doi:10.6041/j.issn.1000-1298.2026.13.026

基于改进 AttentionGAN 与 YOLO v8n 的食用玫瑰夜间成熟度轻量化识别方法

杨启良[1,2]　汪刘严军[1,3]　禹　璐[1]　钟宇晖[1,4]　梁嘉平[1,2]

(1.昆明理工大学现代农业工程学院,昆明 650500;2.云南省农业水资源高效利用与智慧管控重点实验室,昆明 650500;
3.云南省智能农业工程技术与装备国际联合实验室,昆明 650500;
4.季节性旱区水-土-作物系统云南省野外科学观测研究站,昆明 650500)

摘要: 为解决云南食用玫瑰夜间自动化采摘中样本稀缺、成熟度识别精度低、小目标漏检及边缘设备算力受限等问题,本研究设计了改进 AttentionGAN 和 YOLO v8n 方法。在数据增强环节,构建基于改进 AttentionGAN 的日夜图像转换模型,引入多尺度注意力机制(Multi-scale attention)、设计渐进式融合策略,同步对生成器进行轻量化改进;在检测模型环节,以 YOLO v8n 为基础构建 YOLO-NRP 模型,采用 RVG-Ghost(RepVGG-Ghost)模块替换 C2f 主干、在主干网络末端嵌入 ECA-Spatial 注意力机制以增强关键特征表达、采用 PAFPN-Lite 优化颈部网络,强化多尺度特征融合效果。试验结果显示,改进 AttentionGAN 生成图像的 FID(Frechet inception distance)降至 87.97、SSIM(Structural similarity index)提升至 0.588 2,质量显著优于原模型,有效扩充了夜间数据集;YOLO-NRP 模型的精确率、召回率、mAP50 和 mAP50-95 分别为:91.3%、90.2%、94.3%、80.7%,所有指标均优于 YOLO v8n 基线模型,同时模型内存占用量仅 5.49 MB,推理速度达 70.26 f/s,满足边缘设备部署要求,可为自动化采摘装备研发提供可靠技术支撑。

关键词: 食用玫瑰;成熟度识别;AttentionGAN;YOLO v8n;夜间环境;轻量化设计

中图分类号: TP391.4;S641.2　**文献标识码:** A　**文章编号:** 1000-1298(2026)13-0312-15　**OSID:**

Lightweight Method for Nighttime Maturity Recognition of Edible Roses Based on Improved AttentionGAN and YOLO v8n

Yang Qiliang[1,2]　Wang Liuyanjun[1,3]　Yu Lu[1]　Zhong Yuhui[1,4]　Liang Jiaping[1,2]

(1. *Faculty of Modern Agricultural Engineering, Kunming University of Science and Technology, Kunming* 650500, *China*
2. *Yunnan Provincial Key Laboratory of Efficient Utilization of Agricultural Water Resources and Intelligent Management and Control, Kunming* 650500, *China*
3. *Yunnan Provincial International Joint Laboratory of Intelligent Agricultural Engineering Technology and Equipment, Kunming* 650500, *China*
4. *Yunnan Provincial Field Scientific Observation and Research Station of Water-Soil-Crop System in Seasonal Arid Areas, Kunming* 650500, *China*)

Abstract: Aiming to address the challenges in nighttime automated harvesting of edible roses in Yunnan, including limited samples, low maturity recognition accuracy, missed small targets, and constrained edge computing capacity, the targeted technical optimizations were conducted. In the data enhancement stage, an improved AttentionGAN-based day-night image conversion model was constructed, incorporating the multi-scale attention mechanism, designing a progressive fusion strategy, and simultaneously implementing lightweight modification of the generator. For the detection model, the YOLO-NRP model was built based on YOLO v8n: the C2f backbone was replaced with the RVG-Ghost (RepVGG-Ghost) module, the ECA-Spatial attention mechanism was embedded at the end of the backbone network to enhance key feature expression, and the neck network was optimized with PAFPN-Lite to strengthen multi-scale feature fusion. Experimental results showed that the images generated by the improved

收稿日期:2025-11-05　修回日期:2026-01-13

基金项目: 云南省基础研究计划项目(202501AW070013、202501BC070015)

作者简介: 杨启良(1978—),男,教授,博士生导师,主要从事丘陵山区智慧农业技术研究,E-mail:yangqilianglovena@163.com

通信作者: 梁嘉平(1990—),男,副教授,博士,主要从事农业信息化研究,E-mail:liangjpxaut@163.com

AttentionGAN achieved a Frechet inception distance (FID) of 87.97 and a structural similarity index (SSIM) of 0.588 2, with significantly better quality than that of the original model, effectively expanding the nighttime dataset. The YOLO - NRP model achieved 91.3% precision, 90.2% recall, 94.3% mAP50, and 80.7% mAP50 - 95, with all indicators outperforming the YOLO v8n baseline model. Meanwhile, the model weighed only 5.49 MB and had an inference speed of 70.26 f/s, meeting the requirements of edge device deployment and providing reliable technical support for the development of automated harvesting equipment.

Key words: edible rose; maturity recognition; AttentionGAN; YOLO v8n; nighttime conditions; lightweight design

0　引言

食用玫瑰是云南“云花”产业的核心高附加值作物，富含多酚、黄酮等生物活性成分，兼具食用、药用与观赏功能，在高原特色农业中占据重要地位[1-2]。然而，其采摘窗口期短（仅 3 ~ 5 d）、时效性要求高，传统人工采摘模式费时费力，难以适配产业智能化升级需求，研发具备 24 h 连续作业能力的自动化采摘机器人成为刚需。而夜间弱光环境下，花朵成熟度识别精度低、采摘点定位偏差及小目标漏检等问题，成为阻碍机器人实用化的核心技术瓶颈[3]。

为突破上述技术瓶颈，深度学习驱动的机器视觉技术为目标检测提供了有力支撑[4-6]。当前主流目标检测模型可分为二阶段与一阶段两类：二阶段网络以 Mask R - CNN[7-8] 和 Faster R - CNN[9] 为代表，通过分阶段优化候选区域与目标参数，检测精度较高，但其存在计算复杂度高、模型显存占用量大的缺陷，难以适应实时检测场景[10-12]。相比之下，一阶段检测网络凭借高效性与实时性的优势更适配农业场景，典型代表包括 RT-DETR[13] 和 YOLO 系列[14-15]等，在作物检测任务中展现出突出性能。

在农作物检测与成熟度识别领域，相关研究已取得一定进展[16-18]。然而，现有研究多聚焦于白天光照充足场景，对夜间弱光环境下的小目标检测能力不足，且在模型轻量化与检测精度的平衡上仍有优化空间。

夜间环境受光照缺陷、目标特征失真、环境干扰等因素制约，始终是农业目标检测的研究难点[19-21]。相关研究[22-23]表明，基于单阶段网络的夜间识别模型具备一定检测性能，但仍面临三重技术瓶颈：①夜间场景下目标图像采集成本高、数量稀缺，现有数据增强技术生成的夜间图像易出现模糊与颜色偏差，制约模型泛化能力。②部分模型虽能提升检测精度，但参数量与计算复杂度较高，轻量化设计不足，难以适配边缘设备。③夜间弱光环境会弱化目标特征，导致小尺寸目标漏检率较高[24]。

针对上述问题，本文以夜间食用玫瑰成熟度精准识别为核心目标，提出融合改进 AttentionGAN 与改进 YOLO v8n 的技术方案，聚焦夜间食用玫瑰成熟度识别，构建适配夜间环境的技术体系，以期为自动化采摘装备研发提供理论依据与实践参考，推动云南食用玫瑰产业的智能化升级。

1　材料与方法

1.1　数据集构建

1.1.1　数据采集

本文以云南墨红玫瑰为研究对象，植株高度为 40 ~ 120 cm，花径为 8 ~ 12 cm[25]。针对日间环境和夜间环境在云南省呈贡区昆明理工大学食用玫瑰实验基地共进行 3 轮试验数据采集。第 1 轮与第 2 轮试验分别于 2025 年 4 月 20 日 10:00—12:00 与 2025 年 4 月 22 日 14:00—16:00 进行，图像采集设备选用 Canon EOS - R7 型相机搭配适马 18 ~ 35 mm f1.8 镜头与三脚架，搭配参数为：光圈 f2.8，快门 1/200 s，感光度 100，景深 1.2 cm，距离花朵 30 ~ 40 cm，拍摄旨在选择光照强度适中、光线角度平缓的时间段采集墨红玫瑰白天的数据集，既能避开正午（12:00—14:00）过强的紫外线与强光导致的花瓣反光，又能防止清晨（09:00 前）露水未干、傍晚（16:00 后）光线骤弱造成的颜色特征模糊，可完整保留墨红玫瑰不同成熟度的花瓣颜色梯度、舒展层数及花萼形态细节。第 3 轮试验于 2025 年 4 月 24 日 21:00—23:00 进行，夜间图像采集设备额外搭配了神牛 V860 型三代闪光灯并安装柔光罩，搭配参数为：光圈 f3.5，快门 1/200 s，感光度 200，闪光灯功率 1/8，前侧 45°柔光。保存为 6 960 像素 ×4 640 像素的 JPEG 格式图像，作为本研究原始采集的高分辨率图像，完整保留食用玫瑰花瓣纹理、花萼形态等细节特征，为标注、图像转换以及模型训练提供高质量原始数据支撑。日间和夜间数据样本示例如图 1 所示。

1.1.2　数据注释与划分

参考其他花卉的生长特性及已提出的物候阶段



(a) 日间 (b) 夜间

图 1 日间和夜间食用玫瑰示例图像

Fig. 1 Daytime and night images of edible rose

划分方法[26],本研究将食用玫瑰的开放进程划分为 4 个核心阶段,分别为花蕾期、初花期、盛花期与败花期[27]。从生产效益视角来看,花蕾期玫瑰采摘后多用于晒干制成茶饮,初花期与盛花期玫瑰则可通过食品深加工实现更高附加值。基于此,本研究将初花期与盛花期食用玫瑰确定为核心采摘对象[28](这 2 个阶段的玫瑰以下统称为“采摘期(PF)玫瑰”)。对于败花期(WF)玫瑰,由于其是采摘期内未及时采收而枯萎的花朵,若未及时修剪清除,会加剧植株养分消耗,不利于玫瑰后续生长恢复,因此需对该阶段花朵进行精准识别,避免影响正常采摘玫瑰的品质[29]。此外,对花蕾期(CF)玫瑰的识别,除了能避免采摘未成熟花朵外,还具有生产规划价值,可为下一轮采摘工作的统筹安排提供数据支撑。采用 LabelImg 标注工具对食用玫瑰 606 幅图像进行人工标注,标注类别严格对应研究核心分析阶段。标注过程中,以花朵最小外接矩形为标准绘制边界框,确保框选区域完整覆盖目标花朵,同时尽量减少背景区域纳入,提升标注定位精度。标注文件格式统一为 YOLO 系列模型兼容的 TXT 格式,包含目标类别标签及边界框坐标。数据集注释类别如图 2 所示。



花蕾期(CF) 采摘期(PF) 败花期(WF)

图 2 数据集注释类别

Fig. 2 Manual labeling of edible rose

为保证模型训练的泛化能力与评估的客观性,采用随机分层划分法对 606 幅标注图像进行划分,避免单一类别样本在某一数据集过度集中。划分比例遵循农业小数据集常用的比例 7∶2∶1原则。在日间和夜间 2 种环境下,共采集到 606 幅图像,其中日间图像 404 幅,夜间图像 202 幅。样本数量分布如表 1 所示。

表 1 不同花期样本分布数量

Tab. 1 Distribution quantity of samples across different flowering stages 幅

环境	花蕾期	采摘期	败花期
日间	107	159	138
夜间	48	96	58
总计	155	255	196

1.2 基于改进 AttentionGAN 的多尺度注意力方法

针对夜间玫瑰花图像采集成本高、样本量稀缺的问题,采用 AttentionGAN 实现日夜图像域转换以扩充数据集。为适配 GAN 模型训练的输入分辨率,将 6 960 像素 ×4 640 像素高分辨率原始图像缩放得到 512 像素 ×512 像素图像,原始 AttentionGAN 处理 512 像素 ×512 像素图像时存在颜色失真、细节模糊等问题。为此设计改进架构:通过多尺度注意力分支提取多层次特征,采用渐进式融合网络自适应学习各尺度权重,结合轻量化 ResNet 主干网络降低计算开销。改进后网络架构包含 3 个关键技术:①轻量化主干网络设计,采用优化的 ResNet 残差块实现高效特征提取与重建。②多尺度注意力机制,通过全、半、1/4 分辨率 3 个并行分支及渐进式融合策略捕获多层次语义信息。③注意力引导的图像合成,利用内容掩码和注意力掩码实现自适应图像融合,各部分功能与连接关系如图 3 所示,图中,Feature Extraction 为特征提取模块,Downsample 为下采样模块,ResNetBlocks 为残差块,Upsampling 为上采样模块,Resolution Branch 为分辨率分支,Concat 为特征拼接,Fusion Net 为融合网络,Fusion Weights 为融合权重,Fused Attention 为融合后注意力,Attention Map 为注意力图,Final Fusion 为最终融合,Content Mask 为内容掩码,Attention Mask 为注意力掩码。

该改进方案的核心目标是破解夜间样本稀缺与生成质量不足的双重难题:借助三尺度注意力分支与渐进式融合网络,精准保留玫瑰花冠纹理、花萼形态等成熟度识别关键结构;通过生成器轻量化改造,在降低计算开销的同时提升图像生成效率。最终实现将 404 幅真实日间样本图像转换至夜间样本图像。

1.2.1 轻量化主干网络设计

轻量化主干网络采用编码器- ResNet -解码器三段式结构,如图 4 所示,图中 FE 为 Feature Extraction 特征提取模块;DS 为 Downsample 下采样模块;US 为 Upsample 上采样模块;ContentMask 为内容掩码。整体架构包含 3 个核心模块:编码器特征提取模块、ResNet 残差块模块和解码器上采样模块。

1.2.1.1 编码器特征提取

编码器负责将输入图像逐步压缩为高维语义特

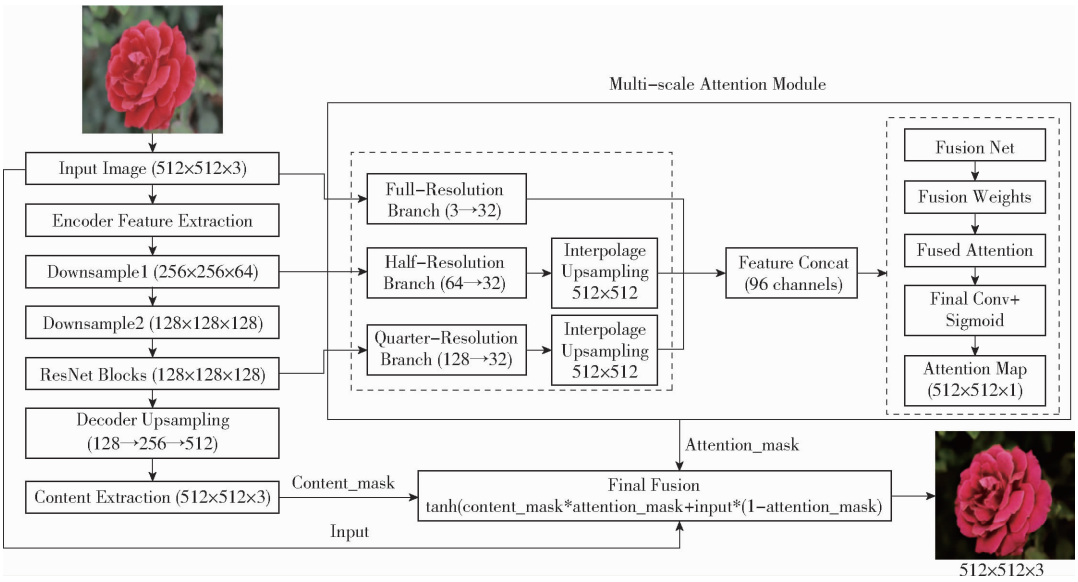

图 3　多尺度 AttentionGAN Lite 架构图

Fig. 3　Architecture of multi-scale AttentionGAN Lite

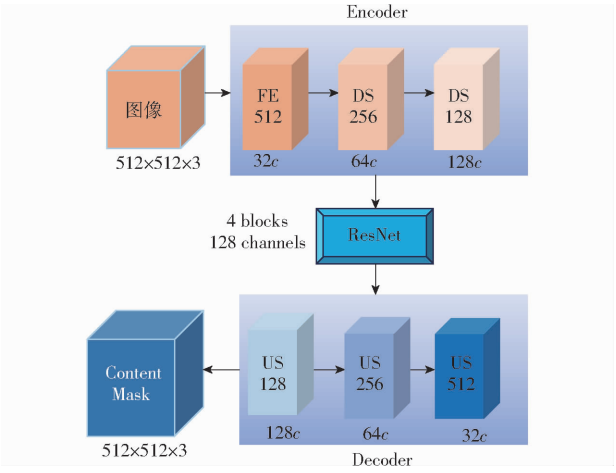

图 4　轻量化主干网络架构图

Fig. 4　Lightweight backbone network architecture diagram

征。具体而言，对于 512 × 512 × 3 的输入图像，首先通过 7 × 7 卷积核提取初始特征，将通道数从 3 扩展至 32。随后经过 2 次步长为 2 的下采样操作，特征图尺寸最终缩减为 128 × 128，通道数相应增加至 128。

与原始 AttentionGAN 相比，改进后的编码器将基础通道数由 64 降低至 32，使得 ResNet 阶段的通道数由 256 降低为 128。最终降低了卷积运算的计算量和参数量。

通过缩减基础通道数与特征图尺寸，显著降低了生成器的计算开销与显存占用量，既适配边缘设备的算力限制，又能快速完成日夜图像转换，为夜间稀缺样本的高效扩充提供技术支撑；同时通道数的合理控制避免了特征冗余，让模型更聚焦于玫瑰成熟度相关特征的学习，提升生成样本与真实夜间样本的特征一致性。

1.2.1.2　ResNet 残差块优化

ResNet 残差块是主干网络的核心计算单元，通过残差连接机制缓解深层网络的梯度消失问题。改进模型采用 4 个 ResNet 残差块，相比原始模型的 9 个残差块大幅减少了网络深度。每个残差块主要由 2 个 3 × 3 卷积层和相应的归一化层构成，并通过跳跃连接将输入直接加到输出上，其整体形式可表示为

$$\boldsymbol{F}_{\text{out}} = \boldsymbol{F}_{\text{in}} + R(\boldsymbol{F}_{\text{in}}) \tag{1}$$

式中　$\boldsymbol{F}_{\text{in}}$——输入特征

$R(\cdot)$——残差映射函数

$\boldsymbol{F}_{\text{out}}$——输出特征

残差块数量的减少原因在于玫瑰花日夜图像转换任务的语义变化相对简单，主要涉及光照和色调调整，不需要过深的网络来学习复杂的结构变换。结合多尺度注意力机制，通过注意力融合弥补网络深度减少可能带来的表达能力损失。

针对玫瑰花日夜转换仅需光照与色调调整的简单语义任务，精简残差块数量可避免网络过度复杂导致的计算冗余与过拟合，同时通过与多尺度注意力机制协同，用注意力融合弥补网络深度不足的表达能力损失，确保在降低计算成本的同时，精准保留玫瑰成熟度识别所需的关键结构特征，避免生成图像出现细节丢失或形态失真，为后续检测模型提供高质量的夜间训练样本。

1.2.1.3　解码器上采样

解码器负责将压缩的特征图逐步恢复至原始分辨率。采用 2 次转置卷积进行上采样，特征图尺寸从 128 × 128 恢复至 512 × 512，通道数相应减少至 32。最终通过 7 × 7 卷积层输出 4 通道特征图，其中

前 3 个通道表示内容掩码,用于后续的注意力引导图像合成。

通过转置卷积实现特征图的精准上采样,在恢复原始图像分辨率的同时,最大限度保留玫瑰花瓣纹理、花萼形态等细粒度特征,避免生成图像出现模糊或拉伸畸变;输出的内容掩码为注意力引导合成提供了目标域风格基础,确保生成的夜间图像在光照调整后仍能清晰呈现成熟度区分特征,让扩充的夜间样本具备与真实样本一致的判别价值,有效缓解夜间样本的稀缺。

1.2.2 多尺度注意力机制

图像的语义信息分布在不同空间尺度上,导致生成图像在整体风格和局部细节之间难以平衡。为此,本研究设计了多尺度注意力机制,如图 5 所示,通过并行处理不同分辨率的特征并自适应融合,实现对图像多层次语义的全面建模。

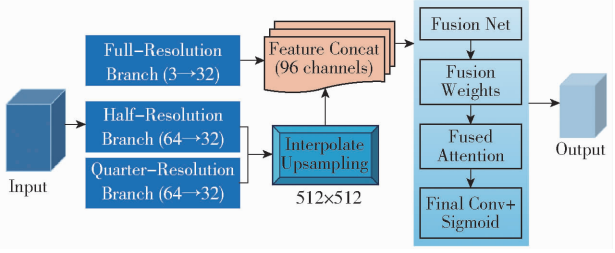

图 5 多尺度注意力网络架构图

Fig. 5 Multi-scale attention network architecture diagram

1.2.2.1 多尺度分支设计

多尺度注意力模块包含 3 个并行分支,分别处理全分辨率(512 × 512)、半分辨率(256 × 256)和 1/4 分辨率(128 × 128)的特征。全分辨率分支直接处理原始输入图像,通过 3 × 3 卷积提取局部纹理和边缘等细粒度特征;半分辨率分支接收编码器第 1 次下采样后的 64 通道特征,捕获中等尺度的结构信息;1/4 分辨率分支则处理 ResNet 输出的 128 通道深层特征,关注全局语义和整体布局。3 个分支采用统一的网络结构,输出通道数统一为 32,便于后续融合。

其次每个分支的输出特征图通过双线性插值统一上采样至分辨率 512 × 512,确保空间维度一致性。

这种多尺度并行处理策略使得网络能够同时关注图像的全局结构和局部细节,避免了单一尺度注意力的局限性,增强了模型对不同语义层次的感知能力。

通过全、半、1/4 分辨率的多分支设计,实现了对玫瑰图像的全方位特征捕获,各分支特征经统一上采样后,为后续融合提供了完整的特征支撑,让生成的夜间图像既符合夜间光照风格,又能精准保留成熟度判别所需的多尺度特征,显著提升扩充样本的质量,为检测模型提供更有效的训练数据。

1.2.2.2 渐进式融合策略

3 个尺度分支的注意力特征需要有效融合才能发挥协同作用。传统的简单加权或拼接方法难以自适应地平衡不同尺度的贡献。为此,本研究提出渐进式融合策略,通过融合网络自动学习各尺度的最优权重。

具体而言,3 个尺度的注意力特征首先在通道维度拼接,形成 96 通道的多尺度特征表示。融合网络采用轻量级设计,包含全局平均池化层(GAP)、2 个 1 × 1 卷积层和 Softmax 激活函数。全局平均池化将空间维度压缩为 1 × 1,提取每个通道的全局统计信息;第 1 个卷积层将 96 通道压缩至 16 通道,引入非线性激活函数(ReLU);第 2 个卷积层输出 3 通道的融合权重,最终经 Softmax 归一化后获得 3 个尺度的权重。融合权重 $\boldsymbol{w}$ 的计算公式为

$$\boldsymbol{w} = \mathrm{Softmax}(\boldsymbol{W}_2 \mathrm{ReLU}(\boldsymbol{W}_1 \mathrm{GAP}(\boldsymbol{F}_{\mathrm{concat}}))) \tag{2}$$

其中

$$\boldsymbol{w} = [w_1, w_2, w_3]^{\mathrm{T}} \tag{3}$$

式中 $\boldsymbol{F}_{\mathrm{concat}}$——拼接后的多尺度特征

$\mathrm{GAP}(\cdot)$——平均池化

$\boldsymbol{W}_1$、$\boldsymbol{W}_2$——2 个卷积层的权重

获得融合权重后,通过加权求和生成融合注意力特征 $\boldsymbol{F}_{\mathrm{fused}}$

$$\boldsymbol{F}_{\mathrm{fused}} = w_1 \boldsymbol{F}_{\mathrm{full}} + w_2 \boldsymbol{F}_{\mathrm{half}} + w_3 \boldsymbol{F}_{\mathrm{quarter}} \tag{4}$$

其中,$\boldsymbol{F}_{\mathrm{full}}$、$\boldsymbol{F}_{\mathrm{half}}$和 $\boldsymbol{F}_{\mathrm{quarter}}$分别为全、半、1/4 分辨率分支的注意力特征。

通过融合网络自动学习各尺度特征的最优权重,避免了人工设定权重导致的尺度失衡问题,能够根据玫瑰图像的具体特征(如遮挡程度、光照强度)自适应调整融合比例,从而让生成图像在不同夜间场景下均能保持高质量,提升与真实夜间样本的分布一致性,有效弥补真实夜间样本不足的缺陷,降低模型过拟合风险。

1.2.2.3 注意力图生成

融合后的 32 通道注意力特征经过最终卷积层转换为单通道注意力图。该卷积层包含 2 个子层:第 1 层将 32 通道压缩至 8 通道并应用 ReLU 激活,第 2 层进一步压缩至 1 通道并通过 Sigmoid 激活函数进行归一化操作。Sigmoid 输出的注意力图表示每个像素位置的重要性权重,数值越接近 1 表示该区域越需要关注和转换,数值越接近 0 则表示应保留原始输入。

生成的单通道注意力图能够精准区分玫瑰花朵主体与背景区域,为图像转换提供空间自适应指导:对花朵区域分配高权重,重点进行光照和色调调整

以适配夜间风格；对背景区域分配低权重，保留原始场景的连贯性以避免生成图像出现违和感。这种精准的空间注意力引导，有效解决了原始模型生成图像中花朵颜色失真、背景与主体风格割裂等问题，让扩充的夜间样本更贴近真实田间场景，为后续成熟度检测模型提供更具泛化能力的训练数据，间接提升检测精度。

1.2.3　注意力引导的图像合成

传统的生成对抗网络直接输出转换结果，缺乏对转换过程的精细控制，容易导致过度转换或转换不足的问题。本研究提出注意力引导的图像合成机制，通过内容掩码和注意力掩码的协同作用，实现对图像转换过程的精准控制。

1.2.3.1　双掩码协同机制

注意力引导的图像合成涉及 3 个关键要素：原始输入图像、内容掩码和注意力掩码。内容掩码由主干网络的解码器生成，包含经过域转换后的图像内容，体现了目标域的风格特征；注意力掩码由多尺度注意力模块生成，指示图像中哪些区域需要重点转换，哪些区域应保留原始特征。2 个掩码的协同作用构成了图像合成的核心机制。

内容掩码的生成过程如下：主干网络的解码器将 128 × 128 的深层特征逐步上采样至 512 × 512，并通过最终的 7 × 7 卷积层输出 4 通道特征图。前 3 个通道构成内容掩码。这一掩码包含了网络学习到的目标域风格信息。

注意力掩码则由多尺度注意力模块输出，是一个单通道的 512 × 512 特征图。该掩码在空间上指示了转换的强度：数值接近 1 的区域表示需要完全转换为目标域风格，数值接近 0 的区域则表示应保留源域特征。这种空间自适应的权重分配使得网络能够针对图像不同区域采取差异化转换策略。

内容掩码提供夜间风格的基础特征，注意力掩码限定风格转换的范围与强度，两者结合有效避免了单一掩码导致的转换不彻底、过度失真等问题。对于玫瑰成熟度识别而言，该机制能确保花朵的颜色梯度、花瓣层数等关键判别特征在转换后仍清晰可辨，同时避免背景环境的不必要干扰，让生成的夜间样本既符合检测任务需求，又具备高度真实性。

1.2.3.2　自适应图像融合

基于双掩码机制，最终的图像合成公式为

$$\boldsymbol{I}_{\text{output}} = \tanh(\boldsymbol{M}_{\text{content}}\boldsymbol{M}_{\text{attention}} + \boldsymbol{I}_{\text{input}}(1 - \boldsymbol{M}_{\text{attention}})) \tag{5}$$

式中　$\boldsymbol{I}_{\text{input}}$——原始输入图像

$\boldsymbol{M}_{\text{content}}$——内容掩码

$\boldsymbol{M}_{\text{attention}}$——注意力掩码

$\boldsymbol{I}_{\text{output}}$——最终输出图像

$\tanh(\cdot)$——激活函数，输出值归一化到[-1,1]区间，与图像预处理的归一化范围保持一致

这种融合机制的优势在于其自适应性。对于玫瑰花日夜图像转换任务，网络自动学习到：花朵主体区域的注意力权重较高，需要显著的光照和色调调整；而背景区域的注意力权重较低，保留更多原始特征以维持场景的连贯性。这种空间自适应的转换策略显著改善了生成图像的自然度和真实感。

自适应融合机制能够根据玫瑰图像的局部特征动态调整融合比例，例如对花瓣边缘等细节区域采用“内容掩码主导 + 注意力掩码强化”的策略，确保夜间场景下细节不丢失；对花萼与枝叶交界处采用“注意力掩码主导 + 内容掩码适配”的策略，避免出现边缘模糊或颜色混淆。这种精细化的融合方式，让生成的夜间图像在风格适配与特征保留之间达到最优平衡，使得扩充的样本不仅数量充足，更具备与真实样本一致的判别价值，能够有效提升后续 YOLO - NRP 模型对夜间玫瑰成熟度的识别精度，同时降低小目标漏检率。

1.3　YOLO v8n 网络改进

针对玫瑰花成熟度检测任务中小数据集过拟合风险高、边缘设备部署算力受限、多尺度特征融合不足 3 大核心问题，以 YOLO v8n 为基础框架，从主干网络轻量化、颈部特征融合优化、注意力机制增强 3 个维度进行系统性改进，构建适用于小数据集场景的高效检测模型。模型基本训练集图像为 424 幅，每个 epoch 都会对这 424 幅图像应用随机增强，300 个 epoch 的总增强数据量为 127 200 幅。改进后的模型通过各模块协同作用，在保证检测精度的同时，显著降低计算复杂度与参数量，整体架构逻辑如图 6 所示。图中，CBS 为基础卷积模块（卷积 + 批归一化 + SiLU 激活函数）；RVG - Ghost 为结合了重参数化与轻量化思想的特征提取模块；ECA - SA 为融合了高效通道注意力与空间注意力的混合注意力模块；SPPF 为快速空间金字塔池化模块；PAFPN - Lite Block 为轻量化的路径聚合网络模块；AFF 为自适应特征融合模块。

1.3.1　RVG - Ghost 主干网络设计

传统 YOLO v8n 采用 C2f 模块作为主干网络的基本单元，虽然在大规模数据集上表现优异，但在小样本场景下存在参数冗余和过拟合风险。其次由于玫瑰花的数据规模有限，本研究提出了 RVG - Ghost 混合架构，旨在通过结构重参数化和特征复用技术，在保持检测精度的同时显著降低模型复杂度。

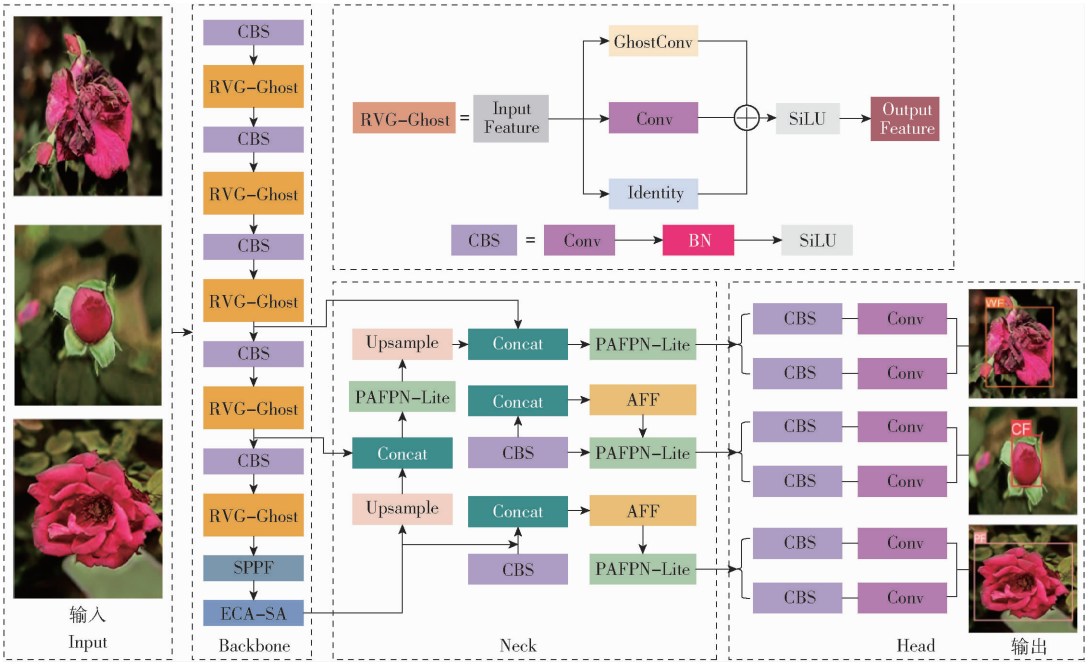

图 6 YOLO－NRP 网络架构

Fig. 6 Network architecture of YOLO－NRP

1.3.1.1 RVG 重参数化机制

RVG 架构的核心思想是训练时采用多分支结构增强模型表达能力，推理时融合为单分支结构提升运行效率。本研究在主干网络中引入 RVG 块，其训练阶段包含 3 个并行分支：3×3 卷积主分支、1×1 卷积辅助分支和恒等映射分支。这种多分支设计使得模型能够同时学习不同尺度的特征表示，其中 3×3 分支捕获局部纹理信息，1×1 分支进行通道间信息交互，恒等分支则保留原始特征以缓解梯度消失问题。

训练完成后，3 个分支可通过代数运算融合为单个 3×3 卷积层。由于每个分支都包含 BatchNorm 层，融合过程需要将 BN 参数吸收到卷积核中。设第 i 个分支的卷积核为 W_i，对应的 BN 参数为：缩放因子 γ_i、偏移量 β_i、均值 u_i、标准差 σ_i 和数值稳定项 ε，则该分支融合后的等效卷积核 $\hat{W}_i$ 和偏置 $\hat{b}_i$ 可表示为

$$\hat{W}_i = \frac{\gamma_i W_i}{\sqrt{\sigma_i^2 + \varepsilon}} \tag{6}$$

$$\hat{b}_i = \beta_i - \frac{\gamma_i u_i}{\sqrt{\sigma_i^2 + \varepsilon}} \tag{7}$$

最终，3 个分支的等效卷积核和偏置通过简单相加得到融合后的参数

$$\boldsymbol{W}_{\text{fused}} = \hat{\boldsymbol{W}}^{3\times3} + \text{pad}(\hat{\boldsymbol{W}}^{1\times1}) + \hat{\boldsymbol{W}}_{\text{identity}} \tag{8}$$

$$\boldsymbol{b}_{\text{fused}} = \hat{\boldsymbol{b}}^{3\times3} + \hat{\boldsymbol{b}}^{1\times1} + \hat{\boldsymbol{b}}_{\text{identity}} \tag{9}$$

式中 $\text{pad}(\cdot)$——将 1×1 卷积核填充至 3×3 尺寸操作函数

$\hat{\boldsymbol{W}}_{\text{identity}}$——恒等映射分支权重矩阵

$\hat{\boldsymbol{W}}^{k\times k}$——$k\times k$ 卷积分支权重矩阵

$\boldsymbol{W}_{\text{fused}}$——融合后的总权重矩阵

$\boldsymbol{b}_{\text{fused}}$——融合后的总偏置向量

$\hat{\boldsymbol{b}}^{k\times k}$——$k\times k$ 卷积分支偏置向量

$\hat{\boldsymbol{b}}_{\text{identity}}$——恒等映射分支偏置向量

这种重参数化策略使得模型在训练阶段享有多分支和 BN 带来的优化优势，而在推理阶段仅需执行单次卷积运算，有效平衡了精度与速度的矛盾。

RVG 重参数化机制针对夜间玫瑰检测的小样本场景和边缘设备部署需求，实现了训练时高精度与推理时高效率的平衡：训练阶段的多分支结构能够充分挖掘有限样本中的特征信息，缓解小样本过拟合，提升对玫瑰不同成熟度特征的区分能力；推理阶段的单分支结构则大幅减少计算步骤，降低模型推理延迟，确保在边缘设备上仍能达到 70.26 f/s 的实时性，完美适配食用玫瑰自动化采摘机器人的连续作业需求。

1.3.1.2 Ghost 卷积轻量化设计

为进一步降低计算复杂度，本研究将 Ghost 卷积机制融入 RVG 主分支。Ghost 卷积的核心理念是通过廉价操作生成冗余特征图，从而减少传统卷积的计算量。具体而言，Ghost 卷积首先使用常规卷积生成一半数量的本征特征图，然后对每个本征特征图应用深度可分离卷积生成对应的 Ghost 特征图，最后将两者拼接得到完整输出。其计算量 $FLOPs_{\text{ghost}}$ 可表示为

$$FLOPs_{ghost} = \frac{C_{out}}{2} HW(C_{in}k^2 + d^2) \quad (10)$$

式中　k——卷积核尺寸

d——深度卷积核尺寸

H、W——特征图高度和宽度

C_{in}、C_{out}——输入、输出特征图通道数

Ghost 卷积在保证玫瑰关键特征提取完整性的前提下，将主干网络的计算量降低至传统卷积的 50% 左右，参数量同步大幅缩减。这种轻量化设计精准适配了边缘设备的算力受限问题；同时，特征复用机制避免了参数冗余导致的过拟合，让模型在小样本夜间数据集中仍能保持稳定的检测精度。

1.3.1.3　RVG－Ghost 架构

本研究提出的 RVG－Ghost 块将上述 2 种技术有机结合，如图 7 所示，图中，GhostConv 为轻量化的特征提取卷积；Conv 为标准卷积；DWConv 为深度可分离卷积。在训练阶段，主分支采用 Ghost 卷积替代传统 3×3 卷积，辅助分支和恒等分支保持不变。这种设计使得模型在多分支训练优势的基础上，进一步使用 Ghost 卷积得到参数和计算量的削减。推理阶段，Ghost 卷积的 2 个子卷积层可与 RVG 的其他分支一同融合为单个 3×3 卷积，实现端到端的高效推理。

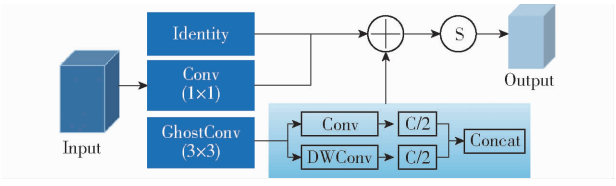

图 7　RVG－Ghost 网络结构图

Fig. 7　Network architecture of RVG－Ghost

RVG－Ghost 架构的有机融合，实现了精度、速度、轻量化的三重优化。该架构既解决了小目标漏检，又解决了边缘设备算力受限，还保证了自动化采摘所需的实时性，为夜间成熟度精准识别提供了核心硬件适配支撑。

1.3.2　PAFPN－Lite 轻量化颈部网络

传统 YOLO v8n 采用路径聚合特征金字塔网络（PAFPN）结构，通过自顶向下和自底向上的双向特征传播实现多尺度信息融合。然而，标准 PAFPN 在特征融合过程中使用了大量的 C2f 模块，在提供丰富特征的同时也带来了显著计算开销。针对玫瑰成熟度检测任务对实时性和轻量化的需求，本研究提出了 PAFPN－Lite 架构，在保持多尺度特征融合能力的同时大幅降低模型复杂度。

1.3.2.1　PAFPN－Lite 块设计

PAFPN－Lite 模块采用深度可分离卷积作为核心构建单元，将标准卷积分解为深度卷积和逐点卷积 2 个步骤。深度卷积对每个输入通道独立进行空间卷积，逐点卷积则通过 1×1 卷积实现通道间的信息交互。这种分解策略显著降低了计算复杂度和参数量。

设输入特征图通道数为 C，空间尺寸为 $H \times W$，卷积核大小为 k^2，深度可分离卷积的计算量 $FLOPs_{depthwise}$ 可表示为

$$FLOPs_{depthwise} = CHW(k^2 + C) \quad (11)$$

通过跟标准卷积计算量 $FLOPs_{standard}$ 比较可得压缩比 C_R（Compression ration）

$$C_R = \frac{FLOPs_{depthwise}}{FLOPs_{standard}} = \frac{1}{C} + \frac{1}{k^2} \quad (12)$$

当通道数 C 较大且卷积核尺寸 $k=3$ 时，压缩比可达到约 1/9，即计算量降低至原来的 11% 左右。其次，PAFPN－Lite 块在深度可分离卷积的基础上，引入了残差连接和通道注意力机制。残差连接通过跳跃连接缓解深层网络的梯度消失问题，使得网络能够学习更深层次的特征表示，最终如图 8 所示，图中，LiteBottleneck 为轻量化瓶颈模块；Residual Connection 为残差连接；Depthwise 为深度卷积；Pointwise 为逐点卷积；Channel Attention 为通道注意力；GAP 为全局平均池化；FC 为全连接层；S 为 Sigmoid 激活函数。

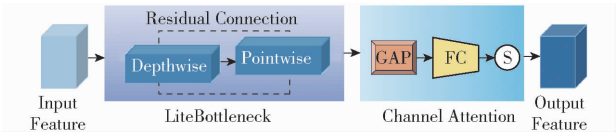

图 8　PAFPN－Lite 内部网络结构图

Fig. 8　Internal network architecture of PAFPN－Lite

PAFPN－Lite 块通过深度可分离卷积将计算量压缩至原始 PAFPN 的 11%，大幅降低颈部网络的算力消耗，适配边缘设备部署需求；同时，残差连接确保了多尺度特征在传播过程中不丢失，通道注意力机制则强化了玫瑰成熟度相关特征的响应，避免了轻量化设计导致的特征弱化。该模块保证了对夜间弱光环境下多尺度玫瑰目标的特征融合质量，为小目标的精准检出提供了关键支撑。

1.3.2.2　自适应特征融合机制

传统的特征融合方法通常采用简单的拼接或相加操作，未能充分考虑不同特征的重要性差异。本研究提出的自适应特征融合（Adaptive feature fusion，AFF）机制通过学习特征权重，实现不同尺度特征的自适应加权融合。

具体而言，AFF 模块首先对输入的多尺度特征进行全局平均池化，得到每个特征图的全局描述向

量。然后通过一个轻量级的多层感知机(MLP)网络学习每个特征的重要性权重,最后将权重应用于原始特征并进行加权求和。设输入特征为{$\boldsymbol{F}_1$,$\boldsymbol{F}_2$,…,$\boldsymbol{F}_n$},AFF 模块的输出 $\boldsymbol{F}_{\text{out}}$可表示为

$$\boldsymbol{F}_{\text{out}}=\sum_{i=1}^{n} w_i \boldsymbol{F}_i \tag{13}$$

其中

$$w_i=\frac{\exp(\text{MLP}(\text{GAP}(\boldsymbol{F}_i)))}{\sum_{j=1}^{n}\exp(\text{MLP}(\text{GAP}(\boldsymbol{F}_j)))} \tag{14}$$

式中 GAP(·)——全局平均池化操作

MLP(·)——多层感知机网络

权重通过 Softmax 归一化确保加权和为 1。这种自适应加权机制使得网络能够根据输入特征的质量动态调整融合权重,提高了特征融合的有效性。

基于 PAFPN-Lite 块和自适应特征融合机制,本研究构建了轻量化的颈部网络。该网络采用标准的 PAFPN 拓扑结构,包含自顶向下的特征金字塔路径和自底向上的特征增强路径。主干网络输出的 3 个尺度特征图(P3、P4、P5)首先经过自顶向下路径进行高层语义信息的传播,然后通过自底向上路径进行低层空间细节的增强。

在自顶向下路径中,高层特征图经过上采样后与对应尺度的主干特征图进行拼接,然后通过 PAFPN-Lite 块进行特征处理。在自底向上路径中,低层特征图经过下采样后与对应尺度的特征图进行融合,融合操作采用本研究提出的自适应特征融合机制。这种双向特征传播设计使得网络能够充分利用不同尺度的特征信息,提高对多尺度目标的检测能力。

自适应特征融合机制能够根据夜间玫瑰图像的特征质量动态分配权重,对清晰的大尺度特征分配合理权重以保证识别精度,对模糊的小尺度特征分配更高权重以强化特征响应,有效解决了传统融合方法导致的小目标漏检问题。同时,轻量级 MLP 的设计避免了额外算力开销,确保融合过程的高效性,让模型在保持实时性的前提下,提升对不同尺度、不同清晰度玫瑰目标的识别能力,尤其适配夜间弱光、枝叶遮挡等复杂场景下的成熟度判别需求。

1.3.3 ECA-Spatial 混合注意力机制

原始 YOLO v8n 在主干网络末端未引入注意力机制,导致深层特征对玫瑰花关键区域的响应不足。传统的注意力机制,如 SENet 和 CBAM 往往引入较多的参数和计算开销。针对这一问题,提出了 ECA-Spatial 混合注意力机制,通过 ECA 通道注意力和空间注意力的有机结合,在保持轻量化的同时显著提升模型的特征表达能力。

1.3.3.1 ECA 通道注意力

ECA(Efficient channel attention)高效通道注意力机制首先对输入特征图进行全局平均池化,得到每个通道的全局描述向量。然后通过一个自适应大小的一维卷积核对通道描述向量进行卷积操作,学习通道间的局部交互关系。最后通过 Sigmoid 激活函数生成通道注意力权重,并将其应用于原始特征图。设输入特征图为 $\boldsymbol{X}\in\mathbf{R}^{C\times H\times W}$,ECA 机制的输出可表示为

$$\text{ECA}(\boldsymbol{X})=\boldsymbol{X}\odot\sigma(\text{Conv1D}_k(\text{GAP}(\boldsymbol{X}))) \tag{15}$$

式中 σ——Sigmoid 激活函数

$\odot$——逐元素乘法

Conv1D_k——卷积核尺寸为 k 的一维卷积操作

该设计使 ECA 的参数量仅为 SE 注意力的约 1/5,有效避免小数据集过拟合。ECA 通道注意力以极低的参数量开销,实现了对玫瑰成熟度相关特征通道的精准校准。这种轻量化的通道注意力设计,既避免了小样本数据集下的过拟合风险,又能让模型在夜间弱光环境下快速聚焦核心特征,为成熟度精准识别奠定基础,同时适配边缘设备的算力限制。

1.3.3.2 空间注意力

Channel Attention(CA)空间注意力机制通过学习特征图不同空间位置的重要性,使网络能够聚焦于目标所在的关键区域。本研究采用的 CA 首先对输入特征图沿通道维度进行最大池化和平均池化操作,得到 2 个空间描述图。然后将这 2 个描述图沿通道维度拼接,通过一个 7×7 卷积层学习空间注意力权重。CA 机制的输出可表示为

$$\text{SA}(\boldsymbol{X})=\boldsymbol{X}\otimes\boldsymbol{M}_s(\boldsymbol{X}) \tag{16}$$

其中

$$\boldsymbol{M}_s(\boldsymbol{X})=\sigma(\text{Conv}_{7\times7}(\boldsymbol{F}_s)) \tag{17}$$

$$\boldsymbol{F}_s=[\text{AvgPool}_C(\boldsymbol{X});\text{MaxPool}_C(\boldsymbol{X})] \tag{18}$$

式中 $\text{MaxPool}_C(\cdot)$——沿通道维度的最大池化

$\text{AvgPool}_C(\cdot)$——沿通道维度的平均池化

$\text{SA}(\boldsymbol{X})$——空间注意力机制的最终输出特征图

$\otimes$——张量逐元素相乘(Hadamard 积)

$\boldsymbol{M}_s(\boldsymbol{X})$——空间注意力权重图

$\text{Conv}_{7\times7}$——卷积核尺寸为 7×7 的二维卷积操作

$\boldsymbol{F}_s$——空间维度的聚合特征

最大池化能够捕获最显著的特征响应,平均池化则提供更平滑的特征分布,两者的结合使得空间注意力能够更全面地理解特征图的空间结构。

CA 机制通过最大池化与平均池化的结合,能够

精准定位夜间玫瑰的目标区域，即使在枝叶遮挡、弱光模糊的场景下，也能快速捕获花朵的空间位置信息，强化目标区域的特征响应，抑制背景环境的干扰。对于小尺寸的花蕾期玫瑰或部分遮挡的采摘期玫瑰，该机制能有效提升目标区域的特征权重，避免其被背景噪声淹没，从而降低小目标漏检率，解决了夜间环境下玫瑰目标定位不准的问题。

1.3.3.3　ECA－Spatial 混合注意力

本研究提出的 ECA－Spatial 混合注意力机制将高效通道注意力和空间注意力串联组合，形成一个统一的注意力模块，如图 9 所示，图中，MaxPool 为最大池化；AvgPool 为平均池化；Concat 为特征拼接。

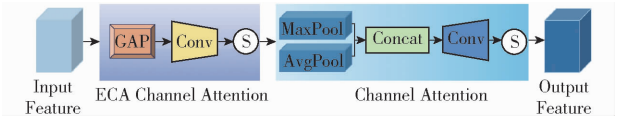

图 9　ECA－Spatial 混合注意力网络架构图

Fig. 9　Network architecture of ECA－Spatial hybrid attention

该模块首先应用 ECA 机制对输入特征图进行通道维度的重新校准，增强重要通道的响应。然后将 ECA 的输出作为空间注意力的输入，进行空间维度的特征增强。这种串联设计使得模型能够实现通道和空间 2 个维度的协同优化。ECA－Spatial 混合注意力的完整计算过程可表示为

$$\boldsymbol{Y} = \mathrm{SA}(\mathrm{ECA}(\boldsymbol{X})) \tag{19}$$

式中　ECA(·)——高效通道注意力操作

SA(·)——空间注意力操作

$\boldsymbol{Y}$——最终输出的增强特征图

这种串联结构的优势在于，通道注意力首先过滤掉冗余通道，减少了后续空间注意力需要处理的信息量，从而提高了整体的计算效率。同时，经过通道增强的特征图能够为空间注意力提供更高质量的输入，使得空间定位更加准确。

ECA－Spatial 混合注意力的串联设计实现了“通道筛选＋空间聚焦”的双重优化，先通过 ECA 过滤冗余通道，让模型聚焦于成熟度相关的核心特征通道，减少无效计算；再通过空间注意力精准定位花朵区域，强化小目标和遮挡目标的特征响应，两者协同既保证了特征质量，又提升了计算效率，让 YOLO－NRP 模型在边缘设备上仍能实现 91.3% 的精确率和 90.2% 的召回率，为夜间自动化采摘提供了精准、高效的成熟度识别支撑。

2　试验设计

2.1　试验环境与参数

本试验的训练与测试平台环境：CPU 为 Intel Core i5－12600kf，GPU 为 NVIDIA GeForce RTX 4070 SUPER 图形处理器（显存 12 GB）。软件环境基于 Windows 11 操作系统，使用 Python 3.8 作为主要编程语言，并选用 PyTorch 2.4.1 作为深度学习框架，CUDA 版本为 12.8。试验数据分为白天域数据集和夜间域数据集。训练轮次为 300 轮，训练批次（batch-size）设为 32，初始学习率为 0.01，优化器采用随机梯度下降法（SGD），权重衰减率为 0.000 5，迭代轮次为 3 900，图像输入尺寸为 640 像素×640 像素（为适配 YOLO－NRP 模型轻量化部署需求，由 6 960 像素×4 640 像素高分辨率原始图像缩放得到，在保证检测精度的同时降低计算开销）。

2.2　性能评估指标

为了评估模型性能，采用精确率（P）、召回率（R）、平均精度（AP）、平均精度均值（mAP）等多个指标来量化评估。

为评估改进 AttentionGAN 生成的样本图像质量，评价指标选用弗雷歇 Inception 距离（Frechet inception distance，FID）、学习感知图像块相似度（Learned perceptual image patch similarity，LPIPS）、结构相似性指数（Structural similarity index，SSIM）、峰值信噪比（Peak signal-to-noise ratio，PSNR）。

3　结果与分析

3.1　数据增强试验

为了证明改进 AttentionGAN 在提升夜间食用玫瑰图像生成质量、缓解夜间样本稀缺问题上的效果，本文设计了原始 AttentionGAN 与改进 AttentionGAN 的对比试验，在构建的日间与夜间食用玫瑰数据集上开展试验，通过 FID、LPIPS、SSIM、PSNR 4 项指标量化对比两者生成的夜间食用玫瑰图像质量差异。

从生成样本质量验证结果来看，各项量化指标均证实合成样本与真实夜间样本的一致性：改进 AttentionGAN 的 FID 由原始模型的 91.61 降至 87.97，表明其生成的夜间食用玫瑰图像与真实夜间样本的特征分布差异更小，图像逼真度与多样性更优；LPIPS 从 0.263 6 降至 0.252 3，意味着生成图像与真实图像在人类视觉感知层面的相似性更高，有效改善了原始模型易出现的颜色失真问题；SSIM 从 0.553 0 提升至 0.588 2，说明改进模型能更精准保留玫瑰花冠纹理、花萼形态等关键结构，结构一致性更优；PSNR 从 11.35 dB 提升至 11.81 dB，体现出生成图像的像素级误差更小，失真程度更低。整体而言，改进 AttentionGAN 显著优化了夜间食用玫瑰图像的生成质量，解决了原始模型生成图像模糊、失真的问题，可为后续 YOLO－NRP 模型训练提供高质

量的夜间样本支撑。

3.2 不同注意力机制对比分析

3.2.1 AttentionGAN 网络注意力机制对比分析

为了验证多尺度注意力设计的优势,分别融入改进 AttentionGAN、SE、CBAM、ECA 与原始 AttentionGAN 网络进行对比。AttentionGAN 网络融入不同注意力机制对比结果如表 2 所示。

表 2 不同注意力机制对比结果

Tab. 2 Comparative results of different attention mechanisms integrated

模型	FID	LPIPS	SSIM	PSNR/dB
原始 AttentionGAN	91.61	0.263 6	0.553 0	11.35
SE	89.39	0.281 8	0.552 0	10.79
CBAM	89.86	0.283 3	0.565 1	10.35
ECA	92.31	0.279 9	0.560 6	10.71
改进 AttentionGAN	87.97	0.252 3	0.588 2	11.81

从指标来看,改进 AttentionGAN 的 FID(较原始模型下降 3.64)、LPIPS(较原始模型下降 0.011 3)、PSNR(较原始模型提升 0.46 dB)、SSIM(较原始模型提升 0.035 2)4 项关键指标,均显著优于原始 AttentionGAN 及 SE、CBAM、ECA 等模型,充分体现多尺度注意力机制对生成图像质量、感知一致性与结构完整性的全面提升;而 SEModel、CBAM 仅在局部指标有小幅优化,ECA 改进效果较弱,整体表现远不及改进 AttentionGAN,这体现了改进的有效性。

该结果进一步验证了多尺度注意力机制在解决样本稀缺问题中的核心作用:通过多尺度特征提取与自适应融合,生成的高质量夜间样本可有效弥补真实夜间数据不足的缺陷,为后续成熟度检测模型提供充足且有效的训练数据支撑,避免因样本量少导致的模型过拟合问题。

3.2.2 YOLO-NRP 模型注意力机制对比分析

为验证 ECA-Spatial Attention 在 YOLO-NRP 模型中的优势,分别将 SE、CBAM、ECA、CA 和 ECA-Spatial Attention 嵌入到 YOLO-NRP 主干网络中。评价指标采用精确率、召回率、mAP50 和 mAP50-95。试验结果如表 3 所示。

表 3 YOLO-NRP 网络融入不同注意力机制对比结果

Tab. 3 Comparative results of YOLO-NRP network with different attention mechanisms integrated %

注意力机制	P	R	mAP50	mAP50-95
SE	86.2	87.0	93.6	79.5
CBAM	83.2	89.4	92.4	78.1
ECA	90.5	86.8	93.7	79.5
CA	85.3	89.4	93.8	76.8
ECA-Spatial Attention	91.3	90.2	94.3	80.7

从表 3 的指标表现来看,ECA-Spatial Attention 在各项核心检测指标中均最优:精确率达 91.3%,能有效抑制枝叶遮挡与露水反光带来的背景干扰,减少误检;召回率达 90.2%,可提升夜间弱光环境下小尺寸花蕾期玫瑰及遮挡玫瑰的检出能力,降低漏检;mAP50 达 94.3%,在常规 IoU 阈值下对食用玫瑰成熟度的识别精度更优;mAP50-95 达 80.7%,在不同 IoU 阈值下均保持稳定高精度,边界框定位鲁棒性更强。

3.3 YOLO-NRP 模型消融试验

为深入检验各改进模块对基线模型的性能提升作用,研究依次将 RVG-Ghost 主干网络、PAFPN-Lite 轻量化颈部网络及差异化注意力机制集成至 YOLO v8n 模型中,并在一致的试验环境与参数配置下设计了 6 组消融试验。其中,YOLO-NRP 网络在验证集上的消融试验结果具体如表 4 所示。

表 4 YOLO-NRP 网络消融试验结果

Tab. 4 Results of YOLO-NRP network ablation experiment %

RVG-Ghost	PAFPN-Lite	ECA-Spatial Attention	P	R	mAP50	mAP50-95
×	×	×	80.5	82.2	85.3	58.3
√	×	×	90.6	86.7	92.2	75.3
×	√	×	87.9	84.5	91.1	77.0
×	×	√	84.6	93.6	94.3	80.1
√	√	×	91.0	86.4	94.6	80.0
√	×	√	87.5	89.1	93.9	78.9
×	√	√	89.8	84.7	91.9	77.3
√	√	√	91.3	90.2	94.3	80.7

注:×表示模型不含该模块,√表示模型包含该模块。

由表 4 可知,单一模块中,RVG-Ghost 单独使用时精确率达 90.6%、mAP50 达 92.2%,能在实现主干轻量化的同时保障基础精度;PAFPN-Lite 单独使用时精确率仅 87.9%,比 RVG-Ghost 低 2.7 个百分点,mAP50 为 91.1%,比 RVG-Ghost 低 1.1 个百分点,召回率 84.5%,各项核心指标均低于 RVG-Ghost,需与其他模块协同才能发挥价值;ECA-Spatial Attention 单独使用时召回率达 93.6%,比 RVG-Ghost 高 6.9 个百分点,mAP50-95 达 80.1%,比 RVG-Ghost 高 4.8 个百分点,综合检测性能显著提升,但精确率仅 84.6%,比 RVG-Ghost 低 6.0 个百分点,易引入误检。多模块组合里,RVG-Ghost 与 PAFPN-Lite 搭配表现最优,精确率提升至 91.0%,比 RVG-Ghost 单独使用时高 0.4 个百分点,mAP50 达 94.6%,比 RVG-Ghost 单独高 2.4 个百分点,均为所有组最高,实现精度与效

率的平衡；RVG－Ghost 与 ECA－Spatial Attention 组合时，召回率降至 89.1%，比 ECA 单独使用时低 4.5 个百分点，精确率提升至 87.5%，比 ECA 单独高 2.9 个百分点，虽缓解误检率但牺牲部分召回率；PAFPN－Lite 与 ECA－Spatial Attention 组合时，mAP50－95 仅 77.3%，比 ECA 单独使用时低 2.8 个百分点，各项指标均无突破，协同效果差。综上，各模块功能定位不同，RVG－Ghost 与 PAFPN－Lite 的组合适合高精度检测场景，单独使用 ECA－Spatial Attention 适合对漏检敏感的场景，而 PAFPN－Lite 不适合单独使用或仅与注意力机制搭配。

3.4　YOLO－NRP 模型对比试验

为了进一步验证本研究提出的改进 YOLO－NRP 模型在夜间食用玫瑰采摘任务上的有效性和先进性，将其与当前主流的目标检测模型进行了横向对比。选取的对比模型包括 RT－DETR、YOLO v8s、YOLO v9n、YOLO v10n、YOLO 11、YOLO 12 以及作为基线模型 YOLO v8n。所有模型均在相同的试验环境和参数设置下，使用本研究构建的数据集进行训练和测试。性能对比结果如表 5 所示。

表 5　不同模型对比试验结果

Tab.5　Comparison experiment results of different models

模型	P/%	R/%	mAP50/%	mAP50－95/%	模型内存占用量/MB	浮点运算量	参数量	帧率/(f·s^{-1})
RT－DETR	85.1	86.5	84.8	73.1	63.11	1.08×10^{11}	3.281×10^{7}	53.86
YOLO v8n	80.5	82.2	85.3	58.3	5.97	8.20×10^{9}	3.01×10^{6}	169.83
YOLO v8s	83.0	85.9	91.7	75.6	21.50	2.87×10^{10}	1.114×10^{7}	180.37
YOLO v9n	74.2	84.1	87.0	51.9	4.44	7.90×10^{9}	2.01×10^{6}	59.70
YOLO v9s	79.8	84.1	90.3	56.2	14.55	2.74×10^{10}	7.29×10^{6}	86.29
YOLO v10n	85.3	88.0	92.1	79.1	5.52	8.40×10^{9}	2.71×10^{6}	119.68
YOLO 11	86.6	86.7	93.5	80.1	5.25	6.40×10^{9}	2.59×10^{6}	67.18
YOLO 12	84.3	85.5	91.9	77.5	5.29	6.50×10^{9}	2.57×10^{6}	65.87
YOLO－NRP	91.3	90.2	94.3	80.7	5.49	6.90×10^{9}	2.73×10^{6}	70.26

由表 5 可见，本研究提出的 YOLO－NRP 模型在夜间食用玫瑰检测任务中综合性能显著优于其他主流目标检测模型，实现了检测精度、轻量化部署与实时性的高效协同。从核心检测精度来看，YOLO－NRP 的精确率达 91.3%，较基线模型 YOLO v8n 提升 10.8 个百分点，同时高于 YOLO v10n 的 85.3%、YOLO 11 的 86.6% 与 YOLO 12 的 84.3%；召回率为 90.2%，较 YOLO v8n 提升 8.0 个百分点，优于 YOLO v9s 的 84.1%、YOLO v10n 的 88.0%、YOLO 11 的 86.7% 及 YOLO 12 的 85.5%；mAP50 达 94.3%，较 YOLO v8n 提升 9.0 个百分点，也高于 YOLO v8s 的 91.7%、YOLO v9n 的 87.0%、YOLO v10n 的 92.1% 与 YOLO 11 的 93.5%；mAP50－95 达 80.7%，较 YOLO v8n 大幅提升 22.4 个百分点，略高于 YOLO 11 的 80.1% 与 YOLO v10n 的 79.1%，体现出对不同成熟度夜间食用玫瑰更强的细节识别能力。

从轻量化特性来看，YOLO－NRP 的模型大小为 5.49 MB、参数量为 2.73×10^{6}，虽略高于 YOLO v9n 的 4.44 MB 与 2.01×10^{6}，但前者 mAP50 较后者提升 7.3 个百分点，精度优势显著，且远轻于 RT－DETR 的 63.11 MB 与 3.281×10^{7}、YOLO v8s 的 21.50 MB 与 1.114×10^{7}，更适配采摘机器人等边缘设备部署。在实时性与计算效率方面，YOLO－NRP 的浮点运算量为 6.90×10^{9}，远低于 RT－DETR 的 1.08×10^{11} 与 YOLO v8s 的 2.87×10^{10}，帧率达 70.26 f/s，高于 RT－DETR 的 53.86 f/s、YOLO v9n 的 59.70 f/s、YOLO 11 的 67.18 f/s 与 YOLO 12 的 65.87 f/s，虽低于 YOLO v9s 的 89.26 f/s、YOLO v8n 的 169.83 f/s 与 YOLO v8s 的 180.37 f/s，但结合精度提升幅度，完全满足夜间食用玫瑰自动化采摘场景的实时检测需求，整体更契合夜间食用玫瑰自动化采摘装备的技术适配需求。

为进一步验证改进模型 YOLO－NRP 的优越性，采用不同模型对夜间草莓进行检测试验，可视化结果如图 10 所示，图 10 展示了不同模型在夜间食用玫瑰检测任务中的可视化对比结果。

从检测效果来看，RT－DETR 虽然具有较高的召回率，但在复杂背景下存在误检现象；YOLO v8n 作为基线模型，对部分小目标和遮挡目标的检测置信度较低；YOLO v10n 和 YOLO 11 在检测精度上有所提升，但对不同成熟度玫瑰花的区分能力仍有待加强；YOLO 12 在轻量化与精度间取得了较好平衡，但在复杂光照条件下的鲁棒性不足。相比之下，本文提出的 YOLO－NRP 模型在所有测试场景中均表现出更高的检测置信度和更准确的边界框定位，能够准确识别并区分花蕾期（绿色框）、采摘期（红色框）和败花期（蓝色框）3 类目标，充分验证了 RVG－Ghost 主干网络、PAFPN－Lite 颈部网络和 ECA－Spatial 注意力机制的协同增效作用。

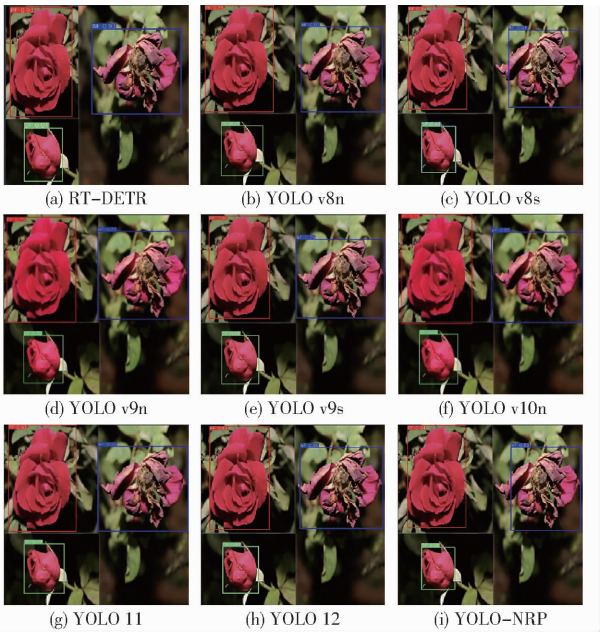

图 10 不同模型对比试验结果

Fig. 10 Comparison experiment results of different models

3.5 成熟度判别对比试验

为了更加直观地体现模型对食用玫瑰成熟度的判别效果，以精确率、召回率、mAP50、漏检率为评价指标，与主流模型进行对比试验，结果如表 6 所示。

表 6 成熟度判别对比试验结果

Tab. 6 Results of comparative experiments on maturity discrimination %

模型	标签	P	R	mAP50	MDR
RT-DETR	CF	86.5	80.3	79.4	19.7
	PF	89.1	89.0	88.0	11.0
	WF	79.5	91.2	88.1	8.8
YOLO v8n	CF	82.1	86.1	87.5	14.0
	PF	91.3	57.6	89.4	42.4
	WF	45.4	85.3	78.8	14.7
YOLO v8s	CF	81.9	85.0	91.9	15.0
	PF	94.9	68.1	94.6	31.9
	WF	75.9	85.3	91.1	14.7
YOLO v9n	CF	71.0	53.6	73.4	46.4
	PF	69.5	99.5	89.9	0.53
	WF	64.1	82.4	86.5	17.7
YOLO v10n	CF	94.3	90.6	93.5	9.4
	PF	86.6	94.2	93.5	5.8
	WF	75.4	79.4	89.2	20.6
YOLO 11	CF	92.8	87.5	93.9	12.5
	PF	89.3	87.3	95.6	12.7
	WF	77.5	85.3	91.1	14.7
YOLO 12	CF	90.2	86.2	91.7	13.8
	PF	86.6	90.9	94.8	9.1
	WF	76.2	79.4	89.2	20.6
YOLO-NRP	CF	93.3	86.5	93.6	13.5
	PF	89.6	93.5	95.2	6.5
	WF	91.1	90.6	94.2	9.4

结合表 6 中成熟度判别对比试验结果可见，YOLO-NRP 模型在食用玫瑰花蕾期、采摘期、败花期 3 类目标的识别上展现出全面且突出的性能优势，花蕾期精确率达 93.3%、mAP50 达 93.6%，漏检率控制在 13.5%，显著优于 YOLO v9n 等模型对小尺寸目标的检出效果，依托 ECA-Spatial 注意力机制对夜间弱光环境下小尺寸目标的特征响应强化，以及轻量化网络结构对特征提取过程中信息损耗的减少，有效改善了枝叶遮挡、弱光干扰下的目标漏检情况，模型整体召回率达 90.2%，较基线 YOLO v8n 提升 8.0 个百分点；采摘期召回率达 93.5%、mAP50 达 95.2%，败花期精确率与召回率均突破 90%、mAP50 达 94.2%，较 RT-DETR、YOLO 12 等主流模型进一步提升。

这一成果得益于以 RVG-Ghost 轻量化主干替换原有模块、PAFPN-Lite 颈部强化多尺度特征融合的模型优化，以及 ECA-Spatial 混合注意力机制对关键特征响应的增强，使得 3 类成熟度目标的 mAP50 均超 93%，成熟度区分的准确性大幅提升；同时，通过改进的 AttentionGAN 构建日夜图像转换模型，引入多尺度注意力分支与渐进式融合策略，404 幅真实日间样本图像被高质量转换为夜间样本图像，生成图像的 FID 从原始模型的 91.61 降至 87.97、SSIM 从 0.553 0 提升至 0.588 2，不仅有效扩充了夜间数据集的规模，更提供了多样化的高质量训练数据，为模型在复杂夜间环境下的稳定表现奠定了坚实基础。

4 结论

(1)改进的 AttentionGAN 有效解决了夜间玫瑰图像样本稀缺与生成质量不足的问题。通过设计全、半、1/4 三尺度无重叠注意力分支与渐进式融合网络，结合 ResNet 生成器残差块精简与损失权重调整，该网络生成的夜间玫瑰图像能精准保留花冠纹理、花萼形态等成熟度识别关键结构，显著优于原始 AttentionGAN 及融入 SE、CBAM、ECA 等注意力机制的模型，为后续检测模型训练提供了高质量样本支撑，缓解了夜间数据采集成本高、样本量不足的痛点。

(2)基于 YOLO v8n 优化构建的 YOLO-NRP 模型实现了夜间玫瑰检测精度、轻量化与实时性的协同提升。该模型采用 RVG-Ghost 为主干网络、PAFPN-Lite 为颈部网络，嵌入 ECA-Spatial 差异化注意力机制，并结合小数据集适配策略，在夜间墨红玫瑰检测任务中表现突出：精确率达 91.3%、召回率达 90.2%、mAP50 达 94.3%、mAP50-95 达 80.7%，较基线 YOLO v8n 分别提升 10.8、8.0、9.0、

22.4 个百分点，且模型内存占用量仅 5.49 MB、参数量 2.73×10^6、帧率 70.26 f/s，轻量化程度优于 RT-DETR、YOLO v8s 等主流模型，实时性满足边缘设备部署需求，有效解决了夜间弱光环境下小目标漏检、复杂场景鲁棒性不足及设备算力受限的问题。

(3)消融试验验证了各改进模块的协同价值。RVG-Ghost 与 PAFPN-Lite 组合时 mAP50 达 94.6%，为各模块组合最优配置，可实现精度与效率的平衡；ECA-Spatial 注意力机制单独使用时虽能提升召回率，但易引入误检，需与其他模块协同以优化精确率，表明各模块功能互补，共同支撑 YOLO-NRP 模型的高性能。

参 考 文 献

[1] 姜丽. 云南安宁食用玫瑰产业现状及可持续发展建议[J]. 中国农业大学学报，2025，30(1)：104-118.
Jiang Li. Current status of edible rose industry in Yunnan Anning and suggestions for its sustainable development[J]. Journal of China Agricultural University，2025，30(1)：104-118. (in Chinese)

[2] 姜丽，龙跃红，杨豪祥，等. 八街食用玫瑰‘滇红’的优势生物活性成分分析[J]. 中国农业大学学报，2025，30(8)：1-11.
Jiang Li，Long Yuehong，Yang Haoxiang，et al. Analysis on the superior bioactive components of Bajie edible rose ‘Dianhong’[J]. Journal of China Agricultural University，2025，30(8)：1-11. (in Chinese)

[3] 王家硕，陈炳舟，杨广召，等. 农业机器人在农业中应用[J]. 安徽农学通报，2020，26(20)：132-133.
Wang Jiashuo，Chen Bingzhou，Yang Guangzhao，et al. Application of agricultural robots in agriculture[J]. Bulletin of Anhui Agricultural Sciences，2020，26(20)：132-133. (in Chinese)

[4] 董建华，宫程翔，刘小刚，等. 基于无人机多光谱遥感的食用玫瑰土壤含水率反演模型[J]. 农业机械学报，2026，57(6)：215-226.
Dong Jianhua，Gong Chengxiang，Liu Xiaogang，et al. Soil moisture content inversion model of edible roses based on UAV multispectral remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery，2026，57(6)：215-226. (in Chinese)

[5] Singh H，Yadav P，Rishiwal V，et al. Localization in WSN-assisted IoT networks using machine larning techniques for smart agriculture[J]. International Journal of Communication Systems，2025，38(5)：e6004.

[6] 周红，刘培杰，谢秋菊，等. 基于深度视觉与类别均衡损失的母猪体况评分方法[J]. 农业机械学报，2026，57(6)：258-270.
Zhou Hong，Liu Peijie，Xie Qiuju，et al. Sow body condition assessment based on deep vision with class-balanced loss[J]. Transactions of the Chinese Society for Agricultural Machinery，2026，57(6)：258-270. (in Chinese)

[7] Shao Y，Guan X，Xuan G，et al. Detection of straw coverage under conservation tillage based on an improved mask regional olutional neural network (Mask R-CNN)[J]. Agronomy，2024，14(7)：1409.

[8] Huang X，Peng D，Qi H，et al. Detection and instance segmentation of grape clusters in orchard environments using an improved Mask R-CNN model[J]. Agriculture，2024，14(6)：918.

[9] Li J，Zhu Z，Liu H，et al. Strawberry R-CNN：recognition and counting model of strawberry based on improved Faster R-CNN[J]. Ecological Informatics，2023，77：102210.

[10] Ragab M G，Abdulkadir S J，Muneer A，et al. A comprehensive systematic review of YOLO for medical object detection (2018 to 2023)[J]. IEEE Access，2024，12：57815-57836.

[11] Vijayakumar A，Vairavasundaram S. Yolo-based object detection models：a review and its applications[J]. Multimedia Tools and Applications，2024，83(35)：83535-83574.

[12] Sapkota R，Ahmed D，Karkee M. Comparing YOLOv8 and Mask R-CNN for instance segmentation in complex orchard environments[J]. Artificial Intelligence in Agriculture，2024，13：84-99.

[13] Li X，Cai M，Tan X，et al. An efficient transformer network for detecting multi-scale chicken in complex free-range farming environments via improved RT-DETR[J]. Computers and Electronics in Agriculture，2024，224：109160.

[14] 潘鹤立，肖松，杨晓霞，等. 基于改进 YOLO v8m 的柑橘花期与花量识别方法[J]. 农业机械学报，2026，57(5)：186-196.
Pan Heli，Xiao Song，Yang Xiaoxia，et al. Dense distribution citrus flowering detection method based on YOLO v8m[J]. Transactions of the Chinese Society for Agricultural Machinery，2026，57(5)：186-196. (in Chinese)

[15] 吴利刚，陈乐，周倩，等. 基于轻量化高效层聚合网络的黄花成熟度检测方法[J]. 农业机械学报，2024，55(2)：268-277.
Wu Ligang，Chen Le，Zhou Qian，et al. Maturity detection method for *Hemerocallis citrina baroni* based on lightweight and efficient layer aggregation network[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(2)：268-277. (in Chinese)

[16] Liu Z，Xiong J，Cai M，et al. V-YOLO：a lightweight and efficient detection model for guava in complex orchard environments[J]. Agronomy，2024，14(9)：1988.

[17] 王永胜，丁宇，张若晨，等. 基于 MSCS-YOLO 的非结构化环境中草莓成熟度识别[J]. 农业机械学报，2026，57(4)：

296 – 308.
Wang Yongsheng, Ding Yu, Zhang Ruochen, et al. Strawberry maturity recognition in unstructured environments based on MSCS – YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2026,57(4):296 – 308. (in Chinese)
[18] 张宝锋,李铮,董雅文,等. YOLO v5s 模型的轻量化及其在草莓检测中的应用[J]. 制造业自动化,2025,47(10):103 – 110.
Zhang Baofeng, Li Zheng, Dong Yawen, et al. The lightweight of YOLO v5s model and its application in strawberry object detection[J]. Manufacturing Automation, 2025,47(10):103 – 110. (in Chinese)
[19] 熊俊涛,霍钊威,黄启寅,等. 结合主动光源和改进 YOLOv5s 模型的夜间柑橘检测方法[J]. 华南农业大学学报,2024,45(1):97 – 107.
Xiong Juntao, Huo Zhaowei, Huang Qiyin, et al. Detection method of citrus in nighttime environment combined with active light source and improved YOLOv5s model[J]. Journal of South China Agricultural University,2024,45(1):97 – 107. (in Chinese)
[20] 吕辉,黄杰,张明容,等. 基于 YOLO v8n 改进的夜间行车目标检测算法研究[J]. 湖南大学学报(自然科学版),2025,52(6):59 – 68.
Lü Hui, Huang Jie, Zhang Mingrong, et al. Research on night-time driving object detection algorithm based on improved YOLO v8n[J]. Journal of Hunan University (Natural Sciences),2025,52(6):59 – 68. (in Chinese)
[21] Zhang Z, Lu Y, Zhao Y, et al. Ts-yolo: an all-day and lightweight tea canopy shoots detection model[J]. Agronomy, 2023, 13(5): 1411.
[22] 卢鹏,孙天文,陈明,等. 基于 YOLO v8 和 CycleGAN 的红掌植株表型参数自动提取方法[J]. 农业机械学报,2024,55(11):154 – 159,319.
Lu Peng, Sun Tianwen, Chen Ming, et al. Automatic extraction of phenotypic parameters from *Anthurium andraeanum* Linden based on YOLO v8 and CycleGAN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(11):154 – 159,319. (in Chinese)
[23] 王新,唐灿,朱建新,等. 基于深度增强与特征抗噪的夜间串番茄成熟度识别方法[J]. 农业机械学报,2025,56(4):363 – 374.
Wang Xin, Tang Can, Zhu Jianxin, et al. Tomato cluster ripeness recognition at night based on depth enhancement and feature noise reduction[J]. Transactions of the Chinese Society for Agricultural Machinery,2025,56(4):363 – 374. (in Chinese)
[24] 王洪雁,彭俊,杨凯. 去噪扩散模型驱动的纹理增强红外–可见光图像融合方法[J]. 电子与信息学报,2025,47(6):1992 – 2004.
Wang Hongyan, Peng Jun, Yang Kai. Texture-enhanced infrared-visible image fusion approach drivenby denoising diffusion model[J]. Journal of Electronics & Information Technology,2025,47(6):1992 – 2004. (in Chinese)
[25] 黄鑫,刘晓海,张雪春,等. 墨红玫瑰营养和化学成分、体外抗氧化和护肤活性研究[J]. 食品与发酵工业,2025,51(4):205 – 213.
Huang Xin, Liu Xiaohai, Zhang Xuechun, et al. Nutritional and chemical composition, vitro antioxidant and skincare activities of Rosa 'Crimson Glory'[J]. Food and Fermentation Industries,2025,51(4):205 – 213. (in Chinese)
[26] 黄宇峰. 基于机器视觉的田间食用玫瑰识别算法研究[D]. 无锡:江南大学,2024.
Huang Yufeng. Research on field edible rose recognition algorithm based on machine vision[D]. Wuxi: Jiangnan University, 2024. (in Chinese)
[27] Zhang Z, Yang T, Liu Y, et al. Haplotype-resolved genome assembly and resequencing provide insights into the origin and breeding of modern rose[J]. Nature Plants, 2024, 10(11): 1659 – 1671.
[28] Mallick S R, Hassan J, Hoque M A, et al. Color, proximate composition, bioactive compounds and antinutrient profiling of rose[J]. Scientific Reports, 2024, 14(1): 21690.
[29] Liu X, Zhang Y, Jiang Z, et al. Micro-moistening irrigation combined with bio-organic fertilizer: an adaptive irrigation and fertilization strategy to improve soil environment, edible rose yield, and nutritional quality[J]. Industrial Crops and Products, 2023, 196: 116487.