

基于实例分割的轻量化玉米作物行实时检测方法

张艳超¹ 陈宇杰¹ 傅霞萍¹ 唐伟² 杨永杰² 陆永良²

(1. 浙江理工大学信息与工程学院, 杭州 310018; 2. 中国水稻研究所水稻生物学国家重点实验室, 杭州 311400)

摘要: 田间机器人是当下研究热点, 机器视觉导航以成本低廉、效果可靠成为农业机器人全自主田间导航重要发展方向。在机器视觉导航线提取方法中, 基于图形学的方法存在适用性差、准确率低、易受噪声干扰等缺点, 基于目标检测方法受制于作物重叠, 识别难度大, 基于语义分割的方法检测速率慢、实时性差、易出现作物行缺失的情况。针对上述问题, 以玉米田间除草机器人视觉导航为研究对象, 本研究提出应用实例分割算法 (Lightweight corn row instance segmentation, LCR-IS) 玉米作物行检测方法展开试验工作。模型融合轻量化主干网络, 引入了边缘注意力融合模块 (C2f edge EMA module, C2f-EEMA) 加强特征融合能力, 同时利用参数共享机制、重参数化原理结合集成卷积模块 (Conv_GN) 设计了轻量化实例分割检测头 (Rep shared-convolutional segmentation head, RSCS)。利用 LCR-IS 模型完成作物行提取后, 通过最小欧氏距离确定作物行中心位置, 最后利用最小二乘法拟合作物行中心线为导航线。试验结果表明, LCR-IS 模型在精度方面, F1 分数为 95.3%、平均交并比为 86.1%, 在导航精度方面平均角度误差仅有 1.17°、距离误差为 6.10 像素, 优于基于 YOLO v8-seg、ESNet、VGG-UNet、Moblie-UNet、Mobile-DeeplabV3+ 的导航线提取方法。LCR-IS 模型在边缘设备 Jetson Nano 上推理速度为 18.6 f/s, 且无需高精度数据标注即可实现较为准确的导航线提取任务。研究结果能满足玉米田间自动除草机器人的实际需求, 在农业机器人实时视觉导航领域具有重要的应用前景。

关键词: 作物行检测; 田间导航; 轻量化; 实例分割; 边缘设备; 农业机器人

中图分类号: TP242.6; S24 文献标识码: A 文章编号: 1000-1298(2026)14-0255-12 **OSID:** 

Real-time Detection of Lightweight Corn Crop Rows Based on Instance Segmentation

Zhang Yanchao¹ Chen Yujie¹ Fu Xiaping¹ Tang Wei² Yang Yongjie² Lu Yongliang²

(1. School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China

2. State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou 311400, China)

Abstract: Field robots are a current research hotspot, and machine vision navigation has emerged as a key development direction for fully autonomous field navigation in agricultural robots due to its low cost and reliable performance. Among machine vision navigation line extraction methods, graphics-based approaches suffer from poor applicability, low accuracy, and susceptibility to noise interference. Object detection-based methods face challenges in crop overlap recognition, while semantic segmentation-based methods exhibit slow detection speeds, poor real-time performance, and frequent crop row omission. To address these challenges, focusing on visual navigation for corn field weeding robots, an experimental approach was proposed by using the lightweight corn row instance segmentation (LCR-IS) algorithm for corn row detection. The model integrated a lightweight backbone network and incorporated the C2f-EEMA (C2f edge EMA module) edge attention fusion module to enhance feature fusion capabilities. Simultaneously, it employed a parameter sharing mechanism and the principle of reparameterization combined with an integrated convolutional module (Conv_GN) to design the lightweight instance segmentation detection head (Rep shared-convolutional segmentation head, RSCS). After extracting crop rows using the LCR-IS model, the center positions were determined via minimum Euclidean distance.

收稿日期: 2025-03-24 修回日期: 2025-05-13

基金项目: 国家自然科学基金项目 (61905219)、浙江省科技合作计划项目 (2024SNJF071)、中国农业科研系统项目 (CARS-01) 和国家重点研发计划项目 (2023YFD140110)

作者简介: 张艳超 (1988—), 男, 副教授, 博士, 主要从事智能农机研究, E-mail: yczhang@zstu.edu.cn

通信作者: 傅霞萍 (1980—), 女, 教授, 博士, 主要从事农产品和食品无损检测技术研究, E-mail: fuxp@zstu.edu.cn

Finally, the centerlines were fitted as navigation lines by using least squares. The model achieved an F1 score of 95.3% and an average intersection-over-union (IoU) of 86.1%. The extracted predicted navigation lines exhibited an average angular error of only 1.17° and a distance error of 6.10 pixel compared with actual navigation lines, outperforming navigation line extraction methods based on YOLO v8-seg, ESNNet, VGG-UNet, Mobile-UNet, and Mobile-DeeplabV3+. The LCR-IS model achieved an inference speed of 18.6 f/s on the Jetson Nano edge device and enabled relatively accurate navigation line extraction without requiring high-precision data annotation. The research results met the practical needs of automated corn field weeding robots and held significant application potential in the field of real-time visual navigation for agricultural robots.

Key words: crop row detection; field navigation; lightweight; instance segmentation; edge devices; agricultural robotics

0 引言

随着全球人口的持续增长,玉米作为全球粮食安全的基石,种植面积和产量的重要性日益凸显^[1-2]。然而,玉米种植过程面临的杂草管理问题仍然是农业生产中的严峻挑战^[3-4]。传统的杂草管理方式主要依赖人工或化学除草剂,不仅效率低,还可能对环境造成负面影响^[5-6]。近年来,为解决这一问题自主除草机器人提供了可持续的替代方案^[7-8]。然而,其实际应用效果高度依赖于机器人在复杂玉米田环境的精确导航能力。

在农业机器人导航领域,传统的导航方式如全球导航卫星系统(GNSS)和激光雷达(LiDAR)在复杂地形和高遮挡环境下存在信号不稳定、成本高等问题^[9-10]。相比之下,基于计算机视觉的导航技术因成本低、适应性强、抗干扰能力好等优势,逐渐成为农业机器人导航的研究热点^[11]。早期的导航线提取方法主要利用机器视觉技术,通常遵循特征图获取、特征点提取和导航线拟合3步流程。Ma等^[12]通过动态阈值分割获得二值图像,利用多线聚类算法执行作物行检测,对移栽6、20、35 d的作物行检测准确率分别达96.79%、90.82%、84.15%。Zhang等^[13]提出了一种基于SUSAN(Smallest univalue segment assimilating nucleus)角点和改进顺序聚类算法的除草机器人导航方法,可用于现场作业。曾宏伟等^[14]基于区域生长算法获取作物已收割和未收割边缘线,利用最小二乘法获取导航线,航向角平均误差小于1.21°。然而,传统计算机视觉方法在面对复杂的田间环境时,如光照变化、杂草干扰和作物行不规则性等问题,往往难以满足实时性和准确性的要求^[15-16]。

近年来,深度学习技术的快速发展为农业机器人视觉导航提供了新的思路^[17]。基于深度学习的目标检测和图像分割算法能够有效识别和分割作物行,从而为导航线提取提供更为可靠的依据。Wang

等^[18]基于改进CS-YOLO v5识别水稻秧苗,随后基于初始聚类和顶部外点消除进行秧苗条带线提取,角度偏差不超过1°。Liu等^[19]提出了一种使用增强型YOLO v5算法提取菠萝田视觉导航线的新方法,行识别准确率、导航线提取准确率和角度误差平均值分别为89.29%、85.74%和3.54°。Diao等^[20]提出了一种基于改进YOLO v8s网络的玉米喷洒机器人导航线提取算法,实现了作物行中心线平均拟合时间45 ms和平均角度误差0.63°。然而基于目标检测的方法,不适用于作物种植较为密集的情况,密集作物容易出现漏检、误检等现象,影响作物特征点提取,导致导航线提取失败。

相较于目标检测方法,语义分割方法不易受作物类别与生长情况的干扰,目前较多应用在田垄导航线检测、作物行间道路检测等领域。苏童等^[21]提出了一种改进Fast-SCNN语义分割模型用于中后期玉米行间路径导航线检测算法,实现了多场景平均角度误差1.15°~3.16°,平均距离误差1.89~3.41像素。侯文慧等^[22]提出了一种轻量化U-Net语义分割模型用于果园垄间路径识别方法,在草莓、蓝莓、葡萄果园场景平均交并比分别为98.06%、97.36%、98.50%。陈子文等^[23]提出了一种轻量化Mobile-Unet语义分割模型用于花椒田场景分割,像素准确率、类别平均像素准确率、平均交并比分别为91.15%、83.34%、70.51%,基于道路和主干的导航线提取航向偏差分别为2.6°和6.7°。俞高红等^[24]引入MobilNetV2结合CA注意力机制改进DeepLabV3+用于提取无作物田垄导航线,改进模型的平均精度均值和平均交并比分别为96.27%和93.18%,在不同田垄环境下最大角度误差为1.2°,最大距离误差为9像素。然而,基于语义分割方法大多依赖于高计算资源,难以在资源受限的嵌入式设备上实时运行,且多用于田垄或单一作物行的航线提取,难以适应复杂多变的田间环境^[25],易出现断行的情况。Yu等^[26]围绕语义分割方法在边缘智

能设备上用于田间农业机器人导航线路提取开展研究,结果表明语义分割方法难以满足边缘智能农业机器人应用。

而在田间机器人应用中,将深度学习算法部署在边缘设备上是实现高效、智能农业自动化的关键,也只有在边缘设备上实现部署,深度学习算法才能有效地为田间机器人赋能。实例分割方法结合了目标检测和语义分割的优势,能够同时预测目标的边界框和分割掩码,具有解决上述问题的潜力,与语义分割仅关注图像中像素信息不同,实例分割不仅能区分语义类别,还能进一步区分同一类别中不同实例,从而提供更丰富的细节信息。与目标检测方法相比,实例分割不会受行间多作物重叠的干扰,具有更好的动态适应性。为此,本研究提出一种新型实例分割算法,用于田间导航的作物行提取并验证其有效性,探索面向田间机器人的轻量化实例分割算法并验证其边缘部署效果。

1 材料与方法

1.1 数据集获取与构建

采用多源数据集与迁移学习策略开展模型训练。数据采集方面,基于 Husky 移动机器人平台搭载的 Intel Realsense D435i 型 RGB 相机,在中国水稻研究所(杭州)2 号试验田完成玉米作物行视频采集,经帧采样处理获得田间实景图像 2 400 幅(分辨率为 1 920 像素×1 080 像素)。为进一步丰富数据多样性,从 Roboflow 平台获取 8 241 幅农业机器人采集的玉米田图像构建基础数据集。针对模型训练效率优化需求,采用渐进式迁移学习方案:首先基于原始数据集(11 241 幅,未经过数据增强)训练获得预训练权重,然后随机抽取 1 450 幅典型样本并统一缩放至 1 024 像素×576 像素(保持原始宽高比),按比例 20: 5: 4 划分为训练集、验证集和测试集。最后对 1 450 幅图像进行数据增强及筛选获取 8 841 幅数据作为最终数据集,训练集 6 628 幅,验证集 1 328 幅,测试集 885 幅(数据集网址: [https://pan. baidu. com/s/1wIpuQD_ SJh9lyuua9I6BfQ? pwd = yp42](https://pan.baidu.com/s/1wIpuQD_SJh9lyuua9I6BfQ?pwd=yp42))。典型数据样本特征如图 1 所示,涵盖不同光照条件、作物生长期及田间遮挡场景。

数据集在 Roboflow 平台进行标注,标注过程使用多边形方框勾勒作物轮廓,确保作物的大部分都在多边形框内。考虑实际应用,仅标注靠近中心的主要作物行,这种标注策略可缩短标注时间,并将作物幼苗不规则生长模式带来的导航挑战降至最低。使用 OpenCV 对标注数据进行处理,生成图像掩码,数据标注样例如图 2 所示。

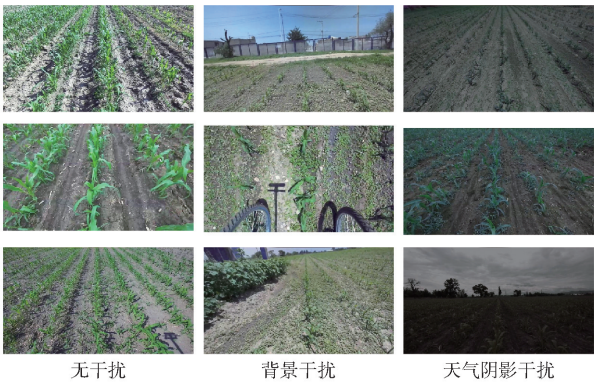


图 1 部分典型玉米作物行场景

Fig. 1 Some typical corn row scenarios



图 2 数据标注样例

Fig. 2 Sample of data labeling

1.2 数据增强

为了有效扩大样本量并提升训练样本的多样性,从而增强模型在多变复杂环境中预测准确性和泛化能力,对原始图像实施包括颜色调整、饱和度调整、色调分离、模糊处理、对比度增强以及伽马校正在内的多种数据增强措施,旨在模拟不同的光照条件、色彩变化和图像质量,使模型能够适应更广泛的实际应用场景。增强效果如图 3 所示,可明显看出增强后图像在视觉特征上更为丰富,有助于提升模型学习效果。

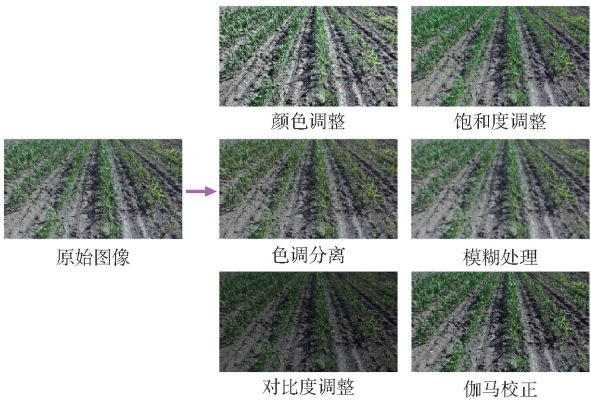


图 3 数据增强效果

Fig. 3 Data enhancement effect

在模型训练过程,采用数据在线增强技术,包括马赛克增强、混合增强、随机扰动以及随机颜色扰动 4 种先进的数据增强方式。为防止模型过拟合,确保模型泛化能力,在训练的最后 20 个轮次,关闭数据在线增强功能,因为在训练后期,模型已学习到足够特征信息,此时更重要的是稳定模型参数,避免对

训练数据中特定噪声或异常值过度敏感。以期模型能对未见数据保持较好的预测性能,同时避免对训练数据过度拟合。

1.3 实例分割模型构建

1.3.1 LCR-IS 分割模型

为构建实例分割模型,采用典型的金字塔特征结构,如图 4 所示。模型架构由特征提取的主干、特征融合的颈部以及生成输出置信度分数和掩码实例的分割头组成。

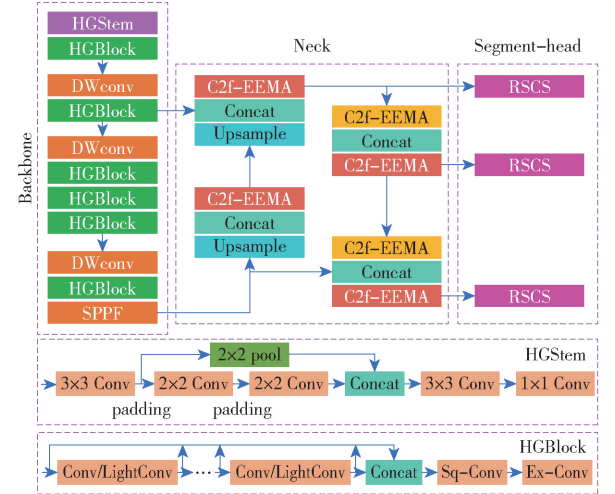


图 4 LCR-IS 网络结构

Fig. 4 Network structure of LCR-IS

对于骨干网络,经过综合评估几个成熟的轻量级网络,包括 StarNet^[27]、MobileNetV4^[28]、EfficientViT^[29]、ResNet18^[30]、FasterNet^[31] 和 HGNetV2^[32],最终选择 HGNetV2 作为骨干网络。在颈部模块,采用金字塔上采样策略,同时为了增强特征融合能力,采用自研边缘轻量化注意力融合模块 C2f-EEMA。考虑边缘设备资源有限,设计一种轻量级分割头(RSCS),可实现高效的模型推理,最终模型被命名为 LCR-IS。DWconv 为深度可分离卷积,SPPF 为空间池化模块,Concat 为拼接模块,用于多个特征图连接。模型中 HGStem 结构在进行 1 次 3×3 卷积后,将一部分特征图进行最大值池化,另一部分进行 2×2 卷积,降低通道数,填充尺寸后

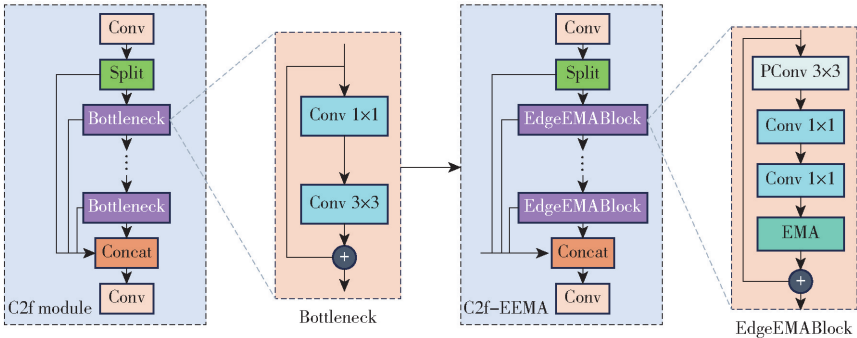


图 5 C2f-EEMA 模块结构

Fig. 5 C2f-EEMA module

再次进行 2×2 卷积,恢复通道数且减半特征图尺寸,2 路结果拼接后进行 2 次卷积,减半尺寸并缩减通道数,输出特征图尺寸为原图的 1/4,降低计算量。HGBlock 则采用多个可选取模块,通过选取标准卷积或者 LightConv 轻量级卷积方式,将每一层卷积结果和输入做拼接,增加目标特征的丰富性和多样性,然后经过 Sq-Conv(压缩卷积层)和 Ex-Conv(激励卷积层)输出特征信息。

1.3.2 边缘注意力融合模块 C2f-EEMA

在特征融合过程,考虑边缘设备的计算能力有限,在保证模型结构不变的基础上实现实时检测,可以使用类似深度可分离卷积的方法代替标准卷积,这样可以在不影响检测精度的前提下减少参数数量和计算复杂度。基于此,设计边缘注意力融合模块 C2f-EEMA,融合 C2f 模块、新式 PConv 与 EMA 注意力机制,能够捕获更丰富的特征信息又避免了复杂计算量,适用于边缘设备,C2f-EEMA 模块结构如图 5 所示。

PConv 在用于空间特征提取工作中,仅应用于选定的输入通道,其余通道保持不变。对于连续或有规律的内存访问,第 1 个或最后 1 个连续的 C_p 通道被视为整个特征映射的计算代表。为保持通用性,假设输入和输出特征映射的通道数相等,因此,PConv 的浮点运算量 L_F (FLOP) 计算式为

$$L_F = h w k^2 C_p^2 \tag{1}$$

式中 h ——通道高度
 w ——通道宽度
 C_p ——选定的连续网络通道
 k ——滤波器窗口尺寸

在典型偏置比 $r = C_p / C = \frac{1}{4}$ (C 为输入的总通道数) 下,PConv 浮点运算量仅相当于标准 Conv 的 $\frac{1}{16}$ 。此外,PConv 内存占用量 L_N 计算式为

$$L_N = h w 2 C_p + k^2 C_p^2 \approx h w 2 C_p \tag{2}$$

与普通卷积相比,PConv 内存占用量仅为普通

卷积的一半。

EMA(Efficient multi-scale attention)是一种用于深度学习模型中的注意力机制,旨在提高模型对不同尺度特征的捕捉能力,从而增强模型性能。EMA 注意力机制适用于处理具有多尺度特征

列数据,结构如图 6 所示。EMA 注意力机制通常需要较少的计算资源,在资源受限的环境中也能有效运行。因此将 EMA 注意力机制添加到 C2f – EEMA 模块中,旨在使用较少计算资源的基础上,实现更好的特征融合。

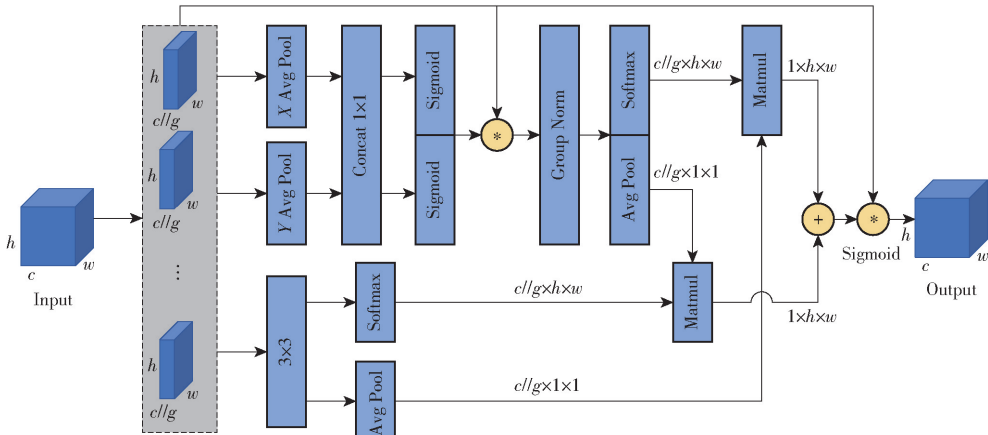


图 6 EMA 工作原理
Fig. 6 Working principle of EMA

在 EMA 模块中,输入被分组后通过不同的分支进行处理,1 个分支进行一维水平 X Avg Pool 和垂直 Y Avg Pool 全局池化,另 1 个分支通过 Conv 3×3 进行特征提取。2 个分支的输出特征通过 Sigmoid 函数和归一化操作进行调制,再通过跨维度交互模块合并,以捕捉像素级的成对关系。最终经过 Sigmoid 调节,输出特征映射以增强或减弱原始输入特征,从而到最终输出。

1.3.3 RSCS 实例分割检测头

实例分割检测头 RSCS 模块结构如图 7 所示,在 P3、P4、P5 层的卷积过程进行参数共享(即使用同一权重),可明显减少参数量,使模型

明显更轻便,更适合资源有限的设备。然而,共享参数可能会限制模型的表现力,不同特征往往需要不同卷积核才能有效捕捉复杂的特征。为减轻共享卷积的潜在缺点,引入可重新设置参数的卷积模块(Diverse branch block, DBB)。通过额外可学习参数,使网络能够更有效地提取特征,并可弥补模型轻量化设计造成的精度损失。可重参数卷积提高了参数利用率,但不会影响推理速度,因为这一阶段功能上等同于标准卷积。为解决各检测头目标尺度不一致的问题,在共享卷积的同时还添加了 1 个尺度层 Proto,以规范特征尺度。

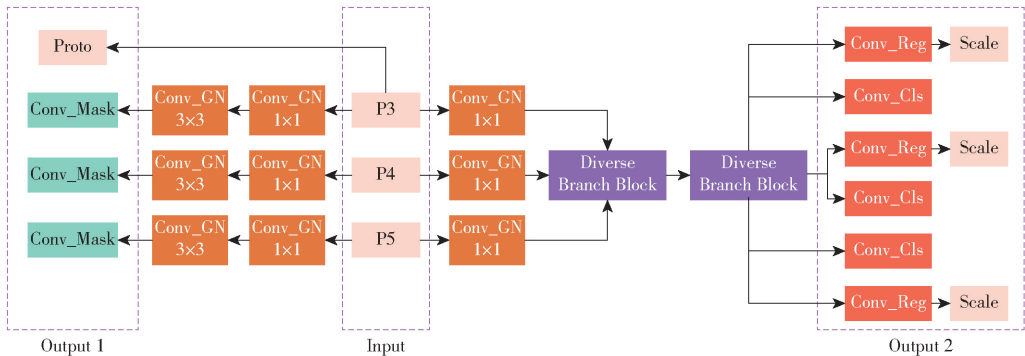


图 7 RSCS 检测头
Fig. 7 RSCS segmentation detection heads

Conv_GN 集成卷积模块是在标准卷积基础上加入组归一化操作(GroupNorm)与激活函数,能提供更多灵活性和方便性。激活函数可以通过设置进行替换,也可选择不激活,默认选择平滑的 SiLU 激活函数,使网络能学习更复杂的特征表示。选取

组归一化操作通过在组内进行归一化,帮助模型学习到更加区分性的特征,对检测头中特征学习尤其重要。且组归一化操作更易并行化,可提高训练效率,同时有助于缓解深层网络中的梯度消失或梯度爆炸问题。集成卷积模块 Conv_GN 结构如图 8 所

示。

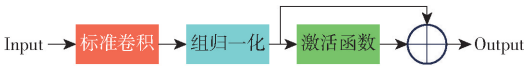


图 8 Conv_GN 模块结构
Fig. 8 Conv_GN module

可重参数卷积模块 DBB 结构如图 9 所示。在训练阶段,模块设计为 1 个类似 Inception 的结构,采用多分支架构丰富卷积块的特征表示能力,包括平均池化和多尺度卷积等结构。在推理之前,通过可重参数化技术将多分支结构融合为单一主分支,在不增加推理时间的前提下,提高模型性能。引入批量归一化(Batch normalization, BN)以稳定训练过程,利用分支求和串联 1×1 卷积调整特征通道数,整合多尺度卷积求和(如 ACNet)以捕获多尺度特征,应用平均池化降低特征图的维度,通过分支拼接合并来自不同层的特征图。这些方法协同作用,旨在减少模型的计算量和参数量,同时增强模型的代表能力和泛化性能。因此,DBB 在训练过程作为 1 个复杂的模块,而在推理过程,其参数被重组为 1 个更简单的卷积层,从而实现性能优化。

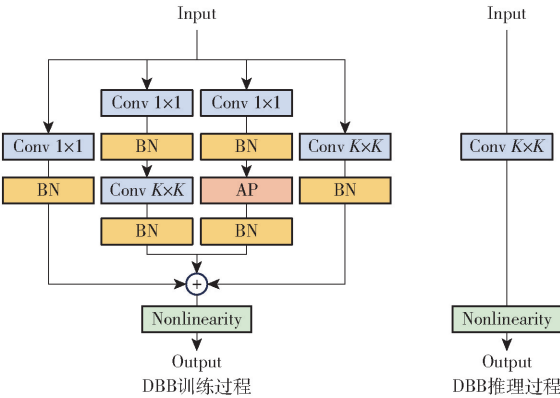


图 9 Diverse branch block 模型
Fig. 9 Diverse branch block module

1.4 损失函数

损失函数采用联合优化框架,由分类损失、边界框回归损失、掩码分割损失 3 部分共同构成,分别对应图 7 检测头的输出 2 (分类损失、边界框回归损失)和输出 1(掩码分割损失),计算式为

$$L = \lambda_{cls} L_{cls} + \lambda_{box} L_{box} + \lambda_{mask} L_{mask}$$
 (3)

式中 L ——总损失
 L_{cls} ——分类损失
 L_{box} ——边界框回归损失
 L_{mask} ——掩码分割损失
 $\lambda_{cls}、\lambda_{box}、\lambda_{mask}$ ——对应超参数权重,取 0.5、7.5 和 1

分类损失采用动态聚焦的二元交叉熵损失,解决多类别预测中前景背景不平衡问题,边界框回归

损失采用改进的 CIoU(Complete IoU)损失函数优化目标框的回归,计算式为

$$L_{box} = 1 - IoU + \frac{\rho^2(b_{pred}, b_{gt})}{l^2} + \alpha v$$
 (5)

其中 $\alpha = \frac{v}{1 - IoU + v}$ (6)

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w_{pred}}{h_{pred}} \right)^2$$
 (7)

式中 IoU ——交并比
 ρ ——预测框中心点到真实框中心点欧氏距离
 l ——最小包围框对角线长度
 $\alpha、v$ ——权重因子与长宽比惩罚项
 $h_{gt}、w_{gt}$ ——真实框长、宽
 $h_{pred}、w_{pred}$ ——预测框长、宽

掩码分割损失函数则采用二元交叉熵 BCE loss 和 Dice loss 的组合,计算式为

$$L_{mask} = \lambda_{bce} L_{bce} + \lambda_{dice} L_{dice}$$
 (8)

式中 $\lambda_{bce}、\lambda_{dice}$ ——二元交叉熵系数

1.5 作物行中心线提取方法

经训练好的模型预测出作物行后,进行二值化处理,获取灰度图像。对图像进行简单形态学操作后,再进行中心线提取。采用一种基于最短距离的航线方法,对整个作物行中心线提取,过程如图 10 所示。



图 10 作物行中心线提取过程
Fig. 10 Extraction process of crop row centerline

对处理后的二值化掩码进行拟合,获取上下端的中点;计算上下端中点与图像中心点的欧氏距离,将距离最短的 2 个中点所在的掩码确定为中心作物行,连接该行上下端中点为中心线。如图 11 所示,红线为预测的作物行中心线,蓝线为实际的作物行中心线。取 2 条中心线与图像上方交点的距离为 L_2 ,与图像下方的交点距离为 L_1 ,取 L_2 与 L_1 的平均

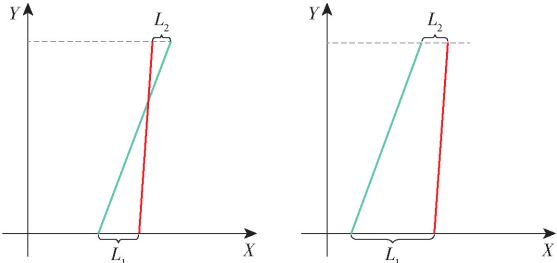


图 11 作物行中心线误差计算示意图
Fig. 11 Schematic for calculating crop row centerline error

值作为作物行中心线的像素误差,角度偏差则通过斜率转换角度作差值求得。

1.6 性能评估指标

选择准确率和推理速度作为评估模型的 2 个指标。准确率评价由分割准确率和导航线准确率组成,导航线准确率通过像素之间的差异以及航线斜率差计算,分割准确率评价采用精确率、召回率、平均交并比、F1 分数、平均精度均值指标。

2 试验与结果分析

2.1 模型训练平台及模型部署平台

模型训练在 Windows 11 上使用 PyTorch 框架进行,硬件主要配置包括 Intel Xeon E5 – 2670 CPU 和 NVIDIA GeForce RTX 4060 图形处理器。编译语言为 Python 3. 10,深度学习框架版本为 PyTorch 2. 0。模型部署则选择 Jetson Nano(英伟达公司,美国)平台,低功耗、经济实惠和适用于资源受限环境。Jetson Nano 功耗仅为 5 ~ 10 W,以最少能源支持更长运行时间,完全符合可持续和经济高效的农业解决方案对智能体现的需求,此外硬件限制与实际边

缘应用场景所面临的挑战非常相似。通过优化算法并评估在 Jetson Nano 上运行效率,以验证算法在受限边缘环境中的实际适用性。

2.2 模型训练参数设置

通过 Adam 优化算法训练模型,不使用预训练权重。初始学习率设置为 0. 01,采用余弦退火法降低学习率,最小学习率为初始学习率的 0. 01 倍。为平衡训练效率与显存占用,将每次训练的批次大小设置为 4,模型训练迭代次数为 150。模型训练过程启用数据在线增强,色调增益设置为 0. 015,饱和度增益设置为 0. 7,增益设置为 0. 4,平移增益设置为 0. 1,尺度增益设置为 0. 5,旋转度、剪切度和透视度数设置为 0,左右翻转概率设置为 0. 5。

2.3 不同主干网络分割性能试验

使用深度学习神经网络对作物行进行分割研究,获得比较高的精度效果。选取 6 种轻量级主干网络进行试验以选择模型的 Backbone, StarNet、MobileNetV4、EfficientViT、ResNet18、FasterNet、HGNetV2 6 种不同主干网络对应的训练损失曲线如图 12 所示。

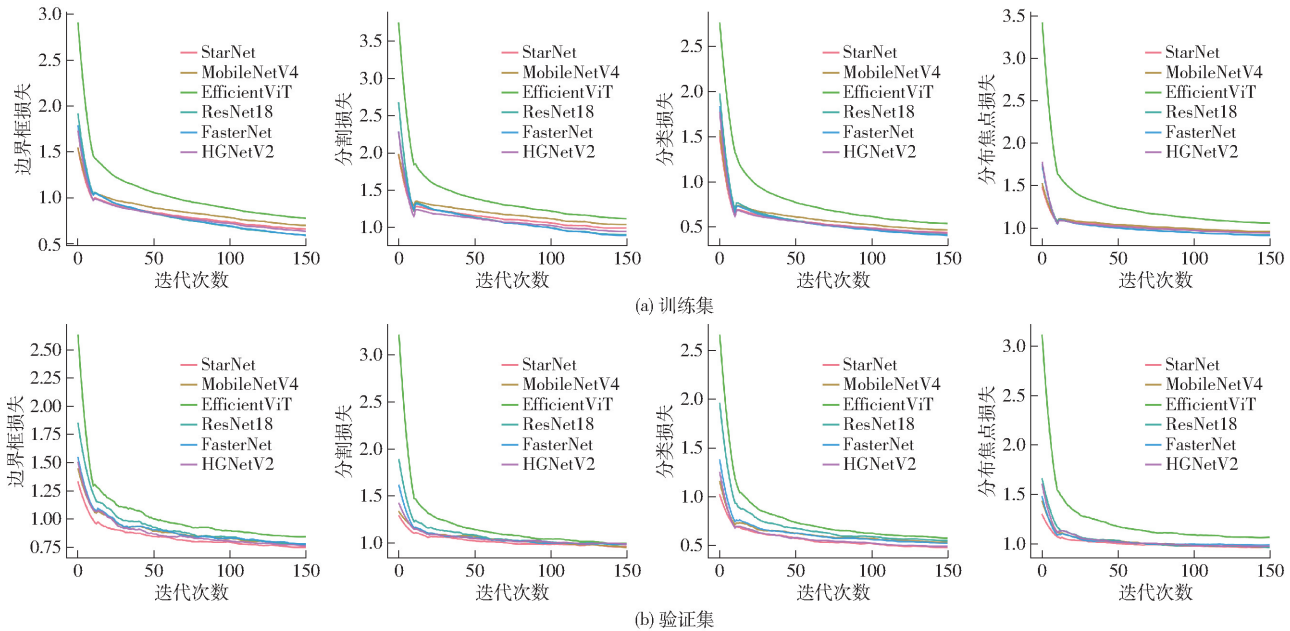


图 12 不同主干网路对应的损失函数曲线

Fig. 12 Loss function curves of different backbone networks

从图 12 可看出,随着迭代次数增加,不同主干网络对应的损失函数曲线都趋于平稳,且无特别明显曲线波动,说明训练策略与模型结构的稳定性。模型分割性能试验结果如表 1 所示。从表 1 可看出,不同主干网络都有着较好的性能,HGNetV2 以突出优势领先于其他主干网络。从实例检测方面来看,HGNetV2 的 mAP50 – 95 为 67. 5%,其次是 StarNet 与 FasterNet,说明 HGNetV2 主干网络对作物行定位有更好的效果。HGNetV2 的 F1 分数为

表 1 不同主干网络分割性能结果

Tab. 1 Segmentation performance results for different

backbone networks					%
主干网络	精确率	召回率	F1 分数	mAP50 – 95	平均交并比
StarNet	95. 1	93. 9	94. 5	66. 1	85. 4
MobileNetV4	94. 1	93. 5	93. 8	64. 5	83. 6
EfficientViT	93. 5	93. 2	93. 3	63. 1	82. 8
ResNet18	95. 1	93. 3	94. 2	64. 6	83. 1
FasterNet	95. 3	92. 8	94. 1	65. 8	83. 1
HGNetV2	95. 9	94. 7	95. 3	67. 5	86. 1

95.3%,平均交并比为86.1%,表明在分割精度与作物行的像素区分方面具有显著优势,其次依然是 StarNet。从模型整体性能来看,HGNetV2 综合效果最佳,其次是 StarNet 主干网络,优于其他 4 种主干网络。

不同主干网络的推理结果如图 13 所示。由图 13 可知,在常见农业机械正向视角获取作物行图像推理结果中,6 种主干网络都表现出不错的分割效果,所有作物行都被完整分割出来,不存在误检现

象。分割图像包含多个场景也由此验证了模型的鲁棒性。在图 13a ~ 13c、13f 场景中,不同的主干网络分割结果并无明显差异。在图 13d、13g 场景中,ResNet18 相较于其他 5 种主干网络,在作物行的分割结果中出现了断行现象,说明 ResNet18 抗干扰能力要逊色于其他主干网络。图 13e 场景中,以 HGNetV2 分割结果最佳,其次是 StarNet 主干网络。综合来看,HGNetV2 与 StarNet 分割效果最好。

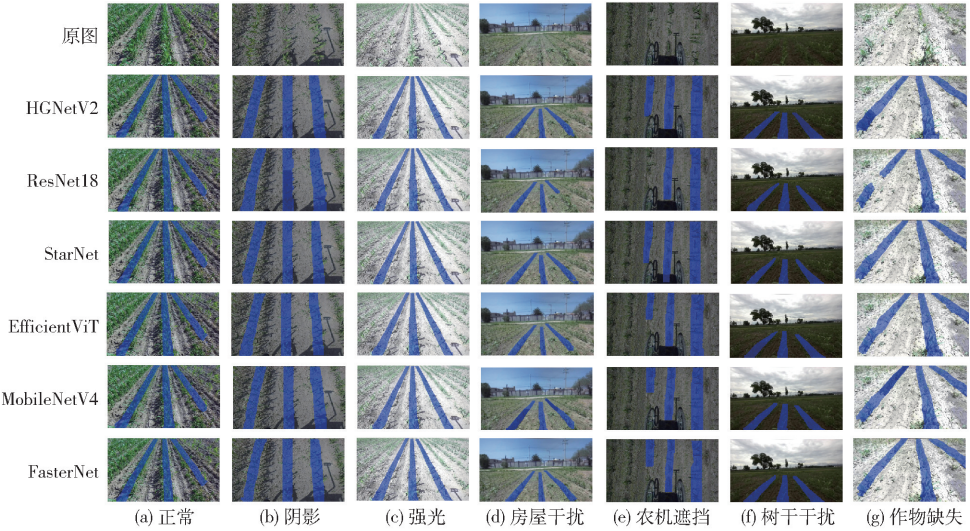


图 13 不同主干网络实际推理结果

Fig. 13 Reasoning results for different backbone structures

在边缘设备的部署效果上 CNN 网络比 ViT 模型更具有优势,这是因为许多边缘设备并没有针对 ViT 的加速部署策略。而且模型实际表现效果可能会受到边缘设备中资源限制,导致实际效果并不理想。为此在边缘设备上进行试验,以验证模型的实际部署效果,选择 Jetson Nano 边缘设备,在模型部署上,采取 Jetson 自带的 tensorrt 进行加速部署,部署结果如表 2 所示。从表 2 可看出,HGNetV2 推理速度最快,达到 18.60 f/s,满足实际实时应用需求。StarNet 尽管模型内存占用量最小,但推理速度不如 HGNetV2,说明 HGNetV2 主干网络更适于边缘设备。综合边缘设备上的推理速度和分割精度,最终确认 HGNetV2 为模型主干网络。

表 2 边缘设备上实际推理结果

Tab. 2 Actual inference results on edge devices

主干网络	内存占用量/	推理时间/	推理速度/
	MB	ms	(f·s ⁻¹)
StarNet	7.92	66.784	14.97
MobileNetV4	14.67	67.693	14.77
EfficientViT	11.26	113.196	8.83
ResNet18	60.24	105.141	9.51
FasterNet	11.57	67.953	14.72
HGNetV2	8.11	53.763	18.60

2.4 作物行中心线提取试验

在分割工作中,图像被分割成一条条完整的作物行条带掩码与无关背景,并很好地提取完整的作物条带。随机选取 250 幅图像,验证作物行中心线提取方法。以最初标定的掩码标签文件作为真实作物行用于导航线提取,得到真实作物行中心线(红色),该导航线作为真实作物行线。同时,用模型推理得出掩码结果进行同样的中心线提取操作获得预测的中心线(绿色),试验结果如图 14 所示。

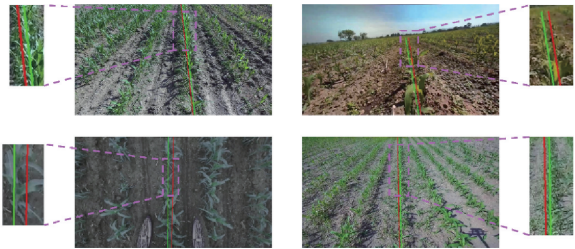


图 14 作物行中心线提取结果

Fig. 14 Extraction results of mid lines

从图 14 可看出,通过模型推理得到的预测中心线与实际作物行中心线高度吻合;然而,根据标签掩码文件提取出的真实作物行中心线(红色)与作物行中心位置存在一定偏差。经分析,该偏差主要源于掩码标注过程中的主观因素:在标注过程未严格以作物行的几何中心为对称轴,导致提取的中心线

(红色)与实际作物行中心之间存在位置和角度的偏差。为验证该判断,图 15 中引入 1 条严格贴合作物行中心的校正中心线(紫色)作为对照,结果表明,紫色中心线能够准确反映实际作物行的走向,进一步证实了标注偏差的存在。尽管存在上述标注误差,模型仍能准确提取接近真实中心线的导航线(绿色),体现了本文算法对标注不精确具有较强的鲁棒性。



图 15 标注过程添加的作物行中心线

Fig. 15 Seedling lines added by labeling process

纠正作物行线偏差后,重新进行了作物行中心线提取,如图 16 所示,红色为真实作物行中心线,绿色为模型推理出的中心线。从图 16 可看出,推理得出的中心线几乎与真实中心线重合,体现本文方法的优越性。

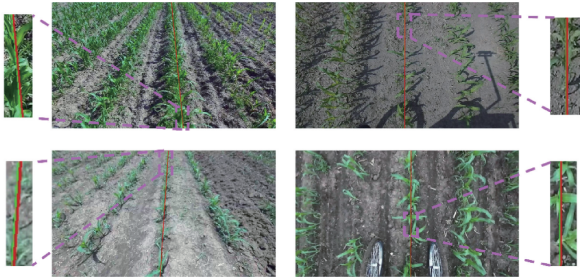


图 16 作物行中心线的最终提取结果

Fig. 16 Final extraction results of seedling lines

为了更直观精确地展示结果,统计了不同主干网络对应的角度误差与距离误差,结果如表 3 所示(图像尺寸为 1 024 像素×576 像素)。从表 3 可看出,不同主干网络得到的作物行中心线误差都很小,角度误差不超过 1.4°,距离误差最大为 7.69 像素,对应 1 024 像素×576 像素的图像,距离误差占比也仅有 0.75%。整体性能以 HGNetV2 表现最好,角度误差为 1.17°,距离误差为 6.10 像素,与分割结果

表 3 不同主干网络导航线误差

Tab. 3 Results of navigation line errors for different backbone networks

主干网络	角度误差/(°)	距离误差/像素
StarNet	1.27	7.03
MobileNetV4	1.37	7.69
EfficientViT	1.26	6.89
ResNet18	1.22	7.14
FasterNet	1.19	6.94
HGNetV2	1.17	6.10

一致,可以满足实际需求。

2.5 消融试验

为了验证 LCR – IS 模型中 C2f – EEMA、RSCS 检测头等模块的有效性,进行消融试验性能测试,试验结果如表 4 所示。从表 4 可看出,在精度提升方面,模块 RSCS 主要提升实例精度,对分割精度没影响,符合 RSCS 模块设计原理,C2f – EEMA 模块对实例精度与分割精度均有提升,2 个模块同时引用起到了更好的精度改进效果,符合设计理念。在推理速度方面,2 个模块均起到速度提升的效果,共同引入同样带来了更好的推理速度提升。

表 4 模型消融试验结果

Tab. 4 Ablation experiment results of the models

模型	mAP50 – 95/ %	MIoU/ %	推理耗时/ ms	推理速度/ (f·s ⁻¹)
HGNetV2	62.5	85.3	59.419	16.83
HGNetV2 + RSCS	65.3	85.3	57.364	17.43
HGNetV2 + C2f – EEMA	64.7	85.8	56.164	17.80
HGNetV2 + C2f – EEMA + RSCS	67.5	86.1	53.763	18.60

为了验证各模块的优越性,进行模块适应性测试,主要将 RSCS、C2f – EEMA 模块迁移到同为特征金字塔结构的 YOLO v8 – seg 模型进行消融试验,结果如表 5 所示。在 YOLO v8 – seg 模型上模块迁移试验与消融试验结果相似,模块的单独作用对模型性能均有改善,共同引入改善效果更佳。

表 5 模块迁移试验结果

Tab. 5 Results of module migration experiments

模型	mAP50 – 95/%	MIoU/%	推理耗时/ms	推理速度/(f·s ⁻¹)
YOLO v8 – seg	67.2	84.8	63.613	15.72
YOLO v8 – seg + RSCS	69.8	84.7	59.767	16.73
YOLO v8 – seg + C2f – EEMA	68.0	85.6	58.972	16.96
YOLO v8 – seg + C2f – EEMA + RSCS	70.2	85.9	57.634	17.35

2.6 不同模型分割试验

为了验证本文方法的先进性,与 YOLO v8 - seg^[33]、ESNet^[25]、VGG - UNet^[34]、Moblie - UNet^[22]、Mobile - DeeplabV3 + ^[24] 进行分割精度对比试验,结果如图 17、表 6 所示。从图 17 可看出,2 种基于实例分割模型在各种场景中均表现出不俗的分割效果,优于其他 4 种语义分割方法。语义分割方法中以 Mobile - UNet 效果最佳,但在图 17b、17c 场景中将部分空地与其他非玉米作物行也识别成了玉米作物行。其余 3 种语义分割方法,特别在图 17e 黑暗环境中,均未识别出作物行。而 2 种实例分割方法,在条带完整性优于其他 4 种语义分割方法,分割结果也更倾向于完整的平滑条带。

在精度方面,LCR - IS 实现了 F1 分数 95.3%,平均交并比 86.1%。在导航线精度方面,LCR - IS 角度误差 1.17°和距离误差 6.10 像素,优于其他 5 种分割方法,其次是 YOLO v8 - seg 方法。在推理速度方面,推理速度最快的为 LCR - IS 方法,达到 18.60 f/s。整体推理速度方面,实例分割方法远优于语义分割方法,其中最快的语义分割方法为 Mobile - DeeplabV3 + ,推理速度仅有 4.43 f/s,远不如实例分割方法。最终试验结果表明,LCR - IS 方法优于 YOLO v8 - seg、ESNet、VGG - UNet、Moblie - UNet、Mobile - DeeplabV3 + 5 种方法,基于实例分割的导航线提取方法总体优于基于语义分割的方法。

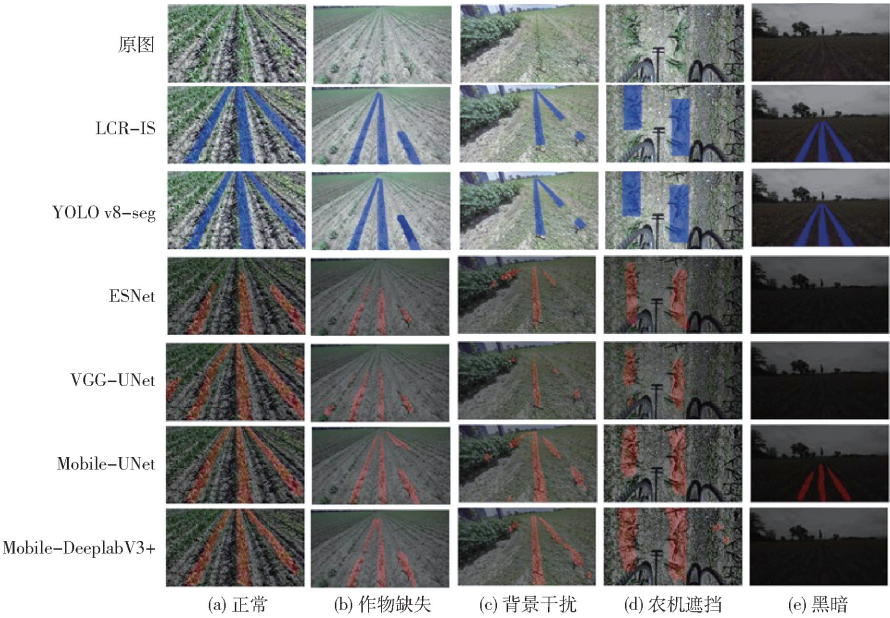


图 17 不同模型分割结果

Fig. 17 Segmentation results of different models

表 6 不同分割模型试验结果

Tab. 6 Results of different segmentation models								
模型	精确率/%	召回率/%	F1 分数/ %	MIoU/%	角度误差/ (°)	距离误差/ 像素	推理时间/ ms	推理速度/ (f·s ⁻¹)
LCR - IS	95.9	94.7	95.3	86.1	1.17	6.10	53.763	18.60
YOLO v8 - seg	96.1	93.6	94.7	84.8	1.21	6.91	63.613	15.72
ESNet	91.1	77.5	83.6	70.3	3.17	17.62	426.395	2.35
VGG - UNet	91.9	76.7	83.6	69.8	3.39	18.82	910.783	1.10
Moblie - UNet	94.5	88.5	91.4	81.4	2.23	12.08	3 537.98	0.28
Mobile - DeeplabV3 +	93.9	88.2	91.0	81.3	2.78	14.14	225.688	4.43

3 结论

(1)提出了一种基于视觉的作物行提取方法用于农业机械的导航线获取。提出的基于实例分割方法的玉米田作物行提取方法,具有较好的鲁棒性与环境适应性,解决了玉米田作物行提取难题,实现了

F1 分数 95.3% 与平均交并比 86.1%,在导航线提取上角度误差仅有 1.17°以及距离误差 6.10 像素,优于 YOLO v8 - seg、ESNet、VGG - UNet、Moblie - UNet、Mobile - DeeplabV3 + 5 种方法,可为满足农业机器人路径导航需求。

(2)提出轻量化实例分割模型融合轻量化主干

网络,引入 C2f – EEMA 边缘注意力融合模块加强特征融合能力,同时利用参数共享机制、重参数化原理结合 Conv_GN 模块设计了轻量化实例分割检测头 RSCS。且开发的 C2f – EEMA 边缘注意力融合模块、RSCS 实例分割检测头可以迁移到其他实例分割模型。

(3)将实例分割算法再边缘设备 Jetson Nano 上部署,推理速度达 18.60 f/s,可满足实际部署需求。此外该模型无需精确的标签标注,提高了模型的通用性,对提高农业机器人的视觉稳定性具有重要意义,为推进跟随作物行线的农业机器人研发提供了重要参考价值。

参 考 文 献

[1] Xiang L, Gai J, Bao Y, et al. Field-based robotic leaf angle detection and characterization of maize plants using stereo vision and deep convolutional neural networks[J]. Journal of Field Robotics, 2023, 40(5): 1034 – 1053.

[2] 张硕, 刘禹, 熊坤, 等. 基于特征工程的大田作物行中心线识别方法[J]. 农业机械学报, 2023, 54(增刊 1): 18 – 26. Zhang Shuo, Liu Yu, Xiong Kun, et al. Center line detection of field crop rows based on feature engineering[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(S1): 18 – 26. (in Chinese)

[3] Komasilovs V, Stalidzans E, Osadcuks V, et al. Specification development of robotic system for pesticide spraying in greenhouse[C] // 2013 IEEE 14th International Symposium on Computational Intelligence and Informatics (CINTI). IEEE, 2013: 453 – 457.

[4] 刘路, 潘艳娟, 陈志健, 等. 高遮挡环境下玉米植保机器人作物行间导航研究[J]. 农业机械学报, 2020, 51(10): 11 – 17. Liu Lu, Pan Yanjuan, Chen Zhijian, et al. Inter-rows navigation method for corn crop protection vehicles under high occlusion environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(10): 11 – 17. (in Chinese)

[5] Montalvo M, Pajares G, Guerrero J M, et al. Automatic detection of crop rows in maize fields with high weeds pressure[J]. Expert Systems with Applications, 2012, 39(15): 11889 – 11897.

[6] 韩长杰, 郑康, 赵学观, 等. 大田甘蓝作物行识别与对行喷雾控制系统设计与试验[J]. 农业机械学报, 2022, 53(6): 89 – 101. Han Changjie, Zheng Kang, Zhao Xueguan, et al. Design and experiment of row identification and row-oriented spray control system for field cabbage crops[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(6): 89 – 101. (in Chinese)

[7] Lippi M, Santilli M, Carpio R F, et al. An autonomous spraying robot architecture for sucker management in large-scale hazelnut orchards[J]. Journal of Field Robotics, 2024, 41(7): 2114 – 2132.

[8] 潘宇镭, 吴雨铎, 李成龙, 等. 杂交水稻制种父本倒播差插秧视觉导航线实时提取方法[J]. 农业机械学报, 2024, 55(增刊 1): 41 – 50. Pan Yulei, Wu Yuhua, Li Chenglong, et al. Real-time guideline extraction method for male parent transplanting in hybrid rice seed production[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(S1): 41 – 50. (in Chinese)

[9] Xie B, Jin Y, Faheem M, et al. Research progress of autonomous navigation technology for multi-agricultural scenes[J]. Computers and Electronics in Agriculture, 2023, 211: 107963.

[10] 江庆, 安东, 韩华宇, 等. 基于激光雷达与 RGB 相机融合的玉米作物行检测算法研究[J]. 农业机械学报, 2024, 55(10): 263 – 274. Jiang Qing, An Dong, Han Huayu, et al. Maize crop row detection algorithm based on fusion of LiDAR and RGB camera[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 263 – 274. (in Chinese)

[11] Hu K, Chen Z, Kang H, et al. 3D vision technologies for a self-developed structural external crack damage recognition robot [J]. Automation in Construction, 2024, 159: 105262.

[12] Ma Z, Tao Z, Du X, et al. Automatic detection of crop root rows in paddy fields based on straight-line clustering algorithm and supervised learning method[J]. Biosystems Engineering, 2021, 211: 63 – 76.

[13] Zhang Q, Chen M S, Li B. A visual navigation algorithm for paddy field weeding robot based on image understanding[J]. Computers and Electronics in Agriculture, 2017, 143: 66 – 78.

[14] 曾宏伟, 雷军波, 陶建峰, 等. 低对比度条件下联合收割机导航线提取方法[J]. 农业工程学报, 2020, 36(4): 18 – 25. Zeng Hongwei, Lei Junbo, Tao Jianfeng, et al. Navigation line extraction method for combine harvester under low contrast conditions[J]. Transactions of the CSAE, 2020, 36(4): 18 – 25. (in Chinese)

[15] Jiang G, Wang X, Wang Z, et al. Wheat rows detection at the early growth stage based on Hough transform and vanishing point[J]. Computers and Electronics in Agriculture, 2016, 123: 211 – 223.

[16] 饶秀勤, 朱逸航, 张延宁, 等. 基于语义分割的作物垄间导航路径识别[J]. 农业工程学报, 2021, 37(20): 179 – 186. Rao Xiuqin, Zhu Yihang, Zhang Yanning, et al. Navigation path recognition between crop ridges based on semantic segmentation[J]. Transactions of the CSAE, 2021, 37(20): 179 – 186. (in Chinese)

[17] Pang Y, Shi Y, Gao S, et al. Improved crop row detection with deep neural network for early-season maize stand count in UAV

- imagery[J]. Computers and Electronics in Agriculture, 2020, 178: 105766.
- [18] Wang S, Yu S, Zhang W, et al. The identification of straight-curved rice seedling rows for automatic row avoidance and weeding system[J]. Biosystems Engineering, 2023, 233: 47–62.
- [19] Liu T H, Zheng Y, Lai J S, et al. Extracting visual navigation line between pineapple field rows based on an enhanced YOLO v5[J]. Computers and Electronics in Agriculture, 2024, 217: 108574.
- [20] Diao Z, Guo P, Zhang B, et al. Navigation line extraction algorithm for corn spraying robot based on improved YOLO v8s network[J]. Computers and Electronics in Agriculture, 2023, 212: 108049.
- [21] 苏童, 王琳, 班超, 等. 基于语义分割的中后期玉米行间路径导航线检测[J]. 农业机械学报, 2024, 55(10): 275–285.
Su Tong, Wang Lin, Ban Chao, et al. Interrow path navigation line detection of maize in middle and late period based on semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 275–285. (in Chinese)
- [22] 侯文慧, 周传起, 程炎, 等. 基于轻量化 U-Net 网络的果园垄间路径识别方法[J]. 农业机械学报, 2024, 55(2): 16–27.
Hou Wenhui, Zhou Chuanqi, Cheng Yan, et al. Path recognition method of orchard ridges based on lightweight U-Net[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 16–27. (in Chinese)
- [23] 陈子文, 杨宇帆, 张海腾, 等. 青花椒田间场景分割与导航线提取方法[J]. 农业工程学报, 2024, 40(23): 217–226.
Chen Ziwen, Yang Yufan, Zhang Haiteng, et al. Extracting navigation line after segmenting the field scene of green peppers [J]. Transactions of the CSAE, 2024, 40(23): 217–226. (in Chinese)
- [24] 俞高红, 王一淼, 甘帅汇, 等. 改进 DeepLabV3 + 算法提取无作物田垄导航线[J]. 农业工程学报, 2024, 40(10): 168–175.
Yu Gaohong, Wang Yimiao, Gan Shuaihui, et al. Extracting the navigation lines of crop-free ridges using improved DeepLabV3 + [J]. Transactions of the CSAE, 2024, 40(10): 168–175. (in Chinese)
- [25] 李秀智, 方会敏, 朱玉垒, 等. 基于语义分割网络的植保机器人视觉伺服控制方法[J]. 农业机械学报, 2024, 55(5): 21–27, 39.
Li Xiuzhi, Fang Huimin, Zhu Yulei, et al. Visual servo control of plant protection robot based on semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(5): 21–27, 39. (in Chinese)
- [26] Yu J, Zhang J, Shu A, et al. Study of convolutional neural network-based semantic segmentation methods on edge intelligence devices for field agricultural robot navigation line extraction[J]. Computers and Electronics in Agriculture, 2023, 209: 107811.
- [27] Ma X, Dai X, Bai Y, et al. Rewrite the stars[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 5694–5703.
- [28] Qin D, Lechner C, Delakis M, et al. MobileNetV4: universal models for the mobile ecosystem[C]//European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 78–96.
- [29] Liu X, Peng H, Zheng N, et al. EfficientViT: memory efficient vision transformer with cascaded group attention[C]//Proceedings of the IEEE/CVF CVPR, 2023: 14420–14430.
- [30] Targ S, Almeida D, Lyman K. Resnet in resnet: generalizing residual architectures[C]//International Conference on Learning Representations, 2016.
- [31] Chen J, Kao S, He H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF CVPR, 2023: 12021–12031.
- [32] Zhao Y, Lv W, Xu S, et al. Detsr beat yolos on real-time object detection[C]//Proceedings of the IEEE/CVF CVPR, 2024: 16965–16974.
- [33] 应仇凯, 程泓超, 马锒宏, 等. 基于 YOLO v8-Seg 的地栽草莓采摘机器人端面视觉导航控制方法[J]. 农业机械学报, 2024, 55(1): 9–17.
Ying Qiukai, Cheng Hongchao, Ma Zenghong, et al. Ridge visual navigation control method for ground-planted strawberry picking robots based on YOLO v8-Seg algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 9–17. (in Chinese)
- [34] 李云伍, 徐俊杰, 刘得雄, 等. 基于改进空洞卷积神经网络的丘陵山区田间道路场景识别[J]. 农业工程学报, 2019, 35(7): 150–159.
Li Yunwu, Xu Junjie, Liu Dexiong, et al. Field road scene recognition in hilly regions based on improved dilated convolutional networks[J]. Transactions of the CSAE, 2019, 35(7): 150–159. (in Chinese)