

基于 YOLO - DCL 的复杂环境油茶果遮挡检测与计数研究

肖伸平^{1,2} 赵倩颖^{2,3} 曾甲元^{2,3} 彭自然^{1,2}

(1. 湖南工业大学电气与信息工程学院, 株洲 412007; 2. 电传动控制与智能装备湖南省重点实验室, 株洲 412007;
3. 湖南工业大学轨道交通学院, 株洲 412007)

摘要: 为解决复杂环境中油茶果因遮挡造成的检测与计数难题, 提出了一种基于双主干网络 (Dual-backbone) 和连续注意力特征融合模块 (Consecutive attention feature fusion, CAFF) 的检测模型。该模型结合了两种不同主干网络的优势, 实现了对不同特征的高效提取。此外, 设计了双输入单输出的连续注意力特征融合模块, 取代了传统的拼接操作 (Concat), 优化了多尺度特征信息的融合策略。为了在精度与模型内存占用量之间取得平衡, 采用了幻影卷积模块 (Ghostconv), 并去除了空间金字塔池化层 (Spatial pyramid pooling fast, SPPF), 加快了训练速度, 减少了参数量。改进后的 YOLO - DCL (YOLO dual-backbone & consecutive attention feature fusion & lightweight) 模型在各类遮挡检测任务上表现优秀, 平均精度均值达到 92.7%, 精确率为 90.7%, 召回率为 84.9%, 而模型内存占用量仅为 5.7 MB。相较 YOLO v8n 模型分别上升 4.0、8.6、2.3 个百分点, 内存占用量下降 9.5%。该模型还具备油茶果遮挡类别的自动计数功能, 可降低人工统计的劳动成本, 适合在野外复杂环境中部署应用。

关键词: 油茶果; 遮挡识别; YOLO v8n; 计数; 双主干网络; 特征融合

中图分类号: TP391.4; S24

文献标识码: A

文章编号: 1000-1298(2024)10-0318-09

OSID:



Camellia oleifera Fruits Occlusion Detection and Counting in Complex Environments Based on Improved YOLO - DCL

XIAO Shenping^{1,2} ZHAO Qianying^{2,3} ZENG Jiayuan^{2,3} PENG Ziran^{1,2}

(1. School of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou 412007, China

2. Key Laboratory for Electric Drive Control and Intelligent Equipment of Hunan Province, Zhuzhou 412007, China

3. School of Railway Transportation, Hunan University of Technology, Zhuzhou 412007, China)

Abstract: To solve the challenges of detecting and counting *Camellia oleifera* fruits with multiple occlusions in complex environments, a detection model was proposed based on a dual-backbone network and a consecutive attention feature fusion module (CAFF). The dual-backbone network combined the advantages of two different backbone networks to achieve efficient extraction of different features. In addition, a dual-input single-output CAFF module was designed. This CAFF module replaced the traditional concat operation and optimizes the fusion strategy for multi-scale feature information. In order to strike a balance between model precision and size, the ghost convolution (Ghostconv) module was used, and the spatial pyramid pooling fast (SPPF) layer was removed. It accelerated training time and reduced the number of parameters. The improved YOLO dual-backbone & consecutive attention feature fusion & lightweight (YOLO - DCL) model performed well on all kinds of occlusion detection tasks, with a mean average precision (mAP) of 92.7%, precision of 90.7%, and recall of 84.9%, while the model size was only 5.7 MB. Compared with the YOLO v8n model, it increased 4.0 percentage points of mAP, 8.6 percentage points of precision, and 2.3 percentage points of recall. At the same time, the model size was decreased by 9.5%. Besides, the model incorporated the ability to automatically count *Camellia oleifera* fruits with occlusion categories, which can reduce labor costs and improve the accuracy

收稿日期: 2024-06-28 修回日期: 2024-07-20

基金项目: 国家重点研发计划项目 (2019YFE0122600)、湖南省教育厅重点科研项目 (22A0423) 和湖南省自然科学基金项目 (2023JJ60267、2022JJ50073)

作者简介: 肖伸平 (1965—), 男, 教授, 博士生导师, 主要从事人工智能与机器人控制、鲁棒控制理论及应用研究, E-mail: xsp@hut.edu.cn

of yield estimation. It was very suitable for deployment in complex environments.

Key words: *Camellia oleifera* fruits; occlusion detection; YOLO v8n; count; dual-backbone network; feature fusion

0 引言

油茶果是我国重要的油料作物,不仅营养价值高,而且具有独特的药用价值,被誉为“东方橄榄油”^[1-2]。然而,目前油茶果的采摘主要依赖于传统的人工方式^[3-5],这一过程不仅劳动强度大、效率低,还伴随着一定的安全风险^[6-7]。尽管已有基于振动或拍打原理的机械化采摘方式被提出,但这些方法往往会对油茶果的花苞造成损伤,影响下一年度的产量^[8-10]。因此,利用机器视觉技术辅助采摘成为当前研究的重点方向。

但是,在自然环境中油茶果的生长状况极为复杂,果实常被周围的枝叶所遮挡^[11-13]。此外,光照的变化也为识别和分类带来了挑战^[14-16]。为了解决这一问题,国内外学者进行了大量的研究,并取得了一定的进展。主要研究方法分为两大类:传统的图像检测技术和现代的深度学习技术。传统的图像检测技术主要依赖于颜色^[17]、纹理^[18-19]和边缘^[20]等特征进行目标检测。这些方法在特定条件下能够实现快速且有效的目标识别,但在复杂的油茶果遮挡环境下,这些方法的泛化性不足,难以满足油茶果的识别需求。现代的深度学习方法,尤其是卷积神经网络,因其强大的特征提取能力和面对复杂数据的识别能力,使其在图像识别和分类任务中展现出了卓越的优势^[21-25]。但以上研究均将油茶果视为单一的检测目标,未根据果实的遮挡状况进行细致分类,这在一定程度上限制了其在采摘机器人中的应用。

本文以油茶果为研究对象,基于 YOLO v8n 模型,提出 YOLO - DCL 检测与计数方法。该方法采用 YOLO v8n 原主干和轻量化卷积网络(ShuffleNet)构建双主干特征提取网络,并提出连续注意力特征融合模块,代替传统的拼接操作,从而优化多尺度信息的融合。通过添加幻影卷积,去除空间金字塔池化层,使模型在保持准确性的同时实现轻量化。针对实际环境中的遮挡问题,对自制数据集进行细致分类,包括无遮挡(No occlusion, NO)、树叶遮挡(Leaves occlusion, LO)、树枝遮挡(Branches and trunks occlusion, BTO)和混合遮挡(Hybrid occlusion, HO)4 种类型。通过 YOLO - DCL 模型检测油茶果的遮挡类型,并进行自动计数,以期为采摘机器人提供更精确的抓取策略。

1 数据集构建与识别方法

1.1 图像获取

本研究所用的数据集采自湖南省株洲市醴陵油茶果基地,油茶果颜色主要有红色、黄色、青色 3 种,形状多为圆形、橄榄形或椭球形。本研究使用佳能 EOS - R50 相机进行图像获取并利用便携式计算机对数据集图像进行监控,如图 1 所示。为了确保数据多样性,从多角度、多方位采集晴天和阴天、不同时段、不同密集程度、不同遮挡强度的油茶果图像、共采集 2 513 幅,分辨率为 1 280 像素 × 1 280 像素,如图 2 所示。



图 1 油茶果数据采集过程

Fig. 1 Data collection process of *Camellia oleifera* fruits

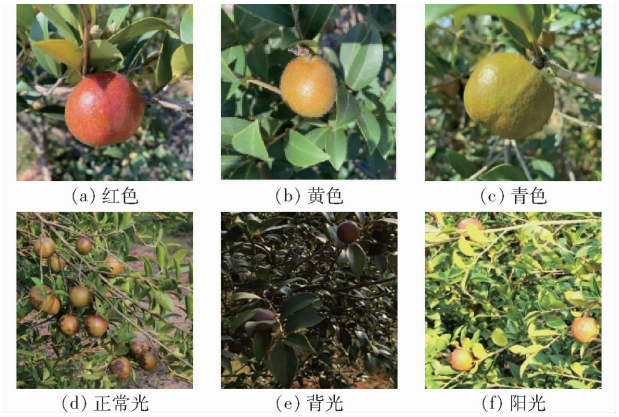


图 2 油茶果样本示例

Fig. 2 Examples of *Camellia oleifera* fruits

1.2 数据处理

从 2 513 幅图像中剔除重复、拖影以及无果实的图像后,将剩余的 1 205 幅图像作为原始数据集。为增强模型对果实角度和光照变化的鲁棒性,以及提升其在复杂条件下的泛化性^[26-27],对所选图像进行了数据增强处理。通过添加椒盐噪声和高斯噪声,模仿运动抖动导致的图像模糊;进行仿射、平移、翻转等变换,模仿相机在一瞬间捕捉到的不

同角度图像。经过数据增强后共获得 2 865 幅图像,使用 Labellmg 工具对图像进行标注。为了指导后续机器人的采摘工作并有效解决识别过程中的遮挡问题,根据遮挡种类将油茶果分为 4 类:无遮挡(NO)、树叶遮挡(LO)、树枝遮挡(BTO)和混合遮挡(HO),如图 3 所示。标注后的数据集以比例 6:2:2 划分为训练集、验证集、测试集,各类别果实数量如表 1 所示。



图 3 油茶果遮挡分类

表 1 各类别果实数量
Tab.1 Specific quantities of fruit by category

果实类别	训练集	验证集	测试集	总计
无遮挡	2 066	582	554	3 202
树叶遮挡	2 103	534	557	3 194
树枝遮挡	305	82	77	464
混合遮挡	1 169	273	316	1 758
总计	5 643	1 471	1 504	8 618

1.3 YOLO – DCL 模型

本文提出了一种轻量化的 YOLO – DCL 模型。该模型采用双主干特征提取网络,通过并行处理来自不同主干的特征,增强了模型对遮挡和复杂背景的适应性。此外,设计了一个连续注意力特征融合模块,替代了传统的特征拼接操作。改进后的模型能够更精准地融合油茶果的关键特征,不仅适用于油茶果的检测,还能用于油茶果计数任务。模型的具体结构如图 4 所示。图中 Ghostconv 为幻影卷积,

C2f 为跨阶段连接模块,CAFF 为连续注意力特征融合模块,ShuffleNet 为轻量化卷积网络,CBRM 为预处理模块。

1.3.1 双主干网络

借鉴多模态数据输入的理念,提出了双主干网络结构,对单一 RGB 图像进行双通道特征提取,而无需改变原始输入图像。这种方法与现有的单一重复双主干网络不同,其采用了不同的主干网络有效整合来自不同通道的信息,克服数据提取过程中的难题。具体而言,本文选用 YOLO v8n 原始主干与轻量化卷积网络^[28] (ShuffleNet) 构成双主干网络。其中 YOLO v8n 原始主干网络使用 C2f 模块^[29] 替代了 YOLO v5 模型主干网络中的 C3 模块^[30]。这一改进不仅使网络结构更加轻量化,而且继承了 CSP Darknet 网络^[31] 的核心思想,在维持高性能的同时,还能够获得更加丰富的梯度信息。

第 2 条主干网络由轻量化卷积网络 (ShuffleNet) 和预处理模块 (CBRM) 构成,其具体结构如图 5 所示。网络首先利用 CBRM 模块进行上下文特征提取和双向关系建模。上下文特征提取主要依赖于卷积层 (Conv) 和全局池化层 (Maxpool),以收集目标周围的上下文信息。双向关系模型在处理特征图时动态地调整注意力,使其集中于不同的空间位置或通道,从而捕捉目标间的相互作用。如图 4 所示,网络串联 6 个 ShuffleNet 模块。ShuffleNet 模块通过轻量级的通道重排操作 (Channel split),将输入特征图按通道数分组并随机排列,打破通道间的信息隔离,增强特征图之间的信息交流。在步长为 1 的 ShuffleNet 模块中,主分支通过卷积层、深度可分离卷积 (Depthwise separable convolution, DWConv) 进行特征提取,而侧分支则将通道分割后

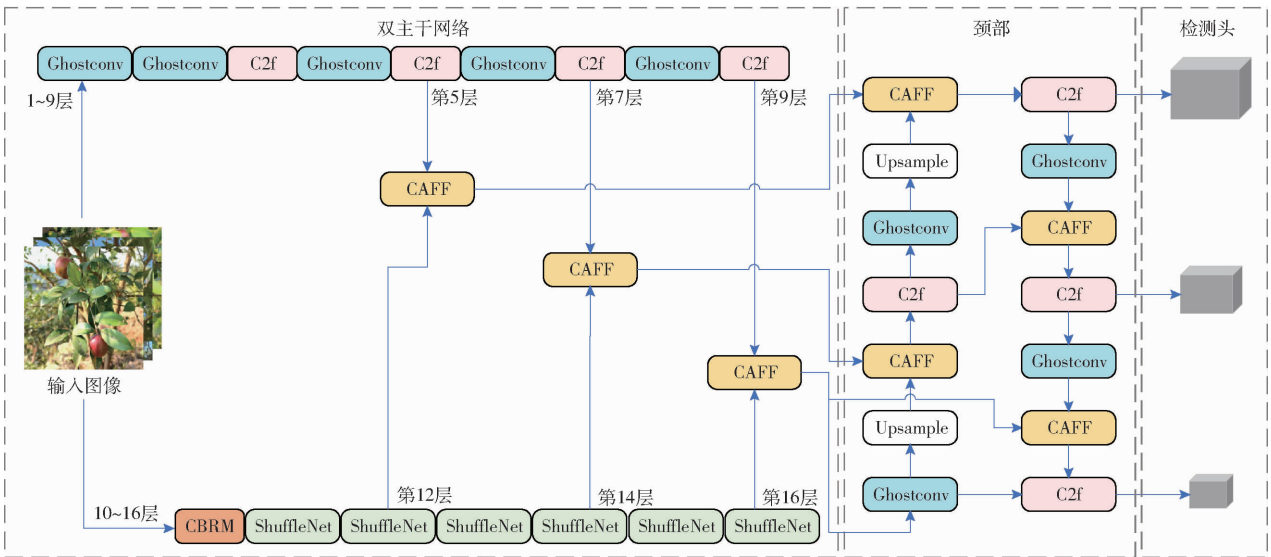


图 4 YOLO – DCL 模型网络结构图

Fig.4 YOLO – DCL network structure diagram

的特征图与主分支的特征图直接进行拼接。对于步长为 2 的 ShuffleNet 模块,主分支结构与步长为 1 的 ShuffleNet 模块相同,而侧分支则由 3×3 深度可分离卷积和 1×1 普通卷积构成。最终,通过该双主干网络,可以减少单一重复双主干网络在特征提取时存在的特征冗余现象,进一步提取图像的不同细粒度特征,促进信息在网络中的充分融合和传递,实现特征互补,从而适应油茶果采摘过程的多变场景。

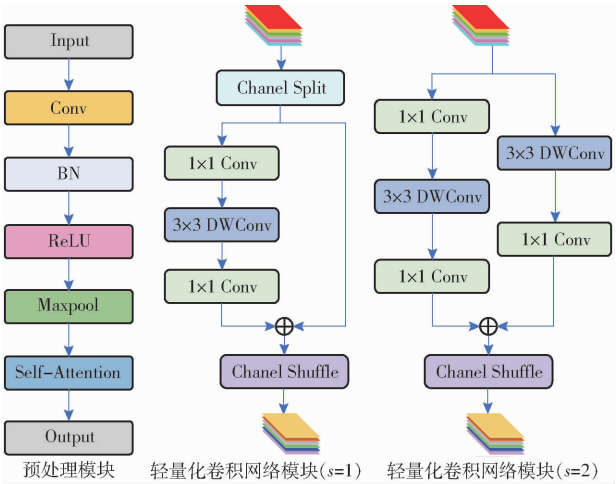


图 5 轻量化卷积网络模块和预处理模块
Fig. 5 ShuffleNet module and CBRM module

1.3.2 连续注意力特征融合模块

为了更充分地融合来自双主干网络的多通道数据,使用特征融合是必要的。然而,传统的拼接操作仅是将输入进行通道方向的简单叠加或者串联,这对于双通道的数据融合并不理想。因此,本文提出了一种双输入单输出的特征融合模块,由于其具有注意力机制的特性,可进行特征连续优化,将其命名为连续注意力特征融合模块 (Consecutive attention feature fusion, CAFF),如图 6 所示。

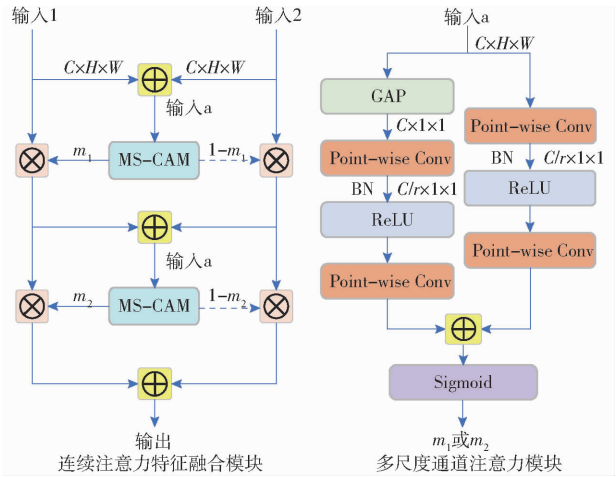


图 6 连续注意力特征融合模块
Fig. 6 Consecutive attention feature fusion module

将双主干网络各自提取的特征分别作为输入 1 与输入 2 传入 CAFF 模块,将两者相加后的结果传

入多尺度通道注意力模块 (Multi-scale channel attention module, MS - CAM) 进行特征融合。其中, C 、 W 、 H 分别表示特征图的通道数、宽度、高度, r 为通道缩减率。由图 6 可知, MS - CAM 模块通过 2 个不同的分支来提取特征信息,其中一个分支使用全局平均池化层 (Global average pooling, GAP) 提取输入图的全局特征,然后再使用逐点卷积 (Point-wise Conv)。另一个分支直接使用 Point-wise Conv 提取输入图的局部特征,然后通过批标准化 (Batch normalization, BN) 对张量进行统一,最后再次使用 Point-wise Conv,将两个维度的数组依次叠加,并将结果输入 Sigmoid 激活函数中。使用 MS - CAM 模块提取的特征 m_1 、 m_2 与输入 1、输入 2 进行相乘,然后再次叠加输入到 MS - CAM 模块中,得到最终的特征图细节信息,实现连续注意力特征融合,提高特征图的整合能力。

在 YOLO - DCL 网络结构中 (图 4), CAFF 模块将第 5 层和第 12 层、第 7 层和第 14 层、第 9 层和第 16 层提取的特征作为输入,代替 Concat 操作进行两两融合,捕捉图像中更丰富的信息,实现多尺度特征互补。这对于检测大小不同和遮挡程度不同的油茶果非常有用,因为较小或遮挡严重的目标可能在高分辨率的特征图中更容易检测,而较大或遮挡轻微的目标可能在低分辨率的特征图中更容易检测,从而提高模型在复杂场景下的性能。

1.3.3 网络轻量化重建

为了缓解双主干网络和连续注意力特征融合模块带来的参数增加和复杂度提升问题,本文对模型进行了轻量化设计。在模型的主干和颈部部分,传统的标准卷积操作虽然有效,但计算成本较高。因此,本文采用一种由标准卷积 (Standard convolution, SC)、混洗单元 (Shuffle Unit) 以及深度可分离卷积 (Depthwise separable convolution, DWConv) 相结合的幻影卷积 (Ghostconv) 来替代标准卷积,该卷积可有效降低运算量,具体结构如图 7 所示。首先,通过 1×1 的标准卷积,对输入图像进行通道数的压缩,生成初步的特征映射。接着,通过 DWConv 进一步减少计算量,生成新的特征图。新生成的特征图与初步的特征映射进行拼接后,通过 Shuffle Unit 操作对来自不同输入通道的数据进行混洗,增强 Ghostconv 的信息捕获能力。

此外,考虑到连续注意力特征融合模块已经引入多尺度信息并实现了特征的有效融合,本文移除原有的空间金字塔池化层 (Spatial pyramid pooling fast, SPPF),以降低结构复杂度。这一轻量化设计不仅保持了较低的数量,而且提升了模型捕获和

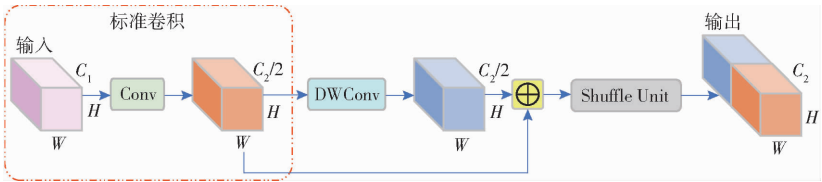


图 7 幻影卷积模块

Fig.7 Ghost convolution module

融合信息的能力,提高了油茶果识别整体运算效率和性能。

1.4 网络模型训练

1.4.1 环境配置

本研究中所有模型的训练与测试均在同一 GPU 服务器上执行。服务器操作系统为 Ubuntu 20.04,搭载 NVIDIA GeForce RTX 2080 Ti 图形处理单元 (GPU) 以及 Intel(R) Xeon(R) Gold 6430 中央处理器 (CPU)。所有程序均在 PyTorch 1.10 深度学习框架下开发,采用 Python 3.8 语言编写,并通过 CUDA 11.1 进行训练过程的加速。

1.4.2 训练策略

为了提升模型性能,通过反复的参数调整和测试对模型训练过程进行细致的优化。在处理输入图像时,统一将其尺寸调整为 640 像素 × 640 像素。使用随机梯度下降方法对部分训练样本进行参数更新,加快训练速度并降低计算复杂度。将初始学习率设为 0.1,权重衰减系数设为 0.000 5,训练次数设为 100 次。最大学习率调整至 1×10^{-3} ,批量大小设为 16。对主干网络使用冻结再解冻的训练方式,在训练的前 25 轮中,主干网络的部分层被冻结,随后逐渐解冻这些层,以促进模型对特征的逐步学习。在训练的最后 10 轮,禁用了 Mosaic 数据增强技术,以加快模型的收敛速度并稳定训练过程。

2 结果与分析

2.1 双主干网络设计改进结果

为了证明双主干网络的优势,进行了一系列消融实验,结果如表 2 所示。分析数据表明,将不同的双主干网络单独应用于 YOLO v8n 模型形成 YOLO-D 模型时,其平均精度均值 (mAP)、精确率 (Precision)、召回率 (Recall) 分别提升 2.2、5.2、3.0 个百分点。当将 YOLO v8n 模型的 C2f 主干简单复制形成 YOLO-C2f 双主干网络时,虽然相较 YOLO v8n 模型精确率有所提升,但平均精度均值和召回率却有所下降,且内存占用量增加。相比之下,本文提出的双主干网络采用两个不同的特征提取网络来执行特征提取任务,有效避免了单一主干在特征提

取时可能出现的相似性问题,减少了信息冗余,通过强化特征之间的互补和融合,提升了油茶果目标检测的综合性能。

表 2 单主干与双主干检测效果对比

Tab.2 Comparison of single backbone and dual-backbone detection effects

模型	平均精度	精确率/	召回率/	内存占用
	均值/%	%	%	
YOLO v8n	88.7	82.1	82.6	6.3
YOLO-C2f	88.5	86.6	80.6	9.1
YOLO-D	90.9	87.3	85.6	6.5

2.2 连续注意力特征融合优化结果

为了验证本文提出的连续注意力特征融合模块的有效性,该模块被单独加入网络中替代拼接操作,并命名为 YOLO-C 模型。根据表 3 可以看出,连续注意力特征融合模块的单独应用对模型的各项评价指标带来了提升,尤其是召回率,增长 3.7 个百分点。进一步地,在双主干网络的基础上结合连续注意力特征融合模块,对模型进行优化,并命名为 YOLO-DC 模型。YOLO-DC 模型与 YOLO v8n 模型内存占用量相当,但平均精度均值、精确率和召回率分别增长 2.5、6.2、0.2 个百分点。尽管召回率相较于单独应用连续注意力特征融合的 YOLO-C 模型略有下降,但 YOLO-DC 模型在各项评价指标上取得了更好的平衡,表现出了更加稳健的整体性。这表明将连续注意力特征融合模块与双主干网络相结合时,模型能够有效地提取和融合特征,增强对果实的关注度。

表 3 连续注意力特征融合优化结果

Tab.3 Optimization results of consecutive attention feature fusion model

模型	平均精度	精确率/	召回率/	内存占用
	均值/%	%	%	
YOLO v8n	88.7	82.1	82.6	6.3
YOLO-C	90.9	84.7	86.3	7.1
YOLO-DC	91.2	88.3	82.8	6.6

2.3 轻量化网络设计改进结果

由上述结果可以发现,使用双主干网络和连续注意力特征融合模块后,模型大小有所增加,因此,

本文对网络结构进行了轻量化设计。由表 4 可知,轻量化后的 YOLO - DCL 模型在内存占用量上,相比 YOLO v8n 模型减少 9.5%, 相较 YOLO - DC 模型下降 13.6%。此外, YOLO - DCL 模型平均精度均值、精确率、召回率相较 YOLO v8n 和 YOLO - DC 模型分别上升 4.0、8.6、2.3 个百分点和 1.5、2.4、2.1 个百分点。这些数据表明,3 项改进措施的结合对各类指标都起到了很大的改善效果,增强了对遮挡目标的识别能力。

表 4 轻量化设计对比结果

Tab. 4 Lightweight design comparison results

模型	平均精度 均值/%	精确率/ %	召回率/ %	内存占用 量/MB
YOLO v8n	88.7	82.1	82.6	6.3
YOLO - DC	91.2	88.3	82.8	6.6
YOLO - DCL	92.7	90.7	84.9	5.7

为了从定性方面验证 3 项改进对模型整体性能的影响,采用 Grad - CAM (Gradient-weighted class activation mapping) 技术对 YOLO - DCL 模型进行评估。改进前后模型对油茶果识别热力图如图 8 所示。热力图颜色深度反映模型对该局部区域的关注程度,其中红色区域表示关注程度最高。通过对比热力图可以发现,改进前的模型容易受到树枝树干等背景信息干扰,导致关注区域杂乱且分散。而本文设计的 YOLO - DCL 模型将主要注意力集中在油茶果果实部分,减少了对背景等干扰信息的关注。因此,本文所设计的 YOLO - DCL 模型可以在很大程度上缓解遮挡对检测的影响,使其在实际应用中具有更高的实用性和可靠性。

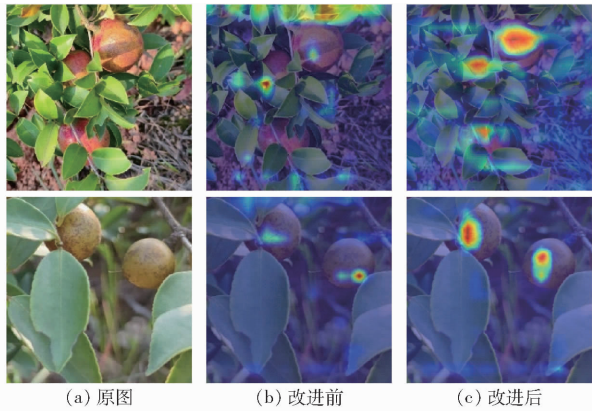


图 8 改进前后模型热力图

Fig. 8 Grad - CAM of model before and after improvement

2.4 YOLO - DCL 模型性能

本文提出的 YOLO - DCL 模型在训练过程中表现出色,特别是在损失曲线和精确率-召回率曲线方面,如图 9 所示。损失曲线显示出快速下降的趋势,并在大约迭代 60 次后开始趋于收敛,最

终稳定在 0.5。这表明该模型具有高效的学习能力,能够快速适应训练数据。此外,精确率-召回率曲线表明 YOLO - DCL 模型在精确率、召回率之间具有良好的平衡性,能够有效地识别无遮挡、树叶遮挡、树枝遮挡以及混合遮挡 4 类情况下的果实。由表 5 可知, YOLO - DCL 模型对这 4 类遮挡果实平均精度均值分别达到 98.3%、94.4%、90.7% 和 87.3%。混合遮挡类别平均精度均值略低于其它类别,这是由于受到多种复杂遮挡的影响,增加了特征提取的难度。但随着采摘的进行,遮挡程度会发生变化,因此这并不影响模型整体识别性能。

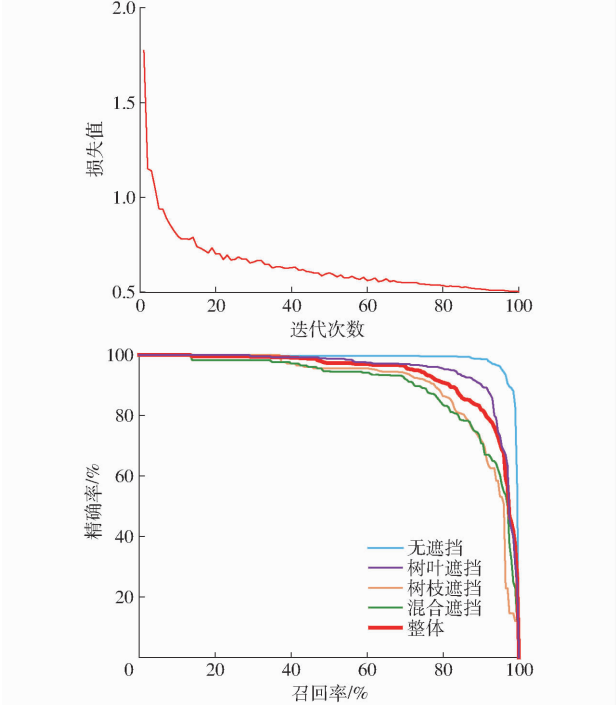


图 9 损失曲线和精确率-召回率曲线

Fig. 9 Loss curve and precision - recall curves

表 5 分类评价指标结果

Tab. 5 Classification evaluation indicator results

果实类别	实际果实 数/个	正确检测 数/个	平均精度 均值/%	精确率/ %	召回率/ %
无遮挡	582	541	98.3	93.0	96.6
树叶遮挡	534	470	94.4	88.1	89.9
树枝遮挡	82	70	90.7	85.1	83.8
混合遮挡	273	238	87.3	84.0	75.5

2.5 不同模型对比实验

为了全面评估改进后的 YOLO - DCL 模型,将其与当前主流的目标检测模型在同一数据集上进行对比,实验结果如表 6 所示。在检测精度方面, YOLO - DCL 模型在平均精度均值和精确率上表现最佳,分别达到 92.7% 和 90.7%。尽管在召回率上, YOLO - DCL 模型略低于 Faster R - CNN 模型,

但 Faster R - CNN 模型的计算复杂性较高,需要更高性能的 GPU,并且在召回率和精确率之间难以实现平衡。通过计算 F1 值可以综合评估模型的整体性能,YOLO - DCL 模型的 F1 值为 87.7%,分别比 YOLO v8n、YOLO v7 - tiny、YOLO v5s、YOLO r50 - csp、SSD、Faster R - CNN 模型高出 5.4、7.7、5.6、14.6、16.4、4.5 个百分点。这一结果表明,YOLO - DCL 模型在油茶果遮挡识别场景中,能够提供良好的整体性能,是完成识别任务的优选模型。

表 6 不同模型检测效果对比

Tab.6 Comparison of detection effect of different models

模型	平均精度	精确	召回率/ 内存占用		F1 值/
	均值/%	率/%	%	量/MB	%
YOLO v8n	88.7	82.1	82.6	6.3	82.3
YOLO v7 - tiny	85.7	76.7	83.6	12.3	80.0
YOLO v5s	88.5	83.0	81.2	5.3	82.1
YOLO r50 - csp	71.0	73.7	72.5	68.3	73.1
SSD	83.4	88.1	59.9	91.2	71.3
Faster R - CNN	81.1	76.3	91.4	108.2	83.2
YOLO - DCL	92.7	90.7	84.9	5.7	87.7

2.6 模型实际应用验证

为了验证模型在不同遮挡状态下的实际检测效果,开展了一系列实验。图 10 为不同光照和环境因素影响下,各目标检测模型的性能表现。在光照充足的明亮环境中,所有参与比较的模型均表现出色,未出现漏检或错检现象。其中,改进后的 YOLO - DCL 模型置信度最高,分别达到 0.92、0.90 和 0.83。在昏暗光线下,YOLO - DCL 模型保持了较高的检测效果,平均置信度达 0.89。相比之下,YOLO v5s 和 YOLO v7 - tiny 模型则出现了漏检或错检现象。例如,YOLO v5s 模型未能识别出图像左下角被前方油茶果遮挡的小目标,而 YOLO v7 - tiny 模型不仅错误地将树叶遮挡(LO)的果实检测为混合遮挡(HO),同样未能识别左下角的小目标果实。在稀疏环境下,改进后的 YOLO - DCL 模型由于添加了双主干网络和 CAFF 模块,增强了特征提取能力,使其置信度超过了其它模型。而在相同情况下,SSD 模型则偶尔出现漏检现象。在密集环境



图 10 不同场景下检测效果对比

Fig.10 Comparisons of detection effect in different scenes

下,遮挡现象严重,主流的目标检测模型出现了较多的漏检或错检。例如:YOLO v7 - tiny 模型发生了 3 次错检和 1 次漏检,SSD 模型则出现了 3 次漏检和 2 次错检,这些错误多发生在混合遮挡(HO)的果实上,这说明遮挡对目标检测精度的影响较大。但本文提出的 YOLO - DCL 模型在检测时,由于不同尺度的特征能够互相补充和增强,提升了模型对小物体和遮挡物体的检测能力,因而较少出现漏检或错检现象。综上所述,YOLO - DCL 模型在泛化能力方面表现出色,特别是在各类遮挡的油茶果检测任务中。

2.7 计数可视化结果

为了提高油茶果目标检测的效率,增加了计数

功能。通过精确计数,可以更好地进行数据分析,了解油茶果的生产情况,减少人工计数所需的时间和劳动强度。在分析过程中,随机抽取验证集中的数据样本进行处理,计数结果清晰地显示在每幅图像的左上角。例如,图 11 中,正常情况下,模型能准确识别出 1 个混合遮挡(HO)果实、1 个无遮挡(NO)果实和 3 个树叶遮挡(LO)果实;而在添加高斯噪声的条件下,模型检测到含有 6 个树叶遮挡(LO)果实。综合所有可视化结果可以观察到,无论在何种环境条件下,改进后的 YOLO - DCL 模型都能够有效识别小型目标、近景目标,以及在各种遮挡环境下的油茶果,提升了目标检测的准确性,减少了其它模型常见的漏检或错检现象,具有良好的鲁棒性。

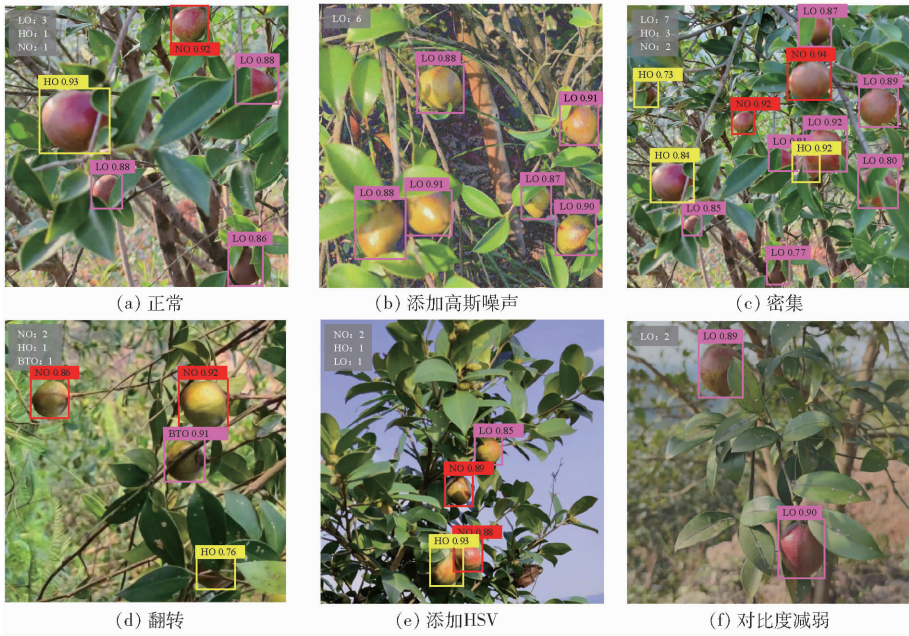


图 11 油茶果计数可视化结果

Fig. 11 Visualization results of *Camellia oleifera* fruits counting

3 结论

(1)为了在复杂遮挡环境下实现油茶果的检测与计数,提出了一种基于 YOLO v8n 的检测方法。该方法设计了一种双主干网络,通过独立提取来自不同网络的多样化特征信息,有效避免了传统重复双主干网络中常见的特征冗余问题。进一步提出了双输入单输出的连续注意力特征融合模块,对不同尺度的特征信息进行连续融合,增强模型对油茶果未遮挡部分的关注能力,降低漏检或错检现象。最后去除空间金字塔池化层,采用幻影卷积模块,削减了模型参数量和计算量,优化了模型整体性能。

(2)对不同光照条件、不同密集程度等复杂场

景进行讨论,实验结果表明,YOLO - DCL 模型改善了因各类遮挡而导致的漏检或错检现象,能够有效地识别无遮挡、树叶遮挡、树枝遮挡以及混合遮挡的油茶果,平均精度均值分别达到 98.3%、94.4%、90.7%和 87.3%,模型内存占用量为 5.7 MB。与基础模型 YOLO v8n 相比,YOLO - DCL 模型平均精度均值、精确率、召回率上分别提升 4.0、8.6、2.3 个百分点。与其他主流目标检测算法相比,YOLO - DCL 模型内存占用较小,整体识别性能最优。

(3)YOLO - DCL 模型具备对输入图像进行油茶果自动计数的功能,这不仅降低了人工计数成本,而且为后续生产统计工作提供了便利,适用于复杂环境下油茶果采摘过程,具有较高的实用价值和推广前景。

参 考 文 献

[1] CHEN S, ZOU X J, ZHOU X Z, et al. Study on fusion clustering and improved YOLOv5 algorithm based on multiple occlusion of *Camellia oleifera* fruit[J]. Computers and Electronics in Agriculture, 2023, 206: 107706.

[2] 伍德林, 赵恩龙, 姜山, 等. 基于双摆模型的油茶果冠层振动参数优化与试验[J]. 农业机械学报, 2021, 52(12): 96 – 104.
WU Delin, ZHAO Enlong, JIANG Shan, et al. Optimization analysis and test of canopy vibration parameters of camellia fruit based on double pendulum model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(12): 96 – 104. (in Chinese)

[3] 伍德林, 杨俊华, 刘芸, 等. 我国油茶果采摘装备研究进展与趋势[J]. 中国农机化学报, 2022, 43(1): 186 – 194.
WU Delin, YANG Junhua, LIU Yun, et al. Research progress and trend of camellia fruit picking equipment in China[J]. Journal of Chinese Agricultural Mechanization, 2022, 43(1): 186 – 194. (in Chinese)

[4] LEI X, WU M, LI Y, et al. Detection and positioning of *Camellia oleifera* fruit based on LBP image texture matching and binocular stereo vision[J]. Agronomy, 2023, 13(8): 2153.

[5] WU D, ZHAO E, FANG D, et al. Experiment and analysis of mechanized picking of *Camellia oleifera* fruit based on energy utilization rate[J]. INMATEH Agricultural Engineering, 2023, 69(1): 177 – 184.

[6] 周宏平, 金寿祥, 周磊, 等. 基于多模态图像的自然环境下油茶果识别[J]. 农业工程学报, 2023, 39(10): 175 – 182.
ZHOU Hongping, JIN Shouxiang, ZHOU Lei, et al. Recognizing *Camellia oleifera* fruits in natural environment using multi-modal images[J]. Transactions of the CSAE, 2023, 39(10): 175 – 182. (in Chinese)

[7] PAL A, LEITE A C, JOHAN P. A novel end-to-end vision-based architecture for agricultural human-robot collaboration in fruit picking operations[J]. Robotics and Autonomous Systems, 2024, 172: 104567.

[8] 伍德林, 李超, 曹成茂, 等. 摇枝式油茶果采摘装置作业过程分析与试验[J]. 农业工程学报, 2020, 36(10): 56 – 62.
WU Delin, LI Chao, CAO Chengmao, et al. Analysis and experiment of the operation process of branch-shaking type *Camellia oleifera* fruit picking device[J]. Transactions of the CSAE, 2020, 36(10): 56 – 62. (in Chinese)

[9] 杜小强, 宁晨, 贺磊盈, 等. 履带式高地隙油茶果振动采收机设计与试验[J]. 农业机械学报, 2022, 53(7): 113 – 121.
DU Xiaoqiang, NING Chen, HE Leiyong, et al. Design and test of crawler-type high clearance *Camellia oleifera* fruit vibratory harvester[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 113 – 121. (in Chinese)

[10] 伍德林, 袁嘉豪, 李超, 等. 扭梳式油茶果采摘末端执行器设计与试验[J]. 农业机械学报, 2021, 52(4): 21 – 33.
WU Delin, YUAN Jiahao, LI Chao, et al. Design and experiment of twist-comb end effector for picking camellia fruit[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 21 – 33. (in Chinese)

[11] 陈锋军, 陈闯, 朱学岩, 等. 基于改进 YOLOv7 的油茶果实成熟度检测[J]. 农业工程学报, 2024, 40(5): 177 – 186.
CHEN Fengjun, CHEN Chuang, ZHU Xueyan, et al. Detection of *Camellia oleifera* fruit maturity based on improved YOLOv7 [J]. Transactions of the CSAE, 2024, 40(5): 177 – 186. (in Chinese)

[12] 肖章, 彭江, 刘俊杰, 等. 基于 YOLOv5 – CP 的复杂环境下油茶果检测[J]. 中国农机化学报, 2023, 44(12): 193 – 199.
XIAO Zhang, PENG Jiang, LIU Junjie, et al. Detection of *Camella oleifera* fruit in complex environment based on YOLOv5 – CP[J]. Journal of Chinese Agricultural Mechanization, 2023, 44(12): 193 – 199. (in Chinese)

[13] ZHENG H, WANG G, LI X. YOLOX-Dense-CT: a detection algorithm for cherry tomatoes based on YOLOX and DenseNet [J]. Journal of Food Measurement and Characterization, 2022, 16(6): 4788 – 4799.

[14] ZHOU Y, TANG Y, ZOU X, et al. Adaptive active positioning of *Camellia oleifera* fruit picking points: classical image processing and YOLOv7 fusion algorithm[J]. Applied Sciences, 2022, 12(24): 12959.

[15] WU D, JIANG S, ZHAO E, et al. Detection of *Camellia oleifera* fruit in complex scenes by using YOLOv7 and data augmentation[J]. Applied Sciences-Basel, 2022, 12(22): 11318.

[16] ZHU X, YU Y, ZHENG Y, et al. Bilinear attention network for image-based fine-grained recognition of oil tea (*Camellia oleifera* Abel.) cultivars[J]. Agronomy, 2022, 12(8): 1846.

[17] LIN G, TANG Y, ZOU X, et al. In-field citrus detection and localisation based on RGB – D image analysis[J]. Biosystems Engineering, 2019, 186: 34 – 44.

[18] JANANI M, JEBAKUMAR R. Detection and classification of groundnut leaf nutrient level extraction in RGB images[J]. Advances in Engineering Software, 2023, 175: 103320.

[19] ALSHEHHI R, MARPU P R. Change detection using multi-scale convolutional feature maps of bi-temporal satellite high-resolution images[J]. European Journal of Remote Sensing, 2023, 56(1): 2161419.

[20] QUAN Y, LI M, HAO Y, et al. Tree species classification in a typical natural secondary forest using UAV-borne LiDAR and hyperspectral data[J]. Giscience & Remote Sensing, 2023, 60(1): 2171706.

[21] 棘玉, 尹显明, 严恩萍, 等. 基于相机拍照的油茶果形状特征提取研究[J]. 南京林业大学学报(自然科学版), 2022, 46(2): 63 – 70.
JI Yu, YIN Xianming, YAN Enping, et al. Research on extraction of shape features of *Camellia oleifera* fruits based on camera photography[J]. Journal of Nanjing Forestry University (Natural Sciences Edition), 2022, 46(2): 63 – 70. (in Chinese)

36(10): 1053 – 1058. (in Chinese)

[26] 牛善帅,王军政,张鹏,等. 基于负载口独立控制的双伺服阀控缸系统[J]. 北京理工大学学报,2019,39(12): 1292 – 1297.
NIU Shanshuai, WANG Junzheng, ZHANG Peng, et al. Dual servo valve control cylinder system based on independent control of load port [J]. Transactions of Beijing Institute of Technology, 2019, 39(12): 1292 – 1297. (in Chinese)

[27] 曾定荣,徐兵,刘英杰. 负载口独立控制系统中防止气穴的理论研究[J]. 机床与液压,2010,38(19):1 – 3.
ZENG Dingrong, XU Bing, LIU Yingjie. Anti-cavitation analysis of separate meter in and separate meter out control system [J]. Machine Tool & Hydraulics, 2010,38(19):1 – 3. (in Chinese)

[28] 柴红强,杨国来,苏华山,等. 液压碎石器内部气穴可视化分析[J]. 振动与冲击,2022,41(21):140 – 147,167.
CHAI Hongqiang, YANG Guolai, SU Huashan, et al. Visual analysis of cavitation inside hydraulic stone crusher[J]. Journal of Vibration and Shock, 2022,41(21):140 – 147,167. (in Chinese)

[29] 齐国宁,田婷,何增强,等. 高压柱塞泵气蚀现象及其控制[J]. 机床与液压,2021,49(9):187 – 191.
QI Guoning, TIAN Ting, HE Zengqiang, et al. Cavitation phenomenon of high-pressure piston pump and control[J]. Machine Tool & Hydraulics, 2021,49(9):187 – 191. (in Chinese)

[30] 刘念,王帆. 阀控非对称缸系统中气穴现象分析及解决方法[J]. 液压与气动,2021,45(8):177 – 182.
LIU Nian, WANG Fan. Analysis and countermeasures about cavitation in valve-controlled-asymmetrical-cylinder system[J]. Chinese Hydraulics & Pneumatics, 2021,45(8):177 – 182. (in Chinese)

(上接第 326 页)

[22] 陈斌,饶洪辉,王玉龙,等. 基于 Faster – RCNN 的自然环境下油茶果检测研究[J]. 江西农业学报, 2021, 33(1): 67 – 70.
CHEN Bin, RAO Honghui, WANG Yulong, et al. Study on detection of camellia fruit in natural environment based on Faster – RCNN[J]. Acta Agriculturae Jiangxi, 2021, 33(1): 67 – 70. (in Chinese)

[23] TANG Y, ZHOU H, WANG H, et al. Fruit detection and positioning technology for a *Camellia oleifera* C. Abel orchard based on improved YOLOv4 – tiny model and binocular stereo vision[J]. Expert Systems with Applications, 2023, 211: 118573.

[24] 宋怀波,王亚男,王云飞,等. 基于 YOLOv5s 的自然场景油茶果识别方法[J]. 农业机械学报, 2022, 53(7): 234 – 242.
SONG Huaibo, WANG Ya’nan, WANG Yunfei, et al. *Camellia oleifera* fruit detection in natural scene based on YOLOv5s [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 234 – 242. (in Chinese)

[25] 王金鹏,何萌,甄乾广,等. 基于改进 COF – YOLOv8n 的油茶果静态与动态检测计数方法[J]. 农业机械学报,2024, 55(4):193 – 203.
WANG Jinpeng, HE Meng, ZHEN Qianguang, et al. *Camellia oleifera* fruit static and dynamic detection counting based on improved COF – YOLOv8n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 193 – 203. (in Chinese)

[26] WANG Y Q, YAO Q M, KWOK J T, et al. Generalizing from a few examples: a survey on few-shot learning[J]. ACM Computing Surveys, 2020, 53(3): 1 – 34.

[27] WANG Z, DU L, MAO J, et al. SAR target detection based on SSD with data augmentation and transfer learning[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(1): 150 – 154.

[28] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6848 – 6856.

[29] SUN Y, CHEN G, ZHOU T, et al. Context-aware cross-level fusion network for camouflaged object detection[J]. arXiv preprint, arXiv: 2105.12555, 2021.

[30] PARK H, YOO Y J, SEO G, et al. C3: concentrated-comprehensive convolution and its application to semantic segmentation [J]. arXiv preprint, arXiv: 1812.04920, 2018.

[31] WANG C Y, LIAO H Y, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390 – 391.