



熊正午,吴瑞梅,黄俊仕,等.深度学习结合快速导向滤波识别自然环境下脐橙果实[J].江西农业大学学报,2022,44(3):736-746.

XIONG Z W,WU R M,HUANG J S,et al.Identification of navel orange fruits in natural environments based on deep learning combined with fast guided filtering[J].Acta agriculturae universitatis Jiangxiensis,2022,44(3):736-746.

深度学习结合快速导向滤波 识别自然环境下脐橙果实

熊正午¹,吴瑞梅²,黄俊仕²,李霸聪¹,戴仕明¹,艾施荣^{1*}

(1.江西农业大学 软件学院,江西 南昌 330045;2.江西农业大学 工学院,江西 南昌 330045)

摘要:【目的】为提高脐橙采摘机器人在自然环境下对脐橙果实进行识别定位的精度,提出采用深度学习结合快速导向滤波方法识别自然环境下脐橙果实。【方法】以赣南脐橙为对象,改进导向滤波方法,去除自然环境下脐橙图像的光照等噪声信息,突出图像颜色和纹理特征。采用带有残差模块的Darknet-53作为特征提取网络,将多尺度融合的3尺度检测网络减少为2尺度检测网络,引入GIoU边界损失函数代替原损失函数,并使用DBSCAN+Kmeans聚类算法,对训练数据集聚类分析,优化预测分支的先验框尺寸,通过迁移学习训练方法建立脐橙果实识别模型,设计单果、向光、背光、果实重叠、枝叶遮挡5组测试集的对比实验,并与其他几种识别模型性能进行比较。【结果】快速导向滤波方法能很好地去除脐橙图像在自然环境下光照及边缘模糊等噪音信息。当优化2个预测分支先验框尺寸时,改进后模型在5种环境下综合性能都优于其他网络,尤其在真实种植环境下识别准确率达到91.22%,召回率为97.30%,F1平均值为94.16%,识别速率约为26.48 fps。【结论】使用快速导向滤波结合深度学习方法建立的脐橙果实识别模型对自然环境下脐橙果实的识别具有较高的鲁棒性和实时性,为脐橙采摘机器人的视觉识别提供了技术支持。

关键词:脐橙果实;深度学习;导向滤波;YOLO v3

中图分类号:S126;S666.4 **文献标志码:**A **文章编号:**1000-2286(2022)03-0736-11

Identification of Navel Orange Fruits in Natural Environments Based on Deep Learning Combined with Fast Guided Filtering

XIONG Zhengwu¹, WU Ruimei², HUANG Junshi²,
LI Bacong¹, DAI Shiming¹, AI Shirong^{1*}

(1.College of Software,Jiangxi Agricultural University,Nanchang 330045,China;2.College of Engineering,Jiangxi Agricultural University,Nanchang 330045,China)

Abstract: [Objective] In order to improve the recognition and orientation accuracy of the navel orange fruit in the natural environment by the picking robot, this study proposes a deep learning combined with fast

收稿日期:2021-09-25 修回日期:2021-11-10

基金项目:国家自然科学基金项目(31460315)和江西省重点研发计划项目(2017ACF60004)

Project supported by the National Natural Science Foundation of China(31460315)and Key R & D Projects in Jiangxi Province(2017ACF60004)

作者简介:熊正午,orcid.org/0000-0001-5691-669X,122982874@qq.com;*通信作者:艾施荣,副教授,orcid.org/0000-0002-7866-0565,主要从事计算机视觉方向研究,aisrong@163.com。

guided filtering method to identify navel orange fruit in natural environment. [Methods] This study selected the Gannan navel oranges as the research material. In order to highlight the color and texture characteristics of the images, the improved guided filtering method was used to remove the noise information such as the illumination of the navel orange image in the natural environment. The Darknet-53 with residual module was used as the feature extraction network, the multi-scale fusion 3-scale detection network was reduced to a 2-scale detection network, the GIoU boundary loss function was introduced to replace the original loss function, the DBSCAN+Kmeans clustering algorithm was used to cluster the training data set, and the anchor box size of the prediction branches were optimized. The recognition model for navel orange fruit was established by the transfer learning training method, five test sets of single fruit, light, backlight, overlap and leaf occlusion were designed for the comparison experiment, and the performance was compared with that of other recognition models. [Results] The fast guided filtering method could remove the noise information of the navel orange image in the natural environment, such as illumination and edge blur. When the sizes of three prediction branches were optimized, the comprehensive performance of the improved model was better than other networks in the five environments, especially in the real planting environment, the recognition accuracy rate reached 91.22%, and the recall rate was 97.30%, the average value of F1-score was 94.16%, and the recognition rate was about 26.48 fps. [Conclusion] The results show that the navel orange fruit recognition model established by fast guided filtering combined with deep learning methods has higher robustness and real-time performance in the recognition of navel orange fruits in natural environments, thus providing technical support for the visual recognition of navel orange picking robots.

Keywords: navel orange fruit; deep learning; guided filter; YOLO v3

【研究意义】脐橙具有味道芳香浓甜、果肉营养丰富等特点,深受消费者喜爱。脐橙种植收益高,而目前脐橙采摘以人工为主,耗时长、时效差、采摘周期短,短期内采摘强度大,采摘成本占到全部生产成本的30%~40%^[1]。实现脐橙自动化智能化采摘已成为亟需解决的问题。随着智慧农业科技的应用,采摘机器人必将成为未来农业机器人发展的重要方向。而要实现采摘机器人工作在自然环境下果实成功采摘,必须能从复杂的自然环境中准确地识别目标,计算其三维空间坐标,将果实进行精确定位,指导机器人末端执行机构精确完成目标果实的采摘。【前人研究进展】国内外研究者采用机器视觉技术结合图像处理技术实现非接触式单个果实的快速识别进行了大量研究^[2-6]。在自然环境下果实生长环境和生长背景较为复杂,传统图像识别算法难以满足识别精度要求,而深度卷积神经网络对如强光、弱光、果实重叠和枝叶遮挡等自然环境下的目标识别和检测都具有较强的鲁棒性和识别性能,基于卷积神经网络的深度学习算法已成为计算机视觉研究的主要方向^[2]。彭红星等^[3]提出一种基于(single shot multibox detector, SSD)改进的目标检测方法,使用ResNet-101主干网络替代原始SSD模型中的VGG16主干网络,该模型对荔枝、皇帝柑、苹果和脐橙等4种水果的平均识别精度达到88.4%。熊俊涛等^[4]利用Faster R-CNN网络模型对自然环境下绿色柑橘果实进行识别研究,其综合识别率达到77.45%。Mao等^[5]提出一种基于多路径卷积神经网络(MPCNN)的黄瓜果实识别新方法,结合颜色特征和支持向量机算法,识别率达90%以上。Tian等^[6]通过对YOLO v3进行改进,实现了对处于生长时期不同阶段苹果的有效识别,该模型对生长期的苹果识别的F1值达到78.3%、膨果期的F1值达到79.1%、成熟期的F1值达到80.9%。

【本研究切入点】面对自然环境下果实采摘的识别问题,自然环境下光照等噪声信息会导致采集图像模糊,影响结果分析。目前,有研究者采用导向滤波方法去除各类噪声信息,该方法能有效平缓图像噪声,同时又能运用于图像去雾^[7]和红外图像的增强和融合等^[8-9]。考虑去除自然环境下光照等噪声因素及所建立果实识别模型的识别速度和准确率,快速导向滤波图像去噪方法和带有残差模块的YOLO v3目标检测方法能解决上述问题。【拟解决的关键问题】为提高在自然环境下脐橙果实的识别准确率,本文以赣南脐橙为研究对象,提出一种基于多尺度卷积神经网络的非接触的脐橙果实识别方法,改进YOLO

v3 目标检测算法,将多尺度融合的 3 尺度检测网络减少为 2 尺度检测网络,设计自然环境下脐橙果实快速识别模型,模型融合快速导向滤波算法,使用 GIoU 边界损失函数替代原有损失函数,通过 DBSCAN+Kmeans 聚类算法,对训练数据集进行聚类分析,优化 2 个预测分支的先验框尺寸。在脐橙真实种植环境下,利用自然环境下脐橙果实快速识别模型,可以从脐橙数据集图像中去除光照等噪声信息,突出脐橙图像的颜色和纹理特征,同时获取目标类别与预测框位置,设计单果、向光、背光、果实重叠、枝叶遮挡 5 组不同环境下测试集的对比实验,验证在复杂自然环境下模型的识别性能。

1 材料与方 法

1.1 脐橙数据采集

脐橙果实图像采集于江西绿萌科技控股有限公司赣南脐橙基地(2019 年 11 月、2020 年 11 月),采集时间为 08:00—17:00,图像采集设备为 Canon EOS 70D,图像格式为 JPG 格式,图像分辨率为 4 032×3 024 像素,脐橙的种植方式为篱壁式和宽行窄株式两种,共采集自然环境下原始图像 2 400 幅,如图 1 所示。



图 1 自然环境下脐橙果实图像样本

Fig.1 Sample image of navel orange fruit in natural environment

1.2 脐橙数据集及标签制作

本研究脐橙图像背景包括向光、背光、独立与相邻、遮挡等情况(图 1 所示),其中独立与相邻是指单个脐橙果实和多个脐橙果实且没有枝叶遮挡或枝叶遮挡小于 1/4 的情况,遮挡表示单个脐橙果实和多个脐橙果实有枝叶遮挡超过 1/4。从拍摄的 2 400 幅脐橙果实图像中筛选出清晰的、具有代表性的脐橙图像 1 500 幅。将经过预处理后的脐橙图像样本存入 orange 文件夹下,并统一批量重命名为 000x.jpg 格式。使用图像标注软件 labelImg 进行标签制作,按照 POSCAL VOC2007 数据集格式划分训练集和测试集。

1.3 脐橙图像快速导向滤波方法预处理

快速导向滤波(fast guided filter)是一种基于局部线性模型的保边滤波算法,该方法在导向滤波基础上进行改进,在导向图像指导滤波过程利用线性可变原理,有效地减少噪声,增强图像信息细节,准确实现边保持,减少光晕现象,同时减少算法的运算时间^[10-12]。快速导向滤波去噪实现原理如下,其中, p 为输入图像, q 为输出图像, I 为导向图像,则滤波模型定义为:

$$q_i = \sum_j W_{ij}(I) p_j \tag{1}$$

式中, i, j 为图像的像素点, W_{ij} 为导向图像 I 和输入图像 p 的函数。

在以像素点 k 为中心、 r 为半径的局部方形窗口 ω_k 中, q 和 I 之间存在线性关系,线性核函数可表示为:

$$q_i = a_k I_i + b_k, \quad \forall i \in \omega_k \tag{2}$$

式中, a_k 表示该线性函数的权重值, b_k 表示该线性函数的偏置值。

a_k 和 b_k 在中心点 k 及半径 r 都不变的局部窗口 ω_k 中为定值, a_k 和 b_k 值由最小二乘法最小化输出图像 q 与输入图像 p 之间的平方差求得,即:

$$E(a_k, b_k) = \sum_{i \in \omega_k} \left((a_k I_i + b_k - p_i)^2 + \epsilon a_k^2 \right) \tag{3}$$

式中, ϵ 表示正则化参数, 参数 ϵ 能防止 a_k 值过大, 且保证输入图像 p 与输出图像 q 之间的差值减小, 该参数值需经大量试验并反复对比图像质量得到, a_k 和 b_k 值由线性变换得到:

$$a_k = \frac{1}{|\omega|} \frac{\sum_{i \in \omega_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \epsilon} \quad (4)$$

$$b_k = \bar{p}_k - a_k \mu_k \quad (5)$$

式中, μ_k 为 I 窗口 ω_k 的均值, σ_k^2 为 I 窗口 ω_k 的方差, $|\omega|$ 表示窗口 ω_k 的像素总数, \bar{p} 表示 p 在窗口 ω_k 的像素均值。

由于像素 i 存在多个窗口 ω_k , 故取所有窗口的均值为系数 a_k 和 b_k :

$$q_i = \bar{a}_i I_i + \bar{b}_i \quad (6)$$

算法是在导向滤波的基础上通过下采样减少像素点, 在计算线性函数的权重值 a_k 与线性函数的偏置值 b_k 的均值后进行上采样恢复到原有尺寸大小。如缩放比例为 S , 则缩小后像素点的个数为 N/S^2 , 时间复杂度为 $O(N/S^2)$ 。

1.4 深度学习算法及优化

YOLO(You only look once)是深度学习中一种基于端到端的单阶段目标检测与识别算法, 其主要思想是将一个分类检测问题转换为一个数学回归问题, 实现目标位置与分类识别的结合, 具有实时性等突出特征。YOLO对输入图像的全局区域进行学习, 可加快训练速度, 且能更好地区分目标和背景。在主干网络完成对脐橙果实的特征提取后, 通过 Bounding box 预测的方式, 能够同时完成目标位置和类别的预测任务。

1.4.1 YOLO 网络模型 YOLO 有一系列网络模型如 YOLO v1、YOLO v2、YOLO v3, 其中 YOLO v1 存在目标定位不准确与网络模型检测精度低等缺陷, YOLO v2 为进一步提升网络模型性能, 通过修改网络结构, 同时加入批量归一化、锚点框和高分辨率分类器等方法进行改进, 提高识别速度, 但对目标重叠的情况识别率较差, YOLO v3 改变网络模型结构, 引入多尺度融合训练、锚点框选择机制和残差模块等操作, 极大提升了网络模型性能。本文选用 YOLO v3 网络模型, 根据脐橙果实在自然环境下的生长特征, 对 YOLO v3 网络模型中的相关参数进行调整和改进, 实现自然环境下脐橙果实的识别, 由于小目标识别存在漏检率高等缺陷, 主干网络相对于 YOLO v2 的 Darknet-19 网络, 采用综合性能更好的 Darknet-53 网络来进行特征提取, 将多尺度融合的 3 尺度检测网络减少为 2 尺度检测网络, 该网络增加了残差模块, 运用残差网络, 可保证网络在结构较深情况下, 模型能够继续收敛且不断被训练, 获得高维特征, 从而提高预测准确率^[13]。

1.4.2 多尺度检测模型改进 为了更好地对脐橙目标进行检测, 本文在尺度融合部分对 YOLO v3 网络进行改进, 改进的 YOLO v3 网络结构如图 2 所示。图中 DBL 模块是该网络的基本组件, 包括卷积、批量归一化(batch normalization, BN)和激励(leaky ReLU)操作; Res Block 表示残差模块, Res_unit 表示残差单元, 使用残差结构, 可以使网络结构更深, 从而获得高维特征。该网络结构中, 每一个 Res Block 由 n 个 Res_unit 构成, 且每个 Res_unit 都需要一个 add 层, 故共含有 23 个 add 层。张量拼接(concat)功能主要是为了实现网络多层特征的复用和融合, 将卷积前后的特征图进行拼接; 相比于其他网络, 该网络通过改变卷积核的步长来实现张量尺寸的变换, 删去所有的池化层和全连接层, 进而减少检测时间。同时为了提高分类准确度, 将用于多标签多分类的逻辑回归层替代原单标签的 softmax 层。该网络为针对果园中脐橙果实进行检测, 将原始高分辨率的图片统一转化成 512×512 的图片, 以 YOLO v3 网络为基础, 采用 Darknet-53 为特征提取网络, 将多尺度融合的 3 尺度检测网络减少为 2 尺度检测网络, 最后在 16×16 和 32×32 尺度的边界框上进行目标检测。

1.4.3 损失函数改进 为了提高自然环境下脐橙果实快速识别模型的识别精度, 通过优化训练模型的损失函数, 使得目标置信度、目标位置和类别三者的训练误差达到平衡^[14]。YOLO v3 的损失函数为:

$$L = k_{box} \delta(box) + k_{conf} \delta(conf) + k_c \delta(c) \quad (7)$$

式中, $\delta(box)$ 为预测边界框 box 的误差函数; $\delta(conf)$ 为预测置信度 $conf$ 的误差函数; $\delta(c)$ 为预测类别 c 的误差函数; k_{box} 、 k_{conf} 、 k_c 为误差权重系数; L 为损失函数。

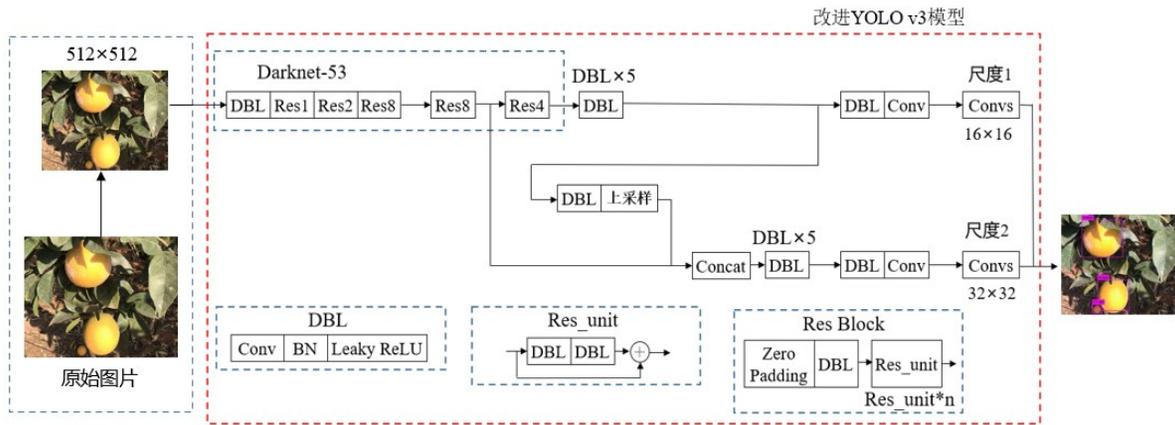


图2 改进的YOLO v3网络结构

Fig.2 Network structure of improved YOLO v3

对于目标检测的边界框回归指标,大多数网络均采用 $L1$ 和 $L2$ 范数作为边界框回归损失函数,其计算损失值实际是通过间接的相对距离计算,但没有考虑在边界框回归中真实框和预测框之间的交并比 IoU 的作用, IoU 定义如下:

$$IoU = \frac{|P \cap G|}{|P \cup G|} \tag{8}$$

式中, P 是目标的预测框; G 是目标的真实框; IoU 是真实框和预测框之间的交并比。

IoU 是比较预测框与真实框之间最常用的度量标准,但 IoU 无法直接作为回归距离,为了提高位置精度,本文使用 $GIoU$ 距离作为边界框的损失评价标准^[15], $GIoU$ 距离计算具体如下:

(1) 输入预测框 $B^p = (x_1^p, y_1^p, x_2^p, y_2^p)$ 和真实框坐标 $B^g = (x_1^g, y_1^g, x_2^g, y_2^g)$ 纵横坐标,按照从小到大顺序;

(2) 计算矩形框 B^p 、 B^g 面积及交集 I 、并集 U 面积;

(3) 计算 IoU : $IoU = \frac{I}{A^p + A^g - I}$;

(4) 对 A, B 寻找最小矩形框 B^c , $x_1^c = \min(\hat{x}_1^p, \hat{x}_1^g), x_2^c = \max(\hat{x}_2^p, \hat{x}_2^g)$, $y_1^c = \min(\hat{y}_1^p, \hat{y}_1^g), y_2^c = \max(\hat{y}_2^p, \hat{y}_2^g)$, 计算其面积 A^c ;

(5) 计算 $GIoU$: $GIoU = IoU - \frac{A^c - U}{A^c}, U = A^p + A^g - I$;

(6) 输出 $GIoU$ 距离 $L_{GIoU} = 1 - GIoU$ 。

将式(7)中 $\delta(box)$ 用 $GIoU$ 距离 $\sum_{i \in Pos} L_{GIoU}(B_i^m - \hat{B}_j^m)$ 进行替代,使用改进的损失函数进行模型训练。

1.4.4 目标先验框聚类算法 原始YOLO v3中先验框尺寸是基于COCO数据集和PASCAL VOC数据集先验框大小设定的,相较于COCO数据集和PASCAL VOC数据集,本文训练所使用的脐橙数据集中样本类型较少、尺度跨度较小,为了保证先验框与其相邻的预测框之间能有更大的 IoU ,本文对聚类算法进行优化,在所构建的脐橙数据集上,使用DBSCAN(density-based spatial clustering of applications with noise)聚类算法+Kmeans聚类算法对训练数据集聚类分析。首先使用DBSCAN聚类算法分析聚类个数和聚类包含的点,可以解决聚类时噪音点或者孤立点对聚类结果的影响,同时可以为Kmeans聚类算法提供准确的 K 值;再通过Kmeans聚类算法对DBSCAN聚类结果分析确定聚类中心^[16]。

本文使用优化后的聚类算法对脐橙数据集进行先验框聚类分析,经DBSCAN聚类算法分析确定 K 值为6,然后使用Kmeans聚类算法进行迭代后选取的先验框尺寸为(37×30)、(43×40)、(50×34)、(57×43)、(68×54)、(86×67),得到2个预测分支的先验框尺寸。表1展示了聚类算法改进前后,不同先验框尺寸和检测结果对比,改进后的聚类算法得到的先验框尺寸对检测结果具有较大的提升,识别速度比原网络提高到27.42 fps, $F1$ 值达到了93.23%,提高了0.52个百分点。

表 1 不同先验框检测结果对比
Tab.1 Comparison of detection results of different anchor box

方法 Method	先验框尺寸 Anchor box size	F1 值/% F1-score	识别速度 S/fps Speed
YOLO v3	(10×13)(16×30)(33×23)(30×36)(62×45)(59×119)(116×90) (156×198)(373×326)	92.71	25.86
只改进聚类算法的 YOLO v3	(37×30)(43×40)(50×34)(57×43)(68×54)(86×67)	93.23	27.42

1.4.5 脐橙果实识别模型架构及评价指标 本文构建的自然环境下脐橙果实快速识别模型主要包含数据预处理和脐橙果实快速检测网络两个模块,数据预处理模块是在采集数据样本后进行图像处理,如减小光照影响,数据增强以及归一化处理等,快速检测网络模块是基于 Darknet-53 为特征提取网络,快速提取训练集和验证集中在自然环境下脐橙果实的主要特征,结合多尺度检测模块,同时获取目标类别与预测框位置,根据改进后的损失函数,使优化器不断调整参数,最后训练出脐橙识别模型。自然环境下脐橙果实快速识别模型架构如图 3 所示。

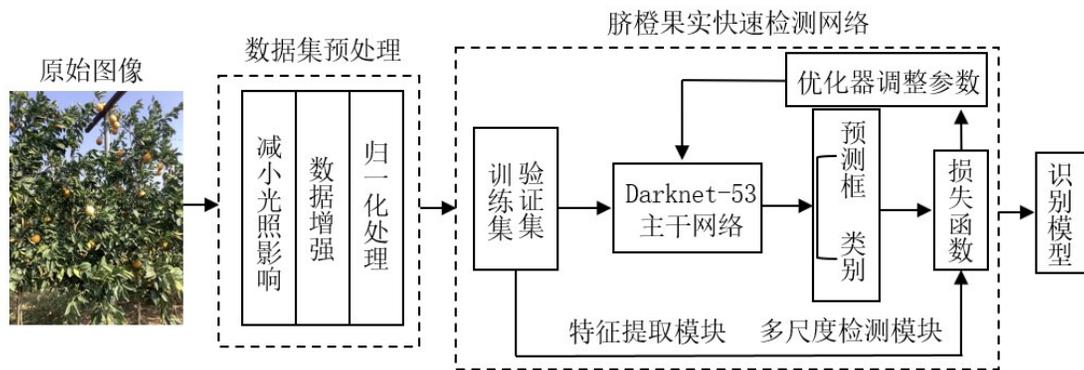


图 3 自然环境下脐橙果实快速识别模型的架构

Fig.3 Framework of fast recognition model for navel oranges under natural environment

本文以准确率(precision, P)、查全率(recall, R)、F1 值($F1$ -score)以及识别速度(speed, S)作为脐橙识别模型的性能评价指标,其中 $F1$ 值能够同时考虑准确率和召回率两者,其计算公式为:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{10}$$

$$F1 - \text{score} = \frac{2\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

式中, TP 表示实际为正样本且被检测为正样本的数量, FP 表示实际为负样本且被检测为正样本的数量, FN 表示实际为正样本且被检测为负样本的数量。

2 结果与分析

2.1 导向滤波结果分析

为有效保留脐橙图像的边缘信息,采用快速导向滤波法去噪时,导向图像 I 和输入图像 p 都采用原始图像。由式(1)~(6)知,参数 a_k 和 r 直接影响方法的去噪程度和边缘保留度,由式(3)知, ϵ 的取值范围在 $0 \sim 1$,且 ϵ 在 a_k 分母部分,直接对 a_k 产生影响,故可改变 ϵ 值来改变 a_k 的大小。而 r 值直接控制窗口的大小,从而影响局部像素点的多少,当 ϵ 和 r 取不同值时,去噪效果不同,经过对比分析,当 $\epsilon=0.01, r=4$ 时,去噪效果最佳(图 4 红框所示)。由图可看出,原始脐橙图像经过快速导向滤波去噪及图像增强后,脐橙表面的颜色和纹理信息更加明显,在背景对比下更加突出,说明采用导向滤波方法能有效去除自然环境下背景噪音对脐橙图像信息的影响。

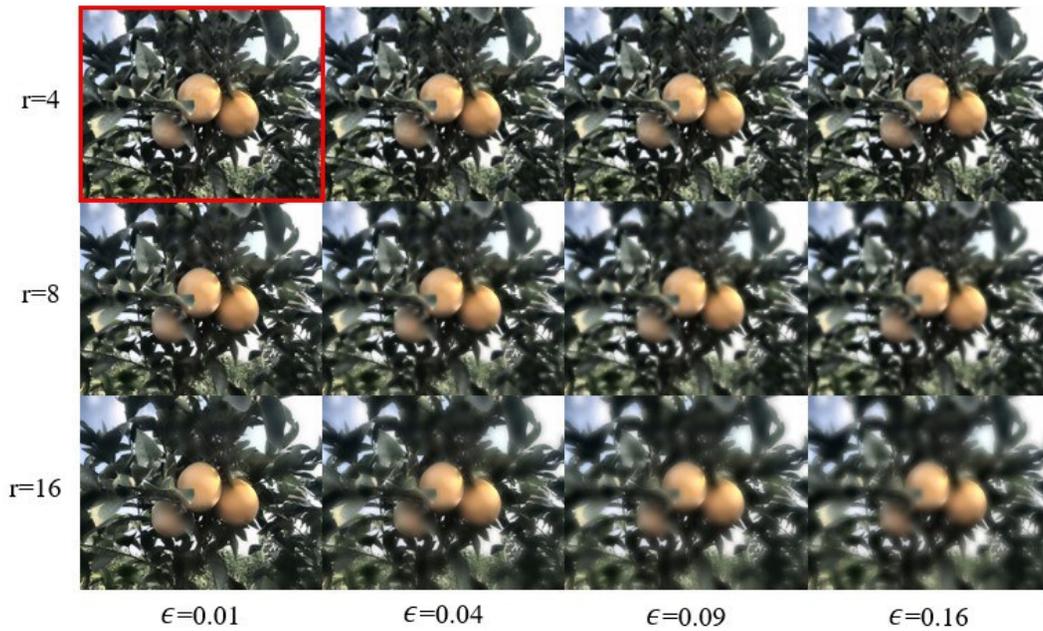


图 4 不同参数下导向滤波去噪效果

Fig.4 The denoising results by guided filtering with different parameters

2.2 不同去噪方法预处理效果对比

常用的图像去噪方法有最小二乘滤波、中值滤波、K近邻平均滤波、小波阈值等,中值滤波是一种基于排序统计理论的经典去噪方法,能有效保护图像的边缘信息,小波阈值是一种基于傅里叶分析的去噪方法,能得到更多的边缘细节。

图 5 为基于中值滤波、小波阈值和导向滤波方法对脐橙图像去噪及增强效果图,由图 5 可看出,由快速导向滤波预处理后的脐橙图像,其颜色和纹理特征明显优于其他两种方法(图 5 中红框所示),说明导向滤波去除自然环境下各种噪音效果优于其他两种方法。

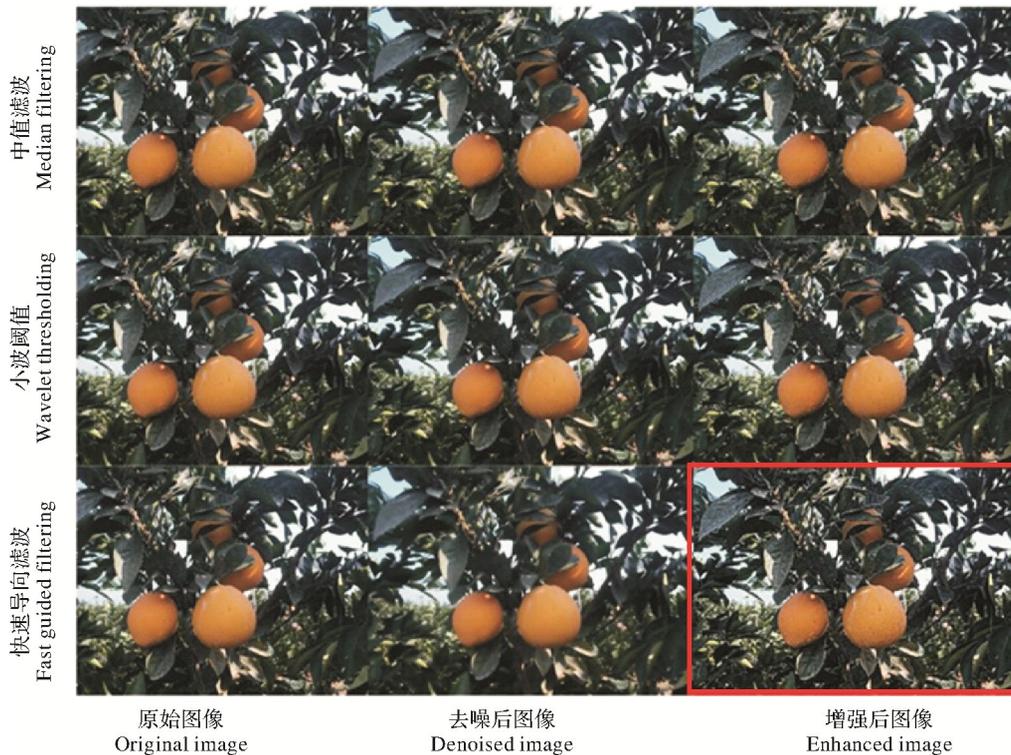


图 5 不同去噪方法对脐橙图像去噪及增强对比

Fig.5 Comparison of denoising and enhancement results with different denoising methods on navel orange leaves images

2.3 网络模型训练与结果分析

2.3.1 参数设置与模型训练 试验采用 Darknet 深度学习框架,进行训练所用的 PC 端配置为:Ubuntu 16.04LTS 操作系统,英特尔 Xeon(至强)W-2133@3.60GHz 六核,显卡为两张 Nvidia Quadro P2000, NVIDIA 440.82 驱动, CUDA 10.2 版本, CUDNN 神经网络加速库版本为 7.6.5。网络参数是对 YOLO v3 论文^[7]中部分参数进行修改,经多次试验后比较选取效果最优的参数。输入样本为 512*512 像素,每次迭代训练的样本数(bath)为 64 个,共 32 批次,动量因子(momentum)为 0.9,权重衰减系数(weight decay)为 0.000 5;训练的最大迭代次数(max_bathes)为 12 000,初始学习率(learning rate)为 0.000 5,学习率调整策略为 steps,学习率分别在迭代次数为 9 600 和 10 800 时减小到之前的 0.1 倍,同时使用对比度调整、图像翻转、虚化等方法对训练集进行数据集增强,能够训练出更适用的脐橙果实识别模型。

采用迁移学习方法对样本集进行训练,训练集与验证集比重为 8:2,使用 Darknet-53 网络在 COCO 数据集上的预训练权重作为模型的初始权重。对脐橙数据集进行模型训练,得到自然环境下脐橙果实的识别模型。对训练集和验证集进行 12 000 次迭代训练后,其误差变化曲线如图 6 所示,训练误差曲线可以清晰看出损失函数值的变化趋势,在前 3 000 次迭代过程中,损失函数值下降较快,网络快速拟合;之后开始缓慢下降,中间有两次较大的波动,5 000 次迭代后损失函数基本趋于平稳,损失函数值最低达到了 0.15 左右(图 6 虚线所示),由此可知改进的 YOLO v3 网络性能优良。

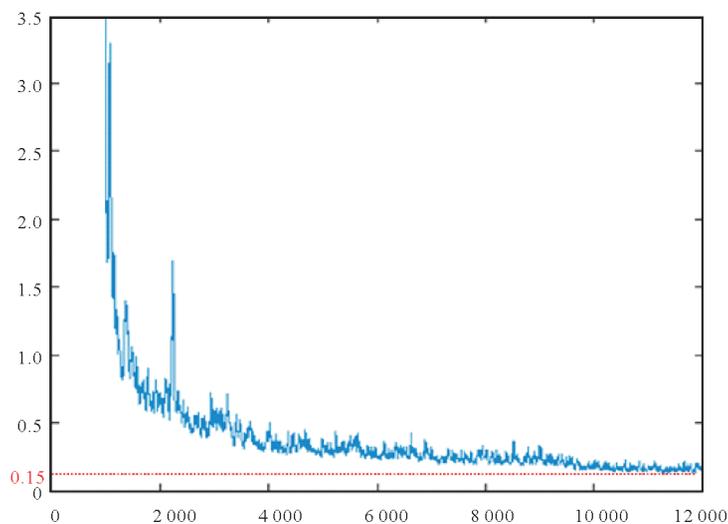


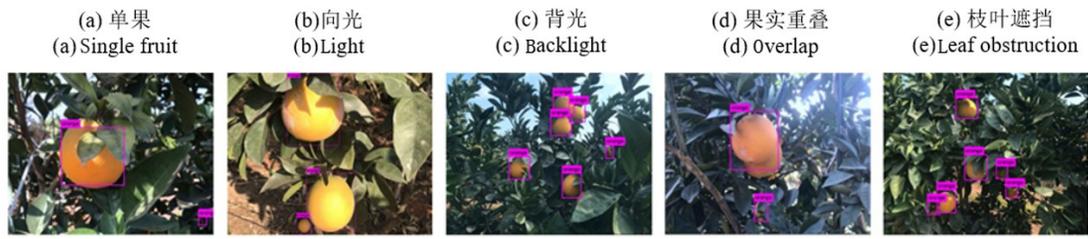
图6 训练误差变化曲线

Fig.6 Training error curve

2.3.2 模型识别效果及对比 为了验证改进后模型对自然环境下脐橙果实的识别效果,设计了改进后模型与 YOLO v3 模型的识别效果对比试验。考虑不同网络在单果、向光、背光、果实重叠、枝叶遮挡等情况,脐橙果实识别效果对比试验结果如图 7 所示。

图 7(a)是 YOLO v3 模型识别效果图,图 7(b)是本文在 YOLO v3 基础上改进后模型识别效果图,由图可看出,YOLO v3 识别模型在单果大目标物体图像上只能大致识别出脐橙主体,但边缘出现了错检情况;改进后模型能够完整的识别出脐橙整体,且没有出现错检情况。YOLO v3 识别模型在向光及背光图像上只能识别出脐橙主体,且出现了漏检情况;改进后模型在这两种图像上均能完整识别出脐橙整体,且对小目标物体识别效果理想,未出现漏检情况。YOLO v3 识别模型对果实重叠及枝叶遮挡图像上识别效果不佳,把重叠的 3 个脐橙识别成 1 个,对枝叶遮挡严重和图像边缘小物体均出现了漏检情况;改进后模型能把重叠的 3 个脐橙都识别出来,对枝叶遮挡严重和图像边缘小物体也能成功识别出来。两组测试结果说明了改进后模型的识别效果更优。

2.3.2 模型识别性能及对比 为进一步准确评估改进后模型对自然环境下脐橙果实的识别性能,将改进后模型与 YOLO v3 模型、单阶段目标检测模型中的代表 SSD 模型以及两阶段目标检测模型中的代表 Faster R-CNN 模型进行识别性能对比,选择了 50 幅未参与训练的脐橙图像对模型性能进行对比试验,表 2 为 4 种识别模型的检测结果。



(a) YOLO v3 模型识别效果
(a) YOLO v3 model recognition renderings



(b) 基于 YOLO v3 改进模型识别效果
(b) Improved model recognition renderings based on YOLO v3

图 7 脐橙果实识别图

Fig.7 Navel orange fruit identification map

由表 2 可以看出,改进后模型相较于同为单阶段模型的 SSD 模型, $F1$ 值提高了 3.39 个百分点,各项指标均优于 SSD 模型,但 Faster R-CNN 模型的 $F1$ 值仍高于改进后模型,这是由于改进后模型与 Faster R-CNN 模型相比,少了生成候选区域这一步骤,所以 $F1$ 值比 Faster R-CNN 低了 1.26 个百分点。虽然两阶段模型相较于单阶段模型而言, $F1$ 值更好,但单阶段模型的识别速度在同等配置情况下远高于两阶段模型, YOLO v3 模型和 SSD 模型的识别速度分别为 25.86 fps 和 22.56 fps,均远超前于 Faster R-CNN 的 1.44 fps,相较于 YOLO v3 模型,虽然加入了图像预处理模块,但由于将 3 尺度检测网络减少为 2 尺度检测网络,改进后模型的识别速度提升达到了 26.48 fps,实时性较好,且 $F1$ 值提高了 1.45 个百分点。试验结果表明,改进后模型综合性能最好,可以同时满足脐橙采摘机器人对脐橙果实识别精度和识别速度的需求。

表 2 4 种识别模型的检测结果

Tab.2 Detection results of four recognition models

模型 Model	准确率 P/% Precision	召回率 R/% Recall	$F1$ 值 I/% $F1$ -score	识别速度 S/fps Speed
YOLO v3	89.45	96.22	92.71	25.86
SSD	87.86	93.87	90.77	22.56
Faster-RCNN	93.44	97.48	95.42	1.44
改进 YOLO v3 模型	91.22	97.30	94.16	26.48

3 结论与讨论

为提高脐橙果实在自然环境下的识别率,本文根据脐橙果实在自然环境下的生长特点,通过改进导向滤波方法去除自然环境下脐橙图像的光照等噪声信息,突出图像的颜色和纹理特征。在 YOLO v3 方法上基础上进行改进,采用带有残差模块的 Darknet-53 作为特征提取网络,将多尺度融合的 3 尺度检测网络减少为 2 尺度检测网络,引入 GIoU 边界损失函数代替原损失函数,使用 DBSCAN+Kmeans 聚类算法,分析优化出两个预测分支的先验框尺寸,采用迁移学习方法进行模型训练,并与其他几种识别模型性能进行比较。研究表明,改进后模型对自然环境下脐橙果实的识别准确率达到 91.22%,召回率为

97.30%, $F1$ 平均值为 94.16%, 识别速度为 26.48 fps, 综合性能明显优于其他模型, 为脐橙采摘机器人提供了可行的视觉识别思路。

在测试脐橙果实识别模型的过程中发现, 模型无法对遮挡较为严重的脐橙果实进行识别, 为进一步提高模型的识别性能, 今后研究通过增强脐橙数据集和优化网络结构等方法, 进行多种显著特征的提取, 提高网络在各类自然环境下的识别能力, 以适应果园中复杂多变的环境。目前改进的 YOLO v3 模型仅应用于成熟脐橙目标检测方面, 而实际应用中往往会有不同成熟度的脐橙目标, 因此将来可扩充不同成熟度的脐橙数据集, 从而对不同成熟度的脐橙进行实时检测还可进一步研究。

致谢: 江西绿萌科技控股有限公司对本研究给予了支持, 谨致谢意!

参考文献 References:

- [1] 彭婷, 黄鹤. 赣南脐橙供应链管理研究[J]. 合作经济与科技, 2017(17): 118-120.
PENG T, HUANG H. Research on supply chain management of gannan navel orange[J]. Co-operative economy & science, 2017(17): 118-120.
- [2] 闫建伟, 赵源, 张乐伟, 等. 基于残差网络的自然环境中刺梨果实的识别[J]. 中国农机化学报, 2020, 41(10): 191-196.
YAN J W, ZHAO Y, ZHAO L W, et al. Recognition of *Rosa roxbunghii* fruit in natural environment based on residual network[J]. Journal of Chinese agricultural mechanization, 2020, 41(10): 191-196.
- [3] 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.
PENG H X, HUANG B, SHAO Y Y, et al. General improved SSD model for picking object recognition of multiple fruits in natural environment[J]. Transactions of the Chinese society of agricultural engineering, 2018, 34(16): 155-162.
- [4] 熊俊涛, 刘振, 汤林越, 等. 自然环境下绿色柑橘视觉检测技术研究[J]. 农业机械学报, 2018, 49(4): 45-52.
XIONG J T, LIU Z, TANG L Y, et al. Visual detection technology of green citrus under natural environment[J]. Transactions of the Chinese society for agricultural machinery, 2018, 49(4): 45-52.
- [5] MAO S, LI Y, MA Y, et al. Automatic cucumber recognition algorithm for harvesting robots in the natural environment using deep learning and multi-feature fusion[J]. Computers and electronics in agriculture, 2019, 170: 105254.
- [6] YUNONG T, GUODONG Y, ZHE W, et al. Apple detection during different growth stages in orchards using the improved YOLO-V3 model[J]. Computers and electronics in agriculture, 2019, 157.
- [7] 田源. 基于均值聚类和循环导向滤波算法的农业图像去雾研究[J]. 计量学报, 2021, 42(3): 308-313.
TIAN Y. Agriculture image dehazing based on mean clustering and loop guided filtering[J]. Acta metrologica Sinica, 2021, 42(3): 308-313.
- [8] 王笛, 沈涛, 孙宾宾, 等. 基于大气灰度因子的红外图像增强算法[J]. 激光与红外, 2019, 49(9): 1135-1140.
WANG D, SHEN T, SUN B B, et al. Infrared image enhancement algorithm based on atmospheric gray factor[J]. Laser & Infrared, 2019, 49(9): 1135-1140.
- [9] 罗晓清, 袁衬衬, 柴鹏飞, 李凯. 面向苹果品质检测的图像融合方法[J]. 食品与生物技术学报, 2019, 38(9): 33-40.
LUO X Q, YUAN C C, CHAI P F, et al. Image fusion for apple quality detection[J]. Journal of food science and biotechnology, 2019, 38(9): 33-40.
- [10] HE K, SUN J, TANG X. Guided image filtering[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 35(6): 1397-1409.
- [11] JIN Z, WEIQI J, LI L, et al. Multiscale infrared and visible image fusion using gradient domain guided image filtering[J]. Infrared physics and technology, 2018, 89.
- [12] 胡潇, 熊爱华, 黄俊仕, 等. 计算机视觉结合引导滤波方法快速量化茶叶叶底品质[J]. 江西农业大学学报, 2019, 41(3): 601-609.
HU X, XIONG A H, HUANG J S, et al. Fast Quantitative Analysis of Quality of Infused Tea Leaves by Computer Vision Combined with Guided Filtering Method[J]. Acta Agriculturae Universitatis Jiangxiensis, 2019, 41(3): 601-609.
- [13] 柳长源, 王琪, 毕晓君. 多目标小尺度车辆目标检测方法的研究[J]. 控制与决策, 2021, 36(11): 2707-2712.

- LIU C Y, WANG Q, BI X J. Research on multi-target and small-scale vehicle target detection method[J]. Control and decision, 2021, 36(11): 2707-2712.
- [14] 刘芳, 刘玉坤, 林森, 等. 基于改进型YOLO的复杂环境下番茄果实快速识别方法[J]. 农业机械学报, 2020, 51(6): 229-237.
- LIU F, LIU Y K, LIN S, 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.
- [15] 杨传栋, 刘桢, 马翰宇, 等. 一种基于改进Yolov3的弹载图像多目标检测方法[J]. 弹箭与制导学报, 2020, 40(4): 149-153.
- YANG C D, LIU Z, MA H Y, et al. A Multi-target detection method for missile-borne images based on improved YOLOv3[J]. Journal of projectiles, rockets, missiles and guidance, 2020, 40(4): 149-153.
- [16] 岳晓新, 贾君霞, 陈喜东, 等. 改进YOLO V3的道路小目标检测[J]. 计算机工程与应用, 2020, 56(21): 218-223.
- YUE X Y, JIA J X, CHEN X D, et al. Road small target detection algorithm based on improved YOLO V3[J]. Computer engineering and applications, 2020, 56(21): 218-223.
- [17] REDMON J, FARHAD I A. YOLOv3: an incremental improvement[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [18] 毕松, 高峰, 陈俊文, 张潞. 基于深度卷积神经网络的柑橘目标识别方法[J]. 农业机械学报, 2019, 50(5): 181-186.
- BI S, GAO F, CHEN J W, et al. Detection method of citrus based on deep convolution neural network[J]. Transactions of the Chinese society for agricultural machinery, 2019, 50(5): 181-186.
- [19] 熊俊涛, 郑镇辉, 梁嘉恩, 等. 基于改进YOLO v3网络的夜间环境柑橘识别方法[J]. 农业机械学报, 2020, 51(4): 199-206.
- XIONG J T, ZHENG Z H, LIANG J E, et al. Citrus detection method in night environment based on improved YOLO v3 network[J]. Transactions of the Chinese society for agricultural machinery, 2020, 51(4): 199-206.
- [20] 王荣, 史再峰, 高荣华, 等. 多变环境下基于多尺度卷积网络的猪个体识别[J]. 江西农业大学学报, 2020, 42(2): 391-400.
- WANG R, SHI Z F, GAO R H, et al. Individual identification of pig based on multi-scale convolutional network in a variable environment[J]. Acta agriculturae universitatis Jiangxiensis, 2020, 42(2): 391-400.
- [21] 唐熔钗, 伍锡如. 基于改进YOLO-V3网络的百香果实时检测[J]. 广西师范大学学报(自然科学版), 2020, 38(6): 32-39.
- TANG R C, WU X R. Real-time detection of passion fruit based on improved YOLO-V3 network[J]. Journal of Guangxi normal university(natural science edition), 2020, 38(6): 32-39.