

基于 ResAC – UNet 网络的玉米作物行识别与导航线提取算法研究

崔永志^{1,2} 刘阳春^{1,2} 毛文华^{1,2} 安麒麟^{1,2} 郭全峰^{1,2} 李广瑞^{1,2}
周 达^{1,2} 周白雪^{1,2} 伟利国^{1,2}

(1. 中国农业机械化科学研究院集团有限公司, 北京 100083; 2. 农业装备技术全国重点实验室, 北京 100083)

摘要: 在复杂非结构化农田环境中,导航线精确提取对农机和农业机器人实现自主作业至关重要。农业环境中普遍存在光照多变、地形起伏和杂草干扰等挑战性因素,传统图像处理方法在适应性、准确性和实时性方面表现不佳,难以满足智慧农业的视觉导航需求。针对这些问题,本文提出一种基于改进 UNet 的 ResAC – UNet 深度学习网络模型。该模型采用 ResNet – 50 网络替代传统 UNet 的编码器结构,增强特征提取能力。通过优化跳跃连接,提高了分割速度和实时响应能力。在网络的瓶颈部分引入 ASPP 模块,实现多尺度感受野建模,保持高分辨率特征的同时捕获更丰富的上下文信息。此外,模型整合 CBAM,增强对作物行边界的精确感知,有效防止了关键特征信息的丢失,进一步提升了分割质量。在分割结果的基础上,采用行锚法和 RANSAC 算法实现了高精度导航线的提取与平滑处理。将获取的前视图像进行鸟瞰图转换,消除透视效应,生成保留 ROI 的作物行俯视图。试验结果表明,ResAC – UNet 模型在精确率、平均像素精度、平均交并比和召回率方面分别达到 99.23%、95.44%、85.23% 和 94.71%,性能优于当前主流的 Segformer、DDRNet、HRNet 及 DeepLabV3+ 等分割网络,ResAC – UNet 的平均推理时间为 15.26 ms,满足智能农机视觉导航识别的实时性需求。在相机的 ROI 区域内可提取 3 条导航线,中间导航线的最大角偏差仅为 0.96°,最大像素偏差为 4.3 像素,实现了高质量导航线可靠提取。对比其他导航路径提取方法,本文方法具有更高的准确性和稳定性。

关键词: 玉米作物行; 视觉导航; 导航线提取; 鸟瞰图; 深度学习; ResAC – UNet

中图分类号: TP391.41 文献标识码: A 文章编号: 1000-1298(2026)01-0348-10 OSID: 

Corn Crop Row Recognition and Navigation Line Extraction Algorithm Based on ResAC – UNet Network

CUI Yongzhi^{1,2} LIU Yangchun^{1,2} MAO Wenhua^{1,2} AN Qilin^{1,2} GUO Quanfeng^{1,2} LI Guangrui^{1,2}
ZHOU Da^{1,2} ZHOU Baixue^{1,2} WEI Ligu^{1,2}

(1. Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing 100083, China
2. State Key Laboratory of Agricultural Equipment Technology, Beijing 100083, China)

Abstract: In complex unstructured farmland environments, accurate extraction of navigation lines is crucial for agricultural machinery and agricultural robots to achieve autonomous operation. Challenging factors such as variable lighting, undulating terrain, and weed interference that are common in agricultural environments, making traditional image processing methods perform poorly in terms of adaptability, accuracy, and real-time performance, it's difficult to meet the visual navigation needs of smart agriculture. To address these issues, a ResAC – UNet deep learning network model was proposed based on an improved UNet. This model used the ResNet – 50 network to replace the encoder structure of the traditional UNet to enhance feature extraction capabilities. The segmentation speed and real-time response capabilities were improved through optimized jump connections. The ASPP module was introduced in the bottleneck part of the network to achieve multi-scale receptive field modeling, while maintaining high-resolution features and capturing richer contextual information. In addition, the model

收稿日期: 2025 – 05 – 16 修回日期: 2025 – 07 – 13

基金项目: 中国机械工业集团有限公司青年科技基金项目(QNJJ – PY – 2022 – 20)和国家重点研发计划项目(2024YFD2301100)

作者简介: 崔永志(1995—),男,博士生,主要从事农业装备智能测控技术研究,E-mail: 18787171287@163.com

通信作者: 伟利国(1978—),男,研究员,博士生导师,主要从事农业智能装备与信息化技术研究,E-mail: weilg78@126.com

integrate CBAM to enhance the accurate perception of crop row boundaries, effectively prevented the loss of key feature information, and further improved the segmentation quality. Based on the segmentation results, the row anchor method and RANSAC algorithm were used to achieve high-precision navigation line extraction and smoothing. The acquired front view image was converted into a bird's-eye view to eliminate the perspective effect, and a top view of the crop row with ROI was generated and retained. The experimental results showed that the ResAC - UNet model achieved 99.23%, 95.44%, 85.23% and 94.71% in precision, MPA, MIoU and recall, respectively, which was better than the current mainstream segmentation networks such as Segformer, DDRNet, HRNet and DeepLabV3+. The average inference time of ResAC - UNet was 15.26 ms, which met the real-time recognition requirements of intelligent agricultural machinery visual navigation. Three navigation lines can be extracted in the ROI area of the camera. The maximum angle error of the middle navigation line was only 0.96°, and the maximum pixel deviation was 4.3, which realized the reliable extraction of high-quality navigation lines. Compared with other navigation path extraction methods, the proposed method had higher accuracy and stability. The research result can provide an efficient and robust visual perception solution for the autonomous navigation of intelligent agricultural machinery in the field, which had certain practical value.

Key words: corn crop row; visual navigation; navigation line extraction; bird's-eye view; deep learning; ResAC - UNet

0 引言

玉米是我国主要粮食作物之一,中耕是其早中期关键田间管理环节,依赖农机导航系统实现作物行的精准识别与稳定导航。视觉导航系统利用相机获取环境图像,结合图像处理与计算机视觉算法生成导航路径,提供精确导航参考。当前农机自主导航依赖 GNSS、IMU、激光雷达与机器视觉等多源传感技术^[1]。GNSS 技术已广泛应用于农业自动驾驶作业,但其定位精度与卫星信号强弱和质量密切相关^[2]。在复杂环境中,容易产生多路径效应,导致导航和定位精度下降且缺乏对周围环境的感知能力,难以满足精细化农业作业需求^[3-4]。惯性导航系统在长时间作业中易累积误差,通常与 GNSS 联合使用获取农机位置信息,实时调整速度和方向^[5]。激光雷达虽具备较高测距精度,但易受田间环境影响,限制了其在农业导航领域的广泛应用^[6]。相较于上述导航技术,视觉导航因其低成本、高分辨率和对环境结构的良好感知能力,能实时提取作物行、田垄、障碍物等结构信息^[7]。不仅为农业机械提供高精度导航路径,还能适应多变光照、杂草干扰等复杂田间环境,特别适合应用于需高度环境感知的智慧农业场景^[8]。因此,研究具备鲁棒性与实时性的视觉导航技术,已成为推动农机装备智能化发展的关键技术路径之一。

图像分割是视觉导航的基础,旨在区分前景植被特征与背景区域,为作物行提取和导航路径规划奠定基础。传统方法主要包括基于颜色指数和阈值分割^[10-11],随着人工智能技术的迅猛发展,深度学

习在图像处理领域展现出显著优势^[9]。近年来,深度学习技术已广泛应用于农业研究的多个方向,作物行特征的识别与检测也日益受到关注。当前主要方法包括:基于 YOLO 系列模型的目标检测方法,通过检测框提取作物位置^[12-13];基于语义分割的区域识别方法,通过像素级分类获取作物冠层的完整区域信息,为高精度导航与行间作业提供有力数据支撑^[14-16]。

针对传统方法易受到光照和杂草等环境的干扰,YOLO 系列检测模型虽然推理速度较快但在作物行密集的场景下精度不足、语义分割算法在作物行弯曲和断行的场景下分割精度存在波动等问题,本文采用基于改进 UNet 骨干网络,融合 ASPP 和 CBAM 模块的方法,开展无人农场玉米作物行分割与导航线提取研究。

1 材料与方法

1.1 数据集构建

本研究采用 CRBD 数据集,如图 1 所示,该数据集包含多种作物类型的图像,涵盖玉米、芹菜、马铃薯、洋葱、向日葵和大豆等多种作物^[17]。原始图像分辨率为 2 560 像素 × 1 920 像素,调整至 512 像素 × 512 像素以降低计算复杂度,同时保证不会造成关键信息丢失。完整的评估图像集共包含 281 幅多样化样本。

1.2 数据增强

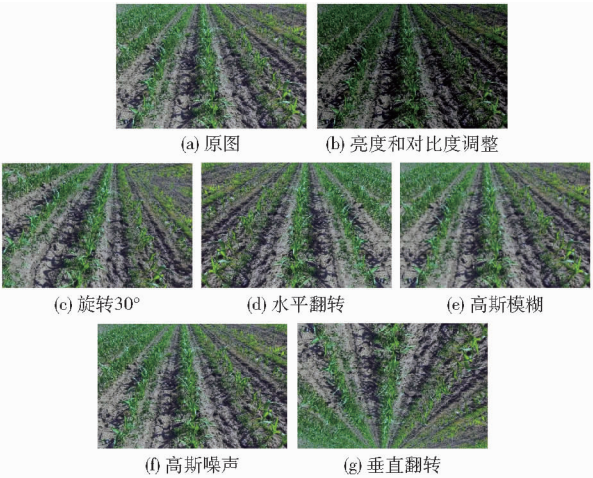
为提高模型的泛化能力和识别鲁棒性,对原始数据集实施了系统化的数据增强策略,如图 2 所示。具体增强方法包括:亮度和对比度调整、随机角度旋



(a) 带有杂草和阴影 (b) 不同作物类型 (c) 不同生长阶段 (d) 作物行弯曲

图 1 CRBD 数据集

Fig. 1 CRBD dataset



(a) 原图 (b) 亮度和对比度调整

(c) 旋转30° (d) 水平翻转 (e) 高斯模糊

(f) 高斯噪声 (g) 垂直翻转

图 2 数据增强方式

Fig. 2 Data enhancement methods

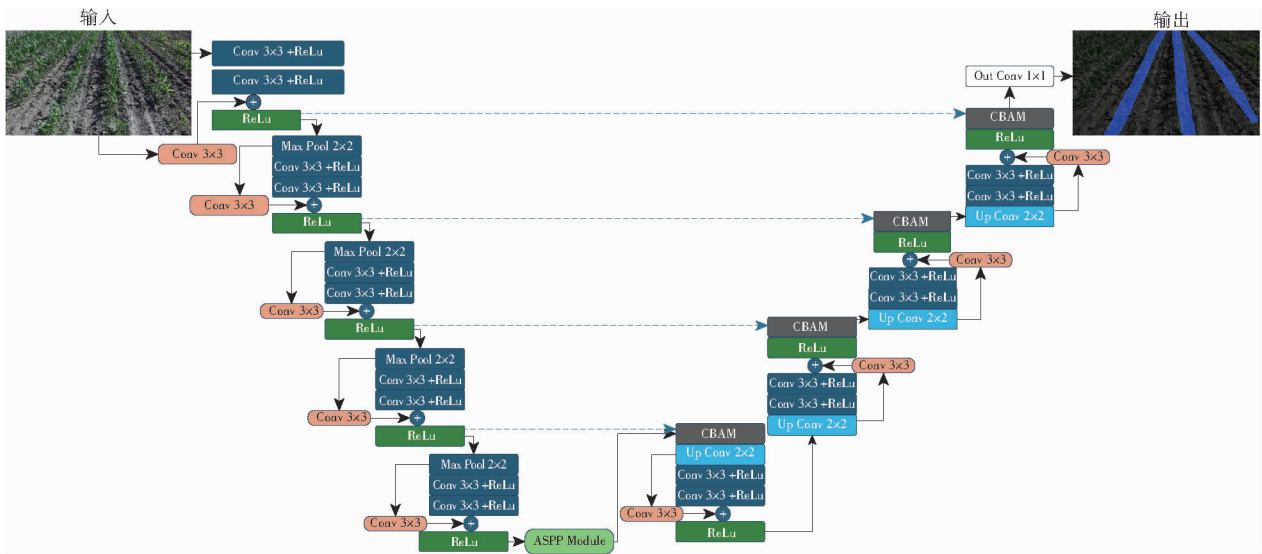


图 3 ResAC-UNet 网络结构图

Fig. 3 ResAC-UNet network structure diagram

解码器部分由转置卷积与上采样操作组成,逐步恢复图像的空间分辨率,实现精细化分割。在解码器结构中,本研究引入了 CBAM,通过精确建模通道和空间注意力,引导网络聚焦于作物行的关键区域和

转(−30°~30°)、水平翻转、高斯模糊处理、添加高斯噪声以及垂直翻转等图像变换技术。通过这些增强技术的综合应用,数据集规模从初始的 800 幅图像显著扩展至 4 000 幅,有效增强了模型应对复杂农田环境的能力。增强后的数据集按照比例 8:1:1 划分为训练集、验证集和测试集,确保模型训练与评估的科学性。

2 基于语义分割的作物行识别模型

2.1 改进的 UNet 网络

UNet 网络采用典型的编码器-解码器结构,在编码器阶段通过逐层下采样提取高层次特征,在解码器阶段通过上采样逐步恢复图像的空间分辨率。其独特之处在于引入跳跃连接机制,有效融合编码器提取的低级细节信息与解码器生成的高级语义特征,从而显著提高模型对边缘和空间结构的感知能力^[18]。如图 3 所示,编码器结构采用重复的卷积层、批量归一化和最大池化操作提取图像特征,逐层压缩信息以增强表达能力。与传统 UNet 的简单卷积结构不同,本研究将改进的 ResNet-50 主干网络引入编码器,显著增强了网络的特征提取深度和环境适应性。输入图像经过最大池化层进行下采样,通过多个残差块提取深层次的作物行结构特征。每级特征通过跳跃连接传递至解码器,有效保留空间细节并提升特征融合连续性。相加后的特征通过 ReLU 激活函数处理,进一步传递至上采样模块。

边界,从而显著提高分割结果的准确性与边缘定位能力。

2.2 ASPP 模块

在网络的颈部网络,引入了空洞空间金字塔池

化 (Atrouspatial pyramid pooling, ASPP) 模块,以增强对多尺度上下文信息的建模能力。该位置处于整个网络特征维度最深、语义信息最丰富的层级,经过多次下采样与 ResNet 主干特征提取后,得到的高层语义特征更适合进行多尺度感受野的融合^[19]。如图 4 所示,ASPP 结构包含 4 个并行的 atrous 卷积分支,其膨胀率分别为 1、6、12、18,用于捕捉从局部到全局的多尺度特征。此外,为增强全局语义感知能力,ASPP 还集成了 1 个全局平均池化分支,有效补充整幅图像的上下文信息。每个 atrous 卷积分支均由 1 个标准卷积层、批量归一化层和 ReLU 激活函数构成。所有分支的输出在通道维度上进行拼接,然后通过 1 个 1×1 卷积层,融合多尺度信息并压缩特征维度。

尽管 ASPP 能够提升特征的全局感知能力,但在处理农田图像中的作物行细节时,仍存在局部信息丢失的问题,特别是在复杂背景环境下易导致边

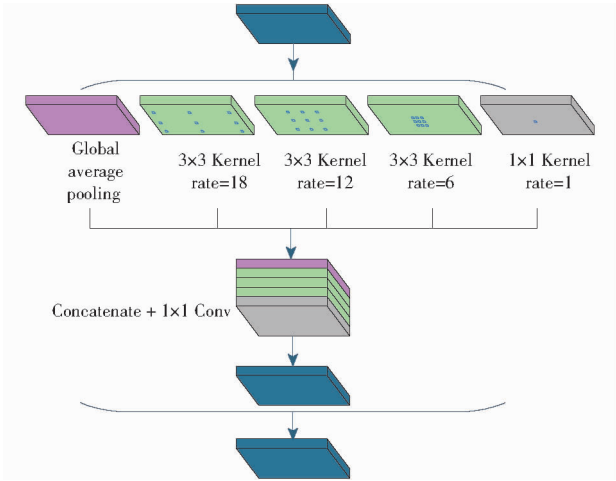


图 4 ASPP 模块

Fig. 4 ASPP module

界模糊、细节识别能力下降。为弥补这一不足,在 ASPP 模块之后引入了 CBAM 注意力机制,以增强特征表达能力。CBAM 由通道注意力模块和空间注意力模块两部分组成,如图 5 所示。

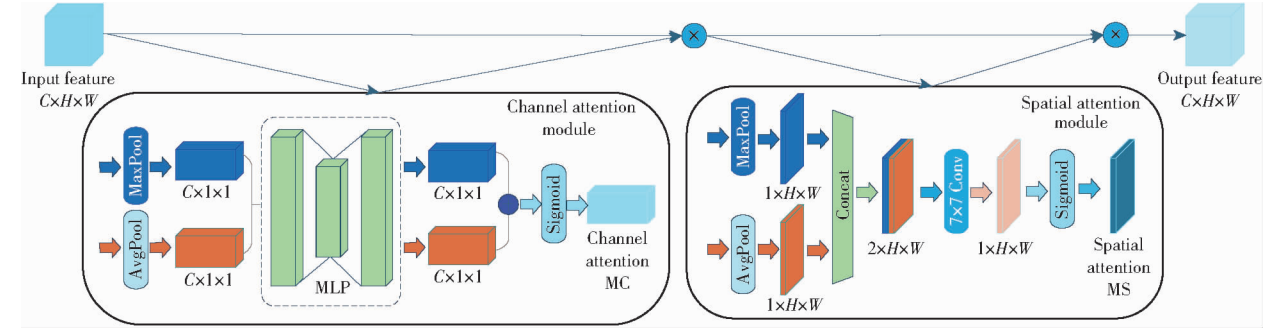


图 5 CBAM 模块

Fig. 5 CBAM module

2.3 CBAM 注意力机制

融合后的多尺度特征随后输入到 CBAM 注意力机制模块中。CBAM 依次通过通道注意力和空间注意力对特征图进行精细加权优化,显著提升网络对关键目标区域的关注能力^[20]。优化后的特征图被送入上采样模块,进入解码路径,逐步恢复空间分辨率并实现精确的图像分割。

通道注意力通过全局平均与最大池化生成描述向量,输入共享至 MLP 后得出权重,与原特征图逐通道相乘,突出关键通道信息。随后,空间注意力通过通道维度池化与拼接,提取空间权重图,并与特征图逐元素相乘,从而精确引导模型聚焦于目标区域的关键空间位置^[21]。CBAM 在计算开销小的前提下增强特征表达能力,强化对作物行边界与纹理的感知。模型在面对田间复杂环境时,能够更有效地区分目标与背景,显著提升整体分割精度与环境适应性。此外,注意力机制有助于模型在面对含有未标注物体或背景干扰区域的农业图像时,有选择地

忽略无关信息,进一步提高作物行识别的准确性与稳定性。

3 模型训练

3.1 试验平台及参数设置

在模型训练过程中,采用 Adam 优化器最小化损失函数,设置动量因子为 0.9,初始学习率为 1×10^{-3} ,并根据批量大小自适应调整学习率,通过余弦退火 (Cosine annealing) 策略实现动态学习率调整。为抑制模型过拟合,引入权重衰减机制,权重衰减系数设定为 5×10^{-4} 。模型训练总计进行 50 个周期的端到端训练,训练批量大小为 8。为提高小样本条件下模型的学习能力,采用迁移学习策略初始化网络权重,加速模型收敛并提升分割性能。本试验的全部训练和验证均在配置为 NVIDIA GeForce RTX 4070 Laptop GPU、32 GB 主存储器、Windows 11 (64 位) 操作系统平台上完成。使用 Python 3.9 与 PyTorch 1.7.0 作为主要开发环境,并基于 CUDA

11.3 与 CUDNN 8.2.1 实现 GPU 加速计算试验。

3.2 损失函数

为实现像素级精确分类,将作物行分割任务建模为典型的二分类问题,综合采用焦点损失函数(Focal loss)与 Dice 损失函数进行训练,以兼顾模型对难分类样本的识别能力与整体分割质量的准确性。Focal loss 旨在解决正负样本分布极度不均衡的场景,通过引入调制因子,自适应地降低易分类样本对总损失的贡献,强化对难分类像素的关注,从而显著提升模型在目标区域上的检测能力与召回率。然而,Focal loss 未能充分反映预测区域与真实区域在图像层面上的重叠程度,可能导致边界定位精度不足。为增强模型对目标结构的整体感知能力,本研究引入 Dice 损失函数作为补充。它直接度量预测区域与真实标签区域之间的重叠程度,能够有效引导模型在训练过程中优化分割边界,特别适用于对目标轮廓和形状结构要求较高的图像分割任务。Focal loss 与 Dice loss 的联合使用,使得模型在面对前景稀疏、类别不均衡的复杂农业场景中,既具备强大的像素级分类能力,又能保障分割结果的整体准确性和边界清晰度,计算公式为

$$Focal\ loss(p_i) = -\alpha(1 - p_i)^\gamma \log(p_i) \tag{1}$$

$$Dice\ loss = \frac{2|A \cap B|}{|A| + |B|} \tag{2}$$

- 式中 γ ——调节难易样本的聚焦因子
 α ——类别不平衡时的类别权重系数
 A ——预测区域
 B ——真实区域
 $|A \cap B|$ ——作物行的像素数量
 p_i ——模型对真实类别的预测概率

ResAC - UNet 与 UNet 模型在训练过程中损失函数的变化趋势如图 6 所示,ResAC - UNet 在训练初期的收敛速度明显快于 UNet,在前 20 轮训练与验证损失迅速下降,并在后期趋于平稳,最终收敛至较低水平,表现出良好的拟合能力和泛化能力。相较而言,UNet 模型训练损失下降较缓,并在中后期出现明显的振荡,训练损失和验证存在一定程度的过拟合。结果表明,ResAC - UNet 结构有效提升了模型训练稳定性与精度收敛能力。

3.3 模型评估指标

本研究聚焦二分类问题,采用二分类混淆矩阵评估模型性能。混淆矩阵包含 4 个关键指标:①真正例(TP):正确地识别作物行。②假正例(FP):将非作物行误识别为作物行。③假负例(FN):将作物行误识别为非作物行。④真负例(TN):正确地识别非作物行。

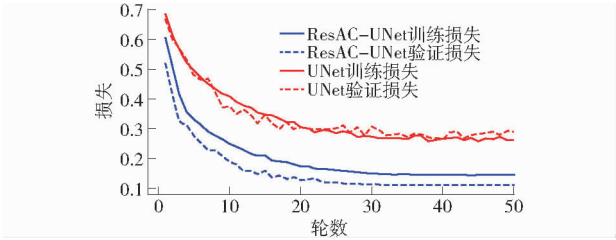


图 6 ResAC - UNet 与 UNet 的训练与验证损失对比曲线

Fig. 6 Comparison curves of training and validation loss between ResAC - UNet and UNet

为了全面评估语义分割模型在农机视觉导航任务中的性能,本研究采用了多种评估指标,包括平均交并比(Mean Intersection over Union, MIoU)、精确率(Precision, P)、召回率(Recall, R)、平均像素精度(Mean pixel accuracy, MPA),这些指标均建立在混淆矩阵的基础上^[22]。在实际应用场景中,模型单次推理的平均耗时通常采用平均推理时间进行衡量。由于模型初始化、GPU 内存分配等原因,首次推理过程往往耗时较长,该时间无法真实反映模型在常规状态下的推理性能,也会对实时处理需求的判断产生显著影响,因此需要将首次推理耗时排除在外。具体而言,模型的平均推理时间是在剔除首次推理耗时后,对剩余 $N - 1$ 次推理耗时取平均值,这一指标在深度学习模型评估中具有重要意义。

4 导航线提取

4.1 基于透视投影的作物图像 ROI 提取

智能农机搭载前向相机在农田中行进时,由于相机的透视效应,现实中平行的作物行在图像中呈现出“近宽远窄,逐渐汇聚”的视觉现象,如图 7 所示,最终汇聚到远处的消失点^[23]。

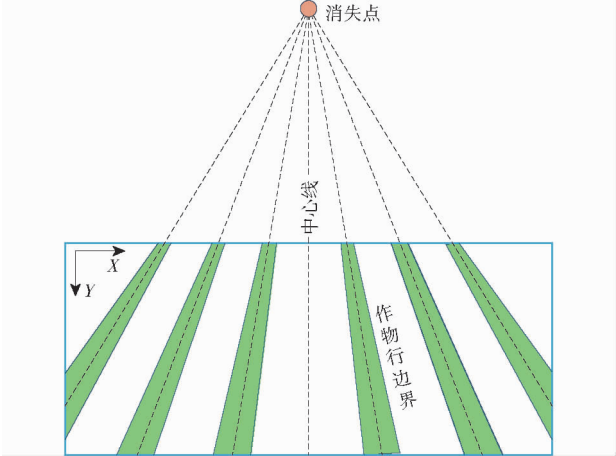


图 7 作物行透视效应

Fig. 7 Crop row perspective effect

基于智能农机在田间作业的行駛速度相对缓慢,系统更应关注前方近距离作物区域,而非远处模

糊信息。如图 8 所示,前向摄像头安装于农机中心位置,以一定俯角采集地面图像。为提高系统的稳定性和实时性能,优先提取并处理靠近农机的作物区域。这些近距离作物具有更清晰的边缘特征和更准确的几何比例,能更有效地支持导航线提取与路径规划。此外,图中展示的区域投影机制进一步说明了如何将三维田间场景映射至二维图像平面,便于对图像进行有效裁剪,仅保留感兴趣区域,排除冗余信息,提高处理效率和精度。

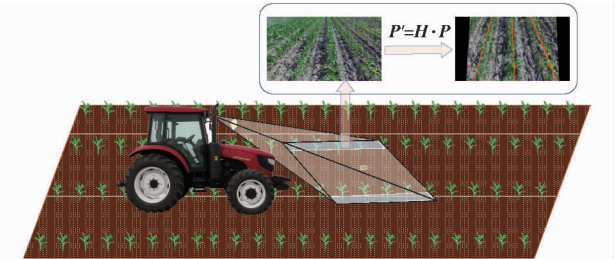


图 8 提取 ROI 区域
Fig. 8 Extracting ROI area

4.2 鸟瞰图转换

本研究通过图像鸟瞰变换,将前视图像转换为能消除透视效应且保留 ROI 的俯视图。将 $P = (x, y, 1)$ 作为输入点,通过 1 个 3×3 的变换矩阵 H ,得到目标输出点 $P' = (x', y', 1)$ 。

$$P' = HP = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3}$$

则有

$$\begin{cases} x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}} \\ y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \end{cases} \tag{4}$$

其中 $\omega = h_{31}x + h_{32}y + h_{33}$ 为缩放因子,需要进行归一化处理,使 $\omega = 1$,得

$$\begin{cases} x' = \frac{x'}{\omega} \\ y' = \frac{y'}{\omega} \end{cases} \tag{5}$$

通过原始 4 个点 (x_i, y_i) 与目标 4 个点 (x'_i, y'_i) , 构建线性方程组求解变换矩阵,即

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x'_1 & -y_1x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y'_1 & -y_1y'_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2x'_2 & -y_2x'_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -x_2y'_2 & -y_2y'_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x_3x'_3 & -y_3x'_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -x_3y'_3 & -y_3y'_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x_4x'_4 & -y_4x'_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -x_4y'_4 & -y_4y'_4 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} = 0 \tag{6}$$

则有

$$\begin{cases} x'_i = \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \\ y'_i = \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \end{cases} \tag{7}$$

最终坐标变换关系如式 (7), 可实现将采集图像转换为鸟瞰图视角, 便于后续作物行检测。

4.3 模型性能验证

4.3.1 主干网络评估

本研究进行了系统化的消融试验, 分别对比采用 VGG - 16 和一系列 ResNet 网络作为主干网络时 UNet 的性能。试验结果如表 1、2 所示。

表 1 主干网络试验对比结果

Tab.1 Comparison of backbone network test results

主干网络	MIoU/%	MPA/%	F1 值/%	Dice/%	参数量	平均推理时间/ms
UNet	83.5	89.7	86.3	87.3	1.28×10^7	21.8
VGG - 16	84.6	90.3	87.2	88.5	2.45×10^7	28.6
ResNet - 18	86.2	91.5	89.1	90.2	1.53×10^7	21.4
ResNet - 34	87.8	92.7	90.3	91.5	2.48×10^7	23.7
ResNet - 50	89.3	93.8	91.6	92.8	3.26×10^7	27.2
ResNet - 101	90.5	94.6	92.4	93.5	4.45×10^7	32.3

表 2 消融试验结果

Tab.2 Ablation experiments results

序号	CA	SA	CBAM	ASPP	MIoU/%	MPA/%	F1 值/%	Dice/%	参数量	平均推理时间/ms
1					89.3	93.8	91.6	92.8	3.26×10^7	27.2
2	√				90.1	94.3	92.2	93.4	3.31×10^7	27.8
3		√			89.9	94.1	92.0	93.2	3.28×10^7	27.6
4			√		90.7	94.9	92.8	94.0	3.34×10^7	28.3
5				√	90.5	94.7	92.5	93.8	3.42×10^7	29.1
6			√	√	91.8	95.6	93.5	94.8	3.51×10^7	30.5

注:“√”表示应用此模块。

ResNet-101 虽然性能最优,但提升幅度相对较小,模型参数量和平均推理时间显著增加,属于边际收益递减阶段。从 MIoU 的数值来分析各网络层级的性能提升与成本增加比,每单位参数量 ResNet-50 比 ResNet-34 提升 0.19%,高于 ResNet-101 的 0.10%,每增加单位参数量带来的性能增益几乎是 ResNet-101 的两倍。ResNet-50 主干网络实现了 89.3% 的 MIoU 和 93.8% 的 MPA,同时 F1 值和 Dice 系数分别达到 91.6% 和 92.8%,在性能和效率之间取得了良好的平衡。

以 ResNet-50 主干网络作为基础模型,完整的 CBAM 结构比单独使用通道注意力或空间注意力效果更好,MIoU 提升至 90.7%;ASPP 模块的引入为模型带来了多尺度特征感知能力,结合 CBAM 后,MIoU 进一步提升至 91.8%,相比仅使用 CBAM、ASPP 分别提高了 1.1、1.3 个百分点。CBAM 仅增加了 8×10^5 参数量和 1.1 ms 平均推理时间,虽然加入 ASPP 模块使参数量增加至 3.51×10^7 ,平均推理时间增加至 30.5 ms,但在作物行多尺度特征的提取上表现出显著优势,综合消融试验结果表明,所提出的 ResAC-UNet 在各项性能指标上均达到最优,相比基础 UNet 的 MIoU 提高了 8.3 个百分点,在农业应用中具备较高的实时性和识别精度。消融试验证明了各模块对性能提升的有效性,特别是 CBAM 和 ASPP 模块之间存在协同效应,共同贡献的性能提升超过了各自独立贡献的简单相加,展示了所提出架构的先进性和合理性。

4.3.2 模型对比试验

分别选用 DeepLabV3+、HRNet、DDRNet、Segformer、原始的 UNet 网络与本研究中的模型进行对比,不同模型的 MIoU、MPA、R、P、平均推理时间如表 3 所示。

表 3 不同语义分割模型对比
Tab.3 Comparison of different semantic segmentation models

模型	P/%	MIoU/%	MPA/%	R/%	平均推理 时间/ms
DeepLabV3+	98.70	83.93	94.54	93.80	71.60
HRNet	98.59	84.68	95.17	92.74	33.32
DDRNet	97.64	81.64	93.32	92.38	15.85
UNet	96.29	77.64	92.36	91.18	12.41
Segformer	96.23	77.09	91.72	90.17	20.95
ResAC-UNet	99.23	85.23	95.44	94.71	15.26

对比其他网络模型,本文模型的 ResAC-UNet 在各项评价指标上均表现出色,如表 3 所示,平均像素精度、平均交并比以及精确率和召回率均为最高。

平均像素精度较次优的 HRNet 提高了 0.27 个百分点;平均交并比提高了 0.55 个百分点。ResAC-UNet 的平均推理时间为 15.26 ms,仅次于 UNet 模型,且满足 60 f/s 以内的大多数作业场景。在精确率方面,相比于其他模型中精确率最高的 DeepLabV3+ (98.70%) 提高了 0.53 个百分点。证明了模型在减少误检和漏检方面的双重优势,为精确的作物行导航提供了可靠保障。

6 种模型的预测结果如图 9 所示,主要包括作物稀疏、作物行间断、作物行弯曲、作物行中包含杂草等场景。在作物稀疏和作物行存在间断的情况下 ResAC-UNet 和 DeepLabV3+ 的作物分割效果相近,但是在作物分割的细节方面 ResAC-UNet 的精度更高;在作物行包含杂草的情况下,ResAC-UNet 的作物分割效果要明显优于其他模型;在作物行弯曲的情况下,虽然 ResAC-UNet 分割作物过程中存在一定的误差,但是在 ROI 区域内仍保持较高的精度。

5 作物行导航线提取与评价

为了将分割完成的作物行实例转换为指导农机导航的有效信息,在进行作物行分割获得作物行的分割掩码后需要进行导航线拟合。本研究设计“行锚-RANSAC-连续化”的流程从分割的作物行掩码中得到可供智能农机视觉导航的连续导航线,如图 10 所示。

5.1 行锚法提取锚点与导航线拟合

通过横向分割算法识别出不同的作物行,采用连通区域分析与 DBSCAN 聚类,根据 x 坐标密度为每条作物掩码行分配唯一实例标识。随后对各实例沿垂直方向施加形态学闭运算,修补断裂区域,获得连贯完整的行掩码。通过垂直等距采样的锚点提取部分作物行中作物的片段信息,然后定位每个实例的横向中心位置,结合行锚点纵坐标形成了不同分割作物行的坐标点信息。针对断行等情况,作物行内某一高度缺少像素,采用上下搜索策略以最近有效中心替补,增强锚点集合完整性。确保每行至少有足够数量的锚点,以保证后续拟合的可靠性。

采用随机抽样一致性算法 (Random sample consensus, RANSAC) 对得到的离散锚点集合进行拟合,得到精确的导航参考线。RANSAC 算法能够有效区分内点与离群点,从而在复杂农业环境中提供可靠的导航参数。RANSAC 算法的核心优势在于能够排除异常锚点的干扰,在存在杂草、土壤不均匀、部分作物缺失等情况下仍能准确拟合导航线。相较

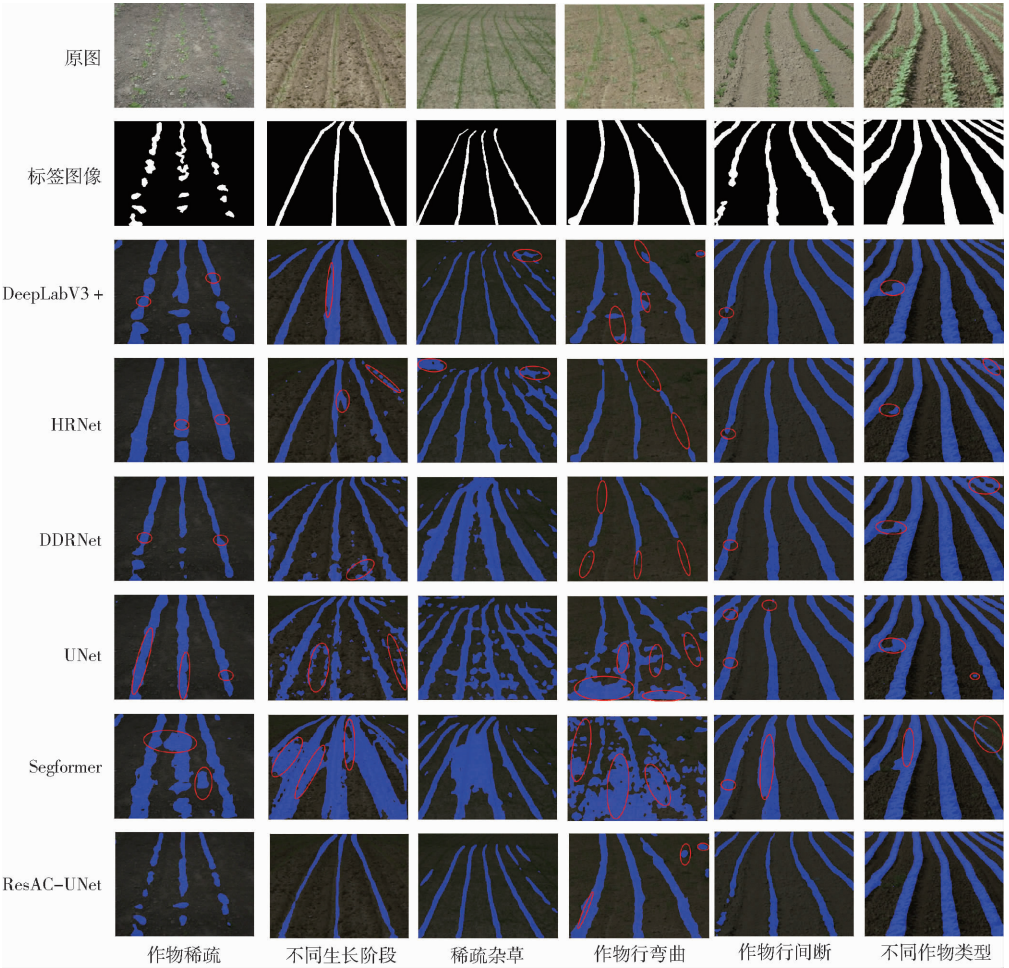


图 9 分割结果

Fig. 9 Segmentation results

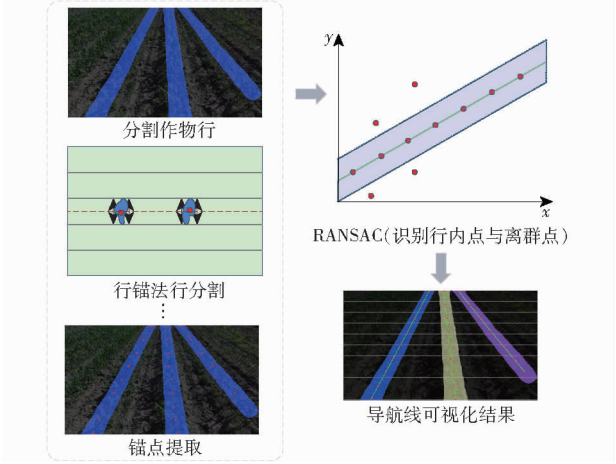


图 10 导航线提取

Fig. 10 Navigation line extraction

于霍夫变换和直接最小二乘法, RANSAC 算法具有较强的抗噪能力和鲁棒性, 能够适应噪声多和离群点多的复杂农业场景。

5.2 导航线提取精度评价

为了有效评估提取的作物行导航线的拟合精度, 对于直线的导航线采用平均横向距离 L 和拟合线上地面真实线之间的角 θ 进行分析。如图 11 所

示为评价方法的计算原理。

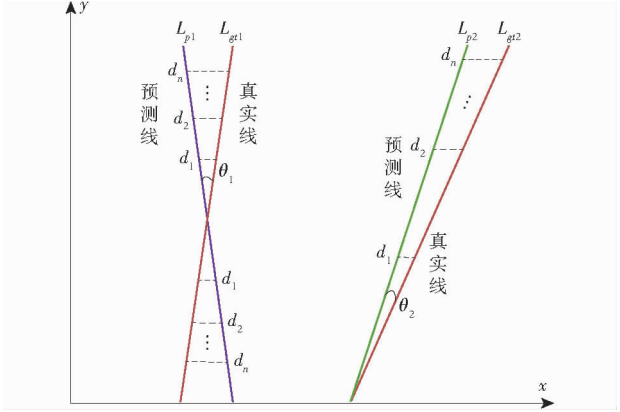


图 11 导航线提取精度评价示意图

Fig. 11 Schematic of navigation line extraction accuracy evaluation

将 CVAT 标注的作物行地面真实线作为真值, 与算法拟合的导航线叠加比较, 二者之间的角度差值为航偏角偏差, 二者之间的横向距离差作为像素偏差。如图 12 所示, 4 种场景下导航线与地面真实线之间的偏差比较小, 直线作物行精度优于弯曲作物行。

如表 4 所示, 平均角偏差在 1.6° 以内, 平均横

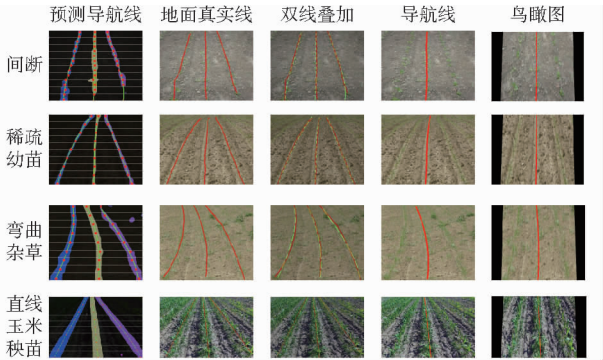


图 12 导航线拟合图

Fig. 12 Navigation line fitting diagram

向距离偏差在 7.4 像素以内。其中直线的玉米秧苗导航线平均偏差最小,3 条导航线的平均角偏差为 0.78°、平均横向距离偏差为 1.82 像素;弯曲杂草的导航线平均偏差最大,分别为 1.52°和7.35 像素。偏差分析表明,图像中心区域的导航线提取精度更高,在实际农机视觉导航过程中,将导航的核心引导区域定位在作物行中间区域。在间断的作物行情景中,导航线的提取算法能够有效预测提取断行之间的导航线,平均偏差为 1.00°和 2.21 像素;在作物行为直线的情境中,导航线的提取精度最高。针对不同的场景,本研究的方法获得导航线偏差均在 8 像素以内,证明了该方法的稳定性和适用性。

表 4 不同场景下导航线平均偏差

Tab. 4 Average error of navigation line in different scenarios

场景	角偏差/(°)				横向距离偏差/像素			
	左	中	右	平均值	左	中	右	平均值
间断	1.14	0.47	1.39	1.00	2.38	1.69	2.57	2.21
稀疏幼苗	1.44	0.45	1.40	1.10	2.07	1.41	3.19	2.22
弯曲杂草	2.09	0.96	1.51	1.52	10.16	4.30	7.58	7.35
直线玉米秧苗	1.03	0.34	0.98	0.78	2.50	1.01	1.94	1.82

将本研究得到的导航线平均误差与文献[24]的平均角偏差及文献[25]的横向距离偏差进行对比,由于本研究与文献[25]的图像分辨率不同,故统一转换成物理距离偏差 (cm) 进行比较。如表 5 所示,本文的作物行导航线提取方法的角偏差平均值和标准差相较于文献[24]降低了 56.8% 和 30.5%;横向距离偏差的平均值和标准差相较于文献[25]降低了 43.7% 和 31.9%,说明本文研究方

法的准确性和稳定性更高,能更好地适应复杂农田环境,满足农机高精度导航的应用需求。

表 5 不同导航路径提取方案对比

Tab. 5 Comparison of different navigation path extraction solutions

作物行检测	角偏差/(°)		横向距离偏差/cm	
	平均值	标准差	平均值	标准差
本文	0.555	0.276	0.410	0.290
文献[24]	1.285	0.397		
文献[25]			0.728	0.426

6 结论

(1)本文针对作物行导航线提取问题,提出了改进 UNet 网络模型,采用变换鸟瞰图解决相机透视效应,联合行锚法和 RANSAC 算法精确提取分割图像中的导航线。

(2)利用轻量化的 ResNet-50 代替 UNet 的编码器主干网络,并在模型的瓶颈网络融合 ASPP 模块,编码器结构中整合 CBAM 注意力机制,相比于原 UNet 的 MIoU 提高了 8.3 个百分点;CBAM 和 ASPP 联合使用比仅单独使用的 MIoU 分别提高了 1.1 个百分点和 1.3 个百分点。能够有效融合作物行的多尺度信息,并显著提升模型的整体分割精度与环境适应性。

(3)通过对相机的前视图像进行鸟瞰图转换,消除透视效应,且生成保留 ROI 的作物行俯视图;对比 DeepLabV3+、HRNet、DDRNet、Segformer 和原始的 UNet 网络,本文改进模型的 P、MIoU、MPA 和 R 分别为 99.23%、85.23%、95.44% 和 94.71%,均为最优,ResAC-UNet 的平均推理时间为 15.26 ms,满足 60 f/s 以内的大多数作业场景中视觉导航的识别实时性需求。

(4)通过行锚法获取作物行的片段信息,获得锚点位置;采用 RANSAC 算法对获取的锚点进行拟合,获得导航线。面对不同的情景,提取相机中间的 3 条导航线,外侧的导航线偏差大于内部,故保留图像中间的导航线作为视觉导航线。不同场景下提取中间的导航线的最大角偏差为 0.96°,最大像素偏差为 4.30 像素,满足视觉导航的精度需求。通过与其他导航路径提取方法对比可知,本文的研究方法准确性和稳定性更高。

参 考 文 献

[1] MA Z, YIN C, DU X, et al. Rice row tracking control of crawler tractor based on the satellite and visual integrated navigation [J]. Computers and Electronics in Agriculture, 2022, 197: 106935.

[2] LOHAN S K, NARANG M K, SINGH M, et al. Design and development of remote-control system for two-wheel paddy transplanter [J]. Journal of Field Robotics, 2022, 39(3): 177-187.

- [3] 袁洪良,杨浚宇,唐睿,等. 基于低成本惯导和运动约束的农机高精度定位系统研究[J]. 农业机械学报,2023,54(7): 17 - 25.
YUAN Hongliang, YANG Junyu, TANG Rui, et al. High-precision localization of autonomous agricultural machinery using low-cost IMU and motion constraints[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(7): 17 - 25. (in Chinese)
- [4] SAHA S, MORITA T, OSPINA R, et al. A vision-based navigation system for an agricultural autonomous tractor[J]. IFAC-PapersOnLine, 2022, 55(32): 48 - 53.
- [5] 钟银,薛梦琦,袁洪良. 智能农机 GNSS/INS 组合导航系统设计[J]. 农业工程学报,2021,37(9): 40 - 46.
ZHONG Yin, XUE Mengqi, YUAN Hongliang. Design of the GNSS/INS integrated navigation system for intelligent agricultural machinery[J]. Transactions of the CSAE, 2021, 37(9): 40 - 46. (in Chinese)
- [6] 邱权,胡青含,樊正强,等. 基于自适应系数卡尔曼滤波的农业移动机器人组合定位[J]. 农业机械学报,2022,53(增刊 1): 36 - 43.
QIU Quan, HU Qinghan, FAN Zhengqiang, et al. Adaptive-coefficient Kalman filter based combined positioning algorithm for agricultural mobile robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(Supp. 1): 36 - 43. (in Chinese)
- [7] 张伟荣,陈学庚,齐江涛,等. 基于深度学习和高斯过程回归的玉米冠下视觉导航路径提取方法[J]. 农业机械学报, 2024, 55(7): 15 - 26.
ZHANG Weirong, CHEN Xuegeng, QI Jiangtao, et al. Deep learning and Gaussian process regression based path extraction for visual navigation under canopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(7): 15 - 26. (in Chinese)
- [8] MOUSAZADEH H. A technical review on navigation systems of agricultural autonomous off-road vehicles [J]. Journal of Terramechanics, 2013, 50(3): 211 - 232.
- [9] 武锦龙,吴虹麒,李浩,等. 基于改进 DeepLabV3 + 的荞麦苗期无人机遥感图像分割识别方法研究[J]. 农业机械学报, 2024, 55(5): 186 - 195.
WU Jinlong, WU Hongqi, LI Hao, et al. Segmentation of buckwheat by UAV based on improved lightweight DeepLabV3 + at seedling stage[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(5): 186 - 195. (in Chinese)
- [10] ZHAO R, YUAN X, YANG Z, et al. Image-based crop row detection utilizing the Hough transform and DBSCAN clustering analysis[J]. IET Image Processing, 2024, 18(5): 1161 - 1177.
- [11] BAN C, WANG L, CHI R, et al. A Camera - LiDAR - IMU fusion method for real-time extraction of navigation line between maize field rows[J]. Computers and Electronics in Agriculture, 2024, 223: 109114.
- [12] RUAN Z, CHANG P, CUI S, et al. A precise crop row detection algorithm in complex farmland for unmanned agricultural machines[J]. Biosystems Engineering, 2023, 232: 1 - 12.
- [13] GONG H, ZHUANG W, WANG X. Improving the maize crop row navigation line recognition method of YOLOX[J]. Frontiers in Plant Science, 2024, 15: 1338228.
- [14] WEI J, ZHANG M, WU C, et al. Accurate crop row recognition of maize at the seedling stage using lightweight network[J]. International Journal of Agricultural and Biological Engineering, 2024, 17(1): 189 - 198.
- [15] QUAN L, GUO Z, HUANG L, et al. Efficient extraction of corn rows in diverse scenarios: a grid-based selection method for intelligent classification[J]. Computers and Electronics in Agriculture, 2024, 218: 108759.
- [16] DE SILVA R, CIELNIAK G, GAO J. Vision based crop row navigation under varying field conditions in arable fields[J]. Computers and Electronics in Agriculture, 2024, 217: 108581.
- [17] VIDOVIĆ I, CUPEC R, HOCENSKI Ž. Crop row detection by global energy minimization[J]. Pattern Recognition, 2016, 55: 68 - 86.
- [18] 韩振浩,李佳,苑严伟,等. 基于 U - Net 网络的果园视觉导航路径识别方法[J]. 农业机械学报,2021,52(1): 30 - 39.
HAN Zhenhao, LI Jia, YUAN Yanwei, et al. Path recognition of orchard visual navigation based on U - Net[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 30 - 39. (in Chinese)
- [19] 张继成,侯郁硕,郑萍,等. 低数据集下基于 ASPP - YOLO v5 的苋菜识别方法研究[J]. 农业机械学报,2023,54(增刊 2): 223 - 228.
ZHANG Jicheng, HOU Yushuo, ZHENG Ping, et al. Method for amaranth identification based on ASPP - YOLO v5 model in low data set[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp. 2): 223 - 228. (in Chinese)
- [20] 王泰华,郭亚州,张家乐,等. 基于改进 YOLO v5s 的水稻害虫识别研究[J]. 农业机械学报,2024,55(11): 39 - 48.
WANG Taihua, GUO Yazhou, ZHANG Jiale, et al. Rice pest identification based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 39 - 48. (in Chinese)
- [21] 袁培森,丁毅飞,徐焕良. 基于深度主动学习与 CBAM 的细粒度菊花表型识别[J]. 农业机械学报,2024,55(2): 258 - 267.
YUAN Peisen, DING Yifei, XU Huanliang. Fine-grained chrysanthemum phenotype recognition based on deep active learning and CBAM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 258 - 267. (in Chinese)

[23] 张娜, 曾琴, 杨小兰, 等. 振动磨机磨介球的离散元法仿真与比较分析[J]. 南京工程学院学报:自然科学版, 2017, 15(1): 23 – 28.
ZHANG Na, ZENG Qin, YANG Xiaolan, et al. Discrete element simulation and comparative analysis of grinding medium in a vibration mill[J]. Journal of Nanjing Institute of Technology: Natural Science Edition, 2017, 15(1): 23 – 28. (in Chinese)

[24] BOND F C. Confirmation of the third theory[M]. Sem. A., 1960: 13 – 153.

[25] 郭杰. 自循环超细粉碎机理论及试验研究[D]. 昆明:昆明理工大学, 2002.
GUO Jie. A study of theory and experiment on auto inner cycling fine crusher[D]. Kunming: Kunming University of Science and Technology, 2002. (in Chinese)

[26] 纪冠亚. 不同尺度机械粉碎对秸秆物性及酶解效率的影响和机理研究[D]. 北京: 中国农业大学, 2018.
JI Guanya. Multi-scales mechanical fragmentation of crop residues: impacts on properties and enzymatic hydrolysis and related mechanism[D]. Beijing: China Agricultural University, 2018. (in Chinese)

[27] KIM S M, DIEN B S, SINGH V. Promise of combined hydrothermal/chemical and mechanical refining for pretreatment of woody and herbaceous biomass[J]. Biotechnology for Biofuels and Bioproducts, 2016, 9: 97.

~~~~~

(上接第 357 页)

[22] CHAMARA N, BAI G, GE Y. AICropCAM: deploying classification, segmentation, detection, and counting deep-learning models for crop monitoring on the edge[J]. Computers and Electronics in Agriculture, 2023, 215:108420.

[23] LI D, DONG C, LI B, et al. Hybrid rice row detection at the pollination stage based on vanishing point and line-scanning method[J]. Precision Agriculture, 2023, 24(3): 921 – 947.

[24] 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)

[25] 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(Supp. 1): 41 – 50. (in Chinese)