLO v3 [J]. Sensors, 2021, 21(14): 4870.

(上接第 180 页)

- [23] 曾文,向梨丽,张小林. 南京市社区服务设施可达性的空间格局与低收入社区空间剥夺研究[J]. 人文地理, 2017(1): 73–81.
ZENG Wen, XIANG Lili, ZHANG Xiaolin. Research in spatial pattern of accessibility to community service facilities and spatial deprivation of low income community in Nanjing[J]. Human Geography, 2017(1):73–81. (in Chinese)
- [24] 周扬,李寻欢,童春阳,等. 中国村域贫困地理格局及其分异机理[J]. 地理学报, 2021, 76(4):903–920.
ZHOU Yang, LI Xunhuan, TONG Chunyang, et al. The geographical pattern and differentional mechanism of rural poverty in China[J]. Acta Geographica Sinica, 2021, 76(4):903–920. (in Chinese)
- [25] 王丹,方斌,陈正富. 基于外部性理论的单中心城市土地经济效益测度与低效用地诊断研究——以扬州市区为例[J]. 中国土地科学, 2018, 32(6):61–67.
WANG Dan, FANG Bin, CHEN Zhengfu. Land economic benefit measurement and underused land diagnosis of single central city based on the externality theory: a case study in Yangzhou City[J]. China Land Science, 2018, 32(6):61–67. (in Chinese)
- [26] 戴尔阜,王亚慧. 横断山区产水服务空间异质性及归因分析[J]. 地理学报, 2020, 75(3):607–619.
DAI Erfu, WANG Yahui. Spatial heterogeneity and driving mechanisms of water yield service in the Hengduan mountain region [J]. Acta Geographica Sinica, 2020, 75(3):607–619. (in Chinese)
- [27] 杨建宇,徐凡,赵龙,等. 基于综合质量空间自相关性分析的耕地保护分区方法[J]. 农业机械学报, 2017, 48(3):156–163.
YANG Jianyu, XU Fan, ZHAO Long, et al. Protection zoning of cultivated land comprehensive quality based on spatial autocorrelation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(3):156–163. (in Chinese)
- [28] 王劲峰,徐成东. 地理探测器:原理与展望[J]. 地理学报, 2017, 72(1):116–134.
WANG Jinfeng, XU Chengdong. Geodetector: principle and prospective[J]. Acta Geographica Sinica, 2017, 72(1):116–134. (in Chinese)