doi:10.6041/j.issn.1000-1298.2020.02.026

基于卷积神经网络的草莓识别方法

刘小刚　范　诚　李加念　高燕俐　章宇阳　杨启良
(昆明理工大学农业与食品学院, 昆明 650500)

摘要: 针对目前草莓识别定位大多在简单环境下进行、识别效率较低的问题,提出利用改进的 YOLOv3 识别方法在复杂环境中对草莓进行连续识别检测。通过训练大量的草莓图像数据集,得到最优权值模型,其测试集的精度均值(MAP)达到 87.51%;成熟草莓的识别准确率为 97.14%,召回率为 94.46%;未成熟草莓的识别准确率为 96.51%,召回率为 93.61%。在模型测试阶段,针对夜晚环境下草莓图像模糊的问题,采用伽马变换得到的增强图像较原图识别正确率有显著提升。以调和平均值(F)作为综合评价指标,对比多种识别方法在不同果实数量、不同时间段及视频测试下的实际检测结果,结果表明,YOLOv3 算法 F 值最高,每帧图像的平均检测时间为 34.99 ms,视频的平均检测速率为 58.1 f/s,模型的识别正确率及速率均优于其他算法,满足实时性要求。同时,该方法在果实遮挡、重叠、密集等复杂环境下具有良好的鲁棒性。

关键词: 草莓; 图像识别; 目标检测; YOLO; RetinaNet; 深度学习
中图分类号: TP391.41　　**文献标识码**: A　　**文章编号**: 1000-1298(2020)02-0237-08　　**OSID**:

Identification Method of Strawberry Based on Convolutional Neural Network

LIU Xiaogang　FAN Cheng　LI Jianian　GAO Yanli　ZHANG Yuyang　YANG Qiliang
(*Faculty of Agriculture and Food*, *Kunming University of Science and Technology*, *Kunming* 650500, *China*)

Abstract: Aiming to solve the problems that strawberry identification and localization were mostly carried out in a simple environment and the identification efficiency was low, the continuous identification and detection of strawberry in a complex environment was studied, and an improved YOLOv3 identification method was proposed. By training a large number of strawberry image data sets, the optimal weight model was obtained. The mean average precision (MAP) of the test set reached 87.51%, among which the average accuracy and the recall rate of mature strawberries was 97.14% and 94.46%, respectively, and that of immature strawberries was 96.51% and 93.61%. In the model test stage, aiming at the problem of strawberry image blurring in the night environment, the recognition accuracy of the original image was significantly improved by using Gamma transform image enhancement. The harmonic mean value (F value) was used as the comprehensive evaluation index, and the actual test results of various identification methods under different fruit numbers, time periods and video tests were compared. The results showed that the improved YOLOv3 algorithm had the highest F value, the average detection time of the picture was 34.99 ms, and the average detection frame rate of the video was 58.1 f/s, indicating that the recognition accuracy and rate of the model were better than that of other algorithms, and it had good robustness in complex environments such as fruit occlusion, overlap and density. This study can provide theoretical basis for continuous operation of strawberry picking robot under actual working conditions.

Key words: strawberry; image recognition; target detection; YOLO; RetinaNet; deep learning

0　引言

智能图像识别与目标检测具有识别精度高、检测速度快等优点。将图像识别系统与机器人结合起来,用于农业生产中,将极大地提高工作效率与作业质量。快速稳定的目标识别系统能够使采摘

收稿日期: 2019-06-11　修回日期: 2019-09-16
基金项目: 国家自然科学基金项目(51769010、51979133、51469010)
作者简介: 刘小刚(1977—),男,教授,博士,主要从事农业信息化工程研究,E-mail: liuxiaogangjy@126.com

机器人进行长时间有效的工作,大大提高了采摘效率。

草莓营养价值丰富,被誉为“水果皇后”。草莓的自动识别是机器人采摘的前提,目前国内外草莓图像识别已经取得了一定的进展[1]。张铁中等[2]采用 LRCD 的方法分割草莓图像,从而确定草莓的位置。谢志勇等[3]通过图像分割对其进行 Hough 变换,实现成熟草莓的识别。研究多采用基于图像分割对草莓进行识别定位,但均未对动态机器视觉下研究的可行性进行验证。张凯良等[4]应用软硬件相结合的方式,以地垄栽培模式下的草莓为作业对象,采用镜像匹配法和激光辅助对草莓进行自动定位,但是该方法平均计算效率较低。李长勇等[5]利用双目定位技术识别完全成熟的草莓,但没有分析复杂环境下该方法的识别正确率。ZHANG 等[6]基于机器视觉的颜色模型对水果进行识别检测。侯贵洋等[7]采用颜色识别和特征识别相结合的方法区分成熟草莓和未成熟草莓,但均未结合实际环境对草莓的识别效率进行分析。上述研究大都针对草莓在简单环境下的识别定位,且研究方法的识别效率较低,对草莓在实际复杂环境中的识别效率鲜有研究。

深度卷积神经网络通过对输入图像进行卷积运算,直接从图像中提取到特定目标的信息特征,并通过前向传播与反向传播调整卷积神经网络运算中的参数。卷积神经网络在对目标进行特征提取时,不要求原始图像具有固定的尺寸,从而不需要对原始图像进行剪裁或压缩等操作,提取到的特征信息更加完整和准确[8-9]。目前深度卷积神经网络主要有两大类,一类是 RCNN[10]、Fast - RCNN[11]、Faster - RCNN[12]、Mask - RCNN[13]等基于候选区域的方法,另一类是 SSD[14]、RetinaNet[15]、YOLO 等基于无候选区域的方法。REDMON 等[16-17]提出了基于一步检测的 YOLO 算法,该算法不同于 RCNN 为代表的两步检测算法,使得网络结构更加简单,同一条件下检测速度比 Faster - RCNN 快 10 倍左右。YOLOv2[18]在 YOLOv1[19]的基础上加入了批量归一化,仅通过输入高分辨率的图像,MAP 提高了 4%[20]。薛月菊等[21]采用 YOLOv2 方法在光照多样性及复杂背景下识别未成熟芒果,具有较高的准确率及召回率,但未说明该方法在夜间识别的效率。赵德安等[22]基于 YOLO 算法在复杂背景下实现多种算法对苹果的识别,验证了在不同环境下 YOLOv3 算法的可行性,但未具体分析在动态检测下目标的检测效率是否变化。可见,基于深度卷积神经网络的 YOLO 识别方法在复杂环境下具有效率高、可靠性好的特点。

针对目前草莓识别存在的问题,基于卷积神经网络下改进的 YOLOv3 算法,本文对草莓静态图像和动态视频进行识别分析,探究草莓在实际复杂环境下机器视觉的识别效率,为草莓采摘机器人识别系统提供理论基础。

1 材料和方法

1.1 检测对象数据采集与标注

图像采集于云南省昆明市呈贡区某草莓种植园,大棚地垄式种植。草莓的品种为章姬,成熟草莓为鲜艳的红色,未成熟的草莓为绿色或是粉色。采集样本的时间分为白昼、傍晚和夜晚。草莓同一茎上会生长形状大小相似的成熟草莓和未成熟草莓,为降低采摘机器人误采摘的概率,将识别对象分为成熟草莓和未成熟草莓两类。

为满足样品环境的多样性,采集白天、傍晚、夜晚、遮挡、重叠、密集环境下的草莓图像,共 620 幅,其中 500 幅图像用于模型训练,120 幅图像用于模型测试。草莓图像采用 LabelImg 进行人工手动标注,对原始图像进行裁剪、旋转及镜像等操作从而扩充数据集,共标注 1 320 幅图像,保存为 PASCAL VOC 格式。标注的矩形框用于草莓的分类和识别。

使用 2 400 万像素的摄像头采集动态视频数据,采用固定位置拍摄、横向拍摄和纵向拍摄 3 种拍摄方式,拍摄环境包含果实重叠、密集、遮挡等复杂情况。视频统一为 10 s 短视频,识别后保存为 avi 格式,每 3 s 截取一幅图像显示对比效果。抽取每类视频 20 个,重复 3 次进行识别分析。

1.2 基于 YOLO 的图像识别算法

YOLOv1 算法采用一步检测,网络结构浅且速度快,但对小目标识别准确度上存在不足[23]。YOLOv2 在 YOLOv1 的基础上移除了全连接层,采用卷积层和先验框使得模型更容易学习,特征层由原来的 7×7 变为 13×13,对于一般的物体,能够更好地预测边界框。YOLOv3 在 YOLOv2 基础上,增加了多尺度预测,对小目标的检测精度进一步提升,通过卷积层提取目标信息,采用池化层降低维度,从而加快模型运算速度。

YOLOv3 算法的训练测试流程如图 1 所示。

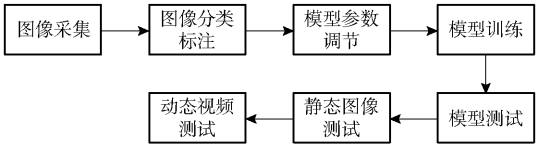


图 1 YOLOv3 训练测试流程

Fig. 1 YOLOv3 training test process

1.2.1　YOLOv3 草莓检测网络与训练

采用 Darknet53 作为预训练模型，训练时设置动量项（Momentum）为 0.9，衰减参数（Decay）为 0.000 5，初始学习率（Learning rate）为 0.001。试验平台 CPU 为 Intel core i7－8750H，主频 2.20 GHz，6 核 12 线程处理器，GPU 为 GeForce GTX 1060，6 GB 显存，8 GB 运行内存，Windows 10 专业版。本文所有模型训练及测试都在同一硬件环境下运行。

由表 1[24] 可知，交并比（Intersection over union，IOU）为 0.5 时，YOLOv3 算法有 3 种像素训练模式，分别为 320 像素 ×320 像素、416 像素 ×416 像素、608 像素 ×608 像素。相对于 320、416 像素，608 像素的平均识别精度最高，相对于 608 像素，416 像素检测时间更短，采用 416 像素 ×416 像素模式。为了对不同模型的性能进行评价，采用准确率 P_i、召回率 R_n、调和平均值 F_j[25]、平均精度（MAP）作为评价指标，计算公式为

$$P_i = \frac{R_{S_i}}{R_{S_i} + F_{S_i}} \times 100\% \quad (i = 1,2) \tag{1}$$

$$R_n = \frac{R_{S_n}}{R_{S_n} + F_{N_n}} \times 100\% \quad (n = 1,2) \tag{2}$$

$$F_j = \frac{2P_jR_n}{P_j + R_n} \quad (j = n = 1,2) \tag{3}$$

$$M_{AP} = \frac{1}{S}\sum_{k=i}^{N} P(k)R(k) \tag{4}$$

式中　R_S——图像中正确检测出的目标草莓数量

F_S——图像中未被检测到的草莓数量

F_N——非草莓物体数量

S——草莓的类别数

k——阈值

N——引用阈值的数量

M_{AP}——成熟草莓和未成熟草莓的平均精度

其中，i、j = 1 表示成熟草莓的相关参数，i、j = 2 表示未成熟草莓的相关参数。

表 1　不同方法的性能比较

Tab. 1　Performance comparison of different methods

方法	平均精度/%	检测时间/ms
SSD321	45.4	61
DSSD321	46.1	85
R－FCN	51.9	85
SSD513	50.4	125
DSSD513	53.3	156
FPN FRCN	59.1	172
RetinaNet－50－500	50.9	73
RetinaNet－101－500	53.1	90
RetinaNet－101－800	57.5	198
YOLOv3－320	51.5	22
YOLOv3－416	55.3	29
YOLOv3－608	57.9	51

YOLOv3 算法共有 53 层卷积层，考虑到本文只识别草莓，过多的卷积层会增加计算量，因此对 YOLOv3 神经网络结构进行改进，参照 VGG－16 搭建如图 2 所示的 14 层网络结构，当输出矩阵分别为 52 ×52 ×64、26 ×26 ×128 时，神经网络并行输出 52 ×52 ×192 和 26 ×26 ×384 的两条网络特征层，最后将 3 种特征层的输出张量合并处理对草莓进行检测。改进后的神经网络减少了卷积层数量及过多的运算，而且对小目标的检测有显著效果。改进网络结构如图 2 所示。图 3 为 YOLOv3 边界框（Bounding boxes）原理图[26]，c_x、c_y 为预测框中心点所在的单元格相对左上角单元格偏移的单元格数

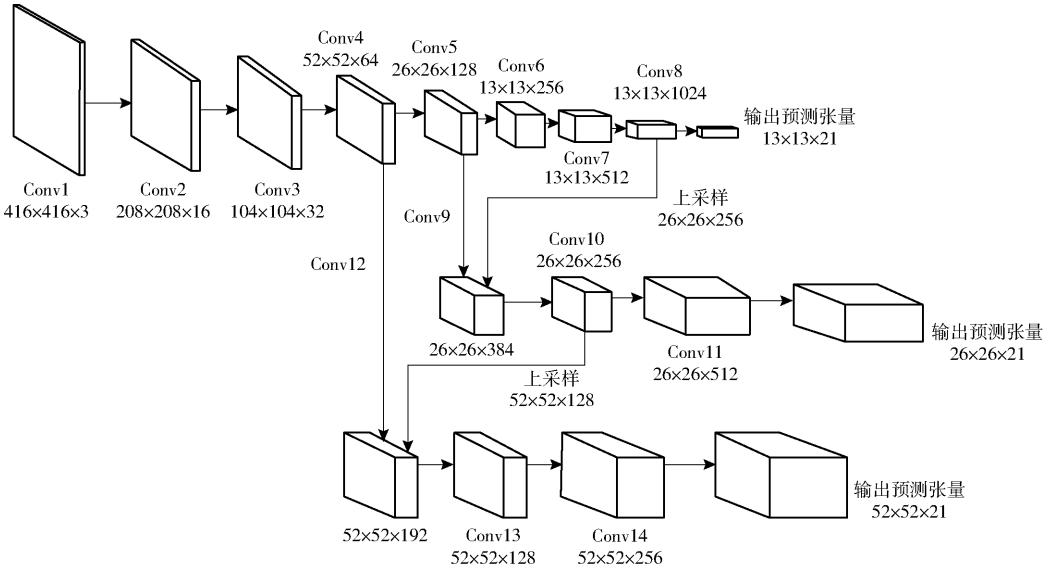


图 2　YOLOv3 网络训练结构

Fig. 2　YOLOv3 network training structure

量;$\sigma(t)$为逻辑函数,将坐标进行归一化,使其在0~1之间变化。P_w、P_h为该单元格对应的先验框(Anchor box)的宽和高。b_w、b_h为预测边界框的宽和高。

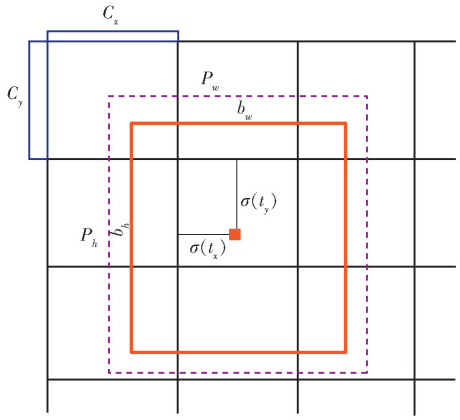


图3 YOLOv3边界框原理图

Fig. 3 Schematic of YOLOv3 boundary frame

1.2.2 训练权值模型

模型共训练8 000次,前1 000次迭代后模型损失值迅速收敛到1以下,训练至约4 000次后损失值逐渐平稳。每1 000次输出一个权重模型,得到8个模型,本文通过评价指标选取最优模型,8个模型的评价指标如表2所示。

表2 权值模型的性能比较

Tab. 2 Performance comparison of weight models

%

权重模型序号	成熟草莓 准确率 P_1	成熟草莓 召回率 R_1	未成熟草莓 准确率 P_2	未成熟草莓 召回率 R_2	成熟草莓 F_1	未成熟草莓 F_2	MAP
1	80.12	50.13	79.75	49.86	61.67	61.67	72.01
2	89.97	86.45	87.99	85.67	88.17	86.81	81.39
3	97.16	89.37	96.84	88.29	93.10	92.37	85.37
4	97.04	89.42	96.87	88.57	93.07	92.53	86.26
5	97.10	90.24	96.27	88.61	93.54	92.28	86.04
6	97.14	94.46	96.51	93.61	95.78	95.04	87.51
7	97.11	93.86	96.39	92.47	95.46	94.39	86.63
8	97.09	92.98	96.34	92.44	94.99	94.35	84.92

从表2可以看出,模型1综合性能最差,模型2~6性能较好且逐渐升高,模型7、8性能逐渐稳定。模型6的调和平均值F_1、F_2最大,说明模型的准确率和召回率综合表现最好,成熟草莓的准确率为97.14%,召回率为94.46%,未成熟草莓的准确率为96.51%,召回率为93.61%,平均精度87.51%最高,综合来看,模型6可以作为后期草莓的测试检验模型。

2 试验结果与分析

2.1 基于OpenCV图像的增强算法

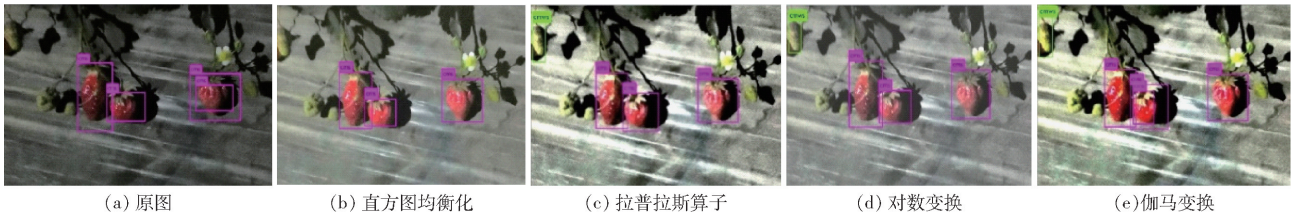
考虑夜晚在只有一个LED灯照明的情况下,采集的草莓图像背景过暗,模型很难将检测目标与背景区分,置信度过低导致无法达到检测目的。因此对夜晚的草莓图像进行图像增强。本文采用基于直方图均衡化[27]、拉普拉斯(Laplus)算子、对数变换、伽马(Gamma)变换[28-29]的4种图像增强方法对图像进行处理。由OpenCV图像增强算法处理后的图像相较原图在亮度、灰度、对比度、锐度等方面都有显著的改善,草莓特征与背景明显区分开。为了选择较好的图像增强方法,随机抽取夜晚草莓图像30幅,重复抽取3次共90幅,采用YOLOv3算法对增强后的图像进行实际效果检测,如图4所示。其中,cms表示成熟草莓,cmws表示未成熟草莓。由图4可以看出,原图背景过暗,识别效果差,甚至无法识别。不同图像增强算法效果对比如表3所示,增强后的图像可以明显看出草莓特征,并且识别精度提高,漏识别和一果多框的现象显著减少。

从表3中可以发现,4种图像增强算法对草莓的正确识别率都有提高,伽马变换图像增强算法比成熟草莓原图的正确识别率高20.15个百分点,比未成熟草莓的正确识别率比原图高35.03个百分点。由此可见,后期模型测试阶段采用伽马变换对夜晚草莓图像进行增强处理,有利于提高夜晚草莓图像识别精度。

2.2 草莓静态图像识别分析

2.2.1 不同草莓数量的试验对比

试验以草莓数量作为变量,测试集包括150幅图像,其中含1~3个草莓的图像共64幅,包含4~6个草莓的图像共51幅,包含7个及以上草莓的图像共35幅。每一类抽取25幅作为检测对象,检测出样本总量、正确识别的草莓数量、未识别出的草莓



(a) 原图 (b) 直方图均衡化 (c) 拉普拉斯算子 (d) 对数变换 (e)伽马变换

图4 不同图像增强算法的YOLOv3检测效果

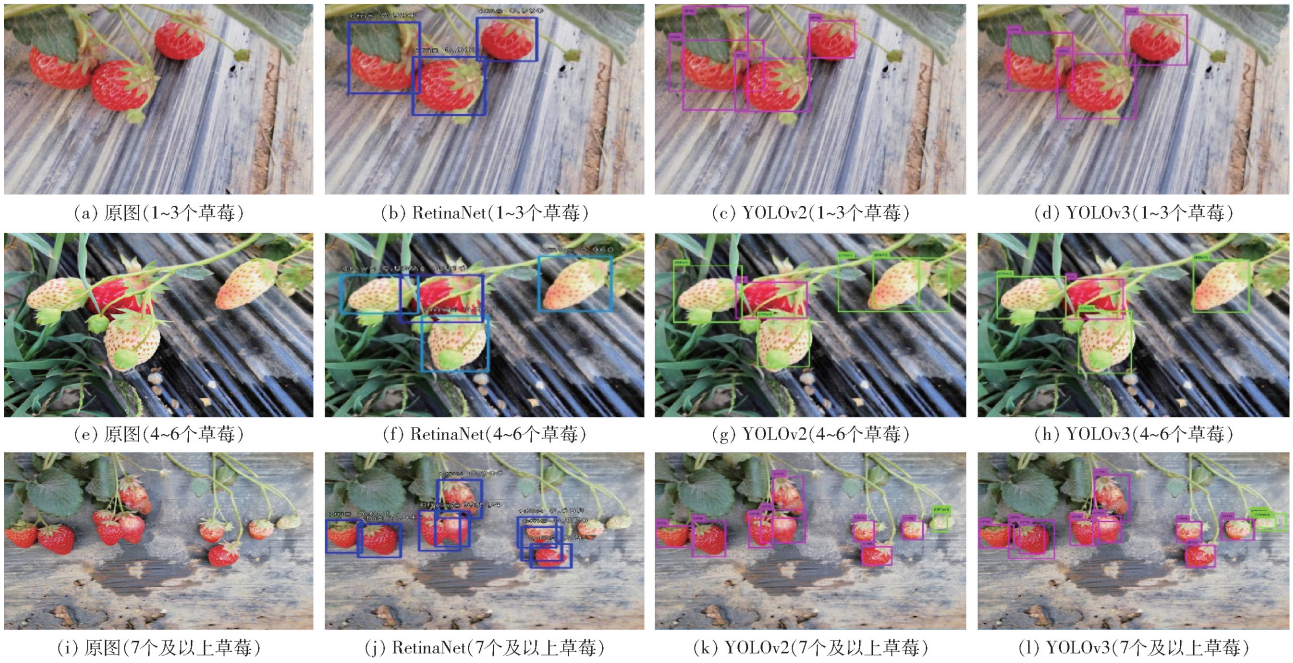
Fig. 4 YOLOv3 detection effects of different image enhancement algorithms

表 3　不同图像增强方法效果对比

Tab. 3　Comparison of effects of different image enhancement algorithms

图像增强方法	成熟草莓正确识别率/%	未成熟草莓正确识别率/%	每帧图像检测时间/ms
原图	70.97	53.21	37.37
直方图均衡化	89.49	81.86	36.40
拉普拉斯算子	89.76	83.14	36.39
对数变换	83.04	77.97	35.54
伽马变换	91.12	88.24	35.60

数量、误识别的草莓数量，计算准确率和召回率，采用 F_1 及 F_2 作为综合评价值，3 次重复取平均 F 值。图像边缘残缺草莓及过小的未成熟草莓忽略不计。从图 5 中可以看出，在草莓果实数量较少的情况下，3 种方法都能正确识别出成熟草莓和未成熟草莓，但草莓数量增加到 7 个以上时，存在一果多框和漏识别的现象，这是由于草莓的尺寸减小、果实重叠与粘连增加了模型识别的难度。表 4 为 3 种算法对不同数量草莓的检测结果。



(a) 原图(1~3个草莓)　(b) RetinaNet(1~3个草莓)　(c) YOLOv2(1~3个草莓)　(d) YOLOv3(1~3个草莓)

(e) 原图(4~6个草莓)　(f) RetinaNet(4~6个草莓)　(g) YOLOv2(4~6个草莓)　(h) YOLOv3(4~6个草莓)

(i) 原图(7个及以上草莓)　(j) RetinaNet(7个及以上草莓)　(k) YOLOv2(7个及以上草莓)　(l) YOLOv3(7个及以上草莓)

图 5　3 种算法对含不同草莓数量图像的检测效果

Fig. 5　Detection effects of three algorithms on different strawberry quantities

表 4　3 种算法对不同草莓数量检测试验结果

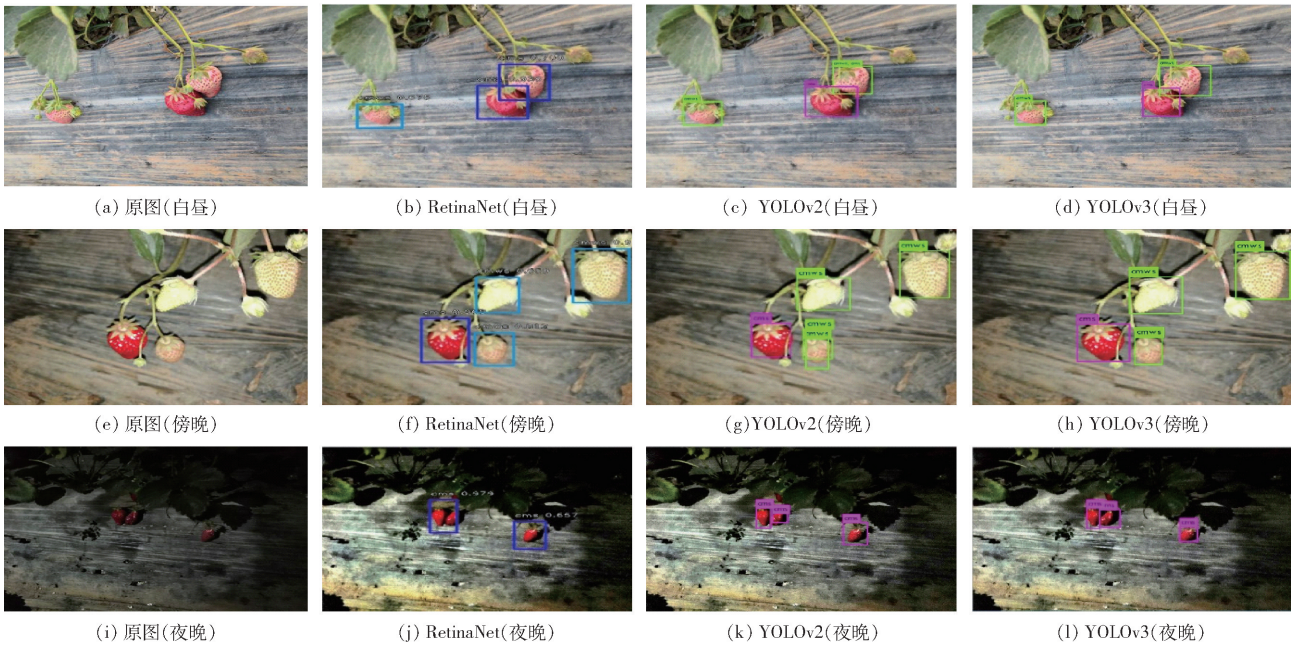
Tab. 4　Comparison of three algorithms for different strawberry quantity detection tests

草莓数量/个	算法	成熟草莓 F_1/%	未成熟草莓 F_2/%	每帧图像检测时间/ms
1~3	RetinaNet	93.97	92.26	81.25
	YOLOv2	93.10	92.27	80.36
	YOLOv3	96.16	95.02	30.48
4~6	RetinaNet	90.45	89.41	80.13
	YOLOv2	91.34	90.60	81.15
	YOLOv3	94.53	93.48	35.53
≥7	RetinaNet	86.16	82.75	95.33
	YOLOv2	87.47	84.98	92.91
	YOLOv3	89.14	86.03	36.64
平均值	RetinaNet	90.19	88.14	85.57
	YOLOv2	91.19	89.28	84.81
	YOLOv3	93.28	91.51	34.16

从表 4 中可看出，随着草莓数量的增多，成熟草莓和未成熟草莓的 F 值均减小，当草莓数量大于 7 时，F 值下降了 7~10 个百分点。随着草莓数量的增多，RetinaNet 算法容易导致一果多框及漏识别的现象，对小物体的检测精度明显下降。而 YOLOv3 采用多层神经特征网络，能够预测更小的栅格，在小物体的检测上取得了较好的检测效果，YOLOv3 算法 F 值最高。YOLOv3 算法选取的阈值为 0.5，YOLOv3 的每帧图像检测时间是 RetinaNet 的 39.92%，是 YOLOv2 的 40.27%。

2.2.2　不同时间段下草莓的识别效果对比

为了满足采摘机器人连续作业的要求，本文从白昼、傍晚和夜晚 3 个时间段来验证模型。采用白昼图像 80 幅，含有 414 个草莓，傍晚图像 65 幅，含有 358 个草莓，夜晚图像 46 幅，含有 255 个草莓。处理方式同 2.2.1 节。由图 6 可知，在同一块草莓区，不同时间段下，白昼下的草莓颜色、形状、表面特征清晰。傍晚时草莓的表面特征明显变模糊，并且背景变暗。而夜晚环境过于黑暗，草莓特征难以与背景明显区分开，极大地增加了识别难度。因此，为了达到较好的识别效果，采用伽马变换对夜晚的草莓图像进行增强。



(a) 原图(白昼) (b) RetinaNet(白昼) (c) YOLOv2(白昼) (d) YOLOv3(白昼)

(e) 原图(傍晚) (f) RetinaNet(傍晚) (g)YOLOv2(傍晚) (h) YOLOv3(傍晚)

(i) 原图(夜晚) (j) RetinaNet(夜晚) (k) YOLOv2(夜晚) (l) YOLOv3(夜晚)

图 6 3 种算法对不同时间段的草莓识别效果

Fig. 6 Strawberry recognition effects of three algorithms in different time periods

表 5 为 3 种算法在不同时间段对草莓识别结果的比较。在不同时间段下,总体上白昼时识别的调和平均值 F_1、F_2最大,比夜晚高近 5 个百分点,说明亮度对识别的效果影响很大,过暗的环境下草莓特征极不明显,增加识别难度。从表 5 中可以看出,YOLOv3 算法 F 值比 YOLOv2 算法平均高 2.79 个百分点,YOLOv2 算法与 RetinaNet 算法在白昼时的 F 值相似,傍晚和夜晚环境下 YOLOv2 算法比 RetinaNet 算法的 F 值略高。同一方法检测下,成熟草莓和未成熟草莓在白昼的 F 值相似,傍晚和夜晚环境下成熟草莓的 F 值比未成熟草莓略高。总体上 YOLOv3 算法每帧图像检测时间是 YOLOv2 算法的 41.13%,是 RetinaNet 算法的 38.57%。

表 5 3 种算法对不同时间段的草莓检测试验结果比较

Tab. 5 Comparison of strawberry test results of three algorithms in different time periods

时间段	算法	成熟草莓 F_1/%	未成熟草莓 F_2/%	每帧图像检测时间/ms
白昼	RetinaNet	93.71	92.84	82.32
	YOLOv2	93.67	92.97	77.55
	YOLOv3	96.19	96.02	34.64
傍晚	RetinaNet	90.72	89.37	92.26
	YOLOv2	91.75	90.87	91.25
	YOLOv3	94.26	93.73	35.27
夜晚	RetinaNet	88.73	86.19	104.01
	YOLOv2	88.96	87.25	92.47
	YOLOv3	91.24	90.76	37.56
均值	RetinaNet	91.05	89.47	92.86
	YOLOv2	91.46	90.36	87.09
	YOLOv3	93.90	93.50	35.82

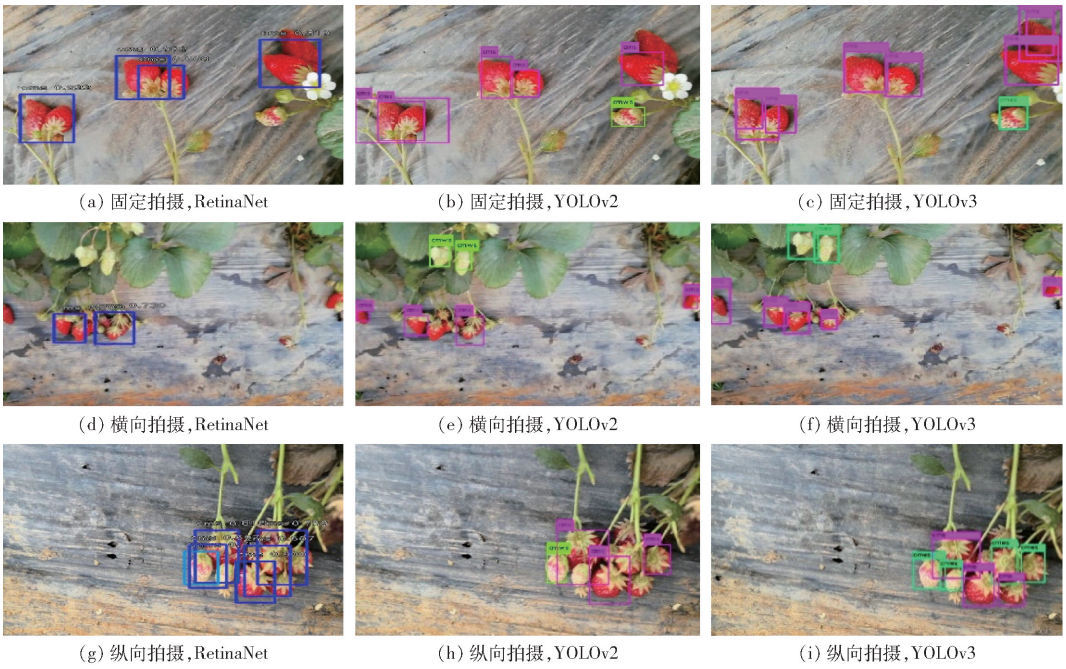
2.3 动态视频中草莓的识别效果对比

图 7 为 3 种算法在不同拍摄方式下的检测效果。由于固定拍摄方式在同一位置停留较长时间,该方式对检测目标的识别精度较其他方式更加准确。而横向拍摄时果实数量相对密集,果实重叠、粘连现象出现频率高,摄像停留时间较短,识别难度相对增加,易出现漏识别、误识别以及一果多框的现象。采用纵向拍摄模式时,草莓数量由少变多,果实遮挡频率高,识别难度先易后难。相对于 RetinaNet 和 YOLOv2 识别算法,YOLOv3 算法在视频识别的精度上有显著优势。因此,YOLOv3 算法不仅在识别精度上保持着较好的准确性,而且在复杂环境下具有良好的鲁棒性。

对于不同的视频拍摄方式,采取固定拍摄时的 F_1和 F_2最高(表 6),算法的准确率和召回率综合表现最好,草莓的识别效果最佳。纵向拍摄的效果略优于横向拍摄,是由于横向拍摄时草莓分布密集,增加了模型识别的时长和运算量。对于不同的识别算法,YOLOv3 算法平均检测速率是 YOLOv2 算法的 2.38 倍,是 RetinaNet 算法的 2.44 倍。在识别效果上,YOLOv3 算法平均 F 值最高,YOLOv2 算法与 RetinaNet 算法相似。综上,YOLOv3 算法在识别精度与检测速率上均满足实时检测的要求,视频平均识别速率达 58.1 f/s。

3 结论

(1)针对夜晚环境下草莓图像模糊的问题,使用伽马变换的图像增强算法处理图像后比原图成熟



(a) 固定拍摄,RetinaNet　(b) 固定拍摄,YOLOv2　(c) 固定拍摄,YOLOv3

(d) 横向拍摄,RetinaNet　(e) 横向拍摄,YOLOv2　(f) 横向拍摄,YOLOv3

(g) 纵向拍摄,RetinaNet　(h) 纵向拍摄,YOLOv2　(i) 纵向拍摄,YOLOv3

图 7　3 种算法在不同拍摄方式下的检测效果

Fig. 7　Detection effects of three algorithms in different shooting modes

表 6　3 种算法在不同拍摄方式下的检测结果比较

Tab. 6　Comparison of detection results of three algorithms in different shooting modes

拍摄方式	算法	成熟草莓 F_1/%	未成熟草莓 F_2/%	检测速率/(f·s^{-1})
固定拍摄	RetinaNet	94.45	93.38	24.5
	YOLOv2	95.75	95.05	25.5
	YOLOv3	97.49	96.95	58.6
横向拍摄	RetinaNet	93.61	92.18	23.4
	YOLOv2	94.08	93.73	24.6
	YOLOv3	96.14	95.55	58.3
纵向拍摄	RetinaNet	93.88	92.75	23.4
	YOLOv2	94.14	93.79	23.1
	YOLOv3	96.83	95.68	57.5
平均值	RetinaNet	93.98	92.77	23.8
	YOLOv2	94.66	94.19	24.4
	YOLOv3	96.82	96.06	58.1

草莓识别正确率提高了 20.15 个百分点,比原图中未成熟草莓识别正确率提高了 35.03 个百分点,平均每幅图像检测时间比原图减少 1.77 ms。

(2)对比不同算法在不同草莓数量、不同时间段下的检测效果, YOLOv3 算法的 F 值最大,综合识别效果最好,每帧图像的平均检测时间为 34.99 ms,是 YOLOv2 算法的 41.13%,是 RetinaNet 算法的 38.57%, YOLOv3 算法在识别精度及速率上优势显著。

(3)采用固定拍摄方式时 F 值最高,识别效果最好,YOLOv3 算法平均检测速率为 58.1 f/s,视频检测流畅,满足实时性要求,在果实遮挡、重叠、密集等复杂情况下具有良好的鲁棒性。

参　考　文　献

[1] 赵德安, 刘晓洋, 陈玉, 等. 苹果采摘机器人夜间识别方法[J/OL]. 农业机械学报, 2015, 46(3):15 - 22.
ZHAO Dean, LIU Xiaoyang, CHEN Yu, et al. Night recognition method of apple picking robot[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(3): 15 - 22. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx? file_no=20150303&flag=1&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2015.03.003. (in Chinese)

[2] 张铁中, 陈利兵, 宋健, 等. 草莓采摘机器人的研究:Ⅱ.基于图像的草莓重心位置和采摘点的确定[J].中国农业大学学报, 2005, 10(1):48 - 51.
ZHANG Tiezhong, CHEN Libing, SONG Jian, et al. Study on strawberry harvesting robot:Ⅱ. Images based identifications of strawberry barycenter and plucking position [J]. Journal of China Agricultural University, 2005, 10(1): 48 - 51. (in Chinese)

[3] 谢志勇, 张铁中, 赵金英. 基于 Hough 变换的成熟草莓识别技术[J].农业机械学报, 2007, 38(3):106 - 109.
XIE Zhiyong, ZHANG Tiezhong, ZHAO Jinying. Ripened strawberry recognition based on Hough transform[J]. Transactions of the Chinese Society for Agricultural Machinery, 2007, 38(3):106 - 109. (in Chinese)

[4] 张凯良, 杨丽, 张铁中. 草莓采摘位置机器视觉与激光辅助定位方法[J].农业机械学报, 2010, 41(4):151 - 156.
ZHANG Kailiang, YANG Li, ZHANG Tiezhong. Object locating method of laser-assisted machine vision for strawberry harvesting[J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(4):151 - 156. (in Chinese)

[5] 李长勇, 房爱青, 谭红, 等. 高架草莓采摘机器人系统研究[J].机械设计与制造, 2017(6):245 - 247.
LI Changyong, FANG Aiqing, TAN Hong, et al. Elevated strawberry picking robot system research[J]. Machinery Design & Manufacture, 2017(6):245 - 247. (in Chinese)

[6] ZHANG Qirong, ZUO Jianjun, YU Tingzhong, et al. Visual recognition system of cherry picking robot based on Lab color model [C]// IOP Conference Series Earth and Environmental Science, 2017, 100(1):1 - 5.
[7] 侯贵洋, 赵贵杰, 王璐瑶. 草莓采摘机器人图像识别系统研究[J]. 软件, 2018, 39(6):184 - 188.
HOU Guiyang, ZHAO Guijie, WANG Luyao. Research on image recognition system of strawberry picking robot[J]. Computer Engineering & Software, 2018, 39(6):184 - 188. (in Chinese)
[8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770 - 778.
[9] LIN T Y, DOLLAR P, ROSS G, et al. Feature pyramid networks for object detection [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:2117 - 2125.
[10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014:580 - 587.
[11] GIRSHICK R. Fast R - CNN[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015:1440 - 1448.
[12] REN S, HE K, GIRSHICK R, et al. Faster R - CNN: towards real-time object detection with region proposal networks[C]// Advances in Neural Information Processing Systems, 2015: 91 - 99.
[13] HE K, GKIOXARI G, DOLLAR P, et al. Mask R - CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 99:2961 - 2969.
[14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// European Conference on Computer Vision. Cham:Springer, Cham, 2016: 21 - 37.
[15] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 99:2999 - 3007.
[16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]// Computer Vision and Pattern Recognition. IEEE, 2016:779 - 788.
[17] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017:6517 - 6525.
[18] 熊俊涛, 刘振, 林睿, 等. 自然环境下树上绿色芒果的无人机视觉检测技术[J/OL]. 农业机械学报, 2018, 49(11):30 - 36.
XIONG Juntao, LIU Zhen, LIN Rui, et al. Unmanned aerial vehicle vision detection technology of green mango on tree in natural environment[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(11):30 - 36. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx? file_no = 20181103&flag = 1&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2018.11.003. (in Chinese)
[19] 王宇宁, 庞智恒, 袁德明. 基于 YOLO 算法的车辆实时检测[J]. 武汉理工大学学报, 2016, 38(10):41 - 46.
WANG Yuning, PANG Zhiheng, YUAN Deming. Vehicle detection based on YOLO in real time[J]. Journal of Wuhan University of Technology, 2016, 38(10):41 - 46. (in Chinese)
[20] 张姗, 逯瑜娇, 罗大为. 基于深度学习的目标检测算法综述[C]// 中国计算机用户协会网络应用分会 2018 年第二十二届网络新技术与应用年会, 2018:123 - 135.
[21] 薛月菊, 黄宁, 涂淑琴, 等. 未成熟芒果的改进 YOLOv2 识别方法[J]. 农业工程学报, 2018, 34(7):173 - 179.
XUE Yueju, HUANG Ning, TU Shuqin, et al. Immature mango detection based on improved YOLOv2[J]. Transactions of the CSAE, 2018, 34(7):173 - 179. (in Chinese)
[22] 赵德安, 吴任迪, 刘晓洋, 等. 基于 YOLO 深度卷积神经网络的复杂背景下机器人采摘苹果定位[J]. 农业工程学报, 2019, 35(3):164 - 173.
ZHAO Dean, WU Rendi, LIU Xiaoyang, et al. Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background[J]. Transactions of the CSAE, 2019, 35(3):164 - 173. (in Chinese)
[23] 李茂晖, 吴传平, 鲍艳, 等. 论 YOLO 算法在机器视觉中应用原理[J]. 教育现代化, 2018, 41(5):180 - 182.
LI Maohui, WU Chuanping, BAO Yan, et al. On the application principle of YOLO algorithm in machine vision[J]. Education Modernization, 2018, 41(5):180 - 182. (in Chinese)
[24] REDMON J, FARHADI A. YOLOv3: an incremental improvement [Z/OL]. Computer Vision and Pattern Recognition (CVPR) (2018 - 4 - 1) [2019 - 04 - 02]. https://arxiv.org/abs/1804.02767v1.
[25] HRIPCSAK G, ROTHSCHILD A S. Agreement, the f-measure, and reliability in information retrieval[J]. Journal of the American Medical Informatics Association, 2005, 12(3): 296 - 298.
[26] 吕铄, 蔡烜, 冯瑞. 基于改进损失函数的 YOLOv3 网络[J]. 计算机系统应用, 2019, 28(2):1 - 7.
LÜ Shuo, CAI Xuan, FENG Rui. YOLOv3 network based on improved loss function[J]. Computer Systems & Application, 2019, 28(2):1 - 7. (in Chinese)
[27] 赵德安, 刘晓洋, 孙月平, 等. 基于机器视觉的水下河蟹识别方法[J/OL]. 农业机械学报, 2019,50(3):151 - 158.
ZHAO Dean, LIU Xiaoyang, SUN Yueping, et al. Detection of underwater crabs based on machine vision[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(3):151 - 158. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20190316&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2019.03.016. (in Chinese)
[28] LIU Bin, WANG Xia, JIN Weiqi, et al. Infrared image detail enhancement based on local adaptive gamma correction[J]. Chinese Optics Letters, 2012, 10(2):25 - 29.
[29] 王术波, 韩宇, 陈建, 等. 基于深度学习的无人机遥感生态灌区杂草分类[J]. 排灌机械工程学报, 2018, 36(11): 1137 - 1141.
WANG Shubo, HAN Yu, CHEN Jian, et al. Classification of weeds in ecological irrigation areas based on deep learning and remote sensing images taken by drone at low-altitude[J]. Journal of Drainage and Irrigation Machinery Engineering, 2018, 36(11): 1137 - 1141. (in Chinese)