

胡健威, 马慧敏, 宁孝梅, 等. 基于改进 YOLOv8 的无人机图像玉米幼苗检测[J]. 江苏农业学报, 2025, 41(6): 1179-1187.
doi: 10.3969/j.issn.1000-4440.2025.06.014

基于改进 YOLOv8 的无人机图像玉米幼苗检测

胡健威, 马慧敏, 宁孝梅, 代腾辉, 戴明宇, 王小申, 吴 琦
(安徽农业大学信息与人工智能学院, 安徽 合肥 230036)

摘要: 无人机技术凭借其高效、精准的优势, 在农业领域发挥着重要作用, 被广泛应用于农田监测、精准施肥、病虫害防治等环节。然而, 在无人机采集的玉米幼苗图像中, 田间杂草等干扰物与玉米幼苗颜色相近, 易导致 YOLOv8 模型误检和漏检; 同时玉米幼苗间相互遮挡, 也会影响模型检测精度。针对这些问题, 本研究提出了 YOLOv8+MultiSEAM+MetaNeXtStag+WIoU 模型(简称 YOLOv8-MMW 模型)。该模型基于 YOLOv8 模型架构, 首先在颈部网络引入 MultiSEAM 注意力机制, 有效提升了模型在复杂场景下的特征提取能力; 其次引入了 InceptionNeXt 中的 MetaNeXtStage 模块; 在此基础上, 采用 Wise-IoU 损失函数以提升模型精度。在测试集上的试验结果表明, YOLOv8-MMW 模型精度和交并比阈值为 0.50 时的平均精度均值(mAP_{50})分别达到 98.9% 和 89.6%, 较原始 YOLOv8n 模型分别提升了 6.1 个百分点和 2.4 个百分点。本研究提出的 YOLOv8-MMW 模型在复杂农田环境下表现出更强的鲁棒性, 能够有效提升对无人机拍摄的玉米幼苗图像的检测准确率, 为农业管理和监测提供了技术支持。

关键词: 玉米幼苗; 无人机; YOLOv8; MultiSEAM; MetaNeXtStage; 损失函数

中图分类号: S513 **文献标识码:** A **文章编号:** 1000-4440(2025)06-1179-09

Corn seedling detection in unmanned aerial vehicle images based on improved YOLOv8

HU Jianwei, MA Huiming, NING Xiaomei, DAI Tenghui, DAI Mingyu, WANG Xiaoshen, WU Yi
(School of Information and Artificial Intelligence, Anhui University of Agriculture, Hefei 230036, China)

Abstract: Unmanned aerial vehicle technology, with its high efficiency and precision, has been playing an important role in the agricultural field and has been widely applied in farmland monitoring, precision fertilization, and pest and disease control. However, in the corn seedling images collected by unmanned aerial vehicle, weeds and other interferents in the field have colors similar to corn seedlings, which can easily lead to false detection and missed detection in the YOLOv8 model. Additionally, the occlusion between corn seedlings can also affect the detection accuracy of the model. In response to these problems, this study proposed the YOLOv8+MultiSEAM+MetaNeXtStag+WIoU (YOLOv8-MMW) model. Based on the YOLOv8 model architecture, this model first introduced the MultiSEAM attention mechanism into the neck network, which effectively enhanced the model's feature extraction ability in complex scenes. Secondly, it incorporated the MetaNeXtStage module from InceptionNeXt. On this basis, the Wise-IoU loss function was adopted to improve model accuracy. The experimental results on the test set showed that the accuracy and the mean average precision at an intersection-over-union threshold of 0.50 (mAP_{50}) of the YOLOv8-MMW model

reach 98.9% and 89.6%, respectively, which were 6.1 percentage points and 2.4 percentage points higher than those of the original YOLOv8n model. The YOLOv8-MMW model proposed in this study demonstrates stronger robustness in complex farmland environments and can effectively improve the detection accuracy of corn seedling images captured by

收稿日期: 2024-11-09

基金项目: 安徽省高校优秀青年人才支持项目(gxyq2022004); 安徽省研究生创新创业项目(2022cxycysj065)

作者简介: 胡健威(1998-), 男, 安徽合肥人, 硕士研究生, 研究方向为作物信息智能感知。(E-mail)jwhu@stu.ahau.edu.cn

通讯作者: 马慧敏, (E-mail)huiminma@ahau.edu.cn

unmanned aerial vehicle, and can provide technical support for agricultural management and monitoring.

Key words: corn seedling; unmanned aerial vehicle; YOLOv8; MultiSEAM; MetaNeXtStage; loss function

玉米是中国种植面积^[1]和产量均居首位的作物^[2],也是重要的经济作物,对国家粮食安全具有重大战略意义^[3]。近年来,中国的玉米生产取得长足发展,但全球玉米生产仍面临诸多挑战。随着全球人口快速增长和老龄化加剧,粮食需求持续增加。联合国粮食和农业组织预测,2050 年全球人口将增至 9.1×10^9 人,为满足人口增长需求,在保障自然资源安全的前提下,粮食产量需实现 70% 的可持续增长^[4]。因此,建立高效的农业系统以提高农业生产效率和质量、降低劳动力投入、改善工作条件势在必行。

卷积神经网络(CNN)作为深度学习的核心技术,具有自动提取图像特征的能力,对位移、缩放和畸变等几何变换具有较强的自适应能力,该技术已被广泛用于计算机视觉领域,包括人脸识别、遥感影像分析和无人驾驶系统等^[5-7]。近年来,卷积神经网络在农业领域的应用研究也取得了显著进展,成为智慧农业发展的重要技术支撑^[8-11]。Malik 等^[12]提出一种深度卷积神经网络模型 TreeNet,该模型的分类准确率、精确度、召回率和 F1 得分分别达到 99.8%、99.0%、100.0%和 99.0%。Liu 等^[13]提出目标检测算法 Agromomic Teacher,用于玉米叶片病害识别,与基线模型相比,该算法在 MaizeData 数据集上的平均精度均值(mAP)提高 6.5 个百分点,在通用目标检测基准 PascalVOC 数据集上的 mAP 提高 8.2 个百分点,表现出良好的泛化能力。李金瑞等^[14]提出一种基于 RTM-Det-tiny 的轻量级雄穗检测模型 MLCE-RTMDet。试验结果显示,改进的 MLCE-RTMDet 模型参数量减少到 3.9×10^6 ,浮点运算数减少到 5.3×10^9 ,测试集上模型平均精度均值(mAP)达到 92.2%,比原模型提高了 1.2 个百分点。Puyenbroeck 等^[15]利用推帚式高光谱成像系统对收获期玉米进行扫描,建立基于偏最小二乘判别分析(PLS-DA)的像素级分类模型,实现了玉米籽粒与其他谷物的精准识别,分类准确率达 95.2%。Mei 等^[16]提出分子光谱深度融合策略,通过结合低级特征融合与中级特征融合,并构建一维卷积神经网络(1D-CNN)模型,实现了玉米油中毒死蜱残留的快速精准检测。

无人机遥感技术的发展推动了农业智能监测系统的完善^[17]。Abel 等^[18]基于无人机影像和全卷积

神经网络,实现了对农作物数量的自动预测。Chen 等^[19]针对小规模玉米穗和强光照条件下的检测问题,开发了 RESAM-YOLOv8n 模型,通过增大输入图像尺寸增强网络检测能力。Xiao 等^[20]整合无人机、地理信息系统(GIS)和深度学习技术,利用 YOLOv5 模型完成对不同管理模式玉米植株的计数。有学者基于无人机光学遥感技术,提出一种深度学习方法,能够有效提取多光谱图像的关键特征,实现了南方玉米锈病的高效监测^[21]。邬开俊等^[22]提出了一种基于 YOLO-TSCAS 模型的玉米植株中心检测算法,试验结果表明,与现有的 YOLO_x 算法相比,该算法的正确率提高 22.73%,F1 分数提高 0.255,平均对数漏检率减少 0.35。Ni 等^[23]设计的玉米混合视觉转换器(MaizeHT)模型,对 224×224 分辨率图像的识别准确率达 97.71%,对 512×512 分辨率图像的识别准确率达到 98.71%。

上述研究结果表明,基于深度学习的模型能够有效识别无人机获取的玉米幼苗图像。但在田间环境下,田间杂草干扰、玉米幼苗间相互遮挡均会影响模型检测精度。针对这些问题,本研究拟提出一种基于 YOLOv8 模型的玉米幼苗检测模型,主要从以下 3 个方面进行优化:首先,重构主干网络架构,采用多尺度特征融合策略增强特征表达能力;其次,嵌入混合注意力机制,提升模型对幼苗关键特征的关注度;最后,改进边界框回归损失函数,提高模型精度。本研究结果旨在为农业管理和监测提供技术支持。

1 材料与方 法

1.1 数据采集和预处理

以幼苗期玉米为研究对象,采用如图 1 所示的大疆无人机 DJI Air 2S 进行图像采集。该设备具有体积小、操作便捷等优点,主要技术参数如下:最大续航时间 31 min,最快飞行速度 19 m/s,悬停精度垂直 ± 0.1 m,水平 ± 0.1 m,最大照片分辨率 5472×3648 ,感光度(ISO)为 100~3200。试验于 2024 年 6 月 25 日(9:30-11:00)在安徽省合肥市安徽农业大学国家高新技术农业园(东经 $117^{\circ}12'30.927''$,北纬 $31^{\circ}56'10.859''$)开展。如图 2 所示,采集过程中,设置图像分辨率为 3840×2160 ,分别在飞行高度 2

m、4 m、6 m、8 m 拍摄图像。根据大田地地形规划航线,共获取不同高度图像 500 余张。



图 1 大疆无人机

Fig.1 DJI unmanned aerial vehicle

本研究对采集的无人机图像进行了系统的预处理,包括数据标注、数据增强和数据集划分 3 个关键步骤。在数据标注环节,采用 LabelImg 工具对玉米

幼苗进行精确标注,确保每株幼苗的边界框选准确无误,同时保持不同拍摄条件和光照环境下标注标准的一致性。在数据增强环节,如图 3 所示,通过 $\pm 20^\circ$ 随机旋转和镜像翻转图像实现几何变换;在图像中添加高斯噪声和椒盐噪声增强模型抗干扰能力;随机调整图像亮度和对比度模拟不同光照条件;通过随机裁剪和缩放图像提升模型空间适应性。这些增强技术的组合应用使训练集样本量扩充至原始数据的 5 倍,显著提升了数据多样性。在数据集划分环节,从所有图像中随机选取 1 920 张图像,构建训练集;同时分别从 2 m、4 m、6 m、8 m 飞行高度拍摄的图像中抽取 60 张图像,共计 240 张图像,构建测试集,用于检验模型评估的全面性和可靠性。

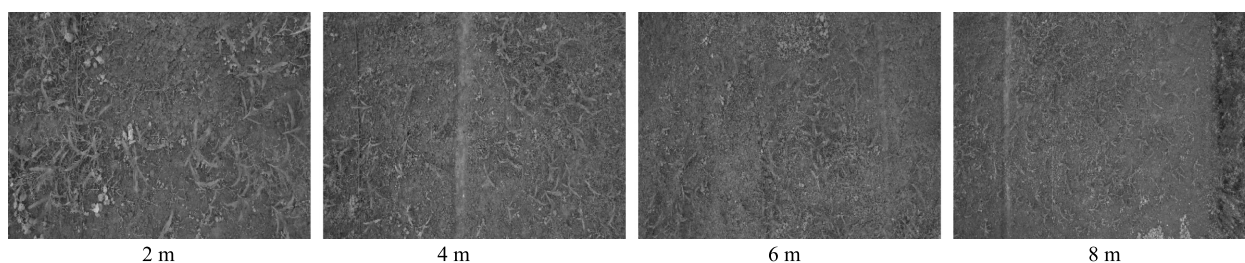


图 2 不同高度无人机拍摄的玉米幼苗图像

Fig.2 Corn seedling images captured by unmanned aerial vehicle at different altitudes

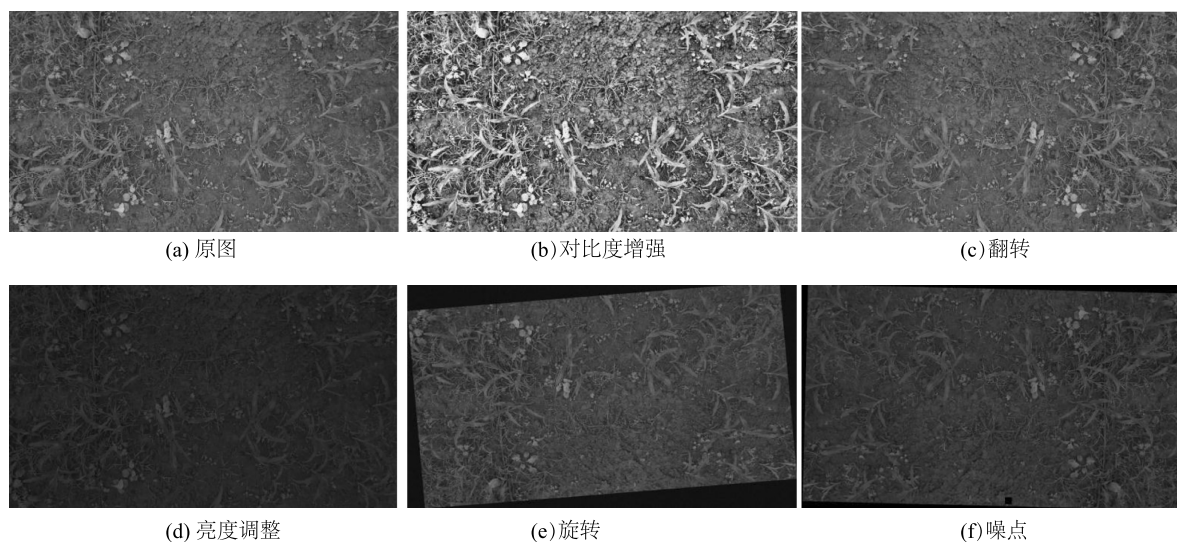


图 3 数据增强方法

Fig.3 Data augmentation methods

1.2 YOLOv8+MultiSEAM+MetaNeXtStag+WIoU 模型的构建

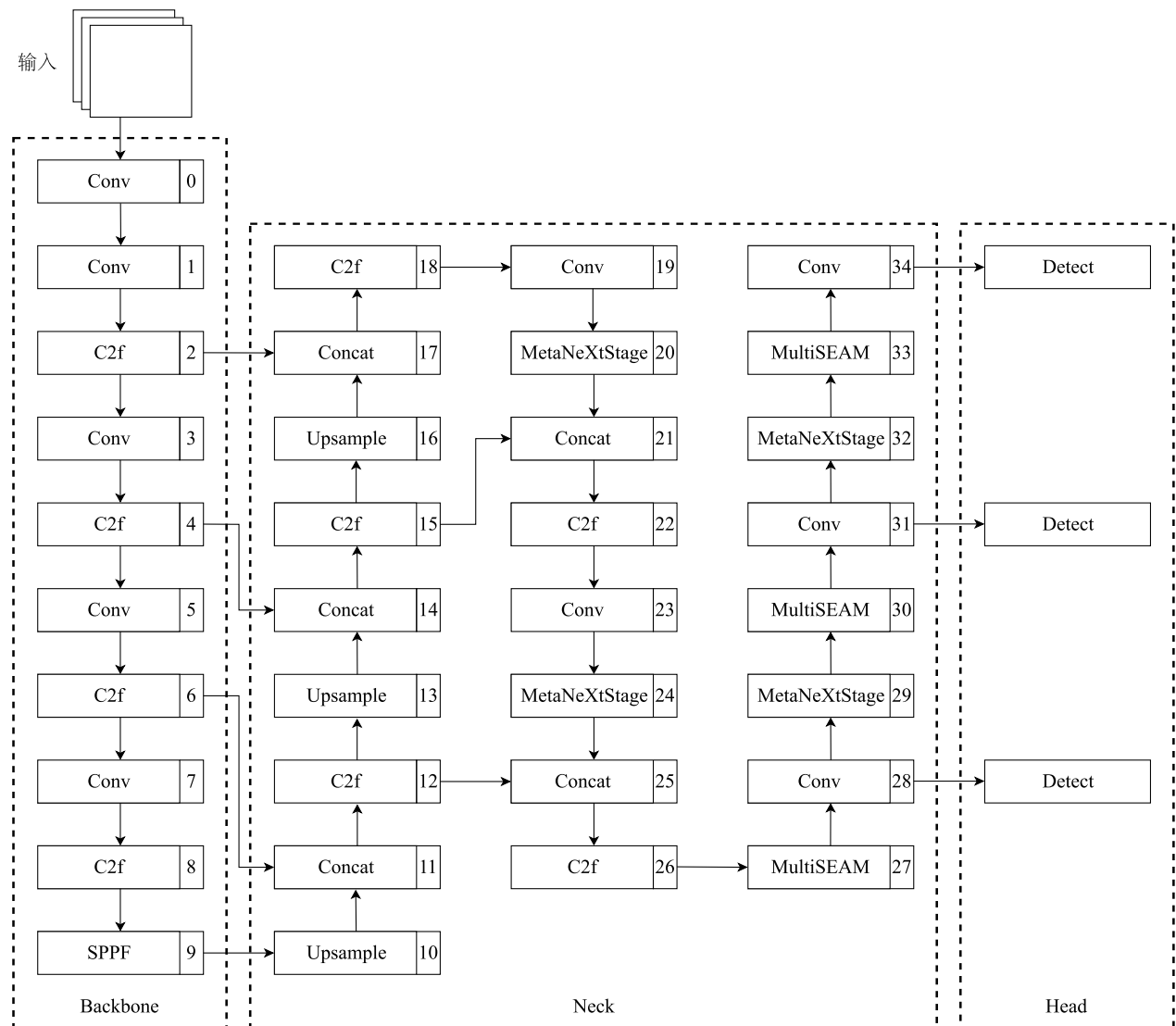
YOLOv8 模型是一种基于单阶段检测框架的实

时目标检测算法,可以直接预测图像中的边界框和类别概率。根据网络深度和宽度的不同配置,YOLOv8 被分为 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l 和

YOLOv8x 5 个版本。本研究选用 YOLOv8n 作为基准模型进行改进。YOLOv8 网络架构由输入模块、主干网络(Backbone)、颈部网络(Neck)和检测头(Head)4 个核心组件构成^[24]。其中,主干网络包含卷积单元(Conv)、空间金字塔池化快速模块(SPPF)、YOLOv5 中的特征提取模块 C3 和 YOLOv7 中的特征提取模块 C2f。颈部网络采用跨阶段局部网络(CSPNet)架构以增强特征融合能力,检测头则通过路径聚合网络与预测头实现特征融合与目标识别。

YOLOv8 模型对玉米幼苗图像的检测存在以下问题:(1)田间杂草等干扰物与玉米幼苗颜色相近,

导致模型误检、漏检;(2)玉米幼苗间相互遮挡,造成模型检测精度下降。基于上述问题,本研究对 YOLOv8 网络结构进行了改进。如图 4 所示,在颈部网络中引入 MultiSEAM 模块和 MetaNeXtStage 模块。MultiSEAM 模块部署于第 27 层、30 层、33 层的输出端,即最终检测层前,用于提升模型的多尺度特征处理能力。MetaNeXtStage 模块则被嵌入第 20 层、24 层、29 层、32 层的卷积层之后,用于增强模型特征表示能力。检测层采用第 28 层、31 层和 34 层的特征图作为输入,有效整合了 MultiSEAM 和 MetaNeXtStage 模块的多尺度特征增强效果。



Conv:卷积层;C2f:跨阶段特征融合模块;SPPF:空间金字塔池化模块;Upsample:上采样模块;Concat:拼接模块;Detect:检测层;Backbone:骨干网络;Neck:颈部;Head:检测头;MetaNeXtStage:一种模块;MultiSEAM:注意力机制。

图 4 YOLOv8-MMW 模型结构

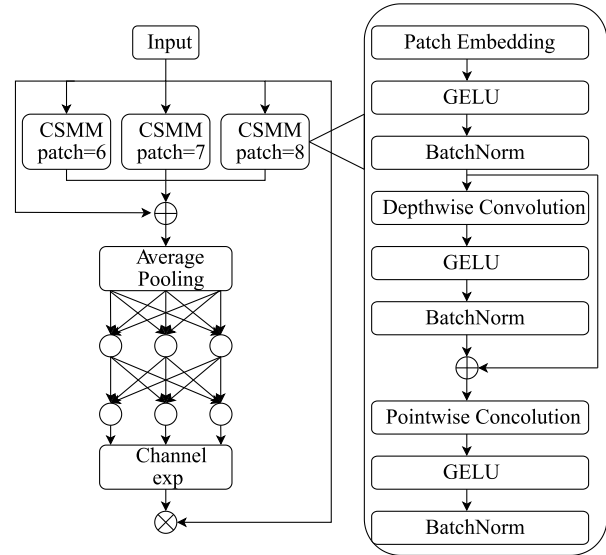
Fig.4 Architecture of the YOLOv8-MMW model

1.2.1 MultiSEAM 注意力机制 挤压和激励注意力模块(SEAM)^[25]通过全局平均池化和全连接层增强特征表达,采用深度可分离卷积与残差连接的组合结构,其中深度可分离卷积按通道进行操作。该模块利用注意力机制突出重要特征,抑制无关特征,从而提升模型识别能力。如图 5 所示,SEAM 模块左侧为整体架构,包含 3 个不同像素块尺寸(patch-6、patch-7、patch-8)的通道和空间混合模块(CSMM),对各模块输出进行平均池化和通道扩展操作,通过特征相乘方式实现特征增强;右侧为 CSMM 模块,该模块通过多尺度特征提取,并采用深度可分离卷积学习空间与通道之间的相关性。

MultiSEAM 模块将小目标检测层集成到颈部网络,生成多尺度特征图,并将这些特征图输入 SEAM 模块进行增强处理。该设计具有以下优势:整合全局与局部上下文信息,提升复杂场景下的特征表达能力;采用多尺度处理策略,增强模型对不同尺寸目标的检测能力。在玉米幼苗检测过程中,MultiSEAM 模块能够有效应对复杂背景和密集目标。

1.2.2 InceptionNeXt 卷积神经网络架构 如图 6 所示,InceptionNeXt^[26]是一种卷积神经网络架构,通过融合 Inception 的多分支结构和 ConvNeXt 的高效设计,实现了对大核深度卷积的优化分解。该架构继承了 Inception 的多尺度特征处理能力,同时采用高效卷积操作,在保持大感受野优势的前提下,显著降低了计算复杂度和参数量。此外,InceptionNeXt 还集成了自适应优化技术,使模型在训练过程中更快收敛。本研究采用 InceptionNeXt 的核心模块 MetaNeXtStage,该模块能够动态调整卷积核参数,增强模型对局部细节的敏感性;自适应优化特征图权重,突出全局关键信息;通过多尺度特征的高效融合,实现性能与计算效率的平衡。

1.2.3 基于动态非单调聚焦机制的边界框损失函数(Wise-IoU) 为提升目标检测模型的性能,本研究系统性地探究了不同边界框回归损失函数对模型性能的影响。原始 YOLOv8 模型采用的完全交并比损失函数(Complete-IoU)通过引入长宽比惩罚项来优化预测框的形状匹配,在模型训练过程中表现出良好的拟合特性。然而,该损失函数在计算预测框宽高比时存在固有缺陷:一方面,宽高比作为相对量计算时具有不确定性;另一方面,其对高质量和低质量锚框的不同影响可能阻碍回归损失的优化过程。



Patch-6:将图像分割成 6×6 像素的块;Patch-7:将图像分割成 7×7 像素的块;Patch-8:将图像分割成 8×8 像素的块。Input:输入;CSMM:通道和空间混合模块;Average Pooling:平均池化层;Channel exp:通道扩展;Patch Embedding:图像块嵌入;GELU:高斯误差线性单元激活函数;BatchNorm:批量归一化;Depthwise Convolution:深度可分离卷积模块;Pointwise Convolution:逐点卷积模块。

图 5 MultiSEAM 注意力机制

Fig.5 The MultiSEAM attention mechanism

为解决上述问题,本研究对以下 4 种边界框回归损失函数的性能进行对比:基于动态非单调聚焦机制的边界框损失函数(WIoU)^[27]、完全交并比损失函数(CIoU)、高效交并比损失函数(EIoU)^[28]、距离交并比损失函数(DIoU)^[29]。WIoU(v3 版本)损失函数计算公式如下:

$$L_{\text{WIoU}} = r \cdot R_{\text{WIoU}} \cdot L_{\text{IoU}} \quad (1)$$

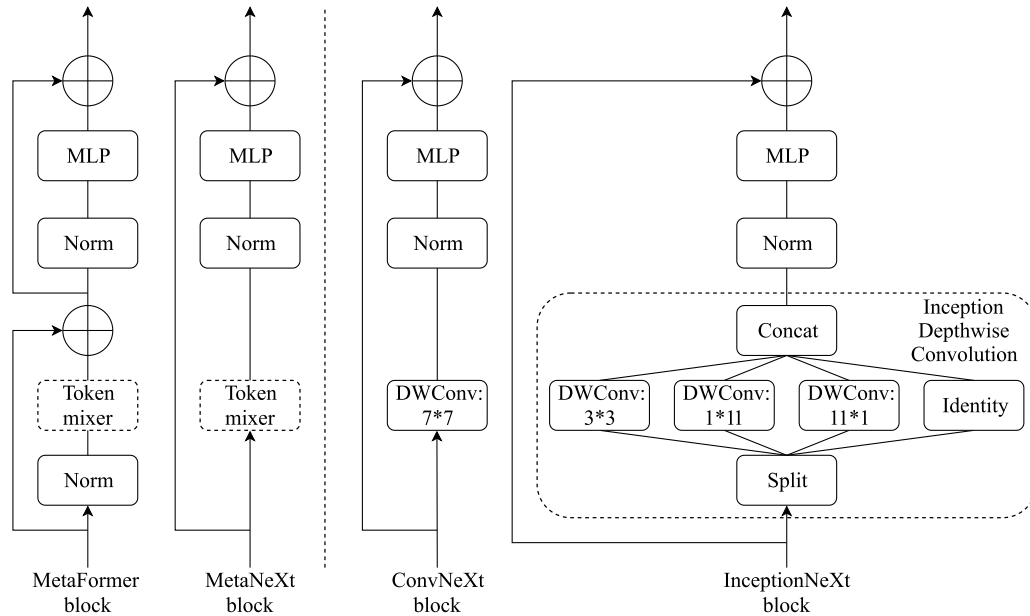
$$R_{\text{WIoU}} = \exp \left[\frac{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2}{(c_w^2 + c_h^2)} \right] \quad (2)$$

$$r = \frac{\beta}{\delta \alpha^{\beta - \delta}} \quad (3)$$

$$\beta = \frac{L_{\text{IoU}}^*}{L_{\text{IoU}}} \in [0, +\infty) \quad (4)$$

式中, L_{WIoU} 为基于动态非单调聚焦机制的边界框损失函数; r 为缩放因子; R_{WIoU} 为距离损失; L_{IoU} 为原始交并比损失函数; $b_{c_x}^{gt}$, $b_{c_y}^{gt}$ 为目标框的中心点坐标; b_{c_x} , b_{c_y} 为预测框的中心点坐标; c_w 为目标框在 x 轴方向的归一化宽度; c_h 为目标框在 y 轴方向的归一化高度; β 为比例系数; δ 和 α 为调整参数; L_{IoU}^* 为

优化后的交并比损失函数。



MLP:多层感知机;Norm:归一化操作;Token mixer:令牌混合器模块;Concat:拼接操作;DWConv:深度可分离卷积模块;Split:分割操作;Depthwise Convolution:深度可分离卷积模块;MetaFormer block:通用视觉模型架构;MetaNeXt:MetaNeXt 卷积神经网络架构;ConvNeXt block:ConvNeXt 卷积神经网络架构;InceptionNeXt block:InceptionNeXt 卷积神经网络架构;Inception Depthwise Convolution:Inception 深度卷积;Identity:恒等映射。

图 6 InceptionNeXt 卷积神经网络

Fig.6 InceptionNeXt convolutional neural network

1.3 模型性能指标

为评估无人机玉米幼苗检测模型的性能,以精度(P)、召回率(R)、平均精度均值(mAP)作为评估指标。

精度(P)计算公式如下:

$$P = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

式中, TP 为正确预测的正例数量, FP 为误判为正例的数量。精度越高,表明模型误检率越低。

召回率(R)计算公式如下:

$$R = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

式中, FN 为漏检的正例数量。召回率越高,表明模型漏检率越低。

平均精度均值(mAP)计算公式如下:

$$mAP = \frac{1}{N} \cdot \sum_{i=0}^N AP_i \quad (7)$$

式中, N 为类别总数, AP_i 为第 i 个类别的平均精度。 mAP 综合了模型对所有类别的检测性能,其值越高,表明模型的整体检测能力越优。

1.4 试验环境

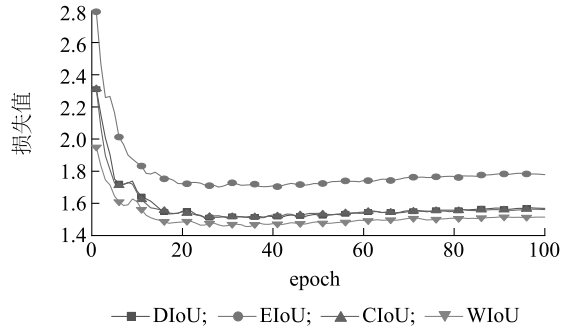
硬件环境为 Intel Xeon Silver 4210 2.40 GHz CPU、RTX 3080 10 G GPU 和 64 GB 内存。软件环境为 PyTorch 1.9.0 和 Python 3.8.0。试验设置如下:迭代次数为 100 次,批次大小为 2,输入图像尺寸调整为 640×640 ,并进行归一化处理。优化算法采用随机梯度下降法(SGD)^[30],学习率调度策略为线性衰减。

2 结果与分析

2.1 Wise-IoU 损失函数对模型性能的影响

如图 7 所示,引入 EIoU 损失函数的模型收敛速度最慢,且最终的损失值最大;引入 DIoU 损失函数和 CIoU 损失函数的模型收敛速度低于引入 EIoU 损失函数的模型,最终损失值小于引入 EIoU 函数的模型。值得注意的是,引入 WIoU 损失函数的模型在前 20 轮迭代中能快速收敛,其最终损失值显著低于引入其他 3 种损失函数的模型。如表 1 所示,引入 WIoU 损失函数不仅加快了模型的收敛速度,还

显著提升了模型的关键性能指标:精度提高 6.9 个百分点,交并比阈值为 0.50 时的平均精度均值 (mAP_{50}) 提高 0.2 个百分点。



DIoU:距离交并比损失函数;EIoU:高效交并比损失函数;CIoU:完全交并比损失函数;WIoU:基于动态非单调聚焦机制的边界框损失函数;epoch:训练的轮数。

图7 各损失函数收敛速度

Fig.7 The convergence speed of each loss function

2.2 引入的各模块对模型性能的影响

为验证各模块对模型性能的影响,基于原始模型 YOLOv8n,分别引入 MultiSEAM、MetaNeXtStag、WIoU。如表 1 所示, YOLOv8 模型的 mAP_{50} 为 87.2%。在 YOLOv8 主干网络的第 27 层、30 层、33 层的输出端和最终检测层前层部署 MultiSEAM 注意力机制, YOLOv8n + MultiSEAM 模型 mAP_{50} 较 YOLOv8 模型提高了 2.0 个百分点;进一步在主干网络的第 20 层、24 层、29 层、32 层的卷积层之后嵌入 MetaNeXtStage 模块, YOLOv8n+MultiSEAM+MetaNeXtStag 模型 mAP_{50} 较 YOLOv8n 模型提高了 2.2 个百分点。表明 MetaNeXtStage 和 MultiSEAM 模块能够有效增强模型对复杂场景中多尺度特征的提取能力。将边界框损失函数替换为 WIoU 后, YOLOv8n+MultiSEAM + MetaNeXtStag + WIoU 模型(简称 YOLOv8-MMW 模型)的 mAP_{50} 达到 89.6%, 较 YOLOv8n 模型提升 2.4 个百分点;精度达到 98.9%, 较 YOLOv8 模型提升 6.1 个百分点。消融试验结果表明,本研究提出的改进策略显著提升了 YOLOv8n 模型的检测性能。

2.3 模型性能对比

基于测试集对 YOLOv8-MMW 模型的性能进行检测。如表 2 所示, YOLOv8-MMW 模型对不同飞行高度的无人机图像均表现出优异的检测性能,当飞行高度为 8 m 时, YOLOv8-MMW 模型的 mAP_{50} 仍

可以保持在 89.4%, 召回率达到 95.0%, 展现出良好的尺度适应性和检测稳定性。

表 1 消融试验结果

Table 1 Ablation study results

模型	mAP_{50} (%)	R (%)	P (%)
YOLOv8n	87.2	94.0	92.8
YOLOv8n+MultiSEAM	89.2	95.0	92.2
YOLOv8n+MetaNeXtStag	87.7	93.0	92.0
YOLOv8n+WIoU	87.8	94.0	93.0
YOLOv8n+MultiSEAM+MetaNeXtStag	89.4	95.0	92.0
YOLOv8n+MultiSEAM+WIoU	88.4	95.0	93.4
YOLOv8n+MetaNeXtStag+WIoU	87.7	93.0	92.9
YOLOv8n+MultiSEAM+MetaNeXtStag+WIoU	89.6	95.0	98.9

mAP_{50} :交并比阈值为 0.50 时的平均精度均值; R :召回率; P :精度;WIoU:基于动态非单调聚焦机制的边界框损失函数。

表 2 YOLOv8n+MultiSEAM+MetaNeXtStag+WIoU 模型对不同高度图像的检测性能

Table 2 Detection performance of the YOLOv8n+MultiSEAM+MetaNeXtStag+WIoU model on images captured at different heights

高度 (m)	mAP_{50} (%)	R (%)	P (%)
2	91.6	98.0	98.9
4	91.0	96.0	95.0
6	90.3	94.0	92.9
8	89.4	95.0	89.6

mAP_{50} :交并比阈值为 0.50 时的平均精度均值; R :召回率; P :精度。

将目标检测模型 YOLOv8s 模型、YOLOv5s 模型和 YOLOv8-MMW 模型进行对比。如表 3 所示,与 YOLOv8s 模型相比, YOLOv8-MMW 模型 mAP_{50} 提升 0.8 个百分点,精度提升 9.3 个百分点;与 YOLOv5s 模型相比, YOLOv8-MMW 模型 mAP_{50} 提升 3.6 个百分点,精度提升 10.1 个百分点。

表 3 不同模型性能对比

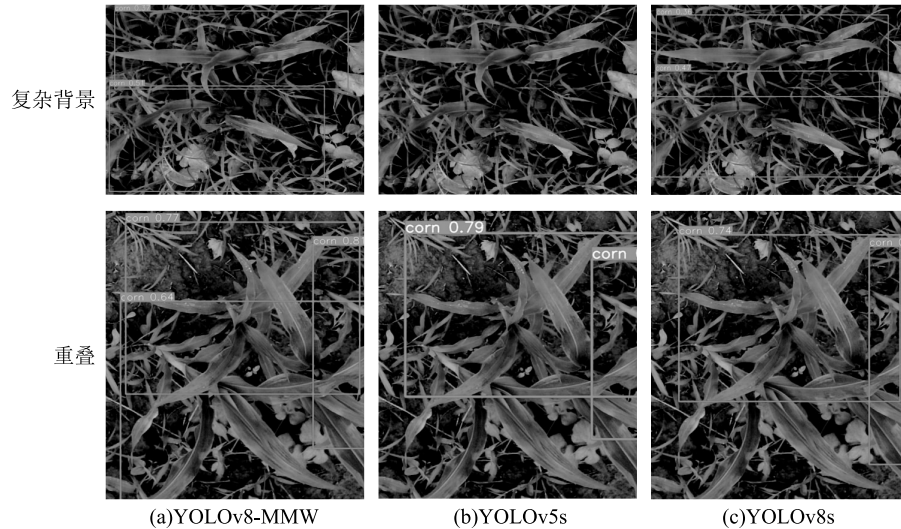
Table 3 Comparison of performance among different models

模型	mAP_{50} (%)	R (%)	P (%)
YOLOv8-MMW	89.6	95.0	98.9
YOLOv8s	88.8	95.0	89.6
YOLOv5s	86.0	93.0	88.8

YOLOv8s、YOLOv5s:当前主流的目标检测网络;YOLOv8-MMW:本研究改进的模型; mAP_{50} :交并比阈值为 0.50 时的平均精度均值; R :召回率; P :精度。

如图 8 所示,总计目标数为 22 213 个,在幼苗重叠和杂草干扰等复杂场景下, YOLOv8s 模型和 YOLOv5s 模型分别漏检了 1 368 个、1 609 个目标,漏检率分别为 6.2%、7.2%,而 YOLOv8-MMW 模型仅漏检了 321 个目标,漏检率为 1.4%。表明 Multi-

SEAM 和 MetaNeXtStage 模块的协同作用不仅能够有效增强目标特征的显著性,还能够抑制复杂背景的干扰。相较于 YOLOv8s 模型和 YOLOv5s 模型, YOLOv8-MMW 模型在复杂场景下展现出更优异的检测性能。



YOLOv8s、YOLOv5s;当前主流的目标检测网络;YOLOv8-MMW;本研究改进的模型;corn:玉米;实线框:正确识别结果。

图 8 不同模型检测结果对比

Fig.8 Comparison of detection results among different models

3 讨论与结论

本研究针对无人机检测玉米幼苗提出了 YOLOv8-MMW 模型。首先,基于 YOLOv8n 原始模型,在模型颈部网络中引入 MultiSEAM 注意力机制,该机制能够在高效捕捉细节的同时,提供更强大的特征表达能力,从而显著提升模型在复杂场景下的表现;其次,引入 InceptionNeXt 网络的 MetaNeXt-Stage 模块;在此基础上,采用 WIoU 损失函数以加快模型收敛速度和精度。与 YOLOv8n 模型相比, YOLOv8-MMW 模型精度和 mAP_{50} 分别提升了 6.1 个百分点、2.4 个百分点。该模型显著提升了复杂背景下玉米幼苗的检测精度和效率,为大田玉米幼苗计数及产量估算提供了有力支持。

未来的研究方向将聚焦于进一步提升模型的检测性能并降低其复杂度,通过轻量化处理优化模型结构,同时深入探索该算法在不同环境条件下的适应性。

参考文献:

[1] 仇焕广,栾昊,李瑾,等. 风险规避对农户化肥过量施用行

为的影响[J]. 中国农村经济,2014(3):85-96.

- [2] 徐文姬,顾道健,张博博,等. 玉米果穗不同部位籽粒激素含量及其与胚乳发育和籽粒灌浆的关系[J]. 作物学报,2013,39(8):1452-1461.
- [3] 赵久然,王荣焕. 中国玉米生产发展历程、存在问题及对策[J]. 中国农业科技导报,2013(3):1-6.
- [4] 陈源源,吕昌河,尚凯丽. 食品安全的内涵、指标与评价方法综述[J]. 中国农学通报,2017,33(22):158-164.
- [5] WANG Y, XIE F, ZHAO C X, et al. Robust face recognition model based sample mining and loss functions[J]. Knowledge-Based Systems,2024,302:112330.
- [6] ZHANG Y S, YAN S K, ZHANG L F, et al. Fast projected fuzzy clustering with anchor guidance for multimodal remote sensing imagery[J]. IEEE transactions on image processing, 2024(33), 4640-4653.
- [7] QIU C Q, TANG H, YANG Y C, et al. Machine vision-based autonomous road hazard avoidance system for self-driving vehicles[J]. Scientific Reports,2024,14(1):12178.
- [8] JIANG H K, CHEN Q C, WANG R J, et al. SWFormer:a scale-wise hybrid CNN-Transformer network for multi-classes weed segmentation[J]. Journal of King Saud University-Computer and Information Sciences,2024,36(7):102144.
- [9] MUQADDAS S, QURESHI W S, JABBAR H, et al. A comprehensive deep learning approach for harvest ready sugarcane pixel

- classification in Punjab, Pakistan using Sentinel-2 multispectral imagery[J]. *Remote Sensing Applications: Society and Environment*, 2024, 35: 101225.
- [10] 周善良, 李锐. 基于卷积神经网络的农作物病虫害识别研究综述[J]. *智慧农业导刊*, 2024, 4(17): 39-45.
- [11] CHEN F, JI X Z, BAI M X, et al. Network analysis of different exogenous hormones on the regulation of deep sowing tolerance in maize seedlings[J]. *Frontiers in Plant Science*, 2021, 12: 739101.
- [12] MALIK M M, FAYYAZ A M, YASMIN M, et al. A novel deep CNN model with entropy coded sine cosine for corn disease classification[J]. *Journal of King Saud University-Computer and Information Sciences*, 2024, 36(7): 102126.
- [13] LIU J Q, HU Y X, SU Q F, et al. Semi-supervised one-stage object detection for maize leaf disease[J]. *Agriculture*, 2024, 14(7): 1140.
- [14] 李金瑞, 杜建军, 张宏鸣, 等. 基于轻量化 MLCE-RTMDet 的人工去雄后玉米雄穗检测算法[J]. *农业机械学报*, 2024, 55(11): 1-14.
- [15] PUYENBROECK E V, WOUTERS N, LEBLICQ T, et al. Detection of kernels in maize forage using hyperspectral imaging[J]. *Computers and Electronics in Agriculture*, 2024, 225: 109336.
- [16] MEI C L, XUE Y C, LI Q H, et al. Deep learning model based on molecular spectra to determine chlorpyrifos residues in corn oil[J]. *Infrared Physics and Technology*, 2024, 140: 105402.
- [17] FAN B K, LI Y, ZHANG R Y, et al. Review on the technological development and application of UAV systems[J]. *Chinese Journal of Electronics*, 2020, 29(2): 5-13.
- [18] ABEL B, PHILIPP L, RAMÓN I Y F, et al. Automatic UAV-based counting of seedlings in sugar-beet field and extension to maize and strawberry[J]. *Computers and Electronics in Agriculture*, 2021, 191: 106493.
- [19] CHEN J H, FU Y S, GUO Y H, et al. An improved deep learning approach for detection of maize tassels using UAV-based RGB images[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2024, 130: 103922.
- [20] XIAO J, SUAB S A, CHEN X Y, et al. Enhancing assessment of corn growth performance using unmanned aerial vehicles (UAVs) and deep learning[J]. *Measurement*, 2023, 214: 112764.
- [21] LV Z G, XU B Y, ZHONG L H, et al. Improved monitoring of southern corn rust using UAV-based multi-view imagery and an attention-based deep learning method[J]. *Computers and Electronics in Agriculture*, 2024, 224: 109232.
- [22] 邬开俊, 白晨帅, 杜建军, 等. 无人机影像的玉米植株中心检测模型和方法[J]. *计算机工程与应用*, 2024(6): 1-17.
- [23] NI X D, WANG F M, HUANG H, et al. A CNN-and self-attention-based maize growth stage recognition method and platform from UAV orthophoto images[J]. *Remote Sensing*, 2024, 16(14): 2672.
- [24] 杨国亮, 盛杨杨, 洪鑫芳, 等. 改进 YOLOv8n 的果园番茄目标检测算法[J]. *计算机工程与应用*, 2024, 60(23): 1-13.
- [25] WANG Y D, ZHANG J, KAN M N, et al. Self-supervised equivariant attention mechanism for weakly supervised semantic segmentation[EB/OL]. (2020-04-09) [2024-10-30]. <https://arxiv.org/abs/2004.04581>.
- [26] YU W H, ZHOU P, YAN S C, et al. InceptionNeXt: when inception meets ConvNeXt[EB/OL]. (2023-03-29) [2024-10-30]. <https://arxiv.org/abs/2303.16900>.
- [27] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. (2023-01-24) [2024-10-30]. <https://arxiv.org/abs/2301.10051>.
- [28] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[EB/OL]. (2021-01-20) [2024-10-30]. <https://arxiv.org/abs/2101.08158.context=cs.CV>.
- [29] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[EB/OL]. (2019-11-19) [2024-10-30]. <https://arxiv.org/abs/1911.08287>.
- [30] CHAKROUN I, HABER T, ASHBY J T. SW-SGD: the sliding window stochastic gradient descent algorithm[J]. *Procedia Computer Science*, 2017, 108: 2318-2322.

(责任编辑:成纾寒)