

基于 YOLO-SSAR 的自然环境下红花检测算法

陈金荣¹, 许燕^{1,2}, 周建平^{1,2*}, 王小荣^{1,2}, 罗鸣³, 徐声彪³

(1. 新疆大学机械工程学院, 乌鲁木齐 830017; 2. 新疆维吾尔自治区农牧机器人及智能装备工程研究中心, 乌鲁木齐 830017;
3. 吉木萨尔县农牧业技术推广中心, 昌吉回族自治州 831700)

摘要: 针对自然环境中红花智能采摘存在红花尺度变化大、遮挡情况复杂的问题, 该研究对 YOLOv5s 模型进行优化, 提出了一种基于多尺度特征提取的 YOLO-SSAR 目标检测算法。首先, 采用 ShuffleNet v2 轻量化结构对 Backbone 层主干特征提取网络进行替换, 减少模型参数量和计算量; 其次, 在 Neck 层添加基于空洞卷积和共享权重的 Scale-Aware RFE 模块, 提高模型对于多尺度特征信息的提取能力; 最后, 为解决目标检测中类内、类间遮挡问题, 在 Head 层引入排斥损失函数对原损失函数进行替换, 减少因非极大值抑制 (non-maximum suppression, NMS) 阈值选取不当造成的漏检或误检, 提高模型的检测精度。试验结果表明, YOLO-SSAR 算法在测试集上的精确率、召回率和平均精度均值分别为 90.1%、88.5%、93.4%, 比 YOLOv5s 原始模型分别提升了 5.9、9.2 和 7.7 个百分点, 推理速度为 115 帧/s, 模型大小为 9.7 MB, 与主流算法 YOLOv4、YOLOv7、YOLOv8s、Faster R-CNN、SSD 相比, 精确率分别高出 6.8、7.2、6.3、16.2 和 10.8 个百分点、召回率高出 9.4、10.3、9.5、17.3 和 59.4 个百分点, 平均精度均值高出 8.8、8.2、8.1、14.9 和 19.4 个百分点。研究表明, YOLO-SSAR 算法在提升综合检测性能的同时也降低了计算复杂度, 研究结果可以为红花智能采摘研究提供算法参考。

关键词: 目标检测; 红花; YOLOv5s; ShuffleNet v2; 多尺度; 排斥损失函数

doi: 10.11975/j.issn.1002-6819.202408148

中图分类号: S225; TP391

文献标志码: A

文章编号: 1002-6819(2025)-02-0215-09

陈金荣, 许燕, 周建平, 等. 基于 YOLO-SSAR 的自然环境下红花检测算法[J]. 农业工程学报, 2025, 41(2): 215-223.

doi: 10.11975/j.issn.1002-6819.202408148 <http://www.tcsae.org>

CHEN Jinrong, XU Yan, ZHOU Jianping, et al. Detecting safflower in the natural environment using YOLO-SSAR[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(2): 215-223. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202408148 <http://www.tcsae.org>

0 引言

红花是一种主要用于药用、油用等工业加工原料的经济作物, 广泛栽培于中国新疆、河南等地区^[1]。通过前期调研发现, 红花不仅具有生长周期短、抗病、易出苗等特点, 而且耐寒、耐贫瘠、适应性较强^[2], 全生命周期的需水量仅为同期小麦的三分之一, 同时除花丝收获环节外已经实现了机械化, 每亩干红花产量在 20~30 kg 左右, 收购价格根据花丝品质在每千克 110~150 元浮动, 种植户所获亩利润可达 1 000 元以上。目前红花采收主要以人工为主, 人工方式存在采收效率低、成本高等问题^[3], 鲜红花的人工采收成本甚至达到了干花收入的一半以上。同时随着季节性劳动力的短缺问题愈发显著, 大量花丝无法在有限的收获期内进行采摘, 未

采摘或未采净的红花花丝又会阻碍新花毛生长, 进而导致了亩产量下降和经济损失。因此, 亟待研制智能化的红花采摘机器人来提高采收效率。

视觉系统作为红花采摘机器人实现探索作业信息快速感知的关键组成, 其首要环节就在于准确识别出非结构自然环境中处于不同尺度、不同光照、不同遮挡情况下的红花。近年来, 随着深度学习与机器视觉技术的不断发展^[4-6], 关于红花目标检测的科学研究已经取得了一定的进展。张新月等^[7]通过结合 Vanillanet 骨干架构和 LSKA 注意力机制, 有效降低了模型复杂度和参数量, 改进后轻量化模型 YOLOv8n-VLWS 相较于原模型, 检测速度提升 7.41%, 模型权重大小仅为原模型的 50.17%, 但是忽视了红花数据集的多样性与代表性, 模型泛化能力有待加强。王小荣等^[8]针对复杂环境中红花图像采集容易受不同光照、遮挡以及成熟度等因素影响, 利用多分类的 Focal Loss 损失函数进行替换, 提高了不均衡样本的关注度, mAP 值达到 88.5%, 与改进前相比, 不均衡类别样本检测的平均精度提高了 15.9 个百分点, 但是对于密集遮挡时红花多目标尺度变化的问题有待进一步研究。张振国等^[9]采用 SPP 空间金字塔池化结构强化目标特征的表现力, 引入 CBAM 注意力机制来解决红花图像的背景信息干扰问题, 平均精度均值达到 91.89%。郭

收稿日期: 2024-08-20 修订日期: 2024-12-02

基金项目: 新疆维吾尔自治区创新团队项目-机器人及智能装备技术科技创新团队 (2022D14002); 新疆维吾尔自治区自然科学基金项目: 非结构化环境下红花采摘机器人环境感知与认知决策关键技术研究 (2023D01C190)

作者简介: 陈金荣, 博士生, 研究方向为智能化检测与控制。

Email: 17798931507@163.com

*通信作者: 周建平, 教授, 博士生导师, 研究方向为机电一体化、农业机械化与信息化。Email: linkzhou@163.com

辉等^[10]提出了一种基于 C-YOLO v5m 的红花花冠目标检测算法,利用 CBAM 注意力模块自适应地调整输入的特征图,从而提升网络的性能和泛化能力,改进后平均精度均值可达 96.73%,但该研究并没有考虑外在田间环境与红花植株内在特征所带来的识别干扰。

虽然基于深度学习的红花目标检测已经取得了一定的效果,但仍存在以下问题:

1) 相机采集过程中,由于果球姿态变化、植株幅展不同及待检目标摆动,导致拍摄到的红花可能以各种角度出现;

2) 红花果球大小不一,同一视野下远近目标的尺度变化大,细节特征容易丢失;

3) 红花遮挡检测任务可以分为无遮挡、类间遮挡和类内遮挡 3 个类别,但现有研究大多基于类间遮挡情况来强化边界框损失或局部特征,忽视了类内遮挡情况对于真实目标回归所造成的干扰。

基于上述问题,本研究针对复杂田间环境和红花数据集的多样性特点,提出了一种基于 YOLO-SSAR 的自然环境下红花检测算法,该算法在 YOLOv5s 网络结构的基础上引入 ShuffleNet v2 网络对原始模型进行轻量化设计,同时在模型颈部网络中加入 Scale-Aware RFE 模块实现多尺度特征提取,并针对类间、类内遮挡问题,采用 Repulsion 边界框损失函数,旨在减少多目标密集场景下红花漏检或误检的情况,提升非结构环境下红花的检测精度,为红花采摘机器人视觉系统开发提供技术支持。

1 数据集制作

1.1 试验数据获取

本研究以新疆吉木萨尔县庆阳湖乡双河村红花千亩连片种植基地种植的平原红花与二工镇红山子村种植的山地红花为研究对象,拍摄时间为 2024 年 6 月 30 日—7 月 15 日,图像采集设备使用 CCD (charge-coupled device) 工业相机,相机像素为 1 920×1 080,相机与红花顶端之间的拍摄高度保持在 15~40 cm。由于红花枝干交叉繁杂且果球高低错落,当相机光轴与相机所处水平面的俯仰角低于 30°时,红花果球的完整性与可见度并不理想,同时容易造成复杂背景中相同类别物体的尺度变化剧烈。因此,为模拟红花采摘机器人的图像采集过程,在俯仰角为 30°~90°的范围内,随机采集了自然环境中不同位姿的红花图像。

考虑到不同天气条件下的光照强度不同,会导致不同时段所拍摄红花图像的整体曝光及局部阴影程度存在差异,因此分别采集上午 07:00-11:00、下午 13:00-17:00、傍晚 19:00-22:00 共 3 个时间段的红花图像。红花数据集包含不同生长形态、不同遮挡情况、不同光照情况及不同尺度的红花图像 2 852 幅,人工剔除模糊失真、尺度变化过大的图像后,最终得到红花数据集总数量为 2 650 幅,经标注后,将数据集按照 8:1:1 划分为训

练集 (2 120 幅)、验证集 (265 幅)、测试集 (265 幅) 进行训练与测试。图 1 为不同场景下的红花样本数据。



注: A 为被类内遮挡的目标, B 为 A 的邻近同类目标, 矩形框表示该框选区域存在目标红花被相邻红花或枝叶遮挡。

Note: A is the target that is occluded by the intra-class, B is the adjacent target of the same kind of A, the rectangular box indicates that the target safflower is in the selected area, obscured by adjacent safflowers or foliage.

图 1 不同场景下的红花样本

Fig.1 Samples of safflower in different scenes

1.2 数据处理

根据红花丝生长状态与不同花期化学成分变化情况^[11],红花开花过程可以分为花蕾期、初花期、盛花期和败花期。本研究将盛花期待采摘红花记为“Wait-Pick”,其他状态记为不可采摘红花“No-Pick”。标注盛花期待采摘红花时,为保证花丝药用经济价值,仅标注顶端生长形态特征为饱满耸立、颜色特征为黄色或桔红色花丝的红花。

由于红花以不规则的任意角度在图像中出现,本研究采用 roLabelImg 工具^[12]进行标注,该工具在支持水平框标注的同时,也能提供旋转框标注位姿状态多变的红花,在进行红花标注位置选取时,必须使标注框尽可能贴合红花花冠及果球,这有利于减少冗余信息和其他物体信息对红花特征提取的干扰。另外对于单幅图像中角度倾斜的对象,标注框的中轴线应尽可能与标注对象的中轴线保持重合或平行。

2 网络模型改进及训练

2.1 YOLOv5 网络模型

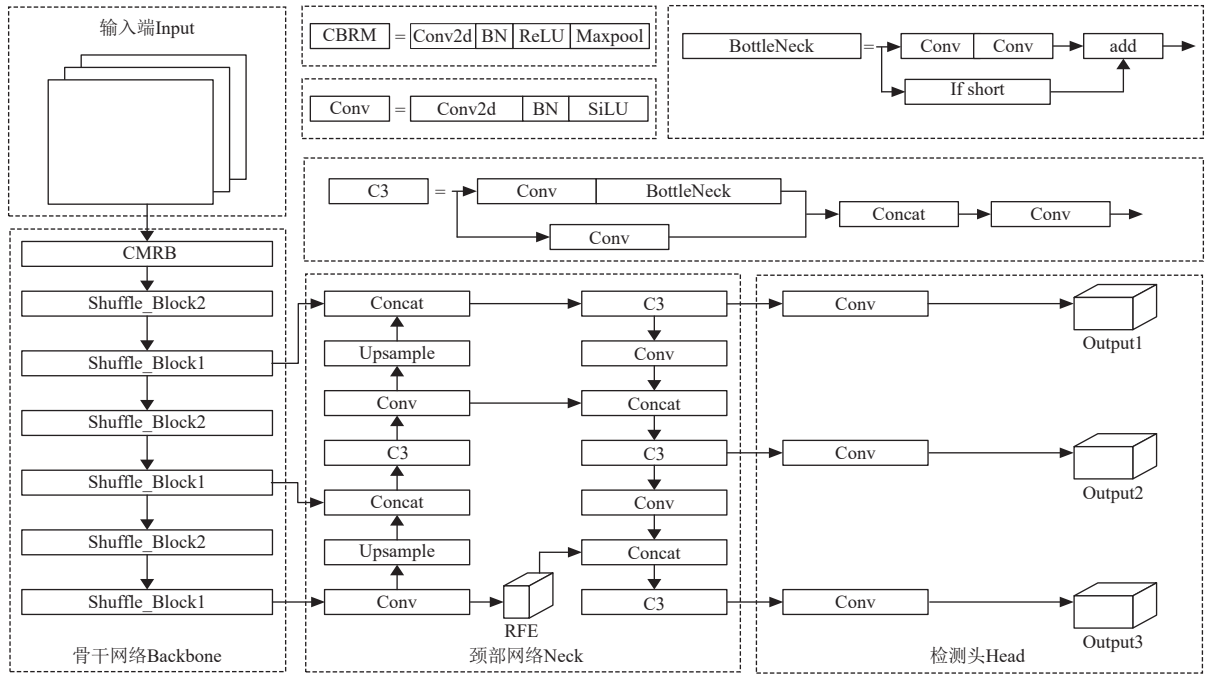
YOLOv5^[13-14]是具有模型小、检测速度快等优势的单阶段经典检测算法，在实时检测场景下更适用于移动端部署。YOLOv5 网络结构主要由 Input 层、Backbone 层、Neck 层和 Head 层组成。Input 层结合 Mosaic 数据增强、自适应锚框计算及自适应图片缩放等方式提高样本集的多样性和小目标在非常规语境中的预测精度；Backbone 层前端充分利用 Focus 结构的切片功能，在信息不丢失的情况下经过卷积获得二倍下采样特征图，并应用 CSP1_X 结构改善主干网络对遮挡物特征信息获取的能力；Neck 层沿袭 YOLOv4 的 FPN 和 PAN 组合，新增基于 CSPNet 思想的 CSP2_X 结构，进而加强网络特征融合能力；Head 层采用 CIOU_LOSS 作为 Bounding

box 损失函数，通过 NMS 非极大值抑制算法消除冗余的重叠边界框。

2.2 网络模型改进

虽然 YOLOv5 具有模型参数量小、推理速度快的优势，但仍旧存在检测精度不高、对模型的特征提取不全面等问题^[15-16]，因此本研究以 YOLOv5 为基础加以改进，提出一种基于 YOLO-SSAR 的红花目标检测算法。

首先在 Backbone 层中利用 ShuffleNet v2 网络结构替换原始骨干特征提取网络，完成模型降重的同时加快推理速度；然后针对尺度变化较大的多目标检测，通过在 Neck 层引入 Scale-Aware RFE 模块改善模型多尺度检测性能；最后替换 Head 层的损失函数^[17]，采用基于样本边界框排斥吸引原则的排斥损失函数 (repulsion loss) 优化回归^[18]，提高密集场景下目标重叠遮挡的检测率。改进后模型结构如图 2 所示。



注：Maxpool 为最大池化操作，CBRM 是由卷积层、BN、ReLU 和最大池化操作组成的模块，Shuffle_Block 是基本的 ShuffleNet 模块，Conv 为卷积操作，BN 为归一化操作，Concat 为通道数相加函数，Upsample 为上采样操作，ReLU 为修正线性单元的激活函数，SiLU 为结合了 ReLU 和 sigmoid 函数的非线性激活函数，RFE 为多尺度特征感知模块，下同。

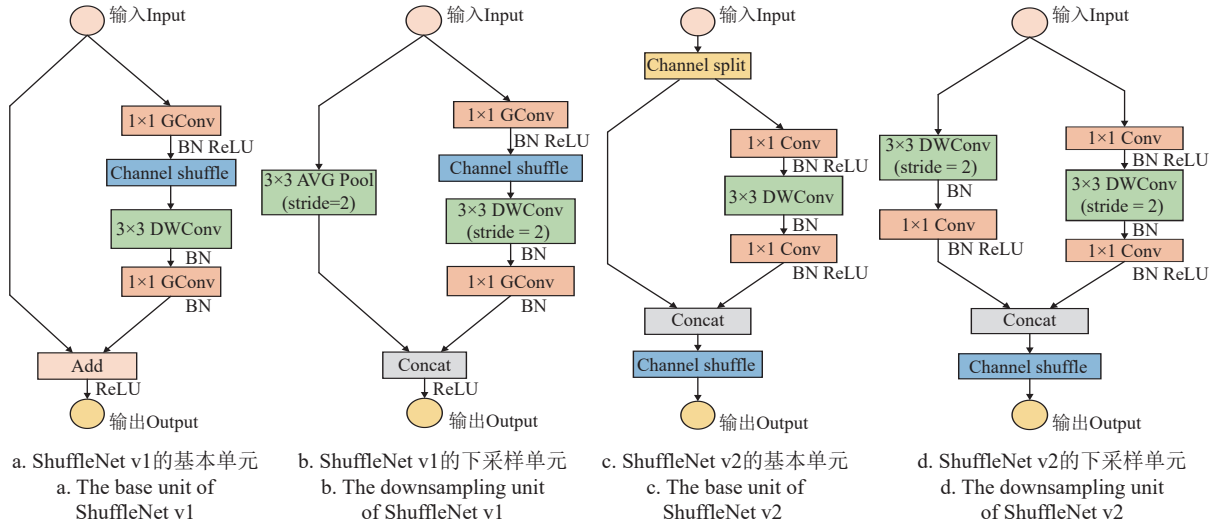
Note: Maxpool is the maximum pooling operation, CBRM is a module consisting of convolutional layer, BN, ReLU, and maximum pooling operation, Shuffle_Block is the basic ShuffleNet module, Conv is convolution, BN is batch normalizing, Concat is channel-wise addition, Upsample stands for feature upsamp, ReLU is the activation function for rectified linear unit, SiLU is the nonlinear activation function that combines ReLU and sigmoid functions, RFE is a multi-scale feature perception module, the same below.

图 2 YOLO-SSAR 模型结构
Fig.2 Model structure of YOLO-SSAR

2.2.1 ShuffleNet v2 轻量化网络

在 YOLOv5 原始网络中，采用 CSP-Darknet53 骨干结构作为特征提取网络^[19]。该结构是在 Darknet53 基础上引入 cross stage network (CSP) 结构，旨在通过跨阶段连接的方式将基础层的特征映射合并，提高目标特征信息的表现力。但考虑到目标标注区域存在着大量的冗余信息，而该网络又会对所有输入特征进行全面表征，进而导致模型复杂度提高，因此 YOLOv5 的骨干特征提取网络并不适用于计算能力有限的移动端设备。

基于此，本研究利用 ShuffleNet v2 轻量化网络^[20-21]替换原 CSP-Darknet53 骨干网络，在保证模型检测性能基本不变的前提下，降低计算和内存访问的成本。从图 3 可以发现，ShuffleNet v2 网络^[22]是 ShuffleNet v1 网络^[23]的改进版本，该网络在特征图输入进入卷积操作前进行了通道二划分，一个分支采取恒等映射方式来消除多分支结构，避免多路结构造成的网络碎片化，另一个分支通过常规卷积层和深度可分离卷积层保证输入与输出通道对等，最小化内存访问成本。

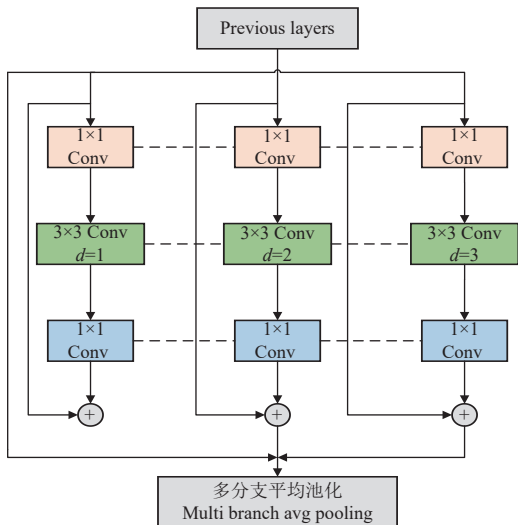


注: GConv 为分组卷积, DWConv 为深度可分离卷积, AVG Pool 为平均池化, Channel shuffle 为通道混洗, Channel split 为通道二划分, stride 为步长。
Note: GConv is the group convolution, DWConv is the deep separable convolution, AVG Pool is the average pooling, Channel shuffle is the channel mixing, Channel split is the channel two division operation, stride is the step size.

图3 ShuffleNet v1 与 ShuffleNet v2 的构建单元
Fig.3 The building blocks of ShuffleNet v1 and ShuffleNet v2

2.2.2 Scale-Aware RFE 模块

对于非结构环境中红花检测任务而言,真实场景中的红花图像由于光照的程度不同、目标错落分布的远近不同,导致红花尺寸跨度与特征差异较大,而 YOLOv5 网络在单一尺寸下容易丢失不同大小、不同尺度的目标特征。本研究在 Neck 层采用一种基于空洞卷积和共享权重的 Scale-Aware RFE 模块^[24]提升多尺度检测性能,使用 4 种不同比例的扩展卷积分支捕获多尺度信息和不同的依赖范围^[25],若检测尺度未知,Scale-Aware 策略将会控制不同尺度的分支负责训练其对应尺度的边界框。虽然基于尺度感知的训练思路可避免在不匹配分支上训练极端尺度的对象,但也可能造成有效样本减少,产生各分支的过拟合问题。Scale-Aware RFE 结构如图 4 所示。



注: d 表示为空洞率, Previous layers 为前端特征层的输出, \oplus 为逐元素相加。
Note: d is represented as dilation rate, Previous layers is the output of the headend feature layers, \oplus is the element-wise addition.

图4 Scale-Aware RFE 模块
Fig.4 Scale-Aware RFE module

由图 4 可知, Scale-Aware RFE 模块是由 3 个空洞卷积支路、收集和加权层两部分构成。多分支实现主要是通过设置不同空洞率 (1、2、3) 的 3x3 空洞卷积与 1x1 普通卷积结合,然后利用加权层收集来自不同分支的信息,并对特征的每个分支进行加权。因此 Scale-Aware RFE 模块将主路分支权值与其他支路共享,减少模型参数数量的同时,通过融合残差连接降低过拟合风险,使不同尺度的物体可以通过相同的表征能力进行统一转换。

2.2.3 损失函数改进

在密集场景检测任务中,红花遮挡可以分为无遮挡、类间遮挡^[26]与类内遮挡^[27]三种情况。类间遮挡主要是由枝干、枝叶交叉生长以及植株倒伏后出现的相邻行红花枝干相互缠绕所导致。类内遮挡大多产生于同类物体在检测区域中聚集分布,如图 1 所示,当红花果球 A 被红花果球 B 遮挡时,由于两物体的纹理特征和颜色特征相似,红花果球 A 容易出现漏检现象。为解决类间遮挡与类内遮挡所导致的漏检、误检问题,本研究通过引入排斥损失函数 (repulsion loss)^[28]在标准边界框回归损失中施加额外惩罚,不仅要求预测框接近待检物体的真实目标框,而且迫使每个预测框远离非待检物体的真实目标框,从而有效防止预测的边界框移动至相邻的重叠目标。

排斥损失函数主要由一个吸引项、两个排斥项和平衡系数组成,吸引项 (L_{Attr}) 为预测框和真实目标框之间的损失值,同类排斥项 (L_{RepGT}) 为预测框和相邻同类真实目标框之间的损失值,异类排斥项 (L_{RepBox}) 为预测框和异类真实目标框之间的损失值, α 和 β 表示平衡两种排斥项损失值的权重系数。排斥损失计算如式 (1) 所示:

$$L = L_{Attr} + \alpha L_{RepGT} + \beta L_{RepBox} \quad (1)$$

式中 L 表示排斥损失函数,吸引项 (L_{Attr}) 构造沿用了 YOLOv4 位置回归中的 $Smooth_{L1}$ 函数,该函数解决了最

小绝对误差在零点处不可导问题，提高收敛精度，同时为解决梯度爆炸问题，在最小平方误差基础上最小化预测值与真实值的误差，降低函数对异常值的敏感度。吸引损失计算如式（2）所示：

$$L_{Attr} = \frac{\sum_{h \in H} \text{Smooth}_{L1}(B^h \cdot G_{Attr}^h)}{|H|} \quad (2)$$

式中 H 表示正候选框集合，当候选框 h 与集合中一个真实框的交并比（IoU）大于 0.5 时，代表给定的一个候选框 $h \in H$ ， B^h 是回归调整所获取的预测框， G_{Attr}^h 表示与 h 有最大 IoU 的真实框。

同类排斥项的设计目标在于排斥邻近非目标类内物体的真实框，对于一个候选框 h 而言，其排斥真实框对象被定义为除其指定目标外具有最大 IoU 区域的对象。为避免模型优化失衡问题，使用 IoG 函数计算候选框 h 与邻近非目标物体真实框的距离，减少候选框 h 与非目标的真实框重叠。 L_{RepGT} 如式（3）所示：

$$L_{RepGT} = \frac{\sum_{h \in H} \text{Smooth}_{ln}(\text{IoG}(B^h \cdot G_{Rep}^h))}{|H|} \quad (3)$$

式中 Smooth_{ln} 是一个在区间（0，1）内连续可微分的平滑函数，如式（4）所示。

$$\text{Smooth}_{ln} = \begin{cases} -\ln(1-x) & x \leq \sigma \\ \frac{x-\sigma}{1-\sigma} - \ln(1-\sigma) & x > \sigma \end{cases} \quad (4)$$

异类排斥项是用于降低检测器对非极大值抑制算法（non-maximum suppression, NMS）敏感度，即避免两个不同类别物体的预测框受阈值影响，其中一个物体的预测框被 NMS 抑制或合并。 L_{RepBox} 如式（5）所示：

$$L_{RepBox} = \frac{\sum_{i \neq j} \text{Smooth}_{ln}(\text{IoU}(B^i, B^j))}{\sum_{i \neq j} I[\text{IoU}(B^i, B^j)] + \varepsilon} \quad (5)$$

式中 I 表示恒等函数， ε 为功能常数，用于避免出现 L_{RepBox} 函数的分母为 0 的情况。 L_{RepBox} 函数是根据图像包含的物体标签，首先将 H 集合划分成互斥的 q 个子集， q 表示目标物体数量。划分形式如式（6）所示：

$$H = h_1 \cap h_2 \cap h_3 \cap \dots \cap h_q \quad (6)$$

由式（5）和（6）可知， h_i 表示同组之间的候选框 h 回归的是同一样本标签， h_j 表示不同组之间候选框 h 对应了不同的样本标签，当 B^i 与 B^j 这两个目标不同的预测框重叠区域越小时，检测器在密集场景下鲁棒性越高。

2.3 模型训练环境与评价

2.3.1 训练环境

本研究以 Pytorch1.11.0 和 CUDA11.3 作为深度学习框架，基于 Python3.6，训练时调用 YOLOv5s.pt 作为预训练权重^[29-30]，采用 Adam 优化器^[31]并设置初始学习率为 0.01。

为避免模型过拟合并获得更好的泛化性能，在训练策略中加入早停机机制^[32]，即当固定周期内模型在验证集上的性能不再提升时，立即终止训练。经测试，训练轮数（epoch）最终设置为 150，数据样本喂入量（batchsize）

与子进程数量（workers）均设置为 8，图像输入大小为 640×640 像素。

2.3.2 评价指标

本文主要采用平均精度均值（mean average precision, mAP_{0.5}）、召回率（recall, R ）、精确率（precision, P ）、推理速度（frames per second, FPS）评价模型的检测性能。

3 结果与分析

3.1 不同改进模块的消融试验结果

为验证本文改进方法的有效性，以自然环境下红花图像为数据集进行消融试验，表 1 为消融试验结果。结果表明，本文提出的 YOLO-SSAP 模型的 mAP_{0.5} 值达到 93.4%，推理速度为 115 帧/s，模型大小为 9.7 MB。

表 1 YOLO 模型的消融试验结果

ShuffleNet v2	Scale-Aware RFE	Repulsion Loss	mAP _{0.5} /%	FPS/(帧·s ⁻¹)	模型大小 Model size/MB
×	×	×	85.7	104	16.4
√	×	×	84.9	132	11.2
×	√	×	90.8	100	14.8
×	×	√	91.5	98	16.5
√	√	×	89.6	128	9.7
√	×	√	90.1	121	11.4
×	√	√	91.7	84	15.8
√	√	√	93.4	115	9.7

注：×表示不使用该改进策略，√表示使用该改进策略，mAP_{0.5} 表示 IoU 阈值取 0.5 时，模型计算得到的平均精度均值，FPS 为单幅图像推理时间。

Note: × indicates that the improvement strategy is not used, √ indicates that the improvement strategy is used, mAP_{0.5} represents the mean average precision calculated by the model when the IoU threshold is set to 0.5, FPS is the inference time for a single image.

由表 1 可知，随着 CSP-Darknet53 特征提取网络被 ShuffleNet v2 网络替换，mAP_{0.5} 值下降了 0.8 个百分点，但模型的推理速度得到提升，这表明 ShuffleNet v2 在保证检测精度基本不变的情况下，能够降低模型计算复杂度。在此基础上，将 Scale-Aware RFE 模块引入 Neck 层，检测精度得到进一步提升，mAP_{0.5} 值增加了 4.7 个百分点，表明新模型在多尺度特征提取方面的检测性能得到进一步优化。基于上述两项改进策略，将 YOLOv5s 的原损失函数替换为排斥损失函数，最终形成了 YOLO-SSAR 算法模型，mAP_{0.5} 值又增加了 3.8 个百分点，说明 repulsion loss 函数有利于提高模型在密集场景下对于遮挡目标的检测成功率，促使 Head 层检测器更加关注类内、类间遮挡问题。综上所述，针对 YOLOv5s 网络中 Backbone 层、Neck 层及 Head 层的不同改进策略均可以改善模型的检测性能，相较于原始模型的 mAP_{0.5} 值最终增加了 7.7 个百分点，模型大小减少 6.7 MB。

3.2 不同算法模型结果对比

当前目标检测算法主要包括 RCNN、YOLO、SSD 三个系列。RCNN 属于 two-stage 方法，即模型先通过生成样本候选框后再进行分类与回归，准确率高但检测效率低。YOLO 与 SSD 属于 one-stage 方法，该方法可直接回归目标的类别及位置信息，能够平衡目标检测任务的实时性与准确性要求。为评估本研究提出 YOLO-SSAR 算法模型的检测性能，在 two-stage 方法中选择

Faster R-CNN 算法, 在 one-stage 方法中选择 YOLOv4、YOLOv7、YOLOv8s、SSD 分别与改进算法模型进行比较, 不同模型的检测性能如表 2 所示。

表 2 不同模型检测性能对比

Table 2 Comparison of detection performance with different models

模型 Model	精确率 Precision P/%	召回率 Recall R/%	$mAP_{0.5}/\%$	FPS/ (帧·s ⁻¹)	模型大小 Model size/MB
YOLO-SSAR	90.1	88.5	93.4	115	9.7
YOLOv5s	84.2	79.3	85.7	104	16.4
YOLOv4	83.3	79.1	84.6	77	25.1
YOLOv7	82.9	78.2	85.2	84	76.8
YOLOv8s	83.8	79.0	85.3	55	22.5
Faster R-CNN	73.9	71.2	78.5	21	228.3
SSD	79.3	29.1	74.0	32	101.4

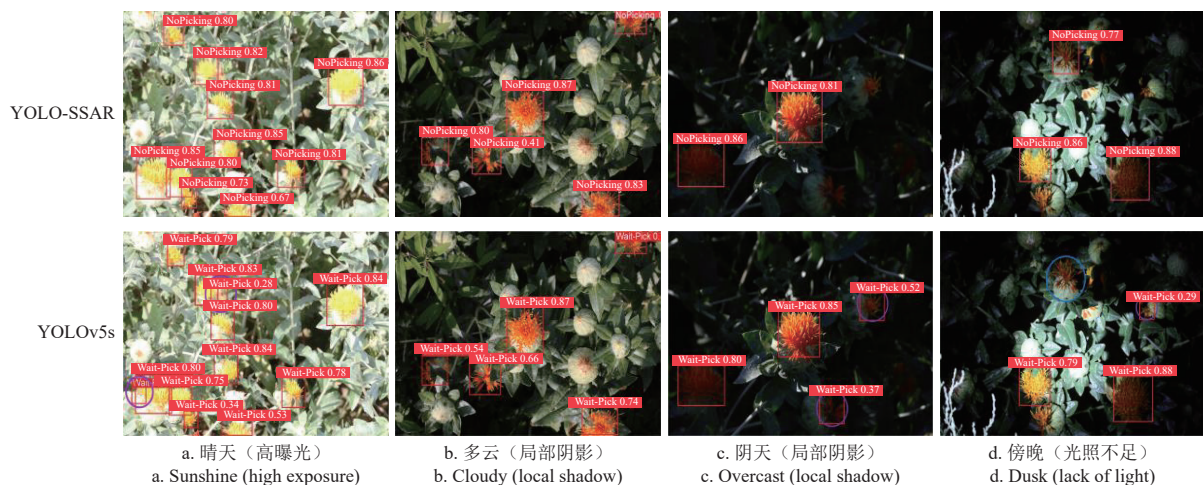
从表 2 可以发现, 改进模型的 P 值、 R 值、 $mAP_{0.5}$ 值较原模型分别提升了 5.9、9.2 和 7.7 个百分点, $mAP_{0.5}$ 值分别高出 YOLOv4、YOLOv7、YOLOv8s、Faster R-CNN、SSD 等主流模型 8.8、8.2、8.1、14.9、19.4 个百分点, R 值分别高出 9.4、10.3、9.5、17.3、59.4 个百分点, P 值分别高出了 6.8、7.2、6.3、16.2、10.8 个百分点。此外, YOLO-SSAR 算法推理速度最快, 模型大小均小

于对比算法。试验结果表明, 本文网络模型参数量最小, 说明 ShuffleNet v2 轻量化结构的加入使模型在保持精度提升的同时, 有效控制了改进模型复杂度的增加。本文算法的推理速度达到 115 帧/s, 均大于其他主流模型, 是两阶段目标检测算法 Faster R-CNN 的 5.5 倍, 是多尺度目标检测算法 SSD 的 3.6 倍, 而模型大小为 9.7 MB, 仅占 Faster R-CNN 的 4% 和 SSD 的 10%, 并优于其他 YOLO 系列模型, 可以满足实时检测的要求。基于此, YOLO-SSAR 目标检测算法的 P 值、 R 值和 $mAP_{0.5}$ 值分别达到 90.1%、88.5% 和 93.4%, 检测准确性和模型大小均优于其他 6 种主流, 更适合部署在算力资源受限的移动端设备中。

3.3 检测效果分析

自然环境中采集到的红花图像主要有以下 3 种光照情况: 自然光照强度极大时图像背景及颜色发白失真; 枝叶交错导致的局部阴影程度不同; 昏暗光线条件下目标边缘模糊、颜色信息丢失。

为了验证 YOLO-SSAR 目标检测算法的识别效果, 挑选不同天气、不同光照的 50 幅红花图像进行测试, 检测效果如图 5 所示。



注: 蓝色标记圈为漏检红花, 紫色标记圈为误检红花, 下同。

Note: The blue marked ring is the missing detection of safflower, the purple marked ring is the false detection of safflower, the same below.

图 5 不同光照强度的检测效果

Fig.5 Detection effect of different light intensity

由于晴天条件下图像曝光现象严重 (图 5a), 易发生区域变亮或褪色, YOLO-SSAR 算法能够准确区分重合度较高的红花目标, 避免重复框选和选框混乱的现象。在阴天条件下 YOLO-SSAR 针对局部阴影情况 (图 5c) 均准确识别出 2 个误检红花。当傍晚光照不足时 (图 5d), 本文提出的改进算法不仅检测出 1 个漏检红花, 而且筛选了原始网络误检出的 1 个败花期衰败红花。虽然在多云天气中 (图 5b) 改进前后的算法均能检测出真实目标, 但 YOLO-SSAR 检测框较原始 YOLOv5s 检测框的平均置信度更高。综上所述, YOLO-SSAR 算法在面对复杂光照背景时综合检测性能更佳。

本研究所提出的目标检测算法针对不同尺度目标的

检测效果如图 6 所示。YOLO-SSAR 算法在俯仰角为 45° 时 (图 6a), 准确识别出 YOLOv5s 算法漏检的 6 个红花目标, 并剔除了 1 个近距离误检红花。在相机俯仰角为 60° 时 (图 6b), 改进后的算法可以在图像边缘检测出相对尺寸较小的漏检目标。红花位姿差异较大时 (图 6c), 虽然改进前后的算法均识别出 7 个目标红花, 但 YOLO-SSAR 较原始网络有 5 个置信度更高的检测框。夜间检测时 (图 6d), YOLO-SSAR 算法能够在远、近距离各识别出 1 个漏检红花。

结果表明, 模型在引入 Scale-Aware RFE 模块后能够准确识别出相对位置较远或较近的目标, 而且针对尺寸不一、位姿各异的红花均可以实现精准检测。

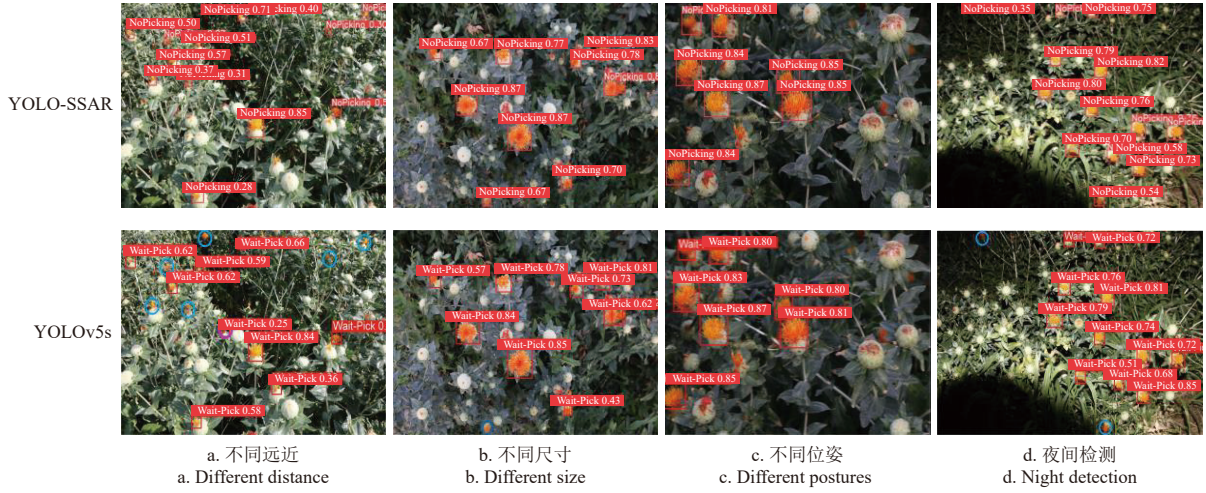


图 6 不同尺度情况的检测效果

Fig.6 The detection effect of different scales situations

不同遮挡情况下的红花检测效果如图 7 所示。当图像中心的相邻红花存在枝叶交错、重叠时（图 7a），原始模型难以保证逻辑回归的准确性，而本文提出的改进算法在筛除 1 个误检框的同时，合理框选了目标区域。针对类内遮挡的两目标重叠问题，YOLO-SSAR 算法展现出更高的置信度，尤其是当图像边缘出现红花重叠且完整度较低的情况（图 7b），改进算法依然能够有效地检测出来。在类内遮挡中存在多目标重合度高且部分果球可见度较低的情况时（图 7d），YOLOv5s 算法将 2 个目标红花识别成 1 个，而本文提出的 YOLO-SSAR 算

法利用排斥损失函数合理选取 NMS 阈值，将 2 个同类簇生遮挡的红花分割成置信度分别为 0.82 和 0.83 的检测框且框选区域较为准确。但由于红花重叠处颜色特征相似，花丝外轮廓边界模糊，YOLO-SSAR 识别出 1 个置信度为 0.35 的误检框。

综上所述，YOLO-SSAR 目标检测算法能够准确识别出受不同遮挡因素影响的待采摘红花，尤其是对于类内遮挡问题，该模型具有较好的检测性能，说明模型在更换排斥损失函数后，可以解决同类、目标重叠情况下边界框混乱问题。

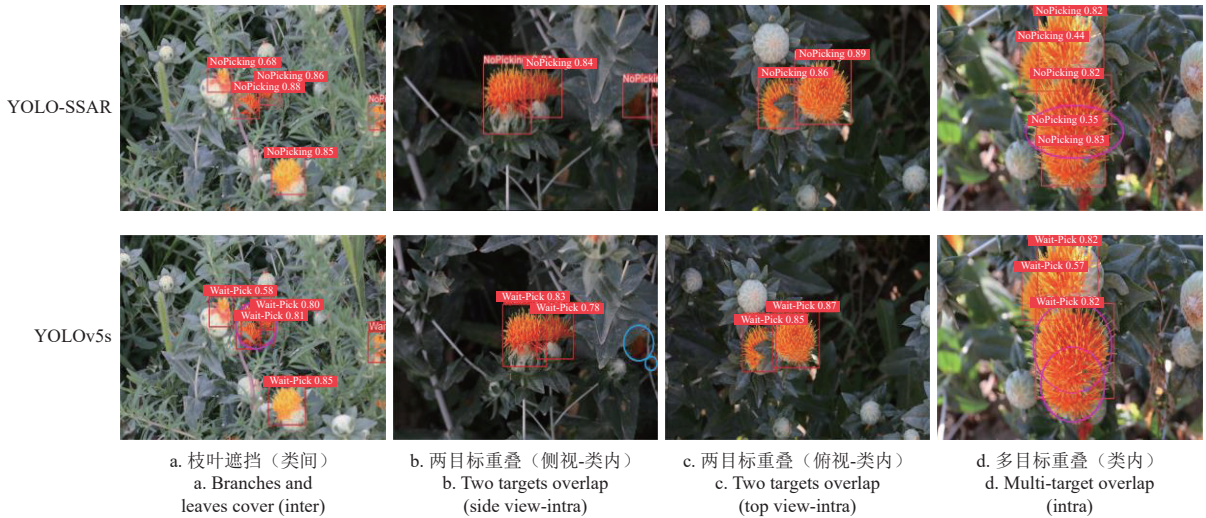


图 7 不同遮挡情况的检测效果

Fig.7 Detection effect of different occlusion conditions

4 结 论

本研究首先针对复杂环境干扰、尺度变化以及二类遮挡对模型检测性能的影响进行分析，提出了一种兼顾检测精度和推理速度的 YOLO-SSAR 红花检测算法，主要结论如下：

1) 使用 ShuffleNet v2 轻量化网络对 YOLOv5s 模型的 Backbone 层加以改进，均衡模型复杂度和检测效率的

同时，降低输入特征提取层的碎片化程度。基于此，在 Neck 层加入 Scale-Aware RFE 模块，提高多尺度目标特征提取能力。引入排斥损失函数替换 Head 层的回归损失函数，提高类间、类内遮挡的检测精度。改进后的算法模型相较于原模型，精确率 P 、召回率 R 和平均精度均值 $mAP_{0.5}$ 分别提高 5.9、9.2 和 7.7 个百分点，模型大小减少 6.7 MB。

2) 与 YOLOv4、YOLOv7、YOLOv8s、Faster R-

CNN、SSD 算法模型相比, YOLO-SSAR 算法平均精度均值 $mAP_{0.5}$ 分别提高 8.8、8.2、8.1、14.9、19.4 个百分点, 精确率 P 分别提高 6.8、7.2、6.3、16.2、10.8 个百分点, 召回率 R 分别提高 9.4、10.3、9.5、17.3、59.4 个百分点, 推理速度为 115 帧/s, 模型大小仅为 9.7MB。

3) 本文提出的 YOLO-SSAR 算法精确率 P 为 90.1%, 召回率 R 为 88.5%, 平均精度均值 $mAP_{0.5}$ 为 93.4%, 能够满足红花复杂背景下实时检测要求, 可为红花智能采收提供支持支撑。

【参 考 文 献】

- [1] 梁慧珍, 董薇, 余永亮, 等. 国内外红花种质资源研究进展[J]. 安徽农业科学, 2015, 43(16): 71-74.
LIANG Huizhen, DONG Wei, YU Yongliang, et al. Advances in studies on safflower (*Carthamus Tinctorius* L.) at home and abroad[J]. Journal of Anhui Agricultural Sciences, 2015, 43(16): 71-74. (in Chinese with English abstract)
- [2] 孟景怡. 新疆红花产量与水分利用效率对旱地坡位与灌溉的响应[D]. 乌鲁木齐: 新疆农业大学, 2023.
MENG Jingyi. Responses of Safflower Yield and Water Use Efficiency to Dryland Slope Position and Irrigation in Xinjiang[D]. Urumqi: Xinjiang Agricultural University, 2023. (in Chinese with English abstract)
- [3] 周远航, 郭建富, 马小龙, 等. 新疆红花生产现状及发展对策研究[J]. 安徽农业科学, 2021, 49(19): 199-201, 217.
ZHOU Yuanhang, GUO Jianfu, MA Xiaolong, et al. Research on current situation and development countermeasures of Xinjiang safflower production, 2021, 49(19): 199-201, 217. (in Chinese with English abstract)
- [4] XU F, YAO X, ZHANG K, et al. Deep learning in cropland field identification: A review[J]. *Computers and Electronics in Agriculture*, 2024, 222: 109042.
- [5] THAKUR P S, KHANNA P, SHEOREY T, et al. Trends in vision-based machine learning techniques for plant disease identification: A systematic review[J]. *Expert Systems with Applications*, 2022, 208: 118117.
- [6] LI Z, YUAN X, WANG C. A review on structural development and recognition-localization methods for end-effector of fruit-vegetable picking robots[J]. *International Journal of Advanced Robotic Systems*, 2022, 19(3): 1-29.
- [7] 张新月, 胡广锐, 李浦航, 等. 基于改进 YOLOv8n 的轻量化红花识别方法[J]. 农业工程学报, 2024, 40(13): 163-170.
ZHANG Xinyue, HU Guangrui, LI Puhang, et al. Recognizing safflower using improved lightweight YOLOv8n[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2024, 40(13): 163-170. (in Chinese with English abstract)
- [8] 王小荣, 许燕, 周建平, 等. 基于改进 YOLOv7 的复杂环境下红花采摘识别[J]. 农业工程学报, 2023, 39(6): 169-176.
WANG Xiaorong, XU Yan, ZHOU Jianping, et al. Safflower picking recognition in complex environments based on an improved YOLOv7[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(6): 169-176. (in Chinese with English abstract)
- [9] 张振国, 邢振宇, 赵敏义, 等. 改进 YOLOv3 的复杂环境下红花丝检测方法[J]. 农业工程学报, 2023, 39(3): 162-170.
ZHANG Zhenguo, XING Zhenyu, ZHAO Minyi, et al. Detecting safflower filaments using an improved YOLOv3 under complex environments[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(3): 162-170. (in Chinese with English abstract)
- [10] 郭辉, 陈海洋, 高国民, 等. 基于 YOLOv5m 的红花花冠目标检测与空间定位方法[J]. 农业机械学报, 2023, 54(7): 272-281.
GUO Hui, CHEN Haiyang, GAO Guomin, et al. Safflower corolla object detection and spatial positioning methods based on YOLOv5m[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(7): 272-281. (in Chinese with English abstract)
- [11] 李彤彤. 红花质量评价和不同花期成分变化研究[D]. 北京: 北京中医药大学, 2019.
LI Tongtong. Study on the Evaluation of Safflower Quality and Compositional Changes at Different Flowering Stages[D]. Beijing: Beijing University of Chinese Medicine, 2019. (in Chinese with English abstract)
- [12] 于俊伟, 陈威威, 郭园森, 等. 基于改进 Oriented R-CNN 的旋转框麦穗检测与计数模型[J]. 农业工程学报, 2024, 40(6): 248-257.
YU Junwei, CHEN Weiwei, GUO Yuansen, et al. Improved Oriented R-CNN-based model for oriented wheat ears detection and counting[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2024, 40(6): 248-257. (in Chinese with English abstract)
- [13] ABDULLAH A, AMRAN G A, TAHMID S M A, et al. A deep-learning-based model for the detection of diseased tomato leaves[J]. *Agronomy*, 2024, 14(7): 1593.
- [14] LI F, ZHENG Y, LIU S, et al. A multi-objective apple leaf disease detection algorithm based on improved TPH-YOLOv5[J]. *Applied Fruit Science*, 2024, 66(2): 399-415.
- [15] 范天浩, 顾寄南, 王文波, 等. 基于改进 YOLOv5s 的轻量化金银花识别方法[J]. 农业工程学报, 2023, 39(11): 192-200.
FAN Tianhao, GU Jinan, WANG Wenbo, et al. Lightweight honeysuckle recognition method based on improved YOLOv5s[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(11): 192-200. (in Chinese with English abstract)
- [16] ZHOU M, LIU S, LI J. Multi-scale forest flame detection based on improved and optimized YOLOv5[J]. *Fire Technology*, 2023, 59(6): 3689-3708.
- [17] 彭炫, 周建平, 许燕, 等. 改进 YOLOv5 识别复杂环境下棉花顶芽[J]. 农业工程学报, 2023, 39(16): 191-197.
PENG Xuan, ZHOU Jianping, XU Yan, et al. Cotton top bud recognition method based on YOLOv5-CPP in complex environment[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(16): 191-197. (in Chinese with English abstract)
- [18] 孙龙, 张荣芬, 刘宇红, 等. 监控视角下密集人群口罩佩戴检测算法[J]. 计算机工程, 2023, 49(9): 313-320.
SUN Long, ZHANG Rongfen, LIU Yuhong, et al. Mask wearing detection algorithm for dense crowds from a monitoring perspective[J]. *Computer Engineering*, 2023, 49(9): 313-320. (in Chinese with English abstract)
- [19] ZHU R, ZOU H, LI Z, et al. Apple-Net: A model based on improved YOLOv5 to detect the apple leaf diseases[J]. *Plants*, 2022, 12(1): 169.
- [20] WIJAYA E S, MIZWAR A, ISLAMI A M, et al. Garbage classification using CNN architecture ShuffleNet v2[C]//2022 IEEE/Seventh International Conference on Informatics and Computing (ICIC), Denpasar, Indonesia, 2022: 1-4.
- [21] 孙月平, 孙杰, 袁必康, 等. 基于改进 YOLOv5s 的轻量化蟹塘障碍物检测与定位方法[J]. 农业工程学报, 2023, 39(23): 152-163.
SUN Yueping, SUN Jie, YUAN Bikang, et al. Lightweight crab pond obstacle detection and location method based on improved YOLOv5s[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(23): 152-163. (in Chinese with English abstract)
- [22] MA N, ZHANG X, ZHENG H, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich, 2018: 116-131.

- [23] 刘浩翰, 樊一鸣, 贺怀清, 等. 改进 YOLOv7-tiny 的目标检测轻量化模型[J]. *计算机工程与应用*, 2023, 59(14): 166-175.
LIU Haohan, FAN Yiming, HE Huaqing, et al. Improved YOLOv7-tiny's object detection lightweight model[J]. *Computer Engineering and Applications*, 2023, 59(14): 166-175. (in Chinese with English abstract)
- [24] YU Z, HUANG H, CHEN W, et al. YOLO-FaceV2: A scale and occlusion aware face detector[J]. *Pattern Recognition*, 2024, 155: 110714.
- [25] 顾清华, 周宇静, 王丹, 等. 基于多尺度特征提取的露天矿区道路坑洼检测模型[J/OL]. *金属矿山*, 1-13[2024-05-17]. <http://kns.cnki.net/kcms/detail/34.1055.td.20240516.1134.006.html>.
GU Qinghua, ZHOU Yujing, WANG Dan, et al. Road pothole detection model in opencast mining area based on multi-scale feature extraction[J/OL]. *Metal Mine*, 1-13[2024-05-17]. <http://kns.cnki.net/kcms/detail/34.1055.td.20240516.1134.006.html>. (in Chinese with English abstract)
- [26] 李艳凤, 薛盼, 陈后金, 等. 一种基于弱监督语义分割和特征弥补的类间遮挡目标检测网络模型的训练方法: CN116503603B[P]. 2024-02-23.
- [27] 鲁岳峻, 张逸飞, 王庆, 等. 基于跨尺度分层特征融合对类内遮挡的鸟鳢检测的方法: CN114863263B[P]. 2022-09-13.
- [28] ZHANG S, WEN L, BIAN X, et al. Occlusion-aware R-CNN: Detecting pedestrians in a crowd[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich, 2018: 637-653.
- [29] 刘永波, 高文波, 何鹏, 等. 基于改进 ResNet50 模型的自然环境下苹果物候期识别[J]. *智慧农业 (中英文)*, 2023, 5(2): 13-22.
LIU Yongbo, GAO Wenbo, HE Peng, et al. Apple phenological period identification in natural environment based on improved ResNet50 model[J]. *Smart Agriculture*, 2023, 5(2): 13-22. (in Chinese with English abstract)
- [30] 兰玉彬, 孙斌书, 张乐春, 等. 基于改进 YOLOv5s 的自然场景下生姜叶片病虫害识别[J]. *农业工程学报*, 2024, 40(1): 210-216.
LAN Yubin, SUN Binshu, ZHANG Lechun, et al. Identifying diseases and pests in ginger leaf under natural scenes using improved YOLOv5s[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2024, 40(1): 210-216. (in Chinese with English abstract)
- [31] WU T, ZHANG Q, WU J, et al. An improved YOLOv5s model for effectively predict sugarcane seed replenishment positions verified by a field re-seeding robot[J]. *Computers and Electronics in Agriculture*, 2023, 214: 108280.
- [32] 张伏, 陈自均, 鲍若飞, 等. 基于改进型 YOLOv4-LITE 轻量级神经网络的密集圣女果识别[J]. *农业工程学报*, 2021, 37(16): 270-278.
ZHANG Fu, CHEN Zijun, BAO Ruofei, et al. Recognition of dense cherry tomatoes based on improved YOLOv4-LITE lightweight neural network[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2021, 37(16): 270-278. (in Chinese with English abstract)

Detecting safflower in the natural environment using YOLO-SSAR

CHEN Jinrong¹, XU Yan^{1,2}, ZHOU Jianping^{1,2*}, WANG Xiaorong^{1,2}, LUO Ming³, XU Shengbiao³

(1. College of Mechanical Engineering, Xinjiang University, Urumqi 830017, China; 2. Agriculture and Animal Husbandry Robot and Intelligent Equipment Engineering Research Center of Xinjiang Uygur Autonomous Region, Urumqi 830017, China; 3. Jimsar County Agricultural and Animal Husbandry Technology Extension Center, Changji Hui Autonomous Prefecture 831700, China)

Abstract: Safflower has often drawn much attention in the field of intelligent harvesting, due to their economic value. The safflower harvesting has also posed the higher requirements for object detection, due to the large-scale variations and complex occlusion in natural environments. Furthermore, missed or false detection has often occurred in traditional object detection, thus seriously affecting picking efficiency and accuracy. In this study, a YOLO-SSAR object detection was proposed to optimize the original YOLOv5s model using multi-scale feature extraction. The effectiveness and rationality of the improved algorithm were verified through ablation experiments, model comparison, and detection effect analysis. Firstly, the ShuffleNet v2 lightweight structure was used to replace the backbone feature extraction network of the backbone layer, in order to reduce the number of model parameters and calculations. The efficient channel mixing and depthwise separable convolution were utilized to improve the efficiency of input feature extraction. Secondly, a Scale-Aware RFE module was added to the Neck layer using dilated convolution and shared weights, in order to extract the multi-scale features. The weights of the main branch were shared with the rest branches, thus lowering the number of model parameters. The risk of overfitting was also reduced to fuse the residual connections, thereby allowing the objects of different scales to be uniformly transformed with the same representation. Finally, the repulsion loss function was introduced into the head layer to replace the original loss function, in order to avoid the intra-class and inter-class occlusion in the object detection. There was a reduction in the missed or false detection caused by improper selection of the non-maximum suppression (NMS) threshold. The detection rate of the target was improved with the overlap occlusion in dense scenes. The experimental results showed that the precision, recall, and mean average precision of the YOLO-SSAR algorithm on the test set were 90.1%, 88.5%, and 93.4%, respectively. Compared with the original YOLOv5s model, the YOLO-SSAR algorithm was improved by 5.9, 9.2, and 7.7 percentage points, respectively. The inference speed reached 115 frames per second, and the model size was 9.7 MB, indicating high efficiency and lightweight in practical applications. Compared with the mainstream algorithms YOLOv4, YOLOv7, YOLOv8s, Faster R-CNN, and SSD, the detection accuracy of the YOLO-SSAR algorithm was in a leading position. The improved model was 5.5 times and 3.6 times that of the two-stage and multi-scale object detection of Faster R-CNN and SSD respectively. Meanwhile, the model size was only 4% of Faster R-CNN and 10% of SSD. The minimum quantity of parameters shared the great prospects in the mobile devices with the limited computing resources. The precision was 6.8, 7.2, 6.3, 16.2, and 10.8 percentage points higher, the recall was 9.4, 10.3, 9.5, 17.3, and 59.4 percentage points higher, and the mean average precision was 8.8, 8.2, 8.1, 14.9 and 19.4 percentage points higher than the mainstream algorithm, respectively. The YOLO-SSAR algorithm improved the detection performance with less computational complexity. The findings can provide the algorithm references for the intelligent harvesting of safflower.

Keywords: target detection; safflower; YOLOv5s; ShuffleNet v2; multi-scale; repulsion loss function