

基于多尺度卷积神经网络特征融合的植株叶片检测技术

李颖¹, 陈龙¹, 黄钊宏², 孙杨², 蔡国榕²

(1. 集美大学诚毅学院, 福建 厦门 361000;

2. 集美大学计算机工程学院, 福建 厦门 361000)

摘要: 植株叶片检测是植株科学培育和精准农业过程中重要的环节之一。传统植株叶片检测的做法对操作人员的专业知识提出了较高要求, 且人工成本高、耗时周期长。基于此, 提出基于多尺度卷积神经网络特征融合(MCFF)的植株叶片检测技术。从深度学习技术辅助植株培育的需求出发, 基于多尺度卷积神经网络特征融合, 针对莲座模式植物、拟南芥和烟草3种不同类型、不同分辨率的植株进行叶片计数检测。经过与其他主流算法的比较, 发现MCFF具备较高的检测精确度, 平均精度均值(mAP)为0.662, 实现了高度竞争的性能(AP=0.946), 各项指标接近实用水平。

关键词: 深度学习; 目标检测; 多尺度卷积神经网络特征融合; 植株叶片检测技术

中图分类号: TP391.41, TN41

文献标识码: A

doi:10.11959/j.issn.2096-6652.202131

Plant leaf detection technology based on multi-scale CNN feature fusion

LI Ying¹, CHEN Long¹, HUANG Zhaohong², SUN Yang², CAI Guorong²

1. Chengyi University College, Jimei University, Xiamen 361000, China

2. College of Computer Engineering, Jimei University, Xiamen 361000, China

Abstract: Plant leaf detection is one of the essential aspects of the scientific plant breeding and precision agriculture process. The traditional practice of plant leaf detection requires professional knowledge of the operators, high labor costs, and long time-consuming cycles. The plant leaf detection technology based on multi-scale CNN feature fusion (MCFF) was proposed. Starting from the needs of deep learning technology assisted plant cultivation, a MCFF was used to detect leaf count for three different types and resolutions of rosette model plants, arabidopsisthaliana, and tobacco. Compared with the other three algorithms, the MCFF has a higher detection accuracy with an average detection rate of mAP 0.662, a highly competitive performance (AP = 0.946) has been achieved for each indicator close to the practical level.

Key words: deep learning, object detection, multi-scale CNN feature fusion, plant leaf detection

1 引言

植株叶片检测是植物培育和精准农业中重要的环节。叶片是植株制造有机物的主要场所, 叶片

的生长状态与植株产量和品质密切相关。植株叶片的检测通过对植株叶片的数量、形状、表型特征等方面进行检测, 对植株的生长状况进行科学系统的评估。根据植株叶片的检测结果, 农林工作者可以

收稿日期: 2021-06-15; 修回日期: 2021-08-18

通信作者: 孙杨, sunyang@jmu.edu.cn

基金项目: 国家自然科学基金资助项目(No.41971424, No.61702251, No.61701191, No.41871380, No.U1605254); 厦门科学技术项目(No.3502Z20183032); 福建省中青年教师教育科研项目(No.JT180877)

Foundation Items: The National Natural Science Foundation of China (No.41971424, No.61702251, No.61701191, No.41871380, No.U1605254), The Science and Technology Project of Xiamen (No.3502Z20183032), The Education and Scientific Research Project for Middle-Aged and Young Teachers of Fujian Province (No.JT180877)

判断植株类型、评估植株发育阶段,定位病害区域并及时进行防治处理。其中,植株叶片数量的检测数据为评估植株生长发育状况提供了有效的支持,进而对科学调控植株生育期的营养具有非常重要的指导意义。根据植株不同的生育时期对养分的不同需求,科学地进行水肥管理,实行精准农业,可以提升植株生长品质,降低种植成本。同时,通过检测样本植株的叶片,得出普适的生长规律,实现植株科学种植和精准农业,对于提高植株的品质和产量以及未来农业发展都具有重要意义。

传统的植株检测方法主要由农林工作者依据个人专业知识和经验,从植株叶片的数量、形状和表型特征等方面对植株的生长状况进行评估,整个研究过程需要耗费大量的人力和时间,且对专业知识有较高要求。事实上,受个人精神状态、身体状况、知识经验甚至天气等外在因素的影响,人工植株叶片检测不够精准。为了避免外界因素对植株叶片检测的影响,提高植株叶片检测效率,在植株叶片检测过程中引入了目前主流的深度学习技术^[1](如基于结构元素的描述符^[2]、基于多因素相关性^[3]、基于检测文本^[4]、基于纹理^[5]等)对图像实现智能检索,提升植株叶片检测的正确率,降低检测工作的劳动强度,并缩短检测时间。

但是,在检测过程中,当遇到样本尺寸多样、分辨率低、重叠遮挡等问题时,容易出现目标误检、漏检等情况,这对检测精确度提出了新的挑战。为了解决以上问题,本文基于多尺度卷积神经网络特征融合(multi-scale CNN feature fusion, MCFF)算法(以下简称多尺度回归算法)对植株叶片进行目标检测,在主干网络结合像素重构网络和梯度集成结构的优势,促进特征融合,满足多尺度目标检测的需求,在预测端采用广义交并比损失函数,提升检测精度。通过与 Faster R-CNN、RetinaNet 和 CenterNet 3 种经典深度学习技术的植株叶片检测结果进行比较可知,综合植株叶片样本尺寸多样、分辨率不同、存在重叠遮挡等多种复杂情况,基于多尺度回归算法的植株叶片检测精度最佳。

2 深度学习技术

2006 年,深度学习技术由 Hinton G E 等人^[6]首次提出,是机器学习的重要分支。深度学习技术结合无监督逐层的预训练和有监督的训练调整^[7],有效地从大量的原始数据中提取出能表达有效特

征的信息,并实现对训练集的最佳逼近。随着科技的迅速发展,深度学习技术在现实生活中的具体应用受到广泛的关注,在人脸识别^[8-9]、自动驾驶^[10]、信息融合中的图像检索^[11-13]等方面展现出强大的生命力。目前深度学习技术主要被应用于自然语言处理、语音识别、计算机视觉及相关的子领域^[14]。其中,计算机视觉包括目标分类、目标定位、目标检测和目标分割四大核心任务。

根据是否有候选框,基于深度学习技术的目标检测可分为双阶段目标检测和单阶段目标检测。双阶段目标检测是基于候选框的目标检测方法^[15]。双阶段目标检测方法首先需要在图像上生成候选框,然后借助深度卷积神经网络对候选框进行特征提取,最后对候选框内的图像类别进行判断、分类和处理。2014 年由 Girshick R 等人^[16]提出的 R-CNN 是基于候选框的典型的识别方法,是第一个用深度学习方法代替传统的目标检测的算法。2015 年微软亚洲研究院的 He K M 等人^[17]提出了 SPP-Net,通过在卷积层和全连接层之间加入空间金字塔结构有效地代替 R-CNN 的区域裁剪和缩放,并且有效地解决了 R-CNN 重复提取特征的问题。随后, Girshick R^[18]借鉴 SPP-Net 的结构提出了 Fast R-CNN,设计了 ROI pooling 的池化层结构,有效地解决了 R-CNN 算法中图像区域尺寸必须统一的问题。但是, Fast R-CNN 仍然没有摆脱选择性搜索算法生成正负样本候选框的问题。由 Ren S Q 等人^[19]提出的 Faster R-CNN 算法设计了区域候选框提取网络——RPN(region proposal network)。首先,判断候选框是否为目标,然后,根据分类定位的多任务损失判断目标类型,减少候选框数量,解决正负样本候选框生成速度慢的问题。事实上,双阶段目标检测策略在速度上不具有优势,因为双阶段网络需要先后对特征提取网络的区域候选框的生成网络和目标分类网络进行优化。于是,以快速轻量检测为特点的单阶段网络应运而生。

单阶段网络目标检测将整个目标检测任务当作一个参数的回归任务,参数包括边界框位置、边界框大小以及边界框内对象类别。RetinaNet、CenterNet 和 YOLO 系列等都是单阶段目标检测的经典网络。Lin T Y 等人^[20]借鉴了 Faster R-CNN 和多尺度目标检测的思想,设计训练出 RetinaNet 目标检测器,该模型的核心思想是通过重塑 focal loss 损失函数来解决训练过程中检测模型的训练样本

出现正负样本不平衡的问题。2019年 Duan K W 等人^[21]提出的 CenterNet 是一种无锚点的单阶段目标检测网络,该网络将待检目标中心视为关键点,然后分别对目标的长宽以及中心点的偏移量进行回归,进而完成目标检测任务,该网络在速度和精度上表现优秀。YOLO 算法是 Redmon J 等人^[22]于 2016 年提出的一种基于回归的单阶段目标检测算法。YOLO 网络首先通过卷积神经网络对输入图像进行特征提取,得到大量降分辨率的特征图谱;其次通过梯度更新算法对参数进行反向更新,然后由网络的最后一层输出预测框;最后用非极大值抑制(non-maximum suppression, NMS)对每个预测框的重叠区域进行概率分析和合并,得到最优检测框。本文在 YOLO 系列框架的基础上提出了多尺度回归算法。

多尺度回归算法模型由输入端、主干网络、连接网络和预测端 4 个部分组成,如图 1 所示,其中 conv 表示卷积,concat 为连接函数。首先,由输入端模块进行数据增强、图像自适应缩放、锚框的计算等工作;其次,主干网络模块中引入像素重构网络(Focus)和梯度集成结构(CSP),以减少漏检情况,降低计算量,提高样本信息提取速度;再次,连接网络模块借助特征金字塔网络(feature pyramid network, FPN)进行上下采样,加强特征融合,提升样本多样性和鲁棒性;最后,在预测端模块中增加 GIoU_Loss 函数的回归训练,优化样本重叠、遮

挡问题,提升目标检测效率。

3 基于多尺度回归算法的植株叶片检测

3.1 输入端

多尺度回归算法将 $608 \times 608 \times 3$ 的植株叶片样本作为输入,在输入端进行多图像拼接的数据增强、自适应图像缩放和自适应锚框计算等操作,以适应网络输入的大小,并进行归一化处理^[23],具体如下。

- 多图像拼接的数据增强每次读取 4 张植株叶片样本,通过随机缩放、随机裁剪、随机排布的方式进行拼接,进而提升多尺度回归网络的训练速度,并降低模型的内存需求。通过多图像拼接的数据增强可提升植株叶片样本的多样性和网络的适应性,大大改善检测过程中出现的错检、漏检等问题。

- 植株叶片样本经过数据增强后,通过自适应图像缩放来适应网络输入大小,减少噪声信息冗余,提升算法推理速度。

- 由于网络训练前需要设定初始锚框,多尺度回归算法提出了自适应锚框计算,针对不同的植株数据集设置特定长宽的锚框。网络模型在基础锚框的基础上训练得到预测框,并与真值框进行比较,根据差值执行反向更新操作,从而迭代整个多尺度回归网络的模型参数。

3.2 主干网络

经输入端处理后输出的 $608 \times 608 \times 3$ 的植株叶片样本通过由像素重构网络和梯度集成结构组成

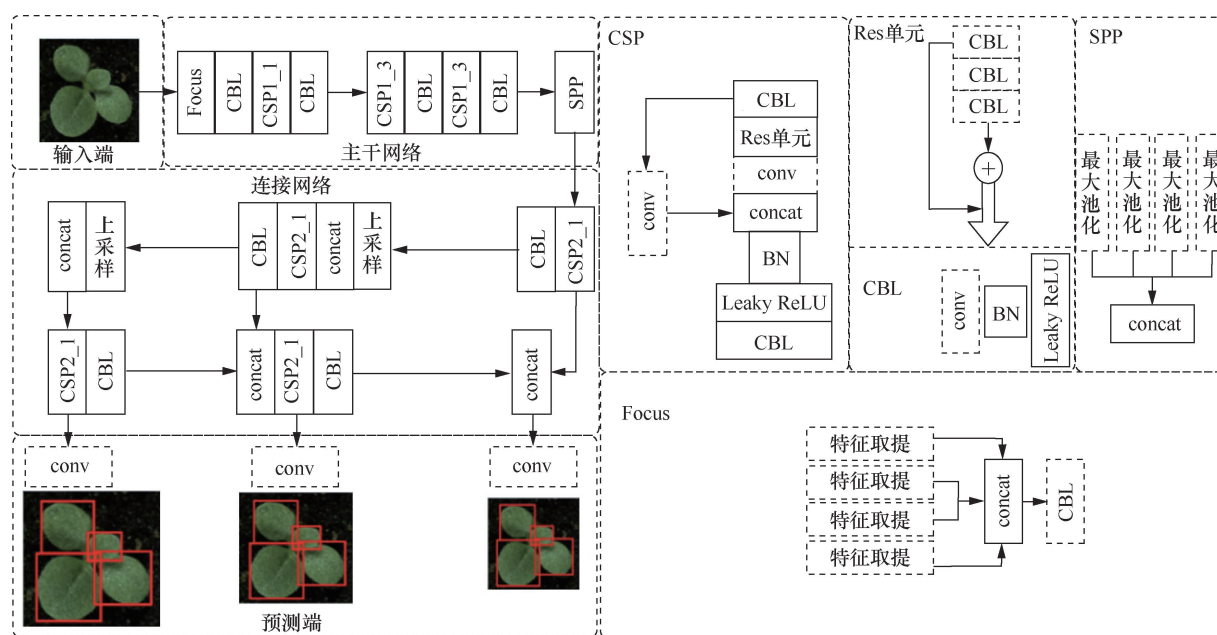


图1 多尺度回归算法流程

的主干网络，输出植株叶片样本的特征^[24]。

像素重构网络首先借助切片与连接操作将 $608 \times 608 \times 3$ 的植株叶片样本输出成维度为 $304 \times 304 \times 12$ 的特征向量，然后将该特征向量作为一个通道数为 32 的卷积层的输入，最后输出 $304 \times 304 \times 32$ 的特征向量。像素重构网络从高分辨率的植株叶片样本中周期性地提取像素点重构到低分辨率的样本中，提高植株图像的感受野，同时减少图像原始信息的丢失。像素重构网络有效地解决了正负样本不平衡的问题，促进了样本特征的融合，不但减少了因植株叶片样本特征分散导致的漏检现象，而且大大减少了网络的计算量。

多尺度回归算法的设计融合了主干梯度和聚合梯度两种梯度集成结构。其中，主干梯度集成结构应用于主干网络。聚合梯度集成结构将网络的梯度变化集成到植株特征图中，整合各植株叶片的信息，提升检测的准确度。多尺度回归算法在主干网络的最后进行了多尺度池化操作，采用 1×1 、 5×5 、 9×9 、 13×13 的最大池化方式对植株叶片样本进行多尺度特征融合，有效地增加了主干网络的植株叶片信息的接受范围，显著地分离了重要的上下文特征，可适应多尺度目标检测的需求。

3.3 连接网络

多尺度回归算法通过上采样特征金字塔和下采样特征金字塔构建连接网络^[25]，并借鉴梯度集成网络中的梯度集成结构来加强网络特征融合的能力。多尺度回归算法在上采样特征金字塔对植株叶片样本进行上采样，自上向下传达强语义特征，以获得预测特征图。

随后通过两个自下向上的特征金字塔结构对植株叶片样本特征图进行下采样。多尺度回归算法通过连接网络进一步提升植株叶片特征的多样性和鲁棒性，有效地解决了低分辨率的植株叶片样本特征弱、检测困难等问题。

3.4 预测端

多尺度回归算法在预测端采用广义交并比损失函数（GIoU_Loss）进行回归训练，广义交并比损失函数相比交并比损失函数（IoU_Loss）增加了相交尺度的融合方法，可有效地解决植株叶片样本预测框与真值框之间边界不重合交并比（intersection over union, IoU）不合理的问题^[26]。当获得一定数量的预测框后，采用 NMS 的加权进行候选框的筛选，进而优化遮挡、重叠目标的检测效

率，解决植株叶片之间相互遮挡、重叠的问题。

3.5 广义交并比损失函数

植株叶片样本在计算损失阶段，分别将多尺度回归网络输出的植株叶片预测框 B^p 和真值框 B^s 的坐标（检测框左上角坐标和右下角坐标）作为输入，通过检测框坐标的大小关系计算相邻最远的两个坐标 (x_1^c, y_1^c) 和 (x_2^c, y_2^c) ，进而计算植株叶片样本预测框与真值框之间的整体框面积 A^c 和植株叶片样本真值框与预测框之间的重叠面积 $\mathcal{I}$ ，同时计算预测框和真值框的总面积 $\mathcal{U}$ 。多尺度回归算法通过计算整体框面积 A^c 和两框总面积 $\mathcal{U}$ 之间的差值获得额外面积，然后将额外面积与整体框面积 A^c 相除，获得浮点型小数；接着，将交并比与获得的浮点型小数相减，得到植株叶片预测框与真值框之间的广义交并比 GIoU；最后，1 减去广义交并比的值，即得到广义交并比损失函数 $\mathcal{L}_{\text{GIoU}}$ ，具体如下。

算法 1 广义交并比损失函数

输入 预测框 B^p 和真值框 B^s 的坐标：

$(x_1^p, y_1^p, x_2^p, y_2^p)$ 和 $(x_1^s, y_1^s, x_2^s, y_2^s)$

输出 广义交并比损失函数 $\mathcal{L}_{\text{GIoU}}$

步骤 1: 对于每个预测框 B^p ，如果 $x_2^p > x_1^p$ 且 $y_2^p > y_1^p$ ，则：

$$\hat{x}_1^p = \min(x_1^p, x_2^p), \hat{x}_2^p = \max(x_1^p, x_2^p)$$

$$\hat{y}_1^p = \min(y_1^p, y_2^p), \hat{y}_2^p = \max(y_1^p, y_2^p)$$

步骤 2: 计算真值框 B^s 的面积 A^s ；

$$A^s = (x_2^s - x_1^s) \times (y_2^s - y_1^s)$$

步骤 3: 计算预测框 B^p 的面积 A^p ；

$$A^p = (\hat{x}_2^p - \hat{x}_1^p) \times (\hat{y}_2^p - \hat{y}_1^p)$$

步骤 4: 计算 B^p 和 B^s 重叠的面积 $\mathcal{I}$ ；

$$x_1^{\mathcal{I}} = \max(\hat{x}_1^p, x_1^s), x_2^{\mathcal{I}} = \min(\hat{x}_2^p, x_2^s)$$

$$y_1^{\mathcal{I}} = \max(\hat{y}_1^p, y_1^s), y_2^{\mathcal{I}} = \min(\hat{y}_2^p, y_2^s)$$

如果 $x_2^{\mathcal{I}} > x_1^{\mathcal{I}}$ 且 $y_2^{\mathcal{I}} > y_1^{\mathcal{I}}$ ， $\mathcal{I} = (x_2^{\mathcal{I}} - x_1^{\mathcal{I}}) \times (y_2^{\mathcal{I}} - y_1^{\mathcal{I}})$ ；

否则， $\mathcal{I} = 0$ 。

步骤 5: 计算整体框 B^c 的面积 A^c ；

$$x_1^c = \min(\hat{x}_1^p, x_1^s), x_2^c = \max(\hat{x}_2^p, x_2^s)$$

$$y_1^c = \min(\hat{y}_1^p, y_1^s), y_2^c = \max(\hat{y}_2^p, y_2^s)$$

$$A^c = (x_2^c - x_1^c) \times (y_2^c - y_1^c)$$

步骤6: 预测框和真值框的总面积 $\mathcal{U}$;

$$\mathcal{U} = A^p + A^g - \mathcal{I}$$

步骤7: 计算交并比, $\text{IoU} = \frac{\mathcal{I}}{\mathcal{U}}$;

步骤8: 计算广义交并比, $\text{GIoU} = \text{IoU} - \frac{A^c - \mathcal{U}}{A^c}$;

步骤9: 计算广义交并比损失函数, $\mathcal{L}_{\text{Giou}} = 1 - \text{Giou}$ 。

4 实验和分析

为了检验多尺度回归算法在植株叶片检测中的检测效果, 本文基于植株叶片数据集, 对 Faster R-CNN、CenterNet、RetinaNet 和本文提出的多尺度回归算法4种目标检测算法的植株叶片检测结果进行比较分析。

4.1 实验数据集

本文基于 Massimo Minervini 团队提供的植物数据集 (347 张)^[27]制作训练数据集和测试数据集。数据集包括3种植株类别, 分别是莲座模式植物、拟南芥和烟草。其中, 莲座模式植物数据集120张, 拟南芥数据集165张, 烟草数据集62张。数据集采用 PASCAL VOC^[28-29]格式, 并在原有数据集的基础上采用翻转、旋转等方式进行扩充, 最终得到4511张图像。采用标图工具 Lable Img 对图像目标进行标注, 如图2所示, 将 XML^[30]作为图像文件的后缀。实验数据集划分如下: 总数据集4511张, 训练集3157张, 测试集1354张。

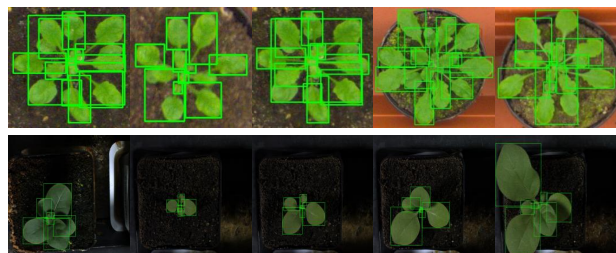


图2 植株叶片数据集中的真实标注框示例

4.2 实验设置

在实验过程中, 植株叶片数据集在构建网络时需要一定数量的训练集。本实验采用莲座模式植物、拟南芥和烟草的数据集, 包含4511张带位置标注的植株叶片图像, 对其进行植株叶片计数检测, 即叶片繁茂程度的预测。假设当预测框与标注的真值框的交并比大于0.5时, 判定预测框的结果为正确。

为了充分验证多尺度回归算法的可靠性, 全面有效地对多尺度回归网络进行评估检测, 实验选择不同分辨率的植株叶片样本, 具体包括: 高分辨率的烟草数据集, 像素为2048×2048; 中分辨率的莲座模式植物数据集, 像素为355×355; 低分辨率的拟南芥数据集, 像素为40×40。

为了保证图像在空间和尺寸上的一致性, 实现像素空间的归一化, 实验将图像调整为像素值为[0, 255]且分辨率为512×512的三通道图像。

4.3 网络训练

在多尺度回归网络模型训练阶段, 迭代批大小设置为16, 衰减系数为0.0005, 总迭代次数为200次, 初始学习率设置为0.001。当迭代次数分别达到100次和150次时, 学习率分别降低至0.0001和0.00001, 大约在180次迭代后, 模型收敛。

RetinaNet、CenterNet、Faster R-CNN 模型的训练阶段比较相似, 迭代批大小设置为16, 衰减系数为0.0005, 总迭代次数为200次, 初始学习率设置为0.001。当迭代次数分别达到100次和150次时, 学习率分别降低至0.0001和0.00001, 大约在170次迭代后, 模型收敛。

4.4 实验结果

4种目标算法基于混合植株、莲座模式植物、拟南芥和烟草4种不同的植株叶片测试数据集进行检测比较, 得出平均精度均值 (mAP)、平均精度 (AP) 和召回率, 如图3所示。由图3可知, 在4种植物数据集上, 相比其他3种算法, 多尺度回归算法的平均精度均值和平均精度都获得最好的效果, 证明本文采用的基于多尺度回归算法的植株叶片检测方法, 结合自适应锚框计算、像素重构网络、梯度集成结构和广义交并比损失函数, 可获得更加可靠的检测精确度。

模型训练完成后, 将测试集输入烟草模型中, 分别用4种目标检测算法得到植株叶片检测框, 进行叶片繁茂程度分析, 如图4所示, 其中绿色框表示真值框, 红色框表示预测框。检测结果表明, 4种算法均能实现目标叶片数量的检测, 但是还存在一些问题。CenterNet 算法在热力图中提取的特征较弱, 导致小目标检测中的精度低, 在植株小叶片的检测中出现误检现象, 如图4(a)所示; 双阶段检测算法 Faster R-CNN 网络结构较落后, 导致在植株叶片检测中, 预测框和真值框之间的重叠度不高, 交并比偏低, 检测精确度低, 如图4(b)所示; RetinaNet 算法正负样本不平衡, 导致漏检了植株最下面的小

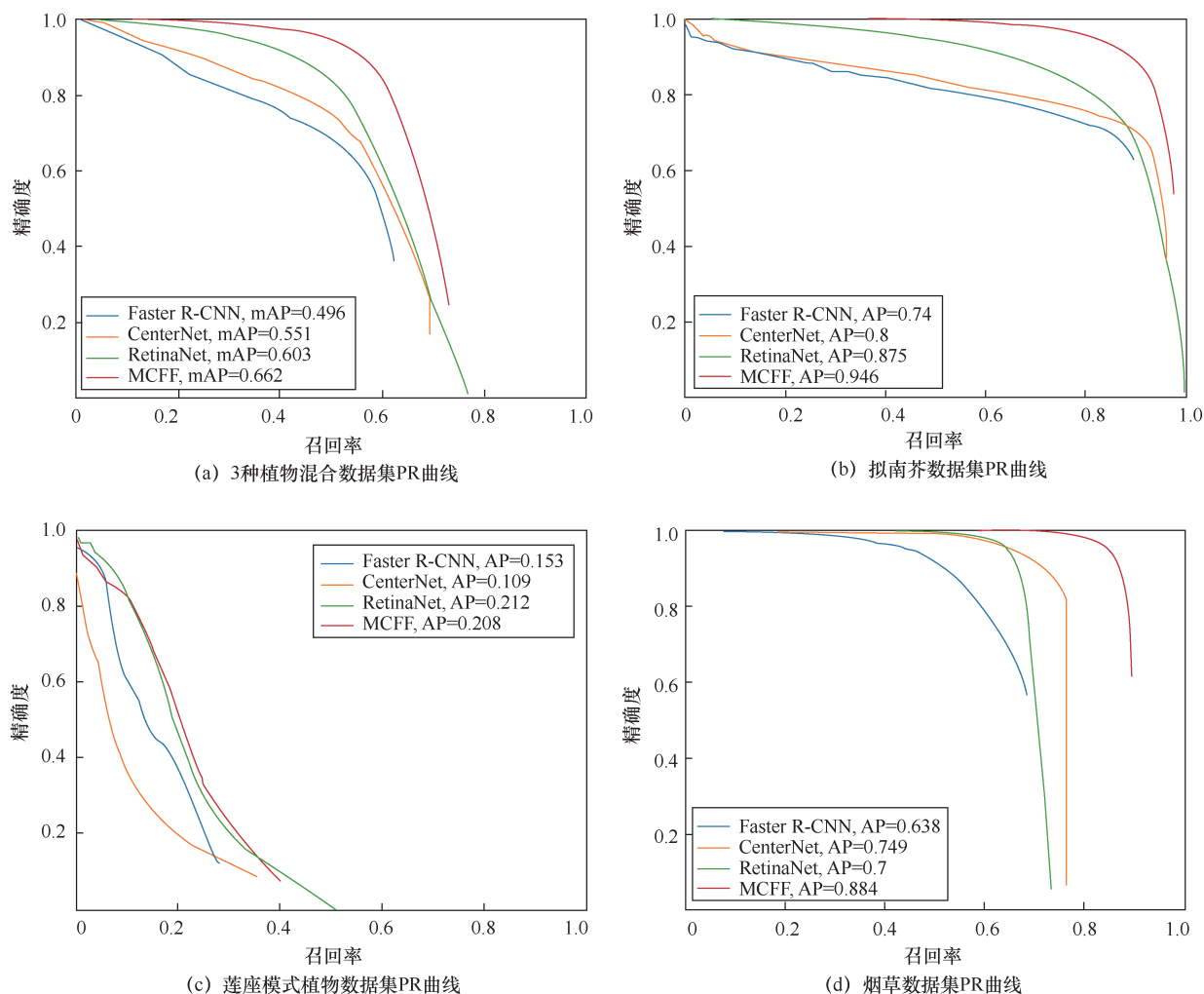


图3 基于4种植物数据集的4种算法的精确度-召回率(PR)曲线

叶片,如图4(c)所示,这说明RetinaNet算法对小目标样本的检测不敏感。多尺度回归算法有效地解决了上述3种算法的问题,如图4(d)所示,预测框和真值框重叠度很高,获得了最高的叶片检测框回归精度。

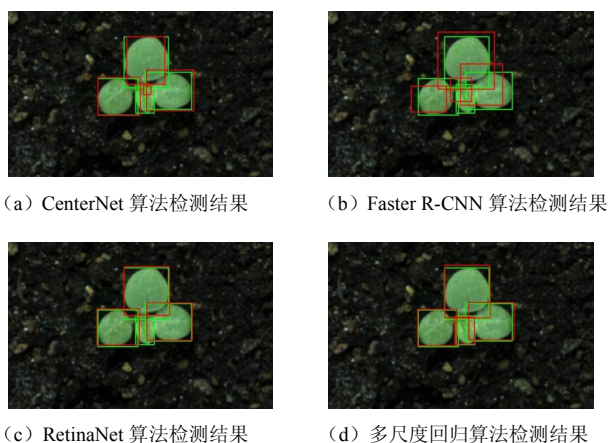


图4 基于烟草数据集的4种算法的检测结果

表1为4种算法在不同植株叶片数据集下的检测精确度评估。由于多尺度回归算法有效地解决了CenterNet算法特征提取不明显、RetinaNet算法正负样本不平衡、Faster R-CNN网络结构落后等问题,其植株叶片检测平均精度大幅度提升,多尺度回归算法平均精度均值达到0.662,在植株叶片检测中具有突出优势。

表1 各种算法的平均精度均值及各植株数据集的平均精度

算法	mAP	烟草 AP	莲座模式植物 AP	拟南芥 AP
Faster R-CNN	0.496	0.638	0.153	0.74
RetinaNet	0.603	0.7	0.212	0.875
CenterNet	0.551	0.749	0.109	0.8
多尺度回归算法	0.662	0.884	0.208	0.946

在高分辨率的烟草数据集检测过程中,烟草植物叶片尺寸多样对叶片的检测提出考验,多尺度回归算法对烟草测试的平均精度高达0.884,这证明

多尺度回归算法中的多尺度融合检测和特征增强在检测多样性的数据集方面具有较强的优势。

在中分辨率的莲座模式植物数据集检测过程中,莲座叶片尺寸多样、重叠较多给植株叶片检测带来了困难。4种算法的检测性能都较弱。但多尺度回归算法的测试平均精度仍达到0.208,说明在极具挑战的低分辨率、多尺度、多重叠的复杂数据集下,多尺度回归算法依旧具有一定程度的检测性能。

在低分辨率的拟南芥数据集检测过程中,多尺度回归算法的平均精度为0.946,远高于其他算法。证明多尺度算法的像素重构网络和梯度集成结构可以促进特征融合、提升检测精度,对检测低分辨率数据集具有明显的优势。

综上,数据集样本像素低、目标尺度复杂、叶片重叠等原因均会对植株叶片检测效果造成一定的影响,导致误检、漏检等问题。相较于其他3种经典的目标检测算法,多尺度回归算法在各种复杂植株叶片数据集下都获得了较优秀的检测效果,能够满足实际应用的需求,但是针对低分辨率的数据集检测仍需做进一步的优化改进。

表2为在3种植物混合的多类别数据集下的检测结果。多尺度回归算法在多类别数据集下达到了0.662的平均精度均值,相比双阶段算法Faster R-CNN提高了0.166,相比单阶段算法CenterNet和RetinaNet分别提高了0.111和0.059。植株叶片样本贴近实际应用场景会出现植株叶片样本上下文特征不明显、目标尺寸多样等问题,进而导致其他3种算法的平均精度均值出现较大幅度的下降,而多尺度回归算法针对植株叶片样本检测中存在的问题进行了网络结构上的优化,在实际应用场景中具备显著优势。

表2 多类别植物数据集的平均精度均值

算法	mAP
Faster R-CNN	0.496
Retinanet	0.603
CenterNet	0.551
多尺度回归算法	0.662

表3为4种算法的准确度和平均检测时间。多尺度回归算法的准确度为88.73%,分别比Faster R-CNN、CenterNet和RetinaNet3种算法提高了52.03%、0.91%、9.59%。多尺度回归算法每检测100张图像的平均时间为30.1 ms,相比双阶段算法

中的FasterR-CNN,效率提高了4倍左右,大大缩短了检测时间。多尺度回归算法在检测准确度和平均检测时间上优于其他3种算法,能够实现在实际环境中对植株叶片繁茂程度进行实时、精准的评估,不仅大大提高了植株叶片检测效率,还减轻了传统植株测量工作的劳动强度。

表3 4种算法的准确度和平均检测时间

算法	准确度	平均检测时间/ms
Faster R-CNN	36.7%	159
CenterNet	87.82%	152
RetinaNet	79.14%	148
多尺度回归算法	88.73%	30.1

5 结束语

为了提高植株叶片检测的准确性和实时性,将人工从烦琐的检测工作中解放出来,本文将深度学习技术融入植株叶片的目标检测工作中,通过实验比较多种经典的目标检测算法可知,基于多尺度回归算法网络模型的植株叶片检测算法的检测准确度达到较高标准,各项指标接近实用水平。该技术可推进植株叶片检测过程的智能化和科学化,对植株的科学培育和推进精准农业具有较好的实际应用价值。未来,笔者计划在多领域应用中实现并验证多尺度回归算法,如人脸识别、工业质检等,考虑引入学习低级特征^[31]等新策略,提升检测模型精度,提高检测效率。

参考文献:

- [1] 张钹. 人工智能进入后深度学习时代[J]. 智能科学与技术学报, 2019, 1(1): 4-6.
ZHANG B. Artificial intelligence is entering the post deep-learning era[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(1): 4-6.
- [2] WANG X Y, WANG Z Y. A novel method for image retrieval based on structure elements' descriptor[J]. Journal of Visual Communication and Image Representation, 2013, 24(1): 63-74.
- [3] WANG X Y, WANG Z Y. The method for image retrieval based on multi-factors correlation utilizing block truncation coding[J]. Pattern Recognition, 2014, 47(10): 3293-3303.
- [4] UNAR S, WANG X Y, ZHANG C, et al. Detected text-based image retrieval approach for textual images[J]. IET Image Processing, 2019, 13(3): 515-521.
- [5] WANG X Y, CHEN Z F, YUN J J. An effective method for color image retrieval based on texture[J]. Computer Standards & Interfaces, 2012, 34(1): 31-35.
- [6] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [7] 袁小锋, 王雅琳, 阳春华, 等. 深度学习在流程工业过程数据建模中的应用[J]. 智能科学与技术学报, 2020, 2(2): 107-115.
YUAN X F, WANG Y L, YANG C H, et al. The application of deep

- learning in data-driven modeling of process industries[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(2): 107-115.
- [8] BARGSHADY G, ZHOU X J, DEO R C, et al. Enhanced deep learning algorithm development to detect pain intensity from facial expression images[J]. Expert Systems with Applications, 2020, 149: 113305.
- [9] HAMMAM A A, SOLIMAN M M, HASSANIEN A E. Real-time multiple spatiotemporal action localization and prediction approach using deep learning[J]. Neural Networks, 2020, 128: 331-344.
- [10] LI Y, MA L F, ZHONG Z L, et al. Deep learning for LiDAR point clouds in autonomous driving: a review[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(8): 3412-3432.
- [11] WANG X Y, CHEN Z F. A fast fractal coding in application of image retrieval[J]. Fractals, 2009, 17(4): 441-450.
- [12] UNAR S, WANG X Y, WANG C P, et al. A decisive content based image retrieval approach for feature fusion in visual and textual images[J]. Knowledge-Based Systems, 2019, 179: 8-20.
- [13] UNAR S, WANG X Y, ZHANG C. Visual and textual information fusion using Kernel method for content based image retrieval[J]. Information Fusion, 2018, 44: 176-187.
- [14] 徐优俊, 裴剑锋. 深度学习在化学信息学中的应用[J]. 大数据, 2017, 3(2): 45-66.
- XU Y J, PEI J F. Deep learning for chemo informatics[J]. Big Data Research, 2017, 3(2): 45-66.
- [15] LIU L, OUYANG W L, WANG X G, et al. Deep learning for generic object detection: a survey[J]. International Journal of Computer Vision, 2020, 128(2): 261-318.
- [16] 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. Piscataway: IEEE Press, 2014: 580-587.
- [17] 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.
- [18] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 1440-1448.
- [19] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [20] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [21] DUAN K W, BAI S, XIE L X, et al. CenterNet: keypoint triplets for object detection[C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 6568-6577.
- [22] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.
- [23] ZHONG Z, ZHENG L, KANG G L, et al. Random erasing data augmentation[J]. arXiv Preprint, 2017, arXiv:1708.04896.
- [24] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2020: 1571-1580.
- [25] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 936-944.
- [26] YU J H, JIANG Y N, WANG Z Y, et al. UnitBox: an advanced object detection network[C]//Proceedings of the 24th ACM international conference on Multimedia. New York: ACM Press, 2016: 516-520.
- [27] MINERVINI M, FISCHBACH A, SCHARR H, et al. Finely-grained

annotated datasets for image-based plant phenotyping[J]. Pattern Recognition Letters, 2016, 81: 80-89.

- [28] VICENTE S, CARREIRA J, AGAPITO L, et al. Reconstructing PASCAL VOC[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 41-48.
- [29] EVERINGHAM M, VAN Gool L, WILLIAMS C K I, et al. The pascal visual object classes (VOC) challenge[J]. International Journal of Computer Vision, 2010, 88(2): 303-338.
- [30] BRAY T, PAOLI J, SPERBERG-MCQUEEN C M, et al. Extensible markup language (XML) 1.0[J]. World Wide Web Journal, 1997, 2(4): 29-66.
- [31] UNAR S, WANG X Y, WANG C P, et al. New strategy for CBIR by combining low-level visual features with a colour descriptor[J]. IET Image Processing, 2019, 13(7): 1191-1200.

[作者简介]



李颖(1983-),女,集美大学诚毅学院副教授,主要研究方向为图像处理、机器学习、计算机视觉、无线传感网络。



陈龙(2000-),男,集美大学诚毅学院在读,主要研究方向为机器学习、目标检测、网络通信。



黄钊宏(2001-),男,集美大学计算机工程学院在读,主要研究方向为深度学习、目标检测,图像识别、知识图谱。



孙杨(1975-),女,集美大学计算机工程学院讲师,主要研究方向为系统结构和计算机图形图像。



蔡国榕(1979-),男,博士,集美大学计算机工程学院教授,主要研究方向为图像处理、机器学习、计算机视觉。