doi:10.6041/j.issn.1000-1298.2025.11.041

基于 ESW - YOLO 的化学疏花苹果花序状态检测方法研究

陈　军[1]　曹晓明[1]　刘光耀[1]　胡广锐[2]　张新月[1]　问仕威[1]
(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100; 2. 西安工业大学设计学院, 西安 710021)

摘要: 为提高复杂果园环境下化学疏花期内小目标苹果花序的检测精度,增强检测综合性能,对 YOLO v7 - tiny 的主干网络、颈部网络以及损失函数进行改进,提出一种化学疏花时期苹果花序状态检测模型 ESW - YOLO(EMA and Slim-neck and WIoU - YOLO)。引入高效多尺度注意力(Efficient multi-scale attention,EMA)模块替换主干网络中的部分卷积层,使模型更好地关注花序,忽略复杂的背景;将 GSConv - Slim - neck 范式设计加入颈部网络,在保证精度的同时降低计算量;采用动态非单调的聚焦机制(Wise intersection of union, WIoU)改善小目标模糊不清、重叠遮挡目标难以检测的问题,提高模型检测精度。试验结果表明,该模型的精确率、召回率和平均精度均值分别为 83.02%、82.04% 和 86.83%,较 YOLO v7 - tiny 模型分别提升 2.62、3.74、3.43 个百分点,参数量减少 5.48%,与 YOLO v5s、YOLO v7 - tiny、YOLO v8s 相比均有所提升。ESW - YOLO 模型在不同光照条件(顺光、逆光)与果树种类(红富士、瑞雪、瑞香红)的检测中 F1 值均高于 YOLO v7 - tiny 模型,且检测示例的 F1 值均大于 84%。采用 Score - CAM 方法对模型主干网络与颈部网络的输出层进行热力图可视化分析,实现了模型的可解释性。

关键词: 苹果花序; 化学疏花; 目标检测; YOLO v7 - tiny; 可视化
中图分类号: S24; TP391　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)11-0433-08　　**OSID**:

Research on Detection Method of Apple Inflorescence Status of Chemical Flower Thinning Based on ESW - YOLO

CHEN Jun[1]　CAO Xiaoming[1]　LIU Guangyao[1]　HU Guangrui[2]　ZHANG Xinyue[1]　WEN Shiwei[1]
(1. *College of Mechanical and Electronic Engineering*, *Northwest A&F University*, *Yangling*, *Shaanxi* 712100, *China*
2. *School of Design*, *Xi'an Technological University*, *Xi'an* 710021, *China*)

Abstract: Aiming to improve the detection accuracy of small apple inflorescences during the chemical thinning period in complex orchard environments and enhance overall detection performance, an improved YOLO v7 - tiny based model named EMA and Slim-neck and WIoU - YOLO (ESW - YOLO) was proposed. The backbone network, neck network, and loss function of YOLO v7 - tiny were modified. An efficient multi-scale attention (EMA) module was introduced to replace some convolutional layers in the backbone, enabling the model to focus more effectively on inflorescences while ignoring complex backgrounds. The GSConv - Slim - neck paradigm was incorporated into the neck network, reducing computational costs without compromising accuracy. Additionally, a dynamic non-monotonic focusing mechanism, wise intersection of union (WIoU), was adopted to address challenges in detecting small, overlapping, and occluded objects, thereby improving detection precision. Experimental results demonstrated that the proposed model achieved precision (P) of 83.02%, recall(R) of 82.04%, and mean average precision (mAP) of 86.83%, representing increases of 2.62, 3.74, and 3.43 percentage points respectively compared with that of the YOLO v7 - tiny model, while reducing the number of parameters by 5.48%. ESW - YOLO outperformed YOLO v5s, YOLO v7 - tiny, and YOLO v8s in all metrics. Furthermore, in detection tasks under different lighting conditions (front light, back light) and for different apple varieties (Red fuji, Ruixue, Ruixianghong), ESW - YOLO consistently achieved F1 scores higher than YOLO v7 - tiny, with all detection examples scoring above 84%. The Score - CAM method was used to visualize the heatmaps of the output layers in the model's backbone and neck networks, enhancing the interpretability of the model.

Key words: apple inflorescence; chemical flower thinning; object detection; YOLO v7 - tiny; visualization

收稿日期: 2024 - 07 - 31　修回日期: 2024 - 09 - 20
基金项目: 国家自然科学基金项目(32272001)和陕西省秦创原"科学家 + 工程师"队伍建设项目(2023KXJ - 016)
作者简介: 陈军(1970—),男,教授,博士生,主要从事智能化农业装备技术研究,E-mail: chenjun_jdxy@nwsuaf.edu.cn

0 引言

疏花是苹果园管理中的关键措施之一[1],旨在提升果实质量、减少养分消耗并确保持续稳定的高产。化学疏花方式因节省劳动力、效率高并且速度快而被大规模使用[2],但疏花时机需要人工判断[3]。因此,进行化学疏花时期的苹果花序状态检测研究对于果园智能化疏花具有重要意义。

在智能化果园管理中,基于图像信息的苹果花序生长状态检测受到了许多学者关注[4]。早期的目标检测模型大多是基于颜色和阈值进行识别,其局限性在于对光照条件依赖性较强,模型鲁棒性较差,且阈值参数的调整为人工判断[5]。这类方法难以在自然环境下区分不同物候时期和形状轮廓的苹果花序,因此在生产实践中并不适用[6]。

深度学习方法在处理复杂自然场景时,展现出更为出色的分类与检测能力[7]。目前检测模型主要包括二阶段检测模型与一阶段检测模型[8],以 Mask R-CNN[9]、Faster R-CNN[10]等为代表的二阶段检测模型首先对目标进行感兴趣部分提取,然后再进行分类以及边界框回归,以 SSD[11]和 YOLO[12]为代表的一阶段模型可以直接预测类别以及边界框定位。TIAN 等[13]采用 U-Net 主干改进 Mask Scoring R-CNN 的模型检测并分割苹果花序的花蕾、半开放和完全开放 3 个生长阶段;DIAS 等[14]结合颜色和形态信息提出一种 CNN + SVM 模型分割图像中的苹果花;MU 等[15]开发了一种基于 Mask R-CNN 的模型识别并计算苹果花中心花开放百分比。

尽管二阶段目标检测模型具有较高的检测精度,但由于其参数量大,处理速度较慢,难以满足实时检测任务需求[16]。一阶段目标检测模型牺牲了一定的准确率但具有更快的检测速度,近年来 YOLO 模型不断发展与改进,检测准确率有较大提升,逐渐被用于实时检测中[17]。陈国防[18]通过加入 CA 注意力机制与 GhostBottleneck 模块、更改 K-means++聚类方法以及 CIoU 损失函数对 YOLO v5 模型进行改进,使用该改进模型对 30~50 cm、100~150 cm 内的苹果花序检测具有较高精度,且检测速度较快;司永胜等[19]通过添加 C-CoTCSP 与 RFB 结构并更改 VariFocal Loss 损失函数对 YOLO v5s 模型进行改进,改进后的模型对苹果中心花与边花的检测性能均有所提高;高昂等[20]通过加入 CBAM 注意力机制、更改不同阶段的损失函数对 YOLOX-s 进行改进并对苹果花序进行检测,取得了较高的精度。

上述研究对苹果花的检测精度较高,但仍难以满足疏花期花序状态检测需求:①检测的苹果花图像均为近距离检测,参与训练与检测的每幅图像中只有少量花序,不能代表化学疏花时期果园内苹果花序状态。②在图像拍摄时基本为俯拍视角,除花序外只包含棕色枝条、地面以及绿色树叶等与白粉色苹果花颜色区分十分明显的背景,不适用于果园复杂背景下跟随化学疏花机工作时的花序检测。综合上述研究,一阶段模型在拥有较高检测精度下模型更小,可以部署至移动端,实现果园检测任务。

本文拟构建化学疏花时期苹果花序状态数据集,提出苹果花序状态检测模型,对现有的 YOLO v7-tiny模型进行改进,使其更适用于复杂田间环境以及小目标检测特点,为确定最佳化学疏花期以及实现精准化学疏花、减少药液浪费提供技术支撑。

1 材料与方法

1.1 数据采集

本文苹果花序图像采集自陕西省宝鸡市凤翔区雨嘉果园,该果园为标准化矮砧密植类果园,果树行间距约 3.5 m。本次采集的苹果品种为红富士、瑞雪和瑞香红,果树年龄分别为 8 年、5 年、4 年,采集方式为距离果树约 1.75 m 处(即行中间)手持设备拍摄,在 10:00—12:00、12:00—15:00、15:00—17:00 拍摄顺光逆光图像,构建化学疏花期苹果花序状态数据集。图像中包含当前树行互相遮挡及树枝树叶遮挡的苹果花序,遮挡程度不同,除完全遮挡外的当前树行正面苹果花序均需要检测,当前树行的背面苹果花序与背景树行中的苹果花序不进行检测,采集图像示例如图 1 所示。

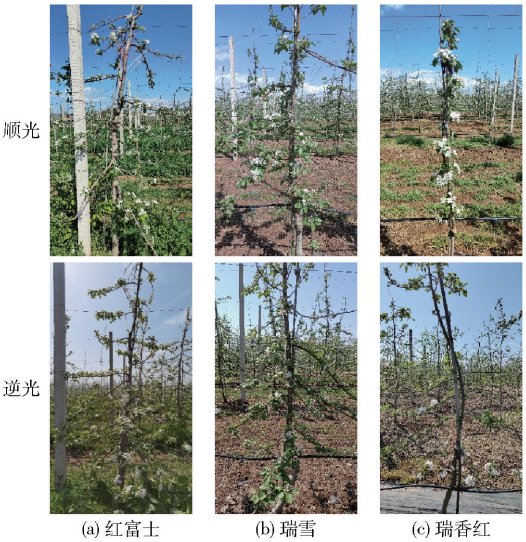


图 1 采集图像示例

Fig. 1 Example of collected images

1.2　数据集预处理

苹果花序中花朵分为中心花与边花，如图 2 所示，在生长过程中，中心花先开、边花后开，化学疏花通过保留中心花和疏除边花来提高坐果品质。

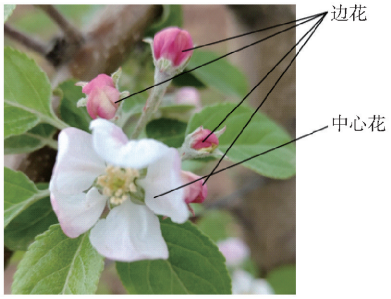


图 2　苹果花序示例

Fig. 2　Example of an apple inflorescence

将标注的苹果花序划分为 2 类标签，分别为中心花未开放（Before blossoming，BB）与中心花已开放（Blossoming stage，BS），图 3 为苹果花序分类示例[21]。将采集到的 2 530 幅数据集通过 LabelImg 软件进行标注，标注示例如图 4 所示，将标签文件保存为. xml 格式。标注完成的数据集内中心花未开放的标签数量为 35 669 个，中心花已开放的标签数量为 67 502 个。



(a) 中心花未开放　　(b) 中心花已开放

图 3　苹果花序分类示例

Fig. 3　Example of apple inflorescence classification



图 4　LabelImg 标注示例

Fig. 4　Example of LabelImg annotation

由于采集数据集时间为化学疏花期，中心花已开放的花序数量偏多，采集到的 BS 标签占总标签数量的近 2/3，造成分类标签数量不平衡。对少数类标签进行扩增，从 2 530 幅图像中挑选 BB 标签数量占 BB 标签数量与 BS 标签数量之和的比例大于 70% 的图像共 354 幅。图像数据增强方式包括旋转、平移、改变亮度、加入噪声以及翻转，将挑选出的图像通过随机添加一种或多种数据增强方式进行数据集扩增，扩增后的数据集图像数量为 4 115 幅，其中 BB 标签数量为 112 669 个，BS 标签数量为 87 817 个，将 4 115 幅图像按比例 7∶2∶1 划分为训练集、验证集与测试集。

为确认化学疏花期内果树状态，采用随机抽样方式，在化学疏花期内采集红富士、瑞雪和瑞香红 3 种果树各 100 幅图像，统计每幅图像中心花开放的花序数量占全部花序数量的百分比并记为 F，生成箱型图如图 5 所示。

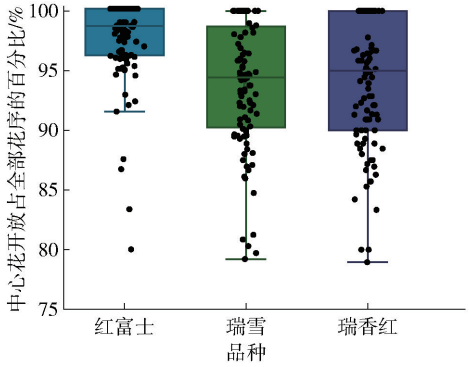


图 5　3 种果树中心花开放花序数量占比图

Fig. 5　Proportion of inflorescences with central flower open in three fruit tree species

由统计可得采集的红富士、瑞雪、瑞香红果树图像中 F 值超过 88% 的图像分别有 96 幅、89 幅以及 85 幅，数量均达 85%，因此本文判定果树进入化学疏花期的标准为 $F \geqslant 88\%$，此时为化学疏花机对苹果树进行疏花作业的最佳时期。

1.3　YOLO v7－tiny 模型改进

YOLO v7 模型保持较高检测精度的同时有着较快的运行速度，常运用于实时检测[22]。本文选用轻量级的 YOLO v7－tiny 作为基础模型，YOLO v7－tiny 网络模型结构分为 4 部分：输入部分（Input）、主干网络（Backbone）、颈部网络（Neck）和头部网络（Head）。在输入部分，图像经过数据增强等预处理，主干网络负责对图像中的特征进行提取，颈部网络融合各层特征并输出不同尺度的特征层，头部网络检测图像中的特征，得到对应的目标框与标签。

在 YOLO v7－tiny 模型的主干网络部分引入高效多尺度注意力（Efficient multi-scale attention，EMA）模块，使特征提取部分更好地关注花序，忽略复杂的背景，使用 GSConv－Slim－neck 范式设计方式降低模型的浮点计算量与参数量，使其能更好地运用至边缘设备，更改损失函数为动态非单调的聚焦机制（Wise intersection of union，WIoU），可以根据目标类别的频率自适应调整各个类别的权重，提

高网络模型的泛化能力和整体性能。本文提出的 ESW - YOLO(EMA and Slim-neck and WIoU - YOLO)网络结构如图 6 所示。

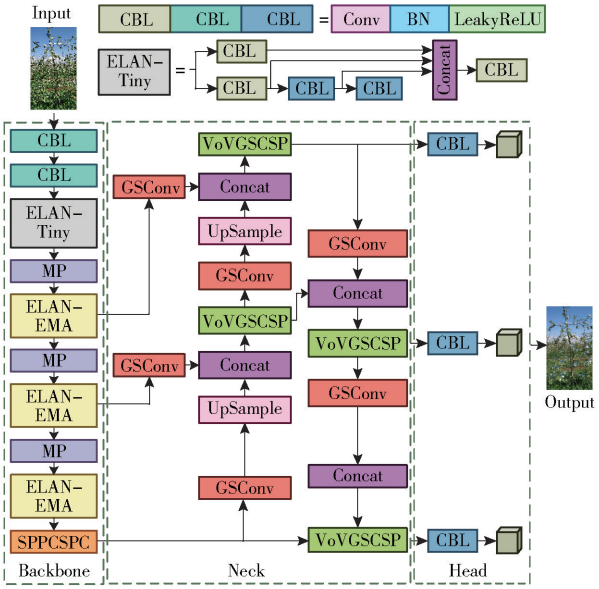


图 6 ESW - YOLO 网络结构图

Fig. 6 ESW - YOLO network architecture diagram

本文主要从以下方面进行改进:

(1)EMA 模块

苹果园内采集到的数据集图像中需要检测的重叠花序较多,在进行检测时目标像素较低,通用检测模型对小目标的适应能力相对较差,因此本文在 YOLO v7 - tiny 的主干网络中加入 EMA 模块[23],提高重要特征的权重,降低对树叶、枝干以及天空等复杂背景信息的关注,从而提高目标检测精度。EMA 模块由并行子网与跨空间学习模块组成,其结构如图 7 所示,该注意力模块将通道维度划分为多个子特征组,采用并行子网的方式收集多尺度空间信息,通过跨空间学习的方式将并行子网的输出进行全局空间信息编码,并对并行子网 2 个分支的输出特征进行点积,将 2 个输出结果进行加权。

YOLO v7 中高效层聚合网络(Efficient layer aggregation networks,ELAN)由多个卷积层(Conv + BN + LeakyRelU,CBL)多分支堆叠构成,YOLO v7 - tiny 中使用的是 ELAN - Tiny 模块,由 5 个卷积模块组成,将 EMA 模块添加到主干网络后 3 个 ELAN - Tiny 结构末端如图 8 所示,将普通的卷积替换为 EMA 模块,通过提取不同尺度的特征以及空间位置信息感知能力,使输入到特征融合部分的特征图保留更多的上下文信息,提高模型检测精度。

(2)GSConv - Slim - neck 范式设计

GSConv(Grouped spatial convolution)与其相对应的设计 Slim - neck 在自动驾驶领域的实时计算平台中有着良好表现[24]。本文将其运用至实时疏

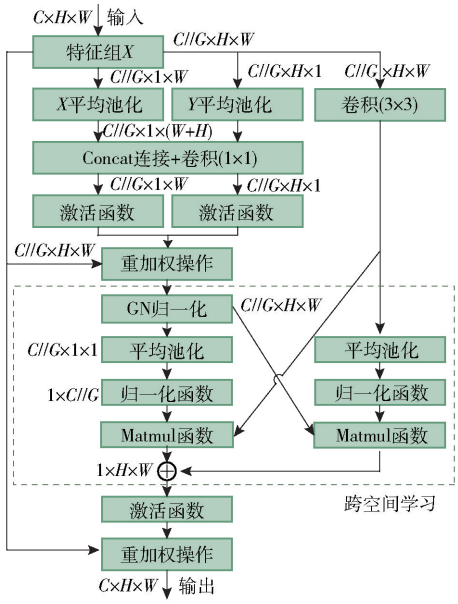


图 7 高效多尺度注意力模块

Fig. 7 Efficient multi-scale attention module

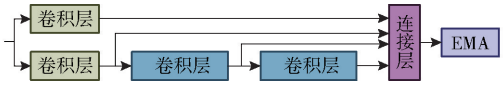


图 8 ELAN - EMA 结构图

Fig. 8 ELAN - EMA structure diagram

花作业中,将 GSConv - Slim - neck 范式设计与 YOLO v7 - tiny 模型相结合。

GSConv 结构如图 9a 所示,GSConv 加入混洗(shuffle)操作,将标准卷积(Standard convolution,SC)后的信息完全混合在主干网络经过深度可分离卷积(Depth-wise separable convolution,DSC)的输出中,缓解了轻量级算法中 DSC 操作由于输入图像的通道信息在计算中会分离导致的特征提取与融合能力下降的缺陷,提高模型感受野大小与语义理解能力。以 GSConv 为基础卷积网络,设计 GSbottleneck 结构,围绕 GSbottleneck 结构设计跨阶段部分网络 VoV - GSCSP,如图 9b 所示。将 GSConv 与 VoV - GSCSP 进行组合运用在 YOLO v7 - tiny 的颈部网络,将 YOLO v7 - tiny 颈部网络中的标准卷积替换为 GSConv,ELAN - Tiny 结构替换为 VoV - GSCSP 结构,使颈部网络保证重要通道连接的同时减少网络参数量与计算量,保证检测精度并提高模型的速度和效率。

(3)Wise - IoU 损失函数

YOLO v7 - tiny 采用的损失函数是 CIoU(Complete intersection over union)[25],该损失函数将中心点距离和宽高比惩罚项纳入考虑范围,评估预测框与真实框相似性学习目标的特征。然而苹果花序因其目标较小、生长形态复杂以及容易被遮挡导致特征较差,宽高比的度量方式会加剧低质量示例

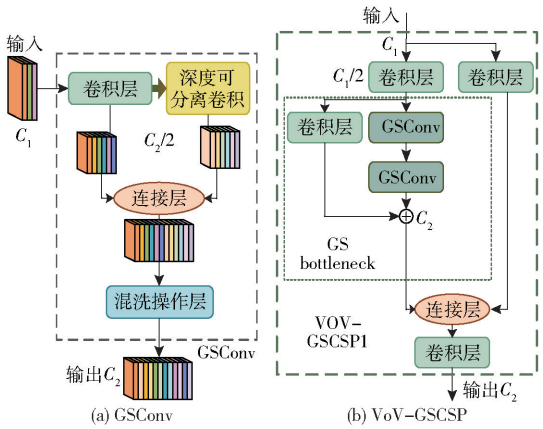


图 9　GSConv 与 VoV－GSCSP 结构图

Fig. 9　GSConv and VoV－GSCSP architecture diagrams

的惩罚，导致模型泛化能力下降。

WIoU[26] 构建基于注意力的锚框损失，弱化宽高比惩罚项对模型训练的影响。WIoU 引入离群度评估锚框质量，动态地给予高质量与低质量锚框一个小的梯度增益，降低高质量锚框竞争力的同时减少低质量锚框产生的有害梯度，提高网络模型的泛化能力，有助于模型聚焦于普通质量锚框，从而提高模型精度及模型检测的整体性能。

2　结果与分析

2.1　试验平台及评价指标

试验运行平台处理器为 Intel（R） Xeon（R） Silver 4 214 CPU@2.20 GHz 2.19 GHz，内存为 32 GB，显卡型号为 NVIDIA Quadro P620，软件环境为 Python 3.7，Windows 10 系统，采用 PyCharm 编辑器、Pytorch 深度学习框架进行训练。

训练采用优化器为 SGD，动量因子设为 0.937，权重衰减设为 0.000 5。选用精确率（Precision，P）、召回率（Recall，R）、平均精度均值（mean Average Precision，mAP）以及 F1 值作为检测精度的评价指标，选用参数量（Parameters）、浮点计算量（Floating point operations per second）以及模型内存占用量作为模型检测性能的评价指标。

2.2　消融试验

为验证 ESW－YOLO 模型的效果，设置 8 组消融试验，与 YOLO v7－tiny 模型进行对比，对比结果如表 1 所示。

由消融试验可知，YOLO v7－tiny 模型的损失函数更改为 WIoU 后，模型的精确率、召回率和平均精度均值分别达到 81.48%、81.52% 以及 86.08%，分别提升了 1.08、3.22、2.68 个百分点，提高了网络的泛化性能。YOLO v7－tiny 模型添加 EMA 模块，提取不同尺度的特征并保留更多的上下文信息，精确

表 1　消融试验结果

Tab. 1　Ablationtest results

WIoU	EMA	GSConv	Slim－neck	精确率/%	召回率/%	平均精度均值/%	参数量	模型内存占用量/MB
				80.40	78.30	83.40	6.02×10^6	12.3
√				81.48	81.52	86.08	6.02×10^6	12.3
	√	√		82.50	80.20	86.20	6.07×10^6	12.4
		√	√	83.20	77.50	84.20	5.63×10^6	11.6
√	√			82.69	81.82	86.49	6.07×10^6	12.4
√		√		81.86	81.88	86.14	5.80×10^6	11.8
√		√	√	83.10	80.20	85.30	5.63×10^6	11.6
√	√	√	√	83.02	82.04	86.83	5.69×10^6	11.7

注：“√”表示添加此模块。

率和召回率以及平均精度均值都有较大提升，分别提高了 2.1、1.9、2.8 个百分点，但是参数量增加了 1%，模型内存占用量增加了 0.1 MB；同时更改损失函数为 WIoU，此时召回率提升了 1.62 个百分点。单独在 YOLO v7－tiny 模型的颈部网络加入轻量化的 GSConv－Slim－neck 范式设计，其参数量减少了 6.5%，降低了模型计算成本，模型内存占用量减少了 0.7 MB，并且精确率和平均精度均值分别提升了 2.8、0.8 个百分点，但是召回率下降了 0.8 个百分点，在更改了 WIoU 损失函数之后精确率下降 0.1 个百分点，但是召回率和平均精度均值分别提升 2.7、1.1 个百分点。在此基础上将 EMA 模块加入模型的主干网络中，在只增加模型内存占用量 0.1 MB 的情况下召回率和平均精度均值分别提高了 1.84、1.53 个百分点，精确率降低了 0.08 个百分点。最终 ESW－YOLO 模型的精确率、召回率和平均精度均值分别为 83.02%、82.04%、86.83%，较 YOLO v7－tiny 模型分别提升 2.62、3.74、3.43 个百分点，参数量减少 5.48%，整体检测性能相较 YOLO v7－tiny 模型均有所提升。

2.3　检测效果与分析

ESW－YOLO 模型与 YOLO v7－tiny 模型在 2 种光照条件（顺光与逆光）下对 3 种苹果树（红富士、瑞雪与瑞香红）进行检测，并将标签与检测目标框个数输出在图像左上角，2 种模型的检测效果如图 10 所示。由图 10 可知在相同图像检测下，ESW－YOLO 模型在不同光照条件与果树品种下可以检测到的目标框数量以及检测精确率均高于 YOLO v7－tiny 模型。

表 2 为经过 ESW－YOLO 模型检测后每个示例图像中花序的误检数、漏检数、目标统计数量以及 F1 值，其中红富士与瑞香红果树的花序误检数和漏检数较少，由于瑞雪果树每个枝干上花序较为密集，

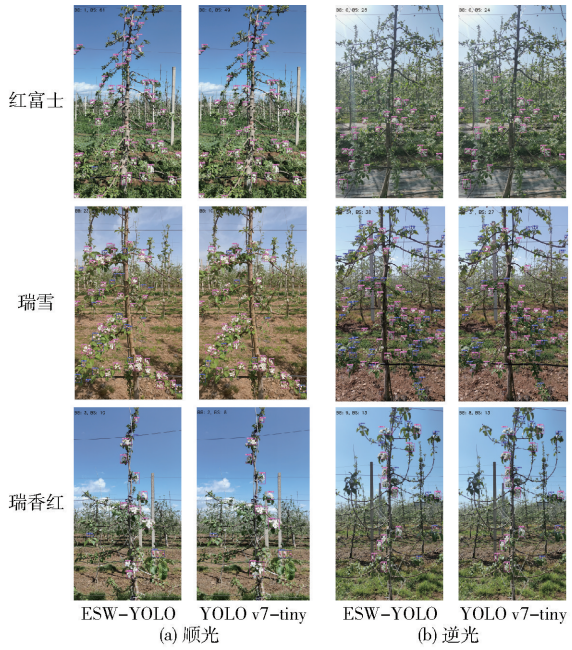


图 10 2 种模型的苹果花序检测结果

Fig. 10 Detection results of apple inflorescences in two models

且相互重叠与遮挡严重,相对其他 2 种果树的花序误检数和漏检数较高。顺光下 3 种果树的花序检测数与实际数相差不大,其中红富士与瑞雪 2 种果树的花序检测数大于实际数,可能是果树背面的花序检测为遮挡花序;逆光下由于光照条件原因各果树种类的检测数均略小于实际数。每幅图像检测的 F1 值均大于 84%,证明 ESW - YOLO 模型对苹果花序的检测性能较高。

本研究将测试集按光照条件分为顺光和逆光 2 种环境,其中逆光图像 136 幅,苹果花序标签 4 748 个,顺光图像 124 幅,苹果花序标签 5 708 个。按照苹果树种类分为红富士、瑞雪、瑞香红,其中红富士果树图像 94 幅,苹果花序 4 410 个,瑞雪果树图像 66 幅,苹果花序 4 243 个,瑞香红果树图像 97 幅,苹果花序 1 718 个。采用 ESW - YOLO 模型对不同品种苹果树以及在不同光照条件下的检测结果如表 3 所示。

由表 3 可知,在顺逆光环境下 3 种苹果树花序

表 2 ESW - YOLO 在示例图像上的检测结果

Tab. 2 Detection results of ESW - YOLO on example images

品种	顺光					逆光				
	误检数/个	漏检数/个	检测数	实际数	F1 值/%	误检数/个	漏检数/个	检测数	实际数	F1 值/%
红富士	8	3	62	57	90.76	2	6	26	30	85.71
瑞雪	11	9	65	63	84.38	4	25	92	13	85.85
瑞香红	2	2	13	13	84.62	2	3	21	22	88.37

表 3 ESW - YOLO 在不同测试集上的检测结果

Tab. 3 Detection results of ESW - YOLO on different test sets

图像分类		T_P	F_P	F_N	F1 值/%
光照条件	顺光	4 757	889	951	83.79
	逆光	3 881	656	867	83.60
果树种类	红富士	3 786	873	624	83.49
	瑞雪	3 458	440	785	84.95
	瑞香红	1 333	229	385	86.96

注:T_P为目标框中识别正确的苹果花序总数;F_P为目标框中识别错误的苹果花序总数;F_N为漏检的苹果花序数量。

的 F1 值均相差不大,而红富士与瑞雪的苹果花序 F1 值低于瑞香红的 F1 值。红富士果树枝叶茂密,树冠较大,花序遮挡较为严重,瑞雪果树上花序生长较为密集,各个花序之间遮挡重叠严重,而瑞香红果树树龄最小,树叶与花序生长较为独立,遮挡与重叠部分较少。

采用 ESW - YOLO 与 YOLO v7 - tiny 模型训练之后的最终权重文件,对顺逆光环境下的不同果树种类的苹果花序测试集进行检测,比较其 F1 值,检测结果如表 4 所示。在不同光照条件与不同果树种类的检测中,ESW - YOLO 模型的 F1 值高于 YOLO v7 - tiny 模型,证明 ESW - YOLO 模型对不同光照条件下与果树种类中花序的检测性能均有较好的提升。

表 4 2 种模型在不同测试集上的 F1 值

Tab. 4 F1 values for both models on different test sets %

模型	光照条件		果树种类		
	顺光	逆光	红富士	瑞雪	瑞香红
ESW - YOLO	83.79	83.60	83.49	84.95	86.96
YOLO v7 - tiny	82.38	81.61	82.82	82.00	80.32

2.4 对比试验

将 ESW - YOLO 模型与目前主流常用一阶段模型 YOLO v5s[27]、YOLO v8s[28] 以及 YOLO v7 - tiny 进行对比,对比结果如表 5 所示。

根据表 5 分析可得,在主流的一阶段模型中,YOLO v7 - tiny 的精确率比 YOLO v8s 低 1.9 个百分点,平均精度均值相差 0.64 个百分点,召回率比 YOLO v8s 高 2.7 个百分点,并且参数量、浮点计算量以及模型内存占用量都比其余模型低,因此选用 YOLO v7 - tiny 为本研究基础模型。ESW - YOLO

表 5　不同一阶段模型对比试验结果

Tab. 5　Comparison experimental results of different single-stage models

参数	YOLO v5s	YOLO v7－tiny	YOLO v8s	ESW－YOLO
精确率/%	80.34	80.40	82.30	83.02
召回率/%	75.13	78.30	75.60	82.04
平均精度均值/%	81.77	83.40	84.00	86.83
参数量	7.06×10^6	6.02×10^6	1.114×10^7	5.69×10^6
浮点计算量	1.61×10^{10}	1.32×10^{10}	2.86×10^{10}	1.29×10^{10}
模型内存占用量/MB	13.6	12.3	22.5	11.7

模型相较于 YOLO v7－tiny 模型参数量减少了 5.48%，模型内存占用量减少了 4.88%，并且精确率和召回率分别提高了 2.62、3.74 个百分点，平均精度均值相较于 YOLO v5、YOLO v7－tiny 以及 YOLO v8s 分别提高了 5.06、3.43、2.83 个百分点。

2.5　模型特征热力图可视化

目前，在深度学习中对网络结构的直观解释常用特征图可视化技术，通过将不同网络层的输出特征转化为可视化的热力图像，将网络重点关注部分进行突出标注，Score－CAM 方法可以防止梯度消失对可视化特征的影响，提高可视化过程的精准度[29]。本研究提取主干网络第 14 层、第 21 层、第 37 层以及颈部网络第 47 层、第 50 层、第 53 层输出特征绘制激活图，并与原图进行尺寸映射得到热力图，其可视化的图像如图 11 所示。

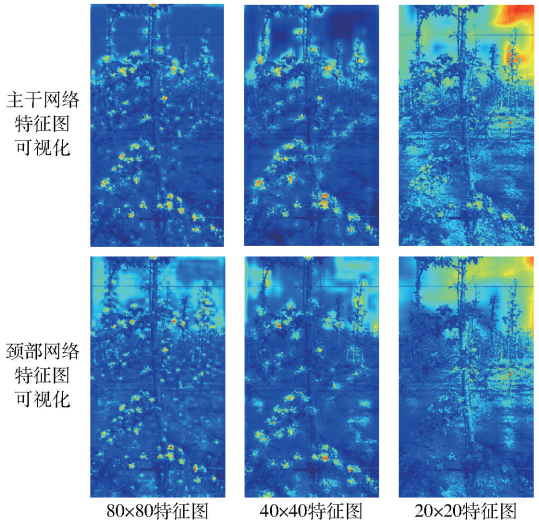


图 11　主干网络与颈部网络可视化特征图

Fig. 11　Visualization features of backbone network and neck network

由图 11 可知，主干网络的浅层部分（80×80、40×40）对苹果花序实现基础的关注，由于该阶段提取到的信息皆为浅层特征，因此也关注了当前行的树干及其附近区域，随着网络的加深，关注区域的边缘变得模糊。网络继续加深，主干网络 20×20 的特征图对不同类别物体进行分别关注，包括天空、地面、树叶以及目标花序。主干网络将提取到的特征输入到颈部网络，通过 GSConv－Slim－neck 范式设计进行特征融合，进一步提高模型的特征提取与融合能力。在颈部网络中，第 47 层输出的 80×80 特征图重点关注目标花序以及背景。第 50 层的 40×40 特征图进一步提取目标花序的轮廓，对当前行苹果树进行重点关注，逐渐平滑背景。第 53 层的输出则重点关注图像中除苹果花序外的背景，有助于过滤复杂背景信息，突出花序特征。

3　结论

（1）为提高化学疏花期苹果花序状态检测性能，提出一种 ESW－YOLO 模型用于检测化学疏花时期的苹果花序状态。更改 YOLO v7－tiny 的损失函数为 WIoU，在主干网络中加入 EMA 模块，在颈部网络中加入 GSConv－Slim－neck 范式设计，ESW－YOLO 模型具有较高的检测精度与较小的模型内存占用量。

（2）ESW－YOLO 模型相较于 YOLO v7－tiny 模型，在不同光照条件下（顺光、逆光）与果树种类（红富士、瑞雪、瑞香红）中检测时均能较好地处理花序遮挡与重叠问题，表现出较好的检测性能。

（3）ESW－YOLO 模型在测试集上的精确率与召回率分别为 83.02% 与 82.04%，平均精度均值为 86.83%，较 YOLO v7－tiny 模型提升 3.43 个百分点，参数量减少 5.48%。与 YOLO v5s 和 YOLO v8s 相比，平均精度均值分别提高 5.06、2.83 个百分点。

（4）该研究采用 Score－CAM 方法对 ESW－YOLO 模型的网络结构进行热力图可视化分析，在主干网络中提取花序与背景细节，在颈部网络中进行特征融合，进一步提取花序轮廓。将特征提取与特征融合网络的输出内容可视化，进一步提高了模型的可解释性。

参　考　文　献

[1]　ALLEN W, KON T, SHERIF S. Evaluation of blossom thinning spray timing strategies in apple[J]. Horticulturae, 2021, 7(9): 308.

[2]　马虽有. 苹果园常用化学疏花疏果剂及其应用技术[J]. 西北园艺(果树), 2024(2): 34－36.

[3]　WANG X, TANG J, WHITTY M. DeepPhenology: estimation of apple flower phenology distributions based on deep learning

[J]. Computers Electronics in Agriculture, 2021, 185: 106123.
[4] 孙丽娟. 苹果园化学疏花精准控制系统研究[D]. 杨凌: 西北农林科技大学, 2024.
SUN Lijuan. Research on precise control system of chemical thinning in apple orchard[D]. Yangling: Northwest A&F University, 2024. (in Chinese)
[5] 郭庆梅, 刘宁波, 王中训, 等. 基于深度学习的目标检测算法综述[J]. 探测与控制学报, 2023, 45(6): 10-20,26.
GUO Qingmei, LIU Ningbo, WANG Zhongxun, et al. Review of on deep learning based object detection algorithms[J]. Journal of Detection & Control, 2023, 45(6): 10-20,26. (in Chinese)
[6] FARJON G, KRIKEB O, HILLEL A, et al. Detection and counting of flowers on apple trees for better chemical thinning decisions[J]. Precision Agriculture, 2020, 21: 503-521.
[7] 陈国防, 陈兆英, 王玉亮, 等. 基于数据增强深度学习的苹果花检测方法研究[J]. 中国农机化学报, 2022, 43(5): 148-155.
CHEN Guofang, CHEN Zhaoying, WANG Yuliang, et al. Research on detection method of apple flower based on data-enhanced deep learning[J]. Journal of Chinese Agricultural Mechanization, 2022, 43(5): 148-155. (in Chinese)
[8] 宋怀波, 尚钰莹, 何东健. 果实目标深度学习识别技术研究进展[J]. 农业机械学报, 2023, 54(1): 1-19.
SONG Huaibo, SHANG Yuying, HE Dongjian. Review on deep learning technology for fruit target recognition[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 1-19. (in Chinese)
[9] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
[10] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.
[11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Computer Vision-ECCV 2016, 2016: 21-37.
[12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
[13] TIAN Y, YANG G, WANG Z, et al. Instance segmentation of apple flowers using the improved Mask R-CNN model[J]. Biosystems Engineering, 2020, 193: 264-278.
[14] DIAS P A, TABB A, MEDEIROS H. Apple flower detection using deep convolutional networks[J]. Computers in Industry, 2018, 99: 17-28.
[15] MU X, HE L, HEINEMANN P, et al. Mask R-CNN based apple flower detection and king flower identification for precision pollination[J]. Smart Agricultural Technology, 2023, 4: 100151.
[16] 王永生, 姬嗣愚. 基于深度学习的目标检测算法综述[J]. 计算机与数字工程, 2023, 51(6): 1231-1237.
WANG Yongsheng, JI Siyu. Review of target detection algorithms based on deep learning[J]. Computer & Digital Engineering, 2023, 51(6): 1231-1237. (in Chinese)
[17] 翟肇裕,张梓涵,徐焕良,等. YOLO 算法在动植物表型研究中应用综述[J]. 农业机械学报,2024,55(11):1-20.
ZHAI Zhaoyu, ZHANG Zihan, XU Huanliang, et al. Review of applying YOLO family algorithms to analyze animal and plant phenotype[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(11): 1-20. (in Chinese)
[18] 陈国防. 基于深度学习的苹果花识别检测方法研究[D]. 泰安: 山东农业大学, 2023.
CHEN Guofang. Research on apple flower recognition and detection methodbased on deep learning[D]. Taian: Shandong Agricultural University, 2023. (in Chinese)
[19] 司永胜, 孔德浩, 王克俭, 等. 基于 CRV-YOLO 的苹果中心花和边花识别方法[J]. 农业机械学报, 2024, 55(2): 278-286.
SI Yongsheng, KONG Dehao, WANG Kejian, et al. Recognition of apple king flower and side flower based on CRV-YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 278-286. (in Chinese)
[20] 高昂, 卢传兵, 任龙龙, 等. 基于改进 YOLOX-s 的苹果花生长状态检测方法及验证分析[J]. 中国农机化学报, 2023, 44(8): 162-167.
GAO Ang, LU Chuanbing, REN Longlong, et al. Detection method and validation analysis based on the improved YOLOX-s apple blossom growth state[J]. Journal of Chinese Agricultural Mechanization, 2023, 44(8): 162-167. (in Chinese)
[21] 薛晓敏, 王金政, 王贵平, 等. 苹果化学疏花疏果剂应用技术规范(试行)[J]. 落叶果树, 2016, 48(6): 57-58.
[22] ZHANG C, MOUTON C, VALENTE J, et al. Automatic flower cluster estimation in apple orchards using aerial and ground based point clouds[J]. Biosystems Engineering, 2022, 221: 164-180.
[23] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning[C]// IEEE International Conference on Acoustics, Speech and Signal Processing, 2023: 1-5.
[24] LI H, LI J, WEI H, et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles[J]. arXiv Preprint arXiv: 2206.02424, 2022.
[25] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
[26] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[J]. arXiv Preprint, arXiv: 2023.10051,2023.
[27] ZHU X, LYU S, WANG X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2021: 2778-2788.
[28] VARGHESE R, SAMBATH M. YOLOv8: a novel object detection algorithm with enhanced performance and robustness[C]// 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), 2024: 1-6.
[29] WANG H, WANG Z, DU M, et al. Score-CAM: score-weighted visual explanations for convolutional neural networks[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020: 111-119.