

基于改进 Faster R-CNN 的百香果自动检测

涂淑琴^a, 黄健^a, 林跃庭^a, 李嘉林^a, 刘浩锋^a, 陈志民^b
(华南农业大学 a. 数学与信息软件学院; b. 基础实验与实践训练中心, 广州 510642)



摘要: 针对自然场景下百香果果实密集, 生长环境相对复杂, 大规模种植带来人工识别、采摘和估计产量困难等问题, 提出了改进 Faster R-CNN 的百香果目标检测算法, 实现无遮挡、遮挡、重叠和背景四类果实自动检测和产量预测。该方法首先采用 ResNet 网络融合 FPN 对百香果进行多尺度特征提取; 然后采用 RPN 网络提取 ROI 区域; 最后, 通过全连接层实现百香果分类和检测。经测试集验证, 该方法在 4 类情况下检测的平均精确率达到 87.98%, 其平均准确率和召回率分别达到 90.79% 和 90.47%, 每幅图片的检测时间在 0.178 s 左右; 产量估算中, 其准确率为 96.80%。结果表明, 基于 FPN + ResNet-101 特征提取的 Faster R-CNN 目标检测算法能应用于自然场景下百香果的快速、准确检测和产量估算。

关键词: 百香果检测; Faster R-CNN; ResNet-50/101; FPN

中图分类号: TP 391.4; S 126 **文献标志码:** A

文章编号: 1006-7167(2021)11-0032-06

Automatic Detection of Passion Fruit Based on Improved Faster R-CNN

TU Shuqin^a, HUANG Jian^a, LIN Yueting^a, LI Jialin^a, LIU Haofeng^a, CHEN Zhimin^b
(a. College of Mathematics and Informatics; b. Experimental Basis and Practical Training Center, South China Agricultural University, Guangzhou 510642, China)

Abstract: Passion fruit is dense in orchards and the growing environment is relatively complex, which makes it difficult to identify, pick and estimate the yield for passion fruits. In this paper, the colour image of the passion fruit in the natural growth environment is taken as the research object, and the fruit is divided into four categories: unobstructed, occluded, overlapping fruits and background, respectively, and the improved faster R-CNN target detection algorithm is proposed for the automatic detection and yield prediction of passion fruit under the deep learning framework. Firstly, residual network fusion FPN is used to extract the multi-scale features of passion fruit, then the ROI region is extracted by RPN network, and finally achieve the classification and detection of passion fruit through full connection layer. In the test set, its best mean average precision reached 87.98%, the average precision and average recall rate reached 90.79% and 90.47%, respectively. The detection time of each picture is about 0.178 s. This algorithm achieves the 96.80% of precision rate in yield estimation. Therefore, the faster R-CNN target detection algorithm based on FPN + ResNet-101 feature extraction can be applied to the fast and accurate detection of passion fruit and yield estimation.

Key words: passion fruit detection; Faster R-CNN; ResNet-50/101; FPN

收稿日期: 2021-03-19

基金项目: 国家自然科学基金项目(61772209); 华南农业大学精品实践课程项目(zlge18010); 华南农业大学质量工程项目(zlge19037); 广东省大学生创新创业训练计划项目(S202010564047)

作者简介: 涂淑琴(1978-), 女, 江西南昌人, 讲师, 研究方向为计

算机视觉、农业信息化技术。

Tel.: 13751838469; E-mail: tsq5_6@scau.edu.cn

通信作者: 陈志民(1981-), 男, 广东湛江人, 高级实验师, 研究方向为智能计算、实验教学与管理。

Tel.: 13570565308; E-mail: zhiminchen@scau.edu.cn

0 引言

百香果,学名“西番莲”,是西番莲科西番莲属的草质藤木。童彤^[1]指出百香果对健康益处多。百香果生长习性良好,环境适应能力较强,不仅可以作为果树大规模种植于果园,还可以作为观赏植物栽种于居民庭院。目前在东南亚和美洲分布较为广泛,在中国主要分布在广东、福建和广西等南方省份。随着百香果产量不断上升,其在创造出巨大财富的同时,也带来了人工采摘的困难。皆彦吉^[2]提到,当前进行果蔬采摘是农业生产中最耗费人力与物力的事情,采摘的成本高为经营者带来重大负担。百香果产量的增加、人工劳动成本的上升和劳动力不足的现状,引发了研究人员对智能采摘机器人研究开发。而目标检测作为果实识别的关键技术,是采摘机器人目标识别的基础。以百香果为目标,研究基于不同特征提取网络以及不同参数环境下的目标检测算法对智能采摘机器人研制,具有重要的意义。

目前我国已经自行研发了苹果采摘机器人,并且吸引了大量研究人员对果蔬机器人进行软硬件技术方面的研究。李昱华^[3]对番茄采摘机器人目标检测与抓取的关键技术进行了研究,重点介绍了果实目标识别的过程,包括图像预处理、模型选择、目标分割算法等。樊艳英等^[4]指出,目标果实的识别和定位是采摘作业的的必要条件和准确采摘的关键,在基于双视觉传感器的基础上,他们设计并研究了一个采摘机器人目标果实识别系统,通过图像处理技术,实现了果实的识别和定位功能,为实现水果采摘的自动化和无人化奠定了软件基础。

基于深度学习的目标检测技术作为采摘机器人软件层面上的重要一环,其算法的优劣对果实的识别和定位起着关键作用。目前国内外对基于深度学习的果实检测展开了一些研究工作^[5-40]。Sa 等^[5]提出将 Faster R-CNN 目标检测算法应用于水果检测中,目标是建立一个准确、快速、可靠的水果检测系统。Stein 等^[6]提出了基于 Faster R-CNN^[7] 的新的多传感器框架,可以有效地对芒果果园的每个果实进行识别、跟踪、定位,利用多视点方法解决了果实遮挡问题,从而解决了耗费大量劳动力在现场进行实际产量估计的问题。陈燕等^[8]利用 YOLOv3-DenseNet34 网络实现对荔枝的目标检测,将传统的目标检测与双目视觉相结合,实现对荔枝串进行定位,引导机器人对荔枝串进行采摘。彭明霞等^[9]提出利用融合特征金字塔网络 (Feature Pyramid Network, FPN) 的 Faster R-CNN 对棉花和杂草的快速、准确识别。熊俊涛等^[10]提出改进 YOLO v3 网络的夜间环境柑橘识别方法,其在测试集下的平均精度 (mean Average Precision, mAP) 为

90.75%。Koirala 等^[11]利用 YOLOv3 和 YOLOv2 对不同场景下的芒果集进行检测,取得了 F1 值为 0.968,平均准确率为 0.983。Tu 等^[12]使用 RGB-D 图像识别技术检测处于不同生长阶段的百香果并对其进行成熟鉴定进行了研究。Bargoti S 等^[13]利用 Faster R-CNN 检测框架实现果园中水果检测应用,提出在产量预测和机器收割等高层次的农业任务中,准确可靠的基于图像的水果检测系统支持是至关重要。

本文提出了基于残差网络 (Residual Network, ResNet)^[14] 和 FPN^[15] 的 Faster R-CNN 百香果检测算法,实现果实无遮挡、遮挡、重叠和背景四类果实自动检测和产量预测。该方法首先采用 ResNet 网络融合 FPN 对百香果进行多尺度特征提取;然后接入到感兴趣区域 (Region of Interest, ROI) 池化层进行尺度归一化;最后通过全连接层实现百香果分类和检测。

1 数据集

本实验研究对象是处于自然生长环境下的百香果,采用华为 Mate10 手机拍摄了约 2 000 张花都区百香果庄园内的彩色百香果照片。经过筛选,挑选了其中较具有代表性的 700 余张图片作为实验数据集,每张图片中包含多个百香果果实,部分图片如图 1 所示。

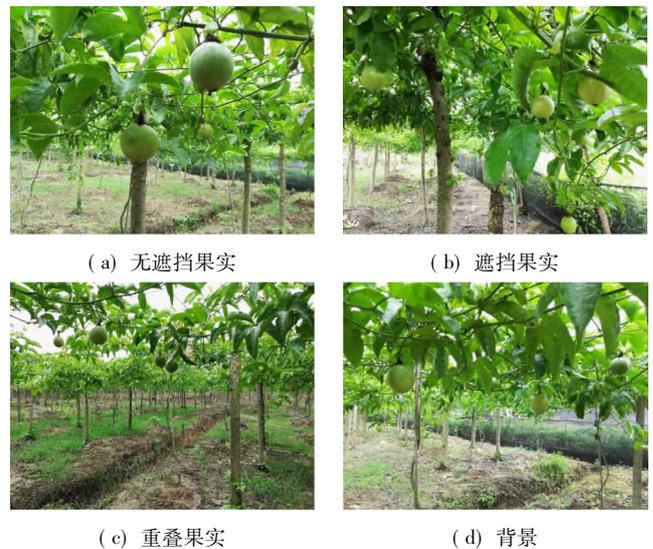


图 1 百香果图片样本

实验中利用 labelImg 对百香果进行标定,同时将百香果果实分为无遮挡果实、遮挡果实、重叠果实和背景。为方便结果的观测,分别用 n_ov、occ、ov 代表无遮挡、遮挡、重叠三类果实标签。通过标注工具标注的约 700 张照片中百香果总数约为 4 000 个,其中无遮挡果实数约为 2 000 个,遮挡果实约为 1 600 个,重叠果实约为 400 个(见图 2)。

2 基于改进的 Faster R-CNN 百香果检测方法

基于 ResNet + FPN 的 Faster R-CNN 百香果检测



图2 不同背景下的百香果

网络检测流程如图3所示,该方法流程包括以下步骤:

(1) 首先,输入任意尺寸大小的百香果 RGB 图像,利用 FasterR-CNN 中的残差网络和 FPN 对输入百香果图像进行多尺度特征提取,形成多尺度特征映射图。

(2) 一方面多尺度特征映射图输入到该网络中的区域候选网络(Region Proposal Network, RPN);另一方面多尺度特征映射图继续前向传播至特有卷积层,产生更高维特征图。

(3) 多尺度特征映射图经过 RPN 网络得到建议

区域,并根据区域得分,区分前景和背景(果实和非果实),对前景区域采用非极大值抑制操作,输出其前 N 个得分最高的候选前景区域作为感兴趣区域(ROI),完成目标框的粗检测。

(4) 多尺度特征映射图及 ROI 输入到 RoiPooling 层。

(5) RoiPooling 层输出至全连接层,对每一个感兴趣区域线性边框回归,得到目标果实的精准检测;用 softmax 回归模型做多分类目标检测,最终识别出不同百香果的类别(果实无遮挡、遮挡、重叠)。

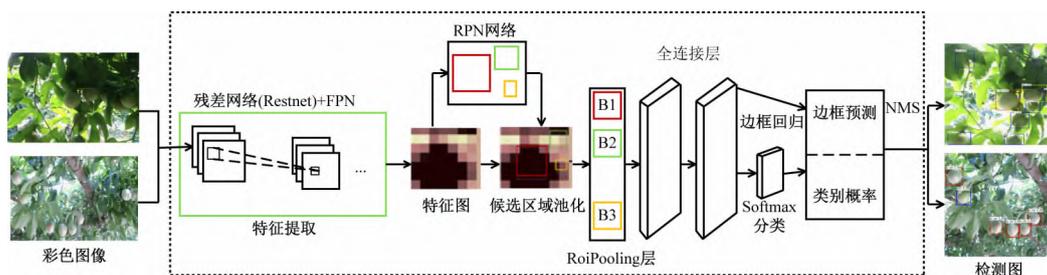


图3 基于 ResNet + FPN 的 Faster R-CNN 百香果检测流程

选择合适的迭代次数,就能够利用 Faster R-CNN 算法进行模型训练。调整训练迭代次数,反复训练模型,建立最佳的基于 ResNet + FPN 的 Faster R-CNN 百香果检测流程。

为实现小对象更好检测,本文在 ResNet + FPN 的 Faster R-CNN 百香果检测网络结构的基础上,对 ResNet + FPN 的 Faster R-CNN 网络框架网络深度和宽度进行优化调整,在已有一定训练基础的参数上,进行迁移学习,用含有较多小果实图像作为训练集,比较不

同层数和不同卷积核的网络模型的检测准确率,确定较优的 ResNet + FPN 的 Faster R-CNN 的网络参数和模型。最终,其网络较基于 ResNet + FPN 的 Faster R-CNN 网络做了两点优化改进: ①针对本文试验场景中果实图像对象,将 Faster R-CNN 中的非极大值抑制(Non Maximum Suppression, NMS)修改为 SoftNMS^[16],这样实现了更多小对象的检测,而且没有增加运行时间。②Faster R-CNN 中分类类别由 80 类改为 4,代表类别数为 4 类(果实无遮挡、遮挡、重叠和背景)。

3 实验结果和讨论

3.1 建立检测模型

具体的训练检测模型搭建过程如下(见图 4):

(1) 为了方便检测效果的对比,实验采用统一的数据集,但是数据集格式上有所改变。在基于 ResNet 的 Faster R-CNN 实验中,使用了 Facebook AI 研究院于 2018 年初公开目标检测框架 detectron,该框架使用的是 COCO 数据集,标注文件类型从 .xml 变为 .json。在实验中,通过简单的代码实现数据集格式的转换。

(2) 分别利用 ResNet50、ResNet101、ResNet50 + FPN、ResNet101 + FPN 四种特征提取器进行模型训练,进一步探究学习率等参数对检测效果的影响,以找出最优模型。ResNet50 的训练时间比较短,迭代次数为 6 000 的情况下训练时间大约为 1 h。

(3) 利用相对最优的模型,进行百香果检测,并对实验结果进行分析和统计。

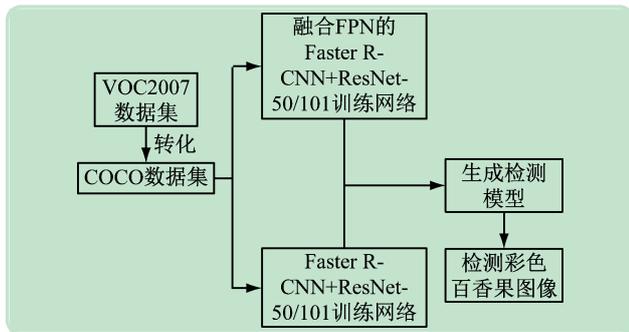


图 4 基于 ResNet-50/101 + FPN 的 Faster R-CNN 百香果检测模型训练过程

3.2 实验结果及分析

由表 1 可以看出,在学习率和迭代次数恒定的前提下,不同特征提取网络的检测算法对检测效果有一

定的影响。ResNet101 比 ResNet50 的检测效果略好,从 mAP 上看,提升了 0.4% 左右,但是由于网络深度的加深,在检测时间上相应要稍逊于 ResNet50,两种网络结构每幅彩色图片的检测时间相差 35 ms 左右。基于 ResNet + FPN 的网络结构有更好的检测结果,从 mAP 上看,比使用 ResNet 提升了 1.3% 左右,这与 FPN 更利于小目标的检测有关;从时间上看,融合了 FPN 的网络明显加快了目标检测的速度,在平均检测时间上缩短了 3/4,可见 FPN 带来了更好检测效果的同时缩短了检测时间,提高了百香果目标检测的效率。从整体上看,使用 ResNet 特征提取网络的 mAP 稳定在区间 86% ~ 87%,无遮挡果实的检测精度保持在 90% 以上。另外,在 ResNet 融合 FPN 后,对比 ResNet 网络结构,其检测平均精确率(Average Precision, AP)和平均耗时上都有显著提升。

表 1 不同网络结构的检测效果

网络结构	AP/%			mAP/%	平均耗时/s
	无遮挡果实	遮挡果实	果实重叠		
ResNet50	90.46	87.81	80.57	86.28	0.525
ResNet50 + FPN	90.00	88.20	84.80	87.70	0.139
ResNet101	90.33	86.70	82.60	86.68	0.560
ResNet101 + FPN	90.30	87.45	86.14	87.98	0.177

图 5(a)、(b)、(c) 分别用 4 种不同颜色的曲线分别描绘了 ResNet50、ResNet101、ResNet50 + FPN 和 ResNet101 + FPN 4 种特征提取网络下的精确率-召回率(Precision-Recall, P-R) 曲线, P 表示精确率; R 表示召回率。图 5(d) 针对 ResNet101 + FPN 网络结构分别对三类果实进行了 P-R 曲线的描绘。在 P-R 图上, P-R 曲线的积分,即 P-R 曲线下的面积越大,代表的检测结果越好。

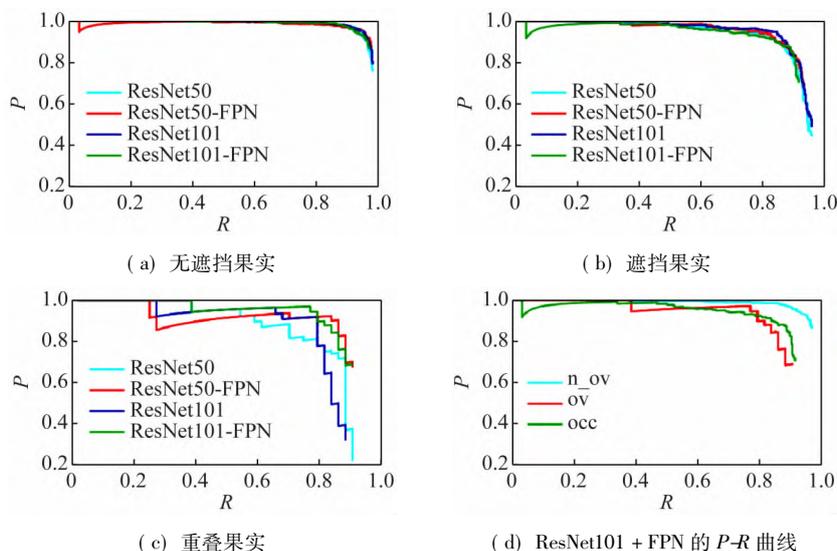


图 5 百香果果实的 P-R 曲线

从图 5 (a) 和 (b) 可以看出,不同特征提取网络下,无遮挡和遮挡果实的检测效果从召回率 R 、准确率 P 和 mAP 上看都非常接近, $P-R$ 曲线近乎贴和,可见特征提取网络对这两类果实的检测效果没有明显影响。而重叠果实受特征提取网络的影响较大,由图 5 (c) 可以看出,重叠果实在 ResNet101 + FPN 下取得最好的检测结果,在 ResNet50 上准确率和召回率较低,检测结果相比其他三种特征提取网络较差。从图 5 (d) 分析,3 类果实中无遮挡果实的检测结果具有明显的优势,其次是遮挡果实,最后是重叠果实。

基于 ResNet101 + FPN 的 Faster R-CNN 目标检测算法的性能最好,利用其对 100 张图片不同类别的百香果的个数、算法检测出的正确、漏检和误检的百香果个数进行统计,其结果如表 2,3 所示。

表 2 基于 ResNet101 + FPN 的 Faster R-CNN 百香果产量预测

百香果	实际个数	正确个数	漏检个数	误检个数
无遮挡	222	214	8	15
遮挡	187	153	17	18
重叠	20	17	3	2

表 3 基于 ResNet101 + FPN 的 Faster R-CNN 检测结果 %

百香果	准确率	召回率	误分率	漏分率
无遮挡	93.40	96.40	6.60	7.70
遮挡	89.50	90.00	10.50	10.00
重叠	89.50	85.00	10.50	15.00

结合表 2 和表 3 分析,在无遮挡果实中,算法共检测出 229 个目标,其中 214 个检测正确,8 个漏检,15 个误检。其准确率和召回率分别达到了 93.4% 和 96.40%,比遮挡和重叠两类都要高,而误分率和漏分率比其他两类都低,这可能与训练样本中无遮挡果实在数量上和质量上的优势相关;在遮挡果实中,算法共检测出 171 个目标,其中正确的 153 个,17 个漏检,18 个误检。该类别的准确率和召回率接近 90% 左右,但是漏检和误检基数大。遮挡果实的遮挡情况多样,遮挡物的不同、遮挡程度的不同加大了特征学习的难度和检测的难度,在一定程度上影响了检测的准确率,增加了漏检和误检的可能性;在重叠果实中,算法共检测出 19 个目标,其中正确的是 17 个,3 个漏检,2 个误检。重叠果实的训练样本数较少,缺乏大量数据的支持使得模型在学习特征时不够充分。另外,在重叠果实中,果实重叠面积的差异对训练和检测产生了一定影响,故准确率和召回率较低,而误分率和漏分率较高。

ResNet101 + FPN 特征提取网络的部分检测结果如图 6 所示。用红框、蓝框、黄框分别表示算法检测的

无遮挡、遮挡、重叠三类果实。用白色箭头指向误检的目标。



(a) 基于 ResNet101 检测效果 (b) 基于 ResNet101 + FPN 检测效果
图 6 两种方法下的场景的百香果检测结果对比

从图 6 可以发现,对比 ResNet101 网络,ResNet-101 + FPN 的 Faster R-CNN 模型在无遮挡、遮挡中不存在漏检和误检,在重叠两类果实上有误检现象,都能较好地检测出无遮挡、遮挡、重叠 3 类果实,其最好 mAP 达到 88.01%,能够对百香果目标进行快速准确的检测。

在产量估计中,采集了 275 张密集小果实图像(共有 7032 个果实)做训练测试数据集。ResNet-101 + FPN 的 Faster R-CNN 检测结果。该方法能检测正确果实个数为 6205,漏检个数为 827,错检个数为 207,精确率为 96.80%,召回率为 88.20%,F1 值为 0.923,能够实现密集果实产量预测要求。

基于 ResNet-101 + FPN 的 Faster R-CNN 的部分密集果实检测结果如图 7 所示,将所有类别都表示为红色框,在图片中,大果实对象都能准确检测出来,而对



图 7 百香果产量估算

部分小对象(像素在 [10,10] 到 [20,20]) 的也能有效检测出来。但是对于像素在 [10,10] 以下的,检测器不能很好工作,出现较多漏检。结果表明,该方法可以实现进行自然场景下果实的产量预估。

4 结 语

本文提出的基于 ResNet101 + FPN 的 Faster R-CNN 百香果检测算法在 mAP 和检测时间上都具有最佳结果,其 mAP 达到了 88.01%,在无遮挡、遮挡、重叠 3 类果实中检测准确率达到 93.4%、89.5%、89.5%,召回率分别为 96.4%、90.0%、89.00%,检测时间低至 0.178 s,检测性能良好,达到了快速准确检测百香果的目标。在产量估算中,该方法取得了准确率为 96.80%,召回率为 88.20%, F_1 值为 0.923。结果表明,该检测算法实现了在自然场景下百香果的快速准确检测,准确地预测百香果产量,可以更好地实时应用于自然场景下采摘机器人的定位检测与产量估算。

参考文献(References):

- [1] 童 彤. 百香果对健康益处多[J]. 中国果业信息, 2018, 35(6): 53-57.
- [2] 皆彦吉. 果蔬采摘机器人的研究现状、问题及对策[J]. 时代农机, 2018, 45(9): 42-45.
- [3] 李昱华. 番茄采摘机器人目标检测与抓取的关键技术研究[J]. 中国新技术新产品, 2018(23): 55-57.
- [4] 樊艳英, 张自敏, 陈冠萍, 等. 视觉传感器在采摘机器人目标果实识别系统中的应用[J]. 农机化研究, 2019, 41(5): 210-214.
- [5] Sa I, Ge Z, Dayoub F, et al. DeepFruits: A fruit detection system using deep neural networks[J]. Sensors-Basel, 2016, 16(8): 1222.

- [6] Stein M, Bargoti S, Underwood J. Image based mango fruit detection, localisation and yield estimation using multiple view geometry[J]. Sensors-Basel, 2016, 16(11): 1915.
- [7] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [8] 陈 燕, 王佳盛, 曾泽钦, 等. 大视场下荔枝采摘机器人的视觉预定位方法[J]. 农业工程学报, 2019, 35(23): 48-54.
- [9] 彭明霞, 夏俊芳, 彭 辉. 融合 FPN 的 Faster R-CNN 复杂背景下棉田杂草高效识别方法[J]. 农业工程学报, 2019, 35(20): 202-209.
- [10] 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.
- [11] Koirala A, Walsh K B, Wang Z, et al. Deep learning-Method overview and review of use for fruit detection and yield estimation[J]. Computers and Electronics in Agriculture, 2019, 162: 219-234.
- [12] Tu S, Xue Y, Zheng C, et al. Detection of passion fruits and maturity classification using Red-Green-Blue Depth images[J]. Biosystems Engineering, 2018, 175: 156-167.
- [13] Bargoti S, Underwood J. Deep fruit detection in orchards[C] // International Conference on Robotics and Automation (ICRA). Singapore: IEEE, 2017: 3626-3633.
- [14] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C] // Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas: IEEE, 2016: 770-778.
- [15] Lin TY, DollárP, GirshickR, et al. Feature pyramid networks for object detection[C] // Conference on Computer Vision and Pattern Recognition(CVPR). Honolulu: IEEE, 2017: 936-944.
- [16] BodlaN, SinghB, ChellappaR, et al. Soft-NMS-Improving object detection with one line of code[C] // International Conference on Computer Vision(ICCV). Venice: IEEE, 2017: 5562-5570.

(上接第 31 页)

- [4] Martin C D. Brittle failure of rock materials: test results and constitutive models[J]. Canadian Geotechnical Journal, 1996, 33(2): 378-385.
- [5] 陈屹然, 秦四清, 薛 雷, 等. 岩石脆性破坏表征与 Weibull 分布适用范围[J]. 地球物理学进展, 2017, 32(5): 2200-2206.
- [6] 陈国庆, 赵 聪, 魏 涛, 等. 基于全应力-应变曲线及起裂应力的岩石脆性特征评价方法[J]. 岩石力学与工程学报, 2018, 37(1): 51-59.
- [7] Zhang C C, Dong D Z, Wang Y M, et al. Brittleness evaluation of the upper ordovician wufeng-lower silurian longmaxi shale in southern sichuan basin, China[J]. Energy Exploration and Exploitation, 2017, 35(4): 430-443.
- [8] 谢和平, 鞠 杨, 黎立云. 基于能量耗散与释放原理的岩石强度与整体破坏准则[J]. 岩石力学与工程学报, 2005, 24(17): 3003-3010.
- [9] Baron L I. Determination of properties of rocks[M]. Moscow: Gogotekhizdat, 1962.
- [10] Kidybinski A. Bursting liability indices of coal[J]. International Journal of Rock Mechanics and mining Sciences and geomechanics

Abstracts, 1981, 18(4): 295-304.

- [11] Tarasov B, Potvin Y. Universal criteria for rock brittleness estimation under triaxial compression[J]. International Journal of Rock Mechanics and Mining Sciences, 2013, 59(4): 57-69.
- [12] 张 军, 艾 池, 李玉伟, 等. 基于岩石破坏全过程能量演化的脆性评价指标[J]. 岩石力学与工程学报, 2017, 36(6): 1326-1340.
- [13] 尤明庆, 华安增. 岩石试样破坏过程的能量分析[J]. 岩石力学与工程学报, 2002, 21(6): 778-781.
- [14] Peng R, Yang J, Wang J G, et al. Energy dissipation and release during coal failure under conventional triaxial compression[J]. Rock Mechanics and Rock Engineering, 2015, 48(2): 509-526.
- [15] GB/T23561.9 煤和岩石物理力学性质测定方法第 9 部分: 煤和岩石三轴强度测定及变形参数测定方法[S]. 北京: 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会, 2009.
- [16] Zhang ZP, Xie H P, Zhang R, et al. Deformation damage and energy evolution characteristics of coal at different depths[J]. Rock Mechanics and Rock Engineering, 2018, 52(5): 1-13.