

化春键, 黄宇峰, 蒋毅, 等. 基于改进 YOLOv5s 模型的田间食用玫瑰花检测方法[J]. 江苏农业学报, 2024, 40(8): 1464-1472.  
doi:10.3969/j.issn.1000-4440.2024.08.011

## 基于改进 YOLOv5s 模型的田间食用玫瑰花检测方法

化春键<sup>1,2</sup>, 黄宇峰<sup>1,2</sup>, 蒋毅<sup>1,2</sup>, 俞建峰<sup>1,2</sup>, 陈莹<sup>3</sup>

(1.江南大学机械工程学院, 江苏 无锡 214122; 2.江苏省食品先进制造装备技术重点实验室, 江苏 无锡 214122; 3.江南大学物联网工程学院, 江苏 无锡 214122)

**摘要:** 为了在田间环境下准确检测食用玫瑰花及其成熟度, 实现花期玫瑰花的自动化采摘, 针对田间光照、遮挡等因素造成识别精度较差的问题, 提出了一种基于 YOLOv5s 的改进模型, 对花蕾期、采摘期、败花期食用玫瑰花的生长状态进行检测。首先, 为了增强多尺度特征融合能力, 对特征融合结构进行改进。其次, 采用多分支结构训练提高精度, 在颈部网络 C3 模块进行改进。最后, 为了提升特征信息的提取能力, 在颈部网络中添加融合注意力模块, 使模型关注检测目标, 减少玫瑰花的误检及漏检现象。改进后的模型检测总体类别平均精度较原始模型提升了 3.6 个百分点, 达到 90.4%, 对 3 个花期玫瑰花的检测精度均有提升。本研究结果为非结构环境下的不同花期食用玫瑰花检测提供了更加准确的方法。

**关键词:** 目标检测; YOLOv5s; 特征融合; 注意力机制; 食用玫瑰花

**中图分类号:** TP391.4 **文献标识码:** A **文章编号:** 1000-4440(2024)08-1464-09

## Detection method of edible roses in field based on improved YOLOv5s model

HUA Chunjian<sup>1,2</sup>, HUANG Yufeng<sup>1,2</sup>, JIANG Yi<sup>1,2</sup>, YU Jianfeng<sup>1,2</sup>, CHEN Ying<sup>3</sup>

(1.School of Mechanical Engineering, Jiangnan University, Wuxi 214122, China; 2.Jiangsu Key Laboratory of Advanced Food Manufacturing Equipment & Technology, Wuxi 214122, China; 3.School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China)

**Abstract:** In order to accurately detect edible roses and their maturity in the field and realize the automatic picking of flowering roses, an improved model based on YOLOv5s was proposed to solve the problem of poor recognition accuracy caused by factors such as light and occlusion in the field. The growth state of edible roses at bud, picking and abortive flowering stages was detected. Firstly, in order to enhance the ability of multi-scale feature fusion, the feature fusion structure was improved. Secondly, multi-branch structure training was used to improve the accuracy, and the neck network C3 module was improved. Finally, in order to improve the ability of feature information extraction, a fusion attention module was added to the neck network to make the model focus on the detection target and reduce the false detection and missed detection of roses. The mean average precision of the improved model was 3.6 percentage points higher than that of the original model, reaching 90.4%, and the detection accuracy of roses in three flowering periods was improved. The results of this study provide a more accurate method for detecting edible roses at different flowering stages in unstructured environment.

**Key words:** object detection; YOLOv5s; feature fusion; attention mechanism; edible roses

收稿日期: 2023-10-09

基金项目: 国家自然科学基金项目(62173160)

作者简介: 化春键(1975-), 男, 北京人, 博士, 副教授, 主要从事机器视觉、图像处理、深度学习等研究。(E-mail) 277795559@qq.com

玫瑰是一种备受人们青睐的观赏花植物, 近年来玫瑰花的食用及药用价值受到人们的关注。玫瑰花产业发展迅速, 玫瑰鲜花饼、花茶、精油等玫瑰花产品深受消费者欢迎<sup>[1]</sup>。墨红玫瑰是常见的可食用玫瑰品种, 墨红玫瑰株高约 56 cm, 植株

较矮,花期短,采摘时效性强,需要短时间投入大量劳动力进行采摘。目前玫瑰花采摘仍然采用人工方式,相较于机械采摘费时费力<sup>[2]</sup>,且采摘效率易受天气影响。随着中国食用玫瑰花产业的发展壮大,建立高准确率的食用玫瑰花检测方法对于确定不同花期玫瑰的分布区域、检测玫瑰的花期具有重要意义。

针对非结构环境下作物识别问题,国内外开展了广泛的研究。传统的图像处理算法普遍采用颜色、纹理和边缘等特征识别作物。陈礼鹏<sup>[3]</sup>以果萼作为识别特征点,采用基于 R-G 的猕猴桃多目标果实识别方法对果实进行检测;吴超<sup>[4]</sup>对玫瑰鲜切花的俯拍及侧拍图像进行边缘提取,并结合支持向量机对玫瑰鲜切花品质进行分级预测;Shao 等<sup>[5]</sup>利用高光谱成像技术对 3 个不同生长阶段的草莓进行图像采集,通过支持向量机结合自适应加权采样方法评估草莓成熟度。近年来,研究学者将深度学习应用于作物的识别和定位中。Chao 等<sup>[6]</sup>利用改进的 YOLOv4 模型检测非结构化环境中茶菊的生长时期,在颈部网络中融合 CSPResNeXt 等模块改进 YOLOv4 模型,对茶菊萌芽期、开花早期和盛花期检测的平均精度达到 89.53%。王小荣等<sup>[7]</sup>为识别复杂环境下的红花,在 YOLOv7 检测模型中引入 Transformer 注意力机制以提高模型的检测精度,检测精确度 (mAP) 达到 88.5%。王彦钧<sup>[8]</sup>为更好地识别玫瑰花,将人工神经网络结合纹理特征识别方法对食用玫瑰进行识别检测,准确率达到 85%,但该方法对光线不敏感。张振国等<sup>[9]</sup>为解决复杂环境下红花丝的检测精度较差问题,引入 GhostNet 结构,并将 CBAM 注意力模块融入特征金字塔结构,关注重要特征的同时弱化背景、枝叶等干扰特征,提高模型的检测性能。龚惟新等<sup>[10]</sup>为提高猕猴桃花朵的检测精度,在 YOLOv5 检测模型中引入 C3HB 和注意力模块,使模型轻量化。顾满局<sup>[11]</sup>在实验室黑色背景下拍摄获得图像数据集,对图像进行预处理后采用 Resnet50 模型对俯视角度的玫瑰鲜切花分级,准确率达到 91%。

针对食用玫瑰花采摘的检测方法较少,目前相关研究主要在实验室中进行,这些检测方法在田间的检测效果较差。墨红玫瑰同一枝干上存在多朵玫瑰,彼此之间往往会互相遮挡。田间环境下,玫瑰花成熟度与图像采集效果受光照等因素的影响较大,

每朵玫瑰成熟度不一致,使得尺度变化较大。这些问题都会导致 YOLOv5 模型对被遮挡玫瑰的检测能力较差,造成误检、漏检等。基于以上问题,本研究建立非结构环境下的食用玫瑰花数据集,以 YOLOv5s 为基础网络模型,根据食用玫瑰实际种植中存在的问题对模型进行改进,为快速、精确识别可食用玫瑰花提供理论依据。

## 1 材料与方法

### 1.1 数据采集与数据集制作

本试验以墨红玫瑰花为研究对象,数据集拍摄于江苏省宜兴市美栖花田(北纬 31°23'57",东经 119°39'40"),花田面积约 40 hm<sup>2</sup>。数据采集时间为 2022 年 9 月下旬至 10 月上旬每天的 9:00–16:00,图像采集设备为 SONYDSC-W830。为识别田间非结构环境下不同花期的食用玫瑰花,在不同时间和角度拍摄玫瑰花照片 1 600 张,部分食用玫瑰花照片如图 1 所示。

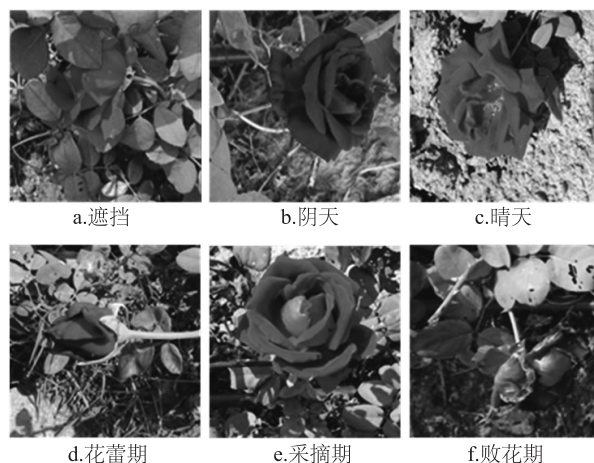


图1 田间环境下食用玫瑰花样本数据

Fig.1 Data of edible rose samples in field environment

通过图片标注工具 LabelImg 对数据集中的食用玫瑰花图像进行标注,用矩形选取框将图像中的玫瑰花框出。根据采收决策,按照花蕾期(CF)、采摘期(PF)、败花期(WF)对样本进行分类及标注,保存标注信息为 txt 格式,并将 3 个花期的玫瑰花图像划分为训练集、测试集和验证集,三者按 8:1:1 比例分配。

### 1.2 YOLOv5 网络模型

随着目标检测模型的不断发展,目前基于深度

学习的目标检测方法已经有 RCNN<sup>[12]</sup>、SSD<sup>[13]</sup>、RetinaNet<sup>[14]</sup> 和 YOLO 系列等。与 YOLOv4 相比, YOLOv5 的检测精度和推理速度较好, 目前已大量应用在农业领域<sup>[15-16]</sup>。近年来, YOLO 系列更新较快, 现在已有 YOLOv7<sup>[17]</sup>、YOLOv8 等版本<sup>[18]</sup>。YOLOv5 模型主要包括主干网络、颈部网络和预测层 3 个部分。主干网络选用基于 CSPDarkNet53 结构的骨干网络对图像特征进行特征提取。颈部网络通过特征金字塔网络 (Feature pyramid networks, FPN)<sup>[19]</sup> 和路径聚合网络 (Path aggregation network, PAN)<sup>[20]</sup> 连接主干网络与预测层, 对特征信息进行多层级的特征融合, 提高网络对不同特征层的识别能力。预测层对不同维度的特征进行预测, 得到预测结果。

## 2 改进的食用玫瑰花检测模型

### 2.1 特征融合结构改进

在卷积神经网络中, 高层特征也指语义特征, 经过多次卷积后从而具有更强的语义信息, 包含丰富的组合信息, 但其分辨率较低<sup>[21]</sup>。低层特征主要指纹理、颜色和形状等特征, 往往具有更高的分辨率, 因此目标位置信息较为准确, 可以较好反映图像的

内容。对食用玫瑰花数据集进行分析, 发现图像中需要识别的玫瑰花尺寸不一, 不少玫瑰花尺寸较小。为了提高食用玫瑰花检测模型的性能, 将部分低层特征与高层特征进行融合, 使其具有更强的语义信息和对细节的感知能力。

传统的特征金字塔网络采用自上而下的路径实现多尺度特征的融合, 如图 2 所示, 路径聚合网络在 FPN 的基础上额外增加了一个自下而上的路径, 而广义特征金字塔网络 (Generalized feature pyramid network, GFPN) 设计了一种新的路径融合, 利用 Queen-Fusion<sup>[22]</sup> 接收来自斜上方和斜下方节点的信息, 通过跨层和跨尺度连接操作, 充分发挥高层语义信息与低层空间信息之间的交换作用, 虽然采用这种方式可以增强特征交互, 但也增加了许多额外的上采样和下采样操作。使用 Concat 方法合并特征通道, 可以有效避免特征信息的丢失, 使模型在检测不同尺寸的物体时具有更好的性能。

考虑到实时检测模型的要求, 移除 Queen-Fusion 中部分上采样操作, 将颈部和骨干网络不同尺度的特征信息相融合, 实现多尺度特征融合, 以提高检测精度。

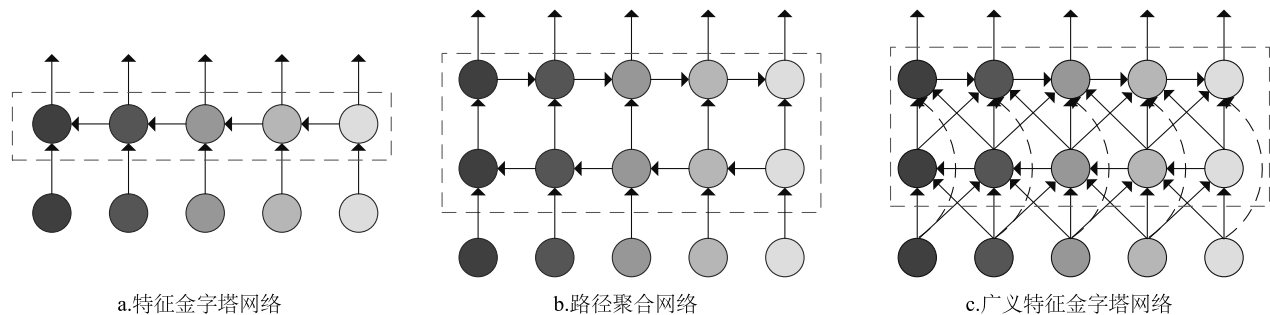


图 2 不同特征融合方式结构

Fig.2 Structure of different feature fusion methods

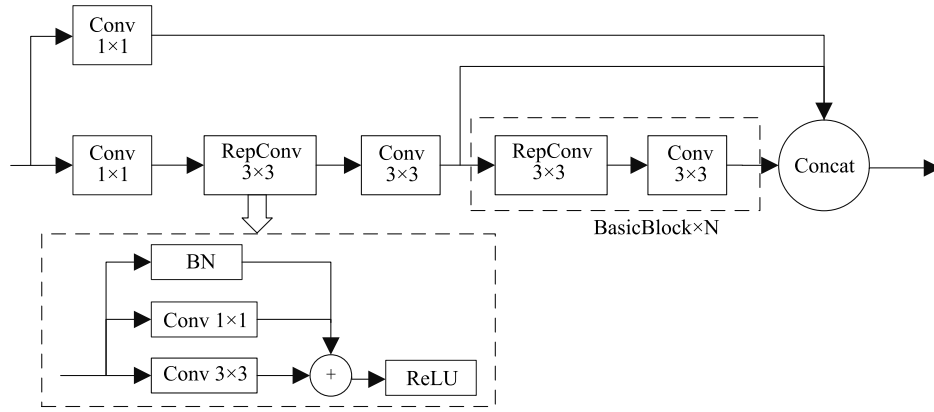
### 2.2 颈部网络改进

使用 CSPStage 结构<sup>[23]</sup> 进行特征融合, 代替基于  $3 \times 3$  卷积的特征融合方式, 提高模型的表征能力, 融合更加丰富的梯度信息。

CSPStage 模块主要由卷积模块 (Conv) 和基础模块 (BasicBlock) 组成。BasicBlock 模块由重参数化卷积 (RepConv  $3 \times 3$ ) 和  $3 \times 3$  卷积 (Conv  $3 \times 3$ ) 组成, 其中 RepConv<sup>[24]</sup> 在  $3 \times 3$  卷积基础上额外添加了并行的 1 个  $1 \times 1$  卷积 (Conv  $1 \times 1$ ) 和一个 BN (Batch

normalization) 层, 在训练时采用多路径拓扑, 在推理时将多分支融合在一起, 从而减少了推理时间, 并且具备了多分支的优点, 增加模型的表征能力, 融合更加丰富的梯度信息。

如图 3 所示, 将前置层的信息作为输入初始特征信息, 首先将输入特征分配到 2 条分支中, 利用包含卷积和 BasicBlock 模块的下分支和独立卷积模块的上分支进行特征提取, 最后使用特征融合方式 Concat 将两个分支的特征信息合并, 得到最终的输出结果。



Conv1×1;1×1 卷积模块;Conv3×3;3×3 卷积模块;ReLU:激活函数;RepConv:重参数化卷积;N:模块叠加次数;Concat:特征融合方式;BasicBlock:网络基础模块;BN:批标准化层。

图3 CSPStage 模块结构

Fig.3 CSPStage module structure

### 2.3 注意力机制模块

为了在进行特征学习时增加特征区域权重,减少背景特征权重,在网络中引入融合注意力机制模块。本研究将通道注意力模块 ECANet<sup>[25]</sup> 和空间注意力模块 GAM<sup>[26]</sup> 组合,形成融合注意力机制模块。融合注意力机制模块的结构如图 4 所示,采用通道-空间注意机制,分别关注特征的含义与位置。

对于给定输入  $F \in R^{C \times H \times W}$ ,  $C$  表示通道数,  $W$  表示输入特征图宽度,  $H$  表示输入特征图高度。首先对输入特征图  $F$  进行全局平均池化 (Global average pooling, GAP), 然后进行卷积核大小为 5 的 1 维卷积操作, 以实现跨通道交互, 提取通道之间的依赖关系。并经过 Sigmoid 激活函数得到各个通道的权重和通道权值矩阵 ( $M_c$ ), 计算公式如下:

$$M_c = \sigma[C1D_k(F)] \quad (1)$$

式中,  $M_c$  为通道权值矩阵,  $F$  为输入特征图, C1D 为一维卷积操作,  $k$  为卷积核大小,  $\sigma$  为 Sigmoid 激活函数。

将  $M_c$  与输入特征图 ( $F$ ) 按通道数相乘得到通道细化后的特征图  $F'$ , 计算公式如下:

$$F' = M_c(F) \otimes F \quad (2)$$

式中,  $F'$  为通道细化特征图,  $F$  为输入特征图,  $\otimes$  为逐元素相乘。

通道细化特征图  $F'$  被用作空间注意力模块输入, 为了聚焦空间信息, 使用 2 个  $7 \times 7$  的卷积操作进行空间信息融合, 先将通道数量减少, 然后再增加通道数量, 保持通道数量一致。经过激活函数输出, 得到空间权值矩阵 ( $M_s$ )。将  $M_s$  与通道细化特征图

( $F'$ ) 相乘得到空间细化特征图 ( $F''$ ), 计算公式如下:

$$F'' = M_s(F') \otimes F' \quad (3)$$

式中,  $F''$  为空间细化特征图,  $F'$  为通道细化特征图,  $\otimes$  为逐元素相乘,  $M_s$  为空间权值矩阵。

综合以上 3 个方面的改进, 本研究提出的食用玫瑰花检测模型如图 5 所示。该模型以 YOLOv5s 作为基础网络模型, 对网络中的特征融合结构进行优化, 并且结合 RepConv 卷积模块对颈部网络中的 C3 模块进行改进, 加强网络的表征能力, 在颈部网络末端引入融合注意力模块, 增加特征区域权重。本研究改进的网络可以实现对复杂场景中不同花期玫瑰的检测, 同时对于中小型玫瑰具有较高的检测精度。

### 2.4 评价指标

本研究主要采用精确度 (P), 召回率 (R), 单个类别平均精度 (AP) 和所有类别平均精度 (mAP) 来反映模型的训练精度; 采用参数量、运算量 (GFLOPs) 及模型权重体现模型的复杂程度; 采用帧率 (FPS) 体现模型实时检测性能。其中 P、R、AP、mAP 的计算公式如下:

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (4)$$

$$R = \frac{T_p}{T_p + F_n} \times 100\% \quad (5)$$

$$AP = \int_0^1 P(r) dr \times 100\% \quad (6)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \times 100\% \quad (7)$$

式中,  $T_p$  表示被正确框选为正样本的数量,  $F_p$



表示被错误框选为正样本的数量,  $F_n$  表示正样本被 错误框选为负样本的数量,  $N$  表示玫瑰样本类别数。

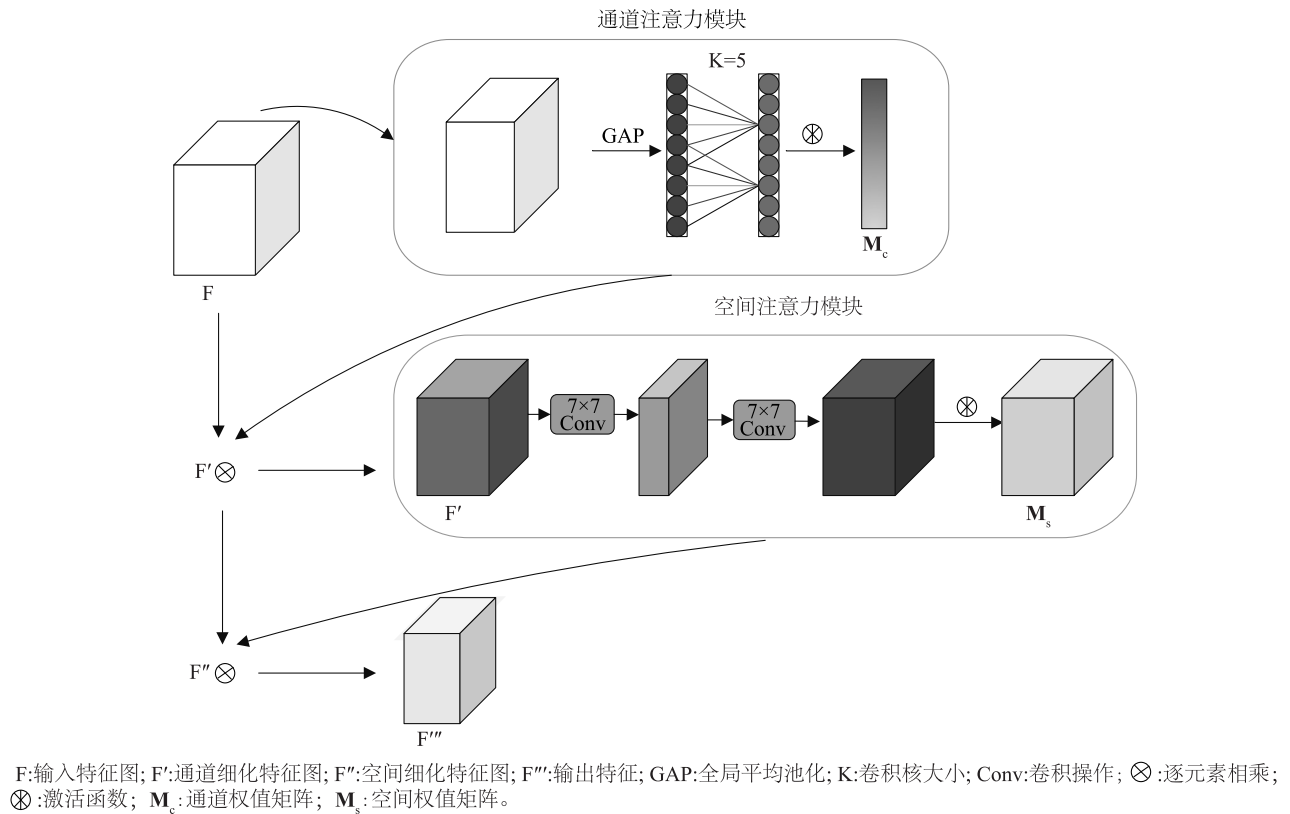


图 4 融合注意力机制结构

Fig.4 Structure of integrating attention mechanism

## 2.5 试验环境与参数设置

试验的运行环境是 64 位 Ubuntu 20.04 LTS 操作系统, CPU 为 AMD EPYC 7232P 8-Core Processor, GPU 为 NVIDIA GeForce RTX 3090, 数量为 2。采用基于 Pytorch 的深度学习框架进行训练, 网络模型以 Python3.8.10 编程语言实现, 软件平台为 Visual Studio Code, 搭配环境 CUDA11.0、Cudnn8005。

初始化学率设置为 0.01, 动量参数设置为 0.937, 批量大小设置为 32, 总迭代次数设置为 300 次, 权重衰退系数设置为 0.1, 设置图片标准尺寸为 640×640 像素。

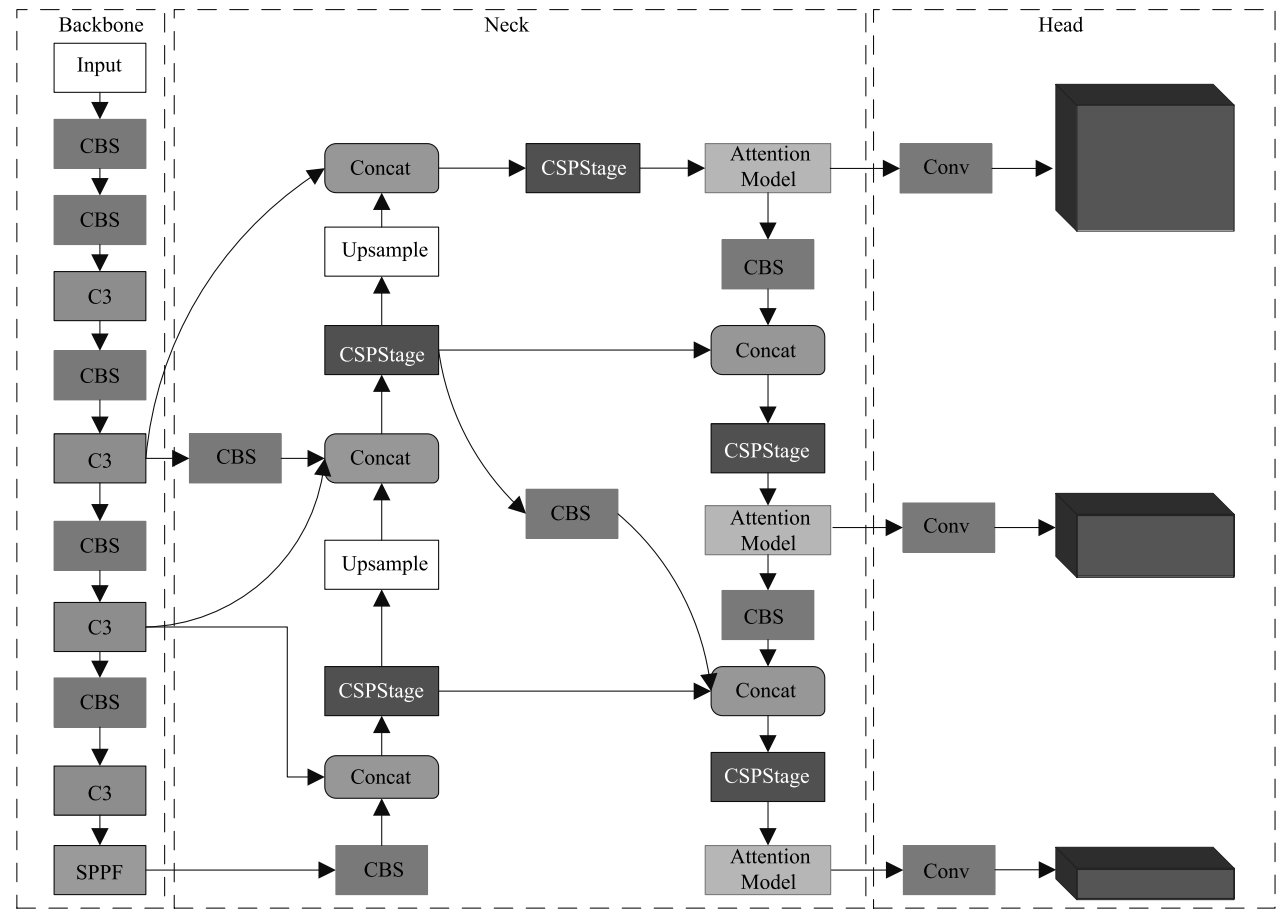
## 3 结果与分析

### 3.1 不同注意力机制对模型的影响

为验证引入的注意力机制对模型的影响, 将 SE (Squeeze-and-excitation)、CBAM (Convolutional block attention module)、ECA (Efficient channel attention) 和 CA (Channel attention), GAM (Global attention

mechanism) 等不同注意力机制添加到模型中展开试验, 试验结果如表 1 所示。

由表 1 可知, 添加不同注意力机制使模型运算量增加; 添加 SE<sup>[27]</sup> 和 CBAM<sup>[28]</sup> 注意力机制都会使模型的总体检测精度下降, 因此不考虑使用; 添加 CA<sup>[29]</sup> 和 ECA 注意力机制提高了模型对采摘期玫瑰花的检测精度, 检测精度提升了 0.7 个百分点, 但是添加 CA 注意力机制的模型对败花期玫瑰花和添加 ECA 注意力机制的模型对花蕾期玫瑰花的检测精度降低; 由于 ECA 注意力机制带来的总体性能提升更显著, 将其作为融合注意力机制模块的通道注意力组成部分, 经对比发现, 添加 ECA+BAM 注意力机制的模型对采摘期玫瑰花的识别精度与添加 ECA+GAM 注意力机制的模型一致, 达到 94.8%, 相较于原模型提升了 1.1 个百分点, 对于花蕾期的花朵识别精度比添加 ECA+GAM 的模型提升了 0.9 个百分点, 但是对败花期的花朵识别精度降低了 6.1 个百分点。



Backbone:主干网络;Neck:颈部网络;Head:预测网络;Input:输入;Conv:卷积;SPPF:空间金字塔池化结构;Upsample:上采样操作;Concat:特征融合方式;Attention Model:融合注意力机制模块。CBS:激活函数为 SiLU 的卷积模块;C3:CSP 架构的 3 卷积模块;CSPStage:改进后的 C3 模块。

图 5 改进后的网络模型结构  
Fig.5 Structure of the improved network model

表 1 不同注意力机制对模型的影响

Table 1 Effects of different attention mechanisms on the model

注意力机制	参数量 (M)	GFLOPs	mAP (%)	AP (%)		
				CF	PF	WF
Baseline	7.01	15.8	86.8	94.1	93.7	72.6
SE	7.03	16.0	84.7	93.5	93.3	67.3
CBAM	7.03	16.0	84.8	93.0	93.9	67.1
CA	7.05	16.0	86.8	94.8	94.4	71.3
ECA	7.03	16.0	87.2	93.7	94.4	73.4
ECA+BAM	7.03	16.0	87.6	96.0	94.8	72.1
ECA+GAM	7.04	16.3	89.4	95.1	94.8	78.2

Baseline:未添加注意力机制的原始模型;SE、CBAM、CA、ECA、GAM、BAM 分别为不同的注意力机制。GFLOPs:运算量;CF:花蕾期;PF:采摘期;WF:败花期;mAP:所有类别平均精度;AP:单个类别平均精度。

仅考虑注意力机制的影响,添加融合注意力机制比添加 CA 等单一注意力机制在所有类别平均精度上更具优势,添加至模型后并没有带来更多的运行负担。

并且相较于部分单一注意力机制只关注通道或空间注意力,本研究的 ECA+GAM 融合注意力机制采用顺序结构,兼顾通道和空间注意力。综合衡量各项精度和

参数量,添加 ECA+GAM 融合注意力机制提升效果更加明显,可有效提高模型特征提取能力。

### 3.2 不同改进方法对模型性能的影响

为验证上述改进后的模型在复杂背景下对食用玫瑰花采摘识别的优化效果,使用原始模型与改进后的模型在数据集上进行训练,进行了一系列消融试验,使用改进方法用“√”表示,未使用改进方法用“—”表示,试验结果见表 2。

根据表 2 可知,改进方法 2 针对制作的玫瑰花数据中存在中小型目标较多的特点,采用多尺度特征融合,将更多低层特征与高层特征进行融合,从而提高模型的特征提取能力,使得改进后的模型相较于原模型 mAP 提升了 0.9 个百分点,对 3 个花期玫瑰的识别都有提升效果,其中对花蕾期花朵单个类别平均精度提高了 1.5 个百分点;改进方法 3 使用 CSPStage 结构替代原有 C3 模块后,在训练时候采

用多路径拓扑,具有多分支的优点,使模型能够更加关注玫瑰花的特征,mAP 达到 88.2%;改进方法 4 为在原始模型中引入融合注意力机制模块,具有空间注意力和通道注意力的优势,增加特征区域权重,减少背景特征权重,提高了对败花期玫瑰花的检测精度,相较于原模型提升了 5.6 个百分点,并且总体类别平均精度达到 89.4%。改进方法 5 为将 YOLOv5s 的特征融合结构及颈部网络 C3 模块进行改进,总体类别平均精度达到 88.8%,改进方法 6 为在改进方法 5 的基础上添加融合注意力机制模块,在满足实时检测需求的情况下增加一定的参数量,使检测精度较大提升,改进后总体类别平均精度达到 90.4%,较原始模型提升了 3.6 个百分点,其中,对于花蕾期、采摘期、败花期的玫瑰花单个类别平均精度分别提高了 2.4 个百分点、2.0 个百分点、6.4 个百分点,有效提高了模型对食用玫瑰识别的精度。

表 2 不同改进方法的消融试验结果

Table 2 The ablation results of different modified methods

改进方法	特征融合结构改进	颈部网络改进	融合注意力机制模块	GFLOPs	mAP (%)	AP (%)		
						CF	PF	WF
1	—	—	—	15.8	86.8	94.1	93.7	72.6
2	√	—	—	17.9	87.7	95.6	94.8	73.1
3	—	√	—	17.1	88.2	95.0	94.9	74.7
4	—	—	√	16.3	89.4	95.1	94.8	78.2
5	√	√	—	19.5	88.8	96.2	95.3	75.0
6	√	√	√	19.9	90.4	96.5	95.7	79.0

改进方法中,1:原始模型 YOLOv5s;2:对原始模型中的特征融合结构进行改进;3:对原始模型中颈部网络中的 C3 模块进行改进;4:在原始模型中引入融合注意力机制模块;5:将原始模型 YOLOv5s 的特征融合结构及颈部网络 C3 模块进行改进;6:在改进方法 5 的基础上添加融合注意力机制模块。√表示已使用改进方法,—表示未使用改进方法。GFLOPs:运算量;mAP:所有类别平均精度;AP:单个类别平均精度。CF:花蕾期;PF:采摘期;WF:败花期。

### 3.3 不同检测模型的检测效果

为了对改进后的网络模型与现有不同模型进行对比,分析不同模型的性能,探究本研究改进模型的优越性,采用 mAP 和帧率以及模型体积为指标,选择目前主流的目标检测模型 Faster R-CNN、YOLO 系列等进行对比分析。由表 3 可知,两阶段检测模型的检测速度较慢,不适用于实时检测。而在与一阶段检测模型的对比中,本研究改进后的模型在检测目标物体时相比于 YOLOv4、YOLOv5s、YOLOv7 所有类别平均精度分别提升了 7.3 个百分点、3.6 个百分点、2.3 个百分点,具有更好的检测性能。本研究改进后的模型使得检测精度得到较大提升,模型体积和检测速率仍能够满足实时检测需要,与其

他模型相比,在检测速度方面具有一定优势,具有较好的综合性能。

表 3 不同检测模型的对比试验

Table 3 Comparative experiments of different detection models

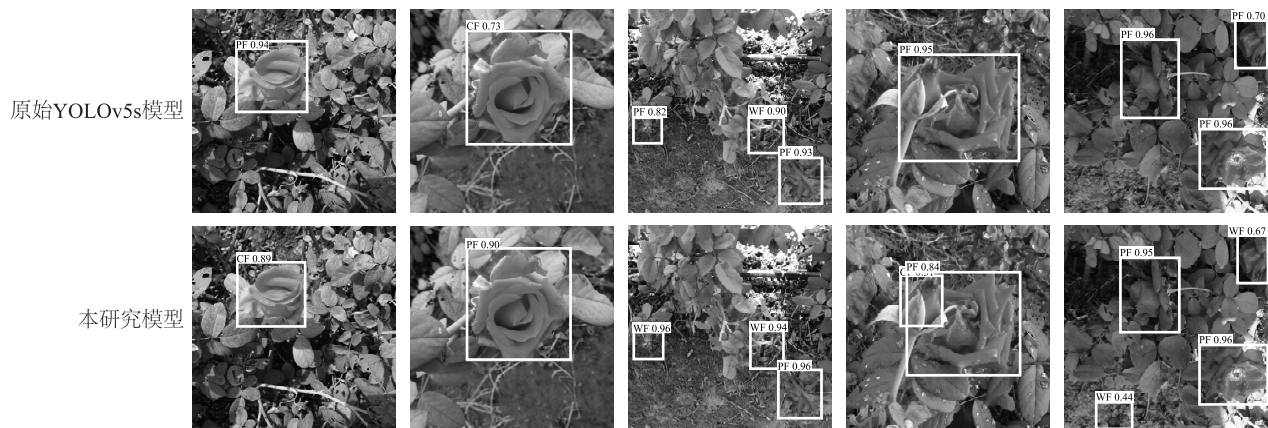
模型	体积 (Mb)	帧率	mAP (%)
Faster R-CNN	108.0	37.0	84.3
YOLOv4	244.0	59.5	83.1
YOLOv5s	14.4	110.5	86.8
YOLOv7	73.0	81.3	88.1
本研究	18.4	95.2	90.4

mAP:所有类别平均精度。

为了进一步比较改进后的模型与其他模型的检

测效果差异,分别对测试集进行检测,检测效果如图 6 所示。改进后的模型识别效果更佳,能够减少相似背景的影响,能够更加有效地识别出不同花期玫

瑰花的位置,并且对重叠目标的识别能力更强。改进后的模型漏检、误检现象减少,对不同尺度的玫瑰花检测效果更好,提升模型检测置信度。



CF:花蕾期;PF:采摘期;WF:败花期。

图 6 不同模型检测效果对比

Fig.6 Comparison of detection effects of different models

## 4 结 论

针对非结构化环境下现有检测模型对不同花期玫瑰花识别准确率低、鲁棒性差等问题,本研究以 YOLOv5s 为基础模型,将主干网络与颈部网络的特征图进行跨层连接,同时进行多尺度特征融合,使特征提取能力提升,使用改进的 CSPStage 模块对颈部网络进行优化,可以在提高精度的同时加快检测速率,并将融合卷积注意力模块添加到 Neck 网络中,采用通道空间顺序结构,提取更全面的有效特征,加强关键信息的提取,提高检测精度。根据实际种植环境,本研究制作了不同自然场景下的不同花期食用玫瑰花数据集用于模型的训练、验证和测试,并分别与 Faster R-CNN、YOLOv4、YOLOv5s 和 YOLOv7 等模型进行对比,结果表明,本研究提出的改进后的模型较原模型对于花蕾期、采摘期、败花期的玫瑰花单个类别平均精度分别提高了 2.4 个百分点、2.0 个百分点、6.4 个百分点。改进后的模型复杂度提高,但检测速率仍能达到每秒 95.2 帧,仍满足实时识别的要求。

本研究所提出的检测模型能够在复杂环境下识别出不同生长状态下的玫瑰花,减少光照以及叶片或其他花朵遮挡等问题对检测效果的影响,使采摘机器人能够更好地完成精准识别任务,为智能化采

摘和科学管理提供技术支撑。

## 参考文献:

- [1] 谢丽美,晏慧君,唐开学,等. 云南 4 个主栽食用玫瑰品种产量评价及营养成分分析[J]. 西南农业学报,2022,35(11):2627-2632.
- [2] 董万鹏,吴楠,吴洪娥,等. 不同食用玫瑰生长特性、花品质及生理变化特征[J]. 热带农业科学,2020,40(8):6-11.
- [3] 陈礼鹏. 基于机器视觉的簇生猕猴桃果实多目标识别方法研究[D]. 杨凌:西北农林科技大学,2018.
- [4] 吴超. 基于计算机视觉的玫瑰鲜切花质量分级评价[D]. 昆明:昆明理工大学,2020.
- [5] SHAO Y, WANG Y, XUAN G, et al. Assessment of strawberry ripeness using hyperspectral imaging[J]. Analytical Letters, 2020, 54(10):1547-1560.
- [6] CHAO Q, NYALALA I, CHEN K J. Detecting the early flowering stage of tea chrysanthemum using the F-YOLO model[J]. Agronomy, 2021, 11(5):834.
- [7] 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.
- [8] 王彦钧. 食用玫瑰花采摘机器人研究[D]. 昆明:昆明理工大学,2015.
- [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.
- [10] 龚惟新,杨珍,李凯,等. 基于改进 YOLOv5s 的自然环境下猕猴桃花朵检测方法[J]. 农业工程学报,2023,39(6):177-185.
- [11] 顾满局. 基于机器视觉的玫瑰鲜切花花形分类研究[D]. 昆



- 明;昆明理工大学,2023.
- [12] REN S, HE K, GIRSHICK R, et al. Faster R-CNN; towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28: 91-99.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD; single shot multibox detector[C]//EUROPEAN CONFERENCE ON COMPUTER VISION. Computer vision-ECCV 2016; 14th european conference. Amsterdam, Netherlands; Springer International Publishing, 2016.
- [14] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//IEEE. 2017 IEEE international conference on computer vision. Venice, Italy; IEEE, 2017.
- [15] 魏天宇, 柳天虹, 张善文, 等. 基于改进 YOLOv5s 的辣椒采摘机器人识别定位方法[J]. 扬州大学学报(自然科学版), 2023, 26(1): 61-69.
- [16] 邢洁洁, 谢定进, 杨然兵, 等. 基于 YOLOv5s 的农田垃圾轻量化检测方法[J]. 农业工程学报, 2022, 38(19): 153-161.
- [17] 王金鹏, 周佳良, 张跃跃, 等. 基于优选 YOLOv7 模型的采摘机器人多姿态火龙果检测系统[J]. 农业工程学报, 2023, 39(8): 276-283.
- [18] JIANG T, CHEN S. A lightweight forest pest image recognition model based on improved YOLOv8[J]. Applied Sciences, 2024, 14(5): 1941.
- [19] LIN T, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//IEEE. 2017 IEEE conference on computer vision and pattern recognition. Honolulu, HI, USA; IEEE, 2017.
- [20] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//IEEE/CVF. 2018 IEEE/CVF conference on computer vision and pattern recognition. Salt Lake City, UT, USA; IEEE, 2018.
- [21] 杨其晟, 李文宽, 杨晓峰, 等. 改进 YOLOv5 的苹果花生长状态检测方法[J]. 计算机工程与应用, 2022, 58(4): 237-246.
- [22] JIANG Y Q, TAN Z Y, WANG J Y, et al. A heavy-neck paradigm for object detection[C]//INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS. The tenth international conference on learning representations. Vienna, Austria; ICLR, 2021.
- [23] 陈范凯, 李士心. 改进 YOLOv5 的无人机目标检测算法[J]. 计算机工程与应用, 2023, 59(18): 218-225.
- [24] DING X, ZHANG X, MA N, et al. RepVGG; making vgg-style convnets great again[C]//IEEE/CVF. 2021 IEEE/CVF conference on computer vision and pattern recognition. Nashville, TN, USA; IEEE, 2021.
- [25] WANG Q, WU B, ZHU P, et al. ECA-Net; efficient channel attention for deep convolutional neural networks[C]//IEEE/CVF. 2020 IEEE/CVF conference on computer vision and pattern recognition. Seattle, WA, USA; IEEE, 2020.
- [26] 肖黎俊, 潘睿志, 李超, 等. 基于改进 YOLOv5s 绝缘子缺陷检测技术研究[J]. 电子测量技术, 2022, 45(24): 137-144.
- [27] 陈超, 齐峰. 卷积神经网络的发展及其在计算机视觉领域中的应用综述[J]. 计算机科学, 2019, 46(3): 63-73.
- [28] LU S, CHEN W, ZHANG X, et al. Canopy-attention-YOLOv4-based immature/mature apple fruit detection on dense-foliage tree architectures for early crop load estimation[J]. Computers and Electronics in Agriculture, 2022, 193: 106696.
- [29] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//IEEE/CVF. 2021 IEEE/CVF conference on computer vision and pattern recognition. Nashville, TN, USA; IEEE, 2020.

(责任编辑:成纾寒)