doi:10.6041/j.issn.1000-1298.2026.22.024

基于轻量化改进 DeepLabV3 + 的果园道路语义分割方法

马佳乐[1]　刘佳玮[1,2]　张付杰[2,3]　蒋仕飞[4]　王法安[2]　张永成[1]

(1. 塔里木大学机械电气化工程学院, 阿拉尔 843300; 2. 昆明理工大学现代农业工程学院, 昆明 650500;
3. 玉溪新天力农业装备制造有限公司, 玉溪 652699; 4. 云南烟草机械有限责任公司, 昆明 650106)

摘要: 针对果园行间道路光影紊乱、纹理模糊、遮挡频繁等现象,导致现有语义分割方法精度低、速度慢与模型轻量化难等问题,本文提出一种基于改进 DeepLabV3 + 的果园道路轻量化分割方法。采用轻量化 MobileNetV4ConvSmall 作为主干网络,提升模型推理效率;引入 ECA 通道注意力机制,增强模型对道路关键特征的鉴别能力;设计 Lite - ASPP 轻量化多尺度上下文模块,融合不同感受野特征,提升对道路边界与细长结构建模性能。自建果园道路数据集上试验结果表明,改进模型平均交并比与平均像素精度分别达 91.07% 和 95.96%,平均推理帧率为 511.08 f/s,参数量仅为 2.62×10^6,相较于 PSPNet、HRNet、SegFormer 及 DeepLabV3 + 等模型,本文方法在分割精度与实时性上均取得较优结果,在不同地表状态与干扰环境下均能实现连续、稳定的道路区域分割。研究结果可为果园作业装备实时视觉导航提供理论依据与技术参考。

关键词: 视觉导航; 语义分割; DeepLabV3 +; 轻量化网络; 果园道路

中图分类号: S126; TP242.6　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)22-0284-11　　**OSID**:

Orchard Road Semantic Segmentation Based on Lightweight Improved DeepLabV3 +

Ma Jiale[1]　Liu Jiawei[1,2]　Zhang Fujie[2,3]　Jiang Shifei[4]　Wang Faan[2]　Zhang Yongcheng[1]

(1. *College of Mechanical and Electronic Engineering*, *Tarim University*, *Alar* 843300, *China*
2. *College of Modern Agricultural Engineering*, *Kunming University of Science and Technology*, *Kunming* 650500, *China*
3. *Yuxi Xintianli Agricultural Equipment Manufacturing Co.*, *Ltd.*, *Yuxi* 652699, *China*
4. *Yunnan Tobacco Machinery Co.*, *Ltd.*, *Kunming* 650106, *China*)

Abstract: Orchard inter-row roads are commonly affected by uneven illumination, blurred ground textures, frequent occlusions, and complex-background interference, which reduce the segmentation accuracy, inference speed, and lightweight deployment capability of existing semantic segmentation methods. To solve these problems, a lightweight orchard road semantic segmentation method was proposed based on an improved DeepLabV3 + network, aiming to improve the operational efficiency of agricultural machinery in orchard inter-row environments. Firstly, the lightweight MobileNetV4ConvSmall network was adopted as the backbone to replace the original feature extraction network, thereby reducing computational complexity and significantly improving model inference efficiency. Secondly, an efficient channel attention (ECA) mechanism was introduced after the high-level features of the backbone network to enhance the discriminative representation of key road features under complex lighting and occlusion conditions. Furthermore, a Lite - ASPP lightweight multi-scale context module was designed to integrate features with different receptive fields, improving the representation capability for road boundaries, slender structures, and weak edge regions. Experimental evaluation was conducted on a self-constructed orchard road dataset. The results showed that the improved model achieved a mean intersection over

收稿日期: 2026 - 03 - 14　修回日期: 2026 - 05 - 26

基金项目: 国家自然科学基金青年项目(52402497)、云南省应用基础研究项目(202501AT070296、202401AU070196)、国家重点研发计划项目(2022YFD2002004)、云南省重大科技专项计划项目(202402AE090013、202402AE090009)、云南省科技厅科技计划项目(202402AE090042)和昆明理工大学人才培养基金项目(KKZ3202423160)

作者简介: 马佳乐(1991—),女,副教授,主要从事农业机械设计研究,E-mail: 120180014@taru.edu.cn

通信作者: 张永成(1991—),男,副教授,主要从事农业机械设计研究,E-mail: yczh0802@taru.edu.cn

union (mIoU) of 91.07% and a mean pixel accuracy (mPA) of 95.96%, with an average inference speed of 511.08 f/s and only 2.62×10^6 parameters. Compared with PSPNet, HRNet, SegFormer, and the DeepLabV3 +, the proposed method achieved better overall performance in segmentation accuracy and real-time inference. Under different ground conditions and environmental disturbances, the model can continuously and stably extract road regions, providing theoretical support and technical reference for real-time visual navigation of orchard operation equipment in practice.

Key words: visual navigation; semantic segmentation; DeepLabV3 +; lightweight network; orchard road

0　引言

果园作业环境存在显著的非结构化特征,包括光照剧烈变化、复杂树冠阴影、相似地表纹理以及频繁遮挡等,对农业装备自主导航环境感知能力提出了更高要求[1-3]。视觉导航作为自主导航中关键环节,需稳定、准确地提取道路或可行驶区域信息,以支持导航线生成与运动控制[4]。

当前较多研究者基于传统图像处理与几何模型进行导航线提取。刘思瑶等[5]通过融合彩色与深度图像,结合纹理-灰度梯度能量模型识别果树行,在复杂光照下取得了较高角度精度。肖珂等[6]利用 Kinect 实现葡萄园叶墙识别与喷施路径规划。张彦斐等[7]基于无人机多光谱影像,融合多种植被指数提取苹果树像元,并通过最小二乘拟合导航线,实现果园复杂环境导航。同时,在规则农田中,樊湘鹏等[8]采用改进 RANSAC 与最小二乘提取棉田作物行;孟庆宽等[9]通过颜色空间与模糊聚类分割田间道路,并结合直线扫描提升稳定性。这些方法在简单场景中有效,但在光照不均、阴影遮挡、杂草干扰及缺苗等复杂果园环境下,因依赖人工特征设计,容易出现分割断裂与导航线不稳定的现象。随着深度学习发展,基于语义分割的视觉导航方法成为研究热点。韩振浩等[10]将 U-Net 用于果园道路分割,并结合样条曲线拟合导航路径。刘慧等[11]基于 DeepLabV3 + 进行果园多类别分割,通过结构优化提升道路与背景的区分能力。周学成等[12]基于改进 DeepLabV3 + 提升了火龙果园道路分割精度。俞高红等[13]引入轻量化主干与注意力机制,实现无作物田垄高精度分割。侯文慧等[14]融合 MobileNetV3 与坐标注意力机制,实现垄间区域分割,并通过迁移学习验证了模型泛化能力。饶秀勤等[15]针对农田作物垄间导航路径识别问题,采用语义分割方法实现垄间可行驶区域提取,为复杂农田环境下农业机器人视觉导航路径识别提供了参考。近年来,研究重点逐渐转向模型轻量化与实时性转化。赵文锋等[16]提出 Orchard YOLOP 多任务感知框架,实现了道路、障碍物与作业区域的联合高精度感知。应仇凯等[17]基于 YOLO v8-Seg 实现草莓垄面的分割与导航控制。毕松等[18]基于 Mask R-CNN 分割道路与树干,并结合几何模型估计位姿参数,实现了导航方法从路径提取到位姿估计的拓展。

尽管现有方法在精度与鲁棒性上取得进展,但在复杂果园环境中仍面临精度、实时性与模型规模难以兼顾的挑战。轻量化模型对道路细长、边界模糊的特征提取能力不足,易导致边缘错分与断裂[19-20],模型常因参数量大、推理速度慢,难以满足实时作业需求[21-22]。

针对果园道路结构弱、遮挡干扰强、分割易断裂及实时性不足等问题,面向行间道路与可通行区域提取需求,本文提出一种基于改进 DeepLabV3 + 的轻量化道路分割方法。以 MobileNetV4ConvSmall 为轻量化主干,在高层特征中引入 ECA 通道注意力机制与 Lite-ASPP 多尺度提取模块,提升道路分割精度,以期为航线提取提供可靠语义支撑。

1　材料与方法

1.1　数据集构建

为满足果园道路语义分割算法在真实农机作业场景中训练与评估需求,构建面向果园道路导航的图像语义分割数据集。数据示例如图 1 所示。

在数据采集阶段,考虑果园道路在实际农机作业中常见的多样性工况,从地表状态、光照条件、植被干扰程度以及道路几何清晰度等多个维度进行系统覆盖。采集场景包括干燥裸土路面、松散土壤路面、车辆长期碾压形成的压实踏面、杂草覆盖路段以及雨后或灌溉后形成的潮湿路面等典型工况,共采集原始图像 1 410 幅。采集得到的原始图像统一按照语义分割任务要求进行精确标注。标注过程中以果园道路可通行区域为目标类别,其余区域统一标注为背景类别,形成二分类语义分割标注结果。

1.2　数据增强

为提升模型对果园道路复杂场景泛化能力,在训练阶段对输入图像进行数据增强处理。数据增强操作主要包括随机水平翻转、随机裁剪以及亮度与对比度轻微扰动,用于模拟不同拍摄角度和光照条件下果园道路场景。通过数据增强将数据集扩充至 4 230 幅,将数据集按比例 9∶1 划分为

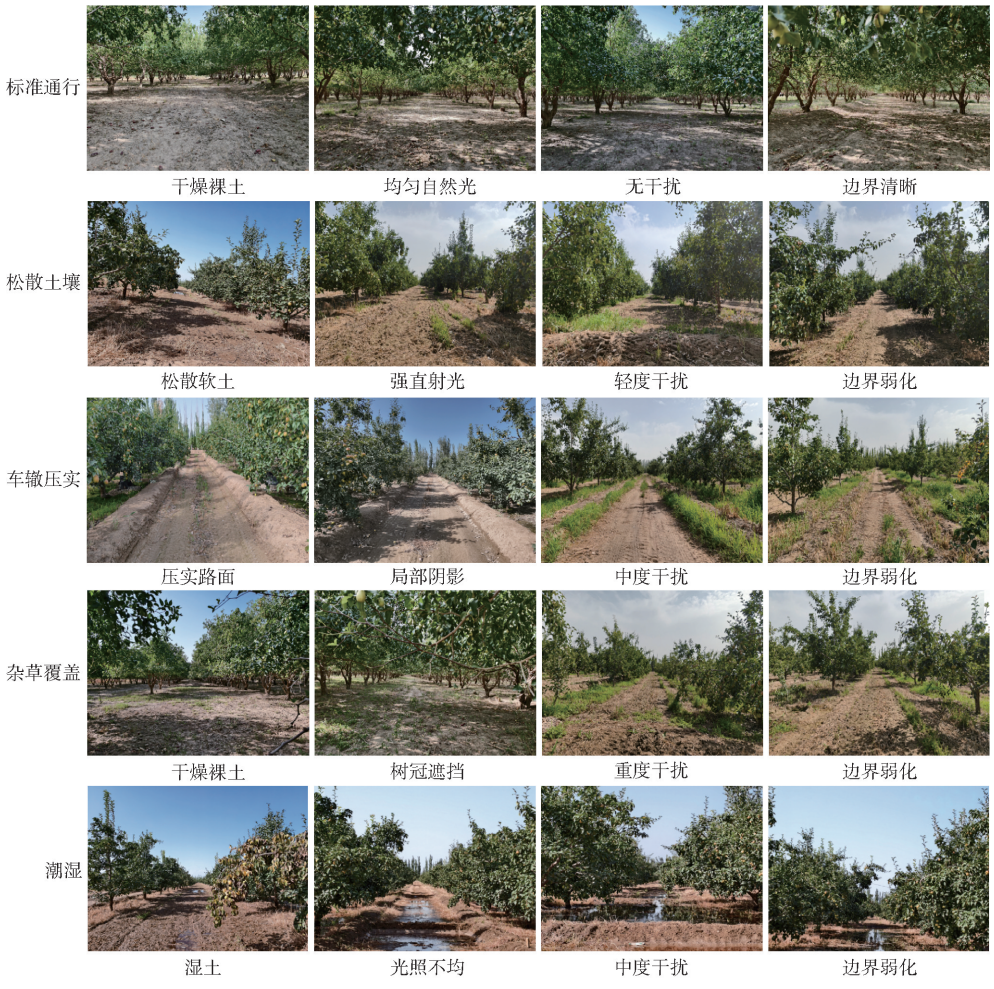


图 1 数据集示例

Fig. 1 Example images from dataset

训练集与验证集。

2 基于语义分割的识别模型

2.1 网络总体结构

DeepLabV3 + 整体网络结构采用编码-解码式语义分割框架,编码器由主干特征提取网络与空洞空间金字塔池化模块 2 部分组成。原始 DeepLabV3 + 网络通常以 Xception 作为主干网络,模型参数量较大、计算复杂度较高,在实际农业导航场景中难以满足实时性要求。为提高模型运行速度并适应农业装备在线导航应用需求,对 DeepLabV3 + 网络结构进行改进,采用轻量级 MobileNetV4ConvSmall 网络替换原有 Xception 主干特征提取网络。MobileNetV4 通过深度可分离卷积有效降低模型参数规模与计算量,使网络在保证基本特征表达能力的同时显著提升推理效率。然而,轻量化主干网络在高层语义特征中对关键通道信息的建模能力相对有限,易导致复杂场景下分割精度下降。针对上述问题,在主干网络输出的高层特征后引入高效通道注意力机制(Efficient channel atention, ECA)。通过对通道维度局部相关性建模,实现对关键特征通道自适应加权,在不增加额外计算量情况下增强特征判别能力,有效弥补轻量化网络在语义建模方面的不足。此外,考虑果园道路与田垄结构在尺度和形态上的多样性,对传统空洞空间金字塔池化模块进行轻量化改进,设计了轻量化空洞空间金字塔池化模块(Lite-ASPP)。通过并行引入不同感受野的卷积分支与全局池化分支,在降低参数量同时保留了多尺度特征融合能力,为后续解码阶段提供更加丰富的上下文信息。改进后 DeepLabV3 + 网络模型,如图 2 所示。

2.2 主干网络

传统卷积神经网络通常通过增加网络深度与宽度提升特征表达能力,但参数规模和计算复杂度显著增加,导致推理延迟增大,难以满足果园导航等嵌入式应用对实时性的要求。针对果园道路场景中计算资源受限、实时性要求高的特点,选用 MobileNetV4ConvSmall 作为网络的主干特征提取网络,以实现分割精度与推理效率之间的平衡。

MobileNetV4 在整体设计上继承了 MobileNetV3

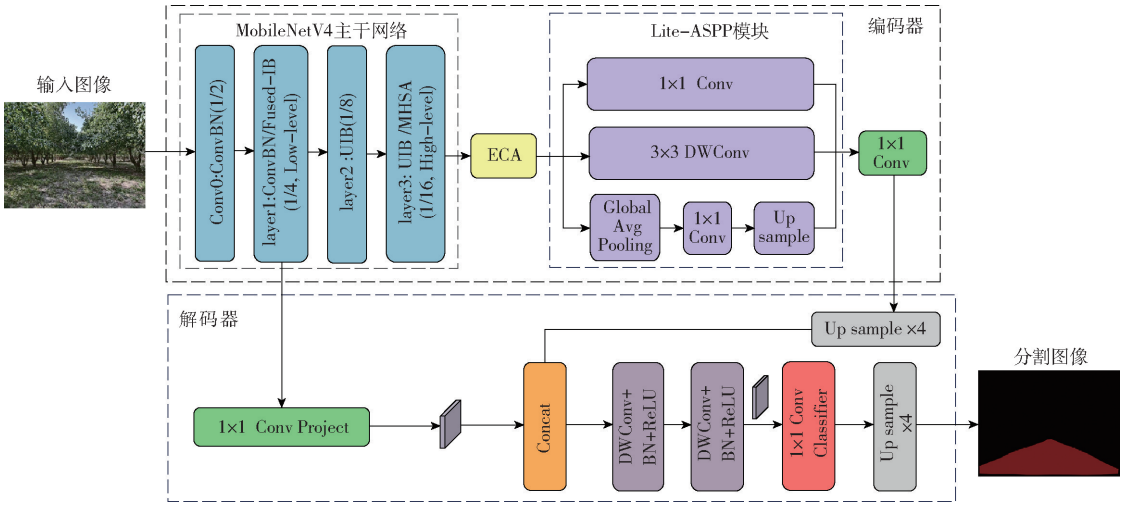


图 2　改进后 DeepLabV3 + 网络结构示意图

Fig. 2　Architecture of improved DeepLabV3 + network

轻量化思想，采用倒残差结构（Inverted residual block），通过“先扩展、再提取、后压缩”方式，在保持较强特征表达能力的同时有效降低模型参数量与计算开销。相比于传统卷积堆叠结构，倒残差结构更加注重通道维度高效利用，适用于光照变化剧烈、纹理复杂的果园道路场景。

在倒残差结构中，输入特征首先通过 1 × 1 点卷积进行通道扩展，将低维特征映射至高维空间；随后采用深度可分离卷积（DSC）进行空间特征提取；最后通过线性点卷积完成通道压缩，并与输入特征形成残差连接。其整体映射关系可表示为

$$F_{\text{out}} = F_{\text{in}} + P(\sigma(\text{BN}(\text{DWConv}(\text{Conv}_{1\times1}(F_{\text{in}}))))) \tag{1}$$

式中　F_{in}、F_{out}——模块输入与输出特征

$\text{Conv}_{1\times1}$——通道扩展操作

DWConv——深度可分离卷积

BN——归一化操作

$\sigma(\cdot)$——非线性激活函数

$P(\cdot)$——道压缩映射

2.3　ECA 高效通道注意力机制

轻量化主干网络显著降低计算量，同时高层特征在通道维度上的判别能力相对有限，容易在复杂背景和光照变化条件下引入冗余特征响应。为增强高层语义特征通道选择性，同时避免引入额外的参数开销，在主干网络输出高层特征后引入高效通道 ECA。

ECA 模块（图 3）以通道维度为建模对象，通过对全局语义信息自适应建模，实现不同通道重要性动态调节。ECA 利用一维卷积直接在通道维度上建模局部跨通道依赖关系，在保持通道信息完整性的同时显著降低模型复杂度。

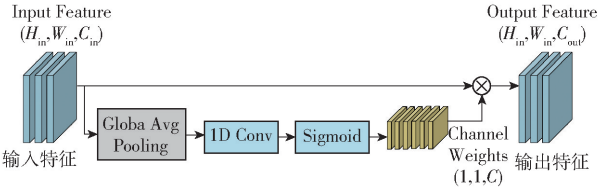


图 3　ECA 模块结构示意图

Fig. 3　Schematic of ECA module

设输入特征图为

$$\boldsymbol{X} \in \mathbf{R}^{C\times H\times W} \tag{2}$$

ECA 模块通过全局平均池化获得各通道全局描述向量

$$z = \text{GAP}(\boldsymbol{X}) \quad (z \in \mathbf{R}^{C}) \tag{3}$$

将 z 映射为一维序列，并通过尺寸为 k 的一维卷积对相邻通道之间的相关性进行建模，其通道权重计算过程可表示为

$$w = \sigma(\text{Conv1D}_k(z)) \tag{4}$$

式中　Conv1D_k——核尺寸为 k 的一维卷积操作

w——通道对应的注意力权重

最终，ECA 模块通过通道加权方式对输入特征进行重标定，其输出特征表示为

$$X_c' = w_c X_c \quad (c = 1,2,\cdots,C) \tag{5}$$

式中　X_c、X_c'——第 c 个通道输入与输出特征

w_c——第 c 个通道权重

2.4　Lite – ASPP 轻量化空洞空间金字塔池化模块

果园道路可通行区域在尺度、形态及空间分布上具有显著差异，仅依赖单一感受野的特征表达难以同时兼顾局部边界细节与整体结构信息。空洞空间金字塔池化（ASPP）模块通过引入多尺度感受野，在语义分割任务中得到广泛应用，但多分支空洞卷积结构往往带来较大的计算开销，不利于轻量化模型在嵌入式平台上的实时部署。

为保证多尺度上下文建模能力的同时进一步降

低计算复杂度,在 DeepLabV3 + 框架基础上设计了一种轻量化空洞空间金字塔池化模块 Lite - ASPP。该模块在保留 ASPP 多分支思想的基础上,通过减少分支数量并引入深度可分离卷积,有效控制参数规模与运算开销。

设输入高层特征图为

$$\boldsymbol{F} \in \mathbf{R}^{C \times H \times W} \tag{6}$$

轻量化空洞空间金字塔池化模块(Lite - ASPP)由 3 并行分支构成,如图 4 所示。第 1 分支采用 1 × 1 卷积对输入特征进行线性变换,用于提取局部语义信息,其输出表示为

$$\boldsymbol{F}_1 = \mathrm{Conv}_{1\times1}(\boldsymbol{F}) \tag{7}$$

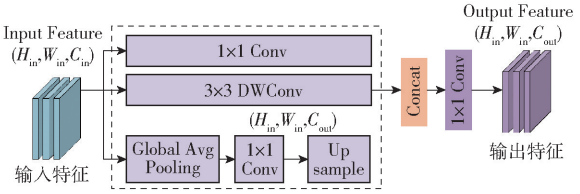


图 4 Lite - ASPP 模块示意图

Fig. 4 Schematic of Lite - ASPP module

第 2 分支引入基于深度可分离卷积的空洞卷积操作,在较低计算成本下扩大感受野,用于捕获中尺度上下文信息,其计算过程可表示为

$$\boldsymbol{F}_2 = \mathrm{DWConv}_{3\times3}^{(d)}(\boldsymbol{F}) \tag{8}$$

式中 $\mathrm{DWConv}_{3\times3}^{(d)}$——膨胀率为 d 的深度可分离空洞卷积

第 3 分支通过全局平均池化提取全局上下文特征,并利用 1 × 1 卷积进行通道映射,其输出为

$$\boldsymbol{F}_3 = \mathrm{Upsample}(\mathrm{Conv}_{1\times1}(\mathrm{GAP}(\boldsymbol{F}))) \tag{9}$$

式中 GAP(·)——全局平均池化操作

Upsample(·)——线性插值上采样至原特征分辨率

将 3 分支的输出特征在通道维度进行拼接,并通过 1 × 1 卷积进行特征融合与压缩,其整体输出可表示为

$$\boldsymbol{F}_{\mathrm{out}} = \mathrm{Conv}_{1\times1}(\mathrm{Concat}(\boldsymbol{F}_1,\boldsymbol{F}_2,\boldsymbol{F}_3)) \tag{10}$$

3 模型训练

3.1 试验平台

使用 PyTorch 深度学习框架完成语义分割模型的训练、验证和测试。试验计算机配置为 AMD 9950X CPU、NVIDIA GeForce RTX 5060ti(16 GB)GPU、48 GB 内存,操作系统为 Windows 11(64-bit),Python 版本为 3.9.24,深度学习框架采用 PyTorch 2.8.0,CUDA 版本为 12.8,cuDNN 版本为 9.1。

3.2 模型训练与参数设置

输入图像分辨率统一调整为 512 像素 × 512 像素,并进行归一化处理后输入网络;模型训练 100 个轮次,batch size 设为 2。优化器采用随机梯度下降(SGD)方法,动量系数设置为 0.9,权重衰减系数设置为 1×10^{-4}。初始学习率设定为 2×10^{-3},并根据 batch size 自适应调整实际学习率;学习率更新策略采用余弦退火衰减方式,在训练过程中逐步降低学习率以促进模型收敛。

3.3 损失函数

针对果园道路可通行区域分割任务属于典型的二分类像素级预测问题,在模型训练阶段采用交叉熵损失函数作为优化目标,对网络输出预测结果进行约束。网络输出经 Softmax 函数归一化后得到像素在各类别上的预测概率分布,记第 n 个像素在第 c 类上的预测概率为 $p_{n,c}$,其对应的真实标签为 $y_{n,c}\in\{0,1\}$。则像素级交叉熵损失函数定义为

$$\mathcal{L}_{\mathrm{CE}} = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c=1}^{C} y_{n,c}\lg p_{n,c} \tag{11}$$

式中 N——图像中像素总数 C——类数

该损失函数通过对每个像素位置预测结果与真实标签进行逐像素约束,引导模型学习道路区域与背景区域之间的判别特征。

3.4 评价指标

为评估所提出模型在果园道路与田垄语义分割任务中分割精度与实时性能,选取平均像素精度(mPA)、平均交并比(mIoU)、每秒传输帧率(FPS)、模型参数量以及单帧推理延迟(Latency)作为综合评价指标。

4 结果与分析

4.1 训练收敛性与稳定性

在相同训练配置与数据划分条件下,对 HRNet、DeepLabV3 +、PSPNet、SegFormer 及本文方法训练损失值进行对比,结果如图 5 所示。

由图 5 可知,各模型在训练初期损失均快速下降,表明网络能够在较少迭代内完成对道路与背景的基本判别特征学习;随后损失值下降速率逐步减缓并趋于平稳,表明优化过程由粗粒度区域分割转入对细节纹理与边界结构的精细拟合阶段。

进一步比较不同模型收敛特性可发现,本文方法在前若干个迭代次数降幅度更大,收敛速度更快,并在整个训练过程中始终维持最低的损失水平;同时,其中后期曲线波动更小、平台期更稳定,反映梯度更新过程更为平滑,优化更不易受局部噪声干扰。相比之下,PSPNet 与 SegFormer 在中期仍保持相对

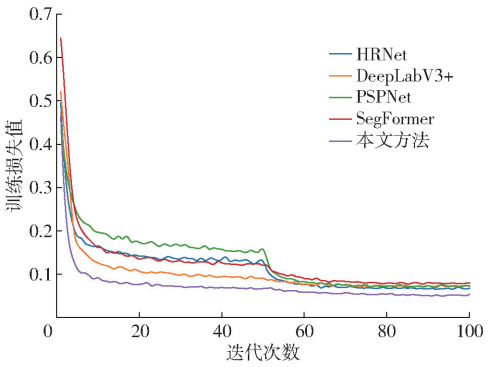


图 5　不同分割模型训练损失值曲线

Fig. 5　Training loss curves of different segmentation models

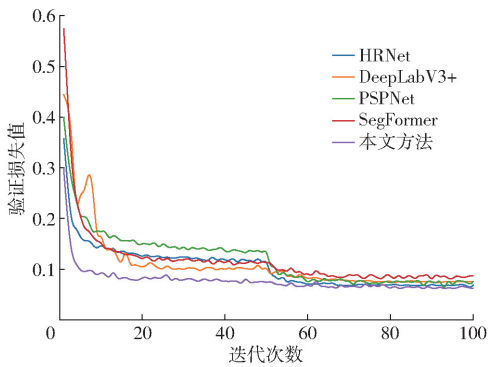


图 6　不同分割模型验证损失值曲线

Fig. 6　Validation loss curves of different segmentation models

较高损失值，收敛效率与最终收敛水平相对不足；HRNet 与 DeepLabV3 + 损失值虽逐渐降低并趋于稳定，但整体仍高于本文方法。

在相同训练策略与数据划分设置下，各模型验证损失值变化曲线如图 6 所示。由图 6 可知，整体趋势与训练阶段相呼应：各方法在前期验证损失快速下降，表明网络能够在较少迭代内形成对道路与背景的稳定判别；随后曲线进入缓慢下降并逐步趋于稳定，说明优化重点由粗粒度区域分割转向对遮挡、阴影与弱边界等困难样本的细节刻画，泛化性能提升呈渐进特征。

对比不同方法收敛水平与稳定性可发现，本文方法在整个训练周期内始终保持更低的验证损失，并在中后期表现出更小的波动幅度，未出现明显回升或持续震荡，反映其对验证分布的适配更加稳定。相比之下，DeepLabV3 + 在训练早期出现短暂的验证损失反弹，说明初期泛化过程存在一定不稳定；PSPNet 与 SegFormer 在中期维持较高的数值，下降速度相对较慢；HRNet 验证损失值曲线较为平滑，但最终收敛水平仍略高于本文方法。

此外，在第 50 个迭代次数附近，各曲线均出现同步下降拐点，与学习率衰减进入精调阶段的特征一致，说明低学习率有助于进一步降低验证误差并细化决策边界。在该阶段，本文方法仍保持最低验证损失并继续拉开与对比方法差距。

4.2　模型有效性验证

在相同试验环境、数据集划分及训练策略条件下，以 MobileNetV3 作为主干网络的 DeepLabV3 + 模型为基线，并逐步引入 MobileNetV4 - CS 主干网络、ECA 通道注意力机制与轻量化空洞空间金字塔池化模块（Lite - ASPP），构建消融试验。各模型 mIoU、mPA、帧率与参数量如表 1 所示。

表 1　不同模型性能

Tab. 1　Performance of different models

网络模型	主干网络	ECA	Lite - ASPP	mIoU/%	mPA/%	帧率/(f·s^{-1})	参数量
DeepLabV3 +	MobileNetV3	×	×	89.69	94.07	200.91	5.81×10^6
DeepLabV3 +	MobileNetV4 - CS	×	×	90.45	95.20	515.05	2.53×10^6
DeepLabV3 +	MobileNetV4 - CS	✓	×	90.74	95.55	500.02	2.53×10^6
DeepLabV3 +	MobileNetV4 - CS	×	✓	91.05	95.70	485.63	2.68×10^6
DeepLabV3 +	MobileNetV4 - CS	✓	✓	91.07	95.96	511.08	2.62×10^6

注：✓表示使用，×表示未使用。

由表 1 可知，将基线模型 MobileNetV3 主干替换为 MobileNetV4 - CS 后，在未引入 ECA 与 Lite - ASPP 情况下，模型 mIoU 相对基线提升 0.76 个百分点，mPA 提升 1.13 个百分点。在运行效率方面，推理帧率相对基线提升 314.14 f/s，由 200.91 f/s 提升至 515.05 f/s，模型参数量减少 3.28×10^6。

在 MobileNetV4 - CS 基线模型基础上引入 ECA 通道注意力机制后，模型 mIoU 相对基线提升 1.05 个百分点，达到 90.74%；mPA 提升 1.48 个百分点，达 95.55%。同时，推理帧率较基线提升 299.11 f/s，参数量仍减少 3.28×10^6。

当仅引入 Lite - ASPP 多尺度上下文模块时，模型 mIoU 相对基线提升 1.36 个百分点，提升至 91.05%，mPA 提升 1.63 个百分点，达到 95.70%。在单模块改进方案中取得了较大的精度提升幅度。在效率方面，模型推理帧率相对基线提升 284.72 f/s，参数量减少 3.13×10^6，多尺度建模在带来精度提升的同时引入了一定计算开销。

当引入高效通道注意力机制与轻量化空洞空间金字塔池化模块后,本文提出的完整模型在综合性能上表现最优,mIoU 与 mPA 相对基线分别提升 1.38、1.89 个百分点。在运行效率方面,推理帧率相对基线提升 310.17 f/s,模型参数量减少 3.19×10^6。与仅引入 Lite - ASPP 模型相比,组合模型在保持更高分割精度的同时,进一步提升了推理帧率并降低了参数量。

4.3 分割性能与效果

将与 PSPNet[23]、HRNet[24]、SegFormer[25] 以及 DeepLabV3 +[26] 等典型语义分割模型进行对比试验。各模型在相同数据集和评价指标条件下进行测试,定量分割结果如表 2 所示。

表 2 不同语义分割模型性能

Tab. 2 Performance of different semantic segmentation models %

模型	mIoU	mPA
PSPNet	89.23	94.10
HRNet	89.88	94.83
SegFormer	89.28	96.79
DeepLabV3 +	89.69	94.07
本文模型	91.07	95.96

由表 2 可知,本文模型 mIoU 和 mPA 分别达 91.07% 和 95.96%,在整体分割精度和类别识别稳定性方面均优于 PSPNet、HRNet、SegFormer 及 DeepLabV3 + 等对比模型。与 PSPNet 相比,本文模型 mIoU 提高 1.84 个百分点,mPA 提高 1.86 个百分点;与 HRNet 相比,mIoU 提高 1.19 个百分点。结果表明,本文模型在果园道路场景中对道路区域整体表达能力更强,能够有效提升分割结果准确性与一致性。

为进一步直观评估不同模型在果园复杂场景下分割性能,选取多种具有代表性的果园道路场景,对本文模型与 DeepLabV3 +、PSPNet、HRNet 及 SegFormer 等主流语义分割模型预测结果进行对比(图 7)。由图 7 可知,在标准通行场景下,各模型均能够较准确地识别果园道路区域,但部分模型在道路边界处仍存在轻微的区域外扩或轮廓不平滑现象。相比之下,本文模型在保持道路主体完整性的同时,道路边界轮廓更加平滑,分割区域与实际道路形态吻合度更高。

在杂草覆盖场景中,道路区域与周围植被纹理相似,容易引起误分割。对比结果表明,DeepLabV3 + 与 PSPNet 在部分区域出现将杂草区域误识别为道路的情况,SegFormer 在局部区域存在分割区域不连续的问题。本文模型能够较好地区分道路与杂草覆盖区域,道路区域整体连贯,误分割现象较少,表现出更强的语义判别能力。

在边界弱化场景中,由于道路与背景区域的颜色和纹理差异减小,多数模型在道路边缘位置出现不同程度的边界模糊或收缩现象。本文模型在该类场景下仍能够较完整地保持道路结构,道路轮廓连续性较好,说明其在多尺度上下文信息建模和边界细节保持方面具有一定优势。

在积水湿滑场景中,道路表面存在反光与局部纹理缺失,增加了分割难度。部分对比模型在积水区域出现道路区域断裂或局部缺失的情况,而本文模型仍能够较为完整地提取道路主体区域,仅在少数边缘位置出现轻微误差,整体分割结果更加稳定。

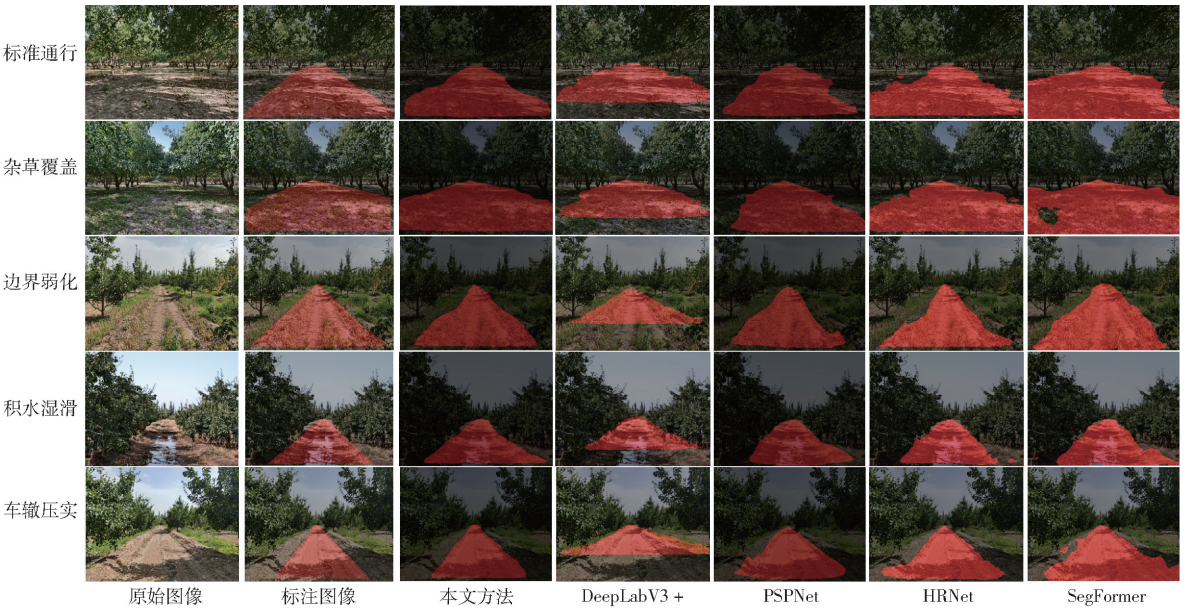


图 7 模型分割效果

Fig. 7 Visual results of model segmentation

在车辙压实场景中，道路表面结构复杂，存在明显的几何形态变化。对比结果显示，部分模型在道路中部或边缘位置出现不规则分割区域，而本文模型能够较好地适应道路形态变化，保持道路区域连续性和结构一致性。

4.4　实时推理性能

为评估不同语义分割模型在果园道路场景下实时推理能力，在相同硬件平台与测试环境条件下，对 PSPNet、HRNet、SegFormer、DeepLabV3 + 以及本文模型推理帧率、单帧推理延迟和参数量进行了对比，试验结果如表 3 所示。

表 3　实时推理性能

Tab. 3　Real-time inference performance

模型	帧率/(f·s^{-1})	单帧推理延迟/ms	参数量
PSPNet	232.55	4.30	2.41×10^6
HRNet	64.19	15.58	9.64×10^6
SegFormer	182.81	5.47	3.71×10^6
DeepLabV3 +	200.91	4.98	5.81×10^6
本文模型	511.08	1.64	2.62×10^6

由表 3 可知，与 DeepLabV3 + 相比，本文模型推理帧率由 200.91 f/s 提至 511.08 f/s，提升 310.17 f/s，约为原模型的 2.54 倍；单帧推理延迟由 4.98 ms 降至 1.64 ms，降低 3.34 ms，降幅约 67.1%；模型参数量由 5.81×10^6 降至 2.62×10^6，减少 3.19×10^6，显著降低了模型存储与计算开销。

与 PSPNet 相比，本文模型推理帧率由 232.55 f/s 提至 511.08 f/s，提升 278.53 f/s，约为 2 倍；单帧推理延迟由 4.30 ms 降至 1.64 ms，降低 2.66 ms；在参数规模方面，本文模型参数量为 2.62×10^6，与 PSPNet 的 2.41×10^6 接近，仍处于同一数量级。

与 SegFormer 相比，本文模型推理帧率由 182.81 f/s 提至 511.08 f/s，提升 328.27 f/s，约为 2.8 倍；单帧推理延迟由 5.47 ms 降至 1.64 ms，降低 3.83 ms；同时，模型参数量由 3.71×10^6 降至 2.62×10^6，减少 1.09×10^6。

与 HRNet 相比，本文模型实时性能优势更为显著。HRNet 推理帧率仅为 64.19 f/s，而本文模型达到 511.08 f/s，提升 446.89 f/s，约为 7.96 倍；单帧推理延迟由 15.58 ms 降至 1.64 ms，降低 13.94 ms；同时，模型参数量由 9.64×10^6 降至 2.62×10^6，减少 7.02×10^6。

5　在线视频推理与部署可行性验证

5.1　视频数据与场景设置

传统静态测试集评估主要反映单帧语义分割精度，但真实农机导航作业通常以视频流形式持续接收图像，系统性能不仅取决于单帧预测质量，更受端到端处理时延、连续帧输出抖动以及结果可持续可用性的影响。因此，仅依赖静态图像指标难以完整刻画模型在在线场景中的部署表现。

基于上述考虑，在无像素级真值标注的视频数据上对模型进行在线推理评测，以模拟实际部署时的输入流形式与处理链路。评测过程中采用逐帧推理与固定后处理策略，记录每帧预处理、网络推理、后处理及总耗时，并输出可视化叠加结果用于直观检验分割边界的连续性。试验采用实采 RGB 视频作为输入，视频分辨率为 1 280 像素 ×720 像素，帧率为 30 f/s，总帧数为 3 300。视频覆盖阴影、弱边界与强光等典型复杂工况如图 8 所示，用于检验模型在连续帧条件下的稳定推理能力。



图 8　视频工况样例

Fig. 8　Examples of video test scenarios

5.2　在线推理流程与部署实现

在线推理采用逐帧处理方式。对每一帧 $\boldsymbol{I}_t$ 执行尺寸归一化与张量化后输入网络，得到前景概率图 $\boldsymbol{P}_t$。为降低噪声并增强区域连通性，采用固定参数的后处理策略，包括分位数阈值化、形态学闭运算与小连通域剔除，得到二值分割结果 $\hat{\boldsymbol{M}}_t$。将 $\hat{\boldsymbol{M}}_t$ 的外轮廓以黄色边界叠加到原始帧中，并输出视频，评测流程如图 9 所示，图 10 展示连续帧分割效果与稳定性。

5.3　试验设置

评测在 batch = 1 条件下进行，输入尺寸固定为 640 像素 ×640 像素。为避免首次算子编译与缓存建立对统计结果造成偏差，采用全流程 warm-up，即在正式计时前重复执行预处理、推理与后处理若干次。计时过程中在 GPU 场景下使用同步操作保证测量精度。

5.4　在线视频推理性能

相比静态测试集评估，视频评测更能够反映模型在连续输入下实时性、稳定性以及系统响应能力。分析模型端到端推理延迟、延迟构成、时序稳定性以及推理帧率。

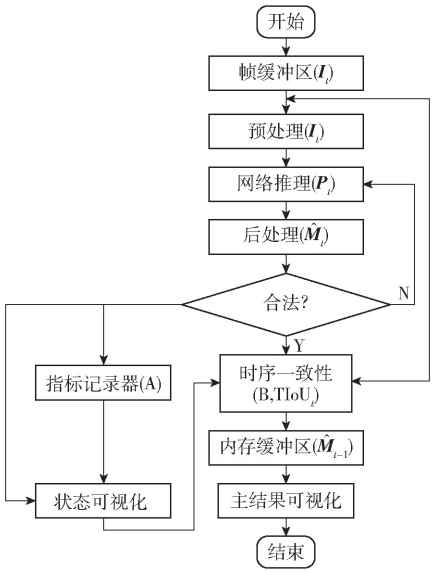


图 9 评测流程

Fig. 9 Evaluation workflow

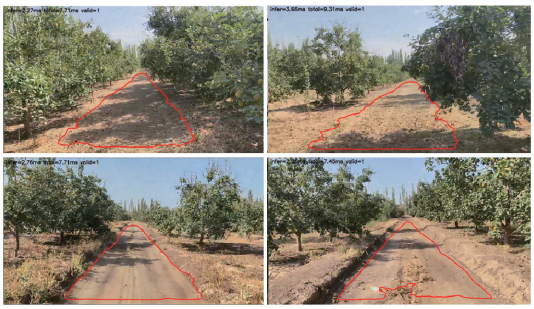


图 10 可视化结果示例

Fig. 10 Examples of segmentation visualization results

从图 11a 可看出,不同模型在视频推理过程中单帧延迟存在明显差异。总体而言,本文方法推理延迟最低,延迟中位数为 8 ms,分别低于 HRNet、PSPNet、SegFormer、DeepLabV3 + 模型 12、9、4、1.5 ms。本文方法箱体高度最小,说明延迟分布更加集中,推理时间波动较小,系统运行稳定性更高。

将端到端推理流程划分为 3 个阶段:预处理阶段、网络推理阶段、后处理阶段,如图 11b 所示。试验结果表明,网络推理阶段是整体延迟的主要来源。对于其他模型而言,该阶段占据总延迟主要比例。HRNet 平均总延迟 22 ms,其中推理阶段 20 ms,SegFormer 平均总延迟 17 ms,推理阶段 14 ms。DeepLabV3 + 平均延迟 12 ms,推理阶段约 10 ms。相比之下,本文方法在各阶段时间消耗均较低:预处理阶段 2.8 ms;推理阶段 2.6 ms;后处理阶段 2.9 ms,整体平均延迟 8.3 ms。本文模型在网络结构设计上具有较高的计算效率,推理阶段开销明显低于其他对比模型。

在视频应用场景中,单帧分割精度并不能完全反映模型实际性能。若模型在连续帧之间产生明显波动,将影响导航系统稳定性。进一步统计了连续帧预测结果之间的 Temporal IoU,用于衡量模型时序稳定性。

试验结果表明,DeepLabV3 + 的 Temporal IoU 中位数为 92%,本文方法为 88%,SegFormer 为 87%,HRNet 为 85%,PSPNet 为 81%。由图 11c 可知,本文方法在保持较高时序稳定性的同时,仍然实现

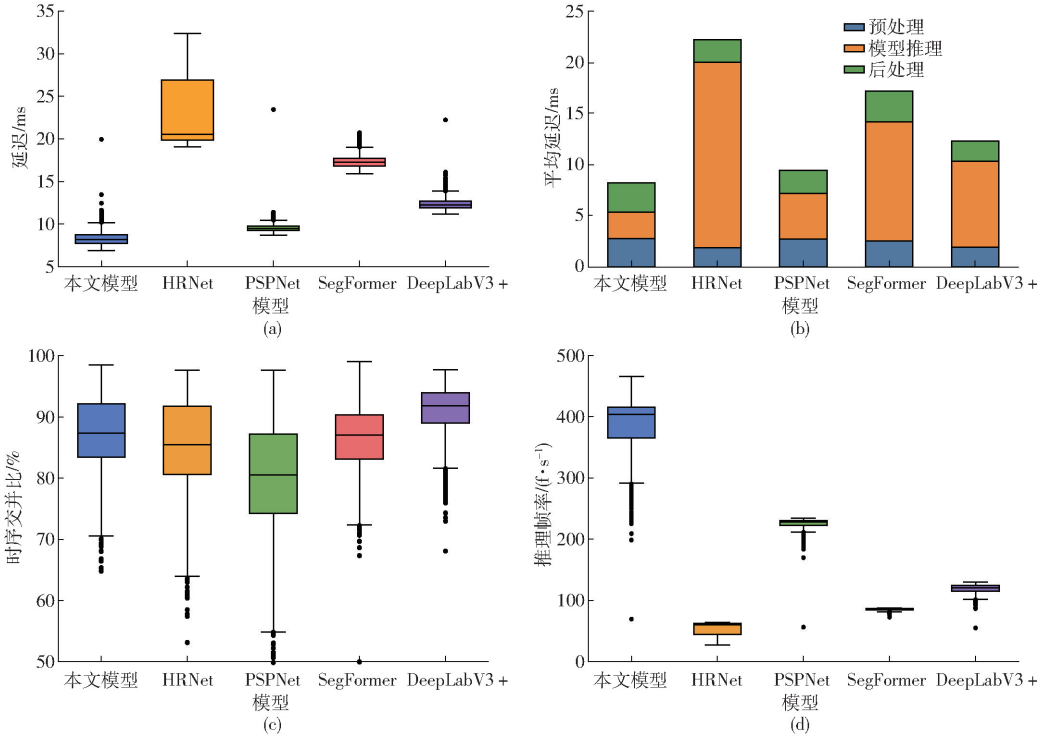


图 11 模型连续视频流性能

Fig. 11 Performance of model on continuous video stream

了更低的计算延迟。从箱线图可知，本文模型离群点较少，预测结果在连续帧之间变化较为稳定。

由图 11d 可知。本文模型与 HRNet、PSPNet、SegFormer、DeepLabV3 + 视频推理帧率分别达到 400、55、225、85、120 f/s。本文方法在推理速度方面具有明显优势。相比 HRNet、SegFormer 和 DeepLabV3 + ，本文模型推理速度分别提升约 7 倍、4.7 倍和 3.3 倍。

6　结论

(1)提出了一种基于改进 DeepLabV3 + 的果园道路轻量化语义分割方法。通过引入 MobileNetV4ConvSmall 轻量化主干网络，并结合 ECA 通道注意力机制与 Lite - ASPP 多尺度上下文模块，在保证模型轻量化与实时性的同时有效提升了道路区域分割精度与边界表达能力。

(2)改进后的网络结构在保持较小参数规模的前提下显著增强了对果园道路细长结构与模糊边界建模能力，平衡了分割精度与推理效率。

(3)自建果园道路数据集试验结果表明，本文模型 mIoU 和 mPA 分别达 91.07% 和 95.96%，推理帧率达到 511.08 f/s，单帧推理延迟仅为 1.64 ms，整体性能优于 PSPNet、HRNet、SegFormer 及 DeepLabV3 + 等对比模型。

(4)在多种典型果园复杂场景下，本文方法能够稳定、连续地提取道路区域，具有较好的鲁棒性和工程应用潜力，可为果园作业装备实时视觉导航理论依据与技术参考。

参　考　文　献

[1] 苗中华，朱子煜，张伟，等. 具身智能农业机器人关键技术与发展趋势[J]. 农业机械学报，2025，56(9)：212 - 239.
Miao Zhonghua，Zhu Ziyu，Zhang Wei，et al. Key technologies and development trends of embodied intelligent agricultural robots[J]. Transactions of the Chinese Society for Agricultural Machinery，2025，56(9)：212 - 239.(in Chinese)

[2] 张智刚，殷珍，吴潇，等. 农业机械自动导航技术研究进展[J]. 农业工程学报，2025，41(23)：1 - 13.
Zhang Zhigang，Yin Zhen，Wu Xiao，et al. Research progress on the automatic navigation technology for agricultural machinery[J]. Transactions of the CSAE，2025，41(23)：1 - 13.(in Chinese)

[3] 窦汉杰，陈震宇，翟长远，等. 果园智能化作业装备自主导航技术研究进展[J]. 农业机械学报，2024，55(4)：1 - 22.
Dou Hanjie，Chen Zhenyu，Zhai Changyuan，et al. Research progress on autonomous navigation technology for orchard intelligent equipment[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(4)：1 - 22.(in Chinese)

[4] 张少侠，闫建伟，蒙超，等. 基于机器视觉的农业车辆及机器人导航技术研究进展[J]. 计算机工程与应用，2025，61(12)：12 - 27.
Zhang Shaoxia，Yan Jianwei，Meng Chao，et al. Research progress of agricultural vehicle and robot navigation technology based on machine vision[J]. Computer Engineering and Applications，2025，61(12)：12 - 27.(in Chinese)

[5] 刘思瑶，王晓燕，陈相安，等. 用动态图层融合与纹理-灰度梯度能量模型的果树行识别[J]. 农业工程学报，2022，38(8)：152 - 160.
Liu Siyao，Wang Xiaoyan，Chen Xiang'an，et al. Fruit tree row recognition with dynamic layer fusion and texture-gray gradient energy model[J]. Transactions of the CSAE，2022，38(8)：152 - 160.(in Chinese)

[6] 肖珂，高冠东，马跃进. 基于 Kinect 视频技术的葡萄园农药喷施路径规划算法[J]. 农业工程学报，2017，33(24)：192 - 199.
Xiao Ke，Gao Guandong，Ma Yuejin. Pesticide spraying route planning algorithm for grapery based on Kinect video technique[J]. Transactions of the CSAE，2017，33(24)：192 - 199.(in Chinese)

[7] 张彦斐，魏鹏，宫金良，等. 复杂环境中苹果树识别与导航线提取方法[J]. 农业机械学报，2022，53(10)：220 - 227.
Zhang Yanfei，Wei Peng，Gong Jinliang，et al. Apple tree recognition and navigation line extraction in complex environment[J]. Transactions of the Chinese Society for Agricultural Machinery，2022，53(10)：220 - 227.(in Chinese)

[8] 樊湘鹏，许燕，周建平. 棉田施药机器人视觉导航方法与田间试验[J]. 农业工程学报，2025，41(6)：52 - 61.
Fan Xiangpeng，Xu Yan，Zhou Jianping. Visual navigation method and field experiment of cotton spraying robots[J]. Transactions of the CSAE，2025，41(6)：52 - 61.(in Chinese)

[9] 孟庆宽，杨晓霞，张漫，等. 基于语义分割的非结构化田间道路场景识别[J]. 农业工程学报，2021，37(22)：152 - 160.
Meng Qingkuan，Yang Xiaoxia，Zhang Man，et al. Recognition of unstructured field road scene based on semantic segmentation model[J]. Transactions of the CSAE，2021，37(22)：152 - 160.(in Chinese)

[10] 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)

[11] 刘慧，姜建滨，沈跃，等. 基于改进 DeepLabV3 + 的果园场景多类别分割方法[J]. 农业机械学报，2022，53(11)：255 - 261.
Liu Hui，Jiang Jianbin，Shen Yue，et al. Multi-category segmentation method of orchard scene based on improved DeepLabV3 +[J]. Transactions of the Chinese Society for Agricultural Machinery，2022，53(11)：255 - 261.(in Chinese)

[12] 周学成，肖明玮，梁英凯，等. 基于改进 DeepLabv3 + 的火龙果园视觉导航路径识别方法[J]. 农业机械学报，2023，54(9)：35 - 43.

Zhou Xuecheng, Xiao Mingwei, Liang Yingkai, et al. Navigation path recognition between dragon orchard using improved DeepLabV3 + [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 35 - 43. (in Chinese)

[13] 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)

[14] 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)

[15] 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)

[16] 赵文锋, 黄袁爵, 钟敏悦, 等. 基于 OrchardYOLOP 的火龙果园多任务视觉感知方法[J]. 农业机械学报, 2024, 55(11): 160 - 170.
Zhao Wenfeng, Huang Yuanjue, Zhong Minyue, et al. Multi-task visual perception method in dragon orchards based on OrchardYOLOP[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 160 - 170. (in Chinese)

[17] 应仇凯, 程泓超, 马锃宏, 等. 基于 YOLO v8 - Seg 的地栽草莓采摘机器人垄面视觉导航控制方法[J]. 农业机械学报, 2024, 55(增刊 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(S1): 9 - 17. (in Chinese)

[18] 毕松, 王宇豪. 果园机器人视觉导航行间位姿估计与果树目标定位方法[J]. 农业机械学报, 2021, 52(8): 16 - 26.
Bi Song, Wang Yuhao. Inter-line pose estimation and fruit tree location method for orchard robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(8): 16 - 26. (in Chinese)

[19] Zhang Liang, Li Ming, Zhu Xinghui, et al. Navigation path recognition between rows of fruit trees based on semantic segmentation[J]. Computers and Electronics in Agriculture, 2024, 216: 108511.

[20] 毕松, 余鑫. 基于 3D 点云分析的果园行间穿梭机器人路径规划方法[J]. 农业机械学报, 2024, 55(10): 37 - 50.
Bi Song, Yu Xin. Path planning method for inter-row shuttle in densely planted orchards[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 37 - 50. (in Chinese)

[21] Liu Hui, Zeng Xiao, Shen Yue, et al. A single-stage navigation path extraction network for agricultural robots in orchards[J]. Computers and Electronics in Agriculture, 2025, 229: 109687.

[22] Guan Xianlu, Wan Huan, He Zixuan, et al. Pomelo - Net: a lightweight semantic segmentation model for key elements segmentation in honey pomelo orchard for automated navigation[J]. Computers and Electronics in Agriculture, 2025, 229: 109760.

[23] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 2881 - 2890.

[24] Sun K, Xiao B, Liu D, et al. Deep high-resolution representation learning for human pose estimation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019: 5693 - 5703.

[25] Xie E, Wang W, Yu Z, et al. SegFormer: simple and efficient design for semantic segmentation with transformers[C]// Proceedings of the 35th International Conference on Neural Information Processing Systems (NeurIPS), 2021: 12077 - 12090.

[26] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]// Computer Vision-ECCV 2018. Cham: Springer, 2018: 833 - 851.