

doi:10.6041/j.issn.1000-1298.2023.05.026

基于 polyphyletic 损失函数的荔枝花检测方法

叶进¹ 吴梦岚¹ 邱文杰¹ 杨娟² 兰伟¹

(1. 广西大学计算机与电子信息学院, 南宁 530004; 2. 广西大学农学院, 南宁 530004)

摘要: 针对密集分布的荔枝花遮挡严重导致检测困难, 现有研究方法忽略了非极大抑制过程中密集建议框之间的相互作用的问题, 为提升检测精度、降低漏检率, 提出了一种基于 polyphyletic 损失函数的检测方法。该方法在损失函数中包含一个聚合损失项, 以迫使建议框接近并紧凑定位相应目标, 同时增加专门为密集作物场景设计的边界框排斥损失, 使预测框远离周围对象, 提高密集荔枝花检测鲁棒性。与 Faster R-CNN、Focus Loss、AdaptiveNMS 和 Mask R-CNN 进行对比, 实验表明, 该方法在标准苹果花数据集上识别精度比其他方法高 2 个百分点, 验证了该方法检测的通用性, 同时, 该方法在自建荔枝花数据集的平均精度均值达到 87.94%, F1 值为 87.07%, 缺失率为 13.29%, 相比 Faster R-CNN、Focus Loss、AdaptiveNMS 和 Mask R-CNN 分别提高 20.09、14.10、8.35、4.86 个百分点, 具有较高检测性能。因此, 本文提出的方法能够高效地对密集荔枝花进行检测, 为复杂场景下的密集作物检测提供参考。

关键词: 荔枝花; 深度学习; 图像识别; 密集目标检测; 损失函数

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298 (2023) 05-0253-08 OSID: 

Litchi Flower Detection Method Based on Polyphyletic Loss Function

YE Jin¹ WU Menglan¹ QIU Wenjie¹ YANG Juan² LAN Wei¹

(1. School of Computer, Electronic and Information, Guangxi University, Nanning 530004, China

2. School of Agriculture, Guangxi University, Nanning 530004, China)

Abstract: The flowering intensity of litchi can directly affect the yield and quality of the fruit, so the detection of litchi flowers is very important for orchard planting strategies. Dense litchi flower detection has important challenges due to serious occlusion. Existing research methods ignore the interaction between dense suggestion boxes in the process of non-maximum suppression. In order to improve the detection precision and reduce the missed detection rate, a detection method was proposed based on polyphyletic loss function. This method included an aggregation loss term in the loss function to force the proposal box to approach and compactly locate the corresponding object. At the same time, the segmentation loss of the bounding box specially designed for the dense crop scene was added to keep the prediction box away from the surrounding objects and improve the robustness of detecting a large number of flowers. Compared with Faster R-CNN, Focus Loss, AdaptiveNMS and Mask R-CNN, the experiment showed that the recognition precision of this method on the standard apple blossom dataset was about 2 percentage points higher than that of other methods, which verified the detection versatility of this method. At the same time, the mean average precision of this method in the self-built litchi flower dataset was 87.94%, the F1 score was 87.07%, and the miss rate was 13.29%. Compared with Faster R-CNN, Focus Loss, AdaptiveNMS and Mask R-CNN, the accuracy of the method was improved by 20.09 percentage points, 14.10 percentage points, 8.35 percentage points and 4.86 percentage points, respectively, with high detection performance. Therefore, the method proposed can effectively detect the dense litchi flowers, and provide an important reference for dense crop detection in complex scenes.

Key words: litchi flower; deep learning; image recognition; dense object detection; loss function

0 引言

荔枝广泛种植在世界各地的亚热带地区,内含的生物活性化合物和矿物质使其具有非常高的营养价值^[1]。据国家荔枝工业技术体系估计,中国大陆荔枝产量近年来保持相对不变,约占世界荔枝生产面积的 52.60%;2020 年中国大陆荔枝产量约为 3.02×10^9 t,比 2019 年增长 54.82%,占全球荔枝产量的近 61.34%^[2]。在果实产量预测中,开花强度(即果园内的花数)指导修剪过程,将直接影响果实产量和质量^[3],因此荔枝花的检测对荔枝产量预测非常重要。

早期许多学者使用传统的图像处理方法对花朵的识别和检测进行了研究。NILSBACK 等^[4]提出了一种通过 RGB 图像自动分割花的算法。在这之后,CHAI 等^[5]提出了一种双层共分割的图像分类方法,通过将图像分割为前景和背景来提高分类性能。

由于农作物通常具有杂乱的背景,传统的图像处理方法对农业场景中花朵的检测效果较差。而荔枝花稠密聚集,在繁花期间花朵粘连成簇,密集花朵容易相互遮挡,这使得荔枝花的检测比传统的花朵检测更具挑战性^[6]。对于存在遮挡的作物检测,近年来的研究方法分为基于特征的手工方法和基于 CNN 的方法。

常见的特征提取方法包括加速鲁棒特征(Speeded-up robust features, SURF)^[7]、基于可变形部分的模型(Deformable part-based model, DPM)^[8]、标度不变特征变换(Scales-invariant feature transformation, SIFT)^[9]、定向梯度直方图(Histogram of oriented gradient, HOG)^[10]和 Haar - Like 特征^[11]。上述特征提取方法通常用于密度较低的图像^[12]。在许多应用中,使用深度学习提取的特征比手工制作的特征更快、更有效^[13]。

基于 CNN 的单级检测器可以有效地考虑分布在图像上的数十万个可能的对象位置,并在一个统一的网络中对其进行分类,YOLO^[14]、RetinaNet^[15]和 EfficientDet^[16]就是很好的例子。相比之下,两阶段探测器的第 1 阶段是从均匀分布的可能的位置过滤一组较小的候选对象,然后在第 2 阶段分类和细化,如 Faster R - CNN^[17]和 Mask R - CNN^[18]。然而,这些方法忽略了在遮挡中检测失败的根本原因:在非极大抑制(Non-maximum suppression, NMS)过程中,相互遮挡的对象的建议框很密集,它们会相互影响,超过交并比(Intersection over union, IoU)阈值的建议框会被过滤掉,导致检测遗漏。

为了解决由遮挡引起的荔枝花漏检问题,提高

对密集荔枝花检测的准确率,本研究提出调整建议框和预测框位置的 polyphyletic 损失函数。polyphyletic 损失迫使预测框紧密地接近和定位相应的物体,并训练一个可能对周围目标的排斥非常有利的最优检测器。为实现聚合能力和排斥能力,本文引入一种聚集损失和排斥损失。将该方法在自建的荔枝花数据集上与其他方法进行比较,验证该方法良好的检测性能。同时,在自建荔枝花数据集上与其他方法进行比较,并对苹果花数据集进行实验,验证该方法的有效性和实用性。

1 材料与预处理

1.1 数据集获取

本研究使用两种不同的数据集进行实验,一种是自建的荔枝花数据集,即本文所研究的对象,第 2 种是标准的苹果花数据集,用于本研究方法的通用性证明。荔枝花的自建图像采集于广西大学农学院试验田(22°50′28.41″N, 108°17′9.00″E)。这些图像拍摄于晴天中午,使用移动设备获得 1 039 幅图像。由于荔枝花小而致密,将图像分割成 9 × 9 个部分。经过筛选,共有 2 600 幅图像,包括荔枝花粘附、不同大小、不同无序角度、不同遮挡的情况。图像尺寸为 997 像素 × 997 像素。选择 2 600 幅图像中的 520 幅作为测试集,并将其余的 2 080 幅图像作为训练集。手动进行图像矩形注释,并保存在“xml”文件中。图 1 显示了荔枝花数据集的示例。

多种果花标准数据集^[19]包括 4 部分:苹果 A、苹果 B、桃子和梨。本文利用苹果 A 进行实验。这些数据是在美国农业部的果园内收集,图像中的苹果树由乔树支撑,并成排种植。

这些图像通过手持相机获得,共 147 幅图像。从这个总数中,随机选择 100 幅图像来构建一个 CNN 的训练集。平均而言,花只占这个数据集中总图像面积的 25%,故将图像进行 16 × 16 分割处理,分割后图像尺寸为 389 像素 × 389 像素,如图 2 所示。

1.2 数据增强

深度学习训练需要大量的数据来提高检测精度,防止模型训练过度拟合和不收敛。而且由于条件不同,人工采集的数据量通常不足,因此有必要增强数据集。本文采用几何变换的数据增强方法,包括随机旋转、水平翻转和垂直翻转。随机旋转是将图像随机在 0°~360°间旋转,本研究选择将图像分别旋转 90°、180°和 270°。水平翻转是指沿着图像的垂直中心线左右翻转图像。垂直翻转是指沿着图像的水平中心线上下翻转图像。



图 1 自建的荔枝花数据集

Fig. 1 Self-built litchi flower dataset

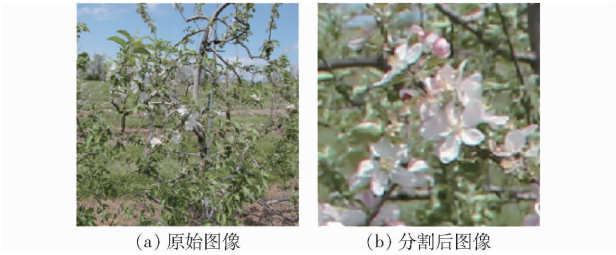


图 2 苹果花数据分割效果

Fig. 2 Apple blossom data segmentation effect

1.3 特征提取

图像特征作为一种可以在图像上识别的属性，一般分为形状特征、颜色特征和纹理特征 3 种。而特征提取就是通过分析处理图像内容中蕴含的信息来达到图像标识与分类的目的，此外特征提取的结果还会直接影响到图像处理效果^[20]。由于花朵图像特征复杂，常需先检测原图像中花朵轮廓，然后提取特征点。

在对数据集中花朵图像使用改进 Sobel 算子检测边缘之后，应用于 SIFT 算法特征提取中，即使用 SIFT 算法描述花朵图像的形状特征。尺度空间概念定义函数为

$$T(x,y,\rho) = S(x,y,\rho)I(x,y) \tag{1}$$

其中
$$S(x,y,\rho) = \frac{1}{2\pi\rho^2}e^{-(x^2+y^2)/(2\rho^2)} \tag{2}$$

式中 x ——横坐标 y ——纵坐标
 ρ ——标准差
 $T(x,y,\rho)$ ——尺度空间变换函数
 $I(x,y)$ ——原图中像素坐标
 S ——高斯函数

通过将不同尺度高斯差分与图像卷积获得高斯差分算子，能够在尺度空间内有效地检测出关键点。

具体函数公式为

$$D(x,y,\rho) = (S(x,y,k\rho) - S(x,y,\rho))I(x,y) = T(x,y,k\rho) - T(x,y,\rho) \tag{3}$$

式中 k ——常数 D ——高斯差分算子

关键点定位就是剔除低对比度、边缘等关键点，准确地定位关键点位置、尺度等信息，增强匹配精度。实验表明，当像素灰度范围为 0 ~ 1 时，可舍弃 0.04 以下的极值点，从而去除了相对不稳定极值点。根据像素在关键点区域内梯度方向的分布特征，可对花朵图像上的特征关键点规定方向参数（位置、尺度、方向），从而使关键点具备平移、缩放及旋转不变性。经过方向分配后，特征描述子可以更加精确地描述该图像中花的形状特征，可以更加完美地提取花的整体轮廓形状。

2 遮挡检测方法

2.1 polyphyletic 损失函数

为了提高预测框生成质量和检测精度，本研究提出 polyphyletic 损失函数，将聚合项和排斥项添加到经典的损失函数中^[17]，polyphyletic 损失函数定义为

$$L = L_{Cls} + \alpha L_{Reg} + \beta L_{Com} + \gamma L_{SegGT} + \delta L_{SegBox} \tag{4}$$

其中
$$L_{Cls} = - \frac{\sum_{i=1}^n y_i \lg Y_i}{n}$$

式中 L_{Cls} ——经典分类损失
 L_{Reg} 、 L_{Com} ——聚合项中的回归项和紧凑项，作用为使预测框和建议框接近真实框
 L_{SegGT} 、 L_{SegBox} ——排斥项，使预测框远离周围

的真实框和预测其他目标的
预测框

- y_i ——第 i 个预测框是否包含一个目标,取 0 或 1
- Y_i ——预测概率
- n ——预测总数
- $\alpha, \beta, \gamma, \delta$ ——权重

图 3 为 polyphyletic 损失值计算的 2 部分:用于将预测框回归到真实框的聚合项,以及用于使预测框远离周围目标真实框和预测框的排斥项。

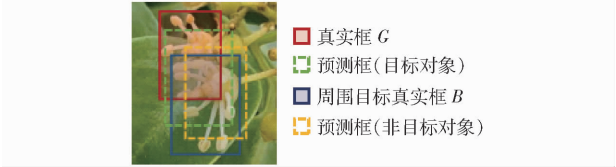


图 3 polyphyletic 损失值计算图

Fig.3 Calculating diagram of polyphyletic loss

设 $F = (x_F, y_F, w_F, h_F)$ 、 $P = (x_P, y_P, w_P, h_P)$ 和 $G = (x_G, y_G, w_G, h_G)$ 为预测框、建议框和真实框, w 和 h 分别为宽度和高度。 $F_+ = \{F_1, F_2, \dots, F_n\}$ 表示所有正样本预测框的集合(交并比(IoU)大于 0.5 的预测框), P_i 表示第 i 个预测目标的所有建议框的集合,其中 $i = 1, 2, \dots, n$ 。 $G_+ = \{G_1, G_2, \dots, G_n\}$ 表示所有真实框的集合。

2.1.1 聚合项

为了减小预测框与真实框之间的距离,常用的方法是欧几里得距离^[21]、Smooth_{L₁}距离^[22]或 IoU 损失^[23]。

Smooth_{L₁}距离是计算预测框和真实框之间的曼哈顿距离,不够精确,不足以计算两点之间的真实距离,而当预测框的面积远大于真实框的面积时,预测框的变化引起的 IoU 损失变化不大,预测框与真实框的重叠度不够准确,因此采用欧氏距离来计算回归损失值。回归项的定义为

$$L_{Reg} = \frac{\sum_{i=1}^{|F_+|} \|F_i - G_{Agg}^{F_i}\|_2^2}{|F_+|}$$

(5)

其中 $G_{Agg}^{F_i} = \operatorname{argmax}_{G \in G_+} IoU(G, F_i)$

(6)

式中 F_i ——计算回归损失的预测框
 $G_{Agg}^{F_i}$ ——预测 F_i 中最大 IoU 值的真实框
 G ——真实框

同时,考虑到靠近目标且存在遮挡的相邻对象对检测的影响,增加一个紧凑的函数项,拉近目标的所有建议框之间的距离,使所有建议框在目标周围紧凑,提高检测精度。 G_i 是集合 $G_+ = \{G_1, G_2, \dots, G_n\}$ 中预测的第 i 个目标的真实框。然后设置 $P_i =$

$\{\tilde{P}_1, \tilde{P}_2, \dots, \tilde{P}_m\}$ 为第 i 个预测目标的所有建议框的集合, $\tilde{P}_j$ 为集合中的第 j 个建议框。 L_{Com} 的计算公式为

$$L_{Com} = \frac{1}{|F_+|} \sum_{i=1}^{|F_+|} \Delta \left(G_i - \frac{\sum_{j \in P_i} \tilde{P}_j}{|P_i|} \right)$$

(7)

式中 $|P_i|$ ——预测第 i 个目标的所有建议框总数
 $|F_+|$ ——所有预测目标数量

$$\Delta \left(G_i - \frac{\sum_{j \in P_i} \tilde{P}_j}{|P_i|} \right)$$

——预测第 i 个目标的建议框的欧几里得距离平均损失值

该函数项计算了图像中每个目标的建议框的平均预测与相应的真实框之间的差值,并描述了建议框与真实框的紧密性。 α 和 β 是平衡 L_{Reg} 和 L_{Com} 项的超参数。

2.1.2 排斥项

当检测有遮挡的密集对象,预测框轻微或显著地移动到相邻的非目标真实对象或结合多个重叠的真实对象时,通常会出现密集检测误差。此外,密集检测误差通常具有相对较高的置信度,导致最高级别的假阳性,如图 4 所示。这表明,为了提高检测器对密集荔枝花检测的鲁棒性,边界框回归需要更多的损失识别。绿色框是正确的预测边界框,而红色框是由于遮挡引起的假阳性。检测器输出的置信度分数也被附加在图像中。当预测框靠近相邻的荔枝花目标或多个荔枝花目标重叠度较高时,通常只会检测为一个荔枝花目标。



图 4 密集荔枝花错误检测的可视化示例

Fig.4 Visual example of dense litchi flower error detection

除指定的目标外,预测框中 IoU 值最大的真值框 G_{Seg}^F ^[24] 被定义为

$$G_{Seg}^F = \operatorname{argmax}_{G \in G_+ \setminus \{G_{Agg}^{F_i}\}} IoU(G, F)$$

(8)

式中 F ——预测框

为了避免并集远大于交集的情况,本文选择真实框交集(IoG),计算公式为

$$I_{oG}(F, G) \triangleq \frac{\operatorname{area}(F \cap G)}{\operatorname{area}(G)}$$

(9)

式中 $I_{oG}(F, G)$ ——预测框与真实框重叠度占比
area——面积函数

为了对预测框 F 和其他目标的真实框 G_{Seg}^F 之间的重叠面积进行惩罚, L_{SegGT} 损失项设计为

$$L_{SegGT} = \frac{\sum_{i=1}^{|F_+|} \text{Smooth}_{\ln z}(\text{IoG}(F_i, G_{Seg}^F))}{|F_+|} \quad (10)$$

其中

$$\text{Smooth}_{\ln z} = \begin{cases} -\ln(1-z) & (z \leq \sigma) \\ \frac{z-\sigma}{1-\sigma} - \ln(1-\sigma) & (z > \sigma) \end{cases} \quad (11)$$

式中 $\text{Smooth}_{\ln z}$ ——光滑 $\ln$ 函数, 在 $(0,1)$ 中连续可微

σ ——平滑参数, 用于调整排斥损失对异常值的敏感性, $\sigma \in [0,1)$

预测框与非目标真实框重叠越多, 边界框回归器添加的惩罚就越多, L_{SegGT} 损失可以有效地阻止预测的边界框转移到非目标的邻近目标。

最后一个 L_{SegBox} 损失项通过排除具有不同指定目标的额外预测框来降低检测器对非极大抑制的灵敏度。 F_+ 被分成 $|G_+|$ 个不相交的子集, $F_+ = F_1 \cap F_2 \cap \dots \cap F_{|G_+|}$ 。对于从两个不同的子集随机抽样的两个方案 $F_i \in F_+ \setminus \{F_j\}$ 和 $F_j \in F_+ \setminus \{F_i\}$, 期望预测框 F_i 和 F_j 的重叠应该尽可能的小, 其中 $i, j = 1, 2, \dots, |G_+|$, 且 $i \neq j$ 。因此, L_{SegBox} 损失项函数表示为

$$L_{SegBox} = \sum_{i=1, i \neq j}^{|F_+|} \left(\frac{\sum_{j=1, i \neq j}^{|F_+|} \text{Smooth}_{\ln}(\text{IoU}(F_i, F_j))}{\sum_{j=1, i \neq j}^{|F_+|} \mathbb{I}[\text{IoU}(F_i, F_j) > 0] + \varepsilon} \right) \quad (12)$$

其中, ε 为一个小常数, 以避免分母为 0 时的误差, 当 $\text{IoU}(F_i, F_j) > 0$ 时, $\mathbb{I}$ 为恒等函数。从式(12)可以看出, 两个具有不同指定目标的预测框之间的 IoU 面积非常小才能最小化 L_{SegBox} 。这意味着在非极大抑制之后, L_{SegBox} 可能会降低具有单独回归目标的预测边界框合并成一个框的可能性, 使检测器对密集和遮挡的情况更具有鲁棒性。 γ 和 δ 是平衡 L_{SegGT} 和 L_{SegBox} 项的超参数。

2.2 漏检分析

使用文献[17]经典方法来测试不同遮挡程度的检测效果, 包括无遮挡、轻微遮挡和严重遮挡。从图 5 可以看出, 遮挡程度越大, 遗漏检测的比例越大。

通过降低 IoU 阈值(从 0.7 减少到 0.5)来减少遗漏检测的数量, 导致遗漏检测的比例更大(从 60.7% 增加到 67.3%)。这是因为在非极大抑制过程中, 由于作物的密集分布, 通过训练产生的候选框也非常密集。只调整 IoU 的阈值并不能产生较高可

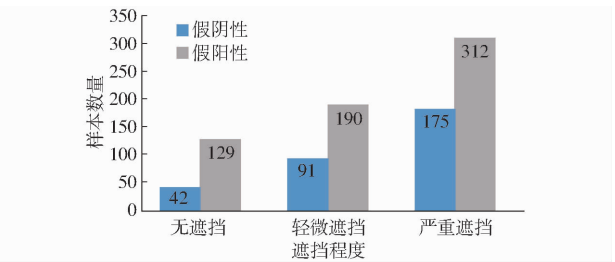


图 5 不同遮挡程度数据集的假阳性数和假阴性数
Fig. 5 False positive and false negative numbers of data sets with different occlusion levels (no occlusion, slight occlusion and severe occlusion)

信度的预测框。为了减少遮挡引起的漏检, 需要调整建议框的位置。本文提出的 polyphyletic 损失就是为此目的而设计。

3 结果与分析

本文实验硬件环境为: Intel i9 - 10900x (3.20 GHz) 10 核 20 线程 CPU, NVIDIA GeForce RTX 2080 Ti 1GB $\times$ 2。实验软件环境为 Windows 10 系统, CUDA 版本为 10.1, cuDNN 版本为 7.0、pycharm 版本为 2018.2。实验框架的前端和后端分别为 Keras 和 Pytorch。

3.1 不同算法的检测结果

选择缺失率 (Miss rate, MR)、召回率 (Recall)、精确率 (Precision)、F1 值和平均精度均值 (mAP) 作为评价指标。精确率为 5 次重复实验中精度最高的取值, 平均精度均值为 5 次重复实验的平均精度。

使用本文制作的荔枝花数据集对 Faster R-CNN^[17]、Focus Loss^[25]、AdaptiveNMS^[26] 和 Mask R-CNN^[18] 进行训练, 得到最优模型之后, 分别设置不同的 IoU 值在测试集上进行测试, 5 种方法的缺失率、召回率、F1 值、精确率和平均准确率如表 1 所示。

表 1 不同研究方法的检测结果						
Tab. 1 Comparison with other research methods						
方法	IoU	缺失率/%	召回率/%	F1 值/%	精确率/ 平均精度	均值/%
Faster R - CNN	0.7	40.03	59.97	61.16	66.91	62.39
	0.5	35.32	64.68	66.23	72.28	67.85
Focus Loss	0.7	34.79	65.21	67.54	70.17	70.04
	0.5	29.65	70.35	72.05	75.65	73.84
AdaptiveNMS	0.7	28.42	71.58	72.95	78.99	74.38
	0.5	22.81	77.19	78.37	82.64	79.59
Mask R - CNN	0.7	33.72	66.28	71.81	79.69	78.34
	0.5	29.02	70.98	76.55	83.46	83.08
本文方法	0.7	17.17	82.83	83.31	86.24	83.80
	0.5	13.29	86.21	87.07	89.93	87.94

使用 ResNet50 作为主干网络, Faster R - CNN 作为基准在检测密集遮挡物上性能较差, 识别精度仅为 60% ~ 70%。Focus Loss 在 Faster R - CNN 的基础上进行损失函数的改进, 降低正样本在损失中的权重, 增加对错误案例的惩罚, 提升了对遮挡物的检测精度, 在本研究的自建荔枝花数据集中 mAP 达到 73.84%。AdaptiveNMS 提出了密度预测模块, 自适应的调整非极大抑制时的阈值, 使得待检测物体重叠时尽可能保留正确的预测框, 数据集上的 mAP 达到 79.59%。Mask R - CNN 在 Faster R - CNN 的基础上添加了一个预测分割 mask 的分支, 将 RoI Pooling 层替换成了 RoIAlign 层, 并添加了并列的 FCN 层(mask 层), 在荔枝花数据集上的 mAP 达到 83.08%, 分别比 Faster R - CNN 和 Focus Loss 提高 15.23、9.24 个百分点, 在其他方法中的准确率最高。而本文方法的识别 mAP 为 87.94%, 比 Mask R - CNN 高约 4 个百分点, 召回率也提高了约 9 个百分点, 证明了本文方法在检测密集遮挡目标方面的有效性。同时将 Faster R - CNN 作为基准, 直观地比较了本文方法和 Mask R - CNN 的检测结果, 如图 6 所示。

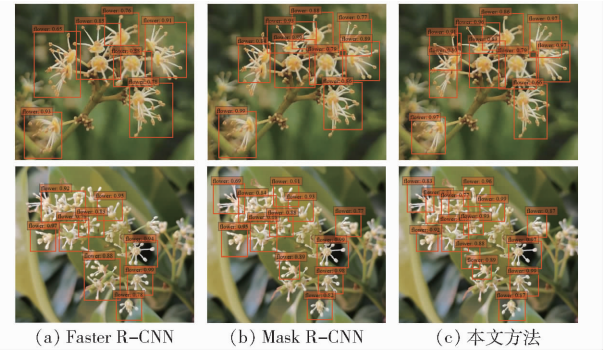


图 6 不同方法在荔枝花数据集上的检测结果
Fig.6 Detection results of different methods on litchi flower dataset

3.2 通用性证明

将本文方法与其他方法分别在苹果花数据集和荔枝花数据集上进行比较。由于苹果花数据集中的苹果花是整个苹果树的图像, 苹果花与苹果花之间的分布较荔枝花相对稀疏; 可以认为苹果花数据集是一般遮挡的对象, 而荔枝花数据集是严重遮挡的对象。在 IoU 为 0.5 的条件下, 以平均每幅图像中的假阳性数量为横轴, 缺失率为纵轴, 结果如图 7 所示。从实验结果可以看出, 本文方法在两个数据集上都表现出了良好的性能。在密集的荔枝花数据集上, 本文方法的缺失率为 2.36%, 这优于另外 4 种方法。本文方法在苹果花数据集上与 Mask R - CNN 相比, MR 仅有微小差异, 反映了本文方法在检

测一般遮挡对象时的鲁棒性。

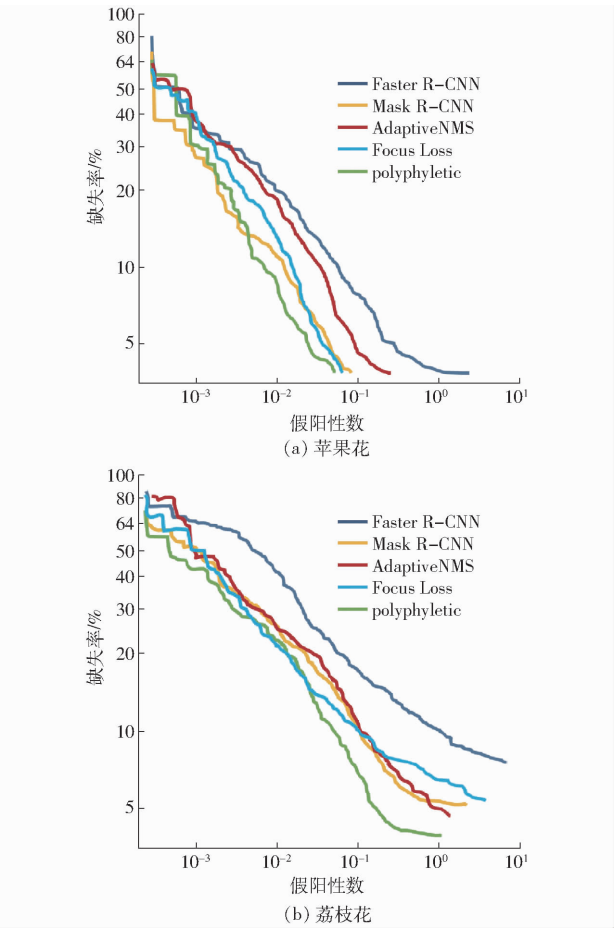


图 7 不同方法对苹果花数据集和荔枝花数据集处理效果比较
Fig.7 Apple flower data set and litchi flower data set studied by different most advanced methods

3.3 消融实验

在 polyphyletic 损失函数中去掉不同的项, 包括去除特征提取 (polyphyletic - F_e)、去除聚合项 (polyphyletic - Com) 和去除排斥项 (polyphyletic - Seg), 以检验每个部分的有效性。其中 F_e 、 Com 和 Seg 分别代表 polyphyletic 损失函数中的特征提取、聚合项和排斥项。表 2 中的实验结果表明, 删除了特征提取后检测精度略有下降, 在删除聚合项后, 检测准确率从 87.94% 下降到 83.76%, 降低 4.18 个百分点, 删除了排斥项后降低 9.37 个百分点。实验结果验证了在损失函数上改进的有效性。图 8 为在基本损失函数中添加一个聚合项 (+ L_{Com}) 前后的建议框的可视化比较, 通过添加聚合项, 比较经典损失函数和非极大抑制前的建议边界框。在经典损失函数 + L_{Com} 结果中, 两个相邻真实目标之间的建议框相交面积较小, 每个真实目标的建议框也更紧凑, 这在密集遮挡场景中是可取的。

3.4 超参数设置

由表 3 可知, 当 σ 为 1.0 时, 通过在合理的评估

表 2 消融实验结果

Tab.2 Ablation experiment results

损失函数	Fe	Com	Seg	召回率/%	mAP/%
polyphyletic _{-Seg}	✓	✓		77.35	78.57
polyphyletic _{-Com}	✓		✓	82.81	83.76
polyphyletic _{-Fe}		✓	✓	84.97	86.69
polyphyletic	✓	✓	✓	86.21	87.94

注:✓表示在经典损失函数中加入此损失项。



(a) 输入图像 (b) 经典损失函数 (c) 经典损失函数+Com

图 8 在经典损失函数中加入聚合项 (+ L_{Com}) 前后的建议框对比

Fig.8 Comparison of suggestion boxes before and after adding aggregate term (+ L_{Com}) to classical loss function

表 3 σ 取不同值时的 L_{SegGT} 和 L_{SegBox} 的缺失率

Tab.3 L_{SegGT} and L_{SegBox} results when σ took different values

排斥项	σ			%
	0	0.5	1.0	
L_{SegGT}	14.3	14.5	13.7	
L_{SegBox}	13.7	14.2	14.3	

设置中添加 L_{SegGT} , 缺失率为 13.7% ,达最佳性能^[17]。

在实验开始时,在荔枝花数据集的验证子集上通过调整损失函数的权重系数来平衡不同项的损失值,并使用缺失率来测量不同权重系数下的模型检测性能。如表 4、5 所示, $\alpha = 0.6$ 、 $\beta = 0.4$ 、 $\gamma = 0.5$ 和 $\delta = 0.5$ 时模型的检测效果最佳。

表 4 α 、 β 取不同值时的缺失率

Tab.4 Missing rate when α and β took different values

α	β	缺失率/%
0.3	0.7	14.1
0.4	0.6	13.9
0.5	0.5	13.3
0.6	0.4	13.2
0.7	0.3	13.9

表 5 γ 、 δ 取不同值时的缺失率

Tab.5 Missing rate when γ and δ took different values

γ	δ	缺失率/%
0.3	0.7	14.1
0.4	0.6	13.9
0.5	0.5	13.2
0.6	0.4	13.4
0.7	0.3	13.9

4 结论

(1) 设计了包含聚集能力和排斥能力的 polyphyletic 损失函数,首先在经典的损失函数中加入聚合项,迫使预测框接近并紧密定位相应的荔枝花目标,从而使预测目标的多个建议框紧密聚集。同时在损失函数中添加了两种类型的排斥项来优化模型,使目标预测框远离其他目标真实框,以减少不同目标之间预测框的重叠。

(2) 实验结果表明,本文设计的 polyphyletic 损失函数对密集荔枝花检测的平均精度均值为 87.94% ,F1 值为 87.07% ,缺失率为 13.29% ,召回率为 86.21% 。与 Faster R - CNN、Focus Loss、AdaptiveNMS、Mask R - CNN 相比,本文方法对密集荔枝花检测具有更高的准确率和更低的缺失率。

参 考 文 献

[1] PUNIA S, KUMAR M. Litchi (Litchi chinensis) seed: nutritional profile, bioactivities, and its industrial applications[J]. Trends in Food Science & Technology, 2021, 108: 58 – 70.

[2] 齐文娥,陈厚彬,罗滔,等. 中国大陆荔枝产业发展现状、趋势与对策[J]. 广东农业科学,2019,46(10):132 – 139. QI Wen'e, CHEN Houbin, LUO Tao, et al. Development status, trend and suggestion of litchi industry in mainland China[J]. Guangdong Agricultural Sciences, 2019,46(10):132 – 139. (in Chinese)

[3] 吴辉,邹华容. 荔枝高产稳产栽培技术[J]. 中国农业文摘-农业工程,2021,33(4):14 – 16.

[4] NILSBACK M E, ZISSERMAN A. Delving into the whorl of flower segmentation[C] // BMVC, 2007: 1 – 10.

[5] CHAI Y, LEMPITSKY V, ZISSERMAN A. Bicos: a bi-level co-segmentation method for image classification[C] // 2011 International Conference on Computer Vision. IEEE, 2011: 2579 – 2586.

[6] 熊俊涛,刘柏林,钟灼,等. 基于深度语义分割网络的荔枝花叶分割与识别[J]. 农业机械学报,2021,52(6):252 – 258. XIONG Juntao, LIU Bolin, ZHONG Zhuo, et al. Litchi flower and leaf segmentation and recognition based on deep semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(6): 252 – 258. (in Chinese)

[7] KARAMI E, PRASAD S, SHEHATA M. Image matching using SIFT, SURF, BRIEF and ORB; performance comparison for distorted images[J]. arXiv preprint, arXiv:1710.02726, 2017.

[8] RASHIDI S, ESFAHANI J A, ELLAHI R. Convective heat transfer and particle motion in an obstructed duct with two side by

side obstacles by means of DPM model[J]. Applied Sciences, 2017, 7(4): 431.

- [9] LIU Y, LIU S, WANG Z. Multi-focus image fusion with dense SIFT[J]. Information Fusion, 2015, 23: 139 – 155.
- [10] NAGARAJA S, PRABHAKAR C J. Low-level features for image retrieval based on extraction of directional binary patterns and its oriented gradients histogram[J]. Computer Applications: an International Journal (CAIJ), 2015, 2(1): 13 – 28.
- [11] MA S, BAI L. A face detection algorithm based on Adaboost and new Haar-Like feature[C]//2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS). IEEE, 2016: 651 – 654.
- [12] 林海波, 卢元栋, 丁荣诚, 等. 基于图像处理与改进 SVM 的苹果多特征融合分级方法[J]. 山东农业科学, 2022, 54(6): 141 – 149.
LIN Haibo, LU Yuandong, DING Rongcheng, et al. A multi-feature fusion classification method for apple based on image processing and improved SVM[J]. Shandong Agricultural Sciences, 2022, 54(6): 141 – 149. (in Chinese)
- [13] 修春波, 孙乐乐. 基于改进 YOLO v4 网络的马铃薯自动育苗叶芽检测方法[J]. 农业机械学报, 2022, 53(6): 265 – 273.
XIU Chunbo, SUN Lele. Potato leaf bud detection method based on improved YOLO v4 network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(6): 265 – 273. (in Chinese)
- [14] 程海超. 基于 YOLO 的番茄病虫害识别算法研究[J]. 农业与技术, 2022, 42(10): 38 – 40.
- [15] TAN L, HUANGFU T, WU L, et al. Comparison of RetinaNet, SSD, and YOLO v3 for real-time pill identification[J]. BMC Medical Informatics and Decision Making, 2021, 21: 1 – 11.
- [16] TAN M, PANG R, LE Q V. EfficientDet: scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781 – 10790.
- [17] 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.
- [18] HE K, GKIOXARI G, DOLLAR R P, et al. Mask R – CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961 – 2969.
- [19] DIAS P A, TABB A, MEDEIROS H. Multispecies fruit flower detection using a refined semantic segmentation network[J]. IEEE Robotics and Automation Letters, 2018, 3(4): 3003 – 3010.
- [20] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770 – 778.
- [21] 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, 2014: 580 – 587.
- [22] GIRSHICK R. Fast R – CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440 – 1448.
- [23] YU J, JIANG Y, WANG Z, et al. Unitbox: an advanced object detection network[C]//Proceedings of the 24th ACM International Conference on Multimedia, 2016: 516 – 520.
- [24] WANG X, XIAO T, JIANG Y, et al. Repulsion loss: detecting pedestrians in a crowd[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7774 – 7783.
- [25] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980 – 2988.
- [26] LIU S, HUANG D, WANG Y. Adaptive NMS: refining pedestrian detection in a crowd[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 6459 – 6468.