

高梦圆,马双宝,董玉婕,等. 基于实例分割苹果采摘机器人视觉定位与检测[J]. 江苏农业科学,2022,50(3):201-208.

doi:10.15889/j.issn.1002-1302.2022.03.033

基于实例分割苹果采摘机器人视觉定位与检测

高梦圆,马双宝,董玉婕,林巍,何毅

(武汉纺织大学机械工程与自动化学院,湖北武汉 430073)

摘要:为了实现采摘机器人对复杂环境中的苹果进行精准的视觉检测与定位,提出了一种基于 Mask RCNN 实例分割模型检测果园苹果的方法。为了降低模型的参数数量并提升检测的速度,将主干特征提取网络由 Resnet101 网络替换为轻量化 Mobilenetv3 网络。该网络结构采用深度可分离卷积代替普通卷积,能有效降低模型参数量。并采用空间金字塔池化(spatial pyramid pooling,简称 SPP)模块,对输出网络进行语义增强,利用空洞卷积掩码特征图增加感受野。模型的参数量相较于 Resnet101 下降了 62.03%,检测速度由 5 帧/s 提高到 9 帧/s,检测速度提高 80%。检测结果准确率为 93.75%,召回率为 89.99%,平均检测精度(mAP)为 93.75%。试验证明,该算法能够对不同场景的苹果边缘位置信息以及类别进行有效的视觉识别检测,可为自动化采摘苹果机器人的视觉系统提供有力的技术支撑。

关键词:Mask RCNN;苹果识别;Mobilenetv3;SPP;空洞卷积

中图分类号:TP391.41;TP242.6⁺2 **文献标志码:**A **文章编号:**1002-1302(2022)03-0201-08

苹果营养丰富,含有大量矿物质、维生素以及膳食纤维,深受人们喜爱。因此,华北、西北地区以及更多地区成为了苹果的主产区。在农业生产中,由于苹果采摘过程的复杂性,采摘作业需要耗费大

量劳动人力,苹果的采摘工作约占据整个生产工作量的 40%^[1]。为了加快农业生产的高效快速发展,推动苹果生产的智能化发展,利用苹果采摘机器人来代替人工采摘具有重大的现实意义和广阔的应用前景^[2-5]。

苹果采摘机器人主要由视觉检测模块、机械结构设计和运动控制 3 个部分组成,主要采用机械臂对其进行采摘^[6-7]。如何快速准确对果园的苹果进行检测是苹果采摘机器人的技术关键。对果园中苹果的视觉检测效果受诸多因素的影响:现场环境的光照变化剧烈;果实之间的差异性较大;果实之

收稿日期:2021-07-22

基金项目:国家自然科学基金(编号:61271008);湖北省数字化纺织装备重点实验室公开项目(编号:DTL2019020)。

作者简介:高梦圆(1995—),男,湖北襄阳人,硕士研究生,主要从事视觉检测与控制。E-mail:2535204189@qq.com。

通信作者:马双宝,博士,副教授,主要从事人工智能技术在农业机器人感知与控制研究。E-mail:2006118@wtu.edu.cn。

[2]邓源喜,张姚瑶,董晓雪,等.花生营养保健价值及在饮料工业中的应用进展[J].保鲜与加工,2018,18(6):166-169,174.

[3]刘娟,汤丰收,张俊,等.国内花生生产技术现状及发展趋势研究[J].中国农学通报,2017,33(22):13-18.

[4]薛然,谢焕雄,胡志超,等.花生荚果分级机械研究现状与发展建议[J].江苏农业科学,2015,43(9):426-428.

[5]陈善峰,王俊.农产品分级技术及品质检测设备的现状与发展趋势[J].粮油加工与食品机械,2000(4):8-10.

[6]毛鹏军,杜东亮,符丽君,等.我国农产品分级的现状、问题及对策[J].农业机械,2006(10):110-111.

[7]Blankenship P D,Dowell F E. A diverging belt screen for farmer stock Peanuts1[J]. Peanut Science,1997,24(1):37-41.

[8]Dowell F E. An automated cleaning, pod sizing, shelling, and kernel sizing system for grading farmers stock Peanuts1 [J]. Peanut Science,1996,23(1):46-54.

[9]刘道奇,刘龙,孙千涛,等.河南省花生机械化收获的现状、存在的问题及对策[J].农业开发与装备,2020(10):29-31.

[10]谢焕雄,王建楠,胡志超,等.我国种用花生机械化脱壳技术路线[J].花生学报,2012,41(2):32-36.

[11]薛然,谢焕雄,胡志超,等.花生荚果圆筒筛分试验研究[J].中国农机化学报,2016,37(2):101-104.

[12]冯金有.一种花生果筛选机:CN2597104[P].2004-01-07.

[13]刘涛,何瑞银,陆静,等.基于 EDEM 的窝眼轮式油菜排种器排种性能仿真与试验[J].华南农业大学学报,2016,37(3):126-132.

[14]薛然.花生荚果圆筒筛分特性研究与参数优化[D].北京:中国农业科学院,2015.

[15]迟明,宋鸿,许杰,等.基于 EDEM 对带式输送机系统的改造与分析[J].起重运输机械,2020(4):86-90.

[16]侯俊铭,杨勇,董帅,等.典型花生种子脱壳特性试验及有限元仿真研究[J].东北农业大学学报,2017,48(11):74-85.

[17]张强,张旭,孙绍安.基于 EDEM 的家用榨油机压榨腔工作性能仿真研究[J].农业工程学报,2018,34(24):283-291.

[18]顿国强,杨永振,郭艳玲,等.不同品种大豆种子充填特性的 EDEM 仿真分析[J].河南农业大学学报,2019,53(1):93-98.

间存在相互遮挡重叠状况;果树枝干、树叶对果实的遮挡;对苹果采摘机器人硬件设备的计算要求高,系统设计复杂。上述问题成为苹果采摘机器人的视觉检测研究的主要障碍。

目前,对苹果机器人视觉检测识别方面的研究工作取得了一定的进展。Fan 等提出了以灰度为中心的 RGB 颜色,利用 K - Means 聚类算法分割苹果的技术,该技术能有效对果园苹果图像进行实时分割^[8]。Liu 等利用颜色空间进行候选区域的提取,再利用 HOG 特征对候选区域的苹果进行检测与定位,在速度和精度上取得很好的检测效果^[9]。张春龙等采用 HSV 颜色空间对绿色苹果进行颜色特征提取,并利用支持向量机(SVM)分类器进行分类,有效实现了青色苹果的识别与定位^[10]。廖崴等利用 RGB 颜色空间和 Ostu 算法将果实和树叶从背景中分割,利用灰度和纹理特征构建随机森林分类器,识别苹果果实与树叶,并利用霍夫圆对苹果边缘进行拟合,在自然光照下对苹果进行有效识别^[11]。

随着人工智能技术的兴起,以深度学习为基础的目标检测技术在苹果检测中开始得到应用。彭红星等采用 Resnet - 101 网络代替 SSD 模型中的 VGG - 16 网络,并利用迁移学习训练方法对 4 种不同水果进行检测^[12]。武星等提出了轻量化同构残差快串联的特征提取网络,采用多种目标损失函数的改进 YOLOv3 模型,对果园中遮挡、重叠情况的苹果取得很好的检测^[13]。在此基础上,李大华等利用 YOLOv3 网络对环境中的苹果目标区域进行检测,并采用颜色空间进行阈值分割,得到苹果的准确位置信息^[14]。岳有军等采用 Mask RCNN 模型对果园的苹果进行检测并提取苹果的像素区域^[15-16]。虽

然上述方法针对果园的苹果检测在实时性和鲁棒性上取得了很好的检测效果,但是模型参数量大,对苹果采摘机器人的计算资源消耗过多,对苹果果实的检测只是采用矩形框进行包围检测,并未对苹果的边缘和形状信息进行有效预测。

针对上述研究中存在问题,本研究提出了以实例分割网络 Mask RCNN^[17]模型作为苹果采摘机器人视觉定位与检测模型。该模型能准确预测采摘果实的边缘和形状信息,但是模型参数量大,检测速度慢,难以达到采摘机器人实时检测要求。针对此问题,在 Mask RCNN 基础上将主干网络改为适用于移动端的 Mobilenetv3 网络模型,为了增强语义信息,使模型对适应复杂的果园环境具有较强的鲁棒性,采用 SPP 网络对特征图进行语义增强,采用空洞卷积增加特征图的感受野和上下文的语义信息。

1 数据采集与处理

1.1 试验数据集采集和标注

试验数据采集地位于山东临沂市沂水县四十里镇。选取红苹果、青苹果和奶油苹果 3 种不同的苹果树进行取样拍摄,拍摄时间为 06:00—16:00,试验选用海康威视、手机等移动端采集设备进行图像采集,拍摄条件包括远距离和近距离、不同角度光照度、果叶和枝干的遮挡、果实之间重叠、不同成熟期果实的果实混杂(图 1),共采集 500 幅图像。试验数据采用随机裁剪、缩放、平移、翻转、颜色扰动,加入马赛克数据增强后为 35 000 幅,对数据进行随机抽取,选用 28 000 幅为训练数据集,3 500 幅为测试数据集,3 500 幅为验证数据集,比例为 8 : 1 : 1。使用 Labelme 软件人工对采集苹果图像中苹果的边界区域进行标注,其余为背景,生成 json 文件。



图1 采集数据

2 苹果果实识别检测模型

2.1 Mask RCNN 模型

Mask RCNN 网络是 He 等在 Faster RCNN 网络模型基础上添加了 Mask 网络预测分支来对检测模

型进行改进^[17],其网络结构见图 2。从图 2 可以看出,该网络模型由特征提取网络(ResNet-101)、特征金字塔(feature pyramid network)、区域建议网络(research proposal network)、ROIAlign 层以及网络预测(类别、边界框、掩码)分支组成。

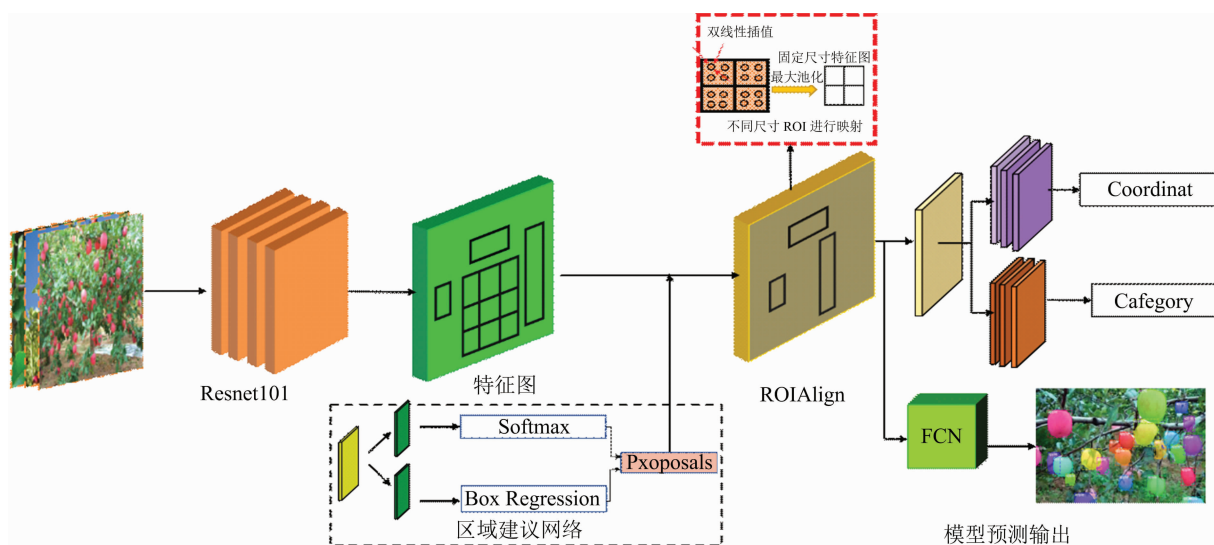


图2 Mask RCNN 网络模型框架

2.2 改进 Mask RCNN 模型

原始 Mask RCNN 模型采用 Resnet-101^[18] 网络作为主干特征提取网络,该网络能够有效提取特征并在实际检测中取得很好的检测效果,但是网络层数过多,模型参数量庞大,在进行特征提取过程中计算参数量大,需要占用过多采摘机器人硬件检测资源,耗费较长时间,导致检测速度过慢。由于农业苹果采摘工作量大,在进行苹果定位检测时需要具有较高实时性以便快速进行农业生产工作,因此需要改进模型参数量提高检测速度。Howard 等提出了致力于嵌入式开发的轻量化网络 Mobilenetv3(图 3)^[19],使用深度可分离卷积^[20]代替标准卷积模块,并将挤压(squeeze)和激励(excitation)2 种轻量级的注意力机制^[21]引入线性瓶颈结构(图 4)。

相对于传统卷积操作,采用深度可分离卷积可明显降低模型运算参数量。如对输入样本的尺寸为 W (行数) $\times H$ (列数) $\times N$ (通道数)的向量,采用 $k \times k$ 的卷积核进行标准的卷积运算,得到 M 通道的输出向量,需要计算量的参数量为

$$q_1 = k^2 NM. \quad (1)$$

若采用深度可分离卷积进行相同的操作,需要的计算量为

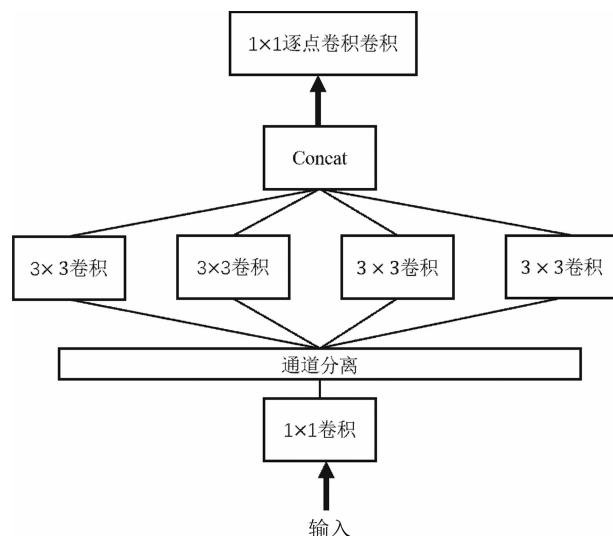


图3 深度可分离卷积

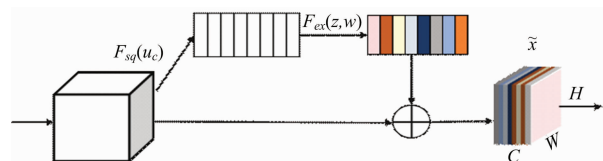


图4 SE 结构

$$q_2 = k^2 N + NM. \quad (2)$$

2 种运算下参数数量的比值为

$$\frac{q_1}{q_2} = \frac{k^2 N + NM}{k^2 NM} = \frac{1}{M} + \frac{1}{k^2}. \quad (3)$$

由上式可知,当输入向量通道数远大于卷积核尺寸 k 时,2 种方式卷积运算的参数量比值与卷积核的尺寸几乎成平方反比关系。假设卷积核的尺寸 k 为 3 时,若采用深度可分离卷积可将计算的参数量减少为原来的 $1/9$ 。由此可见,采用深度可分离卷积可大大降低模型参数量,提高检测速度。

MobilenetV3 网络另一改进之处在于引入了注意力机制,该机制由挤压模块(squeeze block)和激励模块(excitation block)组成(图 4)。利用挤压模块将输入的多通道矩阵压缩成一维向量,然后通过可训练的权重与一维向量进行乘法运算实现激励,通过误差反向传播调整权重值来筛选需要的特征,其中挤压模块的运算公式为

$$z_c = F_{sq}(u_c) = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j)。 \quad (4)$$

式中: z_c 为输出特征向量; u_c 为输入矩阵第 c 个通道特征; F_{sq} 为挤压函数; i 为二维特征图的行坐标; j 为二维特征图的列坐标。

挤压后的向量采用进入全连接网络,输入层和输出层的神经元个数相同,采用 ReLu 激活函数,激励操作采用 Sigmoid 函数进行运算,其表达式为

$$s = F_{ex}(z, w) = \text{ReLu}[g(z, W)] = \text{ReLu}[W_2 \text{Sigmoid}(W_1 z)]。 \quad (5)$$

式中: F_{ex} 为激励函数; g 为全连接层; z 为输出特征值; W_1 、 W_2 为全连接层权重值;ReLu、Sigmoid 为激活函数。

最后,采用 Scale 运算将输出向量与原始特征图的矩阵相乘,获得施加权重的特征图,以增强有效特征,去除或衰减无效特征,使得到的特征更加具有表达性,其表达式为

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c \times u_c。 \quad (6)$$

式中: \tilde{x}_c 为输出特征图的第 c 个通道征; F_{scale} 为注意力权值。

2.2 语义增强模块

Mask RCNN 模型在 Faster RCNN 模型存在 3 个网络预测分支,造成模型精度下降,采用 SPP 网络结构^[22]模型特征图进行语义增强,网络结构见图 5。SPP 结构位于主干网络与 RPN 网络之间,将输入的特征图采用不同的池化操作转变为不同尺度特征图,然后再将特征图进行维度叠加,使通道进行融合,增强网络的鲁棒性,加快模型推理速度。

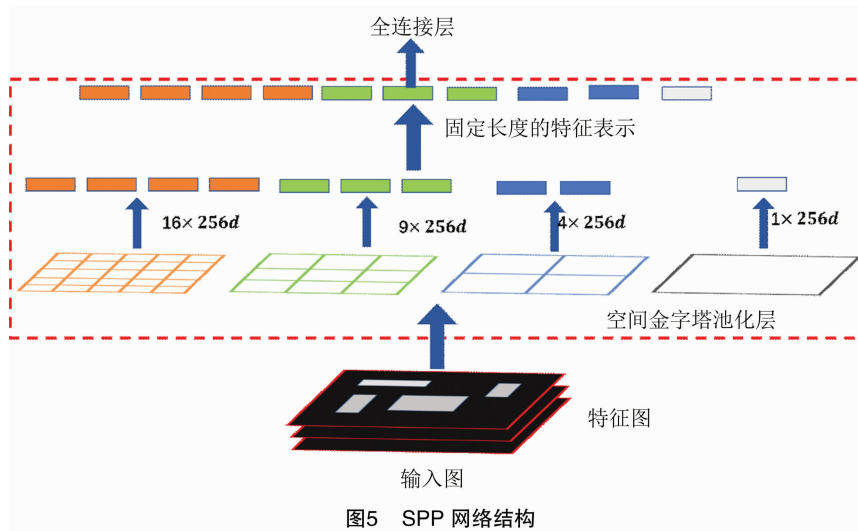


图5 SPP 网络结构

Mask RCNN 模型采用语义分割分支对图像的掩码进行预测,将原始输入图像压缩至 7×7 大小的特征图进行像素的预测,会导致图像特征信息的丢失。采用空洞卷积模块^[23]增加特征图的感受野,获取更加丰富的上下文语义信息来提高对图像中小尺度物体的检测。空洞卷积结构见图 6。

从图 6 可以看出,相对于普通卷积,利用深度可分离卷积不仅大大扩宽特征图的感受野,而且网络

的计算量没有增加。

2.3 损失函数

Mask RCNN 模型训练的损失由 2 个部分组成:RPN 网络损失 L_{RPN} 和最终模型的多分支输出预测损失 $L_{Mul-Branch}$,其表达式示为

$$L_{Total} = L_{RPN} + l_{Mul-Branch}。 \quad (7)$$

式中: L_{RPN} 包含矩形框内物体的分类损失以及矩形框的位置损失,其表达式为

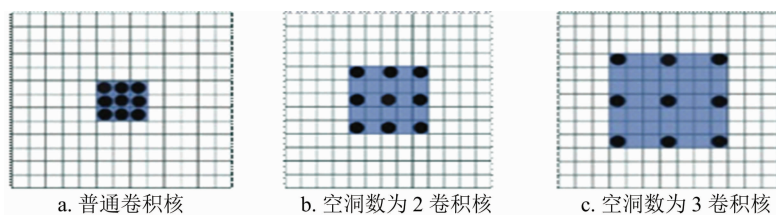


图6 空洞卷积结构

$$L_{RPN} = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda_1 \frac{1}{N_{\text{reg}}} \sum_i p_i^* l_{\text{reg}}(t_i, t_i^*) \quad (8)$$

式中: p_i 为矩形框类别概率; p_i^* 为第 i 个矩形框真实类别概率; t_i 第 i 个矩形框真实坐标信息; t_i^* 第 i 个预测矩形框的坐标信息。

$L_{\text{cls}}(p_i, p_i^*)$ 表示分类损失, 其表达式为

$$L_{\text{cls}}(p_i, p_i^*) = \sum_i -\log p_i p_i^* \quad (9)$$

$L_{\text{reg}}(t_i, t_i^*)$ 表示网络输出的矩形框位置损失, 其表达式为

$$L_{\text{reg}}(t_i, t_i^*) = \text{smooth}_{L1}(t_i^*, t_i) = \begin{cases} 0.5 \times (t_i^*, t_i)^2 & \text{if } |t_i^* - t_i| < 1 \\ |t_i^* - t_i| - 0.5 & \text{otherwise} \end{cases} \quad (10)$$

模型输出的分支网络中回归损失与分类损失与 RPN 网络相同, 掩码损失采用二分类的交叉熵损失函数, 其表达式为

$$L_{\text{mask}} = \sum_i -[s_i^* \log(s_i) + (1 - s_i^*) \log(1 - s_i)] \quad (11)$$

式中: s_i 为第 i 个像素预测属于目标像素的概率; s_i^* 为第 i 个像素属于真实目标像素的概率。

3 结果与分析

3.1 试验环境配置

试验选用 AMD Ryzen 7 3750H CPU, 2.3 GHz 主频, 16 GB 运行内存, GeForce GTX1660Ti GPU, 操作系统为 Ubuntu 18.04。设置模型初始学习率为 0.000 1, 动量因子为 0.001, 权重衰减系数设置为 0.000 1, 动量因子为 0.9, batch_size 设置为 8, 模型 RPN 网络层产生的 Anchor 大小设置为 (16, 32, 64, 128, 256), 非极大抑制系数设置为 0.3, 训练迭代次数设置为 100。

3.2 评价指标

选取准确率 (P)、平均检测精度 (mAP)、召回率 (R) 作为模型检测的评价指标。

3.3 训练数据分析

采用相同的训练集、验证集以及测试集分别在主干网络 Resnet-101、Resnet-50、Mobilenetv3 以及改进网络上训练, 4 种模型训练损失以及验证损失见图 7。

从图 7 可以看出, 改进后的 Mask RCNN 网络模型训练损失值变化波动比较平缓, 幅值下降较快, 反映了网络特征提取能力更强, 模型泛化能力更好, 更能适应果园复杂的训练环境。Resnet101 网络虽然误差曲线波动较小, 但是损失值较慢, 反映了网络特征提取能力弱于改进后的网络。采用 Resnet50 和 Mobilenetv3 网络损失值波动变化大, 下降缓慢, 模型泛化能力弱。改进后的 Mask RCNN 模型在 RPN 网络分类预测损失上比 Resnet101、Resnet50 以及 Mobilenetv3 网络分别下降了 40.62、58.77、617.34 百分点, 在 RPN 网络上预测框的损失分别下降了 52.22、226.34、294.93 百分点, 在输出网络上分类损失分别下降了 75.32、73.30、337.64 百分点, 回归损失分别下降了 32.24、226.94、635.37 百分点, 掩码损失分别下降了 10.38、71.91、237.94 百分点, 模型的总损失分别下降了 23.33、54.24、303.35 百分点。

为了进一步验证采用实例分割网络能有效作为苹果采摘机器人视觉处理模块, 采用 YOLOv4^[23-24]、Faster RCNN^[25-26] 以及改进 Mask RCNN 模型对同一测试数据对比, 检测效果见图 8。从图 8 可以看出, 虽然 YOLOv4、Faster RCNN 模型能有效对遮挡、不同拍照角度、不均匀光照的果园苹果进行有效的检测, 但是只是采用矩形框粗略地对果园苹果进行定位, 难以提取苹果的边缘和形状信息, 对机械臂采摘作业会造成一定的偏差。采用实例分割模型不仅能对苹果区域进行矩形框包围, 而且能采用不同颜色的掩码准确分割目标的像素区域位置, 很好地满足苹果采摘机器人视觉要求。

从表 1 可以看出, 改进后的 Mask RCNN 模型对苹果果园果实检测时在准确率方面比采用 Resnet101

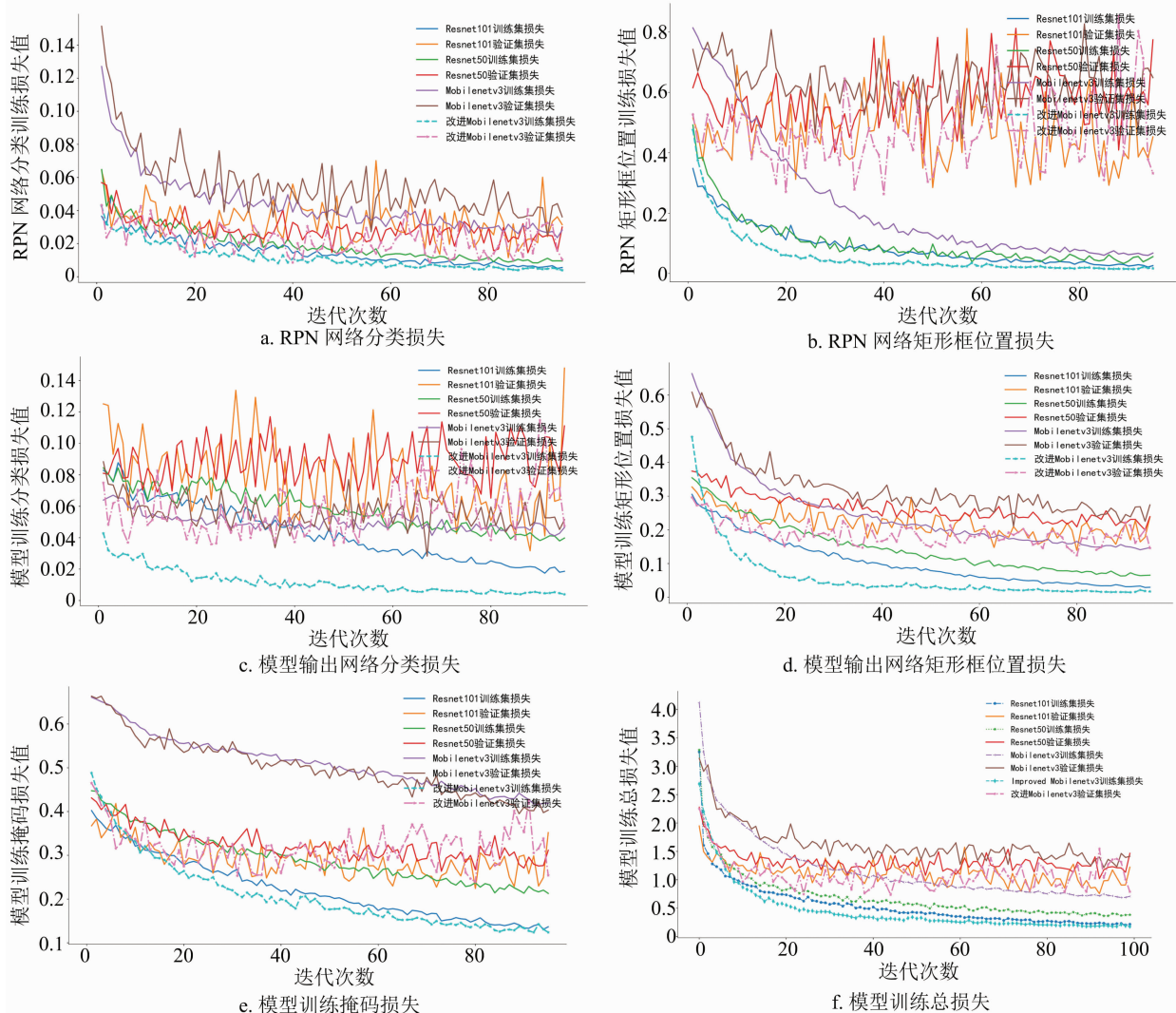


图9 苹果采摘机器人视觉模型损失曲线

网络、Resnet50 网络、Mobilenetv3 网络分别提高了 3.84、10.42、27.08 百分点;在召回率上比采用 Mobilenetv3 网络提高了 6.66 百分点,比 Resnet50 提高了 4.28 百分点,比 Resnet101 网络下降了 0.91 百分点;在 mAP 方面,比 Resnet101 网络提高了 3.30 百分点,比 Resnet50 提高了 26.56 百分点,比 Mobilenetv3 网络提高了 21.56 百分点。

从表 2 可以看出,改进后的模型比采用主干网络为 Resnet101、Resnet50 的网络参数量分别下降了 62.03%、46.24%,比 Mobilenetv3 参量稍微提高了 0.83%。在检测速度方面,比 Resnet101 网络提高了 4 帧/s 的检测速度,与 Resnet50 检测速度相同,比 Mobilenetv3 少 3 帧/s。

为检验改进后的模型对不同种类的苹果果实进行识别检测,在果实遮挡、光照充足、光照不足、远距离检测 4 种常见自然条件下的果园进行测试,

试验检测效果见图 9。

为了进一步验证改进后的模型性能,对模型的参数量、权重文件以及检测速度进行对比,结果见表 2。

从图 9 可以看出,改进后的模型不仅能有效对遮挡的果实进行检测,而且对光照充足和不足、远距离拍摄采样的图片也能进行有效识别,表明该网络在复杂的果园环境下对多目标的果实检测能够有效检测,具有较强的鲁棒性。

4 结论

本研究为了准确对苹果果园的果实进行有效检测,提出了基于 Mask RCNN 的检测模型,为了减少模型参数量,主干网络采用 Mobilenetv3 网络代替 Resnet101 网络。该模型采用深度可分离卷积代替普通卷积,大大减少模型参数量,并采用通道注意

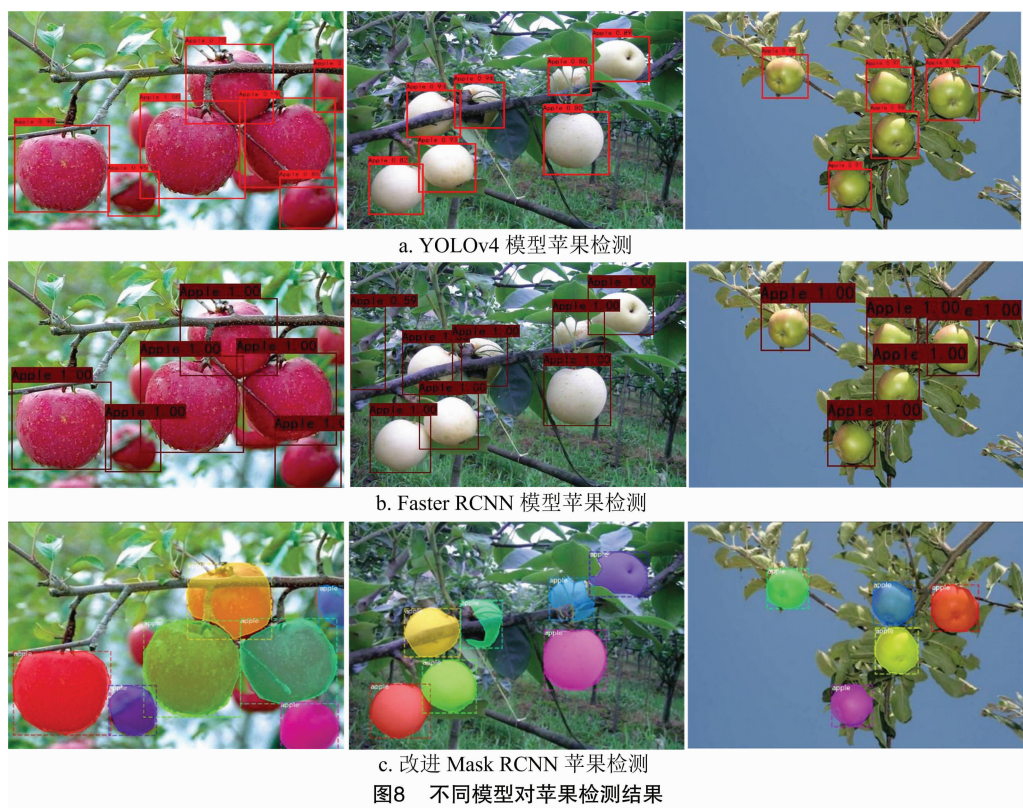


图8 不同模型对苹果检测结果

表 1 不同模型性能对比 %

模型	准确率	召回率	平均检测精度
Resnet101	89.91	90.90	90.45
Resnet50	83.33	85.71	67.19
Mobilenetv3	66.67	83.33	72.19
改进 Mobilenetv3	93.75	89.99	93.75

表 2 模型权重对比

模型	参数量	权重文件 (MB)	检测速度 (FPS)
Resnet101	6.40×10^7	245.50	5
Resnet50	4.52×10^7	171.41	9
Mobilenetv3	2.41×10^7	92.53	12
改进 Mobilenetv3	2.43×10^7	93.48	9

力机制加强模型特征提取能力。为了增强语义信息,采用 SPP 网络结构进行加强,并采用空洞卷积增强特征图的感受野,使模型泛化能力更强。

采用数据集进行模型的验证,在准确率方面比采用 Resnet101 网络、Resnet50 网络、Mobilenetv3 网络分别提高了 3.84、10.42、27.08 百分点;在召回率上比采用 Mobilenetv3 网络提高了 6.66 百分点,比 Resnet50 提高了 4.28 百分点,比 Resnet101 网络下降了 0.91 百分点;在 mAP 方面,比 Resnet101 网络

提高 3.30 百分点,比 Resnet50 提高了 26.56 百分点,比 Mobilenetv3 网络提高了 21.56 百分点。模型参数方面,改进后的参数量将为 2.43×10^7 ,模型的检测速度提高为 9 帧/s。本研究提出的改进 Mask RCNN 模型在精度提升的情况下,模型参数量大幅降低,检测速度有所提升,不仅能判断出苹果的位置,而且能对边缘和形状信息进行预测,更加满足采摘和机器人的实际应用场景。

参考文献:

[1] 祝前峰,陆荣鑑,李奉顺. 苹果采摘机械的研究现状与发展趋势[J]. 林业机械与木工设备,2021,49(5):4-9,15.

[2] 丁一,姬伟,许波,等. 苹果采摘机器人柔顺抓取参数自整定阻抗控制[J]. 农业工程学报,2019,35(22):257-266.

[3] 赵德安,刘晓洋,陈玉,等. 苹果采摘机器人夜间识别方法[J]. 农业机械学报,2015,46(3):15-22.

[4] 任亚婧,张宁宁,徐媛媛,等. 基于视觉识别的成熟苹果识别及采摘定位系统[J]. 现代电子技术,2021,44(11):73-77.

[5] Mazzia V, Khaliq A, Salvetti F, et al. Real-time apple detection system using embedded systems with hardware accelerators; an edge AI application[J]. IEEE Access,2020,8:9102-9114.

[6] 金玉成,高杨,刘继展,等. 采摘机器人深度视觉伺服手-眼协调规划研究[J]. 农业机械学报,2021,52(6):18-25,42.

[7] Jia W K, Zhang Y, Lian J, et al. Apple harvesting robot under information technology: a review[J]. International Journal of Advanced Robotic Systems,2020,17(3):172988142092531.



图9 不同状况下果实检测

- [8] Fan P, Lang G D, Guo P J, et al. Multi - feature patch - based segmentation technique in the gray - centered RGB color space for improved apple target recognition [J]. Agriculture, 2021, 11 (3) :273.
- [9] Liu X Y, Zhao D A, Jia W K, et al. A detection method for apple fruits based on color and shape features[J]. IEEE Access, 2019, 7: 67923 - 67933.
- [10] 张春龙, 张 楫, 张俊雄, 等. 近色背景中树上绿色苹果识别方法[J]. 农业机械学报, 2014, 45(10) :277 - 281.
- [11] 廖 巍, 郑立华, 李民赞, 等. 基于随机森林算法的自然光照条件下绿色苹果识别[J]. 农业机械学报, 2017, 48(增刊 1) :86 - 91.
- [12] 彭红星, 黄 博, 邵园园, 等. 自然环境下多类水果采摘目标识别的通用改进 SSD 模型[J]. 农业工程学报, 2018, 34(16) : 155 - 162.
- [13] 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.
- [14] 李大华, 包学娟, 于 晓, 等. 基于 YOLOv3 网络的自然环境下青苹果检测与识别[J]. 激光杂志, 2021, 42(1) :71 - 77.
- [15] 岳有军, 田博凯, 王红君, 等. 基于改进 Mask RCNN 的复杂环境下苹果检测研究[J]. 中国农机化学报, 2019, 40(10) :128 - 134.
- [16] Jia W K, Tian Y Y, Luo R, et al. Detection and segmentation of overlapped fruits based on optimized mask R - CNN application in apple harvesting robot [J]. Computers and Electronics in Agriculture, 2020, 172:105380.
- [17] He K M, Gkioxari G, Dollár P, et al. Mask R - CNN [C] // 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017.
- [18] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016 .
- [19] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3 [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019:1314 - 1324.
- [20] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/OL]. (2017 - 04 - 13) [2021 - 03 - 05]. <https://arxiv.org/abs/1704.04861>.
- [21] Hu J, Shen L, Albanie S, et al. Squeeze - and - excitation networks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018.
- [22] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37 (9) :1904 - 1916.
- [23] Wu L, Ma J, Zhao Y H, et al. Apple detection in complex scene using the improved YOLOv4 model [J]. Agronomy, 2021, 11 (3) :476.
- [24] Ying B Y, Xu Y C, Zhang S, et al. Weed detection in images of carrot fields based on improved YOLO v4 [J]. Traitement Du Signal, 2021, 38(2) :341 - 348.
- [25] Fu L S, Majeed Y, Zhang X, et al. Faster R - CNN - based apple detection in dense - foliage fruiting - wall trees using RGB and depth features for robotic harvesting[J]. Biosystems Engineering, 2020, 197:245 - 256.
- [26] Gao F F, Fu L S, Zhang X, et al. Multi - class fruit - on - plant detection for apple in SNAP system using Faster R - CNN [J]. Computers and Electronics in Agriculture, 2020, 176:105634.