

DOI:10.19651/j.cnki.emt.2417642

融合轻量化 YOLOv8-Pose 的烟草茎叶角检测算法^{*}

高坤¹ 李军营² 梁虹¹ 马二登² 张宏²

(1. 云南大学信息学院 昆明 650504; 2. 云南省烟草农业科学研究院 昆明 650021)

摘要: 茎叶角检测是烟草表型检测的重要部分,在烟草农业的增产增效和疾病预防方面有重要的意义。针对不同环境下的人工茎叶角检测效率低、周期长、检测不方便等问题,设计并构建了轻量化的烟草茎叶角检测模型 FAL-YOLO。该算法构建 FAI 主干网络结构来充分减少计算量和特征冗余,增加语义信息的利用效率。构建了融合空间注意力和通道注意力 SA 注意力模块的 SAC 检测头模块,进一步减少参数量和增强对茎叶角特征的感知能力。引入 GSCConv 轻量化卷积降低模型复杂度和模型参数量。引入 MPD-IoU 损失函数来提升改进模型整体性能。采用自建的烟草茎叶角检测数据集,开展 FAL-YOLO 模型的对比和消融实验。实验结果表明, FAL-YOLO 模型在自制数据集上的 mAP 达到了 99.2%,相比 YOLOV8-POSE 模型在 GFLOPs, Params 分别降低了 56.7% 和 52%,改进后的模型能够更快更精准的识别烟草植株茎叶角,为烟草农业选种育种智慧化提供支持。

关键词: 烟草茎叶角检测;主干网络;轻量化;金字塔池化;YOLOv8-Pose

中图分类号: TN911.73; TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.6

Tobacco stem and leaf angle detection algorithm integrated with lightweight YOLOv8-Pose

Gao Kun¹ Li Junying² Liang Hong¹ Ma Erdeng² Zhang Hong²

(1. School of Information Science and Technology, Yunnan University, Kunming 650504, China;

2. Yunnan Academy of Tobacco Agriculture Science, Kunming 650021, China)

Abstract: Stem and leaf angle detection is an important part of tobacco phenotype detection, which is of great significance in increasing yield and efficiency and disease prevention in tobacco farming. Aiming at the problems of low efficiency, long cycle time and inconvenience of manual stem and leaf angle detection in different environments, a lightweight tobacco stem and leaf angle detection model, FAL-YOLO, was designed and constructed. The algorithm builds the FAI backbone network structure to sufficiently reduce the amount of computation and feature redundancy, and increase the efficiency of using semantic information. The SAC detection head module, which integrates the spatial attention and channel attention SA attention modules, is designed to further reduce the number of parameters and improve the perception of stem and leaf angle features. GSCConv lightweight convolution is introduced to reduce model complexity and the number of model parameters. The MPD-IoU loss function is introduced to improve the overall performance of the model. A self-constructed tobacco stem and leaf angle detection dataset is used to carry out the comparison and ablation experiments of the FAL-YOLO model. The experimental results show that the mAP of the FAL-YOLO model on the self-constructed dataset reaches 99.2%, compared with the YOLOV8-POSE model in the GFLOPs, the Params are reduced by 56.7% and 52%, respectively, and the improved model is capable of identifying the stem and leaf angles of tobacco plants faster and more accurately, which can support the wisdom of tobacco agricultural seed selection and breeding.

Keywords: tobacco stem and leaf angle detection; backbone network; lightweight; pyramid pooling; YOLOv8-Pose

0 引言

在现代农业生产中,作物表型生长状况的精确监测是

提升产量和优化品质的关键环节。对于烟草这类高经济价值作物^[1],植株的形态结构直接影响着其光合作用效率、养分吸收与分配情况,进而决定最终的产量和品质水平。烟

收稿日期:2024-12-16

* 基金项目:中国烟草总公司重大科技项目(110202401003 (JY-03))资助

草植株的表型主要包括植株高度、茎干粗细、叶片形状与颜色、叶片数量与分布、根系形态以及开花结实特征等,这些特征直接影响其生长状况和环境适应性。其中,烟草茎叶角在植株生长中具有重要作用,它决定了叶片对光照的接收效率,影响光合作用,同时调节空气流通性,降低病害风险。合理的茎叶角还能优化水分与养分传输效率,提高植株稳定性和抗风能力。因此,精确监测茎叶角度具有重要的农业管理价值^[2]。

茎叶角作为烟草植株生长状态的重要表征^[3],对于提升烟草经济效益,促进可持续发展具有重要意义。然而,在大田环境下,人工测量是最早应用的方法之一,其优势在于对于数量较少,环境较差的烟田能够进行较低成本的测量,但在大面积和精度要求较高的情况下,由于烟草植株叶片的多样性、光照变化及复杂的环境干扰,传统手动测量方法存在反应速度慢、劳动强度高、易受其他因素干扰、处理速度缓慢、花费较大等缺点。

近年来,深度学习技术在目标检测领域的快速发展为农业领域的形态结构监测提供了新的解决方案^[4]。YOLOv8-Pose (you only look once version 8-pose) 是 YOLO 系列在姿态估计领域的最新进展,凭借高精度和实时性表现,在动植物姿态检测^[5]、医疗康复和增强现实等领域得到了广泛应用。在烟草表型检测任务中,重要的有两种方法。一是传统的表型检测算法如阈值分割(如 Otsu)、边缘检测(如 Canny),肖恒树等^[6]对 YOLOv8 基础模型进行改进,使用无人机拍摄遥感影像实现大田烟株数量检测模型,能够精准检测大田烟株数量。能够如高理文等^[7]发表一种动态阈值的药用植物叶片图像分割,使用动态阈值分割算法对植物表征进行分类。唐家阳^[8]发表的基于深度学习的叶菜类植物点云分割方法研究与实现,能够对叶菜类的植物进行点云分割提取表型特征。皮鸣寰等^[9]发表在湖北民族大学学报(自然科学版)上,能够实现关于深度学习在植物表型中,茶叶萎凋过程中含水率的检测。深度学习中的目标检测算法可分为多阶段和单阶段两类。多阶段目标检测算法代表有 Faster R-CNN^[10]、Mask R-CNN^[11]等。吴雪梅等^[12]对 Faster R-CNN 网络改进进行改进,实现了对成熟期烟叶分级和准确度提升。王晓婷等^[13]对 Mask R-CNN 进行改进,发表的植物表型智能检测算法能够对西瓜突变体生长情况的性状表型进行检测。袁山等^[14]对基础 Mask R-CNN 模型进行改进,能够解决叶片分割不佳的问题。

相比之下,单阶段目标检测方法如 YOLO (you only look-once)^[15] 和 SSD (single shot multibox detector)^[16] 从整幅图像中提取特征的同时对物体的位置和类别进行预测,具备更快的推理速度。孙君峰等^[17]对 YOLOv8 进行改进,构建车辆与行人检测算法模型,使用心得注意力机制增强算法的多尺度特征提取能力。改进降低模型复杂度,综合考虑不同类型大小的车辆与行人,使其能够精确检测多

类小尺寸目标及行人。向雷等^[18]通过对城市障碍物检测进行研究,对 YOLOv8 算法进行一系列改进,设计全新改进主干网络,降低网络复杂度,使用融合特征金字塔和添加小目标检测头,更好地捕捉小尺寸障碍物的特征。改进模型较原模型提高了检测精度和速度。朱周华等^[19]改进 YOLOv8-POSE 模型,融合 FasterBlock 模块,对模型进行轻量化与精度提升,在模型大小和准确性之间取得了良好的平衡,实现了对不同驾驶行为的识别与分心驾驶行为检测精度的提升。罗智杰等^[20]改进 YOLO 模型,融合 SA 自注意力机制,对 YOLOv8-POSE 模型进行检测头的轻量化,并提升在遮挡情况下模型对运动姿势检测的准确率。实现对关键点遮挡、复杂多变姿态等识别难度较高的情况下,识别精度的提高。

受上述相关研究启发,在对比当前经典检测算法及其轻量化改进模型(Mobilevitv1, Mobilevitv2, Mobilenetv4, Ghostv1, Ghostv2, YOLOv5-pose, YOLOv8-pose, YOLOv10-pose 等)姿态检测算法的基础上,本文针对参数量过多、特征冗余等问题,提出基于 YOLOv8 的融合轻量化的 FAL-YOLO 茎叶角检测模型。主要贡献如下:

1) 通过摄影拍摄四块云南省烟草农业科学院高品质烟田,构建了田间烟草植株的茎叶角图像数据集共计 1 804 张,并进行数据标注与数据增强。

2) 改进算法设计了轻量化高级特征提取 FAI 主干网络(fast-and-improve, FAI),有效屏蔽非烟草茎叶角区域的复杂环境干扰。

3) 融合自注意力机制(self-attention, SA)的主干注意力 SAC(stem attention capturer)检测头模块,提高关键区域位置检测精度。

4) 采用 GSConv 轻量化卷积(gsconv lightweight convolution, GSConv)对 Neck 网络进行优化,在保持模型检测精度的同时有效减少模型复杂度和参数量。

5) 采用 MPDIoU 损失函数(minimum point distance intersection over union, MPDIoU)对模型的回归定位损失进行优化,进一步提升茎叶角检测框定位精度。

通过一系列实验验证,本研究提出的改进算法在烟草茎叶角关键点检测任务上的平均检测精度达到了 99.2%,同时检测帧率达到 128.6 fps,可以进行实时检测。

1 材料和方法

1.1 实验地点

实验地点为中华人民共和国云南省玉溪市的烟草农业科学院高品质烟田。该地区气候湿润,环境对烟草植株的生长成熟状况极为有利。同时该地区有着丰富的烟草种植的经验和广阔的种植面积,为在该区域进行数据收集提供了优质的条件。如图 1 所示为试验区域和数据采集现场。

1.2 数据集构建

实验所需数据为烟草植株正面的多角度摄影影像,为



图 1 试验区域和数据采集现场

Fig. 1 Experimental area and data collection site

确保获取多角度高品质的烟草茎叶角图像数据,分别在晴天、阴天、晌午和傍晚对不同类型烟草茎叶角在试验田中进行拍摄采集,以提高数据集图片质量和图片种类,扩大不同天气情况覆盖范围。摄影设备分辨率为 $3\ 048 \times 4\ 064$,拍摄距离分别设置为距离烟株茎叶角 0.3 m 、 0.5 m 处拍摄,共计采集了烟草农业科学院高品质烟田约1 800张图像。如图2所示为在不同天气和不同距离下拍摄的图像。



图 2 烟草茎叶角数据集

Fig. 2 Tobacco stem and leaf angle dataset

1.3 数据标注与增强

本研究用于神经网络训练的数据集采用COCO(coco annotation format)格式,使用Labelme工具进行烟株茎叶角的标注。Labelme是用于图像标注的开源工具,用户可以通过提供的图像可视化界面添加、编辑、修改数据集的标注和类别。标注框标签为“stem-leaf”,关键点标签为“leaf”、“stem”和“angle”。对所需烟株烟草茎叶角进行关键点和检测框标记,关键点位置和坐标框的标注结果会以

JSON(javascript object notation, JSON)的文件格式进行保存。随后,将数据集中的json文件转化为xml文件,转化完成后对数据集进行划分,将其分为训练集、验证集,其中训练集占比为80%,而验证集占比为20%。所有的图像和标注文件均用于进行姿态检测模型的训练和验证。

数据增强是深度学习中提升模型性能和鲁棒性的重要技术。它可以增加数据多样性、缓解过拟合、优化训练数据分布,并提升模型对未知环境的适应能力。数据增强还能在有限数据集上扩展样本规模,降低对大规模标注数据的依赖,是提高模型泛化能力和迁移学习性能的关键手段。本研究运用数据增强方法为将图像进行添加高斯噪声、随机翻转、亮度变化和随机裁剪等来增加样本的多样性,丰富物体信息。

2 相关算法

2.1 YOLOv8-Pose 算法介绍

YOLOv8-Pose是YOLOv8的一个变体,专门用于关键点检测任务,广泛应用于人体姿态估计等领域。YOLOv8-Pose模型同样由Backbone(骨干)、Neck和Head三个部分组成,但其结构经过调整,通过优化骨干、特征聚合网络以及特定任务的解耦头,YOLOv8-Pose能够实现对多种姿态的高效检测,以适应关键点检测任务的需求。

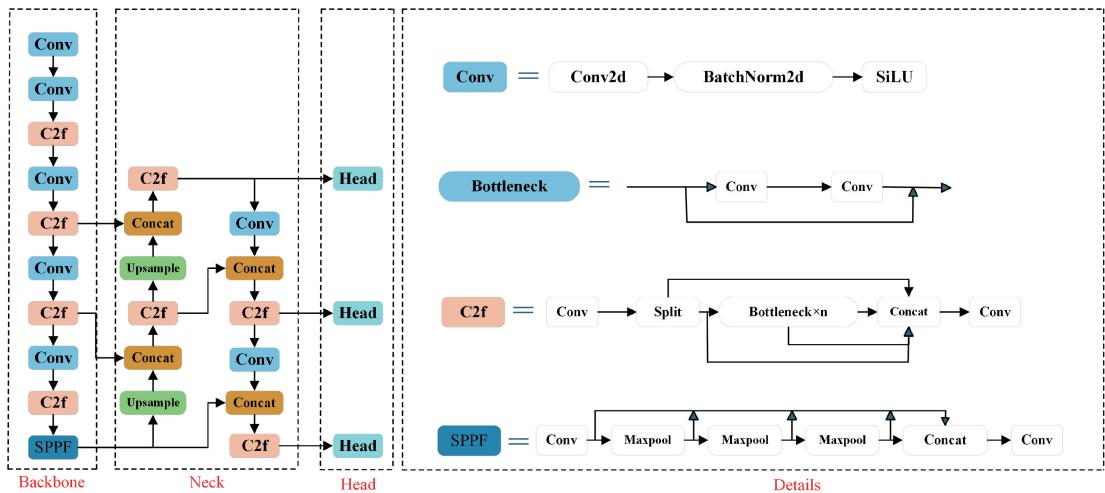
Backbone骨干部分继续继承和使用YOLOv8的CSPDarkNet结构,继续采用Conv模块、C2f模块和SPPF模块对图像进行特征提取。C2f模块通过融合CSPNet的特征分流与残差结构,借助Bottleneck单元,在保证提取高效特征的同时保持轻量化。SPPF模块通过空间金字塔池化对特征图进行处理,使得不同尺度的特征能够统一输出,提升模型在多尺度上的鲁棒性。

Neck部分依旧沿用了YOLOv8的PANet结构,增强了不同层级特征的融合能力,尤其是通过多层特征传递使得低层次的信息能够有效传递至高层,帮助模型在复杂背景下仍能准确地检测关键点。

Head部分针对关键点检测任务进行了调整。与YOLOv8的解耦头类似,YOLOv8-Pose采用了一个专门设计的解耦头,将关键点检测与位置回归任务分离。分类损失继续使用BCE Loss(binary cross entropy loss),而位置回归则引入DFL(distribution focal loss)结合CIoU Loss,对位置和关键点进行更加精确的回归和定位。网络模型如图3所示。

2.2 算法存在的问题

YOLOv8-Pose算法中使用了大量的卷积和池化等模块,虽然在姿态检测方面拥有较高的精度和检测速度,但是大量的卷积和池化增加了模型的复杂度和计算量,降低了运行速度,需要较多的算力才能达到高精度的检测结果。在处理现场烟株茎叶角度检测的问题时,考虑到需要将模型部署到移动设备上,而这类设备的计算能力有限,



因此本文选择了 YOLOv8-Pose 作为基础模型，并进行轻量化改进。在保持模型具有较高检测精度的同时，显著降低模型的参数量和计算量，以适应移动端应用的算力需求。此外，在烟株试验田实际检测的过程中，往往烟草茎叶角处于复杂场景下，而且其角度形态多变、易受遮挡、与背景相似和天气环境影响，YOLOv8-Pose 模型在提取特征时极易忽略小目标和被遮挡对象的信息，时常会造成漏检误检问题。

3 改进模型 FAL-YOLO

3.1 FAI 主干网络

针对 YOLOv8-Pose 算法在复杂场景下由于烟草茎叶角的角度形态多变、易受遮挡和环境影响等问题本文从以下 4 个方面改进模型。

1) 为了解决这些部分语义信息较少导致的特征提取时对高层次特征提取不足问题，设计了一种新型的轻量化主干网络 FAI 结构。利用 FastNet 主干网络削减冗余计算和内存访问，更高效地提取空间特征并结合 AIFG 模块可以关注于处理高级特征和捕捉图像中概念实体之间的联系的能力，提高主干网络面对高级特征图的学习效果。

2) 为解决常规检测头无法适应茎叶角不规则排列和外观与周围环境接近难以分辨的特点，在检测头构建 SAC 检测头模块。本文基于 SA 注意力机制设计了 SAC 检测头对烟草植株特征信息进行有效提取，而且降低并行分支造成额外计算量。

3) 引入 MDP-IoU 损失函数代替 YOLOv8 中原有的损失函数，处理数据时能够对重叠区域、中心点距离以及宽度和高度的偏差进行综合考虑，反映预测框和真实框之间的差异更准确，提升算法的泛化性能。

4) 引入轻量化卷积 GSconv，使模型的计算量和参数量都有明显减少。同时平衡不同样本质量之间的训练过程，

加快模型收敛速度，提升算法的泛化性。

改进后的网络结构如图 4 所示。

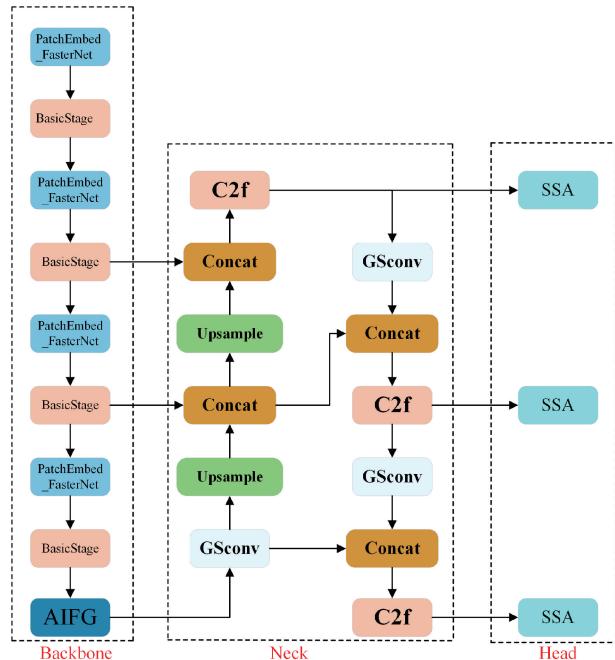


图 4 FAL-YOLO 算法模型

Fig. 4 FAL-YOLO algorithm model

1) 改进的骨干网络结构

在烟草茎叶角检测任务中，YOLOv8-Pose 骨干网络虽然拥有较好的关键点检测能力和多尺度特征融合能力，适用于识别不同的姿势和动作，然而其模型在复杂背景中可能会受到干扰，导致关键点检测的准确性下降，尤其是在背景颜色和被检测事物相似的情况下。为了更聚焦监测关注的烟草茎叶角检测区域，提高模型对烟草茎叶角的高级特征的注意力的同时加快推理速度，本文设计了一种由 FastNet 主干网络和基于注意力的内尺度特征交互 AIFG

模块组建的 FAI 高级特征处理主干网络,可以更有效地保留和处理目标区域的高级特征信息,提升茎叶角检测的检测精度和检测速度。

本文以 FAI 替换原 YOLOv8-Pose 中的主干网络,FAI 由 4 阶层次结构构建而成,利用其中卷积层进行下采样,并利用改进的金字塔池化 AIFG 进行高级特征提取。以 FastNet 主干网络替换原主干网络,以便进行更高效的特征融合以及平衡计算效率和性能。以 AIFG 金字塔池化替换原有的 SPPF 金字塔模块,可以更高效地提取高级特征信息并且减少计算量。由此构成本算法的轻量化高级特征处理网络,用于快速且有效地提取烟株茎叶角特征。FAI 高级特征处理网络整体架构如图 5 所示。

2) FastNet 网络结构

YOLOv8 原有的主干结构在进行连续下采样时,容易导致小目标的细节信息丢失,进而降低目标识别率。此外, YOLOv8 的特征层设计在捕捉复杂茎叶形态的细微变化方面存在不足,影响了检测精度。为有效缓解特征提取过

程中小目标特征信息的流失,提升检测性能,本文引入 FastNet 主干网络^[21] 替代 YOLOv8 中的主干网络。FastNet 网络结构如图 6 所示。

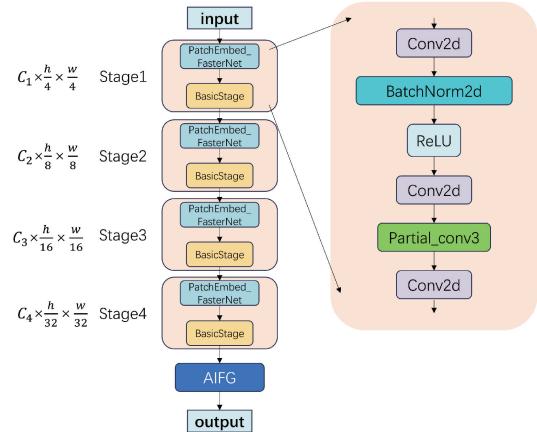


图 5 FAI 算法模型

Fig. 5 FAI algorithm model

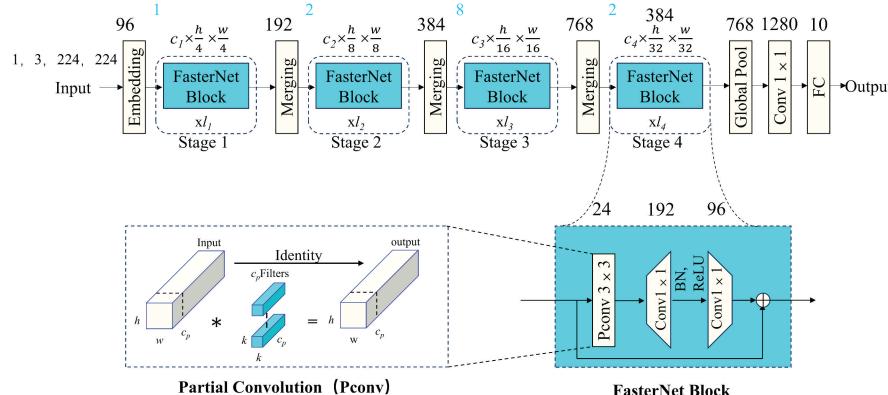


图 6 FastNet 网络结构

Fig. 6 FastNet network architecture

FastNet 主干网络的架构通过结合多种模块实现高效的特征提取。其中 PatchEmbed 模块将输入图像分割为多个小块(patches),通过空间卷积操作,将输入特征图的空间维度转换为更高的深度维度,同时保留重要的空间信息。在特征提取的后续阶段,PatchMerging 模块通过逐层合并特征图来增强特征的深度,并有效降低空间维度。这一过程不仅提高了特征的表达能力,还增强了模型对复杂形态的捕捉能力,使其能够处理多样化的目标姿态和背景。通过使用 FastNet,可以在较低的计算成本和较低的计算量下实现高效的特征提取,并且其结构优化允许更快的推理速度,从而适应实时应用场景。

3) AIFG 金字塔模块

在烟草茎叶角检测中,角度检测识别精度深受背景中其他相似对象干扰,且不同烟草植株之间也存在一定的相似性。YOLOv8 中原有的 SPPF 模块(spatial pyramid pooling fast,SPPF)在处理上述问题时可能因为特征分辨

率的损失、全局上下文信息捕获能力不足以及对几何形状和姿态变化的适应能力弱等原因,导致对目标的区分能力下降,影响检测的准确性和鲁棒性。

为解决上下文信息捕获能力不足,适应能力弱的问题,本文提出自适应交互特征融合网(adaptive interaction feature grid,AIFG)替换 YOLOv8 中的 SPPF,提升网络的表达能力和感知能力。AIFI^[22]计算方法为:

$$F_5 = R(Attn(F(S_5), F(S_5), F(S_5))) \quad (1)$$

式中:Attn()表示多头自注意力,F()为矩阵扁平化变换,R()为F()的逆操作,代表将特征的形状恢复为与S5相同。

AIFG 结合了自注意力机制和选择性特征交互,将 AIFI 模块和轻量化卷积模块融合^[23]。在该结构中,AIFG 模块利用自注意力机制来处理图像中的高级特征。自注意力是一种注意力机制,特别适用于处理具有丰富语义信息的高级图像特征。模型在处理特定部分的数据时,使用

自注意力机制可以同时考虑到其他的数据相关部分。而 AIFG 还可以进行选择性特征交互: AIFG 模块专注于在 S5 级别(即高级特征层)上进行内部尺度交互, 避免在低级特征层进行相同的交互。这是基于认识到高级特征层包含更丰富的语义概念, 能够更有效地捕捉图像中的概念实

体间的联系, 而低级特征缺乏必要的语义深度, 且可能导致数据处理上的重复和混淆。总而言之, AIFG 模块通过自我注意力机制专注于处理高级图像特征, 从而提高模型在对象检测和识别方面的性能, 同时减少不必要的计算消耗。AIFG 金字塔模块结构如图 7 所示。

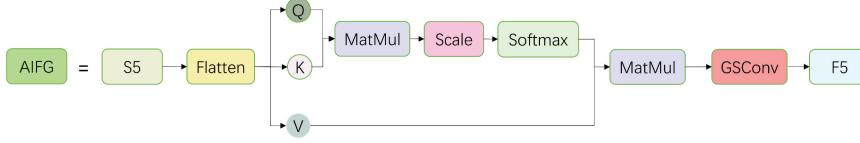


图 7 AIFG 模块结构

Fig. 7 AIFG module structure

3.2 SAC 检测头模块

YOLOv8 的头部采用解耦头结构, 这种结构将分类与回归任务分开完成, 其中每一个分支都是由两个 3×3 和一个 1×1 的卷积完成, 专注于各自分支的任务。然而大量重复的 3×3 卷积导致参数量骤增, 造成大量冗余计算。根据参数共享原理, 基于 GSConv 卷积重新设计的 SAC 模块。该检测头在 YOLOv8n 模型的基础上融入了 SA 自注意力机制, 以通过高效的高分辨率特征提取的方式, 提升烟草植株茎叶角检测的效率和效果。SAC 检测头的构成由内存高效的 SA 自注意力模块^[24]和 Conv 模块进行串联构成。自注意力公式为:

$$Att(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

式中: $Att(Q, K, V)$ 是得到的注意力的值, Q 、 K 、 V 分别代表查询向量(query vector)矩阵、键向量(key vector)矩阵和值向量(value vector)矩阵, 这 3 个矩阵中每一行分别代

表一个对应的向量。

首先, 通过 1×1 卷积层对特征维度进行压缩, 以减少数据的复杂度和自注意力计算的负荷。随后通过多头自注意力层(multi-head self-attention, MHSA), 模块从输入特征中提取长距离依赖关系, 增强全局感知能力。之后在自注意力模块的各层之间加入批量归一化层, 能够有效稳定训练过程, 提高模型的收敛速度。通过残差连接, 输入特征可直接加到输出特征上, 使得信息得以深层传递, 减轻梯度消失问题的影响。同时, 残差连接增强了模型的鲁棒性和训练稳定性。最后与自注意力模块串联的卷积模块, 采用标准的卷积操作, 进一步提取特征局部性信息。该模块负责捕捉运动姿态中的细微变化, 保证模块在关注全局关系的同时, 不忽略重要的局部细节。这一结构不仅增强了模型对全局信息的捕获能力, 还提升了关键点定位的准确性, 同时控制了计算复杂度, 适配 YOLOv8-Pose 的轻量化特性。SAC 检测头模块结构如图 8 所示。

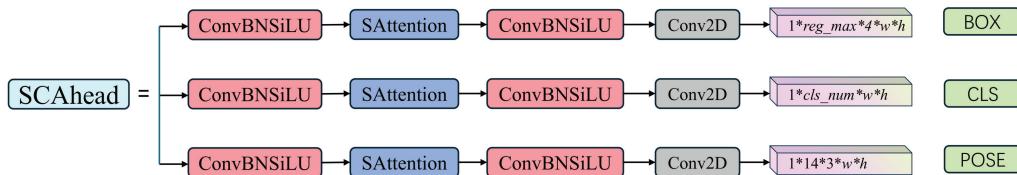


图 8 SAC 模块结构

Fig. 8 SAC module structure

自注意力模块通过 1×1 卷积层降低特征维度, 减少自注意力的计算负荷; 利用多头自注意力层捕捉长距离依赖关系, 增强模型的全局感知能力; 再经第二个 1×1 卷积层恢复特征维度, 保持信息的完整性。同时, 批量归一化层稳定训练, 残差连接促进信息深层传递, ReLu(rectified linear unit)激活函数引入非线性因素, 进一步提升模型的表达能力。自注意力模块结构如图 9 所示。

3.3 GSConv 模块

YOLOv8 检测器使用的传统步长卷积模块的模型大小和计算量对于移动应用部署, 尤其是单片机部署模型方面有着明显的不足。为了在保持模型性能的同时有效减

少模型的大小和计算量, 本文改进算法引入 GSConv 卷积模块替换 YOLOV8 主干网络中的 Conv 卷积模块。GSConv 的时间复杂度公式为:

$$Time_{GSConv} \sim O\left[W \times H \times K_1 \times K_2 \times \frac{C_2}{2}(C_1 + 1)\right] \quad (3)$$

式中: W, H 是输出特征图的宽度和高度。 $K_1 \times K_2$ 是卷积核的大小。 C_1 是每个卷积核的通道数, 也是输入特征图的通道数。 C_2 是输出特征图的通道数^[25]。GSConv 模块结构如图 10 所示。

该模块由空间到深度模块(space-to-depth, SPD)和非跨步卷积 Conv 层组合而成。SPD 模块能在保持空间信息

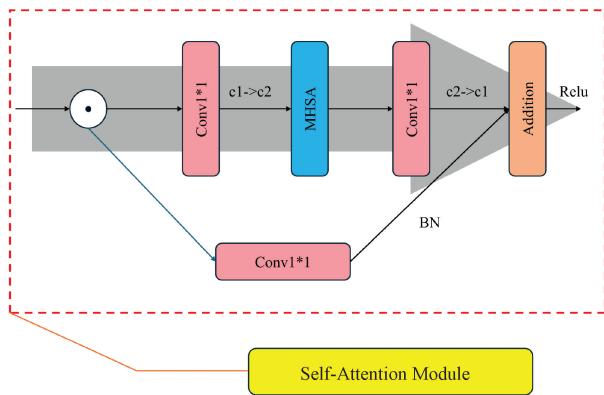


图 9 SA 注意力机制模块结构

Fig. 9 SA attention mechanism module structure

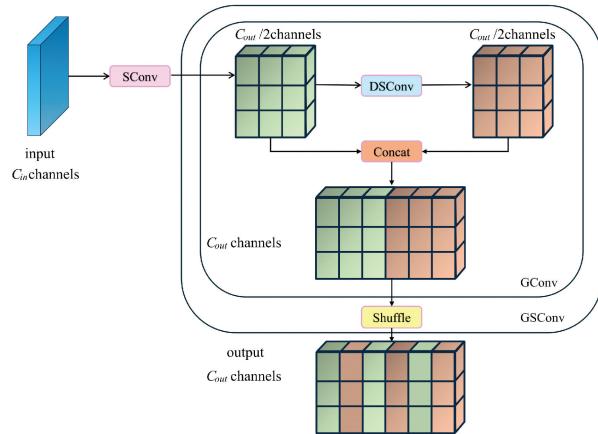


图 10 GSConv 结构

Fig. 10 GSConv structure

完整性的同时,把输入图像的空间维度转换成深度维度,随后的非跨步卷积层则保留了更多空间细粒度信息,增强了小目标检测能力。其具体操作过程为:首先对 $S \times S \times C_1$ 的输入特征图采用 2 倍尺度因子进行隔行隔列的下采样,生成 4 个尺寸均为 $S/2 \times S/2 \times C_1$ 的子特征图,然后将这 4 个子特征图在通道维度上进行拼接重排,最终形成一个尺寸为 $S/2 \times S/2 \times C_2$ 的二倍下采样特征图。

3.4 MPD-IoU 损失函数

在目标检测任务中,损失函数中的辅助边框(auxiliary bounding boxes)可以帮助模型更好地学习如何识别和定位图像中的对象。在烟草茎叶角检测任务中,由于烟草茎叶角部分在图像中可能呈现出多样的方向和与背景的相似性,传统损失函数 CIoU (complete intersection over union) 损失函数在背景相似性的适应性上存在一定局限。CIoU 损失函数在处理背景相似性问题时,容易受到背景区域的影响,导致对评估结果的不精确。MPD-IoU 损失函数^[26] 通过直接最小化预测框和真实框之间的关键点距离,使得评估结果更为准确。它的优势在于即使在存在相同宽高比但不同尺寸或位置的边界框的情况下,也能有效

地评估模型对烟草茎叶角的检测性能。此外,MPD-IoU 损失函数对于实验设计与结果分析相同而大小不同的情况表现出更高的鲁棒性,这使得它在 CIoU 损失函数可能因边缘误差而受影响的情况下,依然能够保持稳定的性能。因此,在 FAL-YOLO 模型中引入 MPD-IoU 损失函数来调整对图像的检测以此来提升模型检测的准确性。MPD-IoU 损失函数框图如图 11 所示。

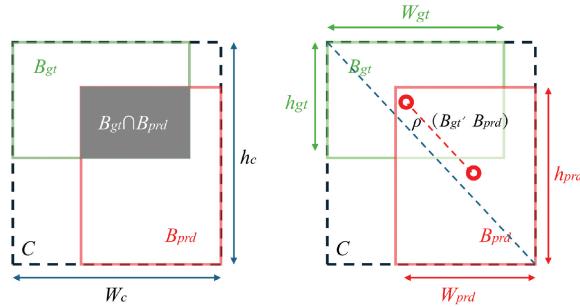


图 11 MPD-IoU 损失函数框图示意图

Fig. 11 Schematic diagram of MPD-IoU loss function

其中, B_{gt} 和 B_{prd} 分别表示预测框和真实值框。MPDIoU 计算公式如下:

输入两个任意矩形: $A, B \subseteq \mathbb{S} \in \mathbb{R}^n$, 输入图像的宽和高: w, h , Output: MPDIoU 对于 A 和 B , (x_1^A, y_1^A) , (x_2^A, y_2^A) 表示 A 的左上角和右下角点坐标, (x_1^B, y_1^B) , (x_2^B, y_2^B) 表示 B 的左上角和右下角点坐标。

$$d_1^2 = (x_1^B - x_1^A)^2 + (y_1^B - y_1^A)^2 \quad (4)$$

$$d_2^2 = (x_2^B - x_2^A)^2 + (y_2^B - y_2^A)^2 \quad (5)$$

$$MPDIoU = \frac{A \cap B}{AB} - \frac{d_1^2}{w^2 + h^2} - \frac{d_2^2}{w^2 + h^2} \quad (6)$$

从式(4)~(6)可以看出,与 IoU 损失相比,当比例大 1 时,MPDIoU 的辅助边界框不仅扩展了回归的有效范围,而且对低样本的回归具有增强效果。

4 实验设计与结果分析

4.1 训练环境与参数配置

本文实验环境配置如表 1 所示。模型训练迭代 300 轮,优化器采用 SGD 优化器。超参数设置如下:批量大小为 16,动量因子为 0.937,初始学习率为 0.01,权重衰减系数为 0.0005。

4.2 数据集介绍

本文使用的烟草茎叶角数据集图片主要来源于网络公开数据集和大田烟株培育基地实地拍摄。为了符合自然条件下烟草茎叶角的生长特性和测量茎叶角遇到的实际问题,筛选数据时不仅关注了各阶段烟草植株的生长状态,而且还对不同光照条件、背景相似等情况进行收集。采用 Labelme 软件对图像进行分类标注,共包含曲茎、成熟、断叶、光照、背景相似 5 个类别。经过添加高斯噪声、随机翻转、亮度变化和随机裁剪等,旨在增加样本的多样性,

并使模型能够识别不同角度和环境的变化,共获得图像 3460 张,按 8:2 比例随机划分数据集,得到训练集 2769 张,测试集 691 张。筛选后的部分数据集如图 12 所示。

表 1 实验环境配置

Table 1 Experimental environment configuration

名称	配置
操作系统	Windows 11
CPU	Intel(R)Core(TM) i5-12600kf
GPU	GeForce RTX 4070S
内存大小	32 G
学习框架	PyTorch 2.3.1
加速环境	Cuda 12.5
编程语言	Python3.12



图 12 部分图像

Fig. 12 Partial images

4.3 评价指标

关键点检测实验的评价指标采用精确率(precision, P),召回率(recall, R),平均精度(mean average precision, mAP),参数量(parameters, Params),模型大小(size)以及浮点运算次数(GFLOPs)作为实验的评价指标。部分指标计算公式如下,其中 mAP@0.5 是指将 IoU 设置为 0.5 时所求得的平均精度。

精确度 P 是指所有预测为正样本的结果中,被预测正确的比例,计算公式如下:

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

召回率 R 是指所有的正样本当中,被正确地预测为正样本的比例,计算公式如下:

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

AP 是指 PR 曲线(precision-recallcurve, PR)下的面积,是对不同召回率点上的准确率进行平均,计算公式如下:

$$AP = \int_0^1 P(r)dr \quad (9)$$

mAP@0.5 是指当 IoU 值等于 0.5 的时候,对所有类

的 AP 值求平均就得到了 mAP,计算公式如下:

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (10)$$

模型参数量 Params 反映模型的大小,参数量越大,占用的内存就越大。能够保证高精度的同时,所占用内存越小说明性能越好。

$$Parameters = [i \times (f \times f) \times o] + o \quad (11)$$

式中:TP 表示预测为真的正样本数量,FP 表示负样本误检测为正样本的数量,FN 是未被检测出的正样本数量。C 表示类别数目。i 表示输入大小,f 表示卷积核大小,o 表示输出大小。

4.4 模型对比评估

1) SAC 模块对比试验

改进 SCA 检测头代替原有 YOLOv8-Pose 中的 POSE 检测头,在 YOLOv8-Pose 检测头的基础上引入 SA 注意力机制,实验结果如表 2 所示。实验结果表明,与原始 YOLOv8-Pose 相比,引入 SA 注意力机制的 SCA 检测头在烟草茎叶角检测任务中表现出了更好的检测精度和鲁棒性。在相同的实验环境下,检测精度 P 提升了 1.0%,R 提升了 2.1%。这验证了 SAC 模块在增强模型特征的表达能力和提高检测性能方面的有效性。

表 2 SAC 模块对模型的影响

Table 2 The impact of SAC module on the model

模型	P/%	R/%	mAP@0.5/%	GFLOPs	Params
POSE	97.5	95.2	99	8.3	3.01
SCA	98.5	97.3	99	7.3	2.94

2) GSConv 卷积对比试验

为研究 GSConv 卷积代替原有 YOLOv8-Pose 中的 Conv 卷积的效果,同时为了平衡检测性能和模型参数量,最终本文改进算法选择对 Neck 中不同层数的 Conv 进行替换 9、16 和 19 层采用 GSConv 卷积模块替换原网络的 Conv 模块,结果如表 3 所示。实验结果显示本文 GSConv 卷积模块用较少的模型参数获得较高的检测性能,与 YOLOv8 基准算法相比烟草茎叶角的 P 提升了 0.1%,R 提升了 2.6%。

3) 损失函数对比

为了验证损失函数 MPD-IoU 的有效性,本文选取当前主流的损失函数 DIoU、SIoU 和 WIoU,在相同实验环境下,分别使用 FAL-YOLO 模型进行对比实验,对比实验结果如表 4 所示。从表 4 中可以看出,使用 MPD-IoU 作为模型边框损失函数时,实验结果具有更好的准确率和精度效果。这说明 MPD-IoU 损失函数的智能权重调整机制能够更好地平衡图像不同比例预测框之间的大小变化,降低一些预测框比例明显缺陷的烟草茎叶角样本产生的干扰,使模型更适合茎叶角检测任务。

表 3 GSConv 模块对比试验

Table 3 GSConv module comparative test

替换位置	mAP50/%	P/%	R/%
YOLOv8-Pose	0.99	97.5	95.2
9	0.989	97.2	95.8
16	98.7	97.7	96.8
19	98.9	97.7	95.8
9,16	98.8	97.3	96.7
9,19	98.9	97.4	97.2
16,19	0.99	96.7	97.8
9,16,19	0.99	97.4	97.8

表 4 损失函数对比试验

Table 4 Loss function comparison experiment

损失函数	P/%	R/%	mAP@0.5%
CIoU	95.7	95.4	98.6
DIoU	95.4	94.7	98.3
SIoU	95.7	96.4	98.4
WIoU	96.5	96.8	99.0
MPDIoU	97.8	98.4	99.2

4) 不同算法对比试验结果

为了更好地验证改进 FastNet 模型的优势,选取几种主流的模型进行一组对比实验,如 mobilevitV2、mobilenetv4、ghostv2、YOLOv5-pose 等模型,文献[9]、文献[17]、文献[19]为对 YOLOv8-Pose 提出的改进方案。实验结果如表 5 所示。由表 5 可知,改进的 FAL-YOLO 模型在精确率和召回率上均优于其他算法。相较于原 YOLOv8-Pose 算法,分别提升了 0.3%、3.2%,参数量减少 1.54 M。并且改进后的算法计算量较原算法降低了 4.7 G。与 YOLOv5-Pose 算法相比,召回率提升了 1.9%,参数量下降了 3.7 M。与 YOLOv10-Pose 算法相比,精确率和召回率提升了 0.5%、2%,参数量下降了 4.4 M。相比于与其余所提方法相比,精确率和召回率均有提升,模型参数量下降较大。文献[9]和文献[11]的模型大小和计算量等相比于原模型有所改进,但模型精度与 FPS 较本文方案仍有差距。文献[19]精度,P、R 等参数较原有模型提升较大,但是参数较原模型增加 16.9%,FPS 较原模型仅提升 7.3%,较改进模型仍有较大差距。而改进后的算法参数量仅为 1.47 M,模型计算量为 3.6 G,更好地兼顾了检测精度、计算量和轻量化三者之间的平衡,更加适用于移动端部署需求。此外,对照模型在训练时检测召回率较低,这是因为所使用的数

据集中包含大量背景相似目标、有缺陷对象。这也印证了传统算法对烟草茎叶角检测任务存在缺陷,而改进后的算法的准确率、召回率分别达到 97.8% 和 98.4%,对背景相似部分的检测效果有了较大提升。因此,本文提出的茎叶角检测算法具有一定的优越性。

表 5 不同算法对比试验

Table 5 Comparative experiments of different algorithms

网络结构	P/%	R/%	mAP@0.5%	Param	GFLOPs	FPS
Mobilevitv1	98.3	96.9	99.1	3.29	10.2	68.3
Mobilevitv2	98.7	95.2	99	3.35	11.5	81.3
Mobilenetv4	97.9	96.5	99.2	4.37	8.2	95.4
Ghostv1	97.9	95.5	98.8	5.43	9.6	57.8
Ghostv2	97.8	96.3	99.2	6.4	8.9	48.9
YOLOv5-pose	98.1	96.5	99.2	2.58	7.3	91.7
YOLOv8-pose	97.5	95.2	99	3.01	8.3	72.6
YOLOv10-pose	97.3	96.4	98.6	2.66	8.0	70.8
文献[9]	97.5	95	99	2.74	7.1	102
文献[17]	97.5	95.9	98.9	2.95	7.9	117.3
文献[19]	97.7	96.8	99.1	3.52	8.7	77.9
FAL	97.8	98.4	99.2	1.47	3.6	128.6

5) 消融实验结果

为了验证本文提出的各个改进模块方法的有效性,以 YOLOv8-Pose 作为基线模型,在相同实验环境下依次添加 FAI 主干网络、SAC 检测头模块、GSconv 卷积以及损失函数 MPDIoU,进行消融实验,具体实验结果如表 6 所示。其中,“√”则表示该模块加入模型当中。从表 6 中可以看出,使用 FAI 替换原 YOLOv8-Pose 主干网络模块,召回率提升了 1.8%,参数量减少近 0.07 M,计算量减少 3.5 G,模型大小减少 2.7 M。使用 SAC 检测头替换原 YOLOv8-Pose 检测头模块时,精确率提升了 1%,召回率提升了 2.1%,参数量减少近 1.29 M,计算量减少 1 G,模型大小减少 0.2 M。由此可以验证主干网络替换和检测头模块替换,可以有效降低神经网络模型的参数量和计算复杂度。当融入 GSConv 模块时,召回率提升了 2.8%,可以验证该模块再提升深层特征图像分辨率后,所获得的高频信道特征对颈部网络融合信息提供了一定帮助。当把 MPDIoU 损失函数嵌入到模型当中,在参数量和计算复杂度不变的情况下,精确率和召回率分别提升了 2.1%、3%,精度也提升了 0.5%,验证了 MPDIoU 自定义预测框大小的策略不仅有效降低了不同预测框比例的大小带来的影响,改善整体模型的检测性能。当 4 个优化模块同时融入,相较于原 YOLOv8-Pose 模型,准确率、召回率以及精度分别提升 0.3%、3.2% 和 0.2%,参数量下降 1.54 M,计算量下降了 4.7 G,模型体积仅为 3.3 M,综合性能得到全面提升。改进前后检测效果如图 13 所示。

表 6 消融实验结果
Table 6 Results of ablation experiment

模型	FAI	SAC	GSCconv	MPDIoU	P/%	R/%	mAP@0.5%	GFLOPs/G	Params/M	Size/M
YOLOv8n-Pose					97.5	95.2	99.0	8.3	3.01	6.4
Model-1	✓				94.9	97.0	98.7	4.8	1.72	3.7
Model-2		✓			98.5	97.3	99.0	7.3	2.94	6.2
Model-3			✓		97.4	97.8	99.0	8.2	2.99	6.3
Model-4				✓	97.5	95.4	98.6	8.3	3.01	6.4
Model-5	✓	✓			97.1	96.2	98.3	3.8	1.56	3.4
Model-6	✓	✓	✓		95.7	95.4	98.7	3.6	1.47	3.3
Model-7	✓	✓	✓	✓	97.8	98.4	99.2	3.6	1.47	3.3



图 13 FAL-YOLO 轻量化检测效果

Fig. 13 FAL-YOLO lightweight detection effect

5 结 论

本文提出的融合轻量化的烟草茎叶角检测算法(FAL-YOLOv8)解决了烟草植株种植环境较复杂,背景相似且干扰严重,茎叶角检测容易造成漏检和错检问题。提出了快速尺度交互 FAI 主干网络,可以有效屏蔽非烟草植株背景区域的复杂环境干扰,减少模型复杂度和参数量。使用自注意力卷积检测头 SAC 模块,能够提高烟草茎叶角的检测精度。同时引入 GSCconv 卷积优化 Neck 网络,在保持改进模型检测精度的同时减少模型复杂度和参数量。最后采用 MPDIoU 损失函数优化算法模型,提升模型精确率,召回率和精度。实验结果表明,FAL-YOLO 模型具有参数量更少、计算量更低、检测精度高等优点,在平均精度 mAP0.5 达到了 99.2%,对比原模型提升了 0.2%,P 和 R 分别提升了 0.3% 和 3.2%。网络模型计算量仅为 3.6 GFLOPs,能够实时、智能化地检测烟草茎叶角,与现有模型相比,不仅提高了检测精度,还降低了对平台计算和存

储能力的要求,便于在资源有限的户外和移动设备上部署。然而,该方法仍存在一些局限性,例如在面对更复杂的植物群体和不同环境条件下的茎叶角定位问题时,需要进一步调整和优化。未来研究将重点关注模型优化、移动端轻量化部署等方面,以期为植物表型检测提供更加准确、高效且实用的智能化解决方案。

参 考 文 献

- [1] 孙伊莹. 新零售视野下烟草经济价值链的整合重塑[J]. 商场现代化, 2024(17):14-16.
SUN Y Y. Integration and reshaping of the tobacco economic value chain under the new retail vision[J]. Mall Modernisation, 2024(17):14-16.
- [2] 卢鹏, 金静静, 曹培健, 等. 植物及烟草表型组学大数据研究进展[J]. 烟草科技, 2021, 54(3): 90-100,112.
LU P, JIN J J, CAO P J, et al. Research progress on big data in plant and tobacco phenomics[J]. Tobacco Science & Technology, 2021, 54(3): 90-100,112.
- [3] 马茹辉. 烟草叶夹角形成的组织结构基础与转录组学分析[D], 北京:中国农业科学院, 2023.
MA R H. The histological basis and transcriptomic analysis of tobacco leaf angle formation[D]. Beijing: Chinese Academy of Agricultural Sciences (CAAS), 2023.
- [4] 傅隆生, 宋珍珍, 李瑞, 等. 深度学习方法在农业信息中的研究进展与应用现状[J]. 中国农业大学学报, 2020, 25(2): 105-120.
FU L SH, SONG ZH ZH, LI R, et al. Research progress and application status of deep learning methods in agricultural information [J]. Journal of China Agricultural University, 2020, 25 (2): 105-120.
- [5] 翁杨, 曾睿, 吴陈铭, 等. 基于深度学习的农业植物表型研究综述[J]. 中国科学:生命科学, 2019, 49 (6): 698-716.

- WENG Y, ZENG R, WU CH M, et al. A review of deep learning-based research on agricultural plant phenotyping[J]. *Scientia Sinica Vitae*, 2019, 49(6): 698-716.
- [6] 肖恒树, 李军营, 梁虹, 等. 基于改进 YOLOv8 和无人机遥感影像的大田烟株数量检测[J]. 电子测量技术, 2024, 47(9): 163-171.
- XIAO H SH, LI J Y, LIANG H, et al. Detection of field tobacco plant count based on improved YOLOv8 and UAV remote sensing images [J]. *Electronic Measurement Technology*, 2024, 47(9): 163-171.
- [7] 高理文, 林小桦. 基于 L* a* b* 彩色空间和局域动态阈值的药用植物叶片图像分割[J]. 计算机应用与软件, 2014, 31(1): 232-235.
- GAO L W, LIN X H. Image segmentation of medicinal plant leaves based on L* a* b* color space and local dynamic threshold [J]. *Computer Applications and Software*, 2014, 31(1): 232-235.
- [8] 唐家阳. 基于深度学习的叶菜类植物点云分割方法研究与实现[D]. 咸阳: 西北农林科技大学, 2024.
- TANG J Y. Research and implementation of point cloud segmentation method for leafy vegetables based on deep learning [D]. Xianyang: Northwest A&F University, 2024.
- [9] 皮鸣寰, 董效杰, 杨艳, 等. 基于 SG-YOLOv8n 算法的茶叶萎凋过程中含水率检测[J]. 湖北民族大学学报(自然科学版), 2024, 42(4): 458-463.
- PI M H, DONG X J, YANG Y, et al. Detection of moisture content during tea withering based on SG-YOLOv8n algorithm [J]. *Journal of Hubei Minzu University(Natural Science Edition)*, 2024, 42(4): 458-463.
- [10] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [11] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]. IEEE International Conference on Computer vision, 2017: 2961-2969.
- [12] 吴雪梅, 刘红芸, 王芳, 等. 基于改进 Faster R-CNN 网络的烟叶分级[J]. 计算机与数字工程, 2023, 51(6): 1422-1427, 1432.
- WU X M, LIU H Y, WANG F, et al. Tobacco leaf grading based on improved Faster R-CNN network[J]. *Computer & Digital Engineering*, 2023, 51(6): 1422-1427, 1432.
- [13] 王晓婷, 赵展, 王阳, 等. 基于改进 Mask R-CNN 的植物表型智能检测算法[J]. 中国农机化学报, 2022, 43(8): 151-157.
- WANG X T, ZHAO ZH, WANG Y, et al. Intelligent plant phenotyping detection algorithm based on improved Mask R-CNN [J]. *Journal of Chinese Agricultural Mechanization*, 2022, 43(8): 151-157.
- [14] 袁山, 汤浩, 郭亚. 基于改进 Mask R-CNN 模型的植物叶片分割方法[J]. 农业工程学报, 2022, 38(1): 212-220.
- YUAN SH, TANG H, GUO Y. Plant leaf segmentation method based on improved Mask R-CNN model [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022, 38(1): 212-220.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision-ECCV 2016: 14th European Conference, 2016: 21-37.
- [17] 孙君峰, 张赵良, 刘云平, 等. 基于改进 YOLOv8 的车辆与行人检测算法[J]. 测控技术, 2025, 44(3): 9-17.
- SUN J F, ZHANG ZH L, LIU Y P, et al. Vehicle and pedestrian detection algorithm based on improved YOLOv8[J]. *Measurement & Control Technology*, 2025, 44(3): 9-17.
- [18] 向雷, 蒋文波. 改进 YOLOv8 的城市行车道路障碍物检测算法研究[J]. 电子测量技术, 2025, 48(1): 29-38.
- XIANG L, JIANG W B. Research on obstacle detection algorithm for urban road driving based on improved YOLOv8 [J]. *Electronic Measurement Technology*, 2025, 48(1): 29-38.
- [19] 朱周华, 侯智杰, 田成源, 等. 基于改进 YOLOv8-pose 的分心驾驶检测与识别[J]. 电子测量技术, 2024, 47(15): 135-143.
- ZHU ZH H, HOU ZH J, TIAN CH Y, et al. Distracted driving detection and recognition based on improved YOLOv8-pose[J]. *Electronic Measurement Technology*, 2024, 47(15): 135-143.
- [20] 罗智杰, 王泽宇, 岑飘, 等. 基于改进 YOLOv8pose 的校园体测运动姿势识别研究[J]. 电子测量技术, 2024, 47(19): 24-33.
- LUO ZH J, WANG Z Y, CEN P, et al. Research on campus physical test posture recognition based on improved YOLOv8-pose[J]. *Electronic Measurement Technology*, 2024, 47(19): 24-33.
- [21] REN SH Q, HE K M, 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.
- [22] ZHAO Y, LYU W, XU S, et al. DETRs beat YOLOs on real-time object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [23] YU H, WAN C, LIU M, et al. Real-time image segmentation via hybrid convolutional-transformer architecture search [J]. ArXiv preprint arXiv: 2403.10413, 2024.
- [24] SHAW P, USZKOREIT J, VASWANI A. Self-attention with relative position representations [J]. ArXiv preprint arXiv:1803.02155, 2018.
- [25] 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.
- [26] MA S, XU Y. MPDIoU: A loss for efficient and accurate bounding box regression[J]. ArXiv preprint arXiv:2307.07662, 2023.

作者简介

高坤,硕士,主要研究方向为计算机视觉、移动开发、图像识别及分类相关应用。

E-mail:gaokun1@stu.ynu.edu.cn

李军营(通信作者),博士,研究员,主要研究方向为烟草栽培。

E-mail:ljy1250@163.com

梁虹,硕士,教授,主要研究方向为深度学习和信号处理。

E-mail:liangh@ynu.edu.cn

马二登,博士,副研究员,主要研究方向为烟草栽培。

E-mail:erdengma@163.com

张宏,本科,农艺师,主要研究方向为烟草学。

E-mail:83754331@qq.com