

基于改进 DeepLabv3 + 的火龙果园视觉导航路径识别方法

周学成^{1,2} 肖明玮¹ 梁英凯¹ 商枫楠¹ 陈 桥¹ 罗陈迪¹

(1. 华南农业大学工程学院, 广州 510642; 2. 南方农业机械与装备关键技术教育部重点实验室, 广州 510642)

摘要: 针对视觉导航系统应用在火龙果园环境中面临干扰因素多、图像背景复杂、复杂模型难以部署等问题, 本文提出了一种基于改进 DeepLabv3 + 网络的火龙果园视觉导航路径识别方法。首先, 采用 MobileNetV2 取代传统 DeepLabv3 + 的主干特征提取网络 Xception, 并将空间金字塔池化模块 (Atrous spatial pyramid pooling, ASPP) 中的空洞卷积替换成深度可分离卷积 (Depthwise separable convolution, DSC), 在提升模型检测速率的同时大幅减少了模型的参数量和内存占用量; 其次, 在特征提取模块处引入坐标注意力机制 (Coordinate attention, CA), 增强了模型的特征提取能力; 最后, 通过设计的导航路径提取算法对网络模型分割出的道路掩码区域拟合出导航路径。实验结果表明: 改进后的 DeepLabv3 + 的平均交并比和平均像素准确率分别达到 95.80% 和 97.86%, 相较原模型分别提升 0.79、0.41 个百分点。同时, 模型内存占用量只有 15.0 MB, 和原模型相比降低 97.00%, 与 Pspnet 和 U-net 模型相比则分别降低 91.57%、91.02%。另外, 导航路径识别精度测试结果表明平均像素误差为 22 像素、平均距离误差 7.58 cm。已知所在果园道路宽度为 3 m, 平均距离误差占比为 2.53%。因此, 本文研究方法可为解决火龙果园视觉导航任务提供有效参考。

关键词: 火龙果园; 导航路径识别; 视觉导航; 语义分割; 坐标注意力机制; 深度可分离卷积

中图分类号: TP242.6 **文献标识码:** A **文章编号:** 1000-1298(2023)09-0035-09

OSID:



Navigation Path Recognition between Dragon Orchard Using Improved DeepLabv3 + Network

ZHOU Xuecheng^{1,2} XIAO Mingwei¹ LIANG Yingkai¹ SHANG Fengnan¹ CHEN Qiao¹ LUO Chendi¹

(1. College of Engineering, South China Agricultural University, Guangzhou 510642, China

2. Key Laboratory of Key Technology on Agricultural Machine and Equipment, Ministry of Education, Guangzhou 510642, China)

Abstract: Visual navigation has the advantages of low cost, wide applicability and high degree of intelligence, so it is widely used in orchard navigation tasks. Therefore, how to quickly and accurately identify the navigation path is a key step to achieve visual navigation. Aiming at the problems of multiple interference factors and complex image background in the application of visual navigation system in dragon orchard environment, a visual navigation path recognition method was proposed for dragon orchard based on improved DeepLabv3 + network. Firstly, the traditional DeepLabv3 + backbone feature extraction network was replaced by MobileNetV2 from Xception, and the atrous convolution in atrous spatial pyramid pooling (ASPP) was replaced with depthwise separable convolution(DSC). While improving the model detection rate, the number and memory footprint of model parameters were greatly reduced. Secondly, coordinate attention (CA) was introduced at the feature extraction module, which was helpful for the model to locate and identify road areas. Then, experiments were conducted on a self-built dragon orchard road dataset containing three different road conditions. The results showed that compared with the traditional DeepLabv3 +, the MIoU and MPA of the improved DeepLabv3 + were increased by 0.79 percentage points and 0.41 percentage points, respectively, reaching 95.80% and 97.86%. Frames per second (FPS) was increased to 57.89 f/s, and the number of parameters and memory footprint were reduced by 92.92% and 97.00%, respectively, to 3.87×10^6 and 15.0 MB. The recognition results of the improved model on the orchard road were verified on the test set, indicating that the model had good robustness and anti-interference. In addition, comparing the proposed model with Pspnet and U-net

收稿日期: 2023-02-24 修回日期: 2023-05-04

基金项目: 国家重点研发计划项目(2017YFD0700602)

作者简介: 周学成(1968—),男,教授,博士,主要从事机器视觉与智能检测研究,E-mail: zxcem@scau.edu.cn

networks, the results showed that the improved models offered significant advantages in detection rate, amount of parameters, and model size, making them more suitable for deployment to embedded devices. According to the segmentation results of the model, the edge information on both sides of the road was extracted, the road boundary line was fitted by the least squares method, and finally the navigation path was extracted by the angle bisector line fitting algorithm. The navigation path recognition accuracy was tested in three different road environments, and the test results showed that the average pixel error was 22 pixels and the average distance error was 7.58 cm. The road width of the orchard in this test was 3 m, and the average distance error accounted for only 2.53%. Therefore, the research result can provide an effective reference for the visual navigation task of dragon orchard.

Key words: dragon orchard; navigation path recognition; visual navigation; semantic segmentation; coordinate attention; depthwise separable convolution

0 引言

水果产业已成为我国继粮食和蔬菜之后的第3大农业种植产业,我国水果种植面积和产量常年稳居世界首位^[1]。尽管我国水果产业已经逐步实现规模化种植,但是机械化采摘程度依然较低,大多仍以人工采摘为主,和发达国家相比存在明显差距。由于果园环境多为半结构化环境,大型机械作业受限,限制了果园机械智能化的发展^[2]。因此我国迫切需要提高果园机械化、智能化水平,提高生产效率^[3]。视觉导航作为果园机械提升智能化水平的关键技术之一,具有成本低、信息丰富等特点,适用于不规则地块^[4]。而且有效弥补了果园因树叶遮挡、卫星信号弱而无法进行导航的情况^[5]。能够实时准确地识别导航路径是当前国内外学者研究的重点。

目前在果园环境下的导航路径识别研究中,研究人员通常将果树树干或者作物行作为获取导航信息的参照物,利用参照物具有的颜色、形态和纹理特征运用传统的图像处理技术提取出导航路径。目前已有针对苹果园^[6]、黄瓜园^[7]、橙园^[8]、枸杞园^[9]、桃园^[10]等环境的视觉导航研究。但是,在设施农业、果园等复杂环境中,图像处理算法易受光照、阴影的影响^[11]。而且果园道路属于非结构化道路,无明显道路边界线,传统图像处理算法不能完全适用。

近年来随着深度学习技术的不断发展,许多研究人员将深度学习技术应用在导航路径的识别研究中^[12-15],与传统的语义分割方法相比,基于深度学习的语义分割方法能获取更多、更高级的语义信息来表达图像中的信息^[16]。SONG等^[17]利用全卷积网络(FCN)对小麦、地面和背景进行语义分割,进而拟合出导航路径。KIM等^[18]提出了一种半结构化环境自主路径检测方法实现路径区域分割。YANG等^[19]提出了一种基于神经网络和像素扫描的可视化导航路径提取方法。韩振浩等^[20]提出了一种基于U-Net网络的果园视觉导航路径识别方法。

以上基于深度学习算法的导航路径识别研究为

本研究提供了借鉴。然而在火龙果园导航路径检测识别任务中,果园道路存在的杂草以及掉落的枝条,使得道路的边界信息模糊,增加了导航路径的识别难度。而且大多数研究较少关注模型的部署问题,所提出的网络结构较为复杂,使得模型的参数量大,不利于部署至硬件条件有限的果园视觉导航系统。为此,针对视觉导航应用在果园环境中面临干扰因素多、图像背景复杂、复杂模型难以部署等问题,本文以火龙果园的自然环境为研究对象,提出一种基于改进DeepLabv3+网络的果园视觉导航路径识别方法。本研究选择轻量化的MobileNetV2替换原网络中的主干特征提取网络,并将空间金字塔池化模块(Atrous spatial pyramid pooling, ASPP)中的空洞卷积替换成深度可分离卷积(Depthwise separable convolution, DSC),以降低模型的复杂度。为了提高模型对道路特征的提取能力,考虑在特征提取模块处引入坐标注意力机制(Coordinate attention, CA);最后利用网络模型生成的道路掩码,得到道路的边界信息并通过最小二乘法拟合道路边界,再运用角平分线算法拟合出导航路径。

1 材料与方法

1.1 图像数据获取

本研究所需的果园图像数据于2022年7月采自广州市番禺区火龙果种植园。拍摄设备为英特尔公司生产的D435i深度相机,图像数据通过USB接口传输并保存在计算机内,自动曝光;图像分辨率为1920像素×1080像素,以PNG格式存储,图像采集帧率为30 f/s。数据采集时将深度相机固定在相机支架上,向下倾斜10°,拍摄方向为道路正前方。根据研究目标,本文只对果园垄间道路进行图像数据采集,不包括果园地头以及果树行间。为了提高样本的多样性,分别在不同道路条件下共采集原始图像456幅,如图1所示。

1.2 果园道路数据集制作

通过相机直接获取的原始图像分辨率较高,在



图 1 火龙果园道路图像示例

Fig. 1 Example of dragon orchard roads

训练时会占用过多显存,降低训练速度,同时图像没有语义标签,需要预先进行标注,才能传入神经网络进行训练。为了减少模型训练时间,将 456 幅原始图像尺寸等比例缩放为 960 像素 $\times$ 540 像素,再使用像素级标注工具 Lableme 进行语义标注,标注后的文件以 .json 格式存储。针对现场采集的数据样本量不足,本文通过几何变换(平移、旋转)与颜色变换(对比度、亮度)进行数据增强。增强后的图像共 1 074 幅,按照 8:1:1 比例划分为训练集、验证集和测试集。训练集用于训练深度网络模型参数权重;验证集用于训练过程中对模型参数进行调优;测试集用于评估最终模型的泛化能力。

2 果园道路场景语义分割

2.1 改进的 DeepLabv3 + 语义分割算法

DeepLabv3 + 网络被称为语义分割网络的新高峰,但也存在不足。首先,为了追求分割精度,选择了网络层数较多、参数量大的 Xception 作为特征提取网络,同时 ASPP 模块中采用空洞卷积,使得模型参数量增加,提高了模型的复杂度;另外,这些特点也对硬件提出了更高要求^[21]。为了使得移动器

人能够实时准确地识别果园道路,确保该网络模型能够部署在嵌入式设备上,本文针对上述问题对传统的 DeepLabv3 + 网络进行了如下改进:首先,为了减小参数计算量并降低模型的复杂度,将 DeepLabv3 + 模型中原本用于主干特征提取的 Xception 网络更换成更为轻量级的 MobileNetV2。为了增强模型学习特征的表达能力,在主干提取网络输出的高层特征层处添加 CA 模块。为了提高模型的检测速率,减小内存占用量,将 ASPP 模块中的空洞卷积替换成深度可分离卷积。改进后的网络结构如图 2 所示。

2.2 轻量化特征提取模块

传统的卷积神经网络通过扩充网络深度和广度,提高网络模型准确性,但也存在复杂度高、运行速度慢等问题。MobileNetV2 是由谷歌团队在 2018 年提出的高性能轻量化的卷积神经网络,相对于 MobileNetV1 而言准确率更高,模型更小^[22]。MobileNetV2 采用一种具有线性瓶颈的残差结构,该模块将输入的低维压缩表示首先扩展到高维并用轻量级深度卷积进行过滤。随后用线性卷积将特征投影回低维表示。最后采用跨连接层将输入特征与输

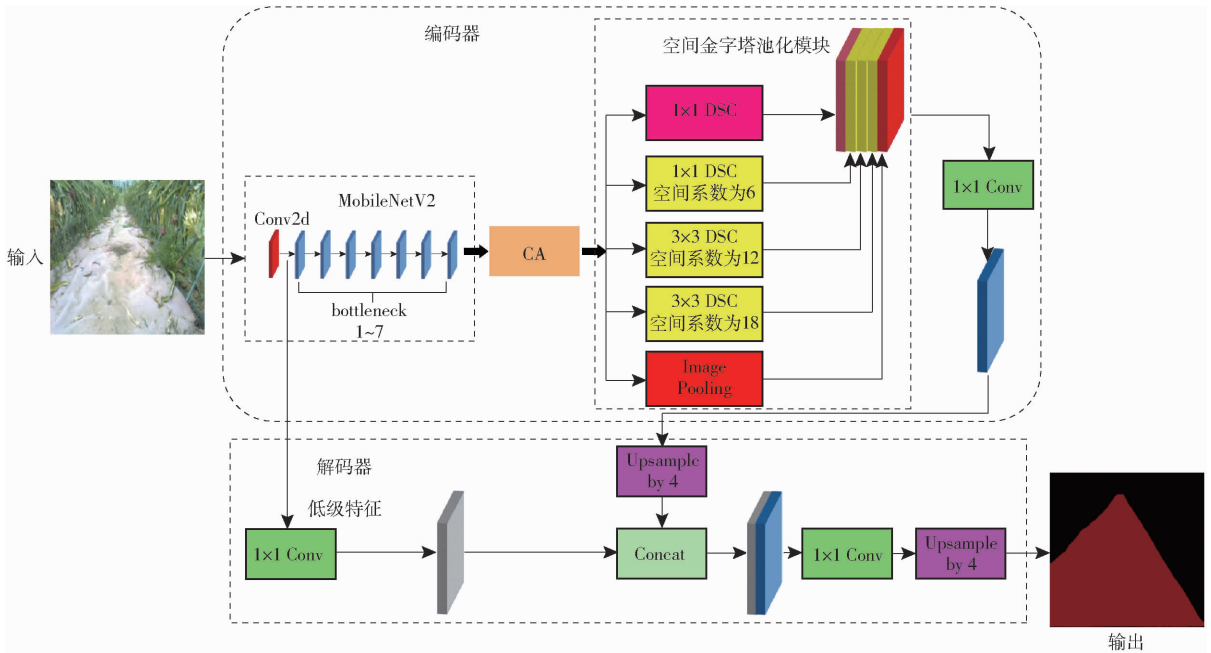


图 2 改进后的 DeepLabv3 + 网络结构图

Fig. 2 Improved DeepLabv3 + network structure diagram

出特征相加,从而增加网络的实时性和准确性。为了让 MobileNetV2 模块能够适用于语义分割,本文对该网络结构做了如下修改:将第 1 部分用于提取特征的 3×3 卷积块以及包含多个深度可分离卷积的中间部分保留,把包含全局平均池化层与特征分类层的第 3 部分去掉。

2.3 CA 模块

注意力机制常用来告诉模型需要更关注哪些内容和哪些位置,已经被广泛使用在深度神经网络中来加强模型的性能。HOU 等^[23]为轻量级网络设计提出了新的注意力机制,该机制将位置信息嵌入到了通道注意力中,称为坐标注意力机制(Coordinate attention, CA)。

不同于通道注意力将输入通过 2 维全局池化转化为单个特征向量,CA 将通道注意力分解为两个沿着不同方向聚合特征的 1 维特征编码过程。这样,可以沿一个空间方向捕获远程依赖关系,同时可以沿另一空间方向保留精确的位置信息。然后,将生成的特征图分别编码,将其互补地应用于输入特征图,以增强关注对象的表示。该模块的结构如图 3 所示,图中 C 表示通道数, H 表示特征图高度, W 表示特征图宽度。

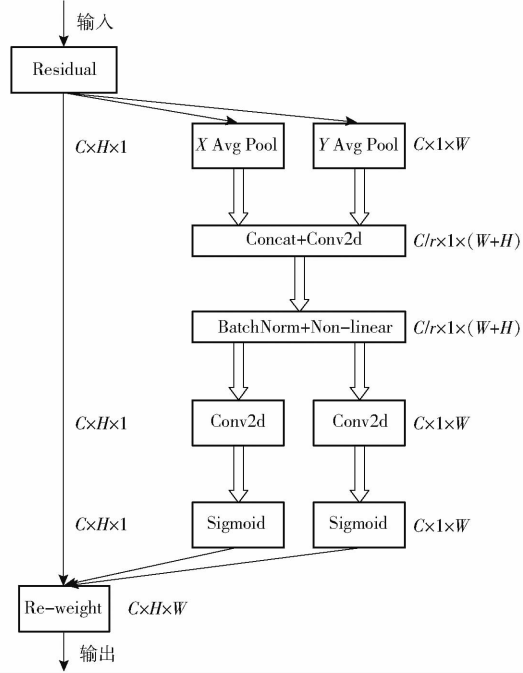


图 3 坐标注意力机制结构图
Fig. 3 Coordinate attention mechanism structure diagram

2.4 网络损失函数

损失函数是一种用来度量模型预测值与真实值之间差异程度的函数,损失函数的值越小,模型的鲁棒性就越好。为了解决处理前景和背景像素数量之间存在严重不平衡的问题,MILLETARI 等^[24]提出

了 DiceLoss。本文将其作为损失函数,其表达式为

$$L_{Dice} = 1 - Dice = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \tag{1}$$

式中 L_{Dice} ——损失函数 $Dice$ ——损失系数
 X ——真实分割图像的像素类别
 Y ——模型预测分割图像的像素类别
 $|X \cap Y|$ —— X 和 Y 的点乘结果之和
 $|X|$ —— X 对应图像中的像素之和
 $|Y|$ —— Y 对应图像中的像素之和

3 导航路径拟合

本文研究所涉火龙果园采用较为规范化的种植模式,果园道路较为平直,采摘机器人可将道路中线作为导航路径。常用的直线拟合算法包括霍夫变换、随机采样一致性、最小二乘法等^[25]。果园道路非结构化,道路边缘信息不规则,为此本文利用网络模型生成的道路掩码,提出一种基于道路边界拟合的角平分线提取导航路径算法。通过逐行扫描提取道路左右边缘信息点,再根据边缘信息点用最小二乘法拟合出道路左右边界线,最后利用角平分线算法拟合导航路径。具体流程如图 4 所示。

3.1 道路边缘信息提取

首先对模型生成的道路掩码区域二值化,使得掩码区域像素值为 255,背景区域像素为 0。道路掩码区域可以近似看成上底边有凸起的梯形区域,如图 5 所示。为了提取道路的边缘信息,首先需要确定掩码区域的顶点。逐行扫描像素,以掩码区域图像的前景首行像素的所有像素的中点位置为顶点,并记录下此时顶点的坐标 (W, H)。其次以顶点的横坐标 W 当作逐行扫描的分界线,以分界线为基准,分别向左、向右扫描。如果向左扫描时某个像素点的相邻两列像素值不同,则该点记为左边界点。同理向右扫描时,如果某个像素点的相邻两列像素值不同,则记为右边界点,直到整幅图像像素遍历完毕。

3.2 基于角平分线算法拟合导航路径

根据得到的道路边缘信息,用最小二乘法拟合出道路边界。但根据果园导航的需求,如果选择道路边线当作导航路径会造成较大的导航偏差,而且也不利于采摘,为了提高导航路径的精度,本研究将道路左右边界线的角平分线当成导航路径。图像坐标系中道路左右边界直线分别为 m_L 和 m_R ,假设 m_L 和 m_R 直线方程分别为

$$y_L = k_L x + b_L \tag{2}$$

$$y_R = k_R x + b_R \tag{3}$$

式中 y_L, y_R ——道路左边界和右边界对应点的纵

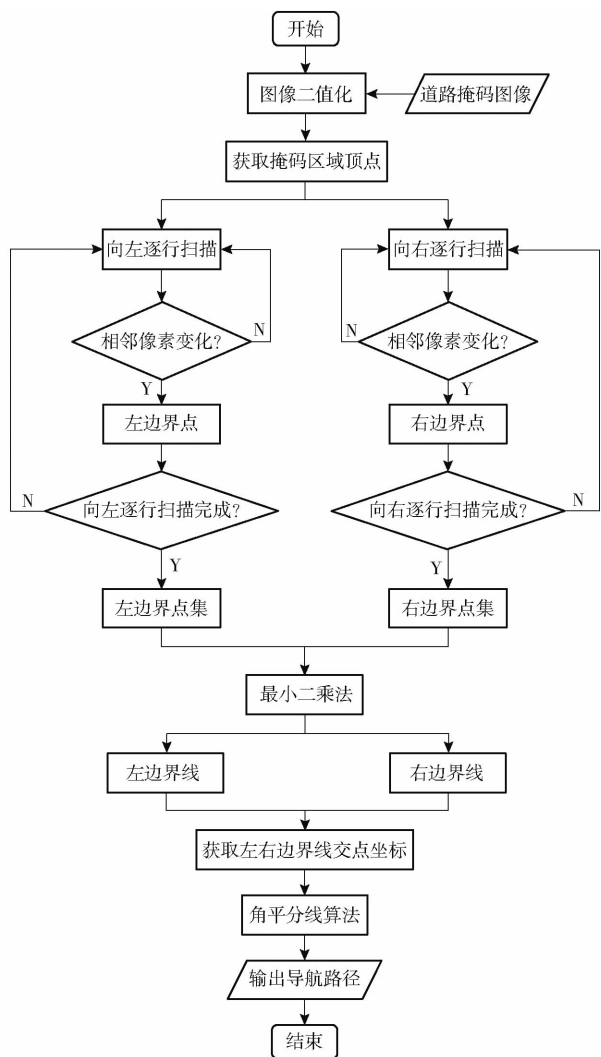


图4 导航路径拟合流程图

Fig. 4 Flowchart of navigation path fit

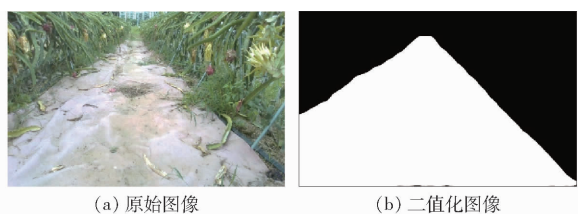


图5 图像分割结果

Fig. 5 Image segmentation results

坐标值

 k_L, k_R ——道路左、右边界直线方程斜率 b_L, b_R ——道路左、右边界直线方程截距

为了得到正确的导航线斜率,本文利用向量来求解。假设左右边界线的交点记为点 $O(x_0, y_0)$, 随机在左右边界线上分别取两个点, 记为点 $A(x_a, y_a)$ 和 $B(x_b, y_b)$, 则有

$$\mathbf{l}_{OA} = (x_a - x_0, y_a - y_0) \quad (4)$$

$$\mathbf{l}_{OB} = (x_b - x_0, y_b - y_0) \quad (5)$$

考虑将 $\mathbf{l}_{OA}$ 和 $\mathbf{l}_{OB}$ 归一化, 需要求出 2 个向量的模长, 根据模长公式则有

$$l_{OA} = \sqrt{(x_a - x_0)^2 + (y_a - y_0)^2} \quad (6)$$

$$l_{OB} = \sqrt{(x_b - x_0)^2 + (y_b - y_0)^2} \quad (7)$$

假设角平分线上存在 $\mathbf{l}_{OC} = (x_c, y_c)$, 使得 $\frac{\mathbf{l}_{OA}}{l_{OA}} +$

$\frac{\mathbf{l}_{OB}}{l_{OB}} = \mathbf{l}_{OC}$, 则有

$$x_c = \frac{x_a - x_0}{\sqrt{(x_a - x_0)^2 + (y_a - y_0)^2}} + \frac{x_b - x_0}{\sqrt{(x_b - x_0)^2 + (y_b - y_0)^2}} \quad (8)$$

$$y_c = \frac{y_a - y_0}{\sqrt{(x_a - x_0)^2 + (y_a - y_0)^2}} + \frac{y_b - y_0}{\sqrt{(x_b - x_0)^2 + (y_b - y_0)^2}} \quad (9)$$

最终可以求得角平分线斜率 $k = \frac{y_c}{x_c}$, 进而求出

截距 b 为

$$b = y_0 - k \quad (10)$$

从而得到导航基准线方程为

$$y = k(x - x_0) + y_0 \quad (11)$$

3.3 导航路径精度测试方法

本文通过人工方法来拟合果园实际导航路径。首先通过卷尺测量同一行左右两棵果株间的距离, 然后在中点位置做好标记, 接着利用白色尼龙绳将测得的中点位置依次连接。为了使得实际导航路径更加精确, 用钉子将尼龙绳加以固定。本文从像素误差、实际距离两个参数指标来评价拟合导航路径的精度。将相机支架放置在道路中间, 拍摄角度和采集图像时一致, 从道路尽头开始, 从拟合导航路径上以 70 像素为间距等距选取 10 个关键点, 利用 D435i 制造商提供的 SDK 来计算每幅图像上 10 个关键点与分别对应具有相同纵坐标的实际导航路径上 10 个点之间的像素误差和实际距离。像素误差的计算公式为

$$e_{\text{pix}} = |X_{\text{fit}} - X_{\text{real}}| \quad (12)$$

式中 e_{pix} ——像素误差

X_{fit} ——拟合导航路径上关键点的像素横坐标

X_{real} ——实际导航路径上关键点的像素横坐标

为了将图像中的像素误差转换为火龙果园中的实际距离, 根据针孔成像模型分别将拟合导航路径以及实际导航路径上关键点的像素坐标转换成相机坐标, 通过计算两点的欧氏距离得到实际距离。相机的内外参数以及畸变参数已标定, 坐标转换公式为

$$\begin{cases} u = f_x x_{\text{corrected}} + u_0 \\ v = f_y y_{\text{corrected}} + v_0 \end{cases} \quad (13)$$

其中

$$\begin{cases} x_{corrected} = x(1 + k_1r^2 + k_2r^4) + 2p_1xy + p_2(r^2 + 2x^2) \\ y_{corrected} = y(1 + k_1r^2 + k_2r^4) + 2p_2xy + p_1(r^2 + 2y^2) \end{cases}$$

(14)

$$(x,y) = \left(\frac{X}{Z}, \frac{Y}{Z}\right)$$

(15)

式中 (u,v) ——像素坐标
 f_x, f_y, u_0, v_0 ——相机内参
 $(x_{corrected}, y_{corrected})$ ——归一化平面纠正坐标
 k_1, k_2, p_1, p_2 ——相机畸变参数
 r ——归一化平面极径
 (X,Y,Z) ——相机坐标系坐标
 (x,y) ——归一化平面直角坐标

4 结果与分析

为了验证基于改进 DeepLabv3 + 的火龙果园导航路径识别算法以及基于角平分线算法拟合导航路径的准确性、稳定性和鲁棒性,分别进行了果园道路分割和导航路径提取实验。

4.1 果园道路分割

4.1.1 实验平台

本研究所使用的训练平台为配备 Windows 10 64 位操作系统的台式图形工作站,其中 CPU 为 Intel i9 10900X @ 3.75 GHz, GPU 为 NVIDIA GeForce GTX3090, RAM 为 128 GB, 软件环境为 PyTorch 1.7.1, CUDA 11.1, 以及 CUDNN 8.0.2。

4.1.2 训练参数设置

训练前,将数据集集中的 1 074 幅图像按照 8:1:1 随机划分为训练集、验证集和测试集,训练集有 860 幅图像,验证集和测试集分别有 107 幅图像。DeepLabv3 + 模型在批量大小为 16、初始学习率为 0.007 时,在 Pacal VOC 2012 和 Cityscapes 数据集上取得了良好的分割效果^[26]。本研究将其作为模型训练超参数,同时根据经验,优化方法选择随机梯度下降优化器(SGD),优化器动量设置为 0.9,权重衰减设置成 0.000 1,采用余弦退火的学习率下降方式。为了提高模型精度,使用 Pytorch 官方提供的 MobileNetV2 预训练权重。由于本文需求是分割道路和背景,故将分类个数设置成 2,并训练 300 个周期。

4.1.3 评价指标

本文采用平均像素准确率 (Mean pixel accuracy, MPA)、平均交并比 (Mean intersection over union, MIoU)、检测速率 (Frames per second, FPS)、模型内存占用量、参数量 (Params) 作为语义分割算法性能评价指标。假设总共有 $k + 1$ 个类,令 P_{ij} 表示第 i 类被预测为第 j 类的像素数量,则 P_{ii} 表示预

测准确的像素数量。

4.1.4 改进前后网络模型训练结果对比

改进前后的 DeepLabv3 + 网络的训练过程如图 6 所示。根据图 6 可知,随迭代次数增加,二者的训练集损失值呈减少趋势,而且随着迭代次数逐渐增加,MIoU 越来越高。当迭代次数到达 100 次时,二者的损失值和精度均趋于稳定。但是与传统 DeepLabv3 + 网络比较,改进后的 DeepLabv3 + 网络的收敛速度更快,精度也更高。

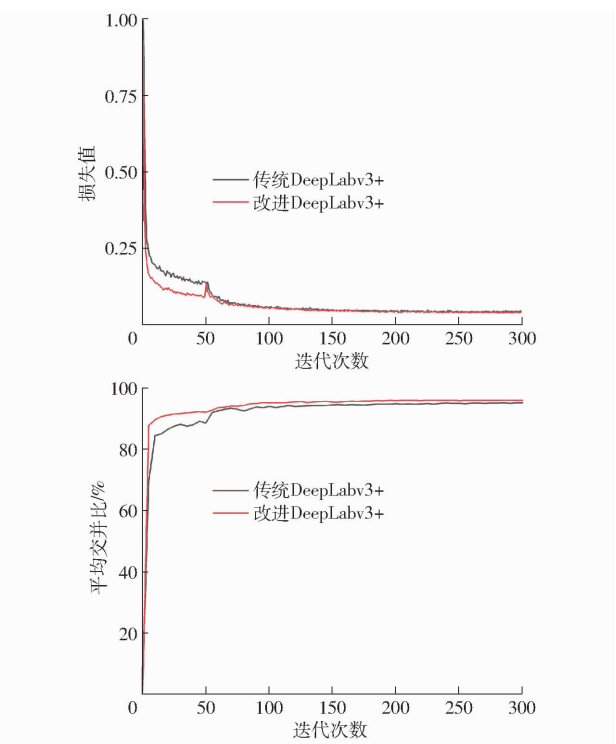


图 6 改进前后网络模型训练结果

Fig. 6 Model training results before and after improvement

4.1.5 模型有效性验证

为了验证在传统 DeepLabv3 + 网络中添加 CA 注意力机制模块的有效性,从测试集中选取了一幅道路条件较为复杂的图像进行验证,如图 7 所示。原始图像的右下方区域包含较多的杂草,同时由于受到光照的影响,路面上存在周边果株的投影区域。对比图中绿色椭圆区域分割结果发现,传统 DeepLabv3 + 网络模型丢失了部分道路边缘信息,而且未能分割出道路存在杂草覆盖的区域。而添加 CA 模块后的 DeepLabv3 + 网络模型对道路边缘信息提取更完整,能准确分割出覆盖杂草的道路,说明添加 CA 模块后提高了网络模型的特征提取能力。

为了验证将传统 DeepLabv3 + 网络中的主干特征提取网络由 Xception 更换为轻量级的 MobileNetV2 的有效性,以及验证添加注意力机制模块 CA、替换 ASPP 模块中的空洞卷积为深度可分离卷积对分割结果的影响,本文根据语义分割评价指

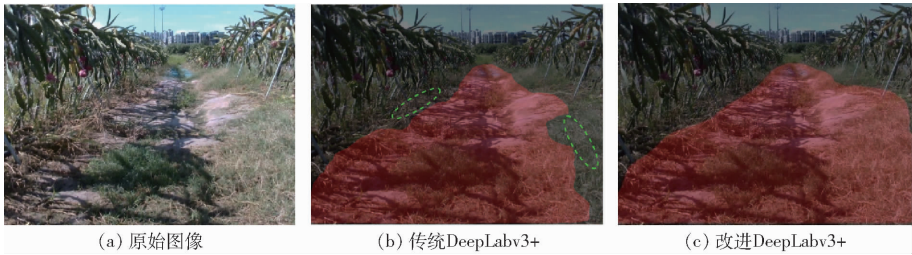


图 7 复杂道路图像的分割结果

Fig. 7 Segmentation results of complex road images

标,设置 4 组不同的改进方案进行消融实验,实验结果如表 1 所示。

表 1 消融实验结果
Tab. 1 Ablation test results

网络模型	主干网络	添加模块	MIoU/%	MPA/%	检测速率/(f·s ⁻¹)	参数量	模型内存占用量/MB
DeepLabv3 +	Xception	无	95. 01	97. 45	33. 19	5. 47 × 10 ⁷	209. 0
DeepLabv3 +	MobileNetV2	无	95. 48	97. 68	56. 74	5. 81 × 10 ⁶	22. 4
DeepLabv3 +	MobileNetV2	CA	95. 94	97. 94	56. 04	5. 83 × 10 ⁶	22. 5
DeepLabv3 +	MobileNetV2	DSC	95. 28	97. 60	59. 33	3. 86 × 10 ⁶	14. 9
DeepLabv3 +	MobileNetV2	CA + DSC	95. 80	97. 86	57. 89	3. 87 × 10 ⁶	15. 0

根据表 1 可得,将 DeepLabv3 + 的主干网络由 Xception 换成 MobileNetV2 后检测速率提升 23.55 f/s,而且极大减小了参数量以及模型内存占用量,参数量和模型内存占用量仅为原来的 1/9。其次,主干网络为 MobileNetV2 的基础上添加 CA 模块后,MIoU 以及 MPA 分别提升 0.46、0.26 个百分点,表明 CA 模块可以在一定程度上提高模型分割精度。另外,将 APSS 模块中的空洞卷积换成 DSC 时,相比原来的训练结果 MIoU 和 MPA 有略微下降,分别降低 0.20、0.08 个百分点,但是参数量和模型内存占用量分别减小 33.70%、33.50%,表明 DSC 可以有效降低模型复杂度。最后,本文提出的改进 DeepLabv3 + 网络模型的训练结果和原模型相比,无论是分割精度还是检测速率都有明显提升,更重要的是改进后的网络模型的参数量和模型内存占用量仅相当于原来的 1/14。因此改进 DeepLabv3 + 模型可以更好适用于果园道路的检测,而且也便于在嵌入式设备和移动设备上部署。图 8 为改进网络在不同环境条件下果园道路场景分割效果图。

从图 8 可以看出,白色地膜道路图像中的道路区域被有效分割,分割结果没有受到地面散落枝条的影响;黑色地膜道路图像中,路面不但覆盖杂乱无章的杂草,而且由于光照的影响,路面上还存在周边果株的投影区域,但分割结果并没有受到复杂路面环境的影响,道路分割完整,同时道路尽头的白色地膜道路没有被错误分割为背景,而是和黑色地膜区域一块识别为道路区域,说明模型具有良好的抗干扰性;在无地膜道路图像中,道路无明显边界,而且无地膜道路的杂草较为密集,同时沙土路面潮湿且

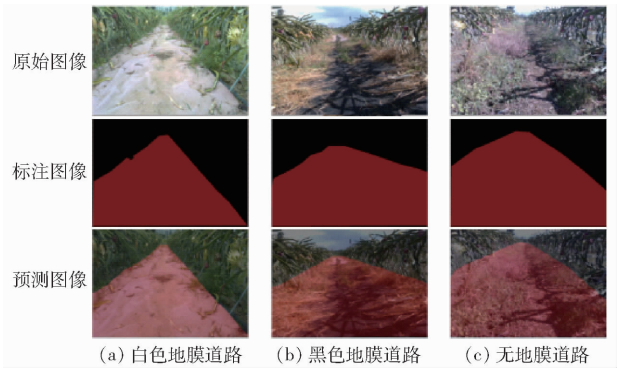


图 8 不同环境条件下果园道路场景分割结果

Fig. 8 Semantic segmentation results of orchard road scene under different environment conditions

不平整,加上光照的影响,这些复杂环境因素会对分割结果造成干扰,但根据模型的预测结果,道路边缘信息完整,道路的左右边界连续且对称分布,分割效果精细,说明模型具有较高的鲁棒性。

4.1.6 不同模型性能对比

为了进一步分析改进 DeepLabv3 + 模型的分割性能,选择分割精度较高的 Pspnet、U-net 与本文模型进行对比,以上网络模型均利用果园道路数据集进行训练,训练结果如表 2 所示。

由表 2 可以看出,本文模型和 Pspnet 相比,无论是在分割精度还是检测速率都优于后者,其中 MIoU 和 MPA 分别提高 0.49、0.25 个百分点,另外检测速率提高 11.7 f/s,参数量和模型内存占用量减小 91%。另外本文模型相较于 U-net,尽管在分割精度上略低于后者,但是在检测速率、参数量和模型内存占用量方面都显著优于 U-net,其中检测速率提

表 2 不同网络模型性能对比

Tab.2 Performance comparison of different network models

网络模型	主干网络	MIoU/%	MPA/%	检测速率/(f·s ⁻¹)	参数量	模型内存占用量/MB
Pspnet	Resnet50	95.31	97.61	46.19	4.67×10 ⁷	178.0
U-net	Resnet50	95.89	97.88	37.31	4.39×10 ⁷	167.0
本文模型	MobileNetV2	95.80	97.86	57.89	3.87×10 ⁶	15.0

升 20.58 f/s, 参数量和模型内存占用量分别减小 4.003×10^7 和 152.0 MB, 相比 U-net, 本文模型更加适合部署至嵌入式设备, 而且保证了模型具有较高的分割精度的同时也更加轻量化。

4.2 导航路径提取

选择白色地膜道路、黑色地膜道路、无地膜道路 3 种不同的道路环境来验证本文设计的导航路径提取算法精度。导航路径识别结果如图 9 所示(上图为原始图像, 下图为结果图像)。从图 9 中可以看出, 本文方法能够完整提取道路两侧的边缘信息, 而且通过最小二乘法拟合出道路边界线与两侧的果树行衔接紧凑, 角平分线算法拟合的导航路径能够有效减小不规则道路边缘对导航路径提取的干扰。

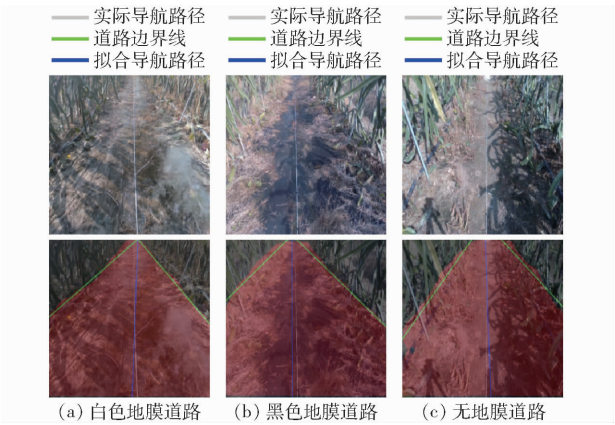


图 9 不同环境条件下果园导航路径识别结果
Fig.9 Semantic segmentation results of orchard road scene under different environment conditions

3 种不同道路条件下的导航路径精度测试结果如表 3 所示。在白色地膜道路条件下, 平均像素误差和平均距离误差分别为 17 像素和 7.42 cm。在黑色地膜道路条件下, 平均像素误差和平均距离误差分别为 32 像素和 10.55 cm。在无地膜道路条件下, 平均像素误差和平均距离误差分别为 17 像素和 4.76 cm。可得不同道路条件下的平均像素误差是 22 像素, 平均距离误差是 7.58 cm。已知所在火龙果园的道路宽度为 3 m, 平均距离误差占比为 2.53%。在当前火龙果园环境中本文方法具有较高

表 3 导航路径精度测试结果

Tab.3 Navigation path accuracy test results

关键点 序号	像素误差/像素			距离误差/cm		
	白色地	黑色地	无地膜	白色地	黑色地	无地膜
	膜道路	膜道路	道路	膜道路	膜道路	道路
1	32	22	15	21.29	14.63	9.98
2	26	23	6	17.30	15.30	3.99
3	22	25	1	11.25	12.78	0.51
4	18	28	8	7.49	11.65	3.33
5	15	30	11	6.24	12.48	4.58
6	10	33	14	3.06	10.10	4.28
7	4	36	22	0.81	7.25	4.43
8	9	38	25	1.81	7.65	5.03
9	11	40	33	1.85	6.73	5.55
10	18	41	35	3.08	6.90	5.89
平均值	17	32	17	7.42	10.55	4.76

的精度, 能够适应不同的火龙果园道路环境。

5 结论

(1) 提出了一种基于改进 DeepLabv3+ 的火龙果园道路识别方法, 通过加入 CA 模块提高了模型的特征提取能力; 另外将原模型中的主干特征提取网络更换成 MobileNetV2, 以及将 ASPP 模块中的空洞卷积更换成 DSC, 极大降低了模型内存占用量, 改进模型和原模型、Pspnet 以及 U-net 相比内存占用量分别减小 97.00%、91.57% 和 91.02%。并且检测速率和平均交并比分别提升至 57.89 f/s 和 95.80%, 说明本文模型更加适合部署至嵌入式设备上。

(2) 根据网络模型分割出的道路掩码, 提出一种基于道路边界拟合的角平分线提取导航路径算法, 该算法能准确提取出导航路径。

(3) 在火龙果园 3 种不同道路环境条件下, 进行了导航路径精度测试实验。实验结果表明, 平均像素误差为 22 像素, 平均距离误差为 7.58 cm。已知所在果园道路宽度约为 3 m, 平均距离误差占比为 2.53%。说明本文提出的导航路径拟合方法具有较好的适应性。因此本研究可以为火龙果园视觉导航提供有效参考。

参 考 文 献

[1] 律秀燕, 秦家辉, 马志伟, 等. 我国果园机械装备应用现状及发展对策[J]. 农业机械, 2021(12): 61-63.
LÜ Xiuyan, QIN Jiahui, MA Zhiwei, et al. Application status and development countermeasures of orchard machinery and

equipment in China[J]. *Agricultural Machinery*,2021(12):61 – 63. (in Chinese)

[2] 李雪峰,李涛,邱权,等. 果园移动机器人自主导航研究进展[J]. *中国农机化学报*, 2022,43(5):156 – 164.
LI Xuefeng, LI Tao, QIU Quan, et al. Research progress of autonomous navigation of orchard mobile robots [J]. *Chinese Journal of Agricultural Mechanization*,2022,43(5):156 – 164. (in Chinese)

[3] 周建军,周文彬,刘建东,等. 果园机器人自动导航技术研究进展[J]. *计算机与数字工程*,2019,47(3):571 – 576
ZHOU Jianjun, ZHOU Wenbin, LIU Jiandong, et al. Research progress of orchard robot automatic navigation technology[J]. *Computer and Digital Engineering*,2019,47(3):571 – 576. (in Chinese)

[4] 张漫,季宇寒,李世超,等. 农业机械导航技术研究进展[J]. *农业机械学报*, 2020,51(4):1 – 18.
ZHANG Man, JI Yuhan, LI Shichao, et al. Research progress of agricultural machinery navigation technology[J]. *Transactions of the Chinese Society for Agricultural Machinery*,2020,51(4):1 – 18. (in Chinese)

[5] 莫冬炎,杨尘宇,黄沛琛,等. 基于环境感知的果园机器人自主导航技术研究进展[J]. *机电工程技数术*, 2021, 50(9):145 – 150.
MO Dongyan, YANG Chenyu, HUANG Peichen, et al. Research progress of autonomous navigation technology of orchard robot based on environmental perception [J]. *Mechanical and Electrical Engineering Technology*, 2021, 50 (9) : 145 – 150. (in Chinese)

[6] OPIYO S, OKINDA C, ZHOU J, et al. Medial axis-based machine-vision system for orchard robot navigation[J]. *Computers and Electronics in Agriculture*, 2021,185:106153.

[7] CHEN J Q, HU Q, WU J H, et al. Navigation path extraction for greenhouse cucumber-picking robots using the prediction-point Hough transform[J]. *Computers and Electronics in Agriculture*, 2021,180:105911.

[8] CHEN X Y, WANG S A, ZHANG B Q, et al. Multi-feature fusion tree trunk detection and orchard mobile robot localization using camera/ultrasonic sensors[J]. *Computers and Electronics in Agriculture*, 2018,147:91 – 108.

[9] MA Y, ZHANG W Q, QURESHI W S, et al. Autonomous navigation for a wolfberry picking robot using visual cues and fuzzy control[J]. *Information Processing in Agriculture*, 2021,8(1):15 – 26.

[10] GAO G D, XIAO K, JIA Y C. A spraying path planning algorithm based on colour-depth fusion segmentation in peach orchards[J]. *Computers and Electronics in Agriculture*, 2020,173:105412.

[11] 谭晨佼,李铁林,王东飞,等. 农业机械自动导航技术研究进展[J]. *农机化研究*, 2020,42(5):7 – 14.
TAN Chenjiao, LI Yilin, WANG Dongfei, et al. Research progress of automatic navigation technology for agricultural machinery[J]. *Journal of Agricultural Mechanization Research*,2020,42(5):7 – 14. (in Chinese)

[12] 马驰,董子扬,陈志军,等. 基于根点替代的猕猴桃果园行间导航线生成方法研究[J]. *干旱地区农业研究*, 2021, 39(5):222 – 230.
MA Chi, DONG Ziyang, CHEN Zhijun, et al. Research on the generation method of interline navigation lines in kiwifruit orchards based on root point substitution[J]. *Agricultural Research in Arid Regions*,2021,39(5):222 – 230. (in Chinese)

[13] 王毅,刘波,熊龙烨,等. 基于深度学习的果园道路导航线生成算法研究[J]. *湖南农业大学学报(自然科学版)*, 2019,45(6):674 – 678.
WANG Yi, LIU Bo, XIONG Longye, et al. Research on orchard road navigation line generation algorithm based on deep learning[J]. *Journal of Hunan Agricultural University (Natural Science Edition)*,2019,45(6):674 – 678. (in Chinese)

[14] 贺正彦. 基于深度学习的智能除草机器人视觉导航与控制[D]. 哈尔滨:哈尔滨工业大学, 2021.
HE Zhengyan. Visual navigation and control of intelligent weeding robot based on deep learning[D]. Harbin: Harbin Institute of Technology,2021. (in Chinese)

[15] CHEN Y, LI G Y, ZHANG X Q, et al. Identifying field and road modes of agricultural machinery based on GNSS recordings: a graph convolutional neural network approach[J]. *Computers and Electronics in Agriculture*, 2022,198:107082.

[16] 王龙飞,严春满. 道路场景语义分割综述[J]. *激光与光电子学进展*, 2021,58(12):44 – 66.
WANG Longfei, YAN Chunman. A review of semantic segmentation of road scenes [J]. *Advances in Lasers and Optoelectronics*, 2021,58(12):44 – 66. (in Chinese)

[17] SONG Y, XU F Y, YAO Q, et al. Navigation algorithm based on semantic segmentation in wheat fields using an RGB – D camera[J/OL]. *Information Processing in Agriculture*. DOI: 10.1016/j.inpa.2022.05.002.

[18] KIM W S, LEE D H, KIM Y J, et al. Path detection for autonomous traveling in orchards using patch-based CNN[J]. *Computers and Electronics in Agriculture*, 2020,175:105620.

[19] YANG Z, OUYANG L, ZHANG Z G, et al. Visual navigation path extraction of orchard hard pavement based on scanning method and neural network[J]. *Computers and Electronics in Agriculture*, 2022,197:106964.

[20] 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. Orchard visual navigation path recognition method based on U – Net network [J]. *Transactions of the Chinese Society for Agricultural Machinery*,2021,52(1):30 – 39. (in Chinese)

[21] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017:1800 – 1807.

[22] MARK S, ANDREW H, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018:4510 – 4520.

[23] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021:13708 – 13717.

[24] MILLETARI F, NAVAB N, AHMADI S A. V – Net: fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV), 2016:565 – 571.

[25] HE Y, ZHANG X Y, ZHANG Z Q, et al. Automated detection of boundary line in paddy field using MobileV2 – UNet and RANSAC[J]. *Computers and Electronics in Agriculture*, 2022,194:106697.

[26] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Computer Vision-ECCV 2018, 2018:833 – 851.