

YOLOv8n-SSND: 改进的无人机航拍藜麦穗目标检测轻量模型

吴婷婷¹, 郭俊睿¹, 陶秋洁¹, 陈世华², 郭善利^{2*}

(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100, 中国;

2. 烟台大学生命科学学院, 山东烟台 264006, 中国)

摘要: [目的/意义] 藜麦穗是藜麦产量评估的重要指标之一。为提高藜麦穗目标检测精度和检测效率, 提出一种适用于无人机搭载的目标识别轻量模型 YOLOv8n-SSND (YOLOv8n with Switchable Atrous Convolution, Slim Neck, and Deformable Attention)。[方法] 以 YOLOv8n 和 YOLOv11n 为基准模型, 考虑藜麦穗形态大小不一及结构复杂, 在 Backbone 层加入可切换空洞卷积模块以提升检测复杂特征的能力; 引入 Slim-Neck 特征融合层, 轻量化主干网络; 添加可变形注意力机制, 使模型能够动态识别藜麦穗部的复杂特征, 同时保持较高的推理效率。[结果和讨论] YOLOv8n-SSND 模型平均精度均值达 94.3%, 相较于 YOLOv11n-SSND、YOLOv11n、YOLOv12n、YOLOv7、YOLOv5s、SSD (Single Shot MultiBox Detector)、Fast R-CNN (Fast Region-Based Convolutional Neural Network) 和 YOLOv8n 模型, 分别提升 0.7、0.9、2.1、1.4、2.0、23.1、19.6、1.8 个百分点。该模型推理速度达 166.7 帧/s, 较基准模型提高 26.7%; 计算量为 6.8 GFLOPs, 较基准模型降低 16.0%。[结论] YOLOv8n-SSND 模型在藜麦穗部识别上表现出更高的精确度、更快的推理速度以及更少的浮点运算量, 为无人机搭载针对藜麦穗部的藜麦目标检测提供了可行方法, 也为藜麦产量评估与智能农业管理提供了高效的技术方案。

关键词: 藜麦穗; 无人机; YOLOv8n; 目标检测; 可变形注意力; 作物表型

中图分类号: TP183; S519; TP391.4

文献标识码: A

文章编号: SA202508021

引用格式: 吴婷婷, 郭俊睿, 陶秋洁, 陈世华, 郭善利. YOLOv8n-SSND: 改进的无人机航拍藜麦穗目标检测轻量模型[J]. 智慧农业(中英文), 2026, 8(2): 59-71. DOI: 10.12133/j.smartag.SA202508021

WU Tingting, GUO Junrui, TAO Qiujie, CHEN Shihua, GUO Shanli. YOLOv8n-SSND: An Improved Lightweight Model for Aerial *Chenopodium quinoa* Willd. Spike Target[J]. Smart Agriculture, 2026, 8(2): 59-71. DOI: 10.12133/j.smartag.SA202508021 (in Chinese with English abstract)

0 引言

藜麦 (*Chenopodium quinoa*) 原产于南美洲安第斯山脉地区, 又称“南美黄金谷物”, 已有 5 000 多年的种植历史^[1]。因其卓越的营养价值和环境适应性, 被联合国粮食及农业组织 (Food and Agriculture Organization of the United Nations, FAO) 誉为“最适宜人类的全营养食品”, 并被美国航空航天局 (National Aeronautics and Space Administration, NASA) 列为宇航员太空食品之一。藜麦耐旱、耐盐碱、适应性强, 适合在干旱或贫瘠的土地

种植^[2-4], 一般种植于高海拔、温差大、种植面积广的区域, 这些特点也使得传统田间调查方法面临效率低、覆盖面有限等挑战。

无人机遥感技术具有快速、便捷、高效的特点^[5,6], 为作物长势评估提供了理想解决方案, 已广泛应用于多种作物的调查研究, 如彭小丹等^[7]提出了基于无人机图像的密集苗木检测和计数方法, 改进后的模型在测试集上的平均绝对误差、均方根误差和平均计数准确率分别为 14.24 株、22.22 株和 91.23%, 准确实现密集种植苗木的检测和计

收稿日期: 2025-08-21

基金项目: 山东省重点研发计划农业良种工程项目 (2023LZGC011)

作者简介: 吴婷婷, 博士, 副教授, 研究方向为作物表型检测技术及装备。E-mail: tt_wu@nwsuaf.edu.cn

*通信作者: 郭善利, 博士, 教授, 研究方向为植物遗传育种、植物基因克隆及功能研究。E-mail: gsl@ytu.edu.cn

copyright©2026 by the authors

(登录 www.smartag.net.cn 免费获取电子版全文)

数。杨福芹等^[8]利用无人机搭载的高光谱相机和数码相机获取冬小麦的冠层光谱数据,根据所得数据构建起冬小麦氮营养指数遥感监测模型。DASH等^[9]利用无人机热成像和多光谱影像评估印度恒河流域此类作物密集型地区的田间尺度作物水分状况,发现无人机估算的地表温度与实地测量高度相关(绝对差异 $\leq \pm 1$ K),且植被温度状况指数与土壤湿度呈显著正相关($r = 0.6 \sim 0.81$),证实无人机遥感可有效评估高分辨率作物水分状况。

基于无人机遥感的这些优势,越来越多研究者把无人机遥感作为主要技术手段引入藜麦长势评估中。井梅秀等^[10]通过无人机多光谱影像进行藜麦长势分析,分析归一化植被指数、绿光归一化植被指数与归一化差值红边指数3种植被指数,证实藜麦全生育期3种指数呈现先升后降的健康生长态势,实现了对藜麦全生育期长势的快速、精准监测。RUIZ等^[11]利用无人机多光谱传感器和地面光谱辐射计,研究了不同归一化植被指数的计算方法与地面实测归一化植被指数在藜麦受控实验和商业种植中的相关性,结果表明,在植被密度不足的区域,由于土壤背景干扰显著,应用归一化植被指数不可行。FLORES^[12]基于无人机获取的多光谱图像,利用机器学习方法实现对藜麦进行分类。这些研究充分证明了无人机遥感技术在藜麦长势监测中的有效性和应用潜力。

近年来,农业目标检测领域的技术发展呈现出从传统人工调查向智能化、精准化快速演进的过程,将机器视觉运用到农业目标检测领域中,采用深度学习等算法获得理想检测精度,成为了技术应用的主要流派^[13,14]。将无人机遥感与先进的深度目标检测技术相结合,实现小麦穗、玉米雄穗、水稻穗等多种作物的大规模、高通量检测与计数也成为目标检测领域的热门方向。翁海勇等^[15]基于无人机遥感和改进YOLOv7_Neckblast模型的水稻穗颈瘟抗性智能鉴定方法,通过引入压缩注意力机制和可变形卷积,在保持模型轻量化的同时将检测精度提升至89.2%,抗性评估准确率达86.67%。JIA等^[16]基于无人机图像与融入坐标注意力机制的CA-YOLO模型,在实现96%检测精度的同时,显著提升了模型对复杂场景下玉米雄穗的检测鲁棒性。为解决单一尺度图像的检测局限,QIU等^[17]提出了一种结合多源图像与基于注意力的内部尺度特征交互(Attention-based Intrascala Feature Interaction, AIFI)等机制的改进YOLOv5模型,有效提升

了水稻穗检测的精度与鲁棒性。张晓勐^[18]基于无人机遥感技术构建玉米雄穗数据集,提出Swin T-YOLO检测模型,实现95.11%的检测精度,并在Jetson Nano边缘设备上实现高速检测。高姻燕等^[19]利用无人机红、绿、蓝三原色(Red, Green, Blue, RGB)图像与YOLOv3算法,开发出高效的小麦穗数检测方法,实现超过90%的识别准确率。

尽管利用无人机对主粮作物(如水稻、小麦)穗部进行高效检测的技术已较为成熟,但针对藜麦穗的精确目标检测研究仍属空白。由于藜麦穗多为分支状花序,形态结构复杂,且在成熟期易出现“垂头”现象,与主粮作物普遍具有的直立或棒状穗型差异显著,导致现有针对主粮作物的穗部检测方法难以直接迁移应用。此外,在无人机航拍场景下,藜麦穗的识别还面临两大挑战:一方面,藜麦种植密度高、穗部目标分布密集、彼此遮挡严重,同时,复杂的花序结构对特征提取的精细化程度提出了更高要求;另一方面,无人机机载平台存在存储容量有限、计算资源受约束等问题,对检测模型的高效性与轻量化提出了严峻考验。因此,针对上述难点,并着眼于目前尚无基于深度学习的藜麦穗目标检测研究的现状,本研究提出一种基于改进YOLOv8n的藜麦穗部检测方法,旨在构建一种精度高、速度快、参数量小、适于无人机平台部署的专用检测模型,以期为实现航拍场景下藜麦生长状态的精准监测提供技术参考,并为后续产量评估提供可靠依据。

1 材料与方法

1.1 藜麦穗数据集构建

1.1.1 数据采集

本研究构建的藜麦穗数据集,采自山东省东营市广饶县农高区藜麦种植基地(平均海拔13 m, 118°39'12.20"E, 37°17'56.34"N)。该区域地处暖温带,年平均气温12.3℃,年平均降水量587.4 mm,年平均无霜期为198 d,年日照极值2 881.4 h。选取DY160、DY154、DY1533、DY1534、DY1537、DY1540、DY1541、DY1381、DY1382、DY1383、DY1384和DY1385共12个藜麦育种品种作为试验材料,上述品种在试验田中采用均匀种植模式。采用大疆Inspire 2无人机作为遥感平台,搭载MicaSense Altum 6波段多光谱相机,获取高精度多光谱遥感数据。多光谱相机包括5个光谱通道(蓝、绿、红、红外、近红外,中心波长分别为475、

560、668、717、840 nm, 带宽分别为20、20、10、10、40 nm) 和1个长波红外(Long Wave Infra Red, LWIR) 通道(中心波长为11 μm , 带宽为6 μm), 多光谱相机各通道中心波长与半波宽如表1所示, 在本研究的模型训练中仅使用了RGB通道数据。

表1 多光谱相机的通道参数

通道	中心波长/nm	半波宽/nm	反射率校正系数/%
Blue	475	20	48.871
Green	560	20	49.028
Red	668	10	49.028
Red-edge	717	10	48.969
NIR	840	40	48.758

无人机设定飞行高度为25 m, 飞行速度2 m/s, 航向和旁向重叠率为80%, 多光谱相机间隔1 s垂直拍摄, 每幅图像的地面覆盖范围约为18.28 m $\times$ 14.19 m。图像采集时间为2023年10月26—28日、31日和11月3日、7日、14日, 于当地时间11:00—14:00进行数据采集。由于每幅图像的地面覆盖范围较大, 包含藜麦株数过多, 因此将原始图像均匀分割为244 $\times$ 244像素大小的图像, 共得1 831幅图像。

1.1.2 数据增强以及标注

为确保数据质量, 从采集到的1 831幅图像中筛选, 剔除因过度曝光或欠曝光、运动模糊和离焦等问题导致的低质量图像, 并剔除目视解译藜麦穗被遮挡率大于40%或边界截断的图像, 针对遮挡率小于等于40%的目标, 采用基于可见部分的矩形标注方案, 标注框严格依据目标可见轮廓绘制, 仅包围实际可见像素区域, 不推测被遮挡部分, 并在保证矩形框完整的前提下使其紧贴目标可见部分的外接矩形, 最大限度减少背景干扰, 最终筛选出113幅的高质量藜麦穗图像数据集, 涵盖不同光照条件、地块位置及植株密度场景。

为了提高模型泛化能力和鲁棒性, 对上述数据集系统实施了亮度调整、翻转与锐化3类数据增强操作, 通过 $\pm 50\%$ 的亮度调整模拟不同光照环境; 采用0.5~1.0强度的拉普拉斯锐化增强图像边缘与细节, 同时保持自然清晰度; 以及通过100%水平翻转平衡样本方向性偏差。经增强后, 数据集规模由113幅扩展至452幅, 为模型训练提供了更丰富、更多样化的样本基础。数据集如图1所示。采用Labelimg软件对藜麦穗数据集中的藜麦麦穗进行标

注。将标记后数据集按7:3比例随机划分为训练集和测试集。

1.2 YOLOv8n基准模型

YOLOv8n是Ultralytics推出的YOLOv8系列中轻量化的模型(Nano版本), 专为实时目标检测设计, 适用于计算资源受限的场景(如边缘设备和移动端), 在保持帧率(Frame Per Second, FPS)较高的同时, 精度优于其他轻量级模型。其核心架构延续了YOLOv8的模块化设计, 由主干网络Backbone层、特征融合Neck层和检测头Head层三大部分组成, 具体结构如图2所示。

YOLOv8n的Backbone层由5个卷积模块、4个C2f模块和1个快速空间金字塔池化(Spatial Pyramid Pooling-Fast, SPPF)模块构成, 卷积模块用于降采样和非线性化表示、C2f模块用于特征提取、SPPF模块采用空间金字塔的方法将不同尺寸的图像传入下1层, 其中C2f可以针对藜麦穗特点进行替换, 以增强其特征提取能力; Neck层的核心作用在于特征融合, 它依托路径聚合网络-特征金字塔网络(Path Aggregation Network with Feature Pyramid Network, PAN-FPN)结构, 对Backbone所输出的不同尺度特征图进行上采样与卷积操作, 以完成融合, 但这种方法的Neck层计算量较大且存在一定的计算冗余; Head层将Neck层输出的特征图进行处理最终产出模型的输出结果, 在Head层前加入注意力机制可以使结果更加注意藜麦穗的关键区域, 可以提升模型对藜麦穗的检测能力。

1.3 轻量化航拍藜麦穗目标检测模型设计

本研究基于YOLOv8n基础架构提出了一种改进的轻量化藜麦穗目标检测模型: YOLOv8n-SSND(YOLOv8n with Switchable Atrous Convolution, Slim Neck, and Deformable Attention), 改进后网络结构如图3所示。首先, 将Backbone部分的C2f模块替换成可切换空洞卷积(Switchable Atrous Convolution, SAC)模块, 对相同输入特征应用不同的空洞率进行卷积, 提高网络处理藜麦穗部不同尺度和复杂度特征时的适应性和准确性。其次, 引入Slim-Neck替换特征融合层, 采用组混洗卷积(Group Shuffle Convolution, GSConv)、组混洗卷积瓶颈层(Group Shuffle Convolution Bottleneck, GS Bottleneck)、基于一次聚合的组混洗卷积跨阶段部分网络(One-Shot Aggregation-based Group Shuffle Convolution Cross Stage Partial Network, VoV-GSC-

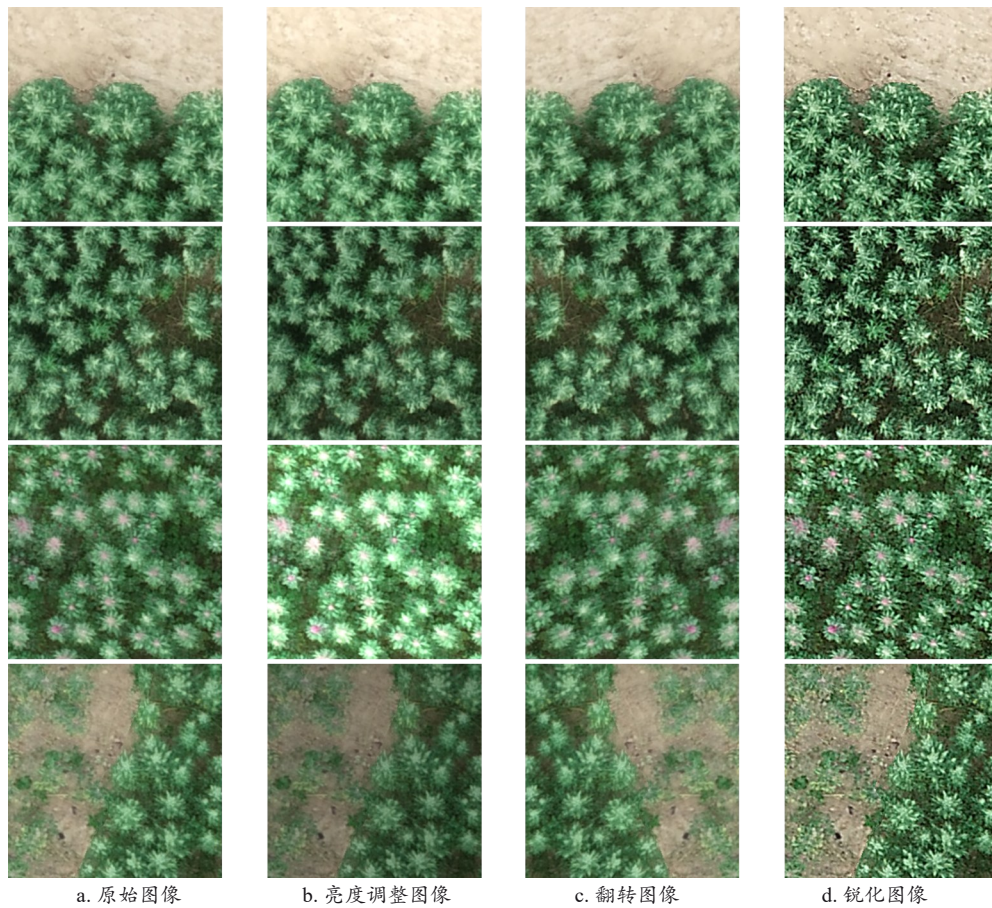


图1 经过增强后的藜麦穗数据集

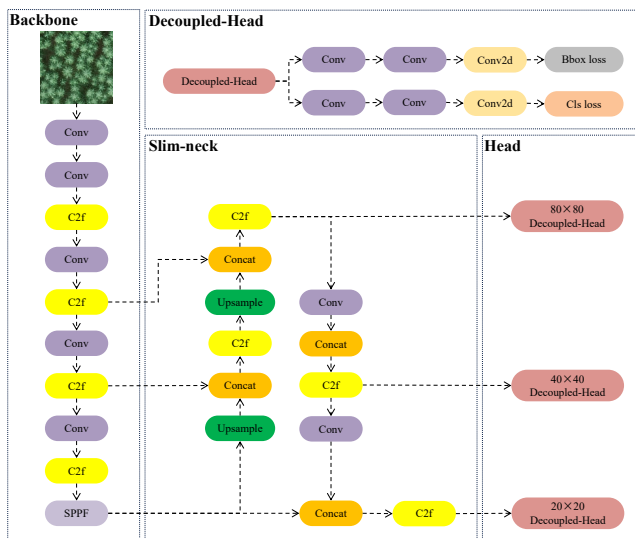
Fig. 1 The enhanced dataset of *Chenopodium quinoa* Willd spikelets

图2 YOLOv8n的网络结构

Fig. 2 Network structure of YOLOv8n

SP) 模块构建 Slim-Neck 层, 降低计算量和参数量的同时保证检测藜麦穗目标准确性。此外, 在检测头前添加可变形注意力 (Deformable Attention, DA) 机制, 使模型在做出最终预测之前, 更加集中注意

力于藜麦穗关键的特征, 从而提升检测精度和鲁棒性。

1.3.1 SAC可切换空洞卷积

在藜麦穗航拍图像中, 由于不同藜麦品种的生长状况差异, 藜麦穗部在图像中具有不同的尺度大小和不同的穗部形状, 导致检测精度下降。因此本研究将 YOLOv8n 基准模型中的 C2F 模块替换为可切换空洞卷积 (Switchable Atrous Convolution, SAC) 模块, 提高网络在处理不同穗部大小和形状的特征时的精确度。SAC 模块是一种动态可切换的空洞卷积机制, 其输入输出特征图维度相同, 通过自适应选择不同空洞率的卷积核, 并通过特别设计的开关函数来融合这些不同卷积的结果, 使得网络能够更灵活地适应不同尺寸的藜麦穗部的特征, 从而更准确地识别和分割图像中的藜麦穗^[20-22]。SAC 模块的实现如图 4 所示。

传统卷积模块可以表示为公式 (1)。

$$y = \text{Conv}(x, w, r_{\text{C2F}}) \quad (1)$$

式中: y 为输出的融合特征图; x 为输入特征

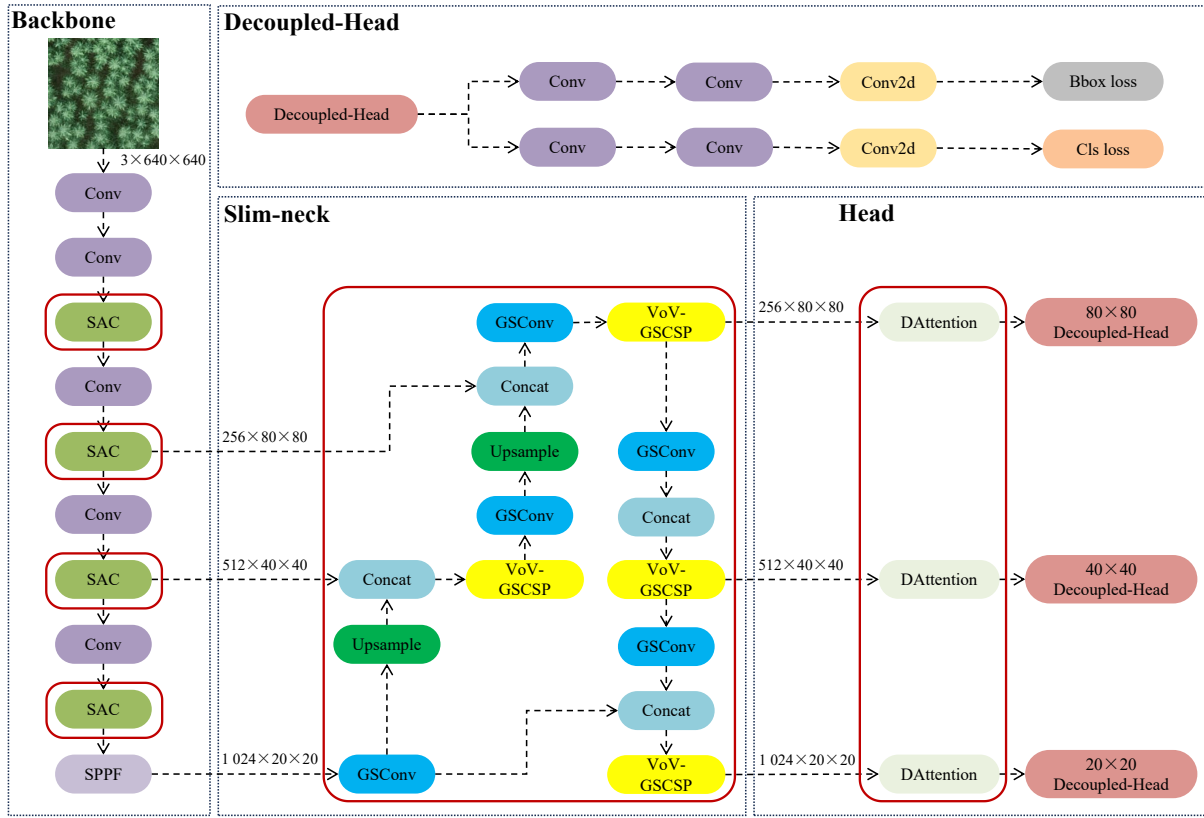


图3 改进的YOLOv8n网络结构图

Fig. 3 Improved structure of YOLOv8n network

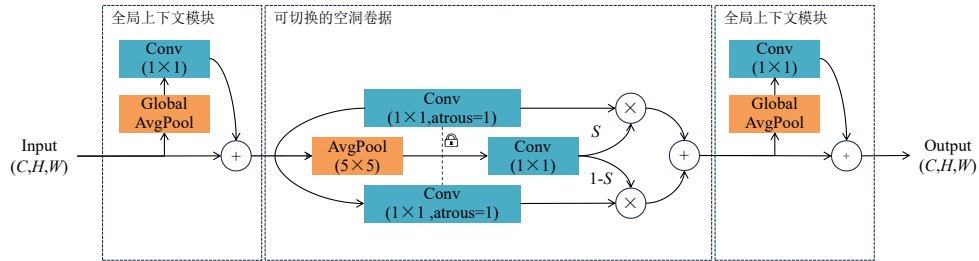


图4 SAC的具体实现方式

Fig. 4 Specific implementation of SAC

图; w 为卷积核权重; r_{c2f} 为空洞率。SAC模块可以用公式(2)表示。

$$y = S(x) \times \text{Conv}(x, w, 1) + (1 - S(x)) \times \text{Conv}(x, w + \Delta w, r_{\text{SAC}}) \quad (2)$$

式中: r_{SAC} 为SAC模块的超参数; Δw 为可训练参数的偏移量; $S(x)$ 为开关函数。

1.3.2 引入Slim-Neck特征融合层

为了提高推理速度并减小计算量和模型大小,本研究将YOLOv8n原特征融合层替换为模块化设计的Slim-Neck轻量化特征融合层。YOLOv8n在网络中使用大量的标准卷积提取图像特征,会消耗大量的时间^[23]。而藜麦穗目标检测可能受光照、种植

密度、叶片间的相互遮挡等复杂环境的影响,对模型的检测速度和检测精度都有很高的要求,因此将一种新范式设计Slim-Neck融入YOLOv8n网络中。Slim-Neck模块是利用GSConv模块和VoV-GSCSP模块设计的轻量化结构,在保持准确性的同时减少模型的复杂度。

GSConv模块用于减少藜麦穗目标检测模型的运行时间,其结构如图5所示,输入图像经过特征提取层得到初步特征图,这些特征图进入GSConv模块,先执行分组卷积,再进行深度可分离卷积,接着利用shuffle操作混合不同通道的信息,增强跨通道交互效果,经过这些变换后的特征会被用于后

续的目标检测或其他任务中。其中, GS bottleneck模块是基于传统卷积模块和 GSConv 模块的一种增强模块, 大幅降低计算成本的同时, 保持甚至提升特征表达能力。

VoV-GSCSP 模块采用一次性聚合机制设计, 作为跨阶段部分网络 (Global Spatial Context Concatenation Pyramid, GSCSP) 的核心组件, 能够充分提取浅层网络与深层网络的特征信息, 从而使输

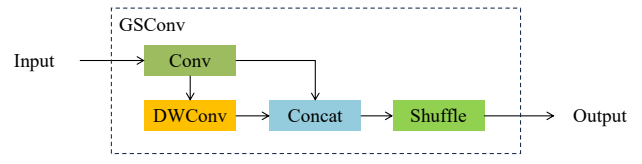


图5 GSConv 模块的结构

Fig. 5 Structure of the GSConv module

出的藜麦穗特征图具备丰富的位置信息和语义信息^[24], 其结构如图6所示。

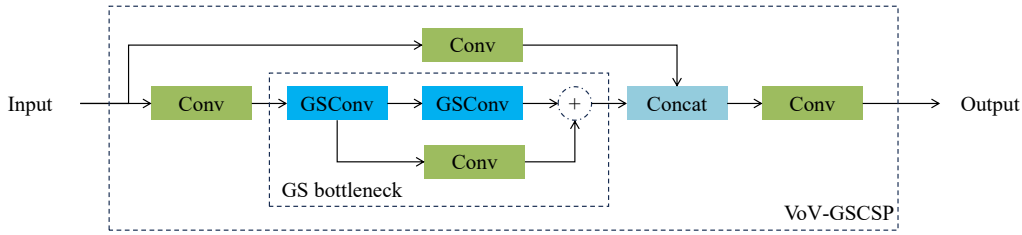


图6 VoV-GSCSP 模块的结构

Fig. 6 Structure of the VoV-GSCSP module

1.3.3 添加 DA 机制

卷积神经网络采用固定的卷积核来进行采样, 因此在对形状不固定的特征图进行建模时存在一定缺陷。现有的视觉 Transformer 在建模长距离依赖时, 存在计算成本高、信息冗余等问题。同时, 为了进一步增强对藜麦不同穗部形状的处理能力, 本研究在模型的探测头前引入 DA 机制, 通过动态采样, 关注藜麦穗部的关键区域, 提高检测的精度。

DA 是一种结合可变形卷积 (Deformable Convolution) 和自注意力机制 (Self-attention) 的混合模块, 其核心思想是通过动态学习特征图中每个位置的偏移量, 自适应调整感受野, 从而更精准地聚焦目标的关键区域。相较于传统注意力机制, DA 的优势在于: 通过偏移量学习, 使模型能适应目标的形变、遮挡等复杂场景; 数据依赖的注意力机制能够动态选择采样点, 自适应地关注重要区域; 通过共享偏移量, 减少了计算开销, 同时保持了线性空间复杂度。

将特征图 $x \in \mathbb{R}^{H \times W \times C}$ 作为输入, 然后生成 1 个均匀的参考点网格 $p \in \mathbb{R}^{H_G \times W_G \times 2}$, 其中, $H_G = \frac{H}{r_{DA}}$, $W_G = \frac{W}{r_{DA}}$, H 为特征图的高, W 为特征图的宽, r_{DA} 为网格大小从输入特征图大小向下采样的因子。为了获得每个参考点的偏移量, 对特征图进行线性投影以获得查询标记 q , 然后将其输入 1 个轻量级子网络偏移量 $\theta_{offset}(q)$ 网络, 以生成参考点 p 的偏移量 Δp 。具体来说, $\theta_{offset}(q)$ 子网络包含两个

非线性激活的卷积模块。首先, 输入特征经过深度卷积以捕捉局部特征。然后, 采用高斯线性误差单元 (Gaussian Error Linear Unit, GELU) 激活和 1×1 卷积来计算偏移量。在生成参考点的偏移量后, 使用采样函数, 如公式 (3) 所示。

$$\phi(z; (p_x, p_y)) = \sum_{(r_x, r_y)} g(p_x, r_x) g(p_y, r_y) z[r_y, r_x, :] \quad (3)$$

式中: $g(a, b) = \max(0, 1 - |a - b|)$; (r_x, r_y) 为 $z \in \mathbb{R}^{H \times W \times C}$ 上的所有位置; p_x 和 p_y 为对应参考点的偏移量。

用双线性插值法对变形点位置的特征进行采样, 然后投影为键值 $\tilde{k}$ 和值 $\tilde{v}$ 。接下来, 对 $q^{(m)}$, $\tilde{k}^{(m)}$, $\tilde{v}^{(m)}$ 进行标准的多头关注, 其中有 M 个头的多头注意力模块如公式 (4) ~ 公式 (6) 所示。

$$q = xW_q, \tilde{k} = \tilde{x}W_k, \tilde{v} = \tilde{x}W_v \quad (4)$$

$$z^{(m)} = \sigma\left(q^{(m)}\tilde{k}^{(m)\top}/\sqrt{d}\right)\tilde{v}^{(m)}, m = 1, \dots, M \quad (5)$$

$$z = \text{Concat}(z^{(1)}, \dots, z^{(M)})W_o \quad (6)$$

式中: $W_q, W_k, W_v, W_o \in \mathbb{R}^{C \times C}$ 为投影矩阵; q 为查询标记; $\tilde{x}$ 为采样特征; $\sigma(\cdot)$ 为 softmax 函数; $d = \frac{C}{M}$ 为每个头部的维度; $z^{(m)}$ 为第 m 个注意力头的输出; $q^{(m)}, \tilde{k}^{(m)}, \tilde{v}^{(m)}$ 为查询、变形键和值嵌入。

最后, 将每个头部的特征串联起来, 并通过 W_o 进行投影, 得到最终输出 z , 如公式 (6) 所示,

整体实现如图7所示。本研究在每层输出前引入DA，通过动态感受野调整 and 自适应特征聚焦，使模型在做出最终预测之前，更加集中注意力于最关

键的特征^[25]，可以显著提升模型对复杂场景的检测能力。

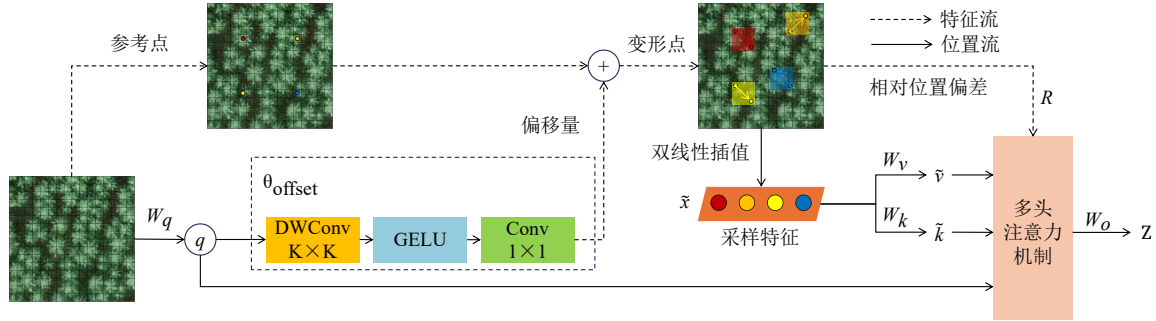


图7 DA实现原理图

Fig. 7 Schematic diagram of DA implementation

1.4 评价指标

为了准确评估模型的有效性，本研究采用精确率 (Precision, P)、召回率 (Recall, R)、计算量 GFLOPs、平均精度均值 (Mean Average Precision, mAP)、FPS、模型体积作为模型性能的评价指标。其中，在评价 mAP 时，检测框与真实框的重叠度 (Intersection over Union, IoU) 阈值为 0.5^[26]。精确率、召回率、平均精度和平均精度均值的计算如公式 (7) ~ 公式 (10) 所示。

$$P = \frac{T_p}{T_p + F_p} \quad (7)$$

$$R = \frac{T_p}{T_p + F_N} \quad (8)$$

$$AP = \int_0^1 P_{\text{interp}}(R) dr \quad (9)$$

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

式中： T_p 为被正确检测到的藜麦穗的数量； F_p 为背景被错误识别成藜麦穗的数量； F_N 为漏检的藜麦穗的数量； $P_{\text{interp}}(R)$ 为召回率 R 处经过插值后的精确率； N 为目标的总数； AP_i 为第 i 个类别的平均精度。

目标检测算法的性能主要涵盖分类能力和定位能力两个方面，最直观的度量方式是 mAP 值。mAP 值越小代表算法的准确性较低，反之表示算法的精确性更高。另外，检测速度通常用模型帧率来衡量，更高的帧率值意味着模型能够处理更多数据，提供更强的实时性^[27]。

1.5 实验环境与参数设置

本研究实验在同一台笔记本电脑上，搭载 Win-

dows11 操作系统，CPU 配置为 13th Gen Intel (R) Core (TM) i7-13620H。GPU 配置为 NVIDIA GeForce RTX 4060 Laptop 显卡，显存为 8 G。具体的软件环境为 Python 3.9，CUDA 版本为 11.8，Pytorch 版本为 2.5.1。改进 YOLOv8n 模型训练的超参数设置如下，初始学习率设置为 0.01，输入图像像素大小为 640×640 ，Batch Size 为 16，Epochs 为 100，优化器设置为 Auto，开启自动混合精度进行训练，未采用余弦学习率调度器，其他参数均为官方推荐参数。

2 结果与分析

2.1 检测结果热力图可视化

为进一步了解 YOLOv8n-SSND 对藜麦穗部的检测效果，将改进模型的检测结果与原基准模型的检测结果通过热力图呈现，如图 8 所示。通过对比，改进模型比原基准模型显著减少了误检和漏检的情况，同时对于图像边缘处的不完整藜麦穗部的检测能力显著提高。这些数据充分证明，本研究的改进方案在检测效果、精度、速度和模型复杂度多个维度实现了最佳平衡。

2.2 DA 与其他注意力机制的对比

为系统验证本研究所采用的 DA 模块性能，本研究选取了当前具有代表性的两种注意力机制——混合注意力 (Hattention, HA) 与可变形大核注意力 (Deformable Large-Kernel Attention, DLKA) 作为对比基准。通过在相同实验环境下对 YOLOv8n 模型进行模块替换与性能测试，从模型大小、计算效

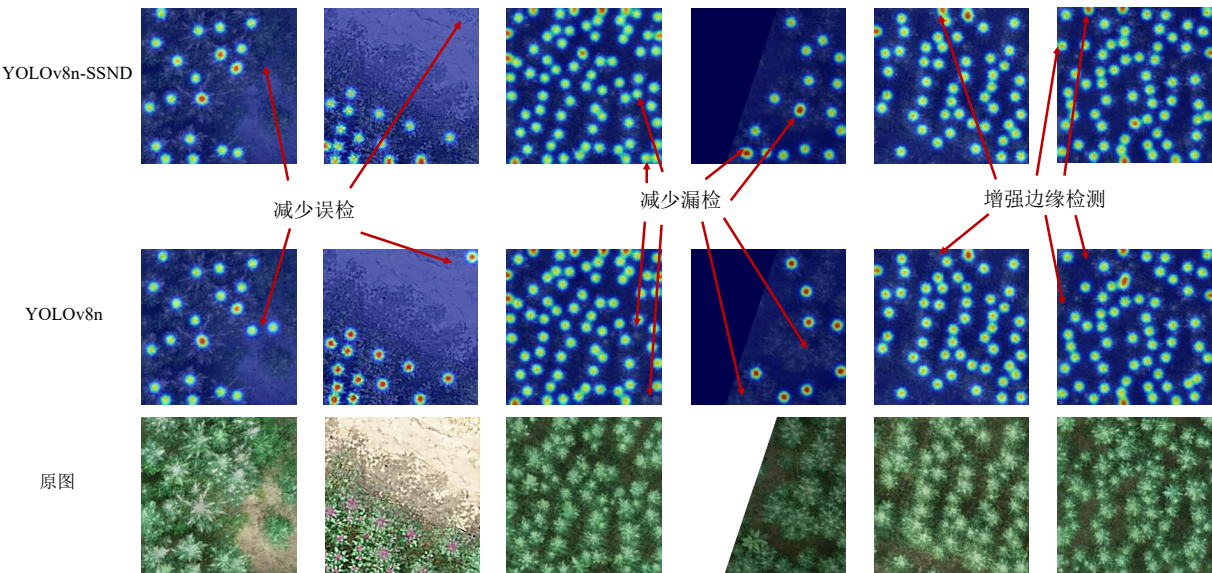


图8 YOLOv8n-SSND与YOLOv8n对航拍藜麦穗检测结果对比图

Fig. 8 Comparison chart of detection results of aerial images of *Chenopodium quinoa* Willd grains by YOLOv8n-SSND and YOLOv8n

率、检测精度等多个维度进行综合评估，以确保比 所示。
较的公平性与结论的可靠性，实验结果如表 2

表2 DA与其他注意力机制对航拍藜麦穗目标检测模型中的对比

Table 2 Comparison of DA and other attention mechanisms in the aerial photography-based *Chenopodium quinoa* willd spike target detection model

模型	模型大小/MB	P/%	R/%	mAP50/%	计算量/GFLOPs	Params/M	FPS/(帧/s)
YOLOv8n-DA	6.8	89.1	86.7	93.6	8.3	3.3	87.7
YOLOv8n-HA	16.9	89.8	85.7	93.4	8.1	8.2	294.1
YOLOv8n-DLKA	9.8	89.0	85.9	93.2	9.5	4.8	185.2

HA采用并行混合注意力架构，通过构建通道与空间双路径实现特征增强。其通道路径通过全局池化与全连接层学习通道间依赖关系，空间路径则通过卷积操作生成空间注意力图，最终通过加权融合机制整合双路径输出。这种设计在保持较高精度的同时显著提升了计算效率，如表2中YOLOv8n-HA实现294.1帧/s的优异表现。

DLKA创新性地将可变形卷积与大核卷积相结合，通过7×7及以上尺寸的大卷积核构建基础感受野，同时引入可变形机制使采样点能够根据特征内容自适应偏移。这种设计既保留了大规模感受野对长距离依赖的捕获能力，又增强了对不规则目标几何变换的适应性，在精度与速度间取得了良好平衡。

相比之下，融合DA模块的YOLOv8n模型展现出更为均衡的综合性能优势。在检测精度方面，该模型取得了93.6%的mAP50，较HA与D-LKA分

别提升0.2和0.4个百分点；在模型复杂度方面，其6.8 MB的模型体积显著小于HA的16.9 MB和DLKA的9.8 MB；同时在计算效率方面，DA以更低的参数量和更优的FPS表现突出。实验结果表明，DA在精度保持、模型轻量化与计算效率3个关键维度上实现了最佳平衡，尤其适合部署于计算资源严格受限的无人机机载平台。

2.3 消融实验

在进行消融实验之前，首先对YOLOv8n、YOLOv11n、YOLOv12n这3个目标检测模型的基准模型进行了对比实验，结果如表3所示。

实验结果表明，YOLOv8n在精确率（90.4%）上表现更为出色，这表明其对于正样本的判断更为谨慎，误检率较低。而YOLOv11n则在综合性能上展现出更大优势，其平均精度、召回率、模型大小、计算复杂度，以及参数量均优于YOLOv8n。

表3 航拍藜麦穗目标检测基准模型对比实验

Table 3 Aerial photography-based comparison experiment of benchmark models for *Chenopodium quinoa* Willd grain target detection

模型	模型大小/MB	P/%	R/%	mAP50/%	计算量/GFLOPs	Params/M	FPS/(帧/s)
YOLOv8n	6.3	90.4	84.0	92.5	8.1	3.0	131.6
YOLOv11n	5.5	89.6	85.0	93.4	6.3	2.6	59.5
YOLOv12n	5.4	87.9	83.9	92.2	5.8	2.5	243.9

鉴于目标检测任务中 mAP 是衡量模型综合性能的核心指标,更高的 mAP50 意味着模型在各种 IoU 阈值下拥有更稳定和准确的平均检测能力。因此,本研究选择 mAP50 值更高的 YOLOv11n 和 YOLOv8n 一同进行后续深入的消融实验。

本研究提出的 3 种改进方法分别是 SAC、DA、SN,为了验证它们对 YOLOv8n 模型的性能优化有效,对不同的改进部分进行消融实验,实验的结果如表 4 所示。

表4 基于 YOLOv8n 的航拍藜麦穗目标检测模型消融实验

Table 4 Ablation experiments on the aerial *Chenopodium quinoa* Willd spike target detection model based on by YOLOv8n

YOLOv8n	SAC	Slim-Neck	DA	模型大小/MB	P/%	R/%	mAP50/%	计算量/GFLOPs	Params/M	FPS/(帧/s)
√	×	×	×	6.3	90.4	84.0	92.5	8.1	3.0	131.6
√	√	×	×	6.9	89.0	87.6	93.8	7.4	3.3	108.7
√	×	√	×	5.9	87.9	84.9	92.7	7.3	2.8	204.1
√	×	×	√	6.8	89.1	86.7	93.6	8.3	3.3	87.7
√	√	√	×	6.5	89.6	85.7	93.4	6.6	3.4	208.3
√	√	×	√	7.5	89.6	86.9	93.9	7.6	3.6	270.3
√	×	√	√	6.5	88.3	86.8	93.2	7.5	3.1	227.3
√	√	√	√	7.1	90.8	86.2	94.3	6.8	3.4	166.7

注:√表示采用此模块;×表示未采用此模块。

可以看出,单独引入 SAC 模块使 mAP50 提升 1.3 个百分点, FPS 提升 17.4%; DA 模块在保持精度的同时召回率显著提高了 2.7 个百分点;而 Slim-Neck 结构则展现出显著的轻量化效果,将模型体积压缩了 6.3%,并将参数量降低了 0.2 M,实现了 204.08 帧/s 的极速推理。3 个模块联合的最终模型将 mAP50 提升 1.8 个百分点到达 94.3% 的同时,计算量降低 16%,达到 6.8 GFLOPs, FPS 提高 26%,达到 166.7 帧/s,验证了各改进组件的互补性和协同方案的技术优势。该优势源于 3 个模块在功能上的高效互补, Slim-Neck 作为轻量化基石,通过优化特征融合结构显著压缩了模型复杂度,为引入更强力但稍显复杂的模块腾出了计算预算, SAC 模块则在此基础上增强了网络在处理不同穗部大小和形状特征时的精确度,而 DA 模块通过对藜麦穗部关键区域进行聚焦。最终,这 3 个模块共同构建了一个“轻量骨架-精准感知-区域聚焦”的协同体系,在藜麦穗部检测任务中实现了精度、轻量与速度的均衡提升。

将 3 种改进方法应用于 YOLOv11n 目标检测模

型,进一步验证改进方法的性能优化有效,对不同的改进部分进行消融实验,实验的结果如表 5 所示。

对 YOLOv11n 的消融实验结果进行分析可得,本研究提出的改进方法在该新型架构上依然有效,但其收益特性与在 YOLOv8n 上有所不同。单独引入 SAC 模块使 mAP50 提升了 0.4 个百分点(从 93.4% 至 93.8%),这验证了其增强模型感受野、捕获多尺度信息的能力在 YOLOv11n 架构上依然具有稳健的有效性。然而, Slim-Neck 和 DA 模块未能带来预期的正向提高,其根本原因在于 YOLOv11n 本身相较于 YOLOv8n 进行了深刻的架构优化。具体而言, YOLOv11n 引入了全新的 C3k2 模块和嵌入金字塔挤压注意力的跨阶段部分网络(CCross Stage Partial Network with Pyramid Squeeze Attention, C2PSA)模块,替代了 YOLOv8n 中的 C2f 模块,同时 YOLOv11n 已经在 SPPF 层后集成了其自研的 C2PSA 注意力模块,该模块已有效地增强了模型对复杂场景的适应能力。这些改进本身就极大地提升了特征提取效率,显著减少了模型参数和计

表5 基于YOLOv11n的航拍藜麦穗目标检测模型消融实验

Table 5 Ablation experiments on the aerial *Chenopodium quinoa* Willd spike target detection model based on by YOLOv11n

YOLOv11	SAC	Slim-Neck	DA	模型大小/MB	P/%	R/%	mAP50/%	计算量/GFLOPs	Params/M	FPS/(帧/s)
√	×	×	×	5.5	89.6	85.0	93.4	6.3	2.6	59.5
√	√	×	×	7.8	90.4	85.8	93.8	6.5	3.7	166.7
√	×	√	×	5.9	88.9	84.7	92.7	6.4	2.8	217.4
√	×	×	√	6.1	89.7	83.3	92.2	6.5	2.9	222.2
√	√	√	×	8.9	89.0	86.1	93.2	7.8	4.2	125.0
√	√	×	√	8.4	89.7	85.1	93.6	6.7	4.0	131.6
√	×	√	√	6.5	88.4	85.7	93.2	6.6	3.1	172.4
√	√	√	√	9.4	89.2	85.9	93.6	8.1	4.5	137.0

注：√表示采用此模块；×表示未采用此模块。

算量。这意味着，旨在压缩 Neck 部分复杂度的 Slim-Neck 模块与 YOLOv11n 本身高度精简和优化的新结构存在功能冗余，其轻量化效果的边际收益因此大幅降低。同样地，DA 机制模块旨在增强模型对关键特征的关注能力，但也因此，再额外引入 DA 模块不仅造成了功能上的重叠，还可能在一定程度上破坏了原架构中精心设计的特征流，导致计算成本增加却未能带来显著的性能增益，甚至引起轻微的精度下降。最终，3 模块联合的模型在 YOLOv11n 上达到了 93.6% 的 mAP50，仍比基线提升了 0.2 个百分点，这表明本研究的改进组合仍具有一定的互补性和有效性。

综合表 4 与表 5 可知，本研究提出的方法在 YOLOv8n 上能实现精度与速度的显著同步提升，而在最新的 YOLOv11n 架构上，其价值则更多体现

在验证核心模块（如 SAC）的泛化能力，以及探索不同架构下的性能权衡上。系列对比实验充分说明，本研究提出的改进方案在主流 YOLOv8 模型上具有极高的应用价值，同时其核心思想对新一代目标检测模型也具备重要的参考意义。

2.4 对比实验

为了充分验证改进 YOLOv8n 模型对航拍图像中的藜麦穗的检测能力，将改进后的 YOLOv8n-SSND 模型与主流检测模型 YOLOv11n、YOLOv12n、YOLOv7、YOLOv5s、SSD（Single Shot MultiBox Detector）、Fast R-CNN（Fast Region-based Convolutional Neural Network），以及 YOLOv11n-SSND 进行对比实验，实验结果如表 6 所示。

表6 不同模型检测的航拍藜麦穗目标性能结果

Table 6 Comparison performance results of different models for target detection of *Chenopodium quinoa* Willd grains

模型	模型大小/MB	P/%	R/%	mAP50/%	计算量/GFLOPs	Params/M	FPS/(帧/s)
YOLOv8n-SSND	7.1	90.8	86.2	94.3	6.8	3.4	166.7
YOLOv11n-SSND	9.4	89.2	85.9	93.6	8.1	4.5	136.7
YOLOv11n	5.5	89.6	85.0	93.4	6.3	2.6	250.0
YOLOv12n	5.4	87.9	83.9	92.2	5.8	2.5	243.9
YOLOv7	71.3	88.0	86.4	92.9	103.2	36.9	53.2
YOLOv5s	3.9	89.4	86.2	92.3	4.1	7.0	119.1
SSD	90.6	79.0	62.6	71.2	30.4	23.3	135.3
Fast R-CNN	108.0	79.3	74.9	74.7	369.7	41.3	13.7

实验数据表明，改进模型 YOLOv8n-SSND 在 mAP50 指标上达到 94.3%，相较于 YOLOv11n-SSND、YOLOv11n、YOLOv12n、YOLOv7、YOLOv5s、SSD、Fast R-CNN 和 YOLOv8n 模型，分

别提升 0.7、0.9、2.1、1.4、2.0、23.1、19.6、1.8 个百分点。在保持高精度的同时，改进模型的 FPS 达到 166.67 帧/s，是 YOLOv7 的 3.1 倍；且模型体积仅 7.1 MB，远小于 YOLOv7 的 71.3 MB；模型参数

量仅为 3.4 M, 远低于 YOLOv7 的 36.9 M。与轻量级 YOLOv5s 相比, 改进模型在 mAP50 上提升 2.0 个百分点的同时, 参数量减少约 52%, 计算效率仍保持竞争优势。

相较于最新的 YOLOv11n 和 YOLOv12n 模型, 虽然 YOLOv11n 模型的 mAP50、模型体积、计算量, 以及 FPS 略微优于 YOLOv8n 模型, 但以 YOLOv11n 为基准模型加入 SAC、DA、Slim-Neck3 种改进方法形成的 YOLOv11n-SSND 模型, 除了模型体积略小于 YOLOv8n-SSND 模型, 精确率、召回率、mAP50、计算量、参数量和 FPS 等指标上都不及 YOLOv8n-SSND 模型。YOLOv12n 的精确度和 mAP50 均低于 YOLOv8n 模型。总体而言, YOLOv8n-SSND 相较于其他主流目标检测模型具有明显优势, 充分验证其对航拍图像中的藜麦穗部的检测能力。

3 结 论

本研究针对无人机航拍场景中, 藜麦穗部性状目标密集、特征复杂, 以及无人机平台机载算力受限等问题, 提出一种基于 YOLOv8n 模型改进后的 YOLOv8n-SSND 目标检测模型, 实现了在自然环境中对藜麦穗目标的准确检测。相对于原模型, YOLOv8n-SSND 主要做了 3 点改进: 1) 在 Backbone 中构建 SAC 模块, 通过动态调整卷积核的空洞率, 使模型能够自适应地捕捉不同尺度的特征信息。2) 引入 Slim-Neck 特征融合层, 采用 GSConv 替代标准卷积, 通过分组可分离卷积和通道重排技术, 在保持特征表达能力的同时, 降低 Neck 部分的计算量。特别设计的 VoV-GSCSP 模块实现了跨阶段特征的高效融合, 同时大幅度压缩了模型体积。3) 在检测头前添加 DA 机制, 通过预测特征点的动态偏移量, 使模型能够更精准地聚焦于藜麦穗部的关键区域。

结果表明, 改进后的 YOLOv8n 模型训练效果更为理想, 藜麦穗部的检测重复率, 以及误检率得到显著下降, 同时对于图片边缘处的不完整藜麦穗的检测能力显著提高。改进模型对测试集的检测精确度达到 94.3%, 召回率达到 86.2%, 表明其能准确识别绝大多数藜麦穗目标, 同时有效降低误检率。此外, 模型在轻量化方面表现优异, 计算量仅为 6.8 GFLOPs, 模型大小压缩至 7.1 MB, 并结合 166.67 帧/s 的高帧率处理能力, 充分满足无人机机载平台对低算力、低存储和高实时性的严苛要求,

为航拍场景下的藜麦穗高效检测提供了可行的轻量化解决方案。在后续研究中, 将针对藜麦不同品种的穗部特点, 提高对其他藜麦品种的检测能力, 从而提出更全面的藜麦目标检测模型。

利益冲突声明: 本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献:

- [1] 姚燕辉. 苗菜型藜麦新品系筛选及遗传多样性评价[D]. 太谷: 山西农业大学, 2022.
YAO Y H. Screening and genetic diversity evaluation of new strains in seedling type quinoa[D]. Taigu: Shanxi Agricultural University, 2022.
- [2] 朱丽丽, 张发玉, 安宁, 等. 116 份藜麦种质资源萌发期抗旱性综合评价[J]. 干旱地区农业研究, 2024, 42(1): 23-31.
ZHU L L, ZHANG F Y, AN N, et al. Comprehensive evaluation of drought resistance of 116 quinoa germplasm resources during germination[J]. Agricultural Research in the Arid Areas, 2024, 42(1): 23-31.
- [3] AHMAD J, KHAN I, MANZOOR A, et al. Quinoa: An underutilized pseudocereal with promising health and industrial benefits [J]. Journal of Agricultural and Food Chemistry, 2025, 73(24): 14722-14741.
- [4] 姜睿, 刘文瑜, 王旺田, 等. 50 份藜麦种质材料萌发期耐低温综合评价[J/OL]. 草业科学, 2025: 1-17. (2025-06-20) [2020-08-15]. <https://kns.cnki.net/KCMS/detail/detail.aspx?filename=CYKX20250613003&dbname=CJFD&dbcode=CJFQ>.
JIANG R, LIU W Y, WANG W T, et al. Comprehensive evaluation of low temperature tolerance of 50 quinoa germplasm materials at germination stage[J/OL]. Pratacultural science, 2025: 1-17. (2025-06-20) [2020-08-15]. <https://kns.cnki.net/KCMS/detail/detail.aspx?filename=CYKX20250613003&dbname=CJFD&dbcode=CJFQ>.
- [5] ZHANG S, WANG X R, LIN H, et al. A review of the application of UAV multispectral remote sensing technology in precision agriculture[J]. Smart Agricultural Technology, 2025, 12: 101406.
- [6] 龙拥兵, 李坤, 刘丹丹, 等. 基于无人机遥感图像的水稻稻穗识别与估产方法研究[J]. 农业机械学报, 2026, 57(3): 109-118.
LONG Y B, LI K, LIU D D, et al. Rice panicle recognition and yield estimation method based on UAV remote sensing images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2026, 57(3): 109-118.
- [7] 彭小丹, 陈锋军, 朱学岩, 等. 基于无人机图像和改进 LSC-CNN 模型的密集苗木检测和计数方法[J]. 智慧农业(中英文), 2024, 6(5): 88-97.
PENG X D, CHEN F J, ZHU X Y, et al. Dense nursery stock detecting and counting based on UAV aerial images and improved LSC-CNN[J]. Smart Agriculture, 2024, 6(5): 88-97.
- [8] 杨福芹, 李昌浩, 张英发, 等. 融合高光谱和数码影像的冬小麦氮营养指数遥感监测[J]. 光谱学与光谱分析, 2025, 45(6): 1719-1728.
YANG F Q, LI C H, ZHANG Y F, et al. Remote sensing monitoring of nitrogen nutrient index in winter wheat by integrating hyperspectral and digital imagery[J]. Spectroscopy and Spectral Analysis, 2025, 45(6): 1719-1728.
- [9] DASH S K, SEMBHI H, LANGSDALE M, et al. Assessing the field-scale crop water condition over an intensive agricultural plain using UAV-based thermal and multispectral imagery[J]. Journal of Hydrology, 2025, 655: 132966.
- [10] 井梅秀, 穆天红, 肖明, 等. 基于无人机多光谱影像的藜麦长势分析[J]. 现代农业科技, 2025(2): 179-182.
JING M X, MU T H, XIAO M, et al. Growth analysis of quinoa based on multi-spectral images of UAV[J]. Modern Agricultural

- Science and Technology, 2025(2): 179-182.
- [11] RUIZ D A C, VILLACÍS M G M, KIRBY E, et al. Correlation of NDVI obtained by different methodologies of spectral data collection in a commercial crop of quinoa (*Chenopodium quinoa*) in central Ecuador[C]// 2020 Seventh International Conference on eDemocracy & eGovernment (ICEDEG). Piscataway, New Jersey, USA: IEEE, 2020: 208-215.
- [12] FLORES A. Classification of organic quinoa crops using multi-spectral aerial imagery and machine learning techniques[C]// 2022 IEEE International Conference on Automation/XXV Congress of the Chilean Association of Automatic Control (ICA-AC-CA). Piscataway, New Jersey, USA: IEEE, 2022: 1-6.
- [13] VELUSAMY P, RAJENDRAN S, MAHENDRAN R K, et al. Unmanned aerial vehicles (UAV) in precision agriculture: Applications and challenges[J]. Energies, 2022, 15(1): 217.
- [14] 赵峻, 聂志刚, 李广, 等. 基于无人机低空近景图像的玉米螟虫害检测方法[J]. 智慧农业(中英文), 2025, 7(6): 111-123.
ZHAO J, NIE Z G, LI G, et al. A study on corn borer detection using low-altitude close-range UAV imagery[J]. Smart agriculture, 2025, 7(6): 111-123.
- [15] 翁海勇, 姚越, 黄德耀, 等. 无人机低空遥感结合 YOLOv7 快速评估水稻穗颈瘟抗性[J]. 农业工程学报, 2024, 40(21): 110-118.
WENG H Y, YAO Y, HUANG D Y, et al. Rapid evaluation of rice neck blast resistance using low altitude remote sensing of UAV combined with YOLOv7[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(21): 110-118.
- [16] JIA Y J, FU K, LAN H, et al. Maize tassel detection with CA-YOLO for UAV images in complex field environments[J]. Computers and Electronics in Agriculture, 2024, 217: 108562.
- [17] QIU F, SHEN X J, ZHOU C, et al. Rice ears detection method based on multi-scale image recognition and attention mechanism[J]. IEEE Access, 2024, 12: 68637-68647.
- [18] 张晓勤. 无人机遥感图像中玉米雄穗检测模型研究及应用[D]. 重庆: 重庆师范大学, 2023.
ZHANG X M. Research and application of maize tassels detection model in UAV remote sensing images[D]. Chongqing: Chongqing Normal University, 2023.
- [19] 高娟燕, 孙义, 李葆春. 基于无人机 RGB 影像估测田间小麦穗数[J]. 中国农业科技导报, 2022, 24(3): 103-110.
- GAO Y Y, SUN Y, LI B C. Estimating of wheat ears number in field based on RGB images using unmanned aerial vehicle[J]. Journal of Agricultural Science and Technology, 2022, 24(3): 103-110.
- [20] QIAO S Y, CHEN L C, YUILLE A. DetectoRS: Detecting objects with recursive feature pyramid and switchable atrous convolution [C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2021: 10208-10219.
- [21] WEI X T, LI Z S, WANG Y T. SED-YOLO based multi-scale attention for small object detection in remote sensing[J]. Scientific Reports, 2025, 15: 3125.
- [22] WANG S Y, LI Q J, YANG T, et al. LSD-YOLO: Enhanced YOLOv8n algorithm for efficient detection of lemon surface diseases [J]. Plants, 2024, 13(15): 2069.
- [23] WEN C M, CHENG Y, LI S P, et al. Slim-YOLO: An improved sugarcane tail tip recognition algorithm based on YOLO11n for complex field environments[J]. Applied Sciences, 2025, 15(8): 4286.
- [24] ANCHA V K, GONUGUNTALA V, VADDI R. GSS-YOLO: An improved YOLOV5 prediction head with slim-neck for defect detection in printed circuit board assembly[J]. Signal, Image and Video Processing, 2025, 19(11): 915.
- [25] XIA Z F, PAN X R, SONG S J, et al. Vision transformer with deformable attention[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2022: 4784-4793.
- [26] 沈德宇, 陈锋军, 朱学岩, 等. 基于无人机航拍与改进 YOLOv5s 的油茶果实检测[J]. 中国农机化学报, 2024, 45(12): 238-244.
SHEN D Y, CHEN F J, ZHU X Y, et al. *Camellia oleifera* fruit detection based on UAV aerial photography and improved YOLOv5s[J]. Journal of Chinese Agricultural Mechanization, 2024, 45(12): 238-244.
- [27] 杨启良, 禹璐, 梁嘉平. 基于改进 YOLOv11 的采后芦笋分级检测方法[J]. 智慧农业(中英文), 2025, 7(4): 84-94.
YANG Q L, YU L, LIANG J P. Grading *Asparagus officinalis* L. using improved YOLOv11[J]. Smart Agriculture, 2025, 7(4): 84-94.

YOLOv8n-SSND: An Improved Lightweight Model for Aerial *Chenopodium quinoa* Willd. Spike Target

WU Tingting¹, GUO Junrui¹, TAO Qiujie¹, CHEN Shihua², GUO Shanli^{2*}

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100, China;

2. College of Life Sciences, Yantai University, Yantai 264006, China)

Abstract: [Objective] The *Chenopodium quinoa* panicle is a critical phenotypic indicator for estimating crop yield and evaluating the growth condition of *Chenopodium quinoa* plants. Accurate and efficient recognition of *Chenopodium quinoa* panicles in complex field environments is therefore of great significance for intelligent agriculture, yield prediction, and automatic crop management. However, unmanned aerial vehicle (UAV)-acquired field imagery often exhibits complex characteristics such as diverse panicle morphology, uneven illumination, overlapping occlusion, and background interference, et al., posing substantial challenges for conventional target detection algorithms. To address these issues, a lightweight target detection model, named YOLOv8n-SSND (YOLOv8n with Switchable Atrous Convolution, Slim Neck, and Deformable Attention) is proposed, and specifically optimized for UAV-based *Chenopodium quinoa* panicle identification to improve the detection accuracy and inference efficiency for *Chenopodium quinoa* panicles while maintaining low computational cost and real-time performance suitable for embedded UAV deployment. **[Methods]** The proposed model was constructed based on the YOLOv8n and YOLOv11n frameworks, and incorporated several improvements tailored for small-ob-

ject agricultural detection tasks. To enhance the ability to capture multi-scale and high-dimensional semantic features, the switchable atrous convolution (SAC) module was embedded into the backbone network. This module dynamically adjusted its receptive field according to spatial context, enabling more precise extraction of local and global texture details of *Chenopodium quinoa* panicles. In order to reduce redundant parameters and maintain high computational efficiency, a slim-neck lightweight feature fusion layer was designed, which effectively strengthened the integration of shallow spatial information and deep semantic features, allowing the network to maintain high accuracy without increasing model complexity. Additionally, a deformable attention (DA) mechanism was introduced to enable adaptive focus on regions with rich panicle-related features while suppressing irrelevant background noise. This attention mechanism assigned dynamic weights across both spatial and channel dimensions, improving the model's robustness against occlusions, illumination variations, and complex field textures commonly encountered in UAV images. **[Results and Discussions]** Comprehensive field experiments were conducted using UAV images of *Chenopodium quinoa* plots collected under different environmental conditions and growth stages. The results demonstrated that the proposed YOLOv8n-SSND model achieved a mean average precision (mAP50) of 94.3%, showing a remarkable improvement over multiple baseline and comparative models. Specifically, compared with YOLOv11n-SSND, YOLOv11n, YOLOv12n, YOLOv7, YOLOv5s, single shot multibox detector (SSD), fast region-based convolutional neural network (Fast R-CNN) and YOLOv8n, the proposed model achieved improvements of 0.7, 0.9, 2.1, 1.4, 2.0, 23.1, 19.6 and 1.8 percentage points respectively (SSD and Fast R-CNN). In terms of computational efficiency, the inference speed reached 166.7 f/s, representing a 26.7% increase over the YOLOv8n baseline, which ensured real-time detection capability for UAV-mounted onboard processors. Moreover, the total operation count was reduced to 6.8 GFLOPs, reflecting a 16.0% reduction compared with the baseline model, thus demonstrating the improved efficiency of the proposed architecture. The experimental comparison also indicated that the integration of SAC enhanced the model's sensitivity to complex spatial patterns, while the DA module effectively improved feature selectivity and prevented overfitting to background textures. The Slim-Neck design contributed significantly to reducing parameter redundancy and facilitated smooth feature propagation across layers. **[Conclusions]** The YOLOv8n-SSND model effectively achieves a balance among detection accuracy, inference speed, and computational cost, making it well-suited for real-time UAV-based agricultural monitoring. The experimental outcomes confirm that the model not only provides high-precision detection of *Chenopodium quinoa* panicles but also offers superior inference efficiency with minimal computational resources. These characteristics make it a promising solution for UAV-deployed intelligent agricultural systems, where power and processing capacity are limited. Furthermore, the proposed method provides a technical foundation for large-scale and automated monitoring of *Chenopodium quinoa* growth, enabling accurate yield estimation, phenotypic analysis, and precision crop management.

Keywords: *Chenopodium quinoa* panicle; UAV; YOLOv8n; object detection; deformable attention; crop phenotype

Foundation items: Key Research Project of Shandong Province for Agricultural Superior Varieties (2023LZGC011)

First author: WU Tingting, E-mail: tt_wu@nwsuaf.edu.cn

***Corresponding author:** GUO Shanli, E-mail: gsl@ytu.edu.cn