doi:10.6041/j.issn.1000-1298.2026.18.027

基于改进 YOLO 11n 的甘薯苗与杂草识别轻量化方法

张万枝[1,2]　王旭阳[1,2]　刘　云[1,3]　刘学龙[1,3]　程　进[1,2]

(1. 山东农业大学机械与电子工程学院, 泰安 271018; 2. 农业装备智能化山东省工程研究中心, 泰安 271018;
3. 山东省园艺机械与装备重点实验室, 泰安 271018)

摘要: 针对甘薯大田种植复杂环境下小目标和重叠遮挡区域的杂草难以识别、移动设备算力有限实时性较差等问题,本研究对 YOLO 11n 模型进行改进,提出了一种兼顾精度和速度的甘薯苗与杂草识别轻量化模型(SCSF-YOLO 11n)。首先,采取 ShuffleNetV2 轻量化网络作为主干特征提取网络,降低模型参数量和浮点运算量;其次,使用 Context Guided Block 替换颈部网络 C3k2 模块的瓶颈(Bottleneck)结构,整合目标区域和周边环境信息,减少田间背景干扰;为解决甘薯苗与杂草重叠遮挡导致局部细节特征丢失的问题,将 SEAM(Separated and enhancement attention module)注意力机制集成至头部网络,进行多尺度特征提取与融合,提升模型遮挡场景下识别能力;最后,采用 Focaler-MPDIoU 作为损失函数,加快模型网络收敛速度,降低其损失值。试验结果表明,改进后模型在验证集上的精确率、召回率、平均精度均值(mAP_{50})分别为 93.6%、94.3%、95.4%,参数量仅为 1.4×10^6,模型内存占用量为 3.2 MB,浮点运算量为 3.1×10^9,处理速度为 112.6 f/s;与 YOLO v5n、YOLOX-tiny、YOLO v7-tiny、YOLO v8n、YOLO v9t、YOLO v10n、YOLO 12n、DINO 等模型相比,mAP_{50} 分别提升 3.0、4.1、3.7、4.6、4.3、5.8、5.1、8.9 个百分点;模型可视化结果显示,改进模型在分散、密集、甘薯苗遮挡场景的识别效果优于其他模型。改进模型部署在 NVIDIA Jetson Orin Nano 上的田间试验结果表明,mAP_{50} 为 94.6%,单幅图像平均推理时间为 33.26 ms,满足甘薯大田复杂环境移动设备部署和实时性需求,可为后续甘薯大田除草机器人产业化应用提供技术支撑。

关键词: 目标识别; 杂草; YOLO 11n; ShuffleNetV2; 注意力机制

中图分类号: S226; TP391.4　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)18-0322-14　　**OSID**:

Lightweight Method for Sweet Potato Seedling and Weed Identification Based on Improved YOLO 11n

Zhang Wanzhi[1,2]　Wang Xuyang[1,2]　Liu Yun[1,3]　Liu Xuelong[1,3]　Cheng Jin[1,2]

(1. *College of Mechanical and Electronic Engineering*, *Shandong Agricultural University*, *Taian* 271018, *China*
2. *Shandong Provincial Engineering Research Center for Intelligent Agricultural Equipment*, *Taian* 271018, *China*
3. *Shandong Provincial Key Laboratory of Horticultural Machinery and Equipment*, *Taian* 271018, *China*)

Abstract: Aiming at the problems that weeds in small targets and overlapping occluded areas were difficult to identify under the complex environment of sweet potato field planting, and that the real-time performance was poor due to limited computing power of mobile devices, the YOLO 11n model was improved and a lightweight model (SCSF-YOLO 11n) was proposed for sweet potato seedling and weed identification that balanced accuracy and speed. Firstly, the ShuffleNetV2 lightweight network was adopted as the backbone feature extraction network to reduce the number of parameters and floating-point operations of the model. Secondly, the Context Guided Block was used to replace the bottleneck structure of the C3k2 module in the neck network, integrating the target area and surrounding environment information to reduce the interference of field background. To solve the problem of missing local detailed features caused by overlapping occlusion between sweet potato seedlings and weeds, the separated and enhancement attention module (SEAM) attention mechanism was integrated into the head network to

收稿日期: 2026-04-14　修回日期: 2026-04-27

基金项目: 山东省薯类产业技术体系建设专项资助项目(SDARS-16)和山东省重点研发计划(重大科技创新工程)项目(2022CXGC020703)

作者简介: 张万枝(1986—),男,副教授,博士,主要从事智能农机装备研究,E-mail: zhangwanzhi@163.com

通信作者: 程进(1993—),男,副教授,博士,主要从事智能农机装备研究,E-mail: ttcheng@163.com

extract and fuse multi-scale features, improving the identification ability of the model in occluded scenes. Finally, Focaler－MPDIoU was adopted as the loss function to accelerate the convergence speed of the model network and reduce its loss value. Experimental results showed that the improved model achieved precision, recall and mAP_{50} (mean average precision at 0.5 IoU) of 93.6%, 94.3% and 95.4% on the validation set, respectively, with only 1.4×10^6 parameters, 3.2 MB memory occupation, 3.1×10^9 floating-point operations and a processing speed of 112.6 frames per second. Compared with YOLO v5n, YOLOX－tiny, YOLO v7－tiny, YOLO v8n, YOLO v9t, YOLO v10n, YOLO 12n and DINO, the mAP_{50} was increased by 3.0, 4.1, 3.7, 4.6, 4.3, 5.8, 5.1 and 8.9 percentage points, respectively. Visualization results demonstrated that the improved model outperformed other models in scattered, dense and sweet potato seedling occluded scenes. Field tests deploying the improved model on NVIDIA Jetson Orin Nano showed that the mAP_{50} was 94.6%, and the average inference time for a single image was 33.26 ms, which met the requirements of mobile device deployment and real-time performance in complex sweet potato field environments. The research result can provide technical support for the subsequent industrial application of sweet potato field weeding robots.

Key words: target recognition; weeds; YOLO 11n; ShuffleNetV2; attention mechanism

0　引言

甘薯田间杂草会和甘薯苗争夺水分、养分以及生长空间,制约了甘薯产量与品质的提升[1-2]。当前,甘薯田间杂草防控主要依赖人工除草、地膜覆盖和化学除草等方式[3]。人工除草效率低,生产成本高[4];地膜覆盖地域限制显著,且残留物易造成环境污染;化学除草会造成药剂残留,危害作物生长,进而破坏土壤生态环境。随着人工智能技术发展,基于机器视觉的智能除草机器人展现出广阔应用前景,但前提需要实现田间杂草的精准识别[5-6]。

近年来,基于深度学习的目标检测模型,国内外学者开展了不同作物田间复杂环境下杂草识别研究。Das 等[7]基于 YOLO v7 模型,引入 CBAM 注意力机制与 BiFPN 加权双向特征金字塔网络,实现了棉花与杂草的精准识别,杂草平均精度可达 84%;García-Navarrete 等[8]针对玉米田间常见 4 类杂草进行多种 YOLO 模型识别性能测试,其中 YOLO v9s 能够降低田间环境干扰,有效区分玉米植株与杂草,平均精度均值可达 83.4%;Cui 等[9]基于 Faster R－CNN 模型,采取优化后的 VGG16 作为特征提取网络,对大豆田间杂草进行了识别,平均精度均值达到 88.69%,单幅图像检测时间为 310 ms。田间杂草识别模型通常需部署在算力资源受限的边缘设备上,因此还需考虑模型轻量化问题。齐江涛等[10]提出了一种基于 YOLO v8n 的田间玉米杂草快速检测算法,平均精度均值为 97.80%,检测速度为 47.81 f/s,模型内存占用量为 4.10 MB;宁姗等[11]针对大豆田间植株与杂草相互遮挡、难以识别的问题,引入感受野增强 SAConv 模块与 Adown 自适应下采样,构建轻量化 YOLO 11n－ADS 模型,平均精度均值为 92.0%,模型检测速度为 137 f/s;杨森森等[12]采取 MobileViT 作为杂草特征提取轻量化网络,模型识别准确率为 99.61%,处理速度为 12.04 f/s,内存占用量为 11.2 MB;冀汶莉等[13]提出一种基于改进 YOLO v5s 的玉米田间杂草识别轻量化方法,改进后模型单幅图像检测时间为 118.1 ms,模型内存占用量为 6.23 MB。目前田间杂草识别仍然面临大田环境复杂、作物与杂草间相互遮挡、轻量化模型的识别精度与速度需平衡等问题,还需进一步研究。

针对上述问题,基于甘薯大田种植复杂环境下小目标和重叠遮挡区域的杂草难以识别、移动设备算力有限实时性较差等现状,本文提出一种兼顾精度和速度的甘薯苗与杂草识别轻量化模型。基于 YOLO 11n 模型,替换 ShuffleNetV2 主干网络和 C3k2_CG 特征提取模块,以降低模型计算资源消耗,提升小目标杂草识别能力;在头部网络添加 SEAM 注意力机制,以增强模型在甘薯苗遮挡场景的鲁棒性;将损失函数替换为 Focaler－MPDIoU,以加快杂草识别网络收敛,提高模型定位精度。

1　数据集构建

1.1　图像采集

采用作物表型感知机器人搭载 Intel RealSense D435i 型深度相机进行甘薯苗与杂草 RGB 图像采集,如图 1 所示。采集地点为山东省泰安市岱岳区鑫腾家庭农场甘薯田,采集时间为 2025 年 5 月 13—20 日,甘薯品种为烟薯 25 号,甘薯移栽 10～17 d(发根返青期至分枝结薯期)。深度相机安装于机器人底部,距离地面高度 75 cm,采取俯视拍摄。考虑到甘薯大田复杂环境,选择不同时段(早晨、中午和傍晚)、不同天气(晴天、阴天和多云)和不同场景(分散、密集和甘薯苗遮挡)进行图像采集,共获得甘薯苗与杂草图像 1 280 幅,部分采集图像如图 2

所示。其中甘薯苗遮挡场景为甘薯苗叶片、茎蔓局部覆盖杂草,二者通常表现为重叠生长。

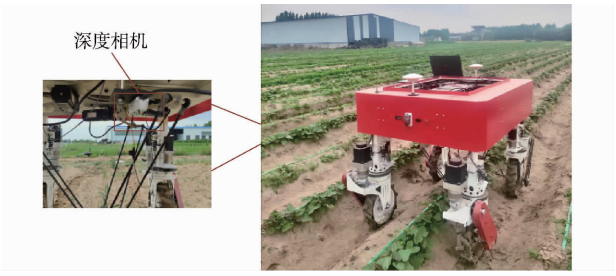


图 1 图像采集场景

Fig. 1 Image acquisition scenario



图 2 样本图像示例

Fig. 2 Examples of sample images

1.2 数据集标注与增强

采用 Labelimg 软件以矩形框进行作物信息注释,保存为 YOLO 格式,标签类别为杂草“weed”和甘薯苗“sweet potato”。为避免训练过程中出现过拟合现象[14],采取添加随机噪声、旋转、平移、镜像、亮度调整等方式进行数据增强,经数据增强后图像为 6 400 幅,按照比例 8∶1∶1划分为训练集(5 120 幅)、验证集(640 幅)以及测试集(640 幅)。部分增强图像如图 3 所示。

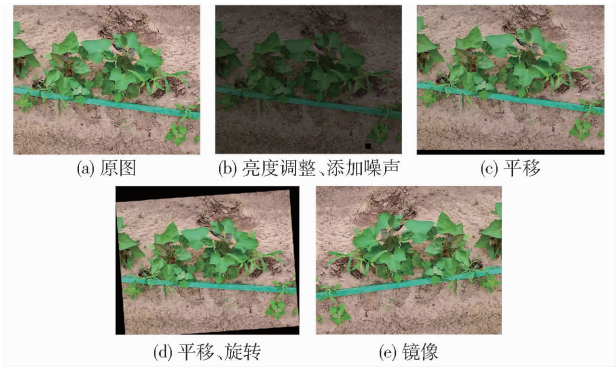


图 3 数据增强图像示例

Fig. 3 Examples of augmented images

2 甘薯苗与杂草识别轻量化模型

2.1 YOLO 11 模型

YOLO 11[15] 模型由输入端(Input)、主干网络(Backbone)、颈部网络(Neck)和头部网络(Head)等组成。Input 负责原始图像输入,通过图像自适应缩放和 Mosic 数据增强方式提升网络模型的鲁棒性;Backbone 引入 C3k2 模块和 C2PSA 模块,使得网络能够有效捕捉多层次的空间信息和语义信息[16];Neck 沿用 FPN 和 PAN 组合架构进行多尺度特征信息融合;Head 引入深度可分离卷积,进一步降低检测头的参数量和浮点运算量。该网络模型根据参数量和复杂度分为 5 个版本,分别为 n、s、m、x、l,以上版本在自建的甘薯苗与杂草数据集上的试验结果如表 1 所示。最终选择浮点运算量最少且保持较高精度的 YOLO 11n 作为基线模型进行改进。

表 1 YOLO 11 不同版本模型试验结果

Tab. 1 Experimental results across different versions of YOLO 11 model

模型	mAP_{50}/%	参数量	浮点运算量
YOLO 11n	93.1	2.60×10^{6}	6.300×10^{9}
YOLO 11s	93.9	9.40×10^{6}	2.170×10^{10}
YOLO 11m	92.6	2.09×10^{7}	6.850×10^{10}
YOLO 11l	93.0	2.53×10^{7}	8.760×10^{10}
YOLO 11x	94.2	5.69×10^{7}	1.965×10^{11}

2.2 模型改进

由于甘薯大田种植环境较复杂、甘薯幼苗和杂草相似、目标之间易重叠遮挡,导致尺度较小杂草的识别率较差,原始模型存在较多错检和漏检现象。另外,YOLO 11n 模型部署在边缘移动设备上时可能会影响实时识别效果。为解决上述问题,本文提出了一种兼顾精度和速度的甘薯与杂草识别轻量化模型 SCSF-YOLO 11n,模型结构如图 4 所示。模型改进主要从 4 个方面进行:①选取 ShuffleNetV2 轻量化特征提取网络重构 Backbone,降低模型的浮点运算量和参数量。②采用 Context Guided Block 替换 C3k2 模块中的瓶颈(Bottleneck)结构,整合杂草目标区域与周围环境信息,降低田间环境干扰,提升模型的特征提取能力。③在 Head 中融入 SEAM 注意力机制,解决甘薯苗与杂草重叠遮挡导致的局部细节特征丢失问题,增强模型对甘薯苗遮挡区域小目标杂草的识别能力。④使用 Focaler-MPDIoU 作为边界框损失函数,调整困难样本权重,加快网络收

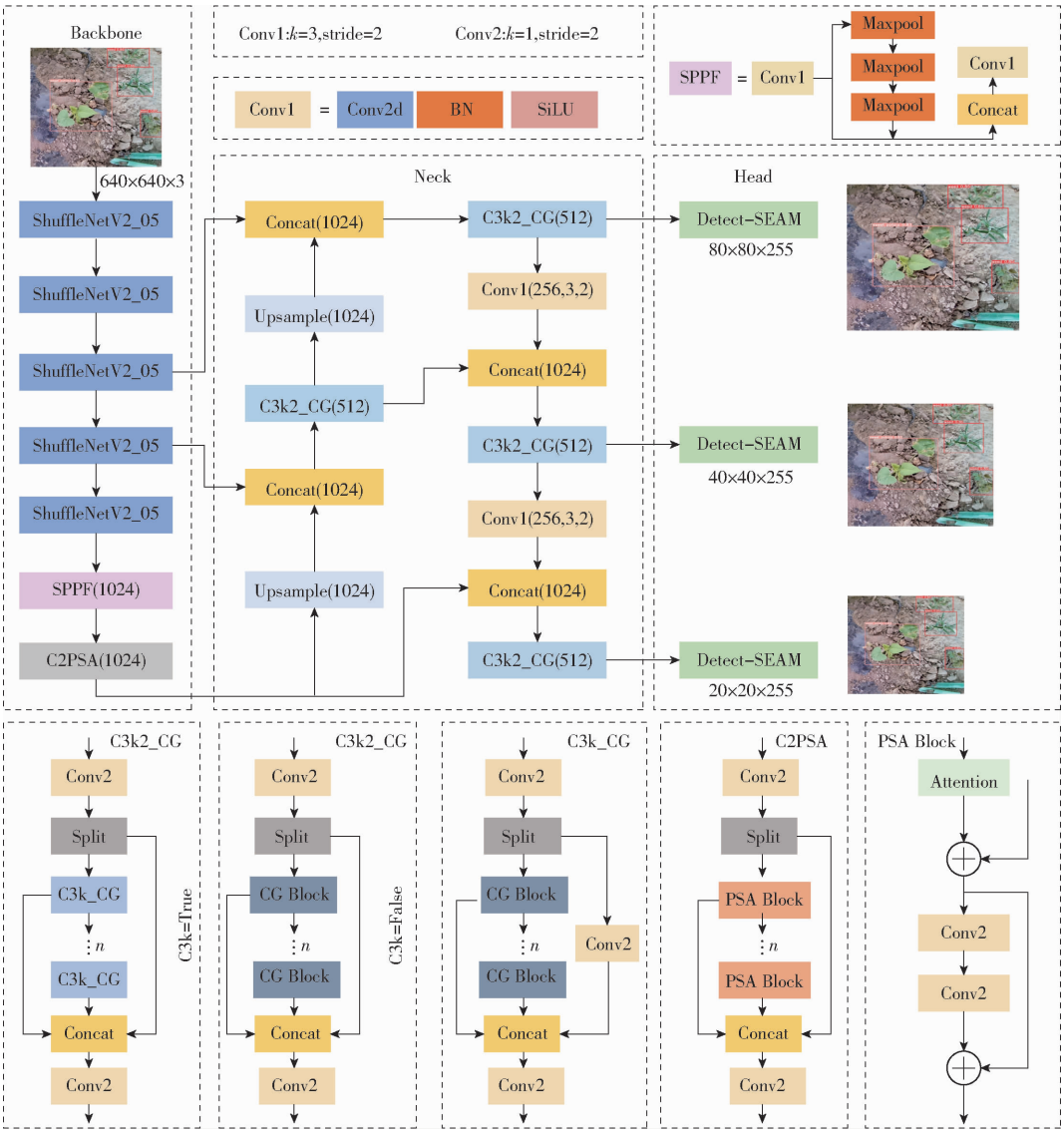


图 4　SCSF－YOLO 11n 模型结构

Fig. 4　SCSF－YOLO 11n model structure

敛速度并提高识别精度。

2.2.1　ShuffleNetV2 特征提取网络

在 YOLO 11 原始网络中，采取 CSP－Darknet 作为主干特征提取网络。由于其参数量过大，后续部署在算力有限的移动设备上会影响设备的识别效率。采用 ShuffleNetV2 重构 CSP－Darknet 主干网络，在保证精度的同时降低模型冗余计算和内存访问成本。

ShuffleNet[17]是一种轻量化卷积神经网络。为解决传统 CNN（Convolutional neural network）在移动设备部署的难题，采用分组卷积（Group convolution，GConv）和通道混洗（Channel shuffle）操作，降低模型浮点运算量和参数量。ShuffleNetV1 网络虽减少了计算开销，但过多采用分组卷积和元素操作（Add、ReLU），削弱了模型特征提取能力。在 V1 网络基础上，ShuffleNetV2 进一步提出了 4 条高效网络设计的实用准则：①保证输入和输出的通道数尽量一致。②避免分组卷积的使用。③降低网络碎片化程度。④减少元素操作[18]。

ShuffleNetV2 网络架构由基本单元（Basic uinit）和下采样单元（Down-sampling unit）组成，如图 5 所示。V2 的基本单元引入通道分离（Channel split）操作，左右两侧分支通道数各占原来的一半，左侧分支采取恒等映射，右侧分支经过 2 次 1×1 普通卷积和 1 次 3×3 的深度可分离卷积（Depthwise separable convolution，DWConv）操作；然后采取 Concat 操作将左右分支输出特征拼接在一起，保证原始特征图的输入和输出通道数一致；最后进行通道混洗操作，重新排列输出层的通道顺序，促进通道之间的信息流动，增强模型表征能力，如图 6 所示。V2 的下采样单元，并未采取通道分离操作，而是直接将特征输送到 2 条分支，左右两侧分支均采用 1×1 普通卷积和 3×3 深度可分离卷积操作，Concat 操作后特征图的分辨率减半，通道数输出为原来的

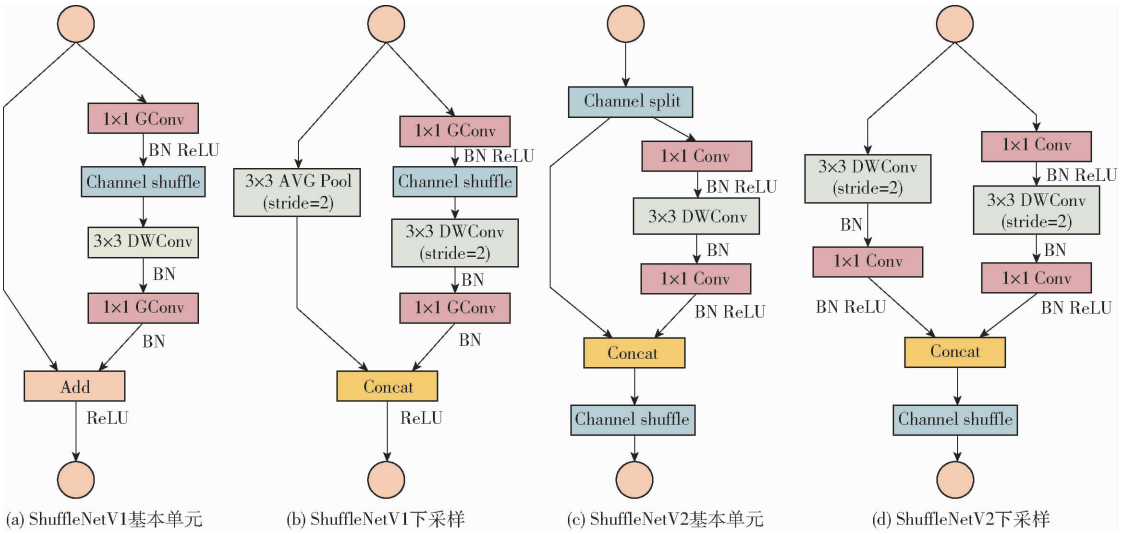


图 5 ShuffleNet 网络架构

Fig. 5 ShuffleNet network architecture

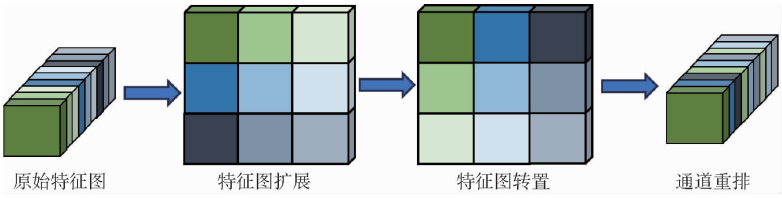


图 6 通道混洗操作

Fig. 6 Channel shuffling operation

2 倍;同时增加网络宽度,进一步提升网络特征提取能力[19]。

2.2.2 C3k2_CG 特征提取模块

C3k2 模块作为 YOLO v8n 特征提取模块(C2f)的变体,引进可调节参数 C3k,当 C3k 为 True 时,会把 Bottleneck 替换为 C3k;当 C3k 为 False 时,C3k2 模块为普通的 C2f 模块。由于甘薯田间背景比较复杂、甘薯苗和杂草在颜色和形态上极为相似,原始模型存在较多杂草错检为甘薯苗的现象。为减少田间背景干扰,提高模型对小目标杂草的关键特征提取能力,采用 Context Guided Block[20] 替换 C3k2 模块的 Bottleneck,并将其命名为 C3k2_CG。该模块由局部特征提取器 $f_{loc}(*)$、周围上下文提取器 $f_{sur}(*)$、联合特征提取器 $f_{joi}(*)$ 和全局上下文提取器 $f_{glo}(*)$ 组成,结构如图 7 所示。

Context Guided Block 工作原理分为 3 个步骤:①采用 $f_{loc}(*)$ 和 $f_{sur}(*)$ 进行目标区域和周围区域信息的整合,帮助模型更好地识别小目标杂草;$f_{loc}(*)$ 采取 3×3 卷积进行局部区域特征提取,保留杂草纹理和边缘信息,实现对甘薯苗和杂草的共同检测;$f_{sur}(*)$ 使用 3×3 空洞卷积(Dilated convolution,DConv),扩大模型感受野从而获取更丰富的全局特征。②使用 $f_{joi}(*)$ 将 $f_{loc}(*)$ 和 $f_{sur}(*)$ 提取到的图像特征拼接在一起,经批量归一化(BN)和参数化线性单元(PReLU)处理,增强目标区域特征质量,提升模型对小目标特征的学习能力。③$f_{glo}(*)$ 通过全局平均池化(GAP)和 2 个全连接层(FC),捕获并利用输入样本图像的全局信息,使

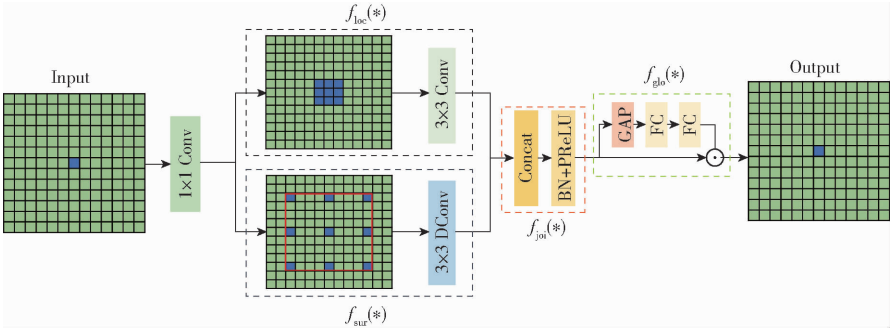


图 7 Context Guided Block 模块结构

Fig. 7 Architecture of Context Guided Block module

模型更准确识别到小目标杂草的位置和特征[21]。这 4 个组件协同合作，通过对全局上下文信息的综合利用，帮助模型更好的区分形态相似的甘薯苗和杂草，进而提升模型对小目标杂草的识别与定位能力。另外，该模块能够优化特征提取过程，减少对冗余信息的处理，降低模型参数量。

2.2.3　SEAM 检测头

在甘薯大田复杂环境下，甘薯苗与杂草易重叠生长、遮挡区域杂草目标尺度过小且像素占比低，原始模型存在较多漏检现象。将 SEAM[22]（Separated and enhancement attention module）注意力机制集成到头部网络，捕捉小目标杂草的局部细节特征，提升模型在遮挡场景下的识别准确率。

SEAM 模块通过深度可分离卷积和残差连接组合来增强未遮挡部分的响应，旨在补偿被遮挡部分的响应损失，从而提高模型在遮挡场景下的泛化能力和感知能力，结构如图 8 所示。首先，采用 3 个不同尺寸（patch = 6、7、8）的 CSMM（Channel and space mixed module）模块，进行多尺度特征提取，大尺度特征用于保持甘薯苗和杂草整体形状信息，小尺度特征用于捕捉杂草叶片的边缘和纹理信息，二者相互补充提升模型的特征提取能力。其次，经平均池化（Average Pooling）和 2 个全连接层（FC）处理，增强甘薯苗、杂草形态和纹理信息通道之间的联系，抑制光照、田间环境对杂草识别的干扰。最后，经过通道扩展（Channel expansion）扩大特征图的通道数，将其输出特征与经 CSMM 模块融合后的特征相乘，可为头部网络提供更丰富的图像信息，使模型能够更精确地聚焦甘薯苗遮挡下的小目标杂草。

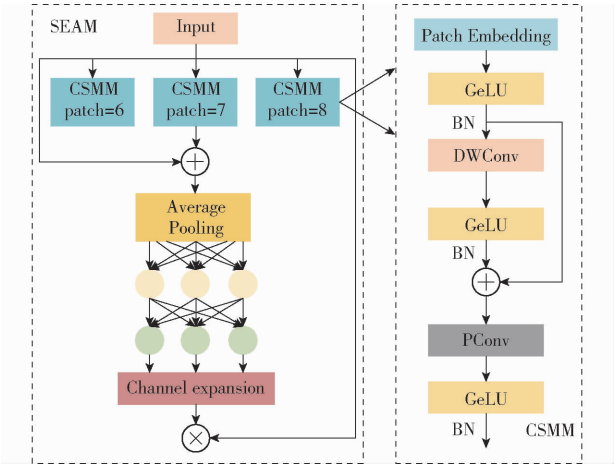


图 8　SEAM 模块结构

Fig. 8　Architecture of SEAM module

CSMM 模块采用 DWConv 进行逐通道卷积，获取较大感受野，在保证特征提取能力的同时降低模型的参数量和计算复杂度。该模块将 DWConv 运算后的特征，经批归一化（BN）和非线性激活函数（GeLU）处理，保留细微纹理特征以及加速模型训练过程；同时，引入残差连接，将输入的原始特征直接添加到输出中，确保甘薯苗和杂草的整体形状和边缘信息不丢失。DWConv 易忽略通道之间的信息交互，在其输出后使用逐点卷积（Pointwise convolution，PConv）对特征图升维，促进跨通道信息的融合。在 Head 中引入 SEAM 注意力机制后，模型通过学习遮挡区域与被遮挡区域的联系，预测并且补偿杂草被遮挡部分的特征，提高了对甘薯苗遮挡区域小目标杂草识别能力。

2.2.4　Focaler－MPDIoU 损失函数

定位损失用于优化预测边界框与真实边界框之间的几何差异，使得模型更准确地预测目标位置和大小。YOLO 11 模型采用 CIoU 计算定位损失，公式为

$$L_{CIoU} = 1 - IoU + \frac{b^2}{C^2} + \alpha v \tag{1}$$

式中　IoU——预测框与真实框交并比

b——预测框与真实框中心点距离

C——预测框与真实框最小外接矩形框对角线距离

α——权重系数

v——修正因子

在实际田间复杂场景下，当甘薯苗和杂草尺度差异较大时，CIoU 无法根据样本的多样性进行合理调整，限制了模型的收敛速度和精度值。因此，本文采用 MPDIoU[23]（Minimum point distance intersection over union）作为边界框损失函数，示意图如图 9 所示。



图 9　MPDIoU 示意图

Fig. 9　Schematic of MPDIoU

MPDIoU 通过简化计算过程，最小化预测框左上角点到真实框右下角点距离，从而提高模型定位精度和收敛速度，公式为

$$MPDIoU = \frac{A \cap B}{A \cup B} - \frac{d_1^2}{W^2 + H^2} - \frac{d_2^2}{W^2 + H^2} \tag{2}$$

其中

$$d_1 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2 \tag{3}$$

$$d_2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \tag{4}$$

式中 A——人工标注真实框

B——模型识别预测框

W——图像宽度

H——图像高度

d_1——预测框左上角顶点(x_1^{gt}, y_1^{gt})与真实框左上角顶点(x_1^{prd}, y_1^{prd})之间距离

d_2——预测框右下角顶点(x_2^{gt}, y_2^{gt})与真实框右下角顶点(x_2^{prd}, y_2^{prd})之间距离

MPDIoU 可以加快模型对预测框与真实框位置关系的拟合,能够更好地适应二者重叠的边界框回归任务,提升模型对于小目标杂草的识别精度。但该损失函数仍然是围绕预测框与真实框的几何关系来提升模型定位精度,忽视了易识别样本和困难样本分布对边界框回归的影响。因此本文选择在 MPDIoU 基础上引入 Focaler-IoU[24](More focused intersection over union loss),通过线性区间映射关注不同的回归样本,使边界框回归效果得到改善。计算公式为

$$L_{Focaler-MPDIoU} = 1 - IoU_{Focaler} + \frac{d_1^2 + d_2^2}{W^2 + H^2} \tag{5}$$

其中

$$IoU_{Focaler} = \begin{cases} 0 & (IoU < d) \\ \dfrac{IoU - d}{u - d} & (d \leqslant IoU \leqslant u) \\ 1 & (IoU > u) \end{cases} \tag{6}$$

式中$[d, u] \in [0, 1]$,调整 d、u 使 Focaler-MPDIoU 关注困难样本,提升模型处理样本不平衡的能力[25]。

在甘薯苗与杂草目标识别任务中,采用 MPDIoU 直接优化预测框与真实框之间的几何差异,加速模型的收敛过程;引入 Focaler-IoU 调整困难样本的权重,增强模型对于小目标和甘薯幼苗形态相似的杂草识别能力。

3 结果与分析

3.1 试验环境与参数配置

试验在 Window 11(64 位)操作系统下进行,处理器为 12th Gen Intel(R) Core(TM) i7-12700k,主频 3.60 GHz,运行内存为 32 GB,搭载 RTX 3080Ti 显卡。在 PyCharm 平台搭建深度学习框架,编程语言为 Python 3.10, CUDA 12.1, Pytorch 2.2.2 + torchvision 0.17.2,所有试验均在此环境下进行。

训练过程参数设置:图像尺寸调整至 640 像素×640 像素,训练轮次(epoch)为 200,批量大小(Batch size)为 8,工作进程数(workers)为 4。采用 SGD 作为优化器,学习率设置为 0.01,动量因子系数为 0.937,权重衰减参数为 0.000 5。

模型性能评估阶段推理设置:图像尺寸为 640 像素×640 像素,批量大小(Batch size)为 8,置信度阈值(conf)为 0.25,IoU 阈值为 0.45。上述参数适用各模型训练和推理过程,并未根据个别模型特点进行单独优化,以确保试验横向比较的公平性。

3.2 评价指标

选用精确率(Precision, P)、召回率(Recall, R)、平均精度(Average precision, AP)、平均精度均值(mAP_{50}、mAP_{50-95})评估模型识别精度,采用参数量、模型内存占用量、浮点运算量评估模型轻量化,使用帧率评估模型识别速度。

3.3 试验结果

3.3.1 改进模型训练过程与结果

相同试验条件下改进前后模型训练过程精确率、召回率、平均精度均值(mAP_{50})表现如图 10 所示。由图 10 可看出,SCSF-YOLO 11n 模型的精确率、召回率以及平均精度均值曲线,在 0~50 轮内迅速上升;当迭代到 180 轮左右,模型逐渐趋于稳定;第 200 轮时,模型呈收敛状态,精确率为 93.6%,召回率为 94.3%,mAP_{50}为 95.4%。

2 种模型在验证集上识别结果混淆矩阵如图 11 所示。由图 11 可看出,在引入 C3k2_CG 模块和

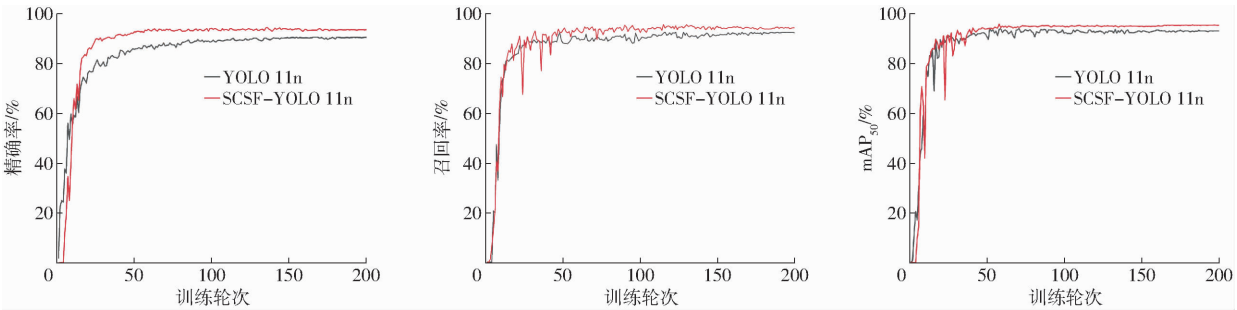
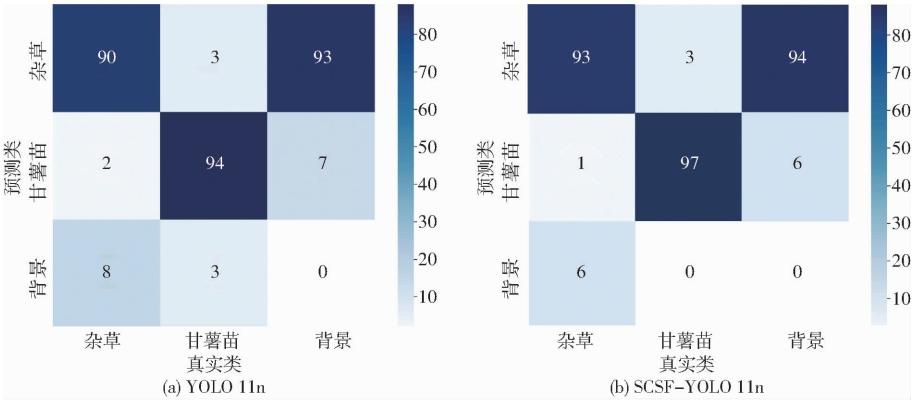


图 10 不同模型训练过程性能曲线

Fig. 10 Performance curves of different model training processes

SEAM 注意力机制后，增强了模型对于小目标遮挡杂草的识别性能，使模型能够更好地区分甘薯苗与杂草，同时 SCSF－YOLO 11n 模型识别误差远小于 YOLO 11n 模型。



(a) YOLO 11n　　(b) SCSF-YOLO 11n

图 11　不同模型在验证集上的混淆矩阵

Fig. 11　Confusion matrix of different models on verification set

3.3.2　不同轻量化主干网络性能测试

对 YOLO 11n 模型主干网络进行重构，选取 6 种轻量级网络进行模型性能测试。由表 2 可知，采用 ShuffleNetV2 作为 YOLO 11n 的主干网络时，精确率为 90.2%，召回率为 91.7%，mAP_{50} 为 92.9%，mAP_{50-95} 为 83.1%，参数量为 1.7×10^6，模型内存占用量为 3.8 MB，浮点运算量为 4.1×10^9。相比 YOLO 11n 原始模型，精确率、召回率、mAP_{50}、mAP_{50-95} 分别下降 0.3、0.6、0.2、0.3 个百分点，但参数量、模型内存占用量、浮点运算量分别为 YOLO 11n 的 65.4%、69.1%、65.1%。MobileNetV4 的精确率提升 1 个百分点，但参数量为 YOLO 11n 的 2 倍多，模型内存占用量增加 5.8 MB，浮点运算量为 2.13×10^{10}，不符合模型后续轻量化部署要求。FasterNet 精确率提升 0.7 个百分点，但召回率下降 2.7 个百分点，增大了模型对于杂草的漏检率，并不适合处理这类甘薯大田复杂环境小目标杂草检测任务。与 ConvNextV2、GhostHGNetV2、HGNetV2、StarNet 网络相比，ShuffleNetV2 精确率分别提升 0.3、0.7、0.2、1.4 个百分点，召回率分别提升 2.4、0.3、2.6、1.2 个百分点，mAP_{50} 分别提升 2.3、0.9、1.4、1.7 个百分点，mAP_{50-95} 分别提升 2.4、1.0、1.8、2.0 个百分点，而参数量为其的 31.5%、54.8%、80.9%、85%，模型内存占用量为其的 34.2%、56.7%、80.8%、90.4%，浮点运算量为其的 32.3%、51.9%、71.9%、82%。因此，选取识别精度较高且模型最为轻量化的 ShuffleNetV2 作为 YOLO 11n 的主干网络。

表 2　轻量化主干网络性能测试结果

Tab. 2　Results of performance testing for lightweight backbone networks

模型	精确率/%	召回率/%	mAP_{50}/%	mAP_{50-95}/%	参数量	模型内存占用量/MB	浮点运算量
YOLO 11n	90.5	92.3	93.1	83.4	2.6×10^6	5.5	6.30×10^9
ConvNextV2	89.9	89.3	90.6	80.7	5.4×10^6	11.1	1.27×10^{10}
FasterNet	91.2	89.6	92.4	82.3	3.9×10^6	8.1	9.30×10^9
GhostHGNetV2	89.5	91.4	92.0	82.1	3.1×10^6	6.7	7.90×10^9
HGNetV2	90.0	89.1	91.5	81.3	2.1×10^6	4.7	5.70×10^9
MobileNetV4	91.5	89.3	91.9	82.0	5.4×10^6	11.3	2.13×10^{10}
StarNet	88.8	90.5	91.2	81.1	2.0×10^6	4.2	5.00×10^9
ShuffleNetV2	90.2	91.7	92.9	83.1	1.7×10^6	3.8	4.10×10^9

3.3.3　不同颈部网络特征提取模块性能测试

YOLO 11n 模型主干网络替换为 ShuffleNetV2 后，精度相比基线模型略有下滑，故选择对颈部网络特征提取模块 C3k2 进行重新设计，提升模型对于杂草的特征提取能力。通过引入 StarBlock（Star）、FasterNetBlock（Faster）、KANBlock（KAN）、EIEM、Context Guided Block（CG），对 C3k2 模块中的瓶颈结构进行替换，试验结果如表 3 所示。由表 3 可看出，采用 Context Guided Block 与其他模块相比，模型精确率分别提升 2.9、1.4、1.0、1.1、0.8 个百分点，召回率分别提升 1.8、2.0、2.4、2.6、3.5 个百分点，mAP_{50} 分别提升 1.9、0.6、0.5、0.9、1.2 个百分点，mAP_{50-95} 分别提升 1.5、0.5、0.3、0.8、0.9 个百分点。相比 ShuffleNetV2 + C3K2 模型，参数量减少

表 3 瓶颈结构性能试验结果

Tab. 3 Experimental result comparing bottleneck structures

模型	精确率/%	召回率/%	mAP_{50}/%	mAP_{50-95}/%	参数量	模型内存占用量/MB	浮点运算量
ShuffleNetV2 + C3k2	90.2	91.7	92.9	83.1	1.7×10^6	3.8	4.1×10^9
ShuffleNetV2 + C3k2_Star	91.7	91.5	94.2	84.1	1.7×10^6	3.7	4.1×10^9
ShuffleNetV2 + C3k2_Faster	92.1	91.1	94.3	84.3	1.6×10^6	3.5	3.8×10^9
ShuffleNetV2 + C3k2_KAN	92.0	90.9	93.9	83.8	1.9×10^6	4.1	4.3×10^9
ShuffleNetV2 + C3k2_EIEM	92.3	90.0	93.6	83.7	1.7×10^6	3.8	4.2×10^9
ShuffleNetV2 + C3k2_CG	93.1	93.5	94.8	84.6	1.5×10^6	3.3	3.8×10^9

2×10^5,模型内存占用量降低 0.5 MB,浮点运算量略有下降。通过以上分析,采用 ShuffleNetV2 + C3k2_CG 组合的综合性能有显著提升。故选用 ShuffleNetV2 轻量化网络和 C3k2_CG 特征提取模块组合。

3.3.4 不同损失函数性能测试

将原边界框损失函数 CIoU 替换为 Focaler - MPDIoU,并引进 WIoUv3、EIoU、MPDIoU 等损失函数进行边界框回归过程性能比较,试验结果如图 12 所示。由图 12 可看出,随着训练轮次增加,以上 5 种边界框损失函数曲线均近乎收敛。在训练初期和末期,EIoU 损失值均远高于其他损失函数,边界框回归速度较慢,表明 EIoU 并不适合处理这类田间环境复杂的目标检测任务。WIoUv3 旨在降低质量较差的杂草样本对模型识别的危害,但随着训练轮次增加,曲线先下降后上升趋于收敛(过拟合),导致模型泛化性变差。采用 CIoU 和 MPDIoU 作为边界框损失函数,二者收敛速度接近;在整个训练过程,MPDIoU 损失值略低于 CIoU。Focaler - MPDIoU 损失函数,在训练 30 轮左右开始收敛且损失值为 1.09,收敛速度远超其他损失函数;在训练末期,损失值为 1.03,其数值远低于其他损失函数。

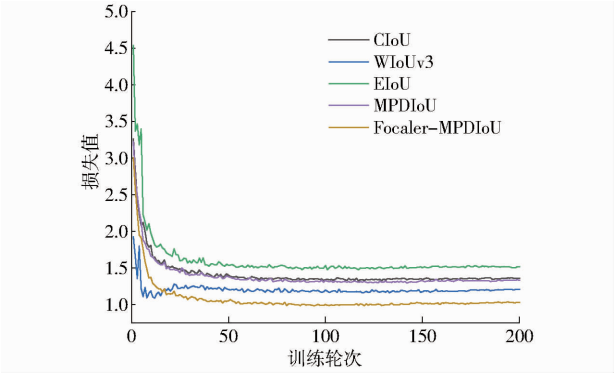


图 12 5 种边界框回归损失函数变化曲线

Fig. 12 Variation curves of five boundary box regression loss functions

3.3.5 消融试验分析

为验证本文改进方法对模型识别性能的影响,在相同试验环境下设置 5 组消融试验,结果如表 4 所示。由表 4 可看出,模型主干网络替换为 ShuffleNetV2 后,精度指标略有降低,模型内存占用量、参数量、浮点运算量均有明显下降。在颈部网络引入 C3k2_CG 模块,整合全局上下文信息,提升模型对于甘薯苗与杂草关键特征的提取能力,甘薯苗与杂草的平均精度分别提升 1.9、1.6 个百分点,mAP_{50}提升 1.9 个百分点,模型参数量、内存占用量以及浮点运算量略有降低。将 SEAM 注意力机制融入到头部网络后,使得模型更加关注被甘薯苗遮挡的杂草特征,提升模型在遮挡环境下的识别能力,甘薯苗与杂草的平均精度分别提升 0.3、0.4 个百分点,mAP_{50}提升 0.3 个百分点,模型参数量和浮点运算量略有降低。替换边界框损失函数为 Focaler - MPDIoU,调整模型对困难样本(小目标遮挡杂草)的权重,加快模型收敛过程,参数量和浮点运算量基本不变,甘薯苗与杂草的平均精度分别提升 0.2、0.4 个百分点,mAP_{50}提升 0.3 个百分点。通过以上改进,模型的识别性能达到最佳,甘薯苗与杂草平均精度、mAP_{50}分别提升至 97.3%、93.5%、95.4%,表明所提改进方法能够有效提升模型识别甘薯苗与杂草的能力;模型参数量、内存占用量、浮点运算量分别减少 46.2%、41.8%、50.8%,使其更适合在农业机器人等计算资源受限的边缘移动设备上部署。

3.3.6 不同模型性能测试

为验证 SCSF - YOLO 11n 模型识别性能,在相同试验环境和条件下,选取多个当前主流目标识别算法进行试验,结果如表 5 所示。由表 5 可看出,改进后模型精确率和召回率均高于其他目标识别模型;改进模型与 YOLO 11n、YOLO v5n、YOLOX - tiny、YOLO v7 - tiny、YOLO v8n、YOLO v9t、YOLO v10n、YOLO 12n、DINO 相比,mAP_{50}分别提升 2.3、3.0、4.1、3.7、4.6、4.3、5.8、5.1、8.9 个百分点,杂草平均精度分别提升 2.3、3.2、4.4、4.0、4.9、4.6、6.0、5.4、8.9 个百分点。在识别速率上,DINO 模型推理速度为 9.4 f/s,远低于 YOLO 同系列模型;相比

表 4　消融试验结果

Tab. 4　Ablation experiment result

ShuffleNetV2	C3k2_CG	SEAM Head	FMPDI	平均精度/%		mAP_{50}/%	参数量	模型内存占用量/MB	浮点运算量
				甘薯苗	杂草				
×	×	×	×	95.1	91.2	93.1	2.6×10^6	5.5	6.3×10^9
√	×	×	×	94.9	91.1	92.9	1.7×10^6	3.8	4.1×10^9
√	√	×	×	96.8	92.7	94.8	1.5×10^6	3.3	3.8×10^9
√	√	√	×	97.1	93.1	95.1	1.4×10^6	3.2	3.1×10^9
√	√	√	√	97.3	93.5	95.4	1.4×10^6	3.2	3.1×10^9

注：√表示采用此模块，×表示未采用此模块。

表 5　不同模型识别效果对比

Tab. 5　Comparison of detection performance across different models

模型	精确率/%	召回率/%	平均精度/%		mAP_{50}/%	帧率/(f·s^{-1})	参数量	模型内存占用量/MB	浮点运算量
			甘薯苗	杂草					
YOLO 11n	90.5	92.3	95.1	91.2	93.1	118.0	2.60×10^6	5.5	6.30×10^9
YOLO v5n	90.9	91.0	94.5	90.3	92.4	44.2	1.70×10^6	3.9	4.20×10^9
YOLOX－tiny	89.5	90.8	93.6	89.1	91.3	105.4	5.00×10^6	62.4	7.50×10^9
YOLO v7－tiny	91.1	90.2	93.9	89.5	91.7	73.5	6.00×10^6	12.3	1.32×10^{10}
YOLO v8n	90.5	88.4	93.1	88.6	90.8	122.0	3.00×10^6	6.3	8.10×10^9
YOLO v9t	89.6	90.5	93.3	88.9	91.1	83.8	2.60×10^6	5.8	1.07×10^{10}
YOLO v10n	89.2	87.5	91.7	87.5	89.6	87.4	2.20×10^6	5.9	6.50×10^9
YOLO 12n	89.6	89.4	92.4	88.1	90.3	97.3	2.50×10^6	5.5	6.30×10^9
DINO	84.7	81.9	88.4	84.6	86.5	9.4	4.70×10^7	546	2.71×10^{11}
SCSF－YOLO 11n	93.6	94.3	97.3	93.5	95.4	112.6	1.40×10^6	3.2	3.10×10^9

YOLO 11n，SCSF－YOLO 11n 模型识别速率降低 5.4 f/s，但仍领先其他目标识别模型，满足移动设备实时识别需求。此外，DINO 模型浮点运算量为 2.71×10^{11}，计算成本过高，并不适合部署在资源受限的移动端设备上。与 YOLO 同系列算法相比，改进模型参数量降低 1.2×10^6、3×10^5、3.6×10^6、4.6×10^6、1.6×10^6、1.2×10^6、8×10^5、1.1×10^6，浮点运算量减少 3.2×10^9、1.1×10^9、4.4×10^9、1.01×10^{10}、5.0×10^9、7.6×10^9、3.4×10^9、3.2×10^9，内存占用量为它们的 58.1%、82.1%、5.1%、26%、50.7%、55.1%、54.2%、58.1%，进一步表明本文改进模型的有效性。在甘薯苗和杂草数据集上，SCSF－YOLO 11n 模型在精度、速度和轻量化指标上都优于其他 8 种模型，满足后续在移动设备上甘薯苗和杂草识别的精度及实时性要求。

3.3.7　模型可视化分析

为更好地展示各项模块前后模型重点关注区域的变化，采取梯度加权类激活映射（Gradient-weight class activation mapping，Grad CAM）[26] 方法生成对应各组消融试验模型热力图，结果如图 13 所示，红色框表示模型漏检区域，黄色框表示模型关注较低区域。

热力图中主要分为红色和淡黄色区域，红色区域表示目标特征对热力图贡献较大，淡黄色区域表示目标特征对热力图贡献较小。由图 13b 可知，存在多处小目标杂草和甘薯苗遮挡区域杂草的漏检，表示 YOLO 11n 模型对杂草的特征提取能力较弱，易被田间环境干扰。相比之下，由图 13f 可知，SCSF－YOLO 11n 模型对小目标杂草和甘薯苗遮挡区域杂草关注度较高，能够有效减少模型漏检现象。

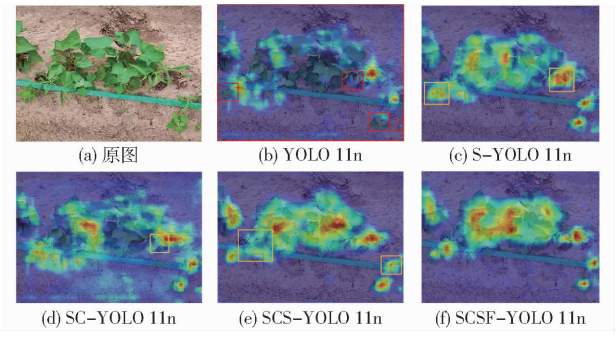


图 13　热力图可视化

Fig. 13　Heat map visualization

3.3.8　模型识别效果分析

为验证 SCSF－YOLO 11n 模型识别效果，针对分散场景、密集场景以及甘薯苗遮挡场景，各随机抽取 1 幅图像进行甘薯苗与杂草目标识别。YOLO v5n、YOLOX－tiny、YOLO v7－tiny、YOLO v8n、YOLO v9t、YOLO v10n、YOLO 11n、YOLO 12n、DINO 等不同模型识别效果，如图 14～16 所示，其中红色椭圆表示漏检杂草，蓝色椭圆表示杂草错检为甘薯苗，黑

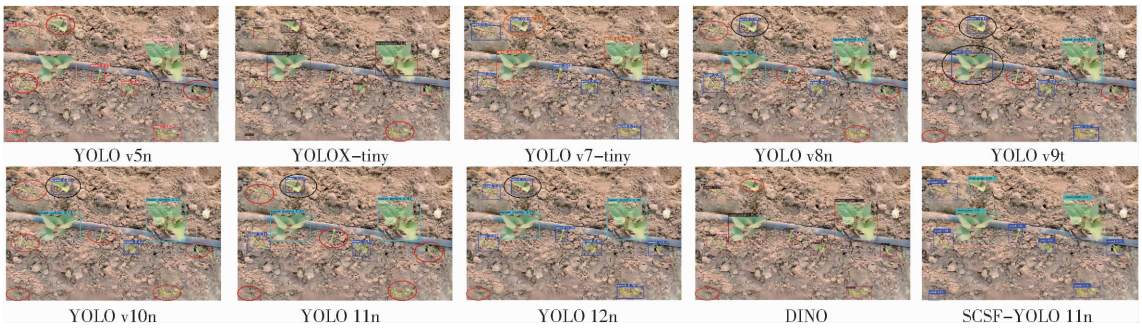


图 14 分散场景不同模型识别效果

Fig. 14 Detection effect of different models in scattered scenarios

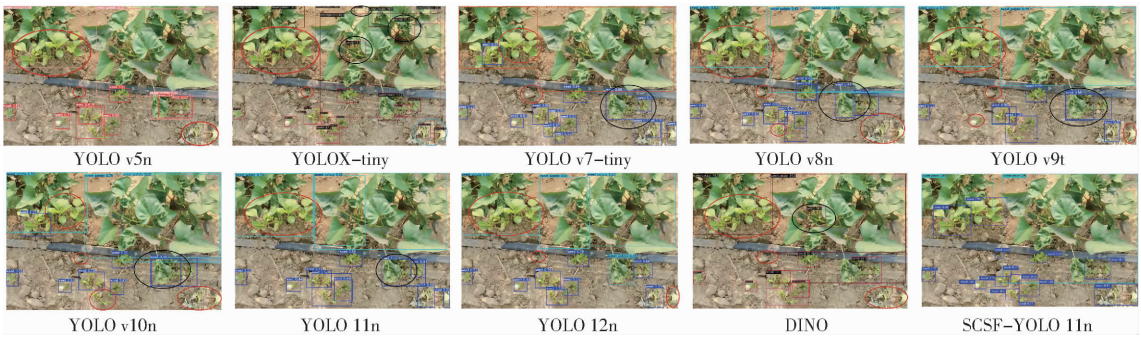


图 15 密集场景不同模型识别效果

Fig. 15 Detection effect of different models in dense scenarios

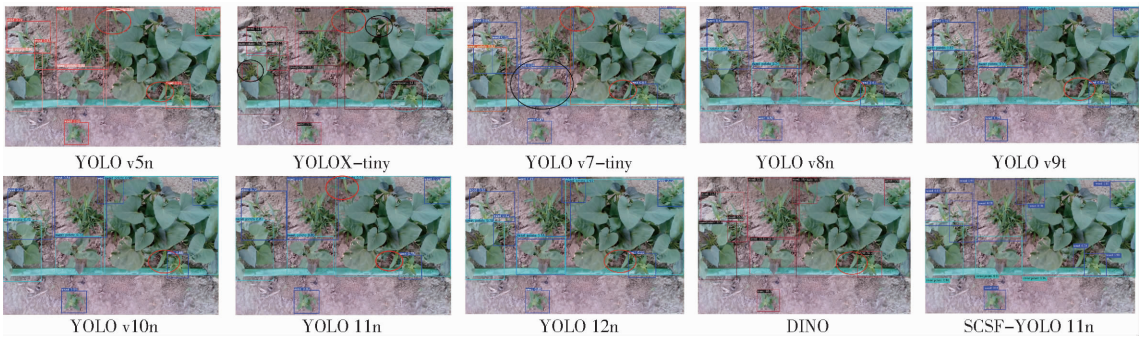


图 16 甘薯苗遮挡场景不同模型识别效果

Fig. 16 Detection effect of different models in sweet potato seedling occlusion scenarios

色椭圆表示甘薯苗错检为杂草，橙色椭圆表示 1 个目标被误识别为多类目标。

在图 14 分散场景下，部分甘薯苗与杂草在形态和颜色相似度较高，YOLO v7 - tiny 局部细节特征提取能力较弱，存在 1 个目标被误识别为 2 类目标（甘薯苗和杂草）的问题；YOLO v9t 模型引入可编程梯度信息（Programmable gradient information，PGI）解决了数据传输过程中的信息丢失问题，但甘薯苗上的细微纹理特征被模糊化，导致出现甘薯苗被误识别为杂草的现象；其他现有目标识别模型对小目标的识别能力有限，存在不同程度误检、漏检现象。相比之下，SCSF - YOLO 11n 模型在 C3K2_CG 和 SEAM 模块协助下，通过捕获输入图像的全局信息和小目标杂草边缘纹理特征信息，不仅提高了对小目标杂草区域的关注程度，还增强了对小目标杂草的定位和识别能力，有效减少了误检和漏检现象。在图 15 密集场景下，田间环境较为复杂，YOLOX - tiny、YOLO v7 - tiny、YOLO v8n、YOLO v9t、YOLO v10n、YOLO 11n、DINO 等模型存在不同程度的误检、漏检现象，而 SCSF - YOLO 11n 模型利用 C3k2_CG 模块的全局上下文信息利用能力，能有效区分甘薯幼苗与杂草并精准定位到各个目标。在图 16 甘薯苗遮挡场景下，现有目标识别模型容易忽略甘薯苗遮挡下尺度较小的杂草，出现较多漏检现象；而 SCSF - YOLO 11n 模型借助 SEAM 注意力机制进行多尺度特征融合，有效捕捉小目标杂草的局部细节特征，表现出更强的识别能力。综上可看出，SCSF - YOLO 11n 模型在分散、密集、甘薯苗遮挡场景下综合性能较优，可以识别出所有甘薯苗与杂草目标，且置信度更高、预测框位置更准确。

3.4　田间试验

3.4.1　田间试验平台

为验证 SCSF－YOLO 11n 模型在实际甘薯大田复杂环境下的识别效果，将其部署在自主设计的作物表型感知机器人上进行田间试验，试验环境涵盖晴天、阴天以及多云等天气状况。感知机器人由移动平台、控制系统、视觉系统和执行装置等组成。考虑到甘薯田间作业路径狭窄和换垄转向困难，选择四轮独立驱动四轮独立转向（4WID－4WIS）底盘作为移动平台，可进行横向平移和定点旋转提升其田间作业的灵活性。控制系统上位机选择算力为 40TOPS 的 NVIDIA Jetson Orin Nano 微型计算机，搭载 32 个 Tensor Core 的 1024 核 NVIDIA Ampere 架构 GPU，具备低功耗、高性能等特点；视觉系统通过 Intel RealSense D435i 型深度相机获取的甘薯苗和杂草图像坐标信息，通过张正友标定法建立二维图像坐标与三维世界坐标系的联系，将其转化为 Delta 机械臂下的空间坐标。控制系统下位机选择 ArduinoMega2560，在接收到 Jetson Orin Nano 发送的坐标指令后，通过求解正逆向运动学方程得到驱动电机旋转角度，从而实现 Delta 机械臂精准控制。田间试验如图 17 所示。

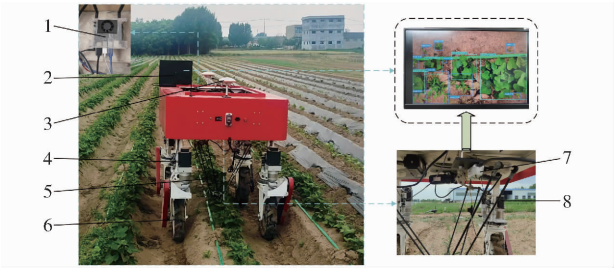


图 17　田间试验

Fig. 17　Field experiment

1. NVIDIA Jetson Orin Nano　2. 显示屏　3. RTK－GPS　4. 行走电机　5. Delta 并联机械臂　6. 驱动轮　7. Intel RealSense D435i 型深度相机　8. 转向电机

3.4.2　边缘设备部署

为实现 SCSF－YOLO 11n 模型在 Jetson Orin Nano 上的高效部署，选择 TensorRT 作为推理框架，有效降低计算资源消耗。在部署过程中，先将训练好 best. pt 导出为 ONNX 文件，并启用 FP16 精度模式以降低设备内存占用和提高模型推理速度。随后，对模型进行节点优化和图结构简化，并将优化后文件传输至 Jetson Orin Nano。最后，基于 TensorRT8. 6. 1 搭建 engine 推理引擎，充分利用 GPU 并行计算能力，实现低延迟田间杂草实时识别。

3.4.3　模型性能验证

田间试验时，机器人工作速度设为 0. 4 m/s，共获得 345 幅样本图像。采取 Python 计数代码，杂草 741 株，甘薯苗 478 株。晴天杂草置信度为 0. 7 ~ 0. 9，阴天杂草置信度大多为 0. 8 以上，多云杂草置信度为 0. 6 ~ 0. 8，表明不同天气状况下光照强度差异对于模型杂草识别性能会有影响。mAP_{50} 为 94. 6%，略低于室内模型训练结果，原因在于田垄地势变化造成相机抖动，影响图像质量和识别效果。为评估改进前后模型在边缘移动设备上的识别速度，随机抽取 40 幅样本图像用于测定推理时间，如图 18 所示。由图 18 可看出，YOLO 11n 模型单幅图像平均推理时间为 33. 03 ms，SCSF－YOLO 11n 模型单幅图像平均推理时间为 33. 26 ms，二者推理时间差异很小，均可满足算力资源受限的移动设备实时性能需求，但改进模型识别精度更优。

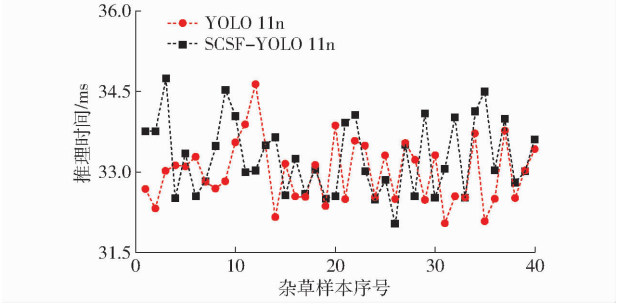
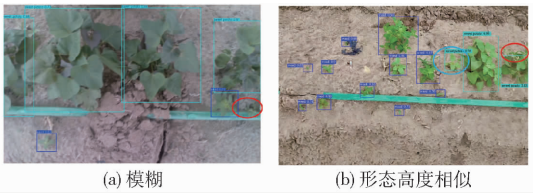


图 18　改进前后模型在边缘设备上的推理时间

Fig. 18　Inference time of models before and after improvement on edge devices

在试验过程中机器人处于移动状态且田间地势多变，导致获取的部分图像出现模糊。模糊图像杂草叶片的边缘轮廓和纹理信息被大幅弱化，导致模型易出现杂草漏检，如图 19a 所示。对于甘薯苗与杂草形态高度相似场景，采用 C3k2_CG 模块和 SEAM 注意力机制增强了模型多尺度特征提取能力，能够识别出大部分小目标杂草，仅个别杂草存在漏检、误检，如图 19b 所示，其原因在于为适应边缘设备需求牺牲了部分特征提取容量，致使模型在该场景下鲁棒性不足。后续将进一步探索模型轻量化与极端场景鲁棒性之间的平衡，提升模型在实际复杂田间环境下的泛化性，以便更好地为甘薯大田除草机器人研发提供更可靠的技术支持。



(a) 模糊　　(b) 形态高度相似

图 19　部分田间场景失败案例

Fig. 19　Failure cases in field scenes

4 结论

(1)提出了一种兼顾精度和速度的甘薯苗与杂草识别轻量化模型(SCSF - YOLO 11n)。通过引入 ShuffleNetV2 轻量化网络重构 YOLO 11n 的主干网络 CSP - Darknet,减少模型的参数量和浮点运算量;采用 CGBlock 替换 C3k2 模块的瓶颈结构完成颈部网络设计,利用更丰富的上下文信息来提高模型对于小目标杂草和甘薯苗关键特征的提取能力;在头部网络中融合 SEAM 注意力机制,提升模型对于遮挡区域杂草的识别能力;引入 Focaler - MPDIoU 作为边界框损失函数,在保证回归速度的同时加速模型网络收敛。

(2) SCSF - YOLO 11n 模型甘薯苗 AP、杂草 AP、mAP_{50} 分别为 97.3%、93.5%、95.4%,相较 YOLO 11n 分别提升 2.2、2.3、2.3 个百分点;改进后模型参数量为 1.4×10^6、内存占用量为 3.2 MB、浮点运算量为 3.1×10^9,帧率为 112.6 f/s。该模型部署在 NVIDIA Jetson Orin Nano 上的田间试验结果表明,mAP_{50} 为 94.6%,单幅图像平均推理时间为 33.26 ms,满足甘薯大田复杂环境移动设备部署和实时性需求,可为后续甘薯大田除草机器人产业化应用提供技术支撑。

参 考 文 献

[1] Luo Z, Yao Z, Yang Y, et al. Genetic fingerprint construction and genetic diversity analysis of sweet potato (*Ipomoea batatas*) germplasm resources[J]. BMC Plant Biology, 2023, 23(1): 355.

[2] Monteiro A, Santos S. Sustainable approach to weed management: the role of precision weed management[J]. Agronomy, 2022, 12(1): 118.

[3] 潘仁秀. 农田杂草的识别与防除技术[J]. 农业灾害研究, 2020, 10(3): 22 - 23, 29.
Pan Renxiu. Identification and control of farmland weeds[J]. Journal of Agricultural Catastrophology, 2020, 10(3):22 - 23, 29. (in Chinese)

[4] 刘亚军, 王文静, 李敏, 等. 不同除草剂减量与助剂复配对杂草防效、甘薯生长发育和产质量的影响[J]. 河南农业科学, 2023, 52(3): 91 - 100.
Liu Yajun, Wang Wenjing, Li Min, et al, The effect of different herbicide reduction and auxiliary compound matching on weed prevention, sweet potato growth and development and yield quality[J]. Henan Agricultural Science, 2023, 52(3): 91 - 100. (in Chinese)

[5] 杨立富. 农业机械设备的智能化技术发展现状与方向[J]. 农业工程技术, 2023, 43(17): 60 - 61.
Yang Lifu. The development status and direction of intelligent technology of agricultural machinery and equipment[J]. Agricultural Engineering Technology, 2023, 43(17): 60 - 61. (in Chinese)

[6] 樊湘鹏, 周建平, 许燕, 等. 基于优化 Faster R - CNN 的棉花苗期杂草识别与定位[J]. 农业机械学报, 2021, 52(5): 26 - 34.
Fan Xiangpeng, Zhou Jianping, Xu Yan, et al. Identification and localization of weeds based on optimized Faster R - CNN in cotton seedling stage[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 26 - 34. (in Chinese)

[7] Das A, Yang Y, Subburaj V H. CB - YOLO v7: a modified YOLO v7 approach for accurate weed detection in complex UAV imagery from cotton fields[J]. AgriEngineering, 2026, 8(6): 235.

[8] García-Navarrete O L, Camacho-Tamayo J H, Bregon A B, et al. Performance analysis of real-time detection transformer and you only look once models for weed detection in maize cultivation[J]. Agronomy, 2025, 15 (4): 796.

[9] Cui J, Zhang X, Zhang J, et al. Weed identification in soybean seedling stage based on UAV images and Faster R - CNN[J]. Computers and Electronics in Agriculture, 2024, 227: 109533.

[10] 齐江涛,吕明阳,王硕,等. 基于改进 YOLO v8n 的田间玉米杂草快速识别方法[J]. 农业机械学报,2026,57(10): 275 - 286.
Qi Jiangtao, Lü Mingyang, Wang Shuo, et al. Corn and weeds faster identification method based on improved YOLO v8n in field[J]. Transactions of the Chinese Society for Agricultural Machinery, 2026, 57(10): 275 - 286. (in Chinese)

[11] 宁姗, 赵伟龙, 乔金友, 等. 基于改进 YOLO 11n 的轻量化大豆田间杂草识别研究[J]. 农业机械学报, 2026, 57(7): 326 - 336.
Ning Shan, Zhao Weilong, Qiao Jinyou, et al. Lightweight weed detection method in soybean field based on improved YOLO 11n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2026, 57(7): 326 - 336. (in Chinese)

[12] 杨森森, 张昊, 兴陆, 等. 改进 MobileViT 网络识别轻量化田间杂草[J]. 农业工程学报, 2023, 39(9): 152 - 160.
Yang Sensen, Zhang Hao, Xing Lu, et al. Improve MobileViT network to identify lightweight field weeds[J]. Transactions of the CSAE, 2023, 39(9): 152 - 160. (in Chinese)

[13] 冀汶莉, 刘洲, 邢海花. 基于 YOLO v5 的农田杂草识别轻量化方法研究[J]. 农业机械学报, 2024, 55(1): 212 - 222, 293.
Ji Wenli, Liu Zhou, Xing Haihua. Lightweight method for identifying farmland weeds based on YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 212 - 222,293. (in Chinese)

[14] Wang Y, Yao Q, Kwok J T, at al. Generalizing from a few examples: a survey on few-shot learning[J]. ACM Computing Surveys, 2020, 53(3): 1 - 24.

[15] Khanam R, Hussain M. YOLO 11: an overview of the key architectural enhancements[PP/OL]. V1. arXiv (2024-10-23)[2026-02-12]. https://doi.org/10.48550/arXiv.2410.17725.

[16] 韩科立, 王振坤, 余永峰, 等. 基于改进 YOLO 11n 模型的棉花田间复杂环境障碍物检测方法[J]. 农业机械学报, 2025, 56(5): 111-120.
Han Keli, Wang Zhenkun, Yu Yongfeng, et al. Obstacle detection method for complex cotton field environments based on improved YOLO 11n model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 111-120. (in Chinese)

[17] Zhang Xiangyu, Zhou Xinyu, Lin Mengxiao, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6848-6856.

[18] Ma Ningning, Zhang Xiangyu, Zheng Haitao, et al. Shufflenet: practical guidelines for efficient CNN architecture design[C]// Proceedings of European Conference on Computer Vision. ECCV, 2018: 116-131.

[19] 彭红星, 何慧君, 高宗梅, 等. 基于改进 ShuffleNetV2 模型的荔枝病虫害识别方法[J]. 农业机械学报, 2022, 53(12): 290-300.
Peng Hongxing, He Huijun, Gao Zongmei, et al. Litchi diseases and insect pests identification method based on improved ShuffleNetV2 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(12): 290-300. (in Chinese)

[20] Wu T, Tang S, Zhang R, et al. CGNet: a light-weight context guided network for semantic segmentation[J]. IEEE Transactions on Image Processing, 2020, 30: 1169-1179.

[21] Yang L, Zhang T, Zhou S, et al. AAB-YOLO: an improved YOLO 11 network for apple detection in natural environments[J]. Agriculture, 2025, 15(8):836.

[22] Yu Z, Huang H, Chen W, et al. YOLO-FaceV2: a scale and occlusion aware face detector[J]. Pattern Recognition, 2024, 155: 110714.

[23] Ma S, Xu Y. Mpdiou: a loss for efficient and accurate bounding box regression[PP/OL]. V1. arXiv (2023-07-14)[2026-02-17]. https://doi.org/10.48550/arXiv.2307.07662.

[24] Zhang H, Zhang S. Focaler-iou: more focused intersection over union loss[PP/OL]. V1. arXiv (2024-01-19)[2026-02-17]. https://doi.org/10.48550/arXiv.2401.10525.

[25] 刘祥, 项若雪, 班成龙, 等. 基于改进 YOLO v8 的轻量化棉铃识别模型与产量预测方法研究[J]. 农业机械学报, 2025, 56(5): 130-140.
Liu Xiang, Xiang Ruoxue, Ban Chenglong, et al. Research on lightweight cotton bell recognition model and yield predictive method based on improved YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 130-140. (in Chinese)

[26] Selvaraju R R, Cogswell M, Das A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 618-626.