

引用格式:

方文博, 郭永刚, 关法春, 张伟, 刘倩倩, 王树文, 张正超, 于皓然. 基于改进 YOLO v5s 算法的大豆叶片虫洞的识别[J]. 湖南农业大学学报(自然科学版), 2023, 49(1): 127–132.

FANG W B, GUO Y G, GUAN F C, ZHANG W, LIU Q Q, WANG S W, ZHANG Z C, YU H R. Soybean leaf wormhole recognition based on improved YOLO v5s algorithm[J]. Journal of Hunan Agricultural University(Natural Sciences), 2023, 49(1): 127–132.

投稿网址: <http://xb.hunau.edu.cn>



基于改进 YOLO v5s 算法的大豆叶片虫洞的识别

方文博¹, 郭永刚¹, 关法春^{2*}, 张伟², 刘倩倩¹, 王树文³, 张正超¹, 于皓然¹

(1.西藏农牧学院水利土木工程学院, 西藏 林芝 860000; 2.吉林省农业科学院农村能源研究所, 吉林 长春 130119; 3.岭南师范学院电子与电气工程学院, 广东 湛江 524048)

摘要: 鉴于对大豆叶片虫洞进行识别有助于及时发现虫情并有针对性的防治虫害, 提出了一种大豆叶片虫洞的识别方法: 以 YOLO v5s 网络作为基础, 在大豆叶片虫洞特征提取过程中引入空洞卷积代替 3 次池化处理, 提取虫洞边缘不规则信息; 将特征信息输入空间注意力机制, 提取时空融合信息, 进而捕获野外不同背景下的颜色信息; 针对大豆叶片虫洞目标远近不一的问题, 重构特征金字塔结构, 增加了 1 层输出层, 将 80 像素×80 像素输出特征图经过上采样后得到 160 像素×160 像素特征图, 并将其与浅层同尺寸特征图进行拼接, 提高虫洞目标识别定位的准确性; 将融合后的总特征输入目标检测模块, 输出单个对象的检测外框, 得到大豆叶片虫洞识别模型。在大豆叶片虫洞样本数据集上对模型进行测试, 结果对大豆叶片虫洞的平均识别准确率最高达 95.24%, 模型存储空间为 15.1 MB, 每秒传输 91 帧。所建立的方法与 Faster R-CNN、YOLO v3、YOLO v5s 对比, 对大豆叶片虫洞识别的平均准确率分别提高 2.50%、12.13%、2.81%。

关键词: 大豆叶片虫洞; 注意力机制; 改进 YOLO v5s; 机器学习; 识别准确率

中图分类号: TP391.4

文献标志码: A

文章编号: 1007-1032(2023)01-0127-06

Soybean leaf wormhole recognition based on improved YOLO v5s algorithm

FANG Wenbo¹, GUO Yonggang¹, GUAN Fachun^{2*}, ZHANG Wei², LIU Qianqian¹,
WANG Shuwen³, ZHANG Zhengchao¹, YU Haoran¹

(1.School of Water Resources and Civil Engineering, Tibet College of Agriculture and Animal Husbandry, Linzhi, Tibet 860000, China; 2.Rural Energy Research Institute, Jilin Academy of Agricultural Sciences, Changchun, Jilin 130119, China; 3.School of Electrical and Electronic Engineering, Lingnan Normal University, Zhanjiang, Guangdong 524048, China)

Abstract: Soybean(*Glycine max*) leaf wormholes seriously affect the quality of crops. However, due to the complex background environment, dense planting and diversified leaf wormhole shapes, traditional manual and machine learning recognition are difficult to meet the requirements in terms of accuracy and speed. In response to this problem, this paper proposes an improved soybean pest identification method. This method is based on the YOLOv5s (You Only Look Once) network, introduces an attention mechanism to improve the recognition ability of wormhole parts, uses the sample transformation method to adapt to the diversity of multi-leaf morphology, and improves the redundant bounding box The elimination mechanism reduces misjudgments and missed judgments. In the experiment, this paper constructed a soybean sample data set as the test data, and compared this method with the traditional deep target recognition method. The average accuracy rate on the test data set is up to 95.24%, and the model storage space is 15.1 MB, the number of frames

收稿日期: 2021-10-25

修回日期: 2022-09-15

基金项目: 国家重点研发计划项目(2018YFD1000905); 广东省教育厅创新人才重点项目(2018KTSCX129); 哈尔滨市创新创业人才项目(HCX06)

作者简介: 方文博(1996—), 男, 浙江温州人, 硕士研究生, 主要从事农业大数据研究, 690704822@qq.com; *通信作者, 关法春, 博士, 副研究员, 主要从事农牧一体化研究, gfc1940@163.com

transmitted per second is 91 f/s. The average accuracy rate is 2.50%, 12.13%, 2.81% higher than Faster R-CNN, YOLO v3, and YOLO v5s respectively. The method proposed in this paper has greatly improved the recognition accuracy and recognition speed, and only requires a small model deployment. The above features make this method more suitable for the practical application of soybean wormhole recognition.

Keywords: wormhole; attention mechanism; YOLO v5s; machine learning; recognition accuracy

大豆叶片出现的虫洞会造成植株长势衰落,虫害严重时会导致干物质无法有效累积输送到豆荚上,使得大豆果实无法鼓粒,从而影响大豆的产量和品质。如能在虫洞出现初期及时进行防治,可以有效提高大豆结荚率^[1-2]。大豆叶片虫洞通常具有不规则形状,与健康叶片相比有明显的视觉特征差异,利用这种差异可实现深度学习对虫洞识别。

计算机视觉技术自 2012 年引入卷积神经网络之后,基于 Region Proposal 的 R-CNN 系算法具有在满足检测精度的同时,大幅度提升检测速度的优势而被广泛应用^[3-5];随后 YOLO v2、YOLO v3、YOLO v4、YOLO v5 逐渐成为计算机视觉目标检测的主流方法^[6-7]。鲍文霞等^[8]引入椭圆型度量,对水稻虫害图像特征提取准确率达 90.5%;赵志焱等^[9]通过改进 Alexnet 模型,在全连接层中减少了神经元数目,在提高运算速度的同时,玉露香梨叶虫害识别的准确率达 81.18%;李静等^[10]通过改进 GoogLeNet 模型,利用迁移学习算法实现了玉米虫害识别,准确率达 96.4%。大豆密集种植使得对虫洞叶片的机器识别面临困难:一方面,遭害虫侵蚀引起虫洞的叶片与健康叶片混杂在一起,且虫洞形状不规则;另一方面,叶片方向的多样性加大了训练数据收集的难度,导致传统识别技术的识别效率及

准确性无法满足基于机器视觉的智能化生产需要。

针对以上问题,笔者基于 YOLO v5s 神经网络模型,在空间池化结构(spatial pyramid pooling, SPP)中添加了卷积注意力模块(CA),使虫洞特征更具分辨性。重构了特征金字塔结构,将 3 个特征尺度改为 4 个特征尺度,提高扩张卷积网络的特征提取能力。引入了特征融合模块,以减少因分辨率降低而导致的信息缺失,融合后的特征尺度包含了之前各个特征尺度的信息,可改善大豆虫洞小目标的识别效果。

1 大豆叶片虫洞数据集的构建

2021 年 8 月,在吉林省农业科学院试验基地采集大豆叶片虫洞图像。选用荣耀 X10 手机,通过录像方式采集图像:拍摄 20 段视频,平均每段 30 s 左右,保存格式为 1920 帧×1020 帧,帧速率为 60 帧/s。通过自编译 Python 脚本,将视频转换成图片,每 20 帧生成 1 张,得到 2120 张图片,删除重复图片后,最终得到 1170 张图片。大豆虫洞图片如图 1 所示。利用 LabelImg 对每幅图像中的虫洞进行人工标注,标签名为 Wormhole,标注时选用产生虫洞部分的最小矩形框,在数据采集时选取不同角度的虫洞图片,以增强机器训练的学习效果。



图 1 大豆叶片虫洞图像

Fig.1 Soybean leaf wormhole atlas

为使模型泛化能力更强,通过随机裁剪、亮度调整、噪声扰动、翻转变换等手段,将训练图像数据集扩充到 4484 幅,其中训练集和数据集比例为

8 : 2。亮度调整采用了伽马变换,伽马系数为 0.7,亮度均值为 0.600,方差为 0.197,使得调整后的图片更加明亮;噪声采用弹性变形处理,根据扭曲场

的平滑度与强度逐一地移动局部像素点实现模糊效果,其中高斯函数中的标准差选取为 34。最终图片大小为 640 像素×640 像素^[11-12]。

2 大豆叶片虫洞的识别

所采用的操作系统为 WIN1064 位,CPU 为 Intel core i7-9850H,GPU 为 RTX 2080,本地内存为 32G,使用语言为 Python,框架采用 Pytorch。初始学习率为 0.01,最终学习率为 0.002,最大训练轮数为 200,IoU 阈值为 0.2。

采用平均精度(mAP)、每秒传输帧数(FPS)、模型存储空间、准确率、召回率、运算时间^[13-14]作为评价指标。

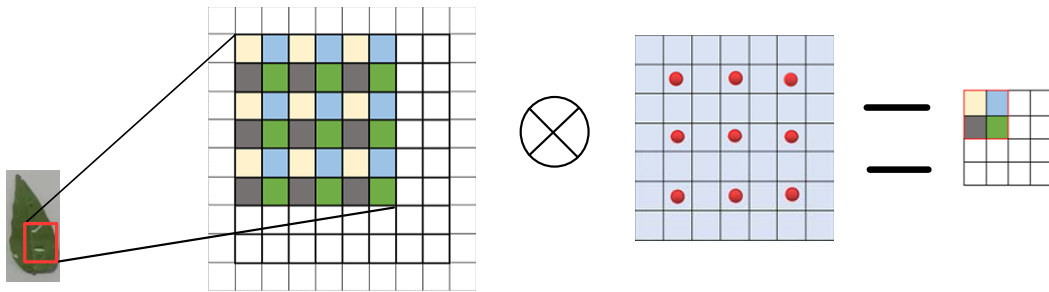


图 2 空洞卷积 3×3 卷积核特征的提取过程

Fig.2 Null convolution 3×3 convolution kernel feature extraction process

在叶片虫洞检测过程中,仅仅依靠目标浅层特征进行判断的准确率较低,而图像上下层信息能有效描述目标的内在联系,应用到视觉检测中能有效减小浅层特征带来的图像细粒度低的影响。空洞卷积引入了膨胀比,重新定义了卷积核之间的空间信息^[17]。在不使用池化操作降低图像分辨率的情况下,扩大作物病害部位敏感信息的覆盖,提高了网络性能。

2.2 基于注意力机制的 YOLO v5s 大豆叶片虫洞的识别

2.2.1 YOLO v5s 网络结构的改进

为解决大豆叶片虫洞发生初期识别准确率较低的问题,对原 YOLO v5s 模型进行了改进。该模型主要分为输入端、Backbone、Neck、Head 等 4 个部分,分别对应图片处理、主干网络、加强特征提取、利用特征进行预测。对原 YOLO v5s 模型的改进分为 4 个步骤。第 1 步,在输入端进行初始锚框优化;第 2 步,在 Backbone 部分引入注意力机制;第 3 步,在 Neck 部分新增 1 个特征尺度;第 4

2.1 特征提取

运用 YOLO v5s 中的空间池化结构提取不同的感受野,将池化后得到的特征进行张量拼接和归一化操作,通过不同次池化得到不同层次的特征^[15-16]。引入空洞卷积代替 3 次池化处理,每个分支中膨胀比逐级增加,分别为 1、3、5,且卷积核为 3×3 的空洞卷积,最终得到 3×3、7×7、13×13 的空洞卷积感受野。添加空洞卷积块后,为了得到小目标更丰富的空间信息,将主干网络的最后一个残差块与空洞卷积的输出进行跳跃连接融合特征,然后经过 1×1 的卷积层下降通道数。空洞卷积 3×3 卷积核特征提取过程如图 2 所示。

步,在 Head 部分使用 DIoU_nms 替换 GIoU_nms。

2.2.2 K-均值聚类算法

由于使用采集的数据集作训练集,需预先修改 Anchors 初始值来匹配数据集,Anchors 大小则通过聚类算法得到。选用 K-means⁺⁺算法^[18],从数据集中随机选取 1 个样本作为第 1 个初始聚类中心,计算每个样本与当前选定的聚类中心之间的最短距离 $D(x)$,再计算每个样本点选为下一个聚类中心的概率 $P(x)$,最后选取最大概率值所对应的样本点作为下一个簇中心。不断重复,直到选取出 K 个聚类中心。

2.2.3 注意力机制模块

大豆的生长环境较为复杂,叶片虫洞必须从多角度观测才能得出虫害的正确结论。为提高识别准确度,引入注意力机制(CA),以适应大豆不同角度的虫洞特征,如图 3 所示。从图 3 中可以看出,这种机制将原有的通道注意力分解为方向和位置 2 个并行的特征编码,将空间坐标信息聚合成新的注意力图,加强对空间的精准定位,用以提高模型的特征提取能力。

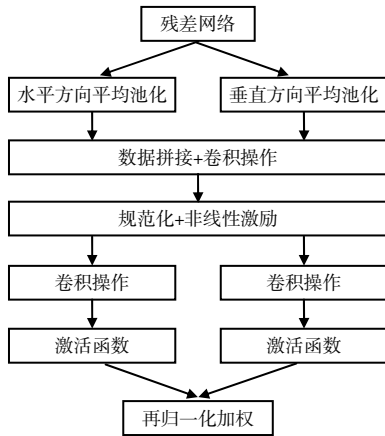


图 3 注意力结构

Fig.3 CA attention structure diagram

2.2.4 多特征尺度的模型结构

单特征尺度模型对原图片进行特征提取后,在该特征图片上直接进行特征检测,这会导致对图片只进行表明的特征提取而缺失高层次的部分。YOLO v5s 通过多特征尺度模型,虽然加深了高层次的特征提取,但依然不适用于大豆虫洞小目标的检测^[19-20],因而对其进行改进(图 4):增加了特征尺度,加强了对虫洞不规则图像边缘的识别,再对×4 下采样图片进行特征融合,减少因分辨率降低而导致的信息缺失,融合后的特征尺度 4 包含了之前各个特征尺度的信息,可改善大豆虫洞小目标的识别效果。

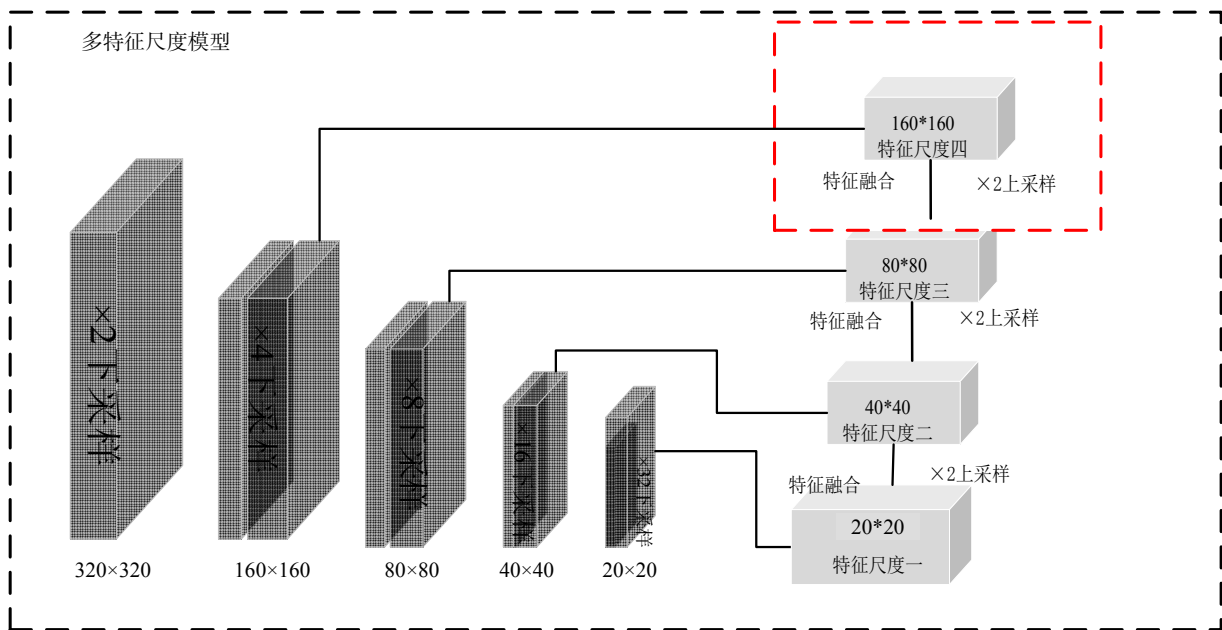


图 4 多特征尺度模型

Fig.4 Multi-feature scale model

2.2.5 非极大值抑制

YOLO v5s 选用了传统的 GIoU_NMS^[21-22],找到最佳的目标边界框,消除冗余的边界框。但在实际应用过程中会出现不同物体距离比较近时,由于交并面积较大,经过处理后只剩下 1 个检测框,进而导致物体检测漏失的情况。针对虫洞这种检测目标相距普遍较近的情况,采用了 DIoU_NMS^[23]来代替传统的 GIoU_NMS。DIoU_NMS 在继承了 IoU 优点的基础上,还考虑了锚定框和预测框中心点距离的问题,当两框中心点相距离远时,判定是 2 个

物体而不会被过滤。

3 结果与分析

3.1 改进 YOLO v5s 模型的训练结果

改进 YOLO v5s 模型训练过程如图 5 所示。基于改进 YOLO v5s 算法检测模型,对 4484 幅图像识别的准确率为 93.85%,召回率为 97.32%,平均精度为 95.24%,检测速度为每秒 91 帧。从识别结果来看,改进后的算法检测模型具有速度快,精度高的特点。

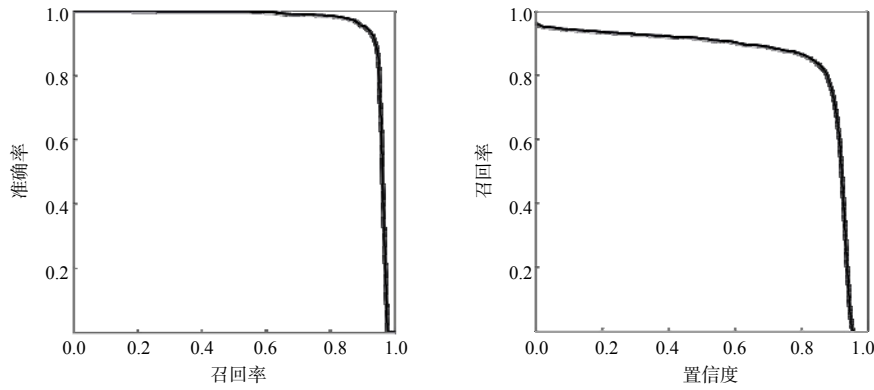


图 5 改进 YOLO v5s 模型训练过程

Fig.5 Improve YOLO v5s model training process

3.2 消融试验

为验证注意力机制和特征尺度 2 个改进方案对 YOLO v5s 算法性能的影响,进行了消融试验^[24-25]。方案 1,对原 YOLO v5s 算法模型不进行任何改动;方案 2,在原 YOLO v5s 算法模型上添加注意力机制;方案 3,在原 YOLO v5s 算法模型上添加特征尺度;方案 4,在原 YOLO v5s 算法模型上同时添加注意力机制和特征尺度。结果(表 1)注意力机制和特征尺度 2 个改进方案平均识别准确率分别提升了 1.72%和 0.83%;同时施加 2 个改进方案平均识别准确率提升 2.81%,这表明 2 个改进方案对算法模型性能的提升效果良好。

表 1 消融试验的召回率和平均准确率

方案	召回率/%	平均准确率/%
1	94.01	92.43
2	97.12	94.15
3	96.54	93.25
4	96.76	95.24

3.3 模型性能比较

为验证所建立方法的有效性,在测试数据集上分别对 Faster R-CNN、YOLO v3、YOLO v5s 这 3 种模型进行算法性能比较,结果如表 2 所示。

表 2 不同模型性能的评价指标对比

模型	图像分辨率	准确率/%	召回率/%	模型存储空间/MB	平均准确率/%	运算时间/s
Faster R-CNN	640×640			512.23	92.74	0.430
YOLO v3	640×640	92.13	90.16	117.64	83.11	0.016
YOLO v5s	640×640	94.27	94.01	13.78	92.43	0.013
改进 YOLO v5s 算法	640×640	94.89	96.76	15.12	95.24	0.011

从表 2 中可以看出,改进算法的 mAP 最高,改进后的 YOLO v5s 模型 mAP 相比未改进的 YOLO v5s 模型提高了 2.81%。YOLO v3 的准确率和召回率虽然和改进方法相近,但在模型存储空间方面,改进模型比 YOLO v3 减少了 102.52 MB。综上,与其他 3 种算法相比,改进后的模型具有较高的 mAP 和较快的运算速度,其 mAP 比 Faster R-CNN、YOLO v3、YOLO v5s 分别提高 2.50%、12.13%、2.81%。

为提高虫洞的识别准确率提供了一种新途径。通过数据增强扩增了样本;使用 K-means⁺⁺聚类优化了初始锚框,减少了拟合时间,加快了网络收敛;添加注意力模块并通过新增特征尺度结合特征融合的方式最终达到虫洞的有效识别。试验结果表明,在测试数据集上检测,对大豆叶片虫洞识别的平均精度达 95.24%,检测速度每秒 91 帧,其准确率和速度均优于 Faster R-CNN、YOLO v3、YOLO v5s 等方法。

4 结论

基于改进 YOLO v5s 的大豆叶片虫洞识别方法

参考文献:

[1] 汤忠琴,尚静,张磊,等.不同田间配置对套作大豆

- 主要虫害的种群分布影响[J]. 四川农业大学学报, 2018, 36(3): 297–302.
- [2] HORIKOSHI R J, DOURADO P M, BERGER G U, et al. Large-scale assessment of lepidopteran soybean pests and efficacy of Cry1Ac soybean in Brazil[J]. Scientific Reports, 2021, 11: 15956.
- [3] GIRSHICK R. Fast r-cnn[C]// 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1440–1448.
- [4] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]// 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2961–2969.
- [5] JIANG H Z, LEARNED-MILLER E. Face detection with the faster R-CNN[C]// 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition. Washington D C, USA: IEEE, 2017: 650–657.
- [6] ZOU D, LONG W J, LING Z. A two-phase plagiarism detection method[C]// 2011 International Conference on Internet Technology and Applications. Wuhan, China: IEEE, 2011: 1–4.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 779–788.
- [8] 鲍文霞, 邱翔, 胡根生, 等. 基于椭圆型度量学习空间变换的水稻虫害识别[J]. 华南理工大学学报(自然科学版), 2020, 48(10): 136–144.
- [9] 赵志焱, 杨华, 胡志伟, 等. 基于 TACNN 的玉露香梨叶虫害识别[J]. 计算机工程与应用, 2021, 57(9): 176–181.
- [10] 李静, 陈桂芬, 安宇. 基于优化卷积神经网络的玉米螟虫害图像识别[J]. 华南农业大学学报, 2020, 41(3): 110–116.
- [11] IJAZ M F, ATTIQUE M, SON Y. Data-driven cervical cancer prediction model with outlier detection and over-sampling methods[J]. Sensors(Basel, Switzerland), 2020, 20(10): 2809.
- [12] VERMA V, LAMB A, BECKHAM C, et al. Manifold mixup: better representations by interpolating hidden states[C]//International Conference on Machine Learning. Long Beach, USA: PMLR, 2019: 6438–6447.
- [13] 侯瑞环, 杨喜旺, 王智超, 等. 一种基于 YOLOv4-TIA 的林业害虫实时检测方法[J]. 计算机工程, 2022, 48(4): 255–261.
- [14] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE, 2019: 658–666.
- [15] NEUBECK A, VAN GOOL L. Efficient non-maximum suppression[C]//International Conference on Pattern Recognition. Gangwon, Korea: IEEE, 2006: 850–855.
- [16] 金龄杰, 林志玮, 洪宇. 融合多尺度特征和梯度信息的云种类识别[J]. 激光与光电子学进展, 2022, 59(18): 155–164.
- [17] 范加利, 田少兵, 黄葵, 等. 基于 Faster R-CNN 的航母舰面多尺度目标检测算法[J]. 系统工程与电子技术, 2022, 44(1): 40–46.
- [18] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS: improving object detection with one line of code[C]// Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 5561–5569.
- [19] 侯志强, 刘晓义, 余旺盛, 等. 使用 GIoU 改进非极大值抑制的目标检测算法[J]. 电子学报, 2021, 49(4): 696–705.
- [20] GADDAM S R, PHOHA V V, BALAGANI K S. K-means+ ID3: A novel method for supervised anomaly detection by cascading K-means clustering and ID3 decision tree learning methods[J]. IEEE Transactions on Knowledge and Data Engineering, 2007, 19(3): 345–354.
- [21] 张诗文, 邓春华, 张俊雯. 各向异性非极大值抑制在工业目标检测中的应用[J]. 计算机应用, 2022, 42(7): 2210–2218.
- [22] CHLUBNA T, MILET T, ZEMČÍK P. Real-time per-pixel focusing method for light field rendering[J]. Computational Visual Media, 2021, 7(3): 319–333.
- [23] 谢娟英, 鲁银圆, 孔维轩, 等. 基于改进 RetinaNet 的自然环境中蝴蝶种类识别[J]. 计算机研究与发展, 2021, 58(8): 1686–1704.
- [24] 卓禹心, 韩素雅, 张榆锋, 等. 基于超声谐波包络 Nakagami 参数图像的微波消融区域自动分割方法[J]. 计算机应用, 2021, 41(10): 3089–3096.
- [25] 李灏, 唐敏, 林建武, 等. 基于改进困难三元组损失的跨模态行人重识别框架[J]. 计算机科学, 2020, 47(10): 180–186.

责任编辑: 罗慧敏
英文编辑: 罗维