

基于语义分割的中后期玉米行间路径导航线检测

苏童¹ 王琳² 班超¹ 迟瑞娟¹ 马悦琦¹

(1. 中国农业大学工学院, 北京 100083; 2. 智能农业动力装备全国重点实验室, 洛阳 471039)

摘要: 中后期玉米行间路径存在光照不足、遮挡等因素的干扰,不利于农业机器人自主作业时导航线的检测。针对此问题,本文提出一种基于改进 Fast-SCNN 语义分割模型的中后期玉米行间路径导航线检测算法。首先,针对目前路径语义分割模型在中后期玉米环境下边缘分割不够准确的问题,提出一种 Edge-FastSCNN 模型,在模型分支中引入本文提出的边缘提取模块(Edge extraction module, EEM)以获取准确的路径边界信息,并在模型中引入空间金字塔池化(Atrous spatial pyramid pooling, ASPP)模块以融合图像边界信息和深层特征。然后,基于模型预测的行间路径掩码,通过像素扫描法检测路径掩码左右边界点,通过加权平均法求得路径掩码中点。最终利用最小二乘法拟合导航线,实现中后期玉米行间路径导航线的检测。为验证所提出方法的性能,基于中后期玉米正常光照无遮挡、光照不足、阴影、杂草遮挡、叶片遮挡等5种环境,进行了模型性能对比实验和导航线检测实验。实验结果表明,模型平均交并比为97.90%,平均像素准确率为98.84%,准确率为99.39%,推理速度为63.0 f/s;模型在上述5种环境下的平均交并比为96.93%~98.01%,平均像素准确率为98.33%~99.03%,准确率为98.53%~99.12%;预测导航线与真实导航线在上述5种环境下的航向角偏差平均值为1.15°~3.16°,平均像素横向距离为1.89~3.41像素;导航线检测算法的单帧图像平均处理时间为90.04 ms。因此,本文提出的导航线检测算法满足中后期玉米行间路径导航任务的准确性和实时性要求。

关键词: 中后期玉米; 行间导航; 导航线检测; 语义分割; 图像处理

中图分类号: TP242.6

文献标识码: A

文章编号: 1000-1298(2024)10-0275-11

OSID:



Interrow Path Navigation Line Detection of Maize in Middle and Late Period Based on Semantic Segmentation

SU Tong¹ WANG Lin² BAN Chao¹ CHI Ruijuan¹ MA Yueqi¹

(1. College of Engineering, China Agricultural University, Beijing 100083, China

2. State Key Laboratory of Intelligent Agricultural Power Equipment, Luoyang 471039, China)

Abstract: The interrow path of maize in the middle and late stages is interfered by factors such as insufficient light and occlusion, which is not favorable to the detection of navigation lines during autonomous operation of agricultural robots. To address this problem, an algorithm based on the improved Fast-SCNN semantic segmentation model for detecting the navigation lines in the interrow path of maize in the mid-late stage was proposed. Firstly, to address the problem that the current path semantic segmentation model was not accurate enough for edge segmentation in the mid-late maize environment, an Edge-FastSCNN model was proposed, and the edge extraction module (EEM) proposed was introduced in the model branch to obtain accurate path boundary information, and spatial pyramid pooling was introduced into the model. Atrous spatial pyramid pooling (ASPP) module was introduced in the model to fuse the image boundary information and deep features. Then based on the interline path mask predicted by the model, the left and right boundary points of the path mask were detected by pixel scanning method, and the midpoint of the path mask was obtained by weighted average method. Finally, the least squares method was used to fit the navigation lines to achieve the detection of the mid- and late-stage maize interline path navigation lines. In order to verify the performance of the proposed method,

收稿日期: 2024-04-08 修回日期: 2024-06-19

基金项目: 国家自然科学基金项目(52172396)和智能农业动力装备全国重点实验室开放项目(SKLIAP2024009)

作者简介: 苏童(2000—),男,硕士生,主要从事智能农机导航技术研究,E-mail: 1321745031@qq.com

通信作者: 迟瑞娟(1969—),女,副教授,博士生导师,主要从事车辆机电控制及智能化研究,E-mail: chiruijuan@cau.edu.cn

model performance comparison experiments and navigation line detection experiments were conducted based on five environments such as normal light without shade, insufficient light, shadows, weeds shade, and leaf shade of maize in the middle and late stages. The experimental results showed that the average intersection and merger ratio of the model was 97.90%, the average pixel accuracy was 98.84%, the accuracy rate was 99.39%, and the inference speed was 63.0 f/s; the average intersection and merger ratio of the model in the five environments mentioned above was ranged from 96.93% to 98.01%, and the average pixel accuracy was ranged from 98.33% to 99.03%, and the accuracy rate was from 98.53% to 99.12%; the average value of heading angle deviation between the predicted navigation line and the real navigation line in the above five environments was $1.15^{\circ} \sim 3.16^{\circ}$, and the average pixel lateral distance was 1.89 ~ 3.41 pixels; the average processing time for a single-frame image of the navigation line detection algorithm was 90.04 ms. Therefore, the navigation line detection algorithm proposed met the mid- and late-stage maize interline path navigation task's accuracy and real-time requirements.

Key words: maize of middle and late period; interline navigation; navigation line detection; semantic segmentation; image processing

0 引言

玉米作为我国三大粮食作物之一,在粮食系统中发挥着关键和日益重要的作用^[1]。精准获取中后期玉米的生长状况并进行相应作业对增加玉米的产量至关重要^[2-4]。为降低劳动成本和提高作业效率,农业机器人可自主行驶入玉米田中获取精确的玉米生长状况并精准作业。目前,农业机器人导航方式包括全球卫星定位导航、激光雷达导航和机器视觉导航等^[5]。全球卫星定位导航方法在高遮挡环境下容易产生信号丢失的问题^[6],并不适用于中后期玉米行间高遮挡环境下的路径导航。3D 激光雷达价格昂贵,在恶劣的工况下不稳定,且获取的点云数据易受不平坦地面和杂草的影响^[7]。相较之下,机器视觉导航技术成本较低,性能稳定^[8-10],更适合较为恶劣的农业环境。而行驶路径导航线的检测是农业机器人基于机器视觉导航技术实现自主导航和作业的关键技术之一。

国内外学者针对农业机器人基于机器视觉的导航线检测算法进行了许多研究,算法思路大都分为作物行提取和导航线拟合 2 个步骤^[11-12]。作物行提取的准确性关系到导航线拟合的准确性。传统的图像处理算法在环境光照较弱、杂草较多时具有较差的作物行提取效果,鲁棒性较差^[12-14]。随着深度学习的不断发展,神经网络模型在作物行提取方面取得了良好的效果。基于深度学习的作物行提取算法包括作物行语义分割和作物目标识别。李秀智等^[15]选用 ESNet 语义分割模型实现了玉米幼苗作物行的感知,在自然农田场景取得了较好的分割效果。LIU 等^[16]在 ERFNet 模型的基础上提出一种 MS-ERFNet 模型,实现了玉米幼苗行语义分割,模型平均交并比 (Mean intersection over union, MIoU)

和像素准确率 (Pixel accuracy, PA) 分别为 93.40% 和 97.54%,分割速度为 15.26 f/s。然而,用于作物行提取的语义分割模型适用于作物植株高度较低且作物行距较宽的情况,当植株高度较高且植株叶片出现交叉的情况时语义分割效果较差。彭书博等^[17]提出一种改进的 YOLO v7 模型以实现果树树根的准确识别,检测精度为 95.21%,检测速度为 42.07 f/s。DIAO 等^[18]基于改进的 YOLO v8s 模型识别玉米生长前期植株冠层,改进后的模型平均像素准确率提升至 90.2%。作物目标识别算法虽然可以适用于植株高度较高的情况,但当作物呈密集种植时植株间距较小,容易导致目标识别错误进而影响最终导航线的检测。可见,基于作物行提取的导航线检测算法受作物类别和生长状况的影响,不适用于中后期玉米植株高度较高且密集种植的情况。

相较于作物行提取算法,基于语义分割模型实现农业机器人行驶路径分割的导航线检测算法不受作物类别和生长状况的影响,目前较多应用于果园等开阔地区和野外道路导航线检测。侯文慧等^[19]在 U-Net 模型的基础上优化,提高了模型对果园垄间路径的识别精度,在草莓、蓝莓和葡萄果园均取得了较好的分割效果。饶秀勤等^[20]在 U-Net 模型的基础上提出一种 Fast-Unet 模型,以提取棉花、玉米和甘蔗田间的导航路径。模型针对上述 3 种作物的 MIoU 分别为 79.1%、88.1% 和 94.0%。赵岩等^[21]提出一种融合 U-Net 和 ResNet 模型优势的 Unet-ResNet34 模型,实现了不同光照和杂草条件下导航路径分割,MIoU 为 91.89%。ZHANG 等^[22]提出一种改进的 U-Net 模型,实现了桃、橙和猕猴桃树行间路径的识别提取,提取精度分别为 97.7%、98.7% 和 95.6%,平均推理速度为 48.8 f/s。然而,目前的路径语义分割模型大多针对路径边缘没有较

多遮挡且光线较为充足的环境,如果园和开放式的农田,在中后期玉米行间路径边缘存在较多遮挡且光照不足的环境下并不适用。

本文针对现有路径语义分割模型因边缘分割不够准确,无法适用于中后期玉米行间环境的问题,采集中后期玉米行间路径图像信息并制作数据集,将 Fast-SCNN 模型改进为 Edge-FastSCNN 模型,使其在行间路径光照不足且存在遮挡的情况下依旧能够准确分割出行间路径边缘,同时保证模型的实时性。在此基础上通过像素扫描法检测预测的路径掩码左右边界点,通过加权平均法计算得出路径掩码中点并利用最小二乘法拟合出行间路径导航线。最后进行模型性能实验和导航线检测实验以验证算法的准确性和实时性。

1 图像采集与数据集制作

1.1 图像采集

本研究所用图像信息是由实验室智能采样车上搭载的 Basler 单目相机采集于中国农业大学上庄实验站,采集时间为 2023 年 5—9 月。为便于采样车进行行间信息采集,玉米采用宽窄行种植模式,其中宽行行距为 80 cm,适于采样车在行间行驶。采集对象为玉米抽雄期之后玉米行间路径信息,采集的路径环境包括正常光照无遮挡、光照不足、阴影、玉米叶遮挡、杂草遮挡等。智能采样车如图 1 所示。

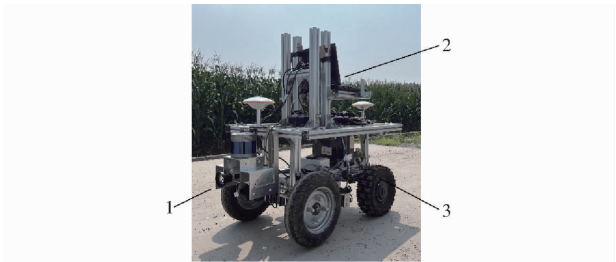


图 1 智能采样车

Fig.1 Intelligent sampling vehicle

1. 单目相机 2. 便携式计算机 3. 采样车

为了采集中后期玉米高遮挡环境下玉米行间路径信息,将相机通过连接件安装在采样车前端高于地面约 20 cm 处。通过手动遥控智能采样车在玉米行间行驶,同时采集图像信息,行驶速度为 5 km/h。采集图像分辨率为 1 920 像素 × 1 200 像素,采集帧率为 10 f/s,视频格式为 AVI。为减少后续模型训练时的计算参数,将图像尺寸缩放至 640 像素 × 480 像素,不同环境下路径图像如图 2 所示。

1.2 数据集制作

基于采集的图像信息选取 600 幅制作数据集,使用 Labelme 对图像中玉米行间路径进行标注。模

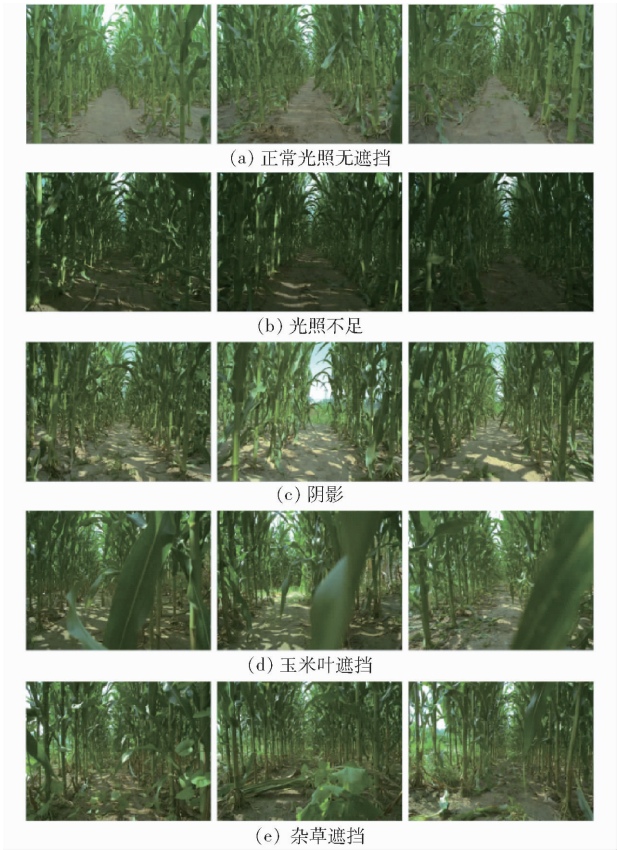


图 2 不同环境下路径图像

Fig.2 Path images in different environments

型训练过程中由于样本数量不足会出现过拟合现象^[23],因此对所选图像进行数据增强,通过改变亮度、水平镜像、上下翻转、旋转随机角度、高斯模糊和添加噪声等方式将图像扩充至 1 800 幅。标注路径掩码和数据增强后图像如图 3 所示。按照比例 4:1:1 将数据集划分为训练数据集、验证数据集和测试数据集,分别用于模型训练、验证和测试。

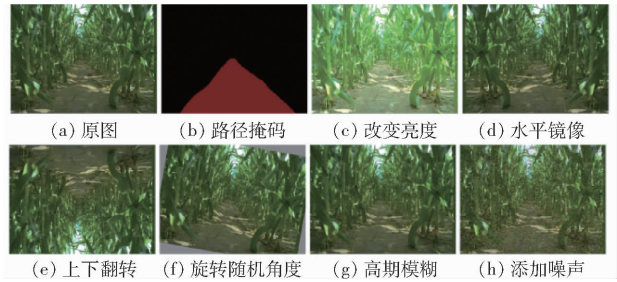


图 3 数据增强样本及图像标签

Fig.3 Data enhancement samples and image labels

2 基于 Edge-FastSCNN 的路径语义分割方法

Fast-SCNN 模型是一种实时语义分割模型,模型结构简单高效,保证其在低存储量的嵌入式设备上具有良好的预测结果^[24]。如图 4 所示, Fast-SCNN 的模型结构由 4 个模块组成:学习下采样模

块、全局特征提取模块、特征融合模块和分类器,其中全局特征提取模块采用双分支结构,提取低分辨率深度特征的同时保留高分辨率信息。

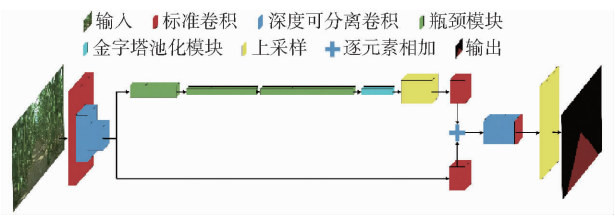


图 4 Fast-SCNN 模型结构

Fig. 4 Fast-SCNN model structure

随着农业深度学习模型不断发展, Fast-SCNN 模型逐渐应用在农业中^[25-27]。Fast-SCNN 模型可以满足一些农业场景语义分割的需求并具有较快的分割速度,但应用场景多数为光照充足且没有较多遮挡的环境,对边界识别的准确性也没有较高要求。而中后期玉米植株较高,导致行间路径光

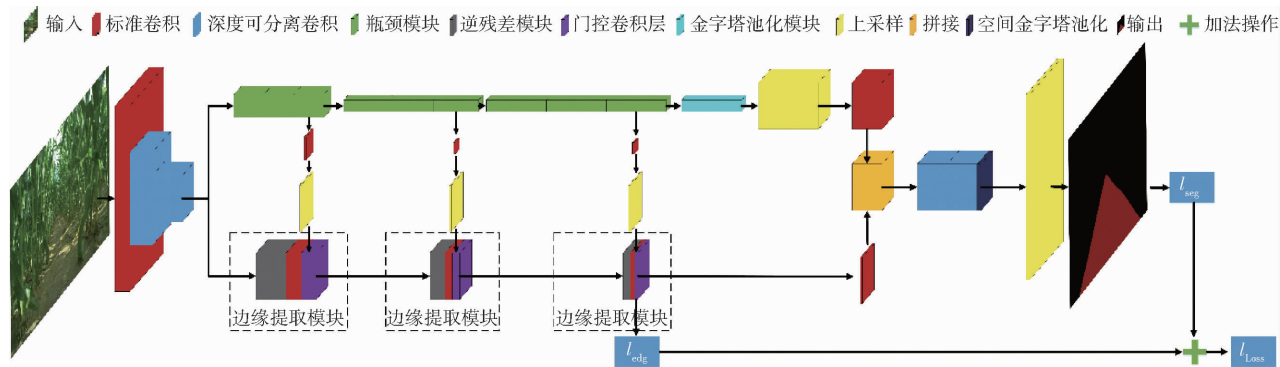


图 5 Edge-FastSCNN 模型结构

Fig. 5 Edge-FastSCNN model structure

全局特征提取模块包括全局特征提取和边缘特征提取两分支,二者共享学习下采样模块输出的图像低级特征。全局特征提取分支旨在捕捉用于图像分割的全局上下文信息,边缘特征提取分支旨在提取图像的边缘信息。全局特征提取分支引入 Mobilenet-V2^[28]中的高效瓶颈模块实现图像的深度特征提取,3 个高效瓶颈模块为一组。前 3 组高效瓶颈模块在降低特征图尺寸的同时改变特征图通道数,第 3 组瓶颈模块仅改变特征图通道数。每组瓶颈模块输出的特征图一方面向后继续传递,经过金字塔池化模块后输出,特征图尺寸为 40 像素 × 30 像素,特征通道数为 128;另一方面经过卷积层和上采样后作为边缘提取模块中 GCL 的输入之一。边缘特征提取分支共享学习下采样模块输出的特征图,集成边缘提取模块用以提取特征图中的边缘信息,经过 3 个边缘提取模块后输出特征图尺寸为 160 像素 × 120 像素,特征通道数为 8。

全局特征提取模块两分支的输出经过双线性插值上采样和标准卷积层后再经拼接操作输入至分类

器模块中,特征图尺寸为 160 像素 × 120 像素,特征图通道数为 65。相较于 Fast-SCNN 模型特征融合模块采用逐元素相加的方式合并 2 个分支的输出,Edge-FastSCNN 对两分支的输出特征进行拼接操作,同时保留深度特征信息和边缘信息,扩展其特征的代表能力。分类器模块由 2 个深度可分离卷积和 ASPP 模块构成,输出特征图尺寸为 160 像素 × 120 像素,特征通道数为 2。最终,通过双线性插值上采样将特征图尺寸恢复至 640 像素 × 480 像素输出,特征通道数不变。

Edge-FastSCNN 模型结构如图 5 所示。学习下采样模块初步学习图像特征的同时,对图像进行下采样以降低图像的分辨率、提高模型的处理速度并扩大接受域。经过学习下采样模块后,图像特征图尺寸降低到原来的 1/4,为 160 像素 × 120 像素;特征图通道数增加为 64。

器模块中,特征图尺寸为 160 像素 × 120 像素,特征图通道数为 65。相较于 Fast-SCNN 模型特征融合模块采用逐元素相加的方式合并 2 个分支的输出,Edge-FastSCNN 对两分支的输出特征进行拼接操作,同时保留深度特征信息和边缘信息,扩展其特征的代表能力。分类器模块由 2 个深度可分离卷积和 ASPP 模块构成,输出特征图尺寸为 160 像素 × 120 像素,特征通道数为 2。最终,通过双线性插值上采样将特征图尺寸恢复至 640 像素 × 480 像素输出,特征通道数不变。

2.1 边缘提取模块 (EEM)

基于路径语义分割的导航线检测算法需要保证路径边缘提取的准确性,但仅保留一定程度高分辨率细节信息的 Fast-SCNN 模型在中后期玉米环境下缺乏对边缘信息的提取能力。因此本文提出一种边缘提取模块并将其集成在 Fast-SCNN 模型中,增强模型提取图像边界信息的能力以解决路径边界分割不准确的问题。如图 5 所示,边缘提取模块由 1 个逆残差模块、1 个标准卷积和 1 个门控卷积层

(GCL)组成。逆残差模块接收边缘特征提取分支的浅层低级特征作为输入,在避免网络退化的同时保留更多的图像特征信息。逆残差模块的输出经过标准卷积层后输入门控卷积层以降低运算量。门控卷积层是边缘提取模块的核心,以标准卷积层的浅层特征和全局特征提取分支的深层特征作为输入,负责提取图像中与边缘相关的信息。

2.1.1 逆残差模块

Edge - FastSCNN 模型的边缘提取分支需要准确提取行驶路径的边缘信息,然而图像的细节特征信息会随着传统卷积层的叠加而丢失,导致无法准确分割出行驶路径。残差结构的引入可以很大程度上缓解这种信息丢失的问题^[29]。传统的残差结构为减少模型参数一般选择将特征图的通道数降低后再通过卷积进行特征提取,但通道数降低后的特征图包含的信息较少,不利于特征信息的获取。因此,为保证充足的特征信息,逆残差模块将特征图的通道数升高后进行特征提取,同时采用深度可分离卷积替代传统卷积来降低模型参数,其结构如图 6 所示^[30]。

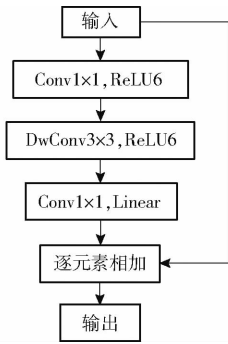


图 6 逆残差模块
Fig. 6 Inverse residual module

因此在边缘提取模块中引入逆残差模块以保留更多的特征信息,保证行驶路径边界提取的准确性。

2.1.2 门控卷积层

GCL 模块可以帮助边界提取模块只处理行间路径相关的信息,过滤掉其他信息,可被视作为一种注意力机制^[31]。GCL 结构如图 7 所示,首先将标准卷积层输出的浅层特征 E_i 和全局特征提取分支的深层特征 G_i 拼接后进行 1×1 卷积,之后利用 Sigmoid 函数形成注意力图 α_i ,计算公式为

$$\alpha_i = \sigma(C_{1 \times 1}(E_i \parallel G_i)) \quad (1)$$

式中 $\parallel$ ——拼接操作

$C_{1 \times 1}()$ —— 1×1 卷积

σ ——Sigmoid 函数

将注意力图 α_i 与 E_i 逐元素点乘,其结果与 E_i 逐元素相加后与权重矩阵相乘得到最终输出,即

$$O_i = ((E_i \alpha_i) + E_i)^T \omega_i \quad (2)$$

式中 O_i ——GCL 模块输出

ω_i ——权重矩阵

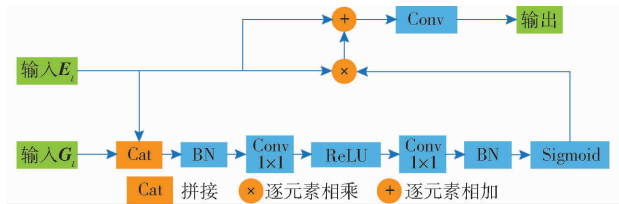


图 7 门控卷积层模块
Fig. 7 Gated convolution layer module

2.2 空间金字塔池化 (ASPP)

为了将图像的深度特征信息和边界信息更好地融合起来,防止因卷积操作导致的某些特征失真或消失,引入了 ASPP 模块来保留更多上下文信息。ASPP 的结构如图 8 所示,由 5 个分支构成,其中第 2、3、4 分支结合了空洞卷积和深度可分离卷积提取不同尺度下的图像特征信息^[32]。空洞卷积通过在卷积核中引入膨胀率来扩大感受野,在不增加卷积核大小的情况下捕获更广阔的上下文信息。深度可分离卷积将传统的卷积操作分解为逐通道卷积和逐点卷积,保证模型特征提取准确性的同时显著减少模型的参数量和计算量。5 个分支的输出结果拼接后经 1×1 卷积层调整通道数后输出。因此,引入 ASPP 模块对两分支的融合特征进行多个尺度的特征提取,保留更多上下文信息的同时增强模型对边缘信息的提取能力,提高了中后期玉米行间路径分割效果。

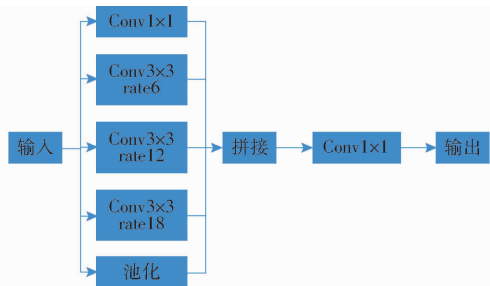


图 8 空间金字塔池化模块
Fig. 8 Atrous spatial pyramid pooling module

3 导航线检测方法

中后期玉米行间路径导航线的检测方法为:基于上文提出的 Edge - FastSCNN 模型预测出的路径掩码,通过像素扫描法检测路径掩码左右边界点,通过加权平均法求得路径掩码中点,最终利用最小二乘法拟合导航线,其流程图如图 9 所示。具体操作如下:

(1)对通过 Edge - FastSCNN 模型预测出的路径掩码进行二值化处理以便于后续左右边界点的检测,其结果如图 9c 所示,图中白色像素的像素值为

255,黑色像素的像素值为 0。

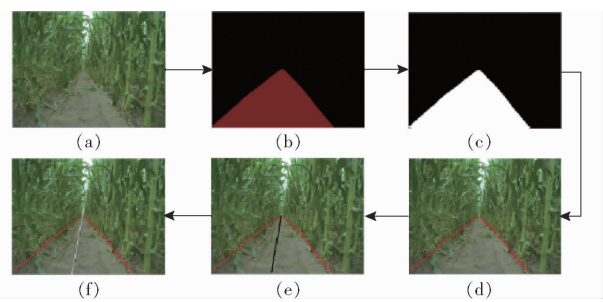


图 9 导航线检测流程图

Fig. 9 Flowchart of navigation line detection

(2)在图 9c 中,以图像左上角为坐标原点,水平向右为 X 轴正方向,垂直向下为 Y 轴正方向建立图像坐标系,设置滑动窗口沿 Y 轴正方向滑动。为减少运算量,根据输出图像尺寸将滑动窗口的高度设为 1 像素,宽度设为 640 像素,滑动间隔设为 5 像素。对滑动窗口内的像素点按照从左到右的顺序扫描,当某个像素点前一个位置的像素值为 0,后一个位置的像素值为 255,则认为该像素点为左边界点。同样,当某个像素点前一个位置像素值为 255,后一个位置像素值为 0,则认为该像素点为右边界点,通过该方法确定滑动窗口内路径掩码左、右两侧的边界点,并记录其位置坐标。之后滑动窗口继续沿 Y 轴方向滑动,直至到达图像像素 Y 轴坐标的最大值,提取的所有左、右边界点如图 9d 所示。

(3)将相同像素行的左右边界点进行加权平均即可得到该 Y 轴坐标值下的路径掩码中点,如图 9e 所示,计算公式为

$$\begin{cases} x_{i,m} = m_1 x_{i,l} + m_2 x_{i,r} \\ y_{i,m} = m_1 y_{i,l} + m_2 y_{i,r} \end{cases} \quad (3)$$

式中 m_1 、 m_2 ——加权平均系数,取 0.5
 $x_{i,m}$ 、 $x_{i,l}$ 、 $x_{i,r}$ ——路径掩码中点、左边界点和右边界点横坐标
 $y_{i,m}$ 、 $y_{i,l}$ 、 $y_{i,r}$ ——路径掩码中点、左边界点和右边界点纵坐标

(4)最后,将所有路径掩码中点通过最小二乘法拟合出导航线,最终生成导航线如图 9f 所示。结果表明,该方法可以基于路径掩码生成准确的导航线。

4 实验

为验证本文提出的 Edge - FastSCNN 模型和导航线检测算法的性能,进行了模型性能对比实验和中后期玉米行间路径导航线检测实验。

4.1 算法实验平台

算法编写和测试所用计算机处理器配置为 Intel i5 - 11600KF,显卡配置为 NVIDIA RTX 3060 -

12G,机带 RAM 为 16.0 GB。操作系统为 Windows 10,通过 Visual Studio Code 和 Anaconda 3 进行算法设计;基于 PyTorch 深度学习框架进行深度学习模型的设计和训练,相应算法环境配置为:PyTorch 1. 12. 0、Python 3. 7、Cuda 10. 0、Cuda DNN。

4.2 Edge - FastSCNN 模型超参数设置及训练

通过 Adam 梯度下降法训练模型,不使用预训练权重。初始学习率设置为 5×10^{-4} ,通过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。为提高训练速度,将每次训练的批次大小设置为 4,模型训练迭代次数为 300。Edge - FastSCNN 模型的训练损失曲线如图 10 所示,从图中可以看出,模型训练损失和验证损失均随着训练次数的增加快速下降,最终达到较低的损失值,说明 Edge - FastSCNN 具有较快的收敛速度且不存在过拟合和欠拟合的问题。为缩短整体训练时间,每经过 5 次迭代评估 1 次 MIoU,共 60 次。基于验证集评估的 MIoU 随着训练迭代次数的增加快速上升,最终达到较高的数值并保持稳定,如图 11 所示。

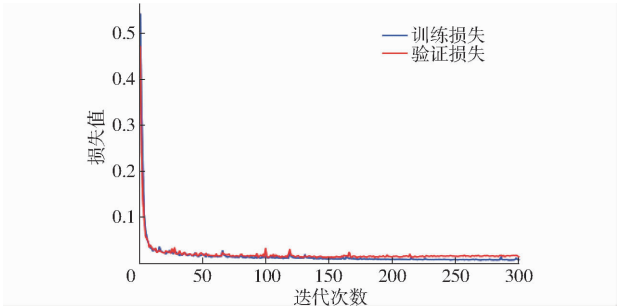


图 10 模型训练损失曲线

Fig. 10 Model training loss

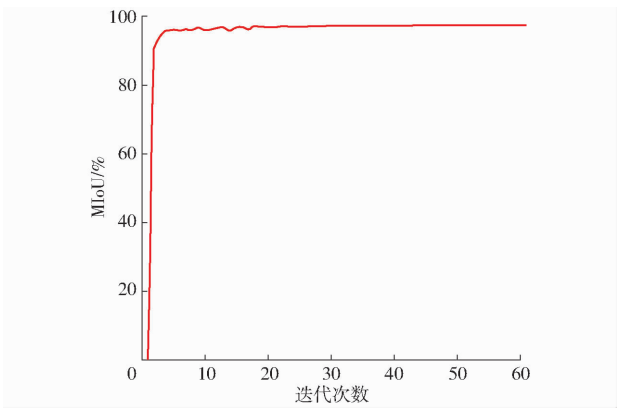


图 11 模型训练精度曲线

Fig. 11 Model training accuracy

4.3 模型性能对比实验

为测试模型的性能,进行了不同模型性能对比实验、模型消融实验和模型泛化性实验。

4.3.1 不同模型性能对比实验

为体现 Edge - FastSCNN 模型的优越性,将

Edge - FastSCNN 模型分割结果与 SegNet、U - Net、Deeplabv3 + 、Fast - SCNN 模型分割效果进行比较。各个模型参数量和浮点数如表 1 所示,其中 SegNet 模型具有最多的参数量和浮点数,为 2.95×10^7 和 1.32×10^{12} ;Fast - SCNN 具有最少的参数量和浮点数,为 1.14×10^6 和 2.07×10^9 。

表 1 模型参数量与浮点数

Tab.1 Model parameters and floating-point numbers					
模型	SegNet	U – Net	Deeplabv3 +	Fast – SCNN	Edge – FastSCNN
参数量	2.95×10^7	1.73×10^7	5.81×10^6	1.14×10^6	1.24×10^6
浮点数	1.32×10^{12}	3.77×10^{11}	6.20×10^{10}	2.07×10^9	1.25×10^{10}

4.3.1.1 模型分割效果对比实验

图 12 为部分中后期玉米行间路径图像信息及路径掩码。SegNet 的分割结果最差,图中几乎看不到被预测为行间路径的像素。U - Net 模型和 Fast - SCNN 模型的分割效果能够大致分割出行间路径,但当图像中存在玉米叶或杂草遮挡时,部分像素会被错误预测导致预测路径区域不完整。Deeplabv3 + 模型可以较好地分割出行间路径,但存在路径边界分割不整齐的问题,影响后续导航线检测的准确性。Edge - FastSCNN 模型预测效果优于其他 4 种模型,实现了准确分割路径且分割边缘整齐,预测路径区域完整。

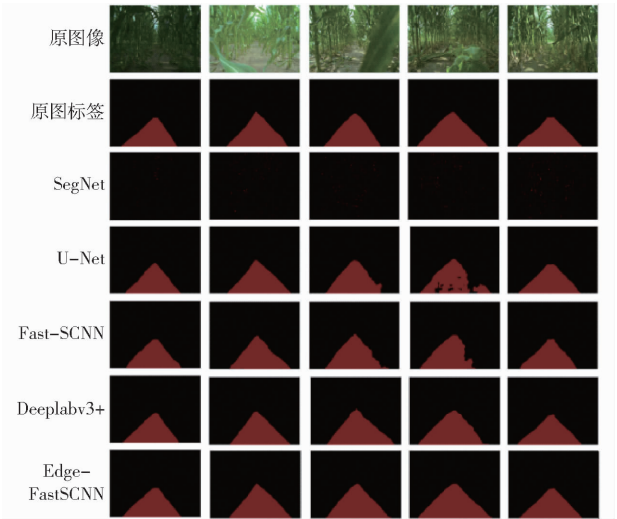


图 12 中后期玉米路径图像信息及路径掩码
Fig.12 Path image information and path mask in middle and late maize

4.3.1.2 模型量化指标对比

为评估 Edge - FastSCNN 模型的性能,使用平均交并比 (MIoU)、平均像素准确率 (mPA)、准确率 (Accuracy, Acc) 和预测速率 (Predict rate, PR) 作为模型评价的量化指标。MIoU 指模型对所有类别的预测掩码与真实掩码交并比的均值。mPA 指所有

类别被正确分类的像素数的平均值。准确率指预测结果中正确分类的像素数与总像素数比值。

为使模型在分割出行驶路径的同时获取更为准确的路径边界,Edge - FastSCNN 模型在训练过程中通过多重损失训练实现模型分割预测和边界图预测的共同监督 (图 5)。由于模型在像素分类过程中采用 Sigmoid 激活函数,因此模型分割预测损失函数 l_{seg} 和边界图损失函数 l_{edg} 均采用二分类交叉熵损失函数 (Binary cross entropy, BCE) l_{bce} ,计算公式为

$$l_{bce} = - [p_{i,j} \ln (y'_{i,j}) + (1 - p_{i,j}) \ln (1 - y'_{i,j})] \quad (4)$$

式中 $p_{i,j}$ ——真实标签图 (i,j) 处像素值
 $y'_{i,j}$ ——模型预测图 (i,j) 处像素值

将 2 个损失函数进行加权求和得到最终的联合损失函数,计算公式为

$$l_{Loss} = \lambda_1 l_{seg} + \lambda_2 l_{edg} \quad (5)$$

式中 λ_1, λ_2 ——损失函数加权系数
 l_{Loss} ——联合损失函数的损失值

表 2 为 5 种模型基于测试集测试后的结果。由表中数据可以看出,改进的 Edge - FastSCNN 模型的 MIoU、mPA 和 Acc 分别为 97.90%、98.84% 和 99.39%,均优于对照组的 4 种模型。SegNet 模型 MIoU、mPA 和 Acc 分别为 42.48%、50.01% 和 82.26%,明显低于其他 4 种模型,不能用于导航线的检测。U - Net 模型相较于 Edge - FastSCNN 模型 MIoU、mPA 和 Acc 仅低 0.97、0.71、0.28 个百分点,但其 PR 为 12.8 f/s,只有后者的 1/5,难以满足导航的实行性要求。Edge - FastSCNN 模型和 Deeplabv3 + 模型的 PR 相近,分别为 63.0 f/s 和 60.0 f/s,但前者的 MIoU、mPA 和 Acc 比后者高 1.39、0.74、0.42 个百分点,相比之下,Edge - FastSCNN 可以更加准确地预测出行间路径。Fast - SCNN 模型因其较小的参数量具有最高的 PR,为 82.2 f/s,但其 MIoU、mPA、Acc 与 Edge - FastSCNN 相比分别低 2.1、1.28、1.39 个百分点。由于农业机器人一般为低速行驶,并不需求过高的预测速率。相较之下,语义分割模型预测掩码的准确性对农业机器人导航更为重要。改进后的 Edge - FastSCNN 模型保证较高预测速率的同时显著改进了模型性能,因此更能

表 2 模型量化指标对比

Tab.2 Model quantitative index comparison				
模型	MIoU/%	mPA/%	Acc/%	PR/(f·s ⁻¹)
SegNet	42.48	50.01	82.25	14.9
U - Net	96.93	98.13	99.11	12.8
Deeplabv3 +	96.51	98.10	98.97	60.0
Fast - SCNN	95.80	97.56	98.00	82.2
Edge - FastSCNN	97.90	98.84	99.39	63.0

满足中后期玉米行间路径导航的需求。综上,改进的 Edge - FastSCNN 模型具有较快推理速度的同时实现了较为准确的路径分割效果。

4.3.2 模型消融实验

为了验证 Edge - FastSCNN 模型中 EEM、ASPP 等模块的有效性,基于中后期玉米测试集对 Fast - SCNN 模型、Fast - SCNN 模型 + EEM、Fast - SCNN 模型 + ASPP 及 Fast - SCNN 模型 + EEM + ASPP 等 3 种结构进行性能测试,实验结果如表 3 所示。由表 3 可知, Fast - SCNN + EEM 和 Fast - SCNN + ASPP 结构的 MIoU、mPA 及 Acc 均高于 Fast - SCNN,说明引入的 EEM 和 ASPP 对 Fast - SCNN 的性能均有促进作用。而主干模型 + EEM + ASPP 结构的 MIoU、mPA 及 Acc 优于其他 3 种结构,说明将 EEM 与 ASPP 共同引入主干模型中能够更好地改善语义分割的结果。

表 3 模型消融实验结果

Tab.3 Model ablation experiment				%
模型	MIoU	mPA	Acc	
Fast - SCNN	95.80	97.56	98.00	
Fast - SCNN + EEM	96.91	98.26	98.79	
Fast - SCNN + ASPP	96.69	98.17	98.53	
Fast - SCNN + EEM + ASPP	97.90	98.84	99.39	

4.3.3 模型泛化性实验

为评估 Edge - FastSCNN 模型的泛化性,选取来自正常光照无遮挡、光照不足、阴影、玉米叶遮挡、杂草遮挡等 5 种环境,共 200 幅中后期玉米行间路径图像对 Edge - FastSCNN 模型进行预测实验。将这 200 幅图像按照不同的环境分为 5 组,每组 40 幅图像。表 4 中数据为 Edge - FastSCNN 模型在上述 5 种不同情况的量化指标结果。图 13 为 Edge - FastSCNN 模型在上述 5 种环境下预测的路径掩码。在光照不足和行间路径存在阴影的环境下,Edge - FastSCNN 的 MIoU、mPA、Acc 相对于正常环境仅分别降低了 0.50、0.16、0.59 个百分点和 0.62、0.36、0.23 个百分点,说明该模型预测结果受光照影响较小。杂草遮挡环境下 Edge - FastSCNN 模型的 MIoU、mPA 和 Acc 与正常环境相比仅分别降低 0.26、0.28、0.11 个百分点。因部分玉米叶遮挡镜头导致路径像素大片丢失的情况下,Edge - FastSCNN 模型的 MIoU、mPA 和 Acc 与正常环境相比仅分别降低 1.08、0.70、0.38 个百分点,同时 Edge - FastSCNN 模型在路径信息丢失的情况下依旧能预测出完整的路径掩码。综上,Edge - FastSCNN 模型对不同的行间路径环境具有良好的适应性,鲁棒性较强。

表 4 不同环境量化指标对比

Tab.4 Comparison of quantitative indicators of different environments				%
环境	MIoU	mPA	Acc	
正常光照无遮挡	98.01	99.03	99.12	
光照不足	97.51	98.87	98.53	
阴影	97.39	98.67	98.89	
玉米叶遮挡	96.93	98.33	98.74	
杂草遮挡	97.75	98.75	99.01	

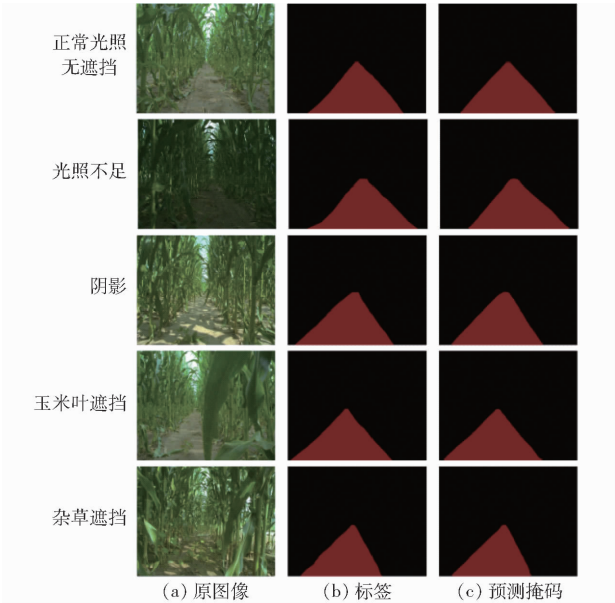


图 13 不同环境下路径掩码

Fig.13 Path mask in different environments

4.4 中后期玉米行间路径导航线检测实验

为验证本文提出的中后期玉米导航线检测算法的准确性和实时性,进行了导航线检测准确性实验和实时性实验。

4.4.1 导航线检测准确性实验

在通过 Edge - FastSCNN 模型预测出中后期玉米正常光照无遮挡、光照不足、阴影、玉米叶遮挡和杂草遮挡等 5 种环境下的行间路径掩码后,通过算法检测出导航线。为验证算法的准确性,选用航向角偏差均值和平均像素横向距离作为评价指标,计算方法如图 14 所示。

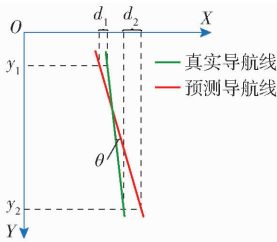


图 14 导航线计算方法

Fig.14 Navigation line calculation method

航向角由导航线斜率的反正切值计算得出,真

实导航线与预测导航线航向角差值即为航向角偏差,如图中 θ 所示,计算式为

$$\theta = \arctan \frac{1}{k_1} - \arctan \frac{1}{k_2} \tag{6}$$

式中 k_1, k_2 ——真实导航线和预测导航线斜率
 θ ——航向角偏差

为防止航向角偏差正负值抵消,将航向角偏差求绝对值后计算航向角偏差平均值,计算公式为

$$\bar{\theta} = \frac{1}{n} \sum_i |\theta_i| \tag{7}$$

式中 $\bar{\theta}$ ——航向角偏差均值
 n ——图像数量 i ——图像编号

像素横坐标偏差为相同纵坐标值下预测导航线与真实导航线横坐标差值的绝对值。考虑真实导航线与预测导航线存在交叉的情况,沿图像 Y 轴方向取 2 个相距较远的点,分别计算其对应在真实导航线与预测导航线上的横坐标,将横坐标差值的绝对值作为像素横坐标距离,如图 14 中 d_1 和 d_2 所示,其具体计算公式为

$$d_i = \frac{1}{n} \sum |x_1 - x_2| \tag{8}$$

式中 x_1, x_2 ——真实导航线和预测导航线对应 y 坐标的像素横坐标值

d_i ——像素横向距离

将 d_1 和 d_2 加权平均后得出像素横向距离平均值 $\bar{d}$,计算公式为

$$\bar{d} = \frac{d_1 + d_2}{2} \tag{9}$$

预测导航线与真实导航线在不同环境下的航向角偏差和平均像素横向距离如表 5 所示。由表 5 可以看出,正常光照无遮挡环境下的航向角偏差和平均像素横向距离最小,分别为 1.15°和 1.89 像素。光照不足和阴影的环境下的航向角偏差分别为 1.40°和 1.55°,相较于正常环境下仅分别增加 0.25°和 0.40°,平均像素横向距离分别为 3.19 像素和 3.14 像素,相较于正常光照下仅分别增加 1.30 像素和 1.25 像素,由此说明光照的变化对导航线检测算法并无明显影响。而在路径被遮挡时,杂草影响下的航向角偏差为 1.60°,平均像素横向距离为 2.97 像素,相较于正常环境下分别增加 0.45°和 1.08 像素;玉米叶遮挡的情况下航向角偏差和平均像素横向距离分别为 3.16°和 3.41 像素,相较于正常环境下分别增加 2.01°和 1.52 像素,说明路径被遮挡时,算法依旧能够较为准确地检测出导航线。实验结果表明,基于 Edge - FastSCNN 模型的中后期玉米导航线检测方法可以有效检测出行间

路径导航线,同时在不同的光照环境和遮挡情况下具有良好的准确性和鲁棒性,满足中后期玉米行间导航任务的需求。

表 5 导航线航向角偏差均值和平均像素横向距离

Tab. 5 Navigation line course angle deviation mean and mean pixel lateral distance

环境	航向角偏差 均值/(°)	平均像素横向距离/像素		
		d_1	d_2	$\bar{d}$
正常光照无遮挡	1.15	1.79	1.99	1.89
光照不足	1.40	2.20	4.17	3.19
阴影	1.55	3.27	3.01	3.14
玉米叶遮挡	3.16	3.95	2.86	3.41
杂草遮挡	1.60	2.88	3.05	2.97

4.4.2 导航线检测实时性实验

为验证基于 Edge - FastSCNN 语义分割模型的导航线检测算法的实时性,拍摄 3 条不同行间路径的视频,将单帧图像的平均处理耗时作为算法实时性的评价标准。通过在算法中设置计时器的方式记录算法开始和检测出导航线的时间,二者相减即为算法处理每帧图像所需要的时间,3 条视频的单帧图像平均处理时间为 90.04 ms,满足中后期玉米行间路径导航的实时性需求。

5 结论

(1)提出了一种可准确分割路径边界的轻量化语义分割模型 Edge - FastSCNN。该模型 MIoU 为 97.90%, mPA 为 98.84%, Acc 为 99.39%,模型参数量为 1.24×10^6 ,预测速率为 63.0 f/s。模型在正常光照无遮挡、光照不足、阴影、杂草、叶片遮挡等 5 种环境下的平均交并比为 96.93% ~ 98.01%,平均像素准确率为 98.33% ~ 99.03%,准确率为 98.53% ~ 99.12%。结果表明,Edge - FastSCNN 模型能够准确预测出中后期玉米行间路径掩码,同时对不同的行间路径环境具有良好的适应性,鲁棒性较强。

(2)基于 Edge - FastSCNN 模型分割出的中后期玉米行间路径掩码,提出一种导航线检测方法。针对不同光照环境和遮挡情况的玉米行间路径进行导航线检测准确性实验,航向角偏差平均值为 1.15°~3.16°,平均像素横向距离为 1.89 ~ 3.41 像素。针对采集的视频进行导航线检测实时性实验,计算得出导航线检测算法的单帧图像平均处理时间为 90.04 ms。结果表明,本文提出的导航线检测算法具有较好的准确性、实时性和鲁棒性,适用于农业机械机器人的中后期玉米行间路径导航任务。

参 考 文 献

[1] 王晨, 张文博, 王济民. 我国粮食供给时空变迁及产能提升路径分析[J/OL]. 中国农业资源与区划. <https://link.cnki.net/urlid/11.3513.S.20231117.1641.004>.
WANG Chen, ZHANG Wenbo, WANG Jimin. The path analysis of spatial-temporal changes and productivity improvement of grain supply in China[J/OL]. Chinese Journal of Agricultural Resources and Regional Planning. <https://link.cnki.net/urlid/11.3513.S.20231117.1641.004>. (in Chinese)

[2] 鹿金秋, 王振营, 何康来, 等. 桃蛀螟在玉米、高粱和向日葵上发生为害与播期的关系[J/OL]. 应用昆虫学报. <https://link.cnki.net/urlid/11.6020.Q.20240322.1500.006>.
LU Jinqiu, WANG Zhenying, HE Kanglai, et al. Relationship between sowing date and occurrence of the yellow peach moth, *conogethes punctiferalis*, on corn, sorghum and sunflower crops[J/OL]. Chinese Journal of Applied Entomology. <https://link.cnki.net/urlid/11.6020.Q.20240322.1500.006>. (in Chinese)

[3] IMOLOAME E, JO O. Weed infestation, growth and yield of maize (*Zea mays* L.) as influenced by periods of weed interference[J]. Advances in Crop Science and Technology, 2017, 5(2):1-7.

[4] 刘爱红. 不同追肥方式对玉米生长和产量的影响[J]. 农业工程技术, 2023, 43(31): 23-24.
LIU Aihong. Effects of different topdressing methods on maize growth and yield[J]. Agricultural Engineering Technology, 2023, 43(31): 23-24. (in Chinese)

[5] 周湘铭. 阔叶类高垄作物视觉导航技术研究[D]. 郑州:河南农业大学, 2023.
ZHOU Xiangming. Research on visual navigation technology of broad-leaved crop for high ridge cultivation model[D]. Zhengzhou: Henan Agricultural University, 2023. (in Chinese)

[6] DE PONTE MÜLLER F. Survey on ranging sensors and cooperative techniques for relative positioning of vehicles[J]. Sensors, 2017, 17(2): 271.

[7] LI Z, XIE D, LIU L, et al. Inter-row information recognition of maize in the middle and late stages via LiDAR supplementary vision[J]. Frontiers in Plant Science, 2022, 13:1024360.

[8] 刘锦浩. 基于单目相机的移动机器人自主人员跟随方法研究[D]. 济南:山东大学, 2024.
LIU Jinhao. Research on autonomous person following method of mobile robot based on monocular camera[D]. Ji'nan: Shandong University, 2024. (in Chinese)

[9] 常江, 李春圣, 王嘉明, 等. 一种基于单目视觉农业机器人提取导航线的方法[J]. 佳木斯大学学报(自然科学版), 2023, 41(1): 92-95.
CHANG Jiang, LI Chunshen, WANG Jiaming, et al. Research on navigation path recognition method of monocular vision agricultural robot[J]. Journal of Jiamusi University(Natural Science Edition), 2023, 41(1): 92-95. (in Chinese)

[10] 苏筠皓. 复杂环境下中耕除草机器人视觉导航路径识别研究[D]. 天津:天津理工大学, 2023.
SU Junhao. Research on visual navigation path recognition of field tillage weeding robot in complex environment[D]. Tianjin: Tianjin University of Technology, 2023. (in Chinese)

[11] 杨洋, 张博立, 查家翼, 等. 玉米行间导航线实时提取[J]. 农业工程学报, 2020, 36(12): 162-171.
YANG Yang, ZHANG Boli, ZHA Jiayi, et al. Real-time extraction of navigation line between corn rows[J]. Transactions of the CSAE, 2020, 36(12): 162-171. (in Chinese)

[12] 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(Supp.1):18-26. (in Chinese)

[13] 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实验[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)

[14] 王侨, 孟志军, 付卫强, 等. 基于机器视觉的玉米苗期多条作物行线检测算法[J]. 农业机械学报, 2021, 52(4): 208-220.
WANG Qiao, MENG Zhijun, FU Weiqiang. Detection algorithm of multiple crop row lines based on machine vision in maize seedling stage[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 208-220. (in Chinese)

[15] 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)

[16] LIU X, QI J, ZHANG W, et al. Recognition method of maize crop rows at the seedling stage based on MS-ERFNet model[J]. Computers and Electronics in Agriculture, 2023, 211: 107964.

[17] 彭书博, 陈兵旗, 李景彬, 等. 基于改进 YOLOv7 的果园行间导航线检测[J]. 农业工程学报, 2023, 39(16): 131-138.

PENG Shubo, CHEN Bingqi, LI Jingbin, et al. Detection of the navigation line between lines in orchard using improved YOLOv7[J]. Transactions of the CSAE, 2023, 39(16): 131 – 138. (in Chinese)

[18] DIAO Z, GUO P, ZHANG B, et al. Navigation line extraction algorithm for corn spraying robot based on improved YOLOv8s network[J]. Computers and Electronics in Agriculture, 2023, 212: 108049.

[19] 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)

[20] 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)

[21] 赵岩, 张人天, 董春旺, 等. 采用改进 Unet 网络的茶园导航路径识别方法[J]. 农业工程学报, 2022, 38(19): 162 – 171.

ZHAO Yan, ZHANG Rentian, DONG Chunwang, et al. Navigation path recognition between tea ridges using improved Unet network[J]. Transactions of the CSAE, 2022, 38(19): 162 – 171. (in Chinese)

[22] ZHANG L, LI M, ZHU X, et al. Navigation path recognition between rows of fruit trees based on semantic segmentation[J]. Computers and Electronics in Agriculture, 2024, 216: 108511.

[23] LIU L, DU Y, LI X, et al. Trailer hopper automatic detection method for silage harvesting based improved U – net[J]. Computers and Electronics in Agriculture, 2022, 198: 107046.

[24] POUDEL R P K, LIWICKI S, CIPOLLA R. Fast – SCNN: fast semantic segmentation network[J]. arXiv Preprint, arXiv: 1902.04502, 2019.

[25] RAHMAN R, INDRIS C, ZHANG T, et al. On the real-time semantic segmentation of aphid clusters in the wild[C] //2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2023: 6299 – 6306.

[26] LU J, YANG Z, SUN Q, et al. A machine vision-based method for tea buds segmentation and picking point location used on a cloud platform[J]. Agronomy, 2023, 13(6): 1537.

[27] CHEN Y, GENG C, WANG Y, et al. Extraction method for centerlines of rice seedlings based on FAST – SCNN semantic segmentation[J]. Inmateh Agricultural Engineering, 2021, 64(2): 335 – 344.

[28] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510 – 4520.

[29] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770 – 778.

[30] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314 – 1324.

[31] TAKIKAWA T, ACUNA D, JAMPANI V, et al. Gated – SCNN: gated shape CNNs for semantic segmentation[C] // Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 5229 – 5238.

[32] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J]. arXiv Preprint, arXiv:1706.05587, 2017.