

基于改进 Tiny-YOLOv5l 算法的串型 番茄定位与计数

Localization and counting of string tomatoes based on
improved Tiny-YOLOv5l algorithm

赵九霄^{1,2} 张馨^{1,2} 史凯丽³ 李晶晶³ 李作麟^{1,2}

ZHAO Jiu-xiao^{1,2} ZHANG Xin^{1,2} SHI Kai-li³ LI Jing-jing³ LI Zuo-lin^{1,2}

(1. 北京市农林科学院信息技术研究中心, 北京 100097; 2. 农业农村部农业信息软硬件产品质量
检测重点实验室, 北京 100097; 3. 北京市农林科学院智能装备技术研究中心, 北京 100097)

(1. *Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China*; 2. *Key Laboratory for Quality Testing of Software and Hardware Products on Agricultural Information, Ministry of Agriculture and Rural Affairs, Beijing 100097, China*; 3. *Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China*)

摘要:目的:提高串型番茄分拣效率,减少误检、错检。方法:采集串型番茄图像数据集,通过数据增强扩充数据并提高模型的泛化性能,将 YOLOv5l 框架内的 Bottleneck 层中的 3×3 卷积替换为改进的 SVM-MHSA 层,通过将 MHSA 中 softmax 分类函数替换为更适用于串型番茄的 SVM 分类函数;将检测框架中剩余 3×3 卷积替换为深度可分离卷积,引入随机纠正线性单元提高网络训练收敛速度。结果:改进后的 Tiny-YOLOv5l 模型可有效实现串型单果识别定位、整串果实计数,检测框损失率由 1.48% 降低至 1.34%,目标损失率由 1.98% 降低至 1.73%,置信度损失降低了 1.4%,精度由 97.36% 提升至 98.89%,召回率由 97.35% 提升至 98.56%。结论: Tiny-YOLOv5l 算法更加精准且兼具轻量化,面对遮挡、背景干扰、光照变化、虚化等挑战具有较高的识别准确率,可为产后串型番茄分拣人员提供准确的单果位置信息以及整串果实数量信息。

关键词:深度学习;分类函数;串型番茄;复杂场景

Abstract: Objective: This study is to improve the sorting

efficiency of string tomatoes, and solve its false detection and false detection. **Methods:** First, collect the image data set of string tomatoes, expand the data and improve the generalization performance of the model through data enhancement, and 3×3 convolution is replaced by improved SVM-MHSA layer. By replacing softmax classification function in MHSA with SVM classification function which is more suitable for string tomatoes, the detection accuracy of string tomatoes is enhanced. Secondly, the remaining 3×3 convolution is replaced by deep separable convolution to reduce the number of parameters and improve the operation efficiency. Finally, random correction linear unit is introduced to improve the convergence speed of network training.

Results: the test results show that the improved tiny YOLOv5l model can effectively realize the string single fruit recognition and positioning and the whole string fruit counting. The detection frame loss rate is reduced from 1.48% to 1.34%, the target loss rate is reduced from 1.98% to 1.73%, the confidence loss is reduced by 1.4%, the accuracy is increased from 97.36% to 98.89%, and the recall rate is increased from 97.35% to 98.56%.

Conclusion: The tiny YOLOv5l algorithm is more accurate and lightweight. It has a high recognition accuracy in the face of challenges such as occlusion, background interference, illumination change and virtualization, and provides accurate information on the location of single fruit and the quantity of the whole string of fruit for post natal string tomato sorters.

Keywords: deep learning; classification function; skewer tomatoes; complex scenes

基金项目:北京市科技计划(编号:Z201100008020013);北京市农林科学院创新能力建设项目(编号:QNJJ202126, KJCX20200430)

作者简介:赵九霄,男,北京市农林科学院信息技术研究中心助理研究员,硕士。

通信作者:李作麟(1982—),男,北京市农林科学院信息技术研究中心助理研究员,硕士。E-mail:5789663@qq.com

收稿日期:2022-04-12 **改回日期:**2022-08-12

2021 年,中国番茄加工生产量为 480 万 t,其中串型番茄由于其植株承果能力强、萼片质量好、果实品质高等优良的商品特性而受到消费者的广泛认可,但在串型番茄市场需求量增加的同时也带来了一系列挑战:① 需要对产后串型番茄进行分级,分级的前提是需要定位串型番茄位置及每一串的果实数量;② 需在该过程中降低人工分拣带来的损失等^[1]。传统的人工定位、计数方法需要消耗大量劳动力资源,且精确度较低、实时性较差,现代化智慧农业的发展需要一种新型视觉无损检测方法来代替传统的人工定位和计数。

Michael 等^[2]提出了一种机器视觉系统,可以准确地定位目标并计算其数量。该方法使用 Fast R-CNN 框架^[3]实现了高效的检测。周胜安等^[4]提出了一种改进 CenterNet 方法,通过替换骨干网络的方法,提升检测中小残缺块的检测能力。Parico 等^[5]提出了一种先进目标检测模型 YOLOv4 的变体,并对多重目标跟踪算法进行深度排序,构建了应用于移动端的鲁棒实时果实计数器。高芳芳等^[6]提出了基于轻量级目标检测网络 YOLOv4-tiny 和卡尔曼滤波跟踪算法的目标检测与视频计数方法,其平均检测精度达 94.47%。Wang 等^[7]提出了一种检测和计数的两阶段方法。该方法可以准确地检测并计算图像中具有高度邻接度的目标,优于基于手工特征的方法和 YOLOv3 方法^[8-9]。Chen 等^[10]提出了一种基于无人机巡线视频的目标检测计数解决方案,结合 YOLO 快速检测的特点,将卷积神经网络应用于极点状态的图像检测。目前,有关农业中作物定位和计数的研究较少,且研究场景大多数为正常情况下的目标检测问题,并未对实际的产后串型番茄定位与计数过程中出现的果实之间遮挡问题、产后分级地点光线变化、拍摄过程中出现的图像虚化问题以及背景干扰问题进行深入研究。

研究拟提出基于深度学习的改进 Tiny-YOLOv5l 串型番茄定位与计数算法,利用适用于样本数据的注意力机制增加分类准确度;通过加入空洞卷积层增大感受野;通过加入深度可分离卷积层减少模型参数,旨在进一步提高串型番茄检测框架的准确率和实时性。

1 试验方法与材料

1.1 图像采集

试验所用的串型番茄由北京宏福国际农业科技有限公司提供,采集的图像是成熟后的串型番茄,品种为 Juanita,生长高度区间为 1.8~2.4 m,种植间隔 45~60 cm。拍摄设备选用佳能 EOS7D,相机采用双数字影像处理器,像素为 1 800 万像素,图像采集工作从 2021 年 8 月—2022 年 1 月,每月采集 2 次串型番茄图像,共采集

12 批,数据样本共 969 张图片。

1.2 数据集构建

通过对训练样本旋转来实现数据扩增,水平旋转角度范围为 $-35^{\circ}\sim+35^{\circ}$ 。在水平旋转的基础上,加入 20% 随机裁剪和高斯噪声。扩增完成后对新增样本进行筛选,去除劣质样本,将增强后的图像和原图混合融入串型番茄数据集中。所有图像统一使用 JPG 格式,分辨率为 $1\,920\times 1\,080$,标注方式使用 labelme 手动标注,数据集标准为 coco 数据集,分为训练集、验证集、测试集,分割比例为 5:3:2。

如表 1 所示,经数据增强及去除劣质样本后,混合数据集共有 2 280 张图片,数据集包含 5 个类别。① 正常情况:为防止叶片、杆径、果实等对串型番茄的干扰,选取串型番茄背面朝上的样本。② 遮挡条件:分别选取被叶片、茎秆、果实等遮挡情况下的串型番茄样本。③ 背景干扰:选取试验器材干扰物干扰的串型番茄样本。④ 光照变化:选取由强到弱 5 种不同光照条件下的串型番茄样本。⑤ 图片虚化:选取虚化的串型番茄样本。

表 1 数据集中不同挑战图片数量

Table 1 Number of different challenge images in the dataset

场景类型	训练集/张	验证集/张
正常条件	1 280	320
遮挡情况	448	112
背景干扰	204	51
光照变化	184	46
图片虚化	164	41

2 Tiny-YOLOv5l 模型建立

2.1 Tiny-YOLOv5l 模型

YOLOv5l 网络结构由 Darknet-53 主干网络、路径聚合网络(PANet)组成^[11-12]。采用 CSP1_X 作为主干,形成两个支路,X 个 Bottleneck 模块串联组成支路一,卷积层组成支路二,采用这种结构使串型番茄深层特征提取能力增强。PANet 结构是由卷积操作、上采样操作、CSP2_X 构成的循环金字塔结构,可以使串型番茄图像不同特征层之间相互融合,以进行掩模预测,经非极大值抑制(NMS)获得最终串型番茄定位预测框。图 1 为改进的 Tiny-YOLOv5l 框架。

YOLOv5l 的损失函数由边框回归损失、置信度损失以及分类概率损失 3 部分组成,采用完全交并比损失函数(CIoU Loss)实现预测^[13],其计算式为:

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

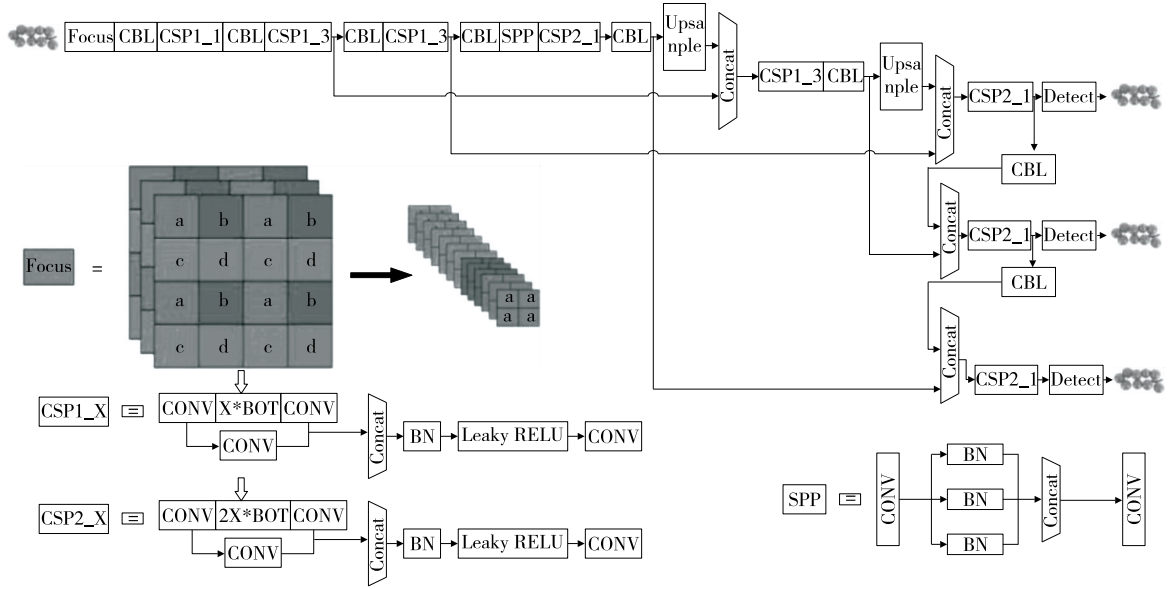


图1 改进的 Tiny-YOLOv5l 框架结构图

Figure 1 Structural diagram of the improved Tiny-YOLOv5l framework

$$a = \frac{v}{(1 - IoU) + v}, \quad (2)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2, \quad (3)$$

$$IoU = \frac{A \cap B}{A \cup B}, \quad (4)$$

式中:

b, b^{gt} —— 预测框和真实框的中心点;

ρ —— 预测框和真实框中心点之间的欧式距离;

c —— 能够同时包含预测框和真实框的最小闭包区域的对角线间距;

w, w^{gt} —— 预测框和真实框的宽度, 像素;

h, h^{gt} —— 预测框和真实框的高度, 像素;

IoU —— 预测框和真实框之间的交并集之比。

2.2 SVM-MHSA

YOLOv5l 包含 Bottleneck CSP 结构, 该结构中的 Bottleneck 层常出现在 ResNet50^[14] 网络结构中, 可以降低深层网络的计算量, bottleneck 结构中第 1 个 1×1 层是降低特征维度, 第 2 个 1×1 层用于提高维度。经过降维和升维后形成类似于沙漏的 Bottleneck 结构, 如图 2(a) 所示。串型番茄个体均为小目标, 且检测环境多为非结构化环境, 果实受到密集分布、遮挡、光照变化等因素影响, 在前向传播过程中果实表层特征信息逐渐减弱, 导致果实的轮廓信息丢失出现漏检情况, 因此加强环境因子影响下的果实目标特征学习尤为重要。在较深的网络层数且需要更高性能的检测网络中使用注意力机制, 其中输入图像的分辨率为 $1\,920 \times 1\,080$ 。考虑到在指数范围内执行自我注意力机制所需内存和计算量, 在主

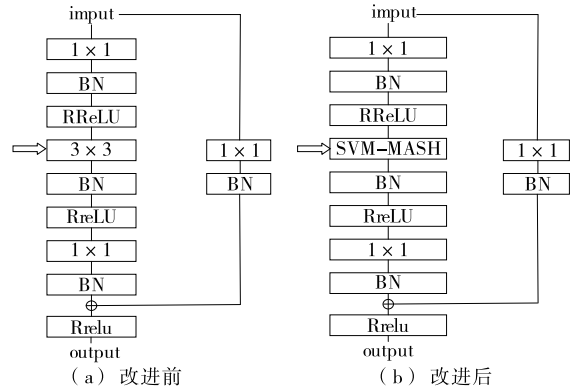


图2 Bottleneck 改进前后结构对比图

Figure 2 Structural comparison diagram before and after Bottleneck improvement

干网络中将深层特征图纳入自我注意力机制模块中。残差网络主干网中的堆栈使用 3 个块, 每个块中有一个 3×3 的空间转换, 使用 MHSA 层进行替换, 构成改进 Bottleneck CSP 体系结构的基础。第 1 个块使用 3×3 的卷积, 步长为 2; 其他两个块步长为 1。对第 1 个改进 CSP 块使用步长为 2 的 2×2 平均池化层, 改进后 Tiny-YOLOv5l 的 bottleneck 层结构如图 2(b) 所示。

支持向量机 (SVM) 的原理是二分类, 其基础是定义在线性空间的间隔最大线性分类器, 同时可以通过核函数解决非线性数据的分类问题, 使用高斯核函数 SVM 进行分类决策。

$$f(x) = \text{sign} \left[\sum_{i=1}^N \alpha_i^* y_i \exp \left(-\frac{\|x - z\|^2}{2\sigma^2} \right) + b^* \right]. \quad (5)$$

SVM 优势: 试验数据集总规模为 969 张, 规模较小, SVM 分类器相较于 softmax 分类器在处理小规模样本方面具有优势; 串型番茄特征相对于背景干扰特征较为明显, 不同于 softmax 函数的精细评分机制, 使用二分类的 SVM 方法只要满足边界值则给出判断, 因此 SVM 分类函数更适用于产后串型番茄; softmax 函数存在过拟合风险, SVM 函数可以解决 softmax 分类器对于非结构环境下串型番茄分类识别能力不足的问题。因此, 使用 SVM 代替 MHSA 结构中的 softmax 分类层, 即提取特征后利用 SVM 分类函数进行分类。多头注意力机制方法可以在不同的投影空间中建立不同的投影信息将多个注意力的方法进行组合, 因此多个注意力的组合效果优于单个注意力机制。SVM-MHSA 机制中, 不同的随机初始化映射矩阵可以将输入向量映射到不同的子空间, 可以使 SVM-MHSA 模型从串型番茄的色泽、大小、位置信息等不同角度理解输入的数据, 通过收集多尺度信息可以使模型从多维度理解串型番茄的表型信息, 以提高检测模型的鲁棒性。

改进后的 SVM-MHSA 层如图 3 所示, 其步骤为:

步骤 1: 初始化算法需要用到的 patch 维度、高度、宽度;

步骤 2: 设置卷积核长度为 1;

步骤 3: 对 query、key、value 值进行 2d 卷积;

步骤 4: 对长和宽进行空间相对位置编码;

步骤 5: 将 nn.SVM 赋值给 self.SVM;

步骤 6: 定义多头注意力机制输入项 x ;

步骤 7: 得到的 $x.size$ 由 batch_size 高度、宽度、通道数组成;

步骤 8: 将长宽拉直后的值、batch_size、通道数三者乘机赋值给 q, k, v ;

步骤 9: 计算 content_content 与 content_new_position 值;

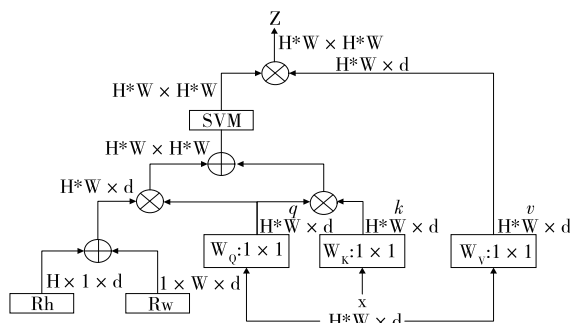


图 3 SVM-MHSA 层结构图

Figure 3 Structure diagram of the SVM-MHSA layers

步骤 10: content_content 与 content_new_position 相加得到的值调用 self.SVM, 最终得到 attention 值;

步骤 11: v 与 attention 乘积得到的值赋值给 out;

步骤 12: 输出。

改进后的 Tiny-YOLOv5l 特征提取图如图 4 所示, Stage2 的 CSP 输出的特征提取图保留了串型番茄的大部分内容信息, 极大程度地突出了串型番茄的轮廓信息, 利用前 3 层浅层特征提取层提取串型番茄的轮廓特征, Stage23 的 CSP 输出的特征提取图突出了串型番茄的深层细节信息, 但对轮廓信息表述较差。使用注意力机制可以较好地结合串型番茄的轮廓信息和细节信息, 使用浅层的轮廓信息恢复最后的检测结果图。

2.3 深度可分离卷积

将 Tiny-YOLOv5l 中 3×3 卷积替换为深度可分离卷积。在保持输入数据集大小稳定的前提下, 改进前卷积计算量 C_1 为:

$$C_1 = D_k \times D_k \times M \times N \times D_F \times D_F, \quad (6)$$

式中:

D_F ——输入与输出特征图的宽度与高度, 像素;

M ——输入特征图的通道数;

N ——输出特征图的通道数;

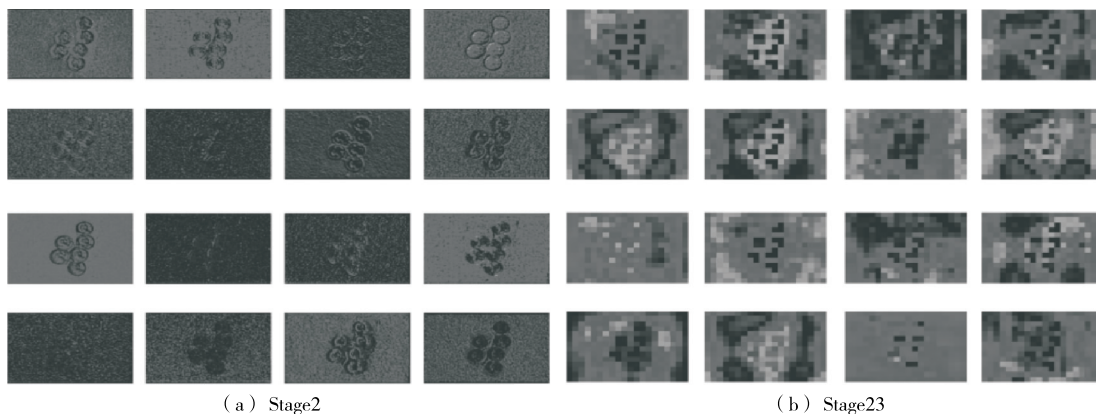


图 4 特征提取图

Figure 4 Feature extraction diagram

D_k ——卷积核的长和宽。

替换可分离卷积后的总计算量 C_2 为:

$$C_2 = D_k \times D_k \times M \times D_F \times D_F + M \times N \times D_F \times D_F. \quad (7)$$

改进后卷积与改进前卷积计算量之比 C_3 为:

$$C_3 = \frac{1}{N} + \frac{1}{D_k^2}. \quad (8)$$

2.4 空洞卷积

在串型番茄检测过程中,对于一些尺寸较小的番茄检测比较困难。由于遮挡、虚化等因素影响,对小番茄的外观特征数据收集匮乏,因而出现漏检、误检的情况。如图 5 所示,空洞卷积可以使卷积囊括更大范围的信息,从而有效利用信息量来提高模型的性能,进而有效捕获边缘特征信息,因此引入空洞卷积来扩大感受野。

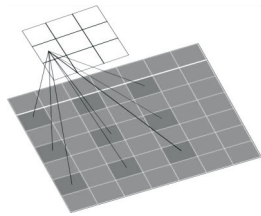


图 5 空洞卷积示意图

Figure 5 Schematic diagram of the void convolution

扩张率为卷积核处理数据时各数据值之间的间距^[15],空洞卷积核和感受野大小为:

$$f_k' = f_k + (f_k - 1) * (r - 1), \quad (9)$$

$$R_m = R_{m+1} + (f_k' - 1) * \prod_{i=1}^{m-1} S_i, \quad (10)$$

式中:

f_k ——原始卷积核大小;

f_k' ——空洞卷积核大小;

r ——卷积核的扩张率;

R_m ——经空洞卷积后第 m 层感受野大小;

S_i ——第 i 层的步长;

$*$ ——卷积。

2.5 Rrelu 激活函数

YOLOv5l 使用的激活函数为 Mish 激活函数与 Leaky_relu 激活函数,Mish 激活函数应用于 Backbone 结构中,Leaky_relu 结构主要应用于 CSP 结构中,Backbone 结构中的 Mish 激活函数不作改变,使用 Rrelu 激活函数^[16],Leaky_relu 表达式为:

$$y_i = \begin{cases} x_i & \text{if } x_i \geq 0 \\ \frac{x_i}{a_i} & \text{if } x_i < 0 \end{cases}, \quad (11)$$

式中:

a_i —— $(1, +\infty)$ 区间内的固定参数。

使用 Rrelu 激活函数代替 Leaky_relu 激活函数,其表

达式为:

$$y_{ji} = \begin{cases} x_{ji} & \text{if } x_{ji} \geq 0 \\ a_{ji} x_{ji} & \text{if } x_{ji} < 0 \end{cases}$$

where

$$a_{ji} \sim U(l, u), l < u \text{ and } l, u \in [0, 1). \quad (12)$$

3 模型训练与测试

3.1 模型运行平台

训练采用的 CPU 为(英特尔)Intel(R) Core(TM) i5-11400,内存为 32 G,操作系统为 Windows 10,显卡为 RTX3060,拥有 12 G 独显。

3.2 参数设置及训练步骤

(1) 配置模型训练所需文件结构及路径。

(2) 用 Anaconda3 配置虚拟环境,调用环境所需 python 版本以及 pytorch 等;根据不同数据集规模及难易程度调整模型阈值参数,将改进后的模型导入。

(3) 总训练迭代次数为 300 轮,为使梯度下降学习法有更好的效率,初始学习率为 0.01。根据串型番茄数据集大小设置每步骤更新公式为:

$$\alpha = 0.95^{\text{epoch_num}} \cdot \alpha_0, \quad (13)$$

式中:

α ——实时学习率;

α_0 ——初始学习率;

epoch_num——epoch 次数。

3.3 评价指标

为了评价 Tiny-YOLOv5l 的性能,主要使用精确度^[17]、召回率^[18]等评价指标。召回率和精度分别按式(14)、式(15)计算。

$$R = \frac{T_P}{T_P + F_N}, \quad (14)$$

$$P = \frac{T_P}{T_P + F_P}, \quad (15)$$

式中:

R ——召回率;

P ——精确度;

T_P ——准确检测出的串型番茄目标数;

F_P ——误检出的串型番茄目标数;

F_N ——漏检的串型番茄目标数。

4 结果与分析

4.1 监测识别结果

在同一数据集上分别训练 YOLOv5l 以及改进的 Tiny-YOLOv5l 两种算法,在正常拍摄条件下改进前的算法检测置信度为 93%,改进后的 Tiny-YOLOv5l 算法对目标检测的置信度为 94%,检测并无显著提高。

由于拍摄角度不同导致取图过程中存在较多遮挡问题(果实被果实遮挡、果实被叶遮挡等)。如图 6(a)所示,蓝色框为 YOLOv5l 算法漏检的串型番茄,其中未识别的



图 6 图片虚化下算法改进前后对比图

Figure 6 The algorithm is improved before and after the picture blur comparison diagram

遮挡串型番茄占未识别总数的 77.8%；如图 6(b)所示，Tiny-YOLOv5l 算法可以 100% 识别所有目标区域中的串型番茄，且置信度方面较 YOLOv5l 提高了 1%~3%。背景干扰下，如图 6(c)中黄色框所示，无关环境因子被误检为串型番茄；如图 6(d)所示，Tiny-YOLOv5l 算法可以识别遮挡面积超过 70% 的串型番茄。选择 5 种光照环境进行对比检测试验。图 6(e)中绿色框为 YOLOv5l 算法漏检的串型番茄，图 6(f)表明 Tiny-YOLOv5l 算法均能有效识别串型番茄个体。在图片虚化情况下，图 6(g)表明 YOLOv5l 算法识别到的串型番茄数量为 0；Tiny-YOLOv5l 算法在加入 Bottleneck 替换层后[图 6(h)]，加强了对模糊特征的学习，可以检测出全部串型番茄。

4.2 模型性能评估

4.2.1 检测速度 由图 7 可知，Tiny-YOLOv5l 算法较 YOLOv5l 所需检测时间少，即使在使用普通电脑配置（酷睿 i5, 16 GB 内存）训练参数规模是 3.76×10^7 的网络模型，训练一步所需时间仅为 8~10 s，总训练时长为 12.77 h。在正常拍摄、遮挡情况、背景干扰、光照变化和图像虚化情况下收敛速度分别缩减，说明在不利拍摄情况下，改进后的算法表现出更好的计算速率。

4.2.2 损失率 由图 8(a)可知，当训练轮数 < 13 时，YOLOv5l 的损失率比 Tiny-YOLOv5l 的低；当训练轮数

增加至 210 时，YOLOv5l 的损失率从 11.13% 收敛至 1.91%，Tiny-YOLOv5l 的损失率从 11.98% 收敛至 1.93%，Tiny-YOLOv5l 的降低幅度明显高于 YOLOv5l；当训练轮数 > 250 时，两种算法的损失率降低幅度逐渐趋于稳定，模型训练效果良好，未出现梯度消失或梯度爆炸现象，Tiny-YOLOv5l 的最终损失率为 1.41%，低于 YOLOv5l 的。由图 8(b)可知，训练开始时 Tiny-YOLOv5l 算法检测框损失率为 10.74%，低于 YOLOv5l 的，随着迭代次数的增加，当迭代次数 > 226 时，损失率趋于稳定，最终 Tiny-YOLOv5l 算法的损失率为 0.65%，优于 YOLOv5l 算法的。

由图 9(a)可知，当训练轮数 < 15 时，改进前算法损失率从 11.9% 上升至 17.83% 后又回落至 10.32%，改进后算法处于不稳定状态，损失率由 8.93% 上升至 14.83% 后又回落至 7.27%；当训练轮数 > 15 时，两种算法均趋于稳定，且 YOLOv5l 算法的目标损失率波动幅度大于 Tiny-YOLOv5l。由图 9(b)可知，Tiny-YOLOv5l 算法的目标损失率趋于稳定，最终稳定在 1.37%，优于 YOLOv5l 算法的。

由图 10 可知，当训练轮数 < 30 时，改进前后算法的置信度损失情况相似，当训练轮数 > 143 时，无论是训练集还是验证集，Tiny-YOLOv5l 算法具有更低的置信度损失率，比 YOLOv5l 低 0.14%。

4.2.3 精度 由图 11 可知，改进前的 YOLOv5l 算法的检测精度稳定于 98.89%，高于改进后的 Tiny-YOLOv5l 算法。

4.2.4 召回率 由图 12 可知，初始时，两种算法的召回率均 < 20%，因为串型番茄形状大小不一，训练开始时漏检情况较多。当训练轮数 < 210 时，两种算法的召回率均在波动上升，而 Tiny-YOLOv5l 的上升趋势较 YOLOv5l 更具优势；当训练轮数 > 210 时，二者召回率均趋近于 1，且 Tiny-YOLOv5l 的召回率高于 YOLOv5l。

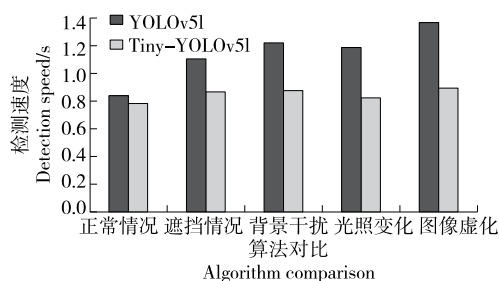


图 7 改进前后算法检测速度比较

Figure 7 Comparison of the detection speed between before and after the improved algorithm

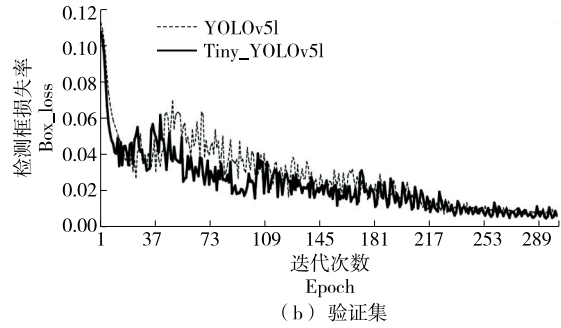
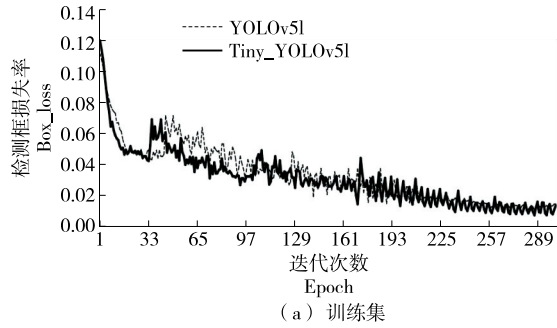


图 8 改进前后检测框损失率对比图

Figure 8 Improve the detection frame loss rate before and after the improvement

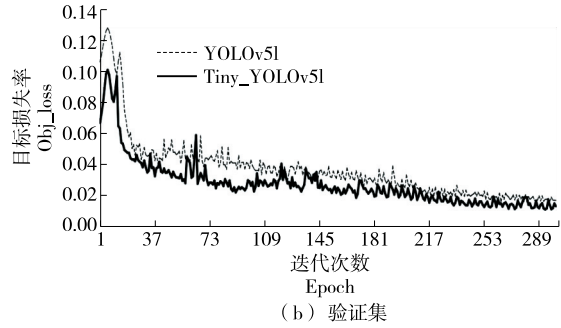
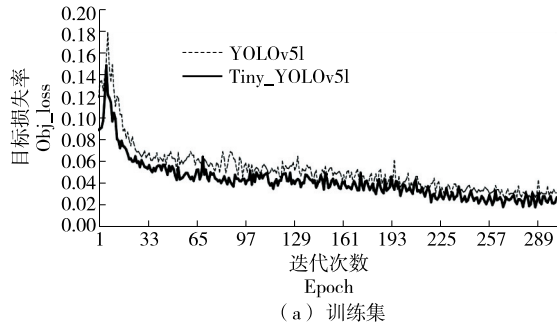


图 9 改进前后目标损失率对比图

Figure 9 Improve before and after the target loss rate comparison chart

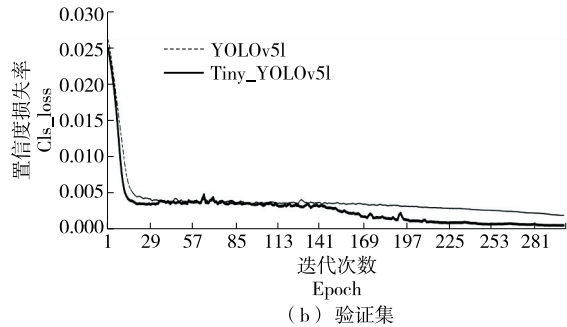
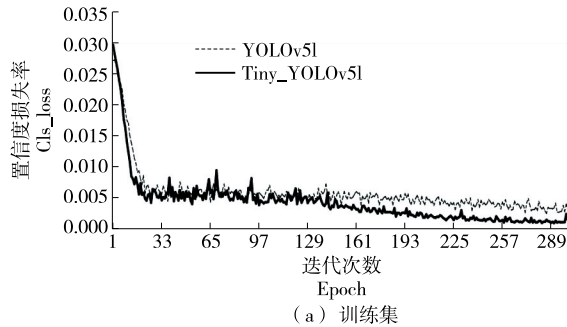


图 10 改进前后训练集和验证集置信度损失率对比图

Figure 10 Comparison of confidence loss rates between before and after training and validation sets

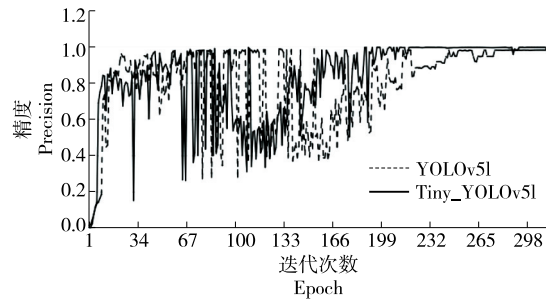


图 11 改进前后精度对比图

Figure 11 Improve the accuracy before and after the accuracy comparison map

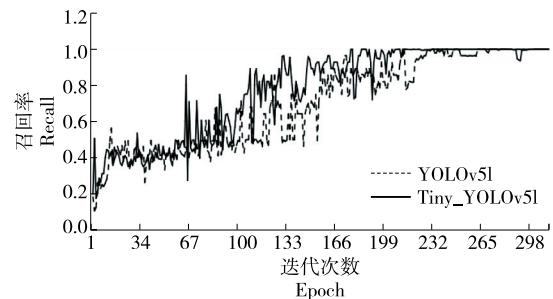


图 12 改进前后召回率对比图

Figure 12 Recall rate comparison chart before and after the improvement

5 结论

研究建立了一种串型番茄实时检测与计数算法模型。在模型中加入注意力机制可以加强对串型番茄目标

识别的准确率;采用 SVM 分类函数代替 softmax 分类函数,可使之更适用于样本数据,并拥有更好的泛化能力;加入空洞卷积增大对样本特征的感受野可以提高检测准

确率;在模型中引入深度可分离卷积,可以减少参数量以增加实时性;使用随机纠正线性单元激活函数可以增加网络训练的收敛速度。在串型番茄检测过程中,计算量仍是一项挑战,如何在保持精度的前提下优化模型是待解决的问题之一;检测的品种种类过于单一,如何提高泛化性能也是一项重要的工作;此外,后续将对串型番茄色泽、残缺种类等其他参数进行研究分类。

参考文献

- [1] 焦方圆, 申金媛, 郝同盟. 一种基于卷积神经网络的烟叶等级识别方法[J]. 食品与机械, 2022, 38(2): 222-227.
JIAO F Y, SHEN J Y, HAO T M. A method of tobacco leaf grade recognition based on convolutional neural network [J]. Food & Machinery, 2022, 38(2): 222-227.
- [2] MICHAEL H, CHRISTOPHER M C, SIMON D, et al. Fruit quantity and ripeness estimation using a robotic vision system[J]. IEEE Robotics & Automation Letters, 2018, 3(4): 2 995-3 002.
- [3] ZHANG Q S, REN J, HUANG G, et al. Mining interpretable AOG representations from convolutional networks via active question answering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(11): 3 949-3 963.
- [4] 周胜安, 黄耿生, 张译匀, 等. 基于深度学习的水果缺陷实时检测方法[J]. 食品与机械, 2021, 37(11): 123-129.
ZHOU S A, HUANG G S, ZHANG Y J, et al. Real time detection method of fruit defects based on deep learning [J]. Food & Machinery, 2021, 37(11): 123-129.
- [5] PARICOA I B, AHAMED T. Real time pear fruit detection and counting using YOLOv4 models and deep SORT[J]. Sensors, 2021, 21(14): 4 803.
- [6] 高芳芳, 武振超, 索睿, 等. 基于深度学习与目标跟踪的苹果检测与视频计数方法[J]. 农业工程学报, 2021, 37(21): 217-224.
GAO F F, WU Z C, SUO R, et al. Apple detection and counting using real-time video based on deep learning and object tracking[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(21): 217-224.
- [7] WANG G, STAPPEN G V, BAETS B D. Automated detection and counting of Artemia using U-shaped fully convolutional networks and deep convolutional networks [J]. Expert Systems with Applications, 2021, 171(4): 114562.
- [8] 陈科羽, 时磊, 刘博迪, 等. 改进 YOLOv3 的输电线路绝缘子检测方法[J]. 科技创新与应用, 2021, 11(34): 79-82, 86.
CHEN K Y, SHI L, LIU B D, et al. Improved YOLOv3 insulator detection method for high voltage transmission and transformation lines[J]. Technology Innovation and Application, 2021, 11(34): 79-82, 86.
- [9] LI Y, HAN Z, XU H, et al. YOLOv3-lite: A lightweight crack detection network for aircraft structure based on depthwise separable convolutions[J]. Applied Sciences, 2019, 9(18): 3 781.
- [10] CHEN B, MIAO X. Distribution line pole detection and counting based on YOLO using UAV inspection line video[J]. Journal of Electrical Engineering and Technology, 2020, 15(1): 441-448.
- [11] LIU Z, WANG K, DONG H, et al. A cross-modal edge-guided salient object detection for RGB-D image [J]. Neurocomputing, 2021, 454(5): 168-177.
- [12] 杨学存, 和沛栋, 陈丽媛, 等. 基于深度可分离卷积的轻量级 YOLOv3 输电线路鸟巢检测方法[J]. 智慧电力, 2021, 49(12): 88-95.
YANG X C, HE P D, CHEN L Y, et al. Bird's nest detection on lightweight YOLOv3 transmission line based on deep separable convolution[J]. Power Grid Analysis & Study, 2021, 49(12): 88-95.
- [13] 王静, 孙紫雲, 郭苹, 等. 改进 YOLOv5 的白细胞检测算法[J]. 计算机工程与应用, 2022, 58(4): 134-142.
WANG J, SUN Z Y, GUO P, et al. Improved leukocyte detection algorithm of YOLOv5[J]. Computer Engineering and Applications, 2022, 58(4): 134-142.
- [14] ELPELTA G M, SALLAM H. Automatic prediction of COVID19 from chest images using modified ResNet50[J]. Multimedia Tools and Applications, 2021, 80(17): 26 451-26 463.
- [15] 袁帅, 王康, 单义, 等. 基于多分支并行空洞卷积的多尺度目标检测算法[J]. 计算机辅助设计与图形学学报, 2021, 33(6): 864-872.
YUAN S, WANG K, SHAN Y, et al. Multi-scale object detection method based on multi-branch parallel dilated convolution [J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(6): 864-872.
- [16] 刘若愚, 刘立波. 基于改进全卷积网络模型的肺结节检测[J]. 激光与光电子学进展, 2020, 57(16): 174-183.
LIU R Y, LIU L B. Detection of pulmonary nodules based on improved full convolution network model [J]. Laser & Optoelectronics Progress, 2020, 57(16): 174-183.
- [17] 张宏鸣, 武杰, 李永恒, 等. 多目标肉牛进食行为识别方法研究[J]. 农业机械学报, 2020, 51(10): 259-267.
ZHANG H M, WU J, LI Y H, et al. Recognition method of feeding behavior of multi-target beef cattle[J]. Transactions of the Chinese Society of Agricultural Machinery, 2020, 51(10): 259-267.
- [18] SHEWELL C, NUGENT C, DONNELLY M, et al. Indoor localisation through object detection within multiple environments utilising a single wearable camera[J]. Health & Technology, 2017, 7(1): 51-60.