

李 颀, 杨 军. 基于多分辨率特征融合的葡萄尺寸检测[J]. 江苏农业学报, 2022, 38(2): 394-402.

doi: 10.3969/j.issn.1000-4440.2022.02.013

基于多分辨率特征融合的葡萄尺寸检测

李 颀¹, 杨 军²

(1. 陕西科技大学电子信息与人工智能学院, 陕西 西安 710021; 2. 陕西科技大学电气与控制工程学院, 陕西 西安 710021)

摘要: 针对葡萄特征提取不够充分且果粒排列密集相互遮挡难以准确检测的问题, 以陕西省鄠邑区户太 8 号葡萄为研究对象, 提出一种基于特征金字塔网络(FPN)特征融合的 Faster R-CNN 卷积神经网络模型完成复杂背景情况下葡萄果粒的检测与识别。以 ResNet50 为主干网络, 引入金字塔结构, 增强网络模型对葡萄果粒不同分辨率特征的提取能力, 同时加入 GA-RPN 网络生成自适应锚框, 引入遮挡补偿机制, 以解决密集葡萄果粒存在的遮挡问题。模型验证结果表明, 本研究提出的模型精度均值(*AP*) 在候选框与原标记框的重叠率(*IOU*) 阈值为 50 时可达 95.9%, 对葡萄果粒、果穗的检测准确率分别为 95.8%、96.1%, 相比于原始 Faster R-CNN 模型识别性能更优。利用双目视觉算法对葡萄果粒进行尺寸测量, 在最佳测量距离(0.6~1.4 m) 其相对误差可控制在 2% 以内。

关键词: 葡萄无损检测; 多分辨率特征融合; 遮挡补偿; 机器视觉

中图分类号: TP391 **文献标识码:** A **文章编号:** 1000-4440(2022)02-0394-09

Grape size detection based on multi-resolution feature fusion

LI Qi¹, YANG Jun²

(1. School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China; 2. School of Electrical and Control Engineering, Shaanxi University of Science and Technology, Xi'an 710021, China)

Abstract: In view of the insufficient feature extraction of grapes and the difficulty in accurate detection of dense occlusion of grape grains, a Faster R-CNN convolutional neural network model based on feature pyramid network(FPN) feature fusion was proposed to detect and identify grape grains under complex background in Hui District of Shaanxi Province. The ResNet50 was used as the main network, and the pyramid structure was introduced to enhance the extraction ability of the network model for different resolution characteristics of grape fruit. At the same time, the GA-RPN network was added to generate the adaptive anchor frame, and the occlusion compensation mechanism was introduced to solve the occlusion problem of dense grape fruit. The model verification results showed that the average precision (*AP*) of the proposed model could reach 95.9% when the threshold of overlap rate between candidate box and original marked box (*IOU*) was 50, and the detection accuracy of grapes and strings was 95.8% and 96.1%, respectively, which was better than that of the original Faster R-CNN model. Using binocular vision algorithm to measure the size of grape seeds, the relative error can be controlled within 2% at the optimal measurement distance

(0.6~1.4 m).

Key words: nondestructive testing of grapes; multi-resolution feature fusion; occlusion compensation; machine vision

收稿日期: 2021-07-06

基金项目: 陕西省农业科技创新工程资助项目[201806117YF05NC13 (1)]; 陕西省科技厅农业科技攻关项目(2015NY028); 陕西科技大学博士科研启动基金资助项目(BJ13-15)

作者简介: 李 颀(1973-), 女, 陕西西安人, 博士, 教授, 主要研究方向为农业智能化信息化、深度学习。(E-mail) liqidq@sust.edu.cn

通讯作者: 杨 军, (E-mail) 1414831375@qq.com

农产品信息快速无损检测是现代农作物精细管理的关键技术之一^[1-4]。传统的小规模种植主要通

过农户肉眼观察田间作物信息以及经验来判断作物的生长状况,难以适应现代化大规模的作物种植。计算机与图像处理技术在国内快速发展^[5-7],被广泛应用于售卖阶段的农产品检测和分级,但在生长状况的实时监测方面应用较少。2016年叶浩等^[8]利用图像数据拟合和像素转换得到实际葡萄果粒尺寸,实现对成穗葡萄等级的自动分类检测,但没有对葡萄果穗进行识别测量。2017年罗陆锋等^[9]建立分界线几何求解与计算模型,完成贴叠葡萄果穗目标的提取,但没有进行葡萄果粒的重叠分割。2019年刘智杭^[10]提出基于改进的图像分割算法和随机变换椭圆检测算法的果粒尺寸检测方法,使用了颜色特征权向量优化 k-means 算法,实现了复杂背景下葡萄果穗图像的分割,但在果粒严重遮挡时的检测结果误差较大。2020年刘平等^[11]通过提取贴叠葡萄果穗图像轮廓的有效属性点信息进行位置判断,区分出重叠轮廓,完成葡萄果穗的识别和分割,但没有对重叠葡萄果粒进行研究。

目前多数基于机器视觉的作物特征提取主要集中在应用于瓜果缺陷等亚健康外观状态以及尺寸检测^[12-15],但在葡萄尺寸检测方面的应用研究还比较少,检测模型容易受到葡萄自然生长条件下诸如复杂背景变化、葡萄果粒尺寸不一、随机遮挡、颜色相似等问题的影响。针对上述问题,本研究使用机器视觉技术对葡萄的本体特征进行提取,通过对 Faster R-CNN 卷积神经网络算法进行改进,实现葡萄果粒、果穗的准确定位与识别,结合双目视觉算法实现葡萄果粒尺寸特征的非接触无损微变检测。

1 材料与方法

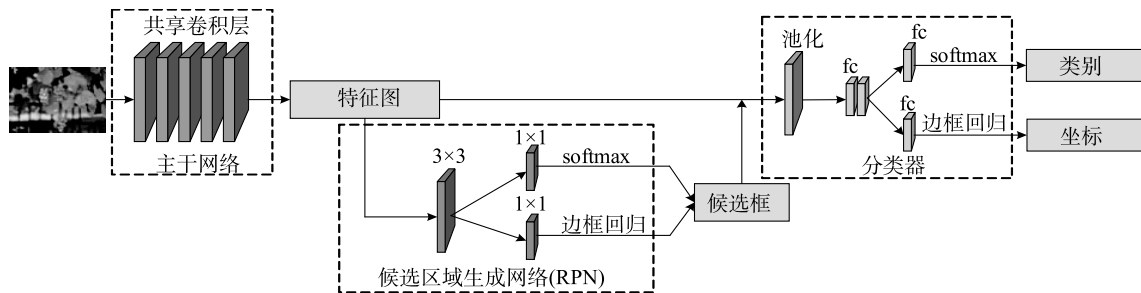
1.1 数据采集

试验样本采集于西安市鄠邑区南部石井镇的户太8号葡萄种植基地,采集时间为2020年4月下旬至2021年4月中旬,采集对象为温室大棚内生长的葡萄果穗。拍摄包含不同角度、距离及光照下的图片,各采集1 000张,使用图像标注工具 LabelImg 对其进行标注,然后通过数据增强对图片进行放大、缩小、旋转、翻转处理,共获得21 600张图片及对应的标注文件,分别随机选择训练集10 500张,测试集300张。

1.2 目标检测网络选择

现有目标检测算法分2大类^[16-18]:一类是基于候选区域建议的二阶段检测算法,如 Faster R-CNN^[19]、SPP-Net 等,首先通过区域建议算法生成含有目标的候选区域,再通过后续网络进行目标分类和位置回归;另一类是基于回归的检测算法,如 YOLOv3^[20]、SSD 等,这类算法不需要生成候选区域,而是直接将预设的滑动窗口映射到生成的特征图上,通过深度特征回归出窗口的类别和位置偏移量,最终完成目标的检测定位。

Faster R-CNN 算法网络结构如图1所示,候选区域生成环节以牺牲时间为代价,Faster R-CNN 虽然没有 YOLOv3 速度快,但是网络的整合度较高,极大提高了检测精度和效率。由于葡萄果粒较小,且生长过程中其本体特征的变化比较迟缓,且对检测的实时性要求不高。因此选择 Faster R-CNN 算法进行葡萄本体特征的检测。



1×1、3×3 分别代表卷积核大小,fc 指全连接层,softmax 指归一化指数函数。

图1 Faster R-CNN 网络结构图

Fig.1 Network structure diagram of Faster R-CNN

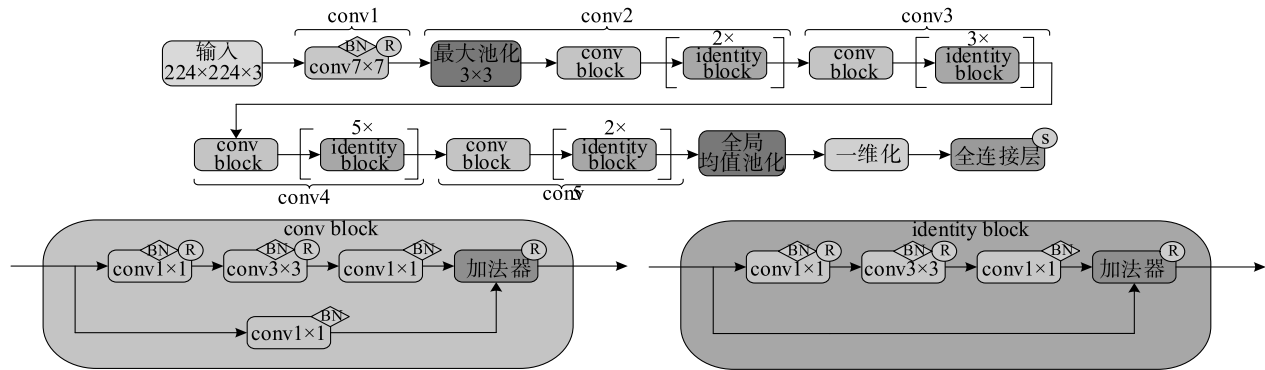
1.3 主干网络改进与多层特征融合

Faster R-CNN 主干网络使用 VGG16,其最大的特点是使用了3×3 的卷积核以及2×2 的池化核,每

个池化层都将图像缩小为原来的 1/2。残差网络 ResNet50 结构如图2 所示,可以解决因卷积层数增多带来的梯度耗散或者爆炸的问题^[21],其中 conv

block 表示添加尺度的残差块,卷积 conv2~conv5 中的 identity block 2×、3×、5×代表的是 2 个、3 个、5 个不改变尺寸的残差块。加入激活函数 R(ReLu) 和批

量正则化处理 BN(Batch Norm) 共同解决反向传播的梯度问题,使得网络在具有较好识别精度的同时还具有优秀的实时性^[22]。

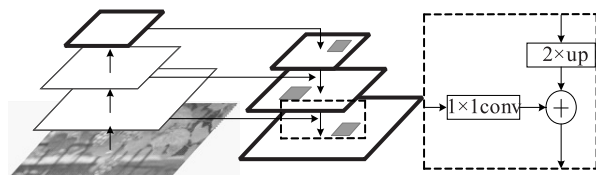


conv1~conv5 表示卷积块 1~卷积块 5,conv1×1 表示卷积核大小,其余同理。2×、3×、5×表示模块数,BN(Batch Norm)表示批量正则化处理,R(ReLu)表示激活函数,S 表示 sigmoid 函数。

图 2 ResNet50 网络结构图

Fig.2 Network structure diagram of ResNet50

对于大棚环境中葡萄的检测,存在不同图像之间、甚至同一张图像里需要被检测出的目标大小相对于整张图像的比例差异较大的问题,图片中葡萄果粒图像的平均占比小于 0.043,葡萄果穗的平均占比大于 0.257。这种尺度差异带来的挑战性,严重限制检测器的整体表现。网络层数越深,表达抽象特征的能力也就越强,但浅层的空间信息也相应丢失。这就导致深层特征图无法提供细粒度的空间信息对目标进行精确定位。同时,小尺寸目标的语义信息也在下采样的过程中逐渐丢失。特征金字塔网络 FPN(Feature pyramid network)综合利用高分辨率的位置信息和高级抽象语义信息,在所有特征上输出预测结果。FPN 网络结构图如图 3 所示,随着层级越来越高,语义信息越来越丰富,红色边界更宽。



1×1conv 表示卷积核大小,2×up 表示 2 倍上采样。

图 3 FPN 网络结构图

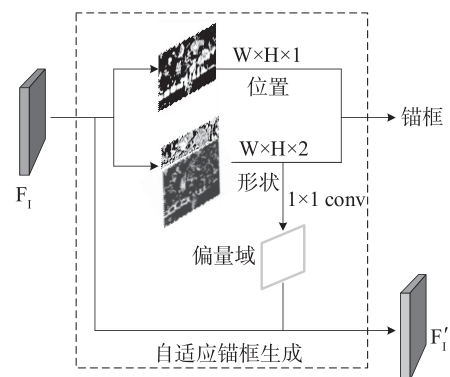
Fig.3 Network structure diagram of FPN

1.4 锚框自适应调整

在 RPN 网络中需要事先根据检测目标的大小

设置多种固定大小和形状的 anchor,然后利用滑动窗口对特征图进行扫描,根据扫描结果对目标进行检测和分类。但这种方法只适用于检测一般形状和大小的目标,对于长宽比悬殊的目标不一定能准确检测,影响模型的检测性能,降低算法的鲁棒性。

GA-RPN 网络结构如图 4 所示,在特征图后面增加 2 条支路用于生成 anchor,与传统滑窗方式不同,一条用于定位目标区域中心点,另外一条用于预测符合目标形状的检测框的最佳长宽比。同时在形状预测分支后面还增加了特征自适应模块,将 anchor 形状信息通过 1×1 的卷积操作融入到特征图中,以解决不同形状大小与特征图不匹配的问题。



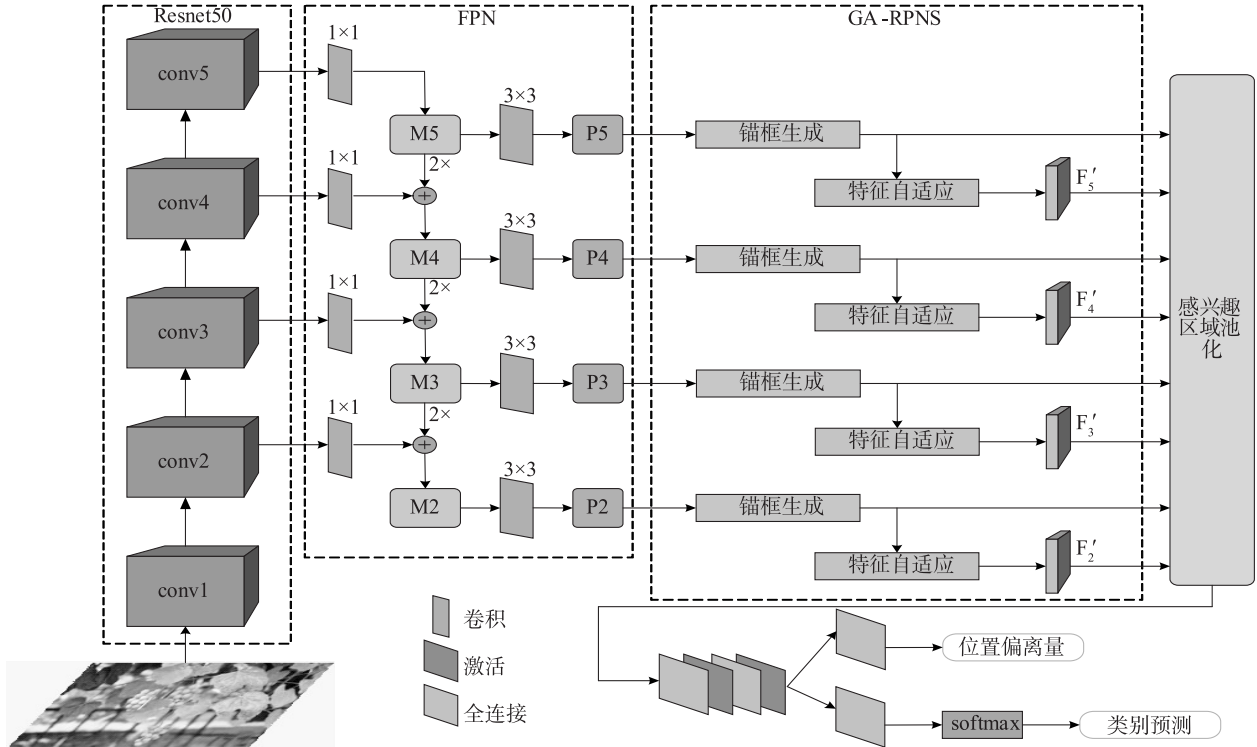
F_1 、 F_1' 分别表示输入、输出特征图。

图 4 GA-RPN 网络结构图

Fig.4 Network structure diagram of GA-RPN

FPN 提取 Resnet50 网络每个阶段最后的特征层作为一个特征层级进行多分辨率特征的融合,分辨率的增加可以获得更多小尺寸葡萄果粒的信息,

同时引入 GA-RPN 网络完成自适应 anchor 的生成,改进的 Faster R-CNN 网络结构如图 5 所示。



Resnet50 表示残差网络,FPN 表示金字塔结构,GA-RPN 表示自适应锚框生成网络,conv1~conv5 表示卷积块,1×1、3×3 表示卷积核大小,2×表示 2 倍上采样,M2-M5、P2-M5 分别表示对应层的特征图,softmax 指归一化指数函数。

图 5 改进的 Faster R-CNN 网络结构图

Fig.5 Network structure diagram of improved Faster R-CNN

1.5 密集葡萄果粒遮挡检测

自然生长条件下的葡萄果粒、果穗难免存在相互遮挡、重叠的关系。葡萄果粒的遮挡情况通常可以分为 3 种:类外遮挡,即其他物体对葡萄果穗的遮挡,如叶片、树枝等遮挡;类内遮挡,目标葡萄果穗上的果粒被其他葡萄果穗遮挡;类间遮挡,目标葡萄果穗上果粒之间的互相遮挡,也称为密集遮挡。其中,葡萄果粒之间的密集遮挡在所有遮挡问题中最常见且占比最大,葡萄果粒密集遮挡示例如图 6 所示。

采用遮挡标记和遮挡补偿来共同解决这一问题,在标注的过程中对于被遮挡的目标也进行完整标注,同时加入遮挡标记,提高模型的通用性。对存在遮挡情况的标注框计算遮挡补偿系数,融入损失函数里进行训练,增强模型对被遮挡葡萄果粒的敏感性,提高模型检测精度。

添加遮挡标记,当葡萄果粒之间存在遮挡情况

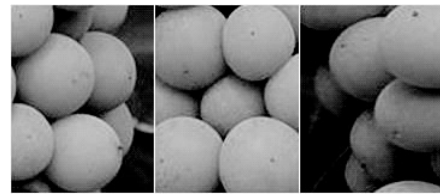


图 6 密集葡萄果粒遮挡示例图

Fig.6 Example diagram of dense grape grain occlusion

时,向被遮挡果粒添加标记信息 $occ_mutual(i)$,将遮挡共分为 3 个等级,当果粒被遮挡区域的占比不超过 25%时, $i=3$;占比 25%~75%时, $i=2$;占比超过 75%时, $i=1$ 。计算遮挡补偿系数 $occ_m(s)$ 如式(1)~式(4)所示, s 为第 s 个标注框。

$$occ_m(s) = \frac{e^{leave(s)} \cdot cross_area(s)}{(e^1 + e^2 + e^3) \cdot union_area} \quad (1)$$

$$cross_area(s) = | (GT_1 \cap GT_s) \cup \dots \cup (GT_n \cap$$

$$GT_s) | \quad (2)$$

$$union_area = |GT_1 \cup GT_2 \cup \dots \cup GT_n| \quad (3)$$

$$leave(s) = \begin{cases} 1, in(s) = '1' \\ 2, in(s) = '2' \\ 3, in(s) = '3' \end{cases} \quad (4)$$

$in(s)$ 代表标记信息对应的 i , GT_s 代表第 s 个真实标注框, $union_area(s)$ 表示第 s 个标注框与所有标注框的相交面积和, $union_area$ 表示当前图片所有标注框的相并面积和, $leave(s)$ 为第 s 个标注框对应的被遮挡等级, n 为当前图片中真实标注总数。将遮挡补偿信息融入在损失函数里来补偿因遮挡造成的目标特征缺失问题, 损失函数定义如式(5)所示。

$$Loss = Loss_{IOU} + Loss_{REP} \quad (5)$$

损失如式(6)、式(7)所示, \bar{C}_i 是预测的有无果粒的置信度, \bar{C}_i 是真实检测的置信度, 有果粒为 1, 无果粒为 0; λ_{no} 是 IOU 误差的权重, 取值同上; $\eta(i)$ 为遮挡补偿系数。

$$Loss_{IOU} = \sum_{i=0}^B l_i [C_i - \eta(i) \bar{C}_i]^2 + \lambda_{no} \sum_{i=0}^B l_i^{no} (C_i - \bar{C}_i)^2 \quad (6)$$

$$\eta(i) = \begin{cases} 0, & \text{无遮挡} \\ occ_m(s), & \text{有遮挡} \end{cases} \quad (7)$$

REP 损失包含 3 个目标: 1 个吸引项和 2 个排斥项, 吸引项的要求是预测框靠近其目标, 而排斥项是需要预测框远离周围其他的标注框和预测框, 如式(8)~式(11)所示。attr 为吸引项, $Smooth_{L_1}$ 作为其回归损失函数, 让预测框更接近葡萄果粒。repgt 的目标是让预测框最大程度远离附近其他果粒的真实标注框, 然后对 IOG 进行 $Smooth_{L_n}$ 损失, 防止预测框向其他临近的真实果粒偏移。repbox 的目标是让预测框最大程度远离附近其他果粒的预测框, 降低检测器对非极大值抑制(NMS)的敏感度, 提高在密集场景下的鲁棒性。

$$L_{rep} = L_{attr} + 0.5 \times L_{repgt} + 0.5 \times L_{repbox} \quad (8)$$

$$L_{attr} = \frac{\sum_{P \in P_+} Smooth_{L_1}(B^P, G_{attr}^P)}{|P_+|} \quad (9)$$

$$L_{repgt} = \frac{\sum_{P \in P_+} Smooth_{L_n}[IOG(B^P, G_{rep}^P)]}{|P_+|} \quad (10)$$

$$L_{repbox} = \frac{\sum_{i \neq j} Smooth_{L_n}[IOG(B^{P_i}, B^{P_j})]}{\sum_{i \neq j} K[IOG(B^{P_i}, B^{P_j}) > 0] + \epsilon} \quad (11)$$

其中, $IOU = area(P \cap G) / area(P \cup G)$, $IOU = area(P \cap G) / area(G)$, $area(P \cap G)$, 表示预测框 P 与真实标注 G 交集的面积, $area(P \cup G)$ 表示 P 与 G 并集的面积, G_{attr}^P 为与预测框匹配的真实标注, G_{rep}^P 为除与预测框匹配后的剩余 IOU 最大的真实标注, $P_+ = P_1 \cap P_2 \cap \dots \cap P_{|K|}$, 是所有正样本的集合, σ 取值范围为 0~1, $Smooth$ 回归函数如式(12)、式(13)所示。

$$Smooth_{L_1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (12)$$

$$Smooth_{L_n}(x) = \begin{cases} -\ln(1-x), & x \leq \sigma \\ \frac{x-\sigma}{1-\sigma} - \ln(1-x), & x > \sigma \end{cases} \quad (13)$$

1.6 双目视觉测量

单目系统常用于固定位置、特定距离下的目标测量, 优势在于成本较低, 对计算资源的要求不高, 系统结构相对简单, 但是无法对非标准障碍物进行判断, 实际中的测量需要真正的空间三维坐标^[23-25]。图 7 所示为一般情况下的双目视觉成像原理图。

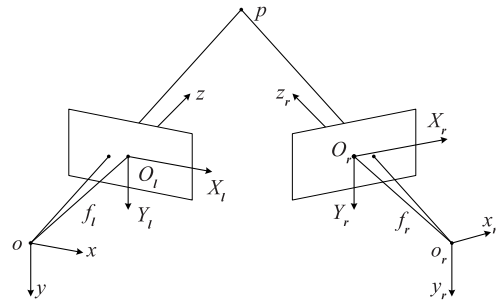


图 7 双目立体视觉成像原理图

Fig.7 Principle diagram of binocular stereo vision imaging

以左相机坐标系 o_xyz 为原点建立世界坐标系, 其中, $o_l-X_l Y_l$ 表示图像坐标系, o_l 为投影点, 有效焦距为 f_l , 右相机同理, 建立相机透视变换模型如式(14)、式(15)。

$$\begin{bmatrix} X_l \\ Y_l \\ 1 \end{bmatrix} = \begin{bmatrix} f_{lx} & 0 & c_{lx} \\ 0 & f_{ly} & c_{ly} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (14)$$

$$\begin{bmatrix} X_r \\ Y_r \\ 1 \end{bmatrix} = \begin{bmatrix} f_{rx} & 0 & c_{rx} \\ 0 & f_{ry} & c_{ry} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} \quad (15)$$

两个坐标系 o_xyz 与 $o_r-x_r y_r z_r$ 之间的相对关系如式(16)所示, 由推导可得, 空间点的三维坐标可

以表示为式(17)所示。

$$\begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = M_{lr} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 & t_x \\ r_4 & r_5 & r_6 & t_y \\ r_7 & r_8 & r_9 & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (16)$$

$$\begin{cases} x = zX_l/f_l \\ y = zY_l/f_l \\ z = \frac{f_l(f_l t_x - X_r t_z)}{X_r(r_7 X_l + r_8 Y_l + r_9 f_l) - f_r(r_1 X_l + r_2 Y_l + r_3 f_l)} \end{cases} \quad (17)$$

已知双目相机基线距离,左右相机焦距 f_l, f_r 和点 P 在左、右相机中的图像坐标已知,再通过相机标定求出的 R 和 T 就可以得到点 P 相对于相机的三维坐标,实现视野范围内任意大小的葡萄特征的测量。

如图8所示,图像拍摄设备为KS4A418-D双目摄像头,3.6 mm定焦镜头,分辨率为 $2\,592 \times 1\,944$,图像格式为MJPG。左右相机空间位置标定结果如表1所示,旋转矩阵的标定结果近似于单位矩阵,说明左、右相机之间的旋转角度很小,光轴接近平行。平移向量表明两个相机的光轴在 x 轴方向上相距59.842 25 mm,接近实际基线距离60 mm,而在 y 和 z 方向上仅存在微小的偏移。



图8 KS4A418-D 双目摄像头

Fig.8 KS4A418-D binocular camera

表1 空间位置标定结果

Table 1 Calibration results of space position

旋转矩阵	平移向量
$\begin{bmatrix} 0.999\,80 & 0.016\,51 & -0.010\,71 \\ -0.016\,51 & 0.999\,86 & 4.677\,24e^{-4} \\ 0.010\,72 & -2.907\,12e^{-4} & 0.999\,94 \end{bmatrix}$	$[59.842\,25 \quad 0.270\,34 \quad 0.115\,86]$

2 结果与分析

2.1 模型训练效果

试验平台的运行环境配置:操作系统 Ubuntu16.04,显卡 NVIDIA GeForce RTX 2080 Ti,算力 7.5,运行内存 128 GB,使用 Python 语言编程,选择 TensorFlow 作为深度学习框架。采用近似联合训练方法进行端到端训练,在保证训练精度的同时提高训练

速度。模型训练参数设置如表2所示,训练结果 Loss 变化如图9所示。从图9中可以看出,网络在前4 000次迭代中迅速拟合,之后趋于稳定,平均损失在0.15左右浮动。模型训练效果良好,最后收敛到稳定值,达到了预期的训练效果。

表2 本研究改进模型的训练参数

Table 2 Training parameters of improved model in this study

参数	数值
Iterations	9 000
Snapshot	100
Learning rate	0.000 1
Batch size	64
Momentum	0.949
Weight decay	0.000 5

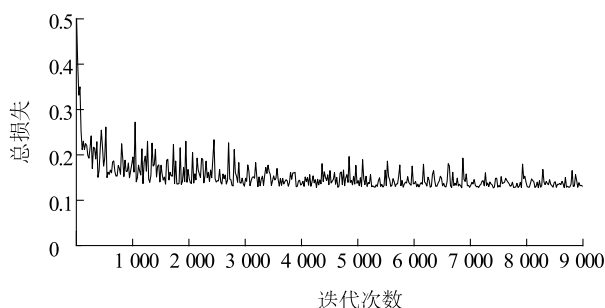


图9 本研究改进模型的训练总损失曲线

Fig.9 Total loss curve of improved model in this study

2.2 遮挡补偿试验结果对比

分别使用原始 Faster R-CNN 模型与引入遮挡补偿机制后的 Faster R-CNN 模型对葡萄果粒进行检测,结果(图10)表明,改进后的模型检测精度高于原始 Faster R-CNN 模型,其中对于葡萄果穗中间被遮挡的果粒检测效果的提升最明显,边缘被遮挡果粒的检测效果也有很大的提升,可以准确识别遮挡面积不超过75%的果粒。果粒检测结果对比示例如图10所示,可以看出,引入遮挡补偿机制后葡萄果穗边缘果粒的漏检情况有明显改善,且中间被漏检的果粒也可以准确识别出来。拍摄多组葡萄果穗图片进行检测,与人眼能观察到的实际果粒数量进行对比分析,结果如表3所示。

2.3 模型检测结果分析

温室大棚场景内葡萄果粒、果穗的检测结果如图11所示,相机距离葡萄0.6~1.4 m。本研究模型可以有效地检测出葡萄果粒、果穗两类目标,尺寸较小的

葡萄果粒也能被精确地检测出。在出现轻微遮挡的情况下,比如树枝、叶片遮挡葡萄果穗,也可以检测出目标,如图 11(b)左侧第 1 个葡萄果穗检测框。

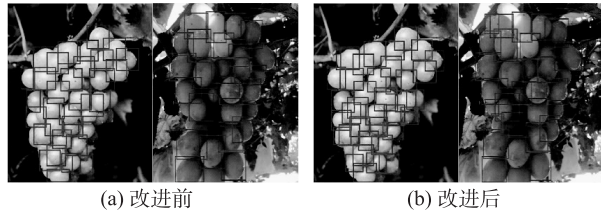


图 10 改进前与改进后模型的葡萄果粒检测结果对比

Fig.10 Comparison of grape fruit detection results between the model before and after improvement

表 3 改进前与改进后模型果粒数量检测结果对比

Table 3 Comparison of the number of grape fruit detection results between the model before and after improvement

组别	实际 果粒数	原始模型 检测结果	误差 (%)	引入遮挡 补偿机制后 模型检测 结果	误差 (%)
1	50	39	22.00	44	12.00
2	47	39	17.02	42	10.64
3	35	30	14.29	31	11.43
4	26	16	38.46	21	19.23
5	78	60	23.08	71	8.97
6	49	37	24.49	44	10.20

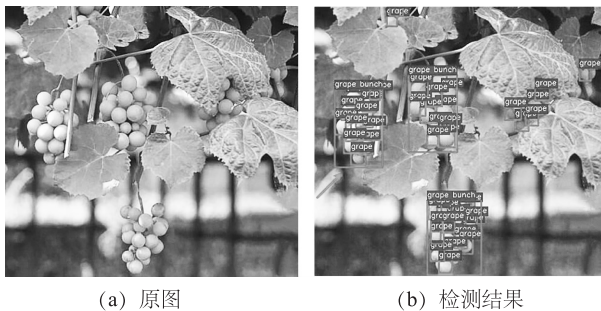


图 11 本研究改进模型实际检测结果

Fig.11 Actual detection results of the improved model in this study

分别将 5 种模型 YOLOv3、Faster R-CNN、Faster R-CNN+ResNet50、Faster R-CNN+FPN 以及本研究改进模型在葡萄果粒、果穗数据集上进行试验评估,并进行结果分析。在多分类任务中,常通过使用混淆矩阵来进行指标评价。本研究中识别种类较少,使用 P-R (Precision-Recall) 曲线来进行指标评价,直接显示分类结果的优劣。

图 12 所示为 P-R 曲线,为了评估算法有效性,计算各个类别在检测阈值 (IOU) 为 50 时的平均精度值 (AP)、不同阈值下的 mAP 和检测速度 (FPS)。AP 是衡量单一类别检测效果的重要指标,可以通过计算准确率 (Precision) 与召回率 (Recall) 曲线所包围的面积得到,如式 (18)~式 (20) 所示。mAP 是所有类别的精度均值,如式 (21) 所示。其中, T_p 表示真的正样本; F_p 表示假的正样本,负样本被错误地预测为正样本; F_n 表示假的负样本,正样本被错误地预测为负样本。表 4 为 5 种模型的精度均值比较结果。

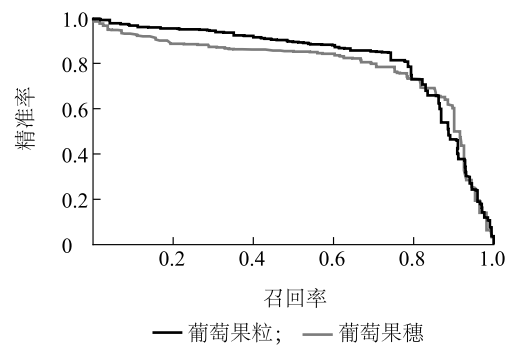


图 12 本研究改进模型检测结果的 P-R 曲线

Fig.12 The P-R curve of the improved model in this study

$$Precision = \frac{T_p}{T_p + F_p} \times 100\% \quad (18)$$

$$Recall = \frac{T_p}{T_p + F_n} \times 100\% \quad (19)$$

$$AP = \int_0^1 (r) dr \quad (20)$$

$$mAP = \frac{\sum_{i=1}^2 AP_i}{2} \quad (21)$$

2.4 葡萄果粒测量结果分析

对模型检测出的葡萄果粒图像进行遍历,找出可以定义果粒横径和纵径的特征点。首先从左往右、从上到下找到的第 1 个点记为 a 点,然后从右往左、从上到下找到的第 1 个点记为 b 点,即找到葡萄果粒的最左侧点和最右侧点。考虑到左右两点可能不在同一水平线上,所以横径大小通过计算 a 、 b 两点投影在 x 轴方向上的距离得到。同理,再分别从上往下和从下往上找到最上方点 c 和最下方点 d ,纵径大小通过计算 c 、 d 两点投影在 y 轴方向上的距离得到。横、纵径特征点示意图如图 13 所示。根据公式 (8) 计算得出 a 、 b 、 c 、 d 4 点的空间三维坐标。设特征点坐标为 $a(X_a, Y_a, Z_a)$, $b(X_b, Y_b, Z_b)$, $c(X_c,$

$Y_c, Z_c), d(X_d, Y_d, Z_d)$, 则果粒横径、纵径的计算如式(22)、式(23)所示, 其中 L 为横径, T 为纵径。

表 4 不同模型测试结果比较

Table 4 Comparison of test results of different models

模型	$AP_{50}(\%)$		$mAP_{50}(\%)$	$mAP_{75}(\%)$	$AP(\%)$	FPS
	葡萄果粒	葡萄果穗				
YOLOv3	62.5	76.2	69.3	39.8	45.4	20
Faster R-CNN	80.4	81.6	81.0	42.7	46.1	7
Faster R-CNN+ResNet50	83.1	84.2	83.6	53.3	67.9	5
Faster R-CNN+FPN	85.7	89.5	87.6	59.1	71.5	6
本研究模型	95.8	96.1	95.9	65.3	84.8	5

AP_{50} 为 IOU 阈值为 50 时每一类的平均精度, mAP_{50} 、 mAP_{75} 分别为 IOU 阈值为 50、75 时所有类别的平均精度, AP 指所有类别下的平均精度, FPS 指模型每秒处理的帧数。

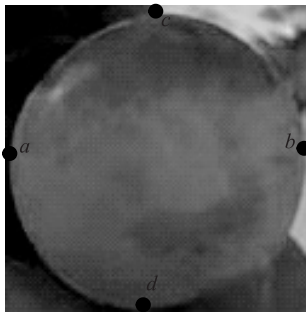


图 13 葡萄果粒横、纵径特征点位置示意图

Fig.13 Location of characteristic points of transverse and longitudinal diameter of grape fruit

$$L = \sqrt{(X_a - X_b)^2 + (Z_a - Z_b)^2} \quad (22)$$

$$T = \sqrt{(Y_c - Y_d)^2 + (Z_c - Z_d)^2} \quad (23)$$

为验证遮挡补偿效果,根据果粒图像被遮挡的面积大小占比不同,进行遮挡补偿试验。分别选取遮挡程度为小于 25%、25%~50%、50%~75%、大于 75%的果粒进行试验,使用游标卡尺测量出葡萄果粒横径的实际值,结果如表 5 所示。为验证本研究模型的实际效果,选取 3 个葡萄果穗分别在不同拍摄距离下进行试验,以葡萄果穗为单位计算横径平均值,果粒测量结果如表 6 所示。

由表 5 可知,引入遮挡补偿机制后,检测误差明显降低,更接近果粒实际大小。当遮挡程度超过 75%时,2 个果粒被漏检,只有尺寸较大的果粒被检测出且误差高于遮挡程度小于 75%的果粒。由表 6 可知,在相机有效拍摄范围内实际检测结果最大误差不超过 4%,当测量距离在 0.6~1.4 m 时误差更小且不超过 2%。因此,本研究提出的模型在最佳测量距离 0.6~1.4 m 对于葡萄果粒尺寸特征的自动测量具有较好的实用性。

表 5 不同遮挡程度葡萄果粒横径检测结果

Table 5 Detection results of grape fruit transverse diameter under different shading degrees

遮挡程度 (%)	果粒横径实际值检测值 (mm)	未加遮挡补偿机制		加入遮挡补偿机制	
		检测值 (mm)	误差 (%)	检测值 (mm)	误差 (%)
25	27.33	28.52	4.36	28.01	2.49
	21.54	20.54	-4.64	21.84	1.37
	25.61	26.61	3.92	25.12	-1.91
25~50	24.37	25.75	5.67	23.96	-1.68
	29.55	28.11	-4.89	28.79	-2.57
	26.18	27.53	5.17	26.74	2.15
50~75	27.42	24.23	-11.62	26.66	-2.77
	25.86	22.19	-14.18	26.56	2.69
	22.16	19.23	-13.24	22.99	3.76
75	27.44	-	-	-	-
	26.56	-	-	-	-
	29.73	23.63	-20.51	31.64	6.41

-表示没有检测到果粒。

3 结论与讨论

本研究提出了一种结合深度学习和视觉测量的葡萄果粒、果穗特征的自动检测模型。通过双目相机拍摄获取数据集,然后利用改进的 Faster R-CNN 算法进行检测,得到目标位置和类别信息。以 ResNet50 为主干网络并引入金字塔结构,增强网络模型在不同分辨率下葡萄果实特征的提取能力,对候选区域生成网络 RPN 进行改进,加入了自适应锚框,同时引入遮挡补偿机制,以解决密集葡萄果粒由于遮挡而被漏检的情况。利用标定结果求出特征点的三维坐标,实现果粒尺寸自动测量。试验结果表明,本研究提出的模型在相机最佳测量距离内对葡萄果粒尺寸的自动测

量具有较高的准确率,能满足葡萄表型参量特征精准测量的需要,也可为其他种类作物的表型参量三维测量提供参考与借鉴,有较好的实用性。在后续的研究中,会对检测模型进行压缩优化,进一步提高检测的实时性和准确率,同时加入更多葡萄表型参量的测量,以更好地分析作物的生长状态。

表 6 不同拍摄距离下葡萄果粒横径检测结果

Table 6 Detection results of grape fruit transverse diameter at different shooting distances

拍摄距离 (m)	横径实际值 (mm)	横径测量值 (mm)	误差 (%)
0.2~0.6	22.11	21.64	-2.13
	24.68	25.12	1.78
	29.34	28.78	-1.91
0.6~1.0	26.43	26.12	-1.17
	27.45	27.02	-1.57
	23.84	23.59	-1.05
1.0~1.4	25.88	25.61	-1.04
	25.46	25.05	-1.61
	22.47	22.78	1.38
1.4~1.8	24.26	24.88	2.56
	23.34	23.87	2.27
	24.77	25.19	1.70
1.8~2.2	28.73	27.76	-3.38
	25.44	25.02	-1.65
	26.91	25.98	-3.46

参考文献:

- [1] 张帅亮. 作物生长过程检测技术研究[D].新乡:河南科技学院,2018.
- [2] 杨涛,李晓晓. 机器视觉技术在现代农业生产中的研究进展[J].中国农机化学报,2021,42(3):171-181.
- [3] ZHAO Y S, GONG L, HUANG Y X, et al. A review of key techniques of vision-based control for harvesting robot[J]. Computers and Electronics in Agriculture, 2016,127(6):312-321.
- [4] XIONG J T, LIN R, LIU Z, et al. The recognition of litchi clusters and the calculation of picking point in a nocturnal natural environment[J]. Biosystems Engineering, 2018,166(8):44-57.
- [5] 李诗禹. 温棚环境及植物生长信息管理分析系统研究与实现[D].银川:宁夏大学,2018.
- [6] 赵献立,王志明. 机器学习算法在农业机器视觉系统中的应用[J].江苏农业科学,2020,48(12):226-231.
- [7] 翁杨,曾睿,吴陈铭,等. 基于深度学习的农业植物表型研究综述[J].中国科学:生命科学,2019,49(6):698-716.
- [8] 叶浩,詹国敏,吴亮,等. 基于图像分割的葡萄果粒径快速测量技术[J].新技术新工艺,2016(5):84-87.
- [9] 罗陆锋,邹湘军,王成琳,等. 基于轮廓分析的双串叠贴葡萄目标识别方法[J].农业机械学报,2017,48(6):15-22.
- [10] 刘智杭. 基于图像的巨峰葡萄果粒尺寸检测[D].哈尔滨:东北林业大学,2019.
- [11] 刘平,朱衍俊,张同勋,等. 自然环境下贴叠葡萄串的识别与图像分割算法[J].农业工程学报,2020,36(6):161-169.
- [12] 李睿. 基于机器视觉的苹果尺寸测量与分类研究[D].唐山:华北理工大学,2020.
- [13] 徐庆. 图像处理技术在现代农业中的应用[J].农村经济与科技,2020,31(1):66-67.
- [14] QURESHI W S, PAYNE A, WALSH K B, et al. Machine vision for counting fruit on mango tree canopies[J]. Precision Agriculture, 2017,18(2):224-244.
- [15] TANG Y C, CHEN M Y, WANG C L, et al. Recognition and localization methods for vision-based fruit picking robots: A review[J]. Frontiers in Plant Science, 2020,5:1-17.
- [16] 李伟山,卫晨,王琳. 改进的 Faster R-CNN 煤矿井下行人检测算法[J].计算机工程与应用,2019,55(4):200-207.
- [17] 张守东,杨明,胡太. 基于多特征融合的显著性目标检测算法[J].计算机科学与探索,2019,13(5):834-845.
- [18] 吕晓琪,吴凉,谷宇等. 基于三维卷积神经网络的低剂量肺结节检测[J].光学精密工程,2018,26(5):1211-1218.
- [19] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017,39(6):1137-1149.
- [20] SOHAIB A, MAHEDI H, SHEIKH M S I. YOLO-based three-stage network for Bangla license plate recognition in Dhaka Metropolitan City[J]. International Conference on Bangla Speech and Language Processing (ICBSLP), 2018,33(15):21-22.
- [21] 毋立芳,汪敏贵,付亨等. 深度目标检测与图像分类相结合的棉花发育期自动识别方法[J].中国科技论文,2018,13(20):2309-2316.
- [22] 张怡,赵珠蒙,王校常,等. 基于 ResNet 卷积神经网络的绿茶种类识别模型构建[J].茶叶科学,2021,41(2):261-271.
- [23] 王珂,王亮亮,李瑞峰,等. 基于折光鱼眼全景相机的双目立体视觉标定[J].中国科技论文,2016,11(8):913-919.
- [24] 卢艳军,王诗宇,张太宁,等. 针对目标跟踪过程的双目视觉测距方法研究[J].电光与控制,2020,27(1):60-63,101.
- [25] MANSOOR A, MUHAMMAD S A, MUHAMMAD R, et al. Real-time machine-learning based Crop and weed detection and classification for variable-rate spraying in precision agriculture[J]. International Conference on Electrical and Electronics Engineering (ICEEE), 2020,10(19):273-280.

(责任编辑:张震林)