

基于 YOLOv8n-Aerolite 的轻量化蝴蝶兰种苗目标检测算法

翟永杰，田济铭，陈鹏晖，王家豪，胡东阳，徐大伟

(华北电力大学自动化系，保定 071003)

摘要：小型植物组织检测对植物自动化培养产业的发展具有重要意义，为了提升蝴蝶兰种苗夹取点视觉检测效率以及解决现有模型参数量较大，检测速度较慢的问题，该研究提出了一种轻量化目标检测算法 YOLOv8n-Aerolite。首先，采用 StarNet 作为主干网络，在此基础上增加嵌入大核可分离卷积的池化层 SPPF_LSKA (large-separable-kernel-attention)，实现轻量化的同时保证准确率；然后在颈部网络中采用结合 StarBlock 的 C2f_Star 模块，提高模型对蝴蝶兰种苗检测的准确率；最后，采用以共享卷积为基础的轻量级检测头 Detect_LSCD (lightweight shared convolutional detection head)，提升模型对小目标检测的精度和速度。在对蝴蝶兰种苗图像数据集的目标检测试验中，YOLOv8n-Aerolite 算法的平均推理速度达到了 435.8 帧/s，精确度达 91.1%，权重文件大小仅为 3.1 MB，对于夹取点所在小目标检测精度达 91.6%，在种苗夹取试验中成功率为 78%，研究结果可为发展小型作物自动化栽培技术提供参考。

关键词：深度学习；YOLOv8n；蝴蝶兰种苗；轻量化；检测速度；小目标检测

doi: 10.11975/j.issn.1002-6819.202408192

中图分类号：S126

文献标志码：A

文章编号：1002-6819(2025)-04-0220-10

翟永杰，田济铭，陈鹏晖，等. 基于 YOLOv8n-Aerolite 的轻量化蝴蝶兰种苗目标检测算法[J]. 农业工程学报，2025, 41(4): 220-229. doi: 10.11975/j.issn.1002-6819.202408192 <http://www.tcsae.org>

Zhai Yongjie, TIAN Jiming, CHEN Penghui, et al. Algorithm for the target detection of phalaenopsis seedlings using lightweight YOLOv8n-Aerolite[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(4): 220-229. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202408192 <http://www.tcsae.org>

0 引言

蝴蝶兰作为一种多年生附生草本植物，因其蝶形花朵和高观赏价值深受喜爱^[1]。然而，传统的人工培养方式劳动强度大、工作枯燥、效率低下。此外，人工分拣和培养过程容易破坏种苗的无菌环境，增加种苗感染病害的风险。随着科技的发展，自动化、人工智能和机器人技术逐步应用于蝴蝶兰种苗分拣、切割、栽种等环节中。在这些自动化生产环节中，以实现种苗组织的自动化培养。对植物组织快速精确的目标检测是其中的关键技术。

对于植物组织的目标检测，早期主要采用图像处理和机器学习方法。马战林等^[2]利用建立的机器学习模型和面域数据反演区域冬小麦产量，实现作物生长模型与机器学习算法的应用耦合。张领先等^[3]利用可见光光谱颜色特征结合支持向量机进行温室黄瓜霜霉病病斑图像分割，采用 SURF (speeded up robust features) 特征及形态学操作对分割结果进行优化。为温室黄瓜霜霉病的科学防治提供参考。张开兴等^[4]将图像处理技术和 BP 神经网络算法引入到玉米叶部病害识别诊断中，实现了玉米叶部常见 6 种病害的分类识别。传统的机器学习方法

在植物组织的目标检测中易受外界环境影响，依赖手动输入调整阈值，训练时间较长且精确度较低。

近年来，基于深度学习的目标检测方法以其较强的适用性与泛化能力在农业自动化领域受到广泛关注。李天华等^[5]采用改进的 YOLOv4 深度学习模型对不同成熟程度的番茄进行检测，结合 HSV (hue saturation value) 方法对番茄的红色区域进行分割，提高识别成熟期番茄的准确性。刘毅君等^[6]采用改进的 Faster-R-CNN 神经网络对马铃薯表面的损伤进行检测，通过构建新的特征提取网络提升检测精度。LE 等^[7]提出了一种用于复杂环境下杂草检测方法，在 Faster R-CNN 模型中引入 Inception-ResNet-V2 特征提取模块，大幅度提升了检测精度。杨坚等^[8]对 YOLOv4-tiny 进行改进，在模型中增加了 76×76 的检测头，并在主干网络中引入了卷积注意力模块，以增强对存在部分遮挡的番茄的识别能力。ZHOU 等^[9]将 MobileNetV3 部署到 Android 平台，通过开发的 APP 实现猕猴桃产量估算。此类研究虽提升了检测精度，但在模型的轻量化程度上关注较少，难以部署在储存能力和算力较差的边缘设备中，在实际生产中的泛化性较差。

在模型轻量化研究领域，YAN 等^[10]提出了一种轻量化的 YOLOv5 目标检测算法，用于实时筛选成熟苹果并优化采摘动作。范天浩等^[11]提出了一种轻量化金银花识别模型。通过引入 EfficientNet 网络结构改进 YOLOv5s，改进后模型参数量仅为原始模型的 55.5%，但在平均精度上仅提升了 0.6 个百分点。此类研究专注于模型的轻量化程度，不适用于小型植物组织目标检测任务。

在种苗组织自动化培养中，需要同时满足检测精度

收稿日期：2024-08-26 修订日期：2024-11-19

基金项目：国家自然科学基金联合基金项目重点支持项目 ((U21A20486)；河北省自然科学基金面上项目 (F2023502010)；河北省高等学校科学技术研究项目资助 (QN2024171)

作者简介：翟永杰，博士，教授。研究方向为电力视觉。

Email: zhaiyongjie@neepu.edu.cn

与检测速度的要求。但现有算法的部署往往受限于边缘设备的存储容量和计算能力^[12]。因此,需要设计一种高精度的轻量化检测算法^[13]。此外,在对蝴蝶兰种苗的后续夹取工作中,选择黑色块茎作为夹取部位对种苗造成损伤较少,但黑色块茎部位目标相对较小,故检测模型应当对小目标检测精确度较高。

针对上述问题,本文在YOLOv8n的基础上设计了一种蝴蝶兰种苗检测模型YOLOv8n-Aerolite,通过对主干网络、颈部网络和头部网络的改进,重点提升模型的轻量化程度,同时实现检测精度与检测速度的平衡,以期为蝴蝶兰种苗的自动化夹取提供一种快速精确的轻量化目标检测算法。

1 材料和方法

1.1 数据采集

本次试验采用在实验室条件下拍摄的蝴蝶兰种苗正视角的高清RGB图像,拍摄设备为Intel深度相机D455,图像存储格式为JPG格式,分辨率为 3072×2048 像素。由于实际部署中很难确定蝴蝶兰种苗摆放的具体姿态,故拍摄时对不同姿态放置的蝴蝶兰种苗都需要拍摄,在灯光下,需要进行夹取的块茎部位为黑色。为提高模型的泛化能力,丰富数据多样性,在多种条件下分别采集了蝴蝶兰种苗图像,样例如图1所示。



图1 蝴蝶兰种苗图像

Fig.1 Phalaenopsis seedling images

1.2 数据集制作

拍摄完成后使用深度学习图像标注软件LabelImage对这些图像进行筛选和标注。在标注过程中,将蝴蝶兰种苗上的叶片、胚根和黑色块茎用矩形框进行标注,矩形框的位置和大小表示目标检测框的信息,检测框的宽度和高度由左上和右下坐标间接表示。标注完成后,将生成的标注信息保存为YOLO模型所需的txt标签文件。总标注图像数量为1630张,按照8:2的比例划分为训练集和测试集,具体数量为训练集图像1304张,测试集图像326张,图像标注示例如图2所示。



图2 蝴蝶兰种苗图像标注示例

Fig.2 Annotation examples of phalaenopsis seedling images

对标注后的训练数据集进行统计,标注框数目如图3所示。共有标注框5716个,其中黑色块茎标注框数目1304个、胚根标注框数目2587个、叶片标注框数目1825个。每株蝴蝶兰种苗一般具有1个黑色块茎、两个胚根、两个叶片,标注中的类别数目比例与真实种苗情况基本一致。

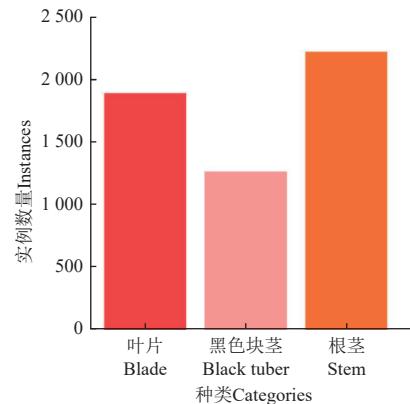


图3 蝴蝶兰种苗数据集类别分布

Fig.3 Phalaenopsis seedling dataset category distribution

2 蝴蝶兰种苗检测方法

2.1 YOLOv8n 目标检测

YOLOv8n是一种高效的目标检测算法^[14]。由主干网络、颈部网络和头部网络组成。主干网络采用C2f模块,使用无锚点检测方法,以实现更准确地边界框预测^[15]。颈部网络通过使用路径聚合网络(PANet)提升对不同尺度对象的检测能力^[16]。头部网络采用解耦设计,提高模型的专门学习能力和精度^[17]。

然而,在应用于特定场景时,YOLOv8n仍存在不足。例如,面对蝴蝶兰种苗目标时,其特征提取能力对小目标的适应性尚显不足;在资源受限的边缘设备上,其模型轻量化水平和推理速度难以完全满足部署需求与实时性要求。针对这些问题,本文在YOLOv8n模型的基础上对主干网络、颈部网络和头部网络分别进行了改进。

2.2 YOLOv8n-Aerolite 算法

首先,在主干网络部分,YOLOv8n-Aerolite引入了

更高效的 StarNet 结构, 结合深度可分离卷积和星形操作, 提高特征提取效率。颈部网络通过采用 C2f_Star 模块增强特征表达能力。此外, 头部网络采用轻量级共享卷积检测头 Detect_LSCD, 通过共享卷积操作减少计算复杂度, 进一步提升小目标的检测精度。YOLOv8n-Aerolite 目标检测算法的整体架构如图 4 所示。

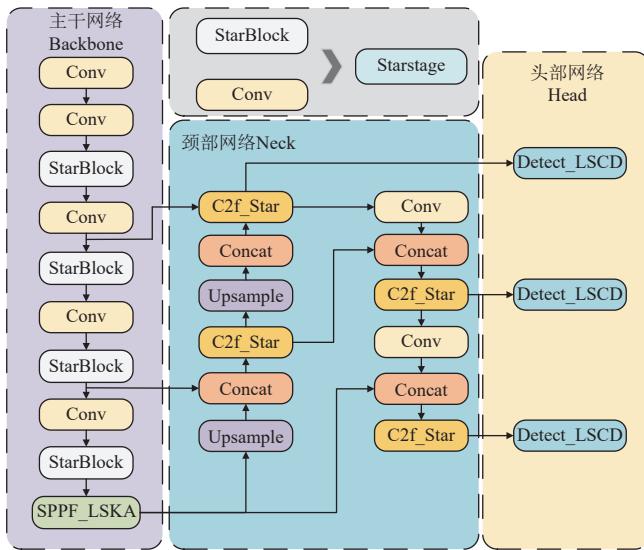


图 4 YOLOv8n-Aerolite 网络结构图

Fig.4 YOLOv8n-Aerolite network architecture diagram

2.2.1 主干网络改进

YOLOv8n-Aerolite 在原始 YOLOv8n 模型的基础上引入了全新的主干网络结构 StarNet, 替换了原有的主干网络。StarNet 的整体架构由多个特征提取阶段 (Starstage) 组成, 每个阶段包括一个卷积下采样层和多个星形块。这些特征提取阶段通过逐步降低特征图的分辨率并增加通道数, 以实现高效的特征提取能力。StarNet 的设计结合了星形操作 (逐元素乘法) 与深度可分离卷积, 有助于减少模型的参数量和计算复杂度。每个星形块由一个 7×7 的深度可分离卷积和多层次感知机制组成, 其中星形操作替代了传统的加法操作, 旨在捕捉特征间的关系。该设计为提升模型轻量化性能提供了潜在可能。

在 StarNet 的核心模块 StarBlock 中, 采用了多个卷积和非线性激活层组合, 以高效地提取和融合图像特征。首先, 通过深度可分离卷积层, 该层利用 7×7 的卷积核以及分组卷积的方式捕捉输入特征图的局部空间特征。具体处理过程如式 (1) 所示。

$$x' = d(x) \quad (1)$$

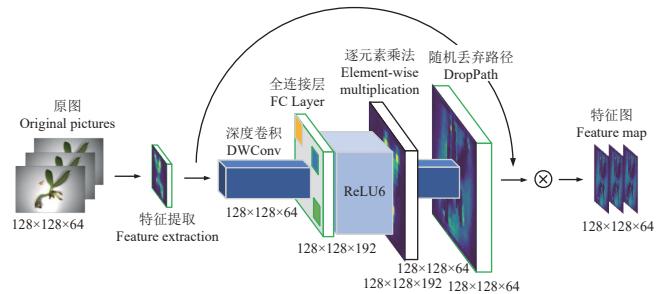
式中 x 表示输入特征图, x' 表示经过深度可分离卷积后的特征图, $d(\cdot)$ 表示深度可分离卷积计算函数。

特征图在经过深度可分离卷积后利用两个全连接层 (FCLayer) 将特征通道数扩展, 通过 ReLU6 激活函数控制输出并通过逐元素乘法^[18] 进行特征融合, 从而增强特征的表示能力, 具体处理过程如式 (2) 所示。

$$\begin{cases} x_1 = \text{FCLayer1}(x') \\ x_2 = \text{FCLayer2}(x') \\ x'' = \text{ReLU6}(x_1) \odot x_2 \end{cases} \quad (2)$$

式中 FCLayer1 和 FCLayer2 是两个全连接层, x_1 和 x_2 是两个全连接层后的特征图, ReLU6^[19] 表示激活函数计算过程, \odot 表示逐元素乘法操作, x'' 表示通过激活函数特征融合后的特征图。

随后, 通过 1×1 卷积层将特征通道数还原至原始维度, 并通过第 2 个 7×7 的卷积层进一步提取高阶特征。为了增强模型的泛化能力, 引入了 DropPath 随机丢弃路径, 通过随机丢弃部分通道来减少过拟合的风险, 从而在保证计算效率的同时, 保持模型的高表达能力和泛化能力。StarBlock 模块的结构如图 5 所示。



注: \otimes 表示残差连接。

Note: \otimes denotes residual connection.

图 5 StarBlock 结构示意图

Fig.5 Schematic diagram of the StarBlock structure

在原始 YOLOv8n 中, 卷积层是主要的计算模块^[20]。假设在卷积计算过程中输入通道数和输出通道数分别为 C_{in1} 和 C_{out1} , 卷积核大小为 $K_{h1} \cdot K_{w1}$, 特征图的高度和宽度分别为 H_1 和 W_1 。令传统卷积网络参数量为 m_1 , 计算量为 S_1 , 具体计算如式 (3) 所示。

$$\begin{cases} m_1 = C_{in1} \cdot C_{out1} \cdot K_{h1} \cdot K_{w1} \\ S_1 = 2 \times H_1 \times W_1 \cdot C_{in1} \cdot C_{out1} \cdot K_{h1} \cdot K_{w1} \end{cases} \quad (3)$$

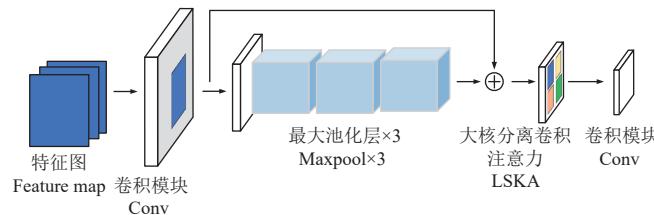
以 C_{in2} 和 C_{out2} 分别表示 StarBlock 的输入通道数和输出通道数, 以 H_2 和 W_2 表示特征图的宽和高。StarBlock 的参数量和计算量主要来源于两部分: 一是深度可分离卷积, 令其卷积核的尺寸为 $K_{h2} \cdot K_{w2}$, 参数量为 $2 \times C_{in2} \cdot K_{h2} \cdot K_{w2}$, 计算量为 $2 \times C_{in2} \cdot K_{h2} \cdot K_{w2} \cdot H_2 \cdot W_2$ 。二是全连接层, 其参数量为 $3 \times C_{in2}^2$, 计算量为 $3 \times C_{in2}^2 \cdot H_2 \cdot W_2$ 。基于此, 将每层特征提取阶段的参数量和计算量汇总, 得出 StarBlock 总的参数量 m_2 和计算量 S_2 , 计算如式 (4) 所示。

$$\begin{cases} m_2 = 2 \times C_{in2} \cdot K_{h2} \cdot K_{w2} + 3 \times C_{in2}^2 \\ S_2 = H_2 \cdot W_2 \cdot C_{in2} \cdot (2 \times K_{h2} \cdot K_{w2} + 3 \times C_{in2}) \end{cases} \quad (4)$$

由式 (3) 与式 (4) 可知, StarBlock 由于引入了深度可分离卷积, 其参数量与计算量不与输出通道数 C_{out2} 成比例关系, 而在传统卷积层计算量与参数量中, 输出通道数 C_{out1} 为乘积项, 当模型网络结构复杂度增加, 通道数较大时, StarBlock 在参数量和计算量上相比传统的卷积模块都有较为明显的减少, 这使得 StarBlock 在处理

高维特征时拥有了轻量化方面的优势。

然后,本文提出了一种改进的SPPF层,称为SPPF_LSKA,将其放置在主干网络的最后。该模块结合了SPPF和LSKA的优势,通过引入LSKA模块在特征提取方面表现出更强的能力,同时保持了高计算效率。具体来说,SPPF_LSKA是一个增强版的空间金字塔池化层。其设计包括4个主要部分:2个卷积层,3个最大池化层(Maxpool),以及1个大型卷积核空间注意力机制LSKA。在前向传递过程中,输入特征图首先通过1个 1×1 卷积层将通道数减半。随后,特征图经过3次连续的最大池化操作,生成多尺度特征图。然后,这些特征图被拼接在一起,并通过LSKA进行空间注意力处理,最后通过1个 1×1 卷积层融合和压缩通道数,生成最终的输出特征图。SPPF_LSKA的主要结构如图6所示。



注: \oplus 表示特征拼接。

Note: \oplus denotes feature splicing.

图6 SPPF_LSKA结构示意图

Fig.6 Schematic diagram of SPPF_LSKA structure

与原始的SPPF相比,SPPF_LSKA的主要改进在于引入了大型卷积核空间注意力机制LSKA^[21]。传统的SPPF仅通过最大池化和卷积操作进行多尺度特征提取,而SPPF_LSKA在此基础上增加了空间注意力模块,能够更好地捕捉全局上下文信息,提高特征图的表示能力。此外,LSKA的大卷积核设计能够有效增强特征图的长距离依赖关系,提高特征提取的质量,从而在目标检测任务中提高检测精度。

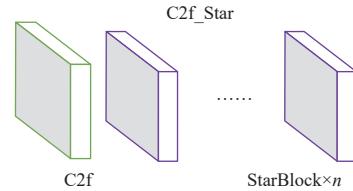
2.2.2 颈部网络改进

在颈部网络中,本研究利用C2f_Star替换了传统的C2f模块,在特征提取和网络结构上进行了优化。该模块在设计上与传统C2f模块有差异。C2f_Star引入了StarBlock,由参数n决定引入StarBlock层数(默认为1),其结合了 7×7 深度可分离卷积和多层次感知机结构,通过更有效的特征提取和交互方式来增强特征表达能力。StarBlock模块采用ReLU6激活函数和随机丢弃路径机制,增强了模型的泛化能力和鲁棒性。C2f_Star在保持计算效率的同时,通过复杂化的特征提取和交互方式,提高模型在复杂任务场景下的性能。C2f_Star模块的主要结构如图7所示。

2.2.3 头部网络改进

在头部网络部分,本文设计了一种新型检测头Detect_LSCD,Detect_LSCD是轻量级共享卷积检测头。此检测头通过共享卷积操作提高特征提取的效率,并降低计算成本。首先通过Conv_GN^[22]进行特征处理,之后将特征图输入到1个共享卷积网络(share_conv)中,

共享卷积网络由2个串联的Conv_GN组成。随后,特征图通过2个不同的卷积层分别计算边界框回归和分类概率,最终通过scale和DFL(分布式Focal Loss)进行后处理。该设计使得多个检测层之间共享特征,有助于提高检测精度和效率。



注: n 为 StarBlock 层数。

Note: n is the number of StarBlock layers.

图7 C2f_Star结构示意图
Fig.7 Schematic diagram of C2f_Star structure

输入特征图的高度和宽度分别为 H_3 和 W_3 ,隐藏层通道数为 C_{idc} ,共享卷积层的参数量 S_3 和计算量 m_3 的计算过程如式(5)所示。

$$\begin{cases} m_3 = 9C_{idc}^2 + C_{idc} \\ S_3 = H_3 \cdot W_3 \cdot (9C_{idc}^2 + C_{idc}) \end{cases} \quad (5)$$

而传统检测头中卷积模块包括2个 3×3 卷积层和1个 1×1 卷积层,假设输入特征图的高度和宽度分别为 H_4 和 W_4 ,第1层卷积输入通道数为 C_{c1} ,第2层为 C_{c2} ,最大回归值为 R_{max} ,其参数量 m_4 与计算量 S_4 如式(6)所示。

$$\begin{cases} m_4 = 2 \times (9C_{c1} \cdot C_{c2} + 9C_{c2}^2) + C_{c2} \cdot 4R_{max} \\ S_4 = H_4 \cdot W_4 \cdot (18C_{c1} \cdot C_{c2} + 18C_{c2}^2 + C_{c2} \cdot 4R_{max}) \end{cases} \quad (6)$$

对比式(5)与式(6)的参数量与计算量可知,当通道数近似时,共享卷积层的参数量和计算量的计算式中仅有项通道数的平方项 C_{idc}^2 ,而传统卷积层因包含多组通道数相乘操作和最大回归值的计算,参数量和计算量较大。传统的检测头通常对每一层特征图分别进行独立处理,而Detect_LSCD通过共享卷积操作,提高了特征提取效率。此外,Detect_LSCD采用了分布式Focal Loss^[23],提高了边界框回归的准确性。图8为Detect_LSCD的结构示意图。

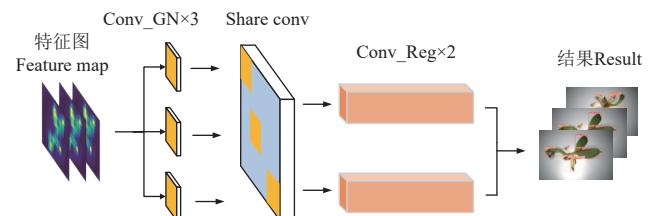


图8 Detect_LSCD结构示意图
Fig.8 Schematic diagram of Detect_LSCD structure

2.3 试验环境与参数

本文使用的视觉传感器为Intel D455深度相机。使用的深度学习模型在训练中使用Ubuntu20.04操作系统,32GB内存,NvidiaGeForce RTX3090显卡。试验采用的

深度学习框架为 Pytorch，在 CUDA12.1+Pytorch2.1.0+Python3.10 的开发环境下实现模型搭建及训练测试工作，深度神经网络 GPU 的加速库为 cuDNN8.9.0。在进行试验之前，对部分参数进行了优化，主要包括图像输入尺寸和每批训练样本数量等。本试验将训练轮数 (Epoch) 设定为 150。具体的环境配置和试验所采用的参数如表 1 所示。系统工作时的识别环境为 Windows10 操作系统，Intel core i7 处理器。最后，将训练好的模型部署到 JAKAZU 机器人上，进行实际场景测试。

表 1 训练参数设置

Table 1 Training parameter setting

参数 Parameter	参数值 Value
图像尺寸 Images size	640
批数量 Batch size	8
迭代次数 Epoch	150
学习率 Learning rate	0.01
动量 Momentum	0.937
权重衰减次数 Weight_decay	0.0005

2.4 评价指标

本文选用的评价指标包括平均精度均值 (mAP)、召回率 (R)、推理速度 (F)、模型参数总量 (m) 以及推理过程中的浮点运算数量 (FLOPs)、权重文件大小。

平均精度均值是用来衡量目标检测模型精度的指标。当混淆矩阵 IoU 阈值为 0.5 时计算所有类别的 AP 的平均值，即 $mAP_{0.5}$ 。召回率为评估目标检测模型能力的重要指标。计算过程如式 (7) 至式 (9) 所示。

$$P = \frac{P_T}{P_T + P_F} \times 100\% \quad (7)$$

$$R = \frac{P_T}{P_T + N_F} \times 100\% \quad (8)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \times 100\% \quad (9)$$

式中 P_T 为模型将实际为正类的样本正确预测为正类的数量， P_F 为模型将实际为负类的样本错误预测为正类的数量。 N_F 实际为正类但模型预测为负类的数量， P 表示精

确度， AP_i 为第 i 个类别的平均精度， N 为类别数。

推理速度 (F) 是用来衡量推理速度的指标，表示每秒钟模型可以推理的图像数目。 T 为推理时间。计算过程如式 (10) 所示。

$$F = \frac{1000}{T} \quad (10)$$

参数量 (m) 指的是神经网络模型中所有需要学习的参数的总数。浮点运算数量 (FLOPs) 表示模型在进行一次前向传播过程中所需的浮点运算次数。

3 试验与结果分析

3.1 消融试验

为验证本文所提出的 YOLOv8n 改进方法的有效性，本研究在蝴蝶兰种苗数据集上进行了多组消融试验，对各部分改进的作用及效果进行了全面评估。

表 2 的数据展示了主干网络、颈部网络和头部网络改进及相互组合后模型性能的变化。主干网络采用的 StarNet 因其星形结构提升了模型的轻量化表现，模型参数量减少了 0.722 M。这一改进有效降低了模型在资源受限设备上的内存占用和计算成本 (试验 2)。颈部网络改进后的模型在检测精度上表现突出，平均精度均值相比基线模型提升了 0.7 个百分点，达到了 91.1%。这一提升表明颈部网络的改进使模型在细粒度特征表达方面得到增强 (试验 3)。头部网络改进的效果主要体现在推理速度的提升，改进后推理速度达到 352.7 帧/s，远高于基线模型，推理速度提升优化了模型在实际应用中的响应效率 (试验 4)。

主干网络和颈部网络改进相结合后的模型在参数量和计算量上有了明显优化，两个指标均优于单独改进的效果 (试验 5)。而主干网络及头部网络的结合为模型带来了推理速度上的提升，较二者单独改进分别提升了 273.8、37.3 帧/s (试验 6)。颈部网络与头部网络改进相结合后的权重文件大小较二者单独改进分别降低了 1.4、0.4 MB (试验 7)。

表 2 消融试验结果

Table 2 Results of ablation experiments

试验编号 Test No.	主干网络改进 Backbone improvements	颈部网络改进 Neck network improvements	头部网络改进 Head network improvements	平均精度均值 mAP _{0.5} /%	参数量 Params/M	浮点运算数 FLOPs/G	权重文件大小 Weights file size/MB	推理速度 FPS/ (帧·s ⁻¹)
1	-	-	-	90.4	3.006	8.1	6.2	182.3
2	√	-	-	90.6	2.284	6.5	4.8	116.2
3	-	√	-	91.1	2.806	7.7	5.9	171.2
4	-	-	√	89.8	2.362	6.5	4.9	352.7
5	√	√	-	89.6	2.084	6.1	4.4	150.2
6	√	-	√	90.5	1.641	5.0	3.5	390.0
7	-	√	√	90.1	2.162	6.1	4.5	106.0
8	√	√	√	91.1	1.440	4.6	3.1	435.8

注: √表示使用该模块, -表示不使用该模块, 基线模型为 YOLOv8n。下同。

Note: √ indicates the use of the module, and - indicates the non use of the module, base model is YOLOv8n. The same below.

综合以上改进，YOLOv8n-Aerolite 取得了较好的检测性能，检测速度较基线模型提升了 253.5 帧/s，平均精度均值提升了 0.7 个百分点，参数量、计算量、权重文件大小均得到了超过 30% 的优化，更适用于要求严格的

自动化生产 (试验 8)。

3.2 超参数及多种网络对比试验

3.2.1 部分超参数对比

为了获得 YOLOv8n-Aerolite 的最佳性能，本文对部

分超参数进行对比试验, 探究超参数变化对模型检测性能的影响。参与对比的超参数包括学习率、动量和权重衰减次数。学习率是影响模型收敛速度和最终性能的关键因素。动量用于加速训练过程, 从而提高模型收敛的效率。权重衰减则通过引入正则化项来防止过拟合, 促进模型的泛化能力。YOLOv8n 模型默认的超参数学习率为 0.01, 动量为 0.937, 权重衰减为 0.0005, 根据默认参数的设置以及模型训练特性, 充分考虑了实际任务需求。学习率范围设定为 0.005~0.020, 是默认值 0.010 的约 2 倍, 确保覆盖较小学习率对模型稳定性的影响和较大学习率加速收敛的潜力; 动量调整为 0.920~0.950 是基于动量值通常接近 1 能够加速梯度下降的理论; 权重衰减范围设置为 0.0002~0.0010, 目的是通过对比该参数变化的影响来评估模型的稳定性。旨在深入探究不同超参数组合对本模型检测效果的影响, 从而确定最优的超参数配置。试验结果如表 3 所示。

表 3 超参数对比试验结果

Table 3 Hyper parameter comparison experiment results

试验编号 Test No.	学习率 Learning rate	动量 Momentum	权重衰减 Weight_decay	平均精度均值 mAP _{0.5} /%	推理速度 FPS/(帧·s ⁻¹)
1	0.005	0.937	0.0005	90.2	367.2
2	0.020	0.937	0.0005	89.4	233.1
3	0.010	0.920	0.0005	91.0	75.6
4	0.010	0.950	0.0005	90.5	79.1
5	0.010	0.937	0.0002	90.3	81.2
6	0.010	0.937	0.0010	90.2	236.4
7	0.010	0.937	0.0005	91.1	435.8

由表 3 可知, 当超参数设置为学习率为 0.01、动量为 0.937、权重衰减为 0.0005 时, 模型在检测精度和稳定性方面表现出最佳性能。这一组合在模型的收敛速度和精度之间达到了良好的平衡, 较高的学习率增加了参数更新幅度, 有助于加快模型收敛, 但未导致明显的过拟合或欠拟合。动量的设置增强了模型训练的稳定性, 使其在较长的训练过程中保持对重要梯度信息的敏感性。同时, 权重衰减有效限制了过大的权重更新, 避免了模型复杂度的无序增长。后续试验与部署将基于此超参数组合, 以确保模型在实际应用中具有高效且稳定的表现。

3.2.2 主干网络对比

为了研究主干网络变化对于模型轻量化程度、检测精度和推理速度的影响, 本研究对比了多种常见的主干网络, 对比过程中保持参数一致, 试验结果如表 4 所示。

由表 4 可知, Ghostnetv1 在精度度、参数量、浮点运算数量以及模型权重文件大小方面接近本文模型, 但在推理速度方面 FPS 仅为 90.2 帧/s, 无法满足实际生产的需求。LSKNet 在模型轻量化方面的表现不佳, 其中参数量达到了 5.051 M, 在实际布置时需要较大的算力。RepHGNetV2 在精度上表现良好, 平均精度均值达到了 90.5%, 但在推理速度上 FPS 仅为 93.2 帧/s, 检测速度较慢。RevColNet 在参数量上仅有 1.433 M, 但准确率不及 90%, 且推理速度较慢。StarNet 在多项指标上均表现突出: 平均精度均值为 91.1%, 推理速度提升至 435.8 帧/s,

同时在参数量和浮点运算量上保持轻量化, 适合实时性要求较高的场景。这表明, StarNet 作为主干网络在确保高精度的同时有效兼顾了轻量化和推理速度的需求, 证明了其在实际生产应用中的潜力。

表 4 主干网络对比试验结果

Table 4 Backbone network comparison experiment results

主干网络 Backbone	平均精度均值 mAP _{0.5} /%	参数量 Params/M	浮点运算数 FLOPs/G	权重文件大小 Weights file size/MB	推理速度 Inference speed FPS/(帧·s ⁻¹)
Ghostnetv1 ^[24]	90.9	1.454	4.2	3.2	90.2
LSKNet ^[25]	90.5	5.051	17.8	10.5	128.0
RepHGNetV2 ^[26]	90.5	2.144	6.5	4.6	93.2
RevColNet ^[27]	89.9	1.433	4.3	3.2	109.0
StarNet	91.1	1.440	4.6	3.1	435.8

3.2.3 头部网络对比

为了对比不同头部网络对蝴蝶兰种苗各部位的检测效果, 本文选取多种类型的检测头对蝴蝶兰种苗进行检测。在其他网络结构一致, 参数一致的条件下做出对比, 结果如表 5 所示。

表 5 头部网络对比试验结果

Table 5 Head network comparison experiment results

头部网络 Head network	平均精度均值 mAP _{0.5} /%	精度 Precision/%			推理速度 FPS/(帧·s ⁻¹)
		叶片 Blade	胚根 Stem	黑色块茎 Black tuber	
Detect_DyHead ^[28]	90.1	98.4	82.4	89.6	118.6
Detect_AFPN ^[29]	89.8	97.5	81.9	89.6	72.1
Detect_Aux ^[30]	89.4	99.2	81.5	87.6	124.8
Detect_Seam ^[31]	89.5	97.5	89.3	81.7	143.2
Detect_LSCD	91.1	97.1	84.5	91.6	435.8

由表 5 试验结果可知, Detect_Seam 对胚根检测精度最高, 但对黑色块茎部位检测的平均精度均值仅为 81.7%, 这表明其对小目标的检测效果较差, 不适用于以黑色块茎为目标的夹取场景。此外, Detect_AFPN 和 Detect_DyHead 的性能表现尽管总体精度较高, 但推理速度分别为 72.1、118.6 帧/s, 较低的推理速度限制了其在实时性要求高的任务中的适用性。Detect_Aux 的平均精度均值为 89.4%, 相对较低, 且在各方面的检测性能并不突出。

本文提出的 Detect_LSCD 头部网络在精度和速度上均具备优势, 平均精度均值达到了 91.1%, 黑色块茎部位检测精度达 91.6%, 检测速度为 435.8 帧/s。这说明 Detect_LSCD 设计更适合对蝴蝶兰种苗不同部位的全面检测任务, 模型将在实时精确性需求高的场景中表现更为优异。

3.2.4 常见模型对比

为了验证模型在蝴蝶兰种苗检测中整体的优越性, 表 6 对比了当前热门的目标检测模型, 包括 YOLOv3^[32]、YOLOv5^[33]、YOLOv6^[34]、Faster R-CNN^[35]、YOLOv8n、YOLOv8n-ghost、YOLOv5-ShuffleNet 和 YOLOv9^[36]。

由表 6 可知, YOLOv3 在检测中 mAP_{0.5} 仅为 84.4%, 但其参数量高达 103.732 M, 资源消耗较大; YOLOv5 和 YOLOv6 的模型参数量和浮点运算数量较小, 推理速

度分别达到 277.3、263.2 帧/s, 但在 $mAP_{0.5}$ 指标上稍显逊色仅有 85.5% 和 84.1%。Faster-RCNN 在检测精度方面表现优异, $mAP_{0.5}$ 达 86.0%, 但作为两阶段检测算法, 模型权重文件较大。YOLOv9 在精度上表现出色, 但其 45.372 M 的参数量较大, 表明其资源需求较高, 在边缘设备部署中可能受限。本文在对比试验中也加入了较为广泛使用的 YOLOv8n-ghost 和 YOLOv5-ShuffleNet 轻量化模型, YOLOv8n-ghost 权重文件大小仅为 3.8 MB, 参

数量 1.714 M, 轻量化优势明显, 但推理速度仅有 80.6 帧/s, 稍显不足。YOLOv5-ShuffleNet 进一步减小了参数量和权重文件大小, 模型权重文件大小仅有 3.4 MB, 但其 $mAP_{0.5}$ 仅有 84.3%, 在复杂任务中的表现有限。相比之下, YOLOv8n-Aerolite 以较小的权重文件大小 (3.1 MB) 实现了较高的 $mAP_{0.5}$ (91.1%) 和检测速度 (435.8 帧/s), 在性能与效率上达成平衡, 是资源受限场景的理想选择。

表 6 多种模型检测效果对比
Table 6 Comparison of detection effect of multiple models

模型 Model	精度 Precision/%			平均精度均值 $mAP_{0.5}/\%$	召回率 Recall/%	权重文件大小 Weights file size/MB	推理速度 Inference speed FPS/ (帧·s ⁻¹)	参数量 Params/M	浮点运算数量 FLOPs/G
YOLOv3	97.9	71.3	84.4	84.4	79.5	207.8	122.3	103.732	282.2
YOLOv5	99.2	73.8	83.5	85.5	81.7	5.3	277.3	2.504	7.1
YOLOv6	98.8	71.7	81.8	84.1	80.2	8.7	263.2	4.234	11.8
Faster-RCNN	98.2	73.1	86.9	86.0	58.9	315.1	1.1	41.761	196.7
YOLOv8n	98.7	88.8	83.7	90.4	86.1	6.2	182.3	3.006	8.1
YOLOv8n-ghost	97.6	81.4	88.5	89.2	85.7	3.8	80.6	1.714	5.0
YOLOv5-ShuffleNet	96.9	80.4	75.5	84.3	79.0	3.4	344.8	1.571	4.6
YOLOv9	99.4	70.9	85.7	85.3	84.1	102.8	206.7	45.372	207.9
YOLOv8n-Aerolite	97.3	84.3	91.6	91.1	86.4	3.1	435.8	1.440	4.6

如图 9 所示, 常见的轻量化算法 YOLOv5-ShuffleNet 在样本 3 的检测中出现了漏检现象未将胚根和黑色块茎部位检出, 在样本 2 中出现错检现象, 将空白区域错误检测为了叶片。同样, YOLOv8n-ghost 的误检与漏检现象也较为明显, 在样本 1 中出现了对黑色块茎重复检测的问题, 在样本 2 中存在对叶片部位的漏检。相比之下, 本文提出的 YOLOv8n-Aerolite 算法在试验中表现优异, 在多个样本上均完成了对蝴蝶兰种苗各部位精确的检测。

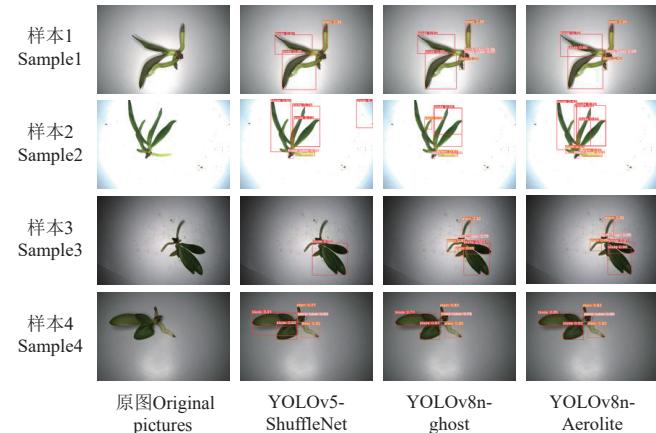


图 9 本模型与主流模型检测效果对比

Fig.9 Comparison of detection performance with mainstream models

在实际部署时, 模型推理速度至关重要, 图像的推理速度决定了机器人抓取种苗的效率。在模型轻量化改进下, 机器人在抓取过程中计算过程的参数量以及内存占用量大幅减少, 夹取点的计算时间也得到优化, 提升了机器人在实际抓取种苗时的效率, 降低运算负担。YOLOv8n-Aerolite 较高的检测速度和其轻量化程度可以支撑在实际生产中使用资源有限的边缘设备对蝴蝶兰种苗进行快速检测。而对夹取点所在黑色块茎小目标的识

别精度提升, 表明 YOLOv8n-Aerolite 算法在蝴蝶兰种苗数据集中的应用更为适宜。

3.3 泛化性试验

为验证模型的泛化能力, 本文对比了 YOLOv8n-Aerolite 和 YOLOv8n 模型在不同蝴蝶兰种苗数据集上的检测性能。通过 3D Max 软件对蝴蝶兰种苗进行三维建模, 使用 KeyShot 软件对模型渲染, 生成了多种样式的蝴蝶兰种苗模型, 并从不同角度采集种苗图像。在模型检测试验中, 使用了共计 2168 张蝴蝶兰种苗图像, 并按照 8:2 的比例将数据集划分为训练集和测试集。试验环境和参数配置与 2.3 节一致, 试验结果如表 7 所示。与 YOLOv8n 相比, YOLOv8n-Aerolite 的 $mAP_{0.5}$ 和召回率分别提升了 1.6 和 1.4 个百分点, 检测速度提升了 84.8 帧/s, 在三维重建蝴蝶兰种苗上的检测结果如图 10 所示, 证明 YOLOv8n-Aerolite 在该数据集上泛化性更好。

表 7 三维重建对比试验
Table 7 Comparison experiment of three-dimensional reconstruction

模型 Model	平均精度均值 $mAP_{0.5}/\%$	召回率 Recall/%	推理速度 Inference speed FPS/ (帧·s ⁻¹)
YOLOv8n	89.2	84.6	206.4
YOLOv8n-Aerolite	90.8	86.0	291.2

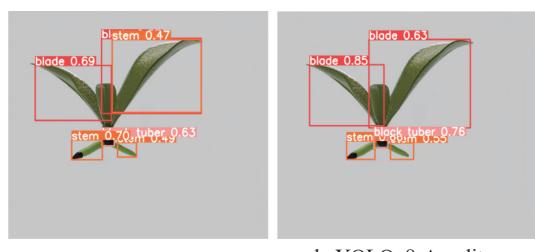


图 10 YOLOv8n 与 YOLOv8n-Aerolite 的种苗三维重建结果
Fig.10 3D-reconstructed seedling detection results of YOLOv8n and YOLOv8n-Aerolite

3.4 实际场景的验证与分析

在传统的机器人夹取作业中, 通常采用示教再现的方法确定夹取点位置, 并通过标定控制对机器人进行操作。然而, 这种方法在处理摆放姿态不固定、形状大小不一的蝴蝶兰种苗时存在局限性。由于夹取点位置不固定, 传统控制方法的夹取成功率较低, 常需人工实时干预以调整夹取位置, 降低了操作的快速性与自动化程度。同时, 人工调整过程中还依赖可视化程序, 增加了对部署设备的计算能力与存储资源的需求。

针对这一问题, 本文提出的 YOLOv8n-Aerolite 模型通过实时视觉检测, 可快速适应不同姿态的蝴蝶兰种苗, 精确定位夹取点。该方法有效提高了夹取效率与准确性, 增强了机器人在复杂工作环境下的自动化水平。此外, YOLOv8n-Aerolite 模型权重文件大小仅为 3.1 MB, 降低了对计算资源的需求, 成功实现了计算夹取点算法的轻量化部署。这一改进使得模型不仅具备较高的检测精度, 还适合资源受限的设备环境, 进一步提升了实际应用中的可操作性。在夹取工作中使用传统的标定控制与部署本模型的结果对比如表 8 所示。

表 8 标定控制与部署本模型检测对比

Table 8 Comparison of calibration control and deployment of this paper's model detection

夹取方案 Gripping strategy	夹取成功率 Success rate/%
标定控制 Calibration control	32
模型部署 Model deployment	78

对于本文在实际部署中采用的 JAKA_ZU3 六自由度机器人在部署时面临实际生产过程中边缘设备处理器计算能力、内存大小及功耗等方面的约束。YOLOv8n-Aerolite 模型的轻量化设计使其参数量和模型权重文件大小降低, 让模型能够很好地适应硬件限制。此外, 该模型在不同硬件平台上的兼容性良好, 且其轻量化特性使得在部署到特殊环境或架构的机器人设备时, 不需要进一步的模型压缩或量化操作。在存储与通信带宽方面, 轻量化模型的权重文件大小减少, 这不仅节省了存储空间, 还降低了在设备间通信时的带宽消耗, 从而提升了整体系统的效率。

为了验证本文提出的方法对在实际夹取时的效果, 选取 50 组蝴蝶兰种苗进行实际夹取试验, 用眼在手上方式布置 D455 相机, 利用本文提出的夹取系统对种苗完成检测, 识别, 夹取等一系列工作。机器人末端金属镊子夹具夹取蝴蝶兰黑色块茎部位成功抬起且姿态良好可用于后续种植即判定为夹取成功。夹取场景如图 11 所示。



图 11 夹取任务实现效果

Fig.11 Implementation results of gripping task

在实际夹取任务中, 选取了不同种类的蝴蝶兰种苗进行实际夹取测试。试验共进行了 50 次, 其中成功夹取 39 次, 失败夹取 11 次, 夹取成功率达到 78%。失败夹取的情况包括种苗在夹取过程中掉落或夹取后种苗的姿态不适宜后续种植工作。

4 结 论

1) 本文提出了一种针对算力有限条件下的轻量化蝴蝶兰种苗目标检测算法 YOLOv8n-Aerolite, 在 YOLOv8n 基础上, 主干网络引入 StarNet 网络结构和 SPPF_LSKA 池化层, 在颈部网络加入 C2f_Star 模块, 并在头部网络中引入 Detect_LSCD 检测头, 提升了模型检测精度。对种苗各部分的平均检测精度达 91.1%, 黑色块茎部位检测精度为 91.6%, 检测速度 435.8 帧/s, 模型权重文件大小为 3.1 MB, 参数量 1.440 M, 浮点运算数量 4.6 G。在提高模型检测精度的同时, 减少了模型的计算复杂度。

2) 本研究在三维重建的蝴蝶兰种苗数据集上进行了泛化性试验, YOLOv8n-Aerolite 相较于 YOLOv8n 模型, mAP_{0.5} 和检测速度分别提高了 1.6 个百分点和 84.8 帧/s, 证明了该方法在其他类似的小目标检测中具有广泛适用性。

3) 将 YOLOv8n-Aerolite 模型部署到边缘设备在实际场景中执行蝴蝶兰种苗检测工作, 配合机器人完成夹取任务, 在实际夹取过程中成功率达 78%。保证后续种植过程的顺利实施。本研究结果可为发展小型作物自动化快速繁育技术提供参考。

[参 考 文 献]

- [1] LI C, DONG N, ZHAO Y M, et al. A review for the breeding of orchids: Current achievements and prospects[J]. *Horticultural Plant Journal*, 2021, 7(5): 380-392.
- [2] 马战林, 文枫, 周颖杰, 等. 基于作物生长模型与机器学习算法的区域冬小麦估产[J]. 农业机械学报, 2023, 54(6): 136-147.
MA Zhanlin, WEN Feng, ZHOU Yingjie, et al. Regional winter wheat yield estimation based on crop growth models and machine learning algorithms[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(6): 136-147. (in Chinese with English abstract)
- [3] 张领先, 田潇, 李云霞, 等. 可见光光谱和机器学习的温室黄瓜霜霉病严重度定量估算[J]. 光谱学与光谱分析, 2020, 40(1): 227-232.
ZHANG Lingxian, TIAN Xiao, LI Yunxia, et al. Quantitative estimation of downy mildew severity in greenhouse cucumbers using visible light spectrum and machine learning[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(1): 227-232. (in Chinese with English abstract)
- [4] 张开兴, 吕高龙, 贾浩, 等. 基于图像处理和 BP 神经网络的玉米叶部病害识别[J]. 中国农机化学报, 2019, 40(8): 122-126.
ZHANG Kaixing, LV Gaolong, JIA Hao, et al. Identification of corn leaf diseases based on image processing and bp neural network[J]. *Journal of Chinese Agricultural Mechanization*, 2019, 40(8): 122-126. (in Chinese with English abstract)
- [5] 李天华, 孙萌, 丁小明, 等. 基于 YOLOv4+HSV 的成熟期番茄识别方法[J]. 农业工程学报, 2021, 37(21): 183-190.
LI Tianhua, SUN Meng, DING Xiaoming, et al. Maturity stage tomato recognition method based on YOLOv4+HSV[J].

- Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(21): 183-190. (in Chinese with English abstract)
- [6] 刘毅君, 何亚凯, 吴晓媚, 等. 基于改进 Faster R-CNN 的马铃薯发芽与表面损伤检测方法[J]. 农业机械学报, 2024, 55(1): 371-378.
- LIU Yijun, HE Yakai, WU Xiaomei, et al. Detection method for potato sprouting and surface damage based on improved Faster R-CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 371-378. (in Chinese with English abstract)
- [7] LE V N T, TRUONG G, ALAMEH K. Detecting weeds from crops under complex field environments based on Faster RCNN[C]//2020 IEEE Eighth International Conference on Communications and Electronics (ICCE). Phu Quoc Island, Vietnam: IEEE, 2021: 350-355.
- 杨坚, 钱振, 张燕军, 等. 采用改进 YOLOv4-tiny 的复杂环境下番茄实时识别[J]. 农业工程学报, 2022, 38(9): 215-221.
- YANG Jian, QIAN Zhen, ZHANG Yanjun, et al. Real-time tomato recognition in complex environments using improved YOLOv4-tiny[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(9): 215-221. (in Chinese with English abstract)
- [9] ZHOU Z X, SONG Z Z, FU L S, et al. Real-time kiwifruit detection in orchard using deep learning on Android™ smartphones for yield estimation[J]. *Computers and Electronics in Agriculture*, 2020, 179: 105856.
- [10] YAN B, FAN P, LEI X, et al. A real-time apple targets detection method for picking robot based on improved YOLOv5[J]. *Remote Sensing*, 2021, 13(9): 1619.
- 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)
- [12] 许鑫, 张力, 岳继博, 等. 农田环境下无人机图像并行拼接识别算法[J]. *农业工程学报*, 2024, 40(9): 154-163.
- XU Xin, ZHANG Li, YUE Jibo, et al. Parallel mosaic recognition algorithm for UAV images in farmland environment[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2024, 40(9): 154-163. (in Chinese with English abstract)
- [13] 王政, 许兴时, 华志新, 等. 融合 YOLOv5n 与通道剪枝算法的轻量化奶牛发情行为识别[J]. 农业工程学报, 2022, 38(23): 130-140.
- Wang Zheng, Xu Xingshi, Hua Zhixin, et al. Lightweight recognition for the oestrus behavior of dairy cows combining YOLOv5n and channel pruning[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(23): 130-140. (in Chinese with English abstract)
- [14] 岳凯, 张鹏超, 王磊, 等. 基于改进 YOLOv8n 的复杂环境下柑橘识别[J]. 农业工程学报, 2024, 40(8): 152-158.
- YUE Kai, ZHANG Pengchao, WANG Lei, et al. Recognizing citrus in complex environment using improved YOLOv8n[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(8): 152-158. (in Chinese with English abstract)
- [15] XIAO B, NGUYEN M, YAN W Q. Fruit ripeness identification using YOLOv8 model[J]. *Multimedia Tools and Applications*, 2024, 83(9): 28039-28056.
- [16] WANG K, LIEW J H, ZOU Y, et al. Panet: Few-shot image semantic segmentation with prototype alignment[C]//Proceedings of the IEEE/CVF international conference on computer vision. Paris, France: IEEE, 2019: 9197-9206.
- [17] 高昂, 梁兴柱, 夏晨星, 等. 一种改进 YOLOv8 的密集行人检测算法[J]. 图学学报, 2023, 44(5): 890-898.
- GAO Ang, LIANG Xingzhu, XIA Chenxing, et al. An improved dense pedestrian detection algorithm for YOLOv8[J]. *Journal of Graphics*, 2023, 44(5): 890-898. (in Chinese with English abstract)
- [18] 吴湘宁, 贺鹏, 邓中港, 等. 一种基于注意力机制的小目标检测深度学习模型[J]. *计算机工程与科学*, 2021, 43(1): 95-104.
- WU Xiangning, HE Peng, DENG Zhonggang, et al. A deep learning model for small target detection based on attention mechanism[J]. *Computer Engineering and Science*, 2021, 43(1): 95-104. (in Chinese with English abstract)
- [19] 陈超, 吴斌. 一种改进残差深度网络的多目标分类技术[J]. *计算机测量与控制*, 2023, 31(7): 199-206.
- CHEN Chao, WU Bin. An improved residual depth network for multi-target classification[J]. *Computerized Measurement and Control*, 2023, 31(7): 199-206. (in Chinese with English abstract)
- [20] 王大阜, 王静, 石宇凯, 等. 基于深度迁移学习的图像隐私目标检测研究[J]. 图学学报, 2023, 44(6): 1112-1120.
- WANG Dafu, WANG Jing, SHI Yukai, et al. Research on image privacy target detection based on deep migration learning[J]. *Journal of Graphics*, 2023, 44(6): 1112-1120. (in Chinese with English abstract)
- [21] DONG X, LIU Y, DAI J, Concrete surface crack detection algorithm based on improved YOLOv8[J]. *Sensors*, 2024, 24(16): 5252.
- [22] ZHANG Z, LI Y, BAI Y, et al. Convolutional graph neural networks-based research on estimating heavy metal concentrations in a soil-rice system[J]. *Environmental Science and Pollution Research*, 2023, 30(15): 44100-44111.
- [23] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. Venice, Italy: IEEE, 2017: 2980-2988.
- [24] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations[C]//Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition. Seattle, WA, USA: IEEE, 2020: 1580-1589.
- [25] WANG J, LIAO X, WANG Y, et al. M-SKSNet: Multi-Scale spatial kernel selection for image segmentation of damaged road markings[J]. *Remote Sensing*, 2024, 16(9): 1476.
- [26] CAI J, LIU X, HUANG R, et al. A review of research intrusion monitoring technology for railway carrier equipment[C]//International Conference on Electrical and Information Technologies for Rail Transportation. Singapore: Springer Nature Singapore, 2023: 546-554.
- [27] 张利丰, 田莹. 改进 YOLOv8 的多尺度轻量型车辆目标检测算法[J]. *计算机工程与应用*, 2024, 60(3): 129-137.
- ZHANG Lifeng, TIAN Ying. Improved multi-scale lightweight vehicle target detection algorithm for YOLOv8[J]. *Computer Engineering and Applications*, 2024, 60(3): 129-137. (in Chinese with English abstract)
- [28] ZHANG C, ZHANG Y, CHANG Z, et al. Sperm YOLOv8E-TrackEVD: A novel approach for sperm detection and tracking[J]. *Sensors*, 2024, 24(11): 3493.
- [29] YANG G, LEI J, ZHU Z, et al. AFPN: asymptotic feature pyramid network for object detection[C]//2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Oahu, Hawaii: IEEE, 2023: 2184-2189.
- [30] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference

- on computer vision and pattern recognition. Music City Center, Nashville TN: IEEE, 2023: 7464-7475.
- [31] LI Y, ZENG J, SHAN S, et al. Occlusion aware facial expression recognition using CNN with attention mechanism[J]. IEEE Transactions on Image Processing, 2018, 28(5): 2439-2450.
- [32] 郝建军, 邝振凯, 杨淑华, 等. 采用改进 YOLOv3 算法检测青皮核桃[J]. 农业工程学报, 2022, 38(14): 183-190.
- HAO Jianjun, BING Zhenkai, YANG Shuhua, et al. Detection of green walnut by improved YOLOv3[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(14): 183-190. (in Chinese with English abstract)
- [33] 徐杨, 熊举举, 李论, 等. 采用改进的 YOLOv5s 检测花椒簇[J]. 农业工程学报, 2023, 39(16): 283-290.
- XU Yang, XIONG Juju, LI Lun, et al. Detecting pepper cluster using improved YOLOv5s[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(16): 283-290. (in Chinese with English abstract)
- [34] NORKOBIL SAYDIRASULOVICH S, ABDUSALOMOV A, JAMIL M K, et al. A YOLOv6-based improved fire detection approach for smart city environments[J]. Sensors, 2023, 23(6): 3161.
- [35] 王歲, 洪学峰, 雷松贵. 基于 MR 的机电装备智能检测维修[J]. 图学学报, 2022, 43(1): 141-148.
- WANG Wei, HONG Xuefeng, LEI Songgui. Intelligent inspection and maintenance of electromechanical equipment based on MR[J]. Journal of Graphics, 2022, 43(1): 141-148. (in Chinese with English abstract)
- [36] LU D, WANG Y. MAR-YOLOv9: A multi-dataset object detection method for agricultural fields based on YOLOv9[J]. Plos One, 2024, 19(10): e0307643.

Algorithm for the target detection of phalaenopsis seedlings using lightweight YOLOv8n-Aerolite

ZHAI Yongjie, TIAN Jiming, CHEN Penghui, WANG Jiahao, HU Dongyang, XU Dawei

(Department of Automation, North China Electric Power University, Baoding 071003, China)

Abstract: This study aimed to improve the efficiency of visual detection for the seedling gripping points in the automated rapid propagation of Phalaenopsis orchids, particularly on edge devices with limited computational resources and storage capacities. A lightweight algorithm of object detection was introduced (named YOLOv8n-Aerolite), in order to balance the high detection accuracy and the low computational complexity. As such, the algorithm was suitable for real-time applications on devices with restricted hardware capabilities. The StarNet was then developed as the backbone network, due to its efficient extraction of various features. An SPPF_LSKA (Large-Separable-Kernel-Attention) layer was incorporated to further optimize the model. The computational demands of the model were significantly reduced to maintain high precision during detection. The large-separable-kernel was designed to enhance the performance of the model, in order to process the key visual features with minimal resource usage. There was a critical advancement for the edge devices. Additionally, a new C2f_Star module was implemented to combine with the StarBlock in the network's neck for better feature fusion. Some finer details were then detected, such as the small and intricate points of seedling gripping. The C2f_Star was also integrated to introduce multi-scale feature processing. The gripping points were distinguished in the dense environments, where the seedlings were closely spaced. The detection head was also redesigned to include a lightweight shared convolutional layer structure, referred to as Detect_LSCD (Lightweight Shared Convolutional Detection Head). There was a notable increase in the detection speed, in order to reduce the overall size of the model. Specifically, the optimizations were fully realized to perform efficiently under resource-limited environments. The improved YOLOv8n-Aerolite algorithm was then tested on the image dataset of Phalaenopsis seedling. Experimental results showed that the improved model was achieved with an average inference speed of 435.8 frames per second, highly suitable for real-time applications. The improved model was also marked as one of the greatest available options for the detection of edge-based seedlings than before. The detection accuracy of the improved model reached 91.1%, particularly with an impressive precision of 91.6% to detect the small targets black tuber. The gripping points of the seedlings also validated the reliability of the improved model in practical deployments. Such high accuracy was achieved in the detection of small targets, indicating better suitability for tasks where the precise targeting of small objects was essential. In addition, the weight file size of the improved model was compressed to just 3.1 MB, particularly for the deployment on the edge devices where the storage capacity was constrained. A series of practical gripping experiments were conducted to further validate the algorithm. A success rate of 78% was obtained for the high efficiency of the improved model in real-world scenarios. The generalizability of the YOLOv8n-Aerolite algorithm was also tested on the 3D reconstructed dataset of phalaenopsis seedlings, similar to the detection of small targets. The results showed that the mAP_{0.5} increased by 1.6 percentage points compared to the original YOLOv8n model. The better performance of the improved model was obtained across different datasets. The cross-dataset testing confirmed that the robustness and adaptability were suitable for a variety of detection tasks. In conclusion, the YOLOv8n-Aerolite algorithm significantly advanced the field of automated crop propagation. A highly efficient, accurate, and lightweight solution was also provided for the visual detection. The finding can serve as a valuable reference to develop scalable and automated technologies, especially for small-scale crops like Phalaenopsis orchids. YOLOv8n-Aerolite can also fully meet the needs of edge computing environments, particularly for the broader applications of agricultural automation.

Keywords: deep learning; YOLOv8n; phalaenopsis seedlings; lightweight; detection speed; small object detection