

基于改进 YOLO v8n 的多育种场景下大豆主茎表型解析方法

饶元^{1,2} 李成宇^{1,2} 万天与^{1,2} 金秀^{1,2} 江丹^{1,2} 王坦^{1,2}
李佳佳³ 张传坤^{1,2}

(1. 安徽农业大学信息与人工智能学院, 合肥 230036; 2. 农业农村部农业传感器重点实验室, 合肥 230036;
3. 安徽农业大学农学院, 合肥 230036)

摘要: 主茎表型准确解析对理解植株生长模式、优化种植密度以及提高作物产量具有重要意义。本文提出了一种基于改进 YOLO v8n 的多育种场景下大豆主茎节点检测模型 (YOLO-SN), 并对主茎节点数量、节间距和主茎长度表型参数进行了解析。在主干构建 DAF 模块 (Deformable attention fusion module), 增强模型对多场景下节点区域关注能力; 在头部设计动态解耦头 (Dynamic decouple-head, DDH), 提升模型对大豆主茎节的感知能力; 引入动态非单调聚焦机制 (Wise intersection over union, WIoU), 提升网络收敛速度与模型泛化能力。试验结果表明, 在单一场景下主茎节检测准确率、召回率、平均精度均值和 F1 分数均有明显提升, 在多场景下综合性能分别达 90.6%、85.1%、90.0% 和 87.8%, 均优于主流目标检测模型。主茎节点数量、节间距和主茎长度的绝对误差、决定系数分别为 0.42、12.6 像素、17.4 像素和 0.88、0.80、0.82, 实现了多场景下大豆主茎表型高精度解析。研究结果为不同育种环境下大豆植株表型准确解析提供了有效方法。

关键词: 多育种场景; 大豆主茎; 表型解析; YOLO v8n

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2026)10-0189-09

OSID:



Soybean Main Stem Phenotype Analysis Method Based on Improved YOLO v8n in Multiple Breeding Scenarios

RAO Yuan^{1,2} LI Chengyu^{1,2} WAN Tianyu^{1,2} JIN Xu^{1,2} JIANG Dan^{1,2} WANG Tan^{1,2}
LI Jiajia³ ZHANG Chuankun^{1,2}

(1. School of Information and Artificial Intelligence, Anhui Agricultural University, Hefei 230036, China
2. Key Laboratory of Agricultural Sensors, Ministry of Agriculture and Rural Affairs, Hefei 230036, China
3. School of Agronomy, Anhui Agricultural University, Hefei 230036, China)

Abstract: Parsing of main stem phenotype is of great significance for understanding plant growth patterns, optimizing planting density and improving crop yield. In this study, one soybean main stem node detection model YOLO-SN based on improved YOLO v8n was proposed for multiple breeding scenarios, and the phenotypic parameters of main stem node number, internode length and main stem length were analyzed. Firstly, aiming at the problems of poor model adaptability and high missed detection rate caused by significant differences in the characteristics of soybean main stem nodes in three typical breeding scenarios, including open field, indoor and artificial climate chamber, respectively. The DAF (Deformable attention fusion module) module is constructed in the backbone to enhance the model's ability to pay attention to node regions in multiple scenarios; Secondly, for the problem that the small pixel share of main stem nodes leads to recognition difficulties, the DDH (Dynamic decouple-head) was designed in the head to improve the model's ability to perceive the main stem nodes of soybeans; Finally, the dynamic non-monotonic focusing mechanism WIoU (Wise intersection over union, WIoU) was designed for further improving the convergence speed and generalization of original network YOLO v8n.

收稿日期: 2025-02-22 修回日期: 2025-06-21

基金项目: 国家自然科学基金项目 (32371993)、安徽省重点研究与开发计划项目 (202204c06020026、2023n06020057) 和安徽省高校自然科学研究重大项目 (2022AH040125、2023AH040135)

作者简介: 饶元 (1982—), 男, 教授, 博士, 主要从事机器视觉和深度学习研究, E-mail: raoyuan@ahau.edu.cn

The experimental results show that the main stem node detection accuracy, recall, average precision mean and F1-score are significantly improved in a single scene, and the comprehensive performance in multiple scenes reaches 90.6%, 85.1%, 90.0% and 87.8%, respectively, which are all better than the mainstream target detection models. The absolute error and coefficient of determination of the number of main stem nodes, the distance between nodes and the length of main stem were 0.42, 12.6 pixels, 17.4 pixels and 0.88, 0.80, 0.82, respectively. This study provide an effective method for the accurate parsing of soybean plant phenotypes in typical breeding environments, as well as offering technical support for intelligent breeding of soybeans and other legumes.

Key words: multi-breeding scenarios; soybean main stem; phenotype parsing; YOLO v8n

0 引言

植物表型是指植物结构可测量的形态参数和性状,由植物遗传物质和生长环境共同决定,对农作物育种和栽培研究至关重要^[1]。近年来,中国大豆进口量逐年增加,2023 年中国全年大豆进口量为 9.941×10^7 t,占全部粮食进口 60% 以上^[2]。大豆生长时期节点数量、节间距和主茎长度反映作物对于生长环境的适应性,对大豆生长状况评估具有重要意义^[3]。大田、室内和人工气候室作为典型的大豆育种场景,对大豆表型监测和解析需求尤为突出。计算机视觉技术能够实现作物表型信息精确、高效提取,远超劳动密集型人工处理方式^[4]。ZHU 等^[5]结合大豆生殖期特点,提出了基于改进 Faster R-CNN 的关键表型识别和计数融合模型,大豆花和豆荚计数决定系数分别达 0.965 和 0.98。管贤平等^[6]采用机载三维激光雷达系统,使用局部邻域特征分割与均值漂移算法提取大田单株大豆株高和冠层体积,实现了一种田间植株冠层几何参数准确获取方法。LIU 等^[7]采用 MobileNetv2 改进 YOLO v5 的主干网络,同时引入高效通道注意模块(Efficient channel attention, ECA)和完整交并比损失函数(Complete intersection over union, CIoU),提升了大豆叶片和花朵识别和计数能力。

加强对主茎表型的关注有助于了解作物生长和发育过程,从而为作物遗传改良和产量提升提供参考^[8]。LI 等^[9]研究大豆节点部位结荚特性、株高与抗倒伏的负相关性以及节点和节间距共同控制株高的关系,有助于培育密植下抗倒伏且高产的理想大豆品种。YAMAMOTO 等^[10]将机器学习用于室内场景单株番茄幼苗节点检测和节间距估计,但未考虑复杂田间种植场景对检测效果的影响。NGUYEN 等^[11]在室内场景下重建单株植物点云模型并进行主茎表型解析,但受叶片形态及植株尺寸影响较大。BOOGAARD 等^[12]采用 YOLO v3 模型检测气候室内黄瓜植株节点位置并计算节间距,然而检测效果受到光照条件以及植株间相互遮挡的限制。SUN

等^[13]利用 3D LiDAR 获取高精度单株田间棉花点云,使用拉普拉斯收缩法提取植物节点骨架,但该方法只适用于植株最高点位于主茎的情况。BOOGAARD 等^[14]重建气候室中黄瓜植株点云并提取主茎表型,但节点识别准确率和召回率远低于二维图像方法。HU 等^[15]基于改进 YOLO v5 模型对多种作物进行主茎节点检测和节间距估计,对复杂环境导致节点误检漏检考虑不够充分。

上述关于主茎节、节间距和主茎长度等作物表型信息提取为大豆主茎表型解析提供了借鉴。但目前针对单一环境中作物生长监测,未充分考虑不同环境因素以及节点固有生长特性对模型检测效果的影响。不同育种环境下大豆主茎节特征差异较大,茎节区域像素占比低且与茎秆纹理相似,特征不够显著。本文提出基于改进 YOLO v8n 的多育种场景下大豆主茎节点检测模型 YOLO-SN,通过构建基于可形变自注意力机制的 DAF 模块,设计 DDH 模块和引入损失函数 WIoU,以提升大豆主茎节点检测模型泛化能力,为不同育种环境下大豆植株表型准确解析提供支持。

1 材料与方法

1.1 数据集采集和建立

构建的大豆数据集包含大田、室内、人工气候室 3 个典型育种场景,使用 2 种常见大豆品种,数量分别为 100、50 株。田间大豆于 2023 年 9 月栽种于安徽农业大学农萃园试验田,室内大豆于 2023 年 9 月在室内环境中单盆栽培,其生长发育过程受到昼夜温度等环境因素影响。人工气候室大豆于 2023 年 10 月种植在安徽农业大学人工气候室内,大豆生长处于较为稳定的温度、相对湿度、光照条件下。采集设备为微软 Azure Kinect DK 深度相机,采集的大豆样本处在大豆生长 V1~R3 时期,拍摄距离为 20~30 cm。数据采集时间为 2023 年 10—11 月,拍摄时间为 09:00—16:00,共采集 1 559 幅图像,包括大田 490 幅、室内 550 幅、人工气候室 519 幅,图 1 为大田落叶杂草复杂场景、室内简单环境和人工气候室内

处在灯光阴影下的部分大豆植株样本。

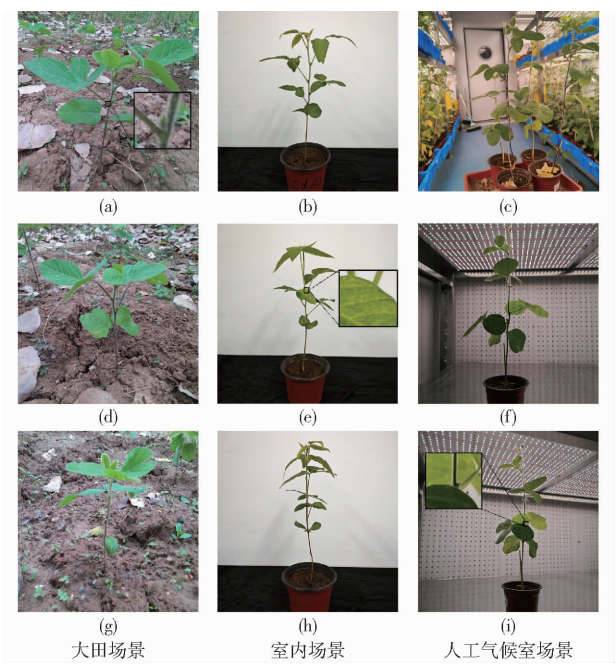


图 1 大豆植株采集图像示例

Fig. 1 Example of soybean plant collection images

采用 Roboflow 数据标记网站分别对 3 个场景大豆主茎节点进行标注,从大豆子叶节开始至顶端节点结束。如图 1 所示,根据主茎节点遮挡类型分为节点无遮挡(图 1a)、节点部分遮挡(图 1i)和节点完全遮挡(图 1e)3 种情况,对于完全遮挡的节点不进行标记处理。将大田、室内、人工气候室 3 个场景混合数据均按照比例 7:2:1 划分为训练集、验证集和测试集,分别构建 3 个单一场景数据集和 1 个多场

景混合数据集,分别用于单一场景模型性能验证和消融试验评估。使用裁剪、旋转、曝光、翻转、调整亮度和增加噪点等方法对训练集进行数据增强。

1.2 大豆主茎表型参数解析方法

1.2.1 YOLO-SN 模型构建

目前,YOLO 系列网络已广泛应用于农业领域表型智能解析和果实检测等任务^[16-17]。近年来提出的 YOLO v8 采用新的主干网络,使用梯度流更丰富的 C2F(CSP bottleneck with 2 convolutions)结构;头部网络采用主流的解耦头结构(DeCouple-Head),同时也从 Anchor-Based 换成了 Anchor-Free;损失函数部分使用正负样本匹配方式,并引入分布焦点损失。这些改进降低了模型复杂度,增强了模型特征提取能力。

针对多育种场景下大豆茎节特征差异大、像素占比低等问题,提出了多育种场景下基于改进 YOLO v8n 的大豆主茎节点检测模型 YOLO-SN,以提高多场景下目标检测模型泛化能力。模型结构如图 2 所示,由主干网络(Backbone)、颈部网络(Neck)、头部网络(Head)构成。其中,主干网络由 CBS(Conv + BatchNorm + SiLU)、C2F、SPPF(Spatial pyramid pooling fast)和 DAF 模块堆叠组成,旨在提高网络特征提取能力;颈部网络由路径聚合网络(Path aggregation network,PANet)构成,以加强网络对不同尺度对象的特征融合能力;Head 部分由 DDH 模块和改进的损失函数 WIoU 构成,提升对小目标节点检测能力以及加快网络收敛速度。

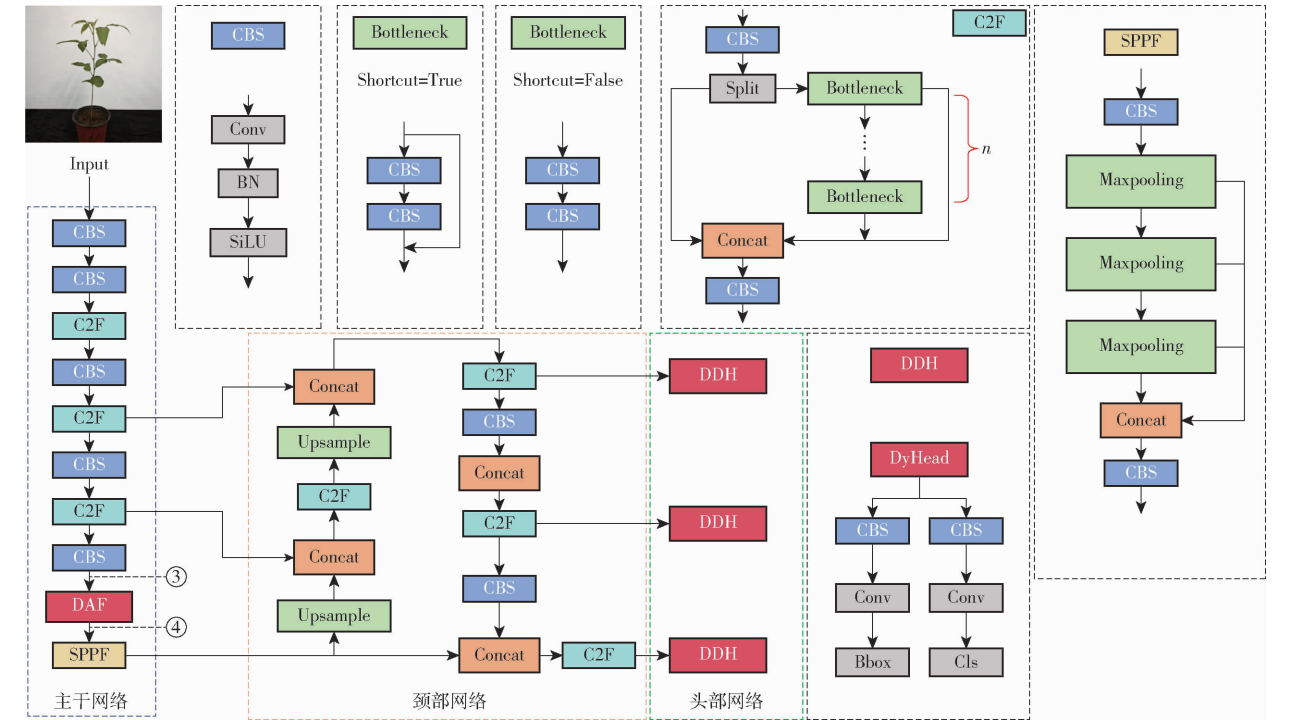


图 2 YOLO-SN 整体结构

Fig. 2 Overall structure of YOLO-SN

(1)DAF 模块

如图 3 所示,DAF 结构由 CBS、Split、Concat 和可形变自注意力^[18] (Deformable attention, DEF - Attention) 模块构成。CBS 模块包含 Conv 层、BatchNorm 层和 SiLU 激活函数,Split 模块在通道维度将输入特征一分为二,Concat 模块在通道维度拼接特征,Bottleneck - DEF 模块由 2 个 CBS 模块和 DEF - Attention 以及残差连接构成。DAF 模块特征提取过程如下:首先,经第 1 个 CBS 模块处理后的

特征由 Split 模块按通道数一分为二。一部分特征进入 n 个 Bottleneck - DEF 模块进行处理,另一部分特征则保留并直接传递至 Concat 模块。在每个 Bottleneck - DEF 模块中,处理的特征再次被分为 2 个分支,1 个分支将处理过的特征传递给下一个 Bottleneck - DEF 模块,另 1 个分支则保留并直接传递至 Concat 模块。最后将所有传递至 Concat 模块的特征在通道维度上拼接,并通过 1 个 CBS 模块进行处理后输出。

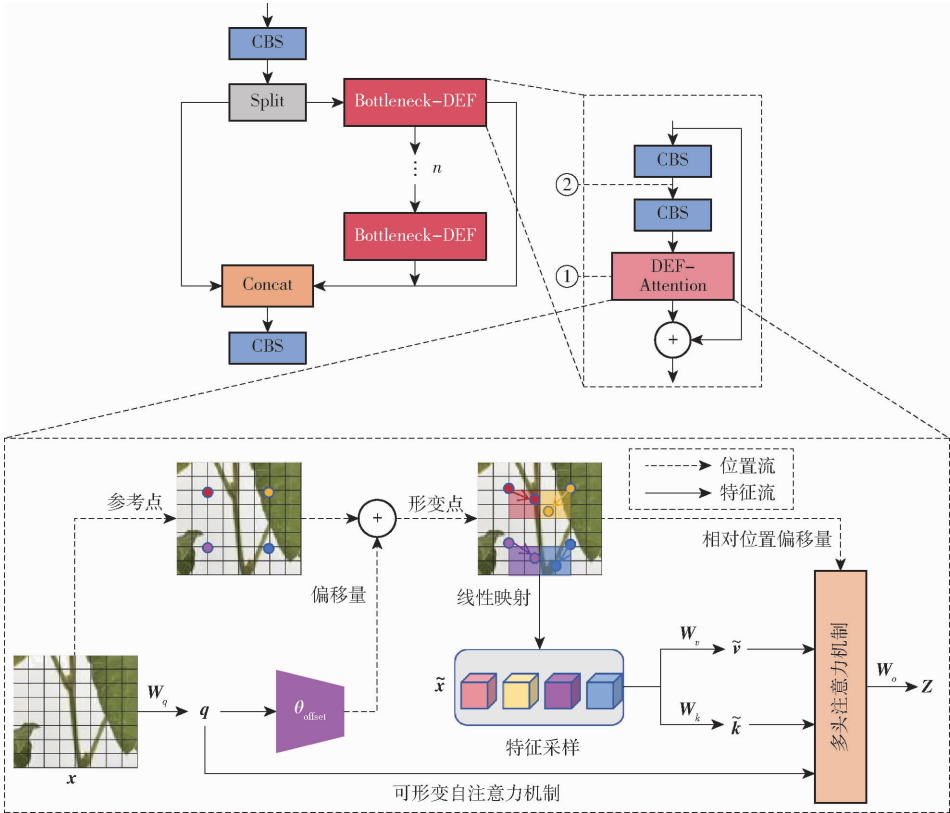


图 3 DAF 模块结构

Fig.3 Structure of DAF

DEF - Attention 通过数据依赖的方式选择自注意力中键值对的位置,使模型能够专注于茎节区域,捕获更多有信息量的特征并减少不同环境背景干扰。根据输入特征图 $\mathbf{x} \in \mathbf{R}^{H \times W \times C}$ 的尺寸和采样因子生成参考网格,参考点的值由线性间隔的二维坐标表示,并进行归一化处理,使参考点坐标数值位于 $[-1,1]$ 。将特征图线性映射得到 $\mathbf{q}$,再经过偏移网络 (Offset network) 生成偏移量 Δp ,即

$$\begin{cases} \Delta p = \theta_{\text{offset}}(\mathbf{q}) \\ \tilde{\mathbf{x}} = \phi(\mathbf{x}; \mathbf{p} + \Delta p) \end{cases} \quad (1)$$

利用双线性插值法对形变点位置进行特征采样,再线性映射得到 $\tilde{\mathbf{k}}$ 和 $\tilde{\mathbf{v}}$,即

$$\begin{cases} \mathbf{q} = \mathbf{x} \mathbf{W}_q \\ \tilde{\mathbf{k}} = \tilde{\mathbf{x}} \mathbf{W}_k \\ \tilde{\mathbf{v}} = \tilde{\mathbf{x}} \mathbf{W}_v \end{cases} \quad (2)$$

式中 $\mathbf{W}_q, \mathbf{W}_k, \mathbf{W}_v$ ——映射矩阵

$\mathbf{q}, \tilde{\mathbf{k}}, \tilde{\mathbf{v}}$ ——查询、形变键和值

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

式中 $\phi(\cdot; \cdot)$ ——双线性插值函数

(p_x, p_y) ——网格点坐标

$g(a, b)$ ——函数 $\max(0, 1 - |a - b|)$

(r_x, r_y) ——索引 $\mathbf{z} \in \mathbf{R}^{H \times W \times C}$ 上的所有位置

由于 g 仅在最接近 (p_x, p_y) 的 4 个积分点上非零,因此简化为 4 个位置加权平均值。使用标准多头注意力并采用相对位置偏移量 R ,输出结果为

$$\begin{aligned} \mathbf{Z}^{(m)} &= \sigma(\mathbf{q}^{(m)} \tilde{\mathbf{k}}^{(m)\text{T}} / \sqrt{d} + \phi(\hat{B}; R)) \tilde{\mathbf{v}}^{(m)} \\ (m &= 1, 2, \dots, M) \end{aligned} \quad (4)$$

其中 $d = D/M$

式中 $\hat{B}$ ——相对位置偏表

$\sigma(\cdot)$ ——softmax 函数

d ——每个头维度

拼接每个头部特征并通过 W_o 映射获得输出 Z , 即

$$Z = \text{Concat}(z^{(1)}, z^{(2)}, \dots, z^{(M)}) W_o \quad (5)$$

(2) 动态解耦头 DDH

在不同距离和角度下采集大豆图像, 导致大豆标记的检测框尺寸不一致, 使得不同尺度节点特征可能出现在同一幅图像中, 需检测器有较高的多尺度感知能力; 生长在不同场景下的大豆节点形状纹

理均有区别, 大豆主茎节点沿着主茎空间分布, 且节点尺寸较小, 需检测器有较高的空间感知能力。DDH 模块由 DyHead^[19] 和解耦头结构组成, 如图 4 所示, DyHead 由尺度感知注意力、空间感知注意力和任务感知注意力 3 部分组成。尺度感知注意力定义为

$$\pi_L(F)F = \varphi\left(f\left(\frac{1}{S_c} \sum_{s,c} F\right)\right)F \quad (6)$$

式中 $f(\cdot)$ —— 1×1 卷积层近似实现的线性函数

$\varphi(\cdot)$ ——hard sigmoid 函数

F ——张量

S_c ——特征图在通道维度上平均值

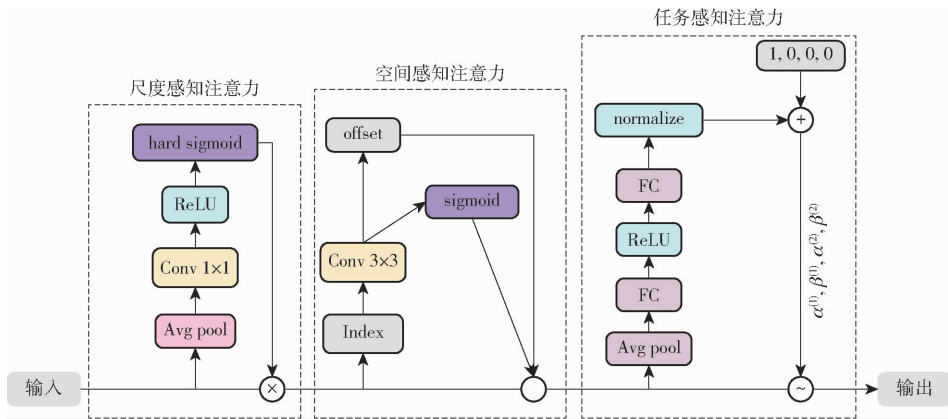


图 4 DyHead 整体结构

Fig. 4 Overall structure of DyHead

空间感知注意力定义为

$$\pi_s(F)F = \frac{1}{L} \sum_{l=1}^L \sum_{k=1}^K w_{l,k} F(l; p_k + \Delta p_k; c) \Delta m_k \quad (7)$$

式中 K ——稀疏采样位置数量

Δm_k ——位置 p_k 处自学习重要性度量因子

任务感知注意力定义为

$$\pi_c(F)F = \max(\alpha^{(1)}(F)F_c + \beta^{(1)}(F), \alpha^{(2)}(F)F_c + \beta^{(2)}(F)) \quad (8)$$

其中 $[\alpha^{(1)} \quad \alpha^{(2)} \quad \beta^{(1)} \quad \beta^{(2)}]^T = \theta(\cdot)$, $\alpha^{(1)}$ 、 $\alpha^{(2)}$ 、 $\beta^{(1)}$ 、 $\beta^{(2)}$ 表示用于学习控制激活阈值超参数。 $\theta(\cdot)$ 首先在 $L \times S$ 的维度上执行全局池化, 再使用 2 个全连接层、1 个归一化层, 并最终使用 shifted sigmoid 函数归一化输出。

(3) 损失函数 WIoU

YOLO v8 网络预测边框坐标损失 L_{CioU} 采用 CIOU^[20] Loss, 在预测框回归过程中预测框与真实框高宽比呈线性关系时, CIOU 惩罚项退化为 0。采用 WIoU v3^[21] 替换原始模型的边界损失函数对模型进行优化, 其损失函数为

$$L_{WIoU} = r R_{WIoU} L_{IoU} \quad (R_{WIoU} \in [1, e], L_{IoU} \in [0, 1]) \quad (9)$$

$$\text{其中 } R_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \quad (10)$$

$$r = \frac{\beta}{\delta \alpha^{\beta - \delta}} \quad (11)$$

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty)$$

式中 R_{WIoU} ——距离聚焦机制

L_{IoU} ——IoU 损失函数

x, y ——预测框中心点坐标

x_{gt}, y_{gt} ——真实框中心点坐标

W_g, H_g ——预测框与真实框最小外接矩形宽和高

$*$ —— W_g, H_g 从计算图中分离

r ——非单调聚焦系数

β ——离群度

$\bar{L}_{IoU}$ ——动态滑动平均值

α, δ ——超参数

通过降低高质量样本对损失值贡献, 动态给予边界框梯度增益, 聚焦于普通梯度质量的锚框, 提高模型检测能力。

1.2.2 主茎表型计算方法和评价指标

大豆主茎节点计数从子叶节以上开始到顶端结

束,不包括子叶节和顶端花序^[22]。通过统计模型生成检测框个数,再排除子叶节点检测框获得主茎节数量。大豆主茎节间距是指大豆主茎上 2 个相邻节点间的距离^[23],采用计算相邻检测框中心欧氏距离作为节间距。大豆主茎长度从子叶节以上至顶端节点^[24],通过累计相邻主茎节点间距离作为主茎长度。

采用准确率 (Precision)、召回率 (Recall)、平均精度均值 (mean Average Precision, mAP) 和 F1 分数作为评价指标。为了定量描述节点数量、节间距和主茎长度计算准确性,采用决定系数 (Coefficient of determination, R^2) 和平均绝对误差 (Mean absolute error, MAE) 进行评价,其定义为预测结果与手工标注结果之间像素点数量差值。

2 试验结果与分析

为了验证模型在多个场景下对大豆主茎节点检测的优越性,消融试验、可形变自注意力不同位置性能试验、不同损失函数性能试验、不同目标检测模型性能试验和主茎表型解析结果分析均采用多场景混合数据集进行训练和验证。而在不同育种场景试验中,为更准确评估模型场景适应性,分别使用大田、室内和人工气候室 3 个单一育种场景的数据集开展模型性能验证。

2.1 消融试验

消融试验结果如表 1 所示,当只添加 DAF 模块,模型准确率、召回率、mAP 和 F1 分数分别提升 1.9、1.2、1.1、1.6 个百分点,表明以数据依赖的方式学习稀疏的注意力模式,使得模型感受野聚焦不同场景下大豆主茎节点区域,提升了模型特征提取能力;只添加 DDH 模块,模型准确率、召回率、mAP 和 F1 分数分别提升 2.8、0.3、1.3、1.5 个百分点,说明在空间分布和尺度上提升了模型对于节点的感知能力;只引入 WIoU v3 损失函数,模型准确率、召回率、mAP 和 F1 分数分别提升 4.6、1.5、2.0、3.0 个百分点,表明 WIoU v3 平衡了高质量锚框和低质量锚框的影响,提高了模型拟合程度。当同时添加 DAF 和 DDH 模块,引入计算参数,模型准确率、mAP 和 F1 分数分别提升 3.5、1.1、0.9 个百分点,召回率下降 1.5 个百分点,模型内存占用量增加 17.7%;同时添加 DAF 模块和引入 WIoU v3,模型准确率、召回率、mAP 和 F1 分数分别提升 1.7、3.2、2.7、2.5 个百分点;同时添加 DDH 和 WIoU v3,模型准确率、召回率、mAP 和 F1 分数分别提升 2.0、3.7、2.9、2.9 个百分点。当同时添加 DAF、DDH 和 WIoU v3,既提升了模型对节点区域的感知能力和小目标的检测能

力,又平衡了锚框质量影响,模型准确率、召回率、mAP 和 F1 分数分别提升 4.5、3.2、3.2、3.9 个百分点。通过加大对节点区域特征的关注、提升对小目标的识别和加快网络收敛速度,相比原始模型,改进后模型对多场景下的大豆主茎节点的检测性能均有较为明显提升。

表 1 消融试验结果
Tab.1 Ablation experiments

DAF	DDH	WIoU	准确	召回	平均精度	模型内存	F1 分
		v3	率/%	率/%	均值/%	占用量/MB	
×	×	×	86.1	81.9	86.8	5.99	83.9
✓	×	×	88.0	83.1	87.9	6.12	85.5
×	✓	×	88.9	82.2	88.1	6.91	85.4
×	×	✓	90.7	83.4	88.8	5.99	86.9
✓	✓	×	89.6	80.4	87.9	7.05	84.8
✓	×	✓	87.8	85.1	89.5	6.12	86.4
×	✓	✓	88.1	85.6	89.7	6.91	86.8
✓	✓	✓	90.6	85.1	90.0	7.05	87.8

注:✓表示使用该模块;×表示未使用该模块。

2.2 可形变自注意力不同位置性能试验

在构建 DAF 模块时,通过在不同位置引入 DEF-Attention,验证其在各个位置效果,最终确定最优插入位置,从而确定 DAF 模块构成。如图 2 中③、④所示,分别在 Backbone 最后一个 C2F 模块前和后添加 DEF-Attention 模块。如图 3 中①、②所示,在 C2F 中 Bottleneck-DEF 的第 1 个 CBS 模块后和第 2 个 CBS 模块后分别添加 DEF-Attention。结果如表 2 所示,在 C2F 中 Bottleneck-DEF 的第 1 个 CBS 模块后和第 2 个 CBS 模块添加 DEF-Attention 准确率、召回率和 F1 分数较为接近,但在第 2 个 CBS 模块后添加 mAP 最高为 87.9%,召回率最高为 83.1%,F1 分数最高为 85.5%,因此选择图 3 中①位置为最优位置。

表 2 可形变自注意力位置消融试验结果
Tab.2 Deformable attention location ablation experiments

模型	准确	召回	平均精度	模型内存	F1 分
	率/%	率/%	均值/%	占用量/MB	数/%
① + YOLO v8n	88.0	83.1	87.9	6.12	85.5
② + YOLO v8n	87.8	82.9	87.4	6.12	85.3
③ + YOLO v8n	88.3	80.3	86.4	6.50	84.1
④ + YOLO v8n	87.9	81.0	86.9	6.50	84.3

2.3 不同损失函数性能试验

通过在 YOLO v8n 模型中引入不同损失函数进行性能试验,可以体现 WIoU v3 的优越性。由图 5 可知,采用 CIoU、DIoU^[20]时,模型训练损失收敛速度和最终收敛的损失函数值相近;采用 GIoU^[25]时,损失收敛速度慢于 CIoU 和 DIoU,但最终收敛的损

失函数值与 CIoU 和 DIoU 相近;采用 WIoU v3 后,训练时模型损失收敛速度最快,收敛后损失值远小于另外 3 种损失函数。由于节点目标区域较小,生成的预测框可能存在完全重叠的情况,在这种情况下 GIoU 不能反映出实际情况退化为 IoU,并且 GIoU 对每个预测框与真实框均计算最小外接框,故损失函数计算及收敛速度受到限制。DIoU 考虑重叠面积和中心点距离,但没有考虑到回归过程中检测框的纵横比。CIoU 没有考虑低质量锚框和高质量锚框的数量不平衡问题,因此在训练时会降低模型泛化能力。WIoU v3 使用动态非单调评估锚盒质量的机制,使模型更加关注锚定普通质量的框,并提高了模型定位对象的能力。

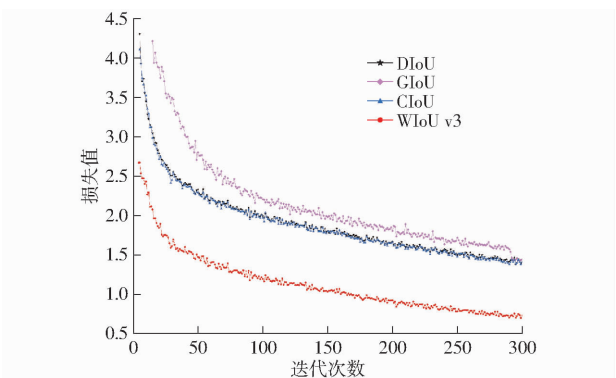


图 5 不同损失函数收敛曲线

Fig. 5 Comparison of different loss functions

2.4 不同目标检测模型性能试验

由表 3 可知,YOLO-SN 准确率、召回率、平均精度均值和 F1 分数均高于其他模型,其中准确率分别比 SSD^[26]、Faster R-CNN^[27]、YOLO v5n、YOLO v7-tiny^[28]、YOLO v8n、YOLO v9-t^[29]、YOLO v10n^[30] 高 33.0、6.8、6.3、9.1、4.5、7.8、5.1 个百分点,召回率分别高 48.8、39.8、8.2、9.3、3.2、13.5、5.8 个百分点,mAP 分别高 38.9、8.9、8.4、7.6、3.2、11.1、5.6 个百分点,F1 分数分别高 44.3、33、7.4、9.3、3.9、11.0、5.5 个百分点。模型内存占用量相较 YOLO v8n、YOLO v9-t、YOLO v10n 较为接近,但明显小于 YOLO v7-tiny 模型内存占用量。两阶段网络模型 Faster R-CNN 先生成建议框再进行后续的处理操作,因此模型内存占用量和计算量较大。一阶段网络模型 SSD、YOLO v5n、YOLO v6、YOLO v7-tiny、YOLO v8n 等仅靠 1 个网络同时输出包含坐标和类别置信度,因此模型内存占用量和计算量小于 Faster R-CNN。YOLO-SN 相较于其他主流网络,构建 DAF 模块使得模型可以根据输入动态调整模型内存占用量,因此对于多场景下大豆节点目标检测效果更好;头部 DDH 模块分别在尺度、空间分布和任务上对目标进行特征处理,因此对于不同

尺度、不同位置 and 不同形状的目标更加敏感;WIoU v3 平衡高质量和低质量锚框对回归的影响,提升了检测效果。

表 3 不同目标检测模型的大豆主茎节点检测结果
Tab.3 Comparison of soybean main stem node detection results of different models

模型	准确	召回	平均精度	模型内存	F1 分
	率/%	率/%	均值/%	占用量/MB	数/%
SSD	57.6	36.3	51.1	182.28	44.5
Faster R-CNN	83.8	45.3	81.1	320.03	58.8
YOLO v5n	84.3	76.9	81.6	3.68	80.4
YOLO v7-tiny	81.5	75.8	82.4	11.70	78.5
YOLO v8n	86.1	81.9	86.8	5.99	83.9
YOLO v9-t	82.8	71.6	78.9	5.81	76.8
YOLO v10n	85.5	79.3	84.4	5.53	82.3
YOLO-SN	90.6	85.1	90.0	7.05	87.8

2.5 不同育种场景性能试验

通过在 3 个单一场景数据集上进行模型试验,以证明 YOLO-SN 在不同育种场景下模型对主茎节检测效果的提升。试验结果如表 4 所示,相较原始 YOLO v8n 模型,YOLO-SN 在单一大田场景下准确率、召回率和 mAP 分别提升 6.2、5.5、5.1 个百分点;在单一室内场景下准确率、召回率和 mAP 分别提升 4.6、2.7、1.4 个百分点;在单一人工气候室场景下的准确率、召回率和 mAP 分别提升 1.1、3.8、3.6 个百分点;在多场景下准确率、召回率和 mAP 分别提升 4.5、3.2、3.2 个百分点。由于单一场景中部分遮挡的节点样本少且不同生长期节点特征差异大,导致模型检测效果不佳、漏检率高;而融合多场景数据显著增加了部分遮挡节点和同一生长期节点的特征数量,同时不同场景下同生长期植株的形态、纹理、节点空间分布等关键特征具有相似性,因此多场景节点检测效果优于单场景检测效果。通过构建 DAF 模块、DDH 头部和引入损失函数 WIoU v3,减少了环境对节点检测的影响,单一场景和多个场景下大豆主茎节点检测性能显著提升。

表 4 不同育种场景性能试验
Tab.4 Different scenes comparison experiments %

场景	模型	准确率	召回率	平均精度 均值
大田	YOLO v8n	79.5	71.6	80.6
	YOLO-SN	85.7	77.1	85.7
室内	YOLO v8n	84.3	81.0	89.4
	YOLO-SN	88.9	83.7	90.8
人工气候室	YOLO v8n	87.5	77.5	84.3
	YOLO-SN	88.6	81.3	87.9
多场景	YOLO v8n	86.1	81.9	86.8
	YOLO-SN	90.6	85.1	90.0

2.6 主茎表型解析结果分析

如图 6 所示,首先输入数据经过 YOLO – SN 模型进行主茎节点检测输出检测框,根据检测框数量确定主茎节点个数,计算相邻检测框中心之间的欧氏距离作为节间距,再计算所有节间距作为主茎长度,最后对预测表型与真实标记框进行误差分析。

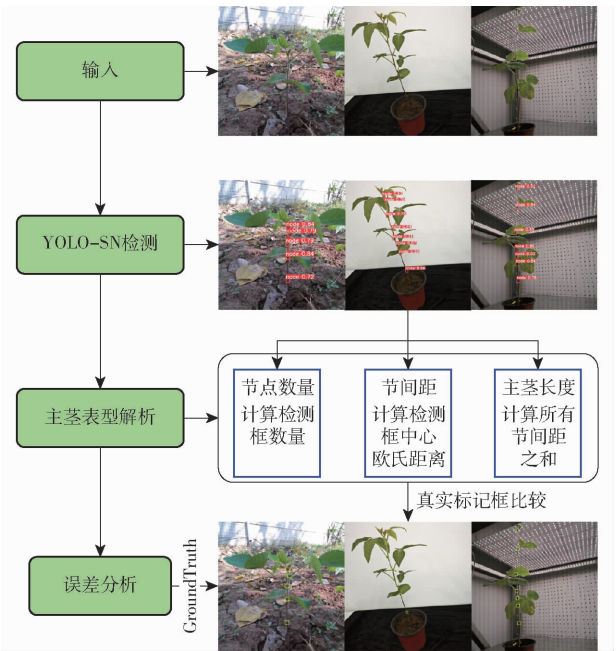


图 6 主茎表型解析流程

Fig.6 Main stem phenotype analysis process

将改进模型和原始模型节点检测效果与实际标注情况进行误差分析,如表 5 所示,YOLO – SN 模型主茎节点数量 MAE 和 R^2 分别为 0.42、0.88,相较原始模型分别减少 0.15 和提升 0.09,表明改进后模型对于不同场景下节点识别的准确性有所提升。节间距 MAE 和 R^2 分别为 12.6 像素、0.80,相较原始

表 5 改进模型和原始模型误差分析结果
Tab.5 Error analysis of improved model and original model

模型	节点数量		节间距		主茎长度	
	MAE	R^2	MAE/ 像素	R^2	MAE/ 像素	R^2
YOLO v8n	0.57	0.79	14.3	0.78	21.2	0.81
YOLO – SN	0.42	0.88	12.6	0.80	17.4	0.82

模型分别减少 1.7 像素和提升 0.02。主茎长度 MAE 和 R^2 分别为 17.4 像素、0.82,相较原始模型分别减少 3.8 像素和提升 0.01。鉴于 3 种表型性状的提取依赖于对主茎节点的准确定位,提高主茎节点识别精度有助于显著减少表型解析误差。在大田场景下,由于复杂田间环境和周围大豆植株干扰,出现误检非目标植株主茎节点,导致节点检测数量偏大,相应的节间距和主茎长度解析误差偏大;在室内场景下背景干扰因素较少,主茎节点检测数量更为准确,表型解析结果精度更高;在人工气候室内,由于植株顶端光照的影响,部分光照下节点存在漏检,导致检测到节点数量偏少,造成节间距和主茎长度解析误差偏大。此外,随着大豆生长过程中节点形态和颜色变化,子叶节点特征相较于其余主茎节点差异较大,从而导致表型解析误差偏高。

3 结论

- (1)提出了一种基于改进 YOLO v8n 的多育种场景下大豆主茎节点检测模型 YOLO – SN。通过构建 DAF 模块、设计 DDH 头部结构以及引入损失函数 WIoU v3,提升了单一育种场景和多育种场景下大豆主茎节点检测效果。
- (2)试验结果表明,YOLO – SN 在大田场景下准确率、召回率和平均精度均值分别提升 6.2、5.5、5.1 个百分点,在室内场景下准确率、召回率和平均精度均值分别提升 4.6、2.7、1.4 个百分点,在人工气候室场景下准确率、召回率和平均精度均值分别提升 1.1、3.8、3.6 个百分点。在多场景混合训练中主茎节检测准确率、召回率、平均精度均值和 F1 分数分别达 90.6%、85.1%、90.0% 和 87.8%,均优于主流模型检测效果。
- (3)基于主茎节点检测结果对大豆节点数量、节间距和主茎长度表型进行解析。节点计数绝对误差、决定系数分别为 0.42 和 0.88,节间距计算绝对误差、决定系数分别为 12.6 像素和 0.80,大豆主茎长度绝对误差、决定系数分别为 17.4 像素和 0.82,具有较高精度,可用于大豆主茎表型的定量分析。

参 考 文 献

[1] KOLHAR S, JAGTAP J. Plant trait estimation and classification studies in plant phenotyping using machine vision—a review [J]. Information Processing in Agriculture, 2023, 10(1): 114 – 135.

[2] WU Y, WANG E, GONG W, et al. Soybean yield variations and the potential of intercropping to increase production in China [J]. Field Crops Research, 2023, 291: 108771.

[3] OKADA M, BARRAS C, TODA Y, et al. High-throughput phenotyping of soybean biomass: conventional trait estimation and novel latent feature extraction using UAV remote sensing and deep learning models[J]. Plant Phenomics, 2024, 6: 0244.

[4] JIANG Y, LI C. Convolutional neural networks for image-based high-throughput plant phenotyping: a review [J]. Plant Phenomics,2020, 2020: 4152816.

[5] ZHU R, WANG X, YAN Z, et al. Exploring soybean flower and pod variation patterns during reproductive period based on

fusion deep learning[J]. *Frontiers in Plant Science*, 2022, 13: 922030.

[6] 管贤平, 刘宽, 邱白晶, 等. 基于机载三维激光扫描的大豆冠层几何参数提取[J]. *农业工程学报*, 2019, 35(23): 96 – 103. GUAN Xianping, LIU Kuan, QIU Baijing, et al. Extraction of geometric parameters of soybean canopy by airborne 3D laser scanning[J]. *Transactions of the CSAE*, 2019, 35(23): 96 – 103. (in Chinese)

[7] LIU L, LIANG J, WANG J, et al. An improved YOLO v5-based approach to soybean phenotype information perception[J]. *Computers and Electrical Engineering*, 2023, 106: 108582.

[8] LI S, YAN Z, GUO Y, et al. SPM – IS: an auto-algorithm to acquire a mature soybean phenotype based on instance segmentation[J]. *The Crop Journal*, 2022, 10(5): 1412 – 1423.

[9] LI S, SUN Z, SANG Q, et al. Soybean reduced internode 1 determines internode length and improves grain yield at dense planting[J]. *Nature Communications*, 2023, 14(1): 7939.

[10] YAMAMOTO K, GUO W, NINOMIYA S. Node detection and internode length estimation of tomato seedlings based on image analysis and machine learning[J]. *Sensors*, 2016, 16(7): 1044.

[11] NGUYEN T T, SLAUGHTER D C, MAX N, et al. Structured light-based 3D reconstruction system for plants[J]. *Sensors*, 2015, 15(8): 18587 – 18612.

[12] BOOGAARD F P, RONGEN K S A H, KOOTSTRA G W. Robust node detection and tracking in fruit-vegetable crops using deep learning and multi-view imaging[J]. *Biosystems Engineering*, 2020, 192: 117 – 132.

[13] SUN S, LI C, CHEE P W, et al. High resolution 3D terrestrial LiDAR for cotton plant main stalk and node detection[J]. *Computers and Electronics in Agriculture*, 2021, 187: 106276.

[14] BOOGAARD F P, VAN HENTEN E J, KOOTSTRA G. The added value of 3D point clouds for digital plant phenotyping—a case study on internode length measurements in cucumber[J]. *Biosystems Engineering*, 2023, 234: 1 – 12.

[15] HU J, LI G, MO H, et al. Crop node detection and internode length estimation using an improved YOLO v5 model[J]. *Agriculture*, 2023, 13(2): 473.

[16] 翟肇裕, 张梓涵, 徐焕良, 等. YOLO 算法在动植物表型研究中应用综述[J]. *农业机械学报*, 2024, 55(11): 1 – 20. ZHAI Zhaoyu, ZHANG Zihan, XU Huanliang, et al. Review of applying YOLO family algorithms to analyze animal and plant phenotype[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(11): 1 – 20. (in Chinese)

[17] 王金鹏, 何萌, 甄乾广, 等. 基于改进 COF – YOLO v8n 的油茶果静态与动态检测计数方法[J]. *农业机械学报*, 2024, 55(4): 193 – 203. WANG Jinpeng, HE Meng, ZHEN Qianguang, et al. *Camellia oleifera* fruit static and dynamic detection counting based on improved COF – YOLO v8n[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(4): 193 – 203. (in Chinese)

[18] XIA Z, PAN X, SONG S, et al. Vision transformer with deformable attention[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 4794 – 4803.

[19] DAI X, CHEN Y, XIAO B, et al. Dynamic head: unifying object detection heads with attentions[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 7373 – 7382.

[20] ZHENG Z, WANG P, LIU W, et al. Distance – IoU loss: faster and better learning for bounding box regression[C] // *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020: 12993 – 13000.

[21] TONG Z, CHEN Y, XU Z, et al. Wise – IoU: bounding box regression loss with dynamic focusing mechanism[J]. *arXiv Preprint arXiv:2301.10051*, 2023.

[22] YANG Q, LIN G, LV H, et al. Environmental and genetic regulation of plant height in soybean[J]. *BMC Plant Biology*, 2021, 21: 1 – 15.

[23] ALLEN JR L H, ZHANG L, BOOTE K J, et al. Elevated temperature intensity, timing, and duration of exposure affect soybean internode elongation, mainstem node number, and pod number per plant[J]. *The Crop Journal*, 2018, 6(2): 148 – 161.

[24] SHAN F, SUN K, GONG S, et al. Effects of shading on the internode critical for soybean (*Glycine max*) lodging[J]. *Agronomy*, 2022, 12(2): 492.

[25] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 658 – 666.

[26] SUN H, XU H, LIU B, et al. MEAN – SSD: a novel real-time detector for apple leaf diseases using improved light-weight convolutional neural networks[J]. *Computers and Electronics in Agriculture*, 2021, 189: 106379.

[27] LI Z, LI Y, YANG Y, et al. A high-precision detection method of hydroponic lettuce seedlings status based on improved Faster RCNN[J]. *Computers and Electronics in Agriculture*, 2021, 182: 106054.

[28] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLO v7: 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*, 2023: 7464 – 7475.

[29] WANG C Y, YE H I H, MARK L H Y. YOLO v9: learning what you want to learn using programmable gradient information [C] // *European Conference on Computer Vision*, 2025: 1 – 21.

[30] WANG A, CHEN H, LIU L, et al. YOLO v10: real-time end-to-end object detection[J]. *Advances in Neural Information Processing Systems*, 2025, 37: 107984 – 108011.