

吴伟斌,唐婷,刘强,等.基于语义分割的山地果茶园道路识别技术研究[J].华中农业大学学报,2022,41(1):246-254.  
DOI:10.13300/j.cnki.hnlkxb.2022.01.025

# 基于语义分割的山地果茶园道路识别技术研究

吴伟斌<sup>1,2</sup>,唐婷<sup>1</sup>,刘强<sup>1</sup>,赵新<sup>1,2</sup>,韩重阳<sup>1</sup>,李杰<sup>1</sup>

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研究室/国家柑橘产业体系机械化研究室/

广东省山地果园机械创新工程技术研究中心,广州 510642

**摘要** 针对果、茶园规模不断扩张并逐渐向智能农业机械化发展的趋势以及常用道路语义分割数据集缺少果、茶园道路场景等问题,将语义分割技术应用到部分果、茶园道路中,以实现对果、茶园道路的像素级分割。以道路、人和车为分类对象,建立果、茶园道路场景图像数据集(包括 6 032 张图像),将数据集按照 9:1 比例随机划分为训练集(5 429 张图像)和测试集(603 张图像)。以 PSPNet (pyramid scene parsing network, 金字塔场景解析网络)分割模型为基础进行优化,构建 MS-PSPNet 语义分割模型;训练结果显示,MS-PSPNet 模型的 MIoU (mean intersection over union, 平均交并比)为 83.41%, FPS(frames per second, 每秒传输帧数)为 22.31。将 MS-PSPNet 模型应用在果、茶园不同道路条件和光照强度下进行现场试验,并进行准确度评估,结果显示,MS-PSPNet 模型类别 MPA (mean pixel accuracy, 像素准确率)均超过 92%, MIoU 在除非硬化道路条件情况均超过 91%, 表明 MS-PSPNet 模型在果、茶园道路识别中具有较好的有效性和适用性。

**关键词** 山地果、茶园; 果园智能化; 道路识别; 语义分割; 超像素分割; 精准农业; 自主导航

中图分类号 TP 391 文献标识码 A 文章编号 1000-2421(2022)01-0246-09

我国果、茶园种植面积大,并且主要分布在丘陵山区,受地形条件和种植模式影响,丘陵地区果、茶园普遍存在着农业机械化程度低、管理模式落后以及农村劳动力老龄化严重等亟待解决的问题<sup>[1-2]</sup>。现代智能农业机械装备的自动化和智能化是山区果、茶园未来发展的必然趋势,对提升农产品及农用物资在田间运输工作效率、降低劳动成本和促进山区农业经济的发展具有重要意义<sup>[3-4]</sup>。

农业机械智能化发展的关键是自动导航技术,该技术的关键在于通过准确快速地对图像进行处理进而对导航路径进行可靠、稳定地识别,从而引导机具行进,因此,根据相应的图像处理算法可实现导航路径规划<sup>[5-7]</sup>。由于丘陵山区果、茶园道路场景复杂、迂回多变且无明显车道边界,使用神经网络分类器、像素块特征学习等方法识别道路与非道路的算法不能完全地适用这种道路环境状况<sup>[8]</sup>。图像语义分割在遵循人类视觉特点下基于对象的图像分割,可以提供简单且可靠的图像特征信息<sup>[9-11]</sup>。韩振浩

等<sup>[5]</sup>建立了一种基于 U-Net 语义分割算法的果园视觉导航路径识别方法,结果表明,临界阈值为 0.4 时,语义分割模型在弱光、普通光以及强光照条件下的分割交并比分别为 89.52%、86.45%、86.16%,能够平稳实现果园道路像素级分割。李云伍等<sup>[12]</sup>构建了基于空洞卷积神经网络的田间道路场景图像语义分割模型,该模型的道路识别准确率达到 86%,对道路场景的阴影干扰有良好的适应性。基于卷积网络和完全连接的 CRF (conditional random fields, 条件随机场)组合的语义分割模型可以实现精确的目标分割,也是解决自动驾驶和精准农业等领域目标场景识别的重要技术手段<sup>[13-14]</sup>。

为实现对果、茶园道路的识别,本研究提出一种基于 MS-PSPNet 语义分割算法的山地果、茶园道路识别方法,通过快速、准确识别出果园中人、车、物等,实现农业机械在果园中的自主导航,对于山地果、茶园的监控管理和农业装备的智能化发展具有重要作用。

收稿日期: 2021-09-12

基金项目: 2021 年广东省农业科研项目和农业技术推广项目 (2020-440000-02100200-8418); 2021 年广东省现代农业产业技术体系创新团队建设项目 (2021NO74-CJXG); 广东省科技计划重点研发项目 (2019B020223001)

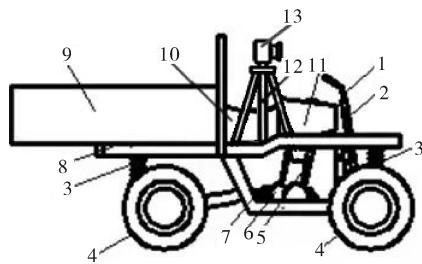
吴伟斌, E-mail: wuweibin@scau.edu.cn

通信作者: 赵新, E-mail: xinxin@scau.edu.cn

# 1 材料与方法

## 1.1 图像数据集采集

1) 采集环境。合适的采集高度和角度对采集到的图像能否反映果茶园运输机械的真实状态至关重要。综合考虑采集图像的空间分辨率需求和避障需求确定了摄像头的安装位置, 同时选择笔者所在实验室前期研制的山地果、茶园运输车搭载相机进行视频拍摄, 如图 1 所示。图像数据采集过程中驾驶员驾驶运输车以 1.2 m/s 的平均速度行驶, 通过相机采集道路场景视频数据后进行存储, 图 2 为采集的部分果、茶园道路图像。



1. 手扶转向车把 Turn the handlebars with your hand; 2. 转向轴 Steering shaft; 3. 悬架组件 Suspension components; 4. 特殊山地轮胎 Special mountain tyre; 5. 汽油发动机 Gasoline engine; 6. 循环挡变速器和独立倒挡器联合布置机构 Combined arrangement mechanism of circulating transmission and independent reverse gear; 7. 传动轴 The shaft; 8. 车架 The frame; 9. 车厢 The carriage; 10. 驾驶座 In the driver's seat; 11. 油箱 Oil compartments; 12. 三脚架 The tripod; 13. 摄像机 The video camera.

图 1 山地果、茶园道路图像采集系统示意图

Fig.1 Schematic diagram of mountain fruit and tea garden road image acquisition system



A: 果园硬化道路 Orchard hardened road; B: 果园非硬化道路 Orchard non-hardened road; C: 茶园硬化道路 Hardened road in tea garden; D: 茶园非硬化道路 Non-hardened road in tea garden.

图 2 果、茶园硬化与非硬化道路现场图

Fig.2 Scene diagram of hardened and non-hardened roads in fruit and tea garden

通过剪辑视频, 按每 10 帧选取 1 幅图像, 共得到图片 6 032 张。其中, 果园道路环境数据 3 886 张, 其中硬化道路图像 2 176 张, 占图像总数的 36%, 非硬化道路图像 1 710 张, 占图像总数的 28.3%; 茶园道路图像 2 146 张, 其中硬化道路图像 1 179 张, 占图像总数 19.6%, 非硬化道路 967 张, 占图像总数 16.1%。

2) 采集类别。采集的类别根据对丘陵山区果、茶园道路环境的特点及出现的频率分为 4 类, 分别是背景、道路、行人和车, 其对应类型图片数量分别为 6 032、6 032、4 562、1 796 张, 平均占比分别为 100%、100%、75.6%、29.8%。

## 1.2 果、茶园道路数据集处理

由于相机直接获取的图片分辨率高, 在训练时占用显存过高, 影响训练速度; 图像没有语义标签, 需要手工标注, 才可用于神经网络的训练。将采集到的果、茶园道路图像导入到 Photoshop 中, 并将图像尺寸大小统一调整到 960 像素  $\times$  540 像素, 再使用像素级标注工具 LabelMe 进行语义标注, 实现对图像语义信息的赋予, 最后将其保存为 json 文件, 作为模型训练时读取的信息文件。将标注好的数据按照 9:1 的比例随机划分为训练集和测试集, 其中训练集 5 429 张, 测试集 603 张。训练集用来训练网络模型, 使模型能够反映现实, 进而预测未来或其他未知的信息, 测试集则用来评估模型的预测性能。部分人工标注图和原图像对比如图 3 所示。



A: 原图 1 Original image 1; B: 标注图 1 Annotation Fig.1; C: 原图 2 Original image 2; D: 标注图 2 Annotation Fig.2.

图 3 原图像和人工标注图标注前后对比

Fig.3 Image comparison between manual annotation map and original image before and after annotation

# 2 结果与分析

## 2.1 PSPNet 的语义分割模型训练与测试

本研究选定 PSPNet (pyramid scene parsing

network,金字塔场景解析网络)模型作为参考模型对果、茶园道路进行优化适配,从而实现果、茶园道路的识别。将训练集、测试集和标注好的文件放入指定位置,并把图像名称和标注数据写入到 txt 文本格式的训练文件和测试文件中,用于模型在训练(图 4)及测试过程中的读取(图 5)。

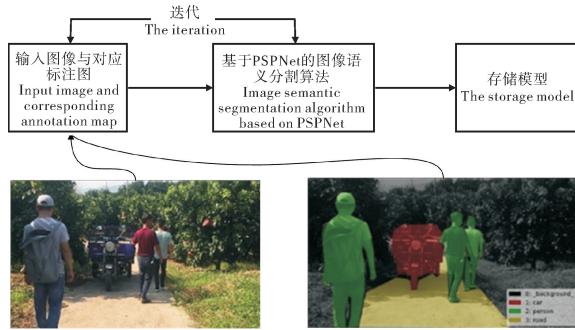


图 4 果、茶园道路场景模型的训练过程

Fig.4 Training process of road scene model of fruit and tea garden

训练开始后,损失函数的 loss 值随着迭代次数的增多而逐渐减少,最后 loss 值趋于稳定,表示训练完成。用准备的测试数据集完成测试过程。

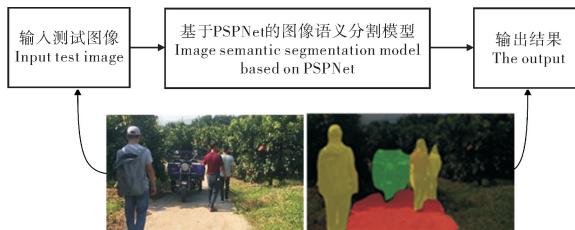


图 5 果、茶园道路场景模型的测试过程

Fig.5 Test process of road scene model of fruit and tea garden

利用 MPA (mean pixel accuracy, 像素准确率) 值和 MIoU (mean intersection over union, 平均交并比) 值对训练结果做出评估以反映图像语义分割的准确度, 利用 FPS (frames per second, 每秒传输帧数) 对图像实时语义分割的速度做出评估。PSPNet 模型在试验时, 测得 FPS 值、MPA 值和 MIoU 值评估结果如表 1 所示。从表 1 可以看出, PSPNet 模型在果、茶园道路数据集取得了平均 MPA 值 91.07% 和平均 MIoU 值 78.49% 的效果, 但是其 FPS 值为 15.04, 表明该模型虽然能做到对果、茶园道路的分割, 但画面卡顿较为严重。此外由于 PSPNet 语义分割模型中的主干特征提取网络 Resnet 模块过于复杂, 导致其在真实的应用场景下难以被应用。主要是因为这种网络模型比较占用内存, 对硬件要求较高; 某些真实场景要求较低的延迟和较快的响应速度, 比如果、茶园道路识别环境。因此, 小而高效的 PSPNet 模型对于目标检测算法实际应用非常重要。

## 2.2 M-PSPNet 模型测试结果

由于选用 MobileNet-V3 卷积神经网络作为特征提取网络运行速度快, 所以采用其替换 PSPNet 原有的主干网络 Resnet, 可以获得更大的 FPS 值, 在语义分割算法的效率和速度上取得突破, 使其更加适用于果、茶园道路的实时识别<sup>[15-16]</sup>。将基于 PSPNet 与 MobileNet-V3 改进的果、茶园道路语义分割模型简称为 M-PSPNet, 其模型流程图如图 6 所示。从图 6 可知, 在果、茶园道路语义分割模型的结构流程图中, M-PSPNet 模型首先将制作的果、茶园道路数据集导入到 MobileNet-V3 主干特征提取网络中进行初步的特征提取, 然后将得到的特征层

表 1 3 种语义分割模型的分割结果

Table 1 Segmentation results of three semantic segmentation models

模型类型	Model type	项目	Item	背景	Background	路	Road	车	Car	人	Human	平均值	Mean
PSPNet		MPA/%		98.50		91.47		87.92		86.39		91.07	
		MIoU/%		96.23		87.36		56.37		74.01		78.49	
		FPS					15.04						
M-PSPNet		MPA/%		97.40		89.54		86.51		85.34		89.69	
		MIoU/%		95.28		85.69		54.72		72.69		77.09	
		FPS					29.62						
MS-PSPNet		MPA/%		98.70		94.51		92.68		92.69		94.64	
		MIoU/%		97.28		90.96		66.90		78.51		83.41	
		FPS					22.31						

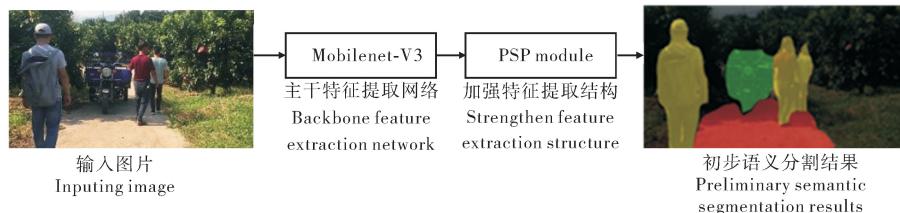


图 6 M-PSPNet 模型流程图

Fig.6 M-PSPNet model flow chart

数据导入到原 PSPNet 模型的 PSP 模块中进行池化处理,提高果、茶园道路语义分割模型获取图片全局信息的能力,最后将生成的果、茶园道路语义分割模型导出进行测试集的测试。通过以上这些流程,完成对果、茶园道路的阶段识别。

对 M-PSPNet 语义分割模型进行训练,其果、

茶园道路的语义分割图像与 PSPNet 模型分割图像对比结果如图 7 所示。两者均能很好地识别出道路及路上物体的基本轮廓,但 PSPNet 的语义分割结果较贴近于标注图,M-PSPNet 的分割结果相对粗糙,并且 2 个模型对于物体边缘分割效果均不理想,锯齿状较为严重。



A:原图 Original image; B:标注图 Label map; C:PSPNet 识别图 M-PSPNet identification map; D:M-PSPNet 识别图 M-PSPNet identification map.

图 7 识别效果对比

Fig.7 Recognition effect comparison

M-PSPNet 模型具体评估指标结果如表 1 所示。从表 1 可知,2 种语义分割模型对于果、茶园道路均有较不错的分割结果,但 2 种模型对行人和车的识别效果稍差。数据显示,模型 MPA 值要略高于 MIoU 的值,这主要是因为 MIoU 值对语义分割结果进行评估时会考虑全局类别语义的识别精确度,导致 MIoU 值略小。其中,PSPNet 模型在分割准确性方面,MPA 平均值为 91.07%,MIoU 平均值为 78.49%,均要优于 M-PSPNet 模型的 89.69% 和 77.09%,可见在分割准确性方面 PSPNet 模型原

本的 Resnet 模块要略优于 MobileNet-V3 模块。实时分割速度方面,M-PSPNet 的 FPS 值达到了 29.62,约是 PSPNet 的 FPS 值(15.04)的 2 倍。分割画面可以实时显示、流畅不卡顿,相较于 PSPNet 模型,M-PSPNet 更加适合应用于果、茶园道路的实时识别环境。

### 2.3 MS-PSPNet 模型测试结果

为解决在物体边缘部分的语义分割效果不够理想的问题,进一步恢复果、茶园道路图像的边缘等图像细节信息,将图像超像素分割算法引入 M-PSP-

Net 模型中,提出通过融合 M-PSPNet 模型分割出的果、茶园道路图像高层语义特征和果、茶园道路超像素信息来进行模型优化<sup>[17-18]</sup>。

在超像素优化算法中,优化方法具体分为 3 个步骤:首先输入图像,使用 SLIC 对其进行超像素

分割,获得果、茶园道路的超像素分割图;其次,获取由 M-PSPNet 模型输出的初步语义分割图;最后将 2 张图融合在一起,具体流程如图 8 所示。本研究将这种融合 SLIC 的 M-PSPNet 模型简称为 MS-PSPNet。

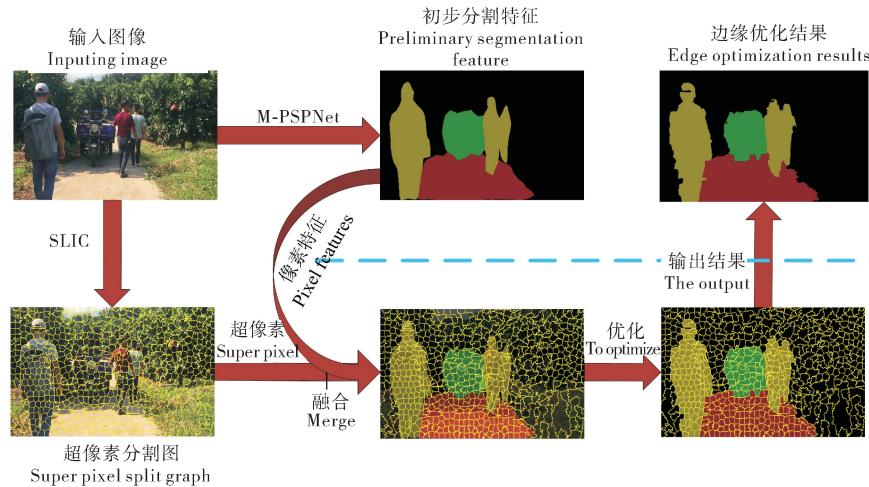


图 8 超像素优化算法流程图

Fig.8 Super pixel optimization algorithm flow chart

MS-PSPNet 模型在超像素优化过程存在 2 种情况,第 1 种是超像素中是否包含图像的边缘,又可以分为有边缘和没有边缘两种情况;第 2 种是超像素内部如果包含图像的边缘,则可以根据超像素分割区域内的像素是否具有相同的语义标签,共存在 4 种情况:情况 A 为超像素内部不包含图像边缘,像素点为相同的果、茶园道路的语义标签;情况 B 为超像素内部不包含图像边缘,像素点包含果、茶园道路的多种语义标签;情况 C 为超像素内部包含图像边缘,像素点为相同的果、茶园道路的语义标签;情况 D 为超像素内包含图像边缘,像素点包含果、茶园道路的多种语义标签。

MS-PSPNet 具体算法为:①通过 M-PSPNet 输入图像  $I$  和初步的语义分割结果  $L$ 。②令  $C = \{C_1, C_2, \dots, C_N\}$ ,是指需要分配给每个像素的所有  $N$  种语义分类。③使用 SLIC 算法将输入图像  $I$  分割为  $n$  个超像素,因此有  $S = \{S_1, S_2, \dots, S_n\}$ ,并且超像素  $S_i$  中有  $m_i$  个像素。④输入图像  $L$ ,同时输入 2 个变量  $i$  和  $j$ ,其中,  $i$  的范围是  $(1 : n)$ ,  $j$  的范围是  $(1 : m)$ ,首先查找  $P_j$  属于  $C$  中的哪种语义分类,其中  $P_j$  表示  $m_i$  中的第  $j$  个像素。例如,如果  $P_j$  表示  $C_K$ ,则令  $M_K = M_K + 1$ ,其中  $M_K$  表示引用  $C_K$  的像

素。进而找到语义分类  $C_{max}$ ,该语义分类  $C_{max}$  在超像素  $S_j$  中有对应的像素,并将  $C_{max}$  分配给该超像素。⑤得到最终的图像  $L'$ ,它是指超像素细化后的语义分割结果,并输出图像  $L'$ 。

对于情况 A,选择使用原有的语义标记;对于情况 B,使用超像素内占比最大的语义标签;对于情况 C,仍然选择使用原有的语义标签;对于情况 D,如果发现在该超像素分割区间内,某个果、茶园道路语义标签被分配到了 80% 及以上的像素,那么就将这个果、茶园道路的语义标签作为整个超像素分割区间内的语义标签,若不存在则放弃对该区域的优化。

SLIC 算法的  $K$  值是人为设定的,  $K$  值设定过小,则会导致无法准确分割图像;  $K$  值设定较大,则会造成分割过程中时间和内存的浪费。因此,为了确定不同的超像素总数  $K$  对分割准确率的影响,使用 MS-PSPNet 模型根据果、茶园道路数据集分辨率的实际情况进行试验(图 9)。由图 9 可知,当  $K$  值小于 1 000 时,图像的分割准确率会随着  $K$  值增大而升高;当  $K$  值大于 1 000 时,虽然分割准确率略有起伏,但平稳的趋势没有改变,而且会浪费较多的处理时间和内存。综上所述,选择  $K = 1 000$ ,超像素分割结果如图 9 所示。对 MS-PSPNet 语义分

割模型进行训练,并与 PSPNet 和 M-PSPNet 模型的语义分割效果进行对比,分析 MS-PSPNet 模型的有效性(图 10)。

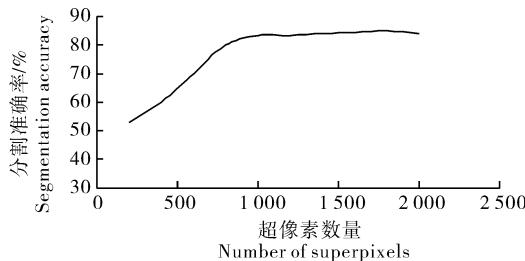
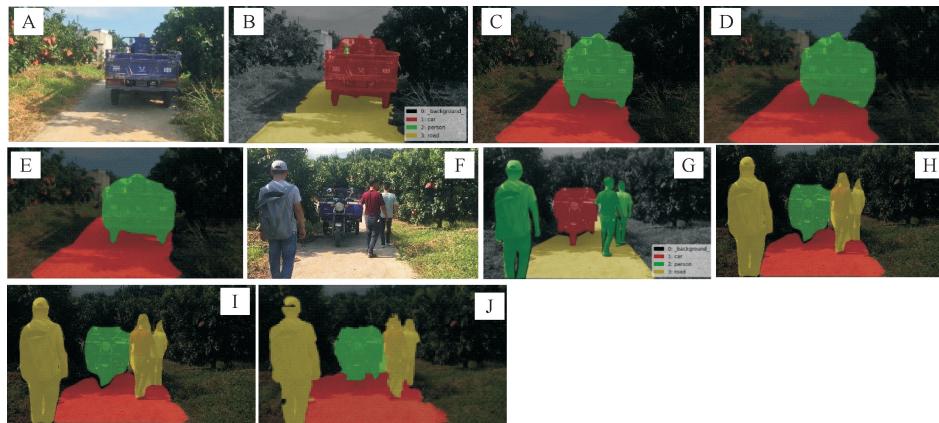


图 9 超像素数量对分割准确率的影响

Fig.9 Effect of the number of super pixels on the accuracy of segmentation

为更好地模拟果、茶园道路环境,分别从测试集中选取了上坡和下坡 2 种道路进行展示,且 2 张道路上均有障碍物(图 10)。从 3 种模型的测试结果来看,虽然 3 种模型均识别出了图片中的语义类别,但是 MS-PSPNet 模型在一些细节处的分割效果要明显优异于其他 2 种模型,具体表现在下坡图中位于最左侧的行人的头部, PSPNet 和 M-PSPNet 均有部分缺失,未能准确地识别出该语义标签,而 MS-PSPNet 实现了对该行人较好地识别效果。同样在道路的识别方面,道路两边的杂草由于其体积较小,均被前两种算法所忽略,而 MS-PSPNet 也能对其进行较好地识别。因此,基于 SLIC 超像素优化算法的 M-PSPNet 模型的确提高了图像边缘的分割效果(表 1)。



A: 原图 1 Original image 1; B: 标注图 1 Label image 1; C: PSPNet 模型识别图 1 PSPNet model recognition Fig.1; D: M-PSPNet 识别图 1 M-PSPNet recognition diagram 1; E: MS-PSPNet 识别图 1 MS-PSPNet recognition Fig.1; F: 原图 2 Original image 2; G: 标注图 2 Label image 2; H: PSPNet 模型识别图 2 PSPNet model recognition Fig.2; I: M-PSPNet 识别图 2 M-PSPNet recognition diagram 2; J: MS-PSPNet 识别图 2 MS-PSPNet recognition Fig.2.

图 10 不同模型的语义分割情况

Fig.10 Semantic segmentation under different models

在融合了 SLIC 以后,MS-PSPNet 的分割精确度超过了 M-PSPNet,也略超 PSPNet。在 MPA 评估指标的平均值达到了 94.64%,在 MIoU 评估指标的平均值达到了 83.41%,均要高于 PSPNet 的 91.07% 和 78.49%,取得了较为不错的分割效果和精确度。在分割速度方面,MS-PSPNet 的 FPS 值达到了 22.31,优于 PSPNet 的 15.04,略差于 M-PSPNet 的 29.62。由于在果、茶园环境中运输车的实际行驶速度不高,因此可牺牲部分分割速度来提高其分割精确度,实际使用中更加适应于果、茶园道路的实时分割。

综上所述,果、茶园道路语义分割模型 MS-PSPNet 可以有效地对果、茶园道路进行实时识别。

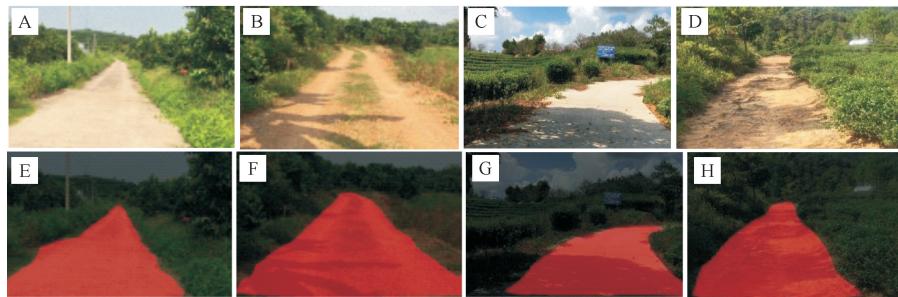
## 2.4 典型道路条件试验

选取果园硬化道路、果园非硬化道路、茶园硬化道路和茶园非硬化道路进行测试来检验 MS-PSPNet 模型在该道路环境下的适用性(图 11)。由图 11 可见,MS-PSPNet 模型对 4 种道路均进行了较好地分割,结果比较理想。但是在非硬化道路条件下的分割,相较硬化道路条件下的分割效果稍差,其主要原因是道路边界上存在不同类型的杂草。非硬化道路两边的杂草形态一般较小,而在硬化道路两边的杂草多较粗壮,导致模型在分割不同路面时出现分割效果差异。

MS-PSPNet 在硬化道路下的 MPA 值为 96.37%,MIoU 值为 93.59%;非硬化道路下的 MPA 值为 92.65%,非硬化道路条件下的 MIoU 值

达 88.33%，基本可以有效地识别不同路面，具有较强的适用性。在典型环境条件下，根据果、茶园道路的实际情况，选取了光照过强、光照适当和光照不足 3 种情况下的图像进行试验(图 12)。由图 12 可知，在 3 种光照环境下，MS-PSPNet 模型对道路均取得了较为不错的分割结果。实质是模型内核之一深度

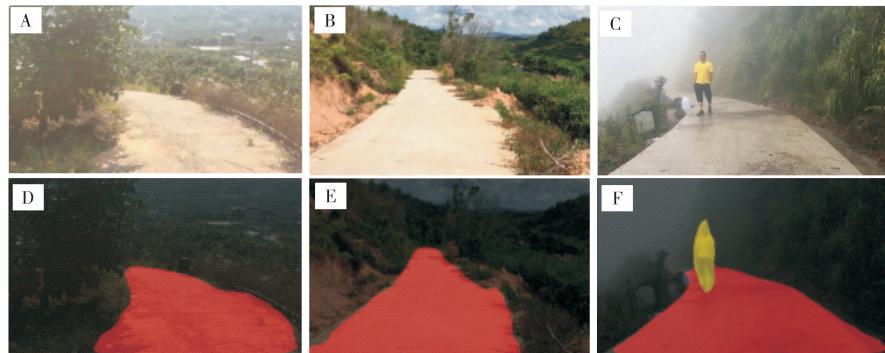
学习神经网络对几何变换、变形和光照变化具有高度不变性，减少了环境光照影响。由于在光照适当的条件下，果、茶园道路上的阴影少，获取的果、茶园道路图像更为清晰，成像质量较高，不同对象类别的颜色差异更为明显，所以光照适当语义分割结果略好。



A: 果园硬化道路 Orchard hardened road; B: 果园非硬化道路 Non-hardened road in orchard; C: 茶园硬化道路 Tea garden hardened road; D: 茶园非硬化道路 Non-hardened road in tea garden; E: 果园硬化道路分割结果 Orchard hardened road segmentation results; F: 果园非硬化道路分割结果 Segmentation result of non-hardened road in orchard; G: 茶园硬化道路分割结果 Tea garden hardened road segmentation results; H: 茶园非硬化道路分割结果 Segmentation result of non-hardened road in tea garden.

图 11 不同道路条件下的语义分割情况

Fig.11 Semantic segmentation under different road conditions



A: 光照过强道路 Road with too much light; B: 光照适当道路 Light the proper road; C: 光照不足道路 Low-lit road; D: 光照过强道路分割结果 Road segmentation results with too much light; E: 光照适当道路分割结果 Road segmentation results with proper lighting; F: 光照不足道路分割结果 Insufficient lighting road segmentation results.

图 12 不同光照下的语义分割情况

Fig.12 Semantic segmentation under different lighting conditions

表 2 为不同光照强度下的分割结果，该结果表明在不同的光照条件下，光照适当的环境中 MPA 值达到了 95.65%，比光照过强的环境高出了 0.6 个百分点，比光照较弱的环境高出了 0.9 个百分点。光照适当的环境中 MIoU 值达到了 92.33%，比光线过强的环境高出了 0.4 个百分点，比光线不足的环境高出了 1.08 个百分点。实质差别不大。光照过强与不足 2 种情况均是果、茶园道路环境中的两种极端条件，MS-PSPNet 模型仍能对道路做出较好的识别效果，表明建立的果、茶园道路识别模型 MS-PSPNet 在不同的光照强度下仍具有较好的适

用性。

表 2 不同光照强度下的分割结果

Table 2 Different light intensity segmentation results %

项目 Item	MPA	MIoU
光照过强 Too much light	95.05	91.93
光照适当 Proper lighting	95.65	92.33
光照不足 Insufficient light	94.75	91.25

### 3 讨 论

本研究以果、茶园道路为研究对象，采用图像语义分割方法，根据果、茶园道路场景的实际情况，对果、茶园道路进行了语义分割及优化，建立了有效和

可行的果、茶园道路语义分割优化模型 MS-PSP-Net,并开展了 MS-PSPNet 模型在典型果、茶园道路上的识别试验。以语义分割模型 PSPNet 为基础,使用神经网络 Mobilenet-V3 取代其中主干特征提取网络 Resnet 模块,构造 M-PSPNet;在 M-PSP-Net 的基础上融合超像素分割算法 SLIC,构建 MS-PSPNet。MS-PSPNet 模型试验数据分析显示:硬化路面的 MIoU 值为 93.59%,非硬化路面的 MIoU 值为 88.33%;光照强度过大的 MIoU 值为 91.93%,光照强度适宜的 MIoU 值为 92.33%,光照强度不足的 MIoU 值为 91.25%。

本研究试验结果表明,MS-PSPNet 模型在果、茶园道路识别中具有较好的识别性能,其在多个数据集上的分割精确度都超越了 FCN、DeepLab-V2、DPN、CRF-RNN 等模型,接近 DeepLab-V3<sup>[16-17]</sup>,可以应用在山地果、茶园智能农业装备导航领域。同时也存在以下不足:因为采集时间不是收获的季节,果、茶园道路上的情况较为单调,所以本研究中定义的果、茶园道路的语义类别信息相对较少,仅为 3 类(不包括空类)。后续研究可以采集更多的果、茶园道路图片来拓展标注数量及语义类别,以提高精确度及适用性。MS-PSPNet 语义分割模型虽然在果、茶园道路识别的准确度和速度上取得了一定的平衡,但从分割结果来看,仍存在提升空间。此外,后期可添加激光雷达或双目摄像头等手段获取深度信息,为农业自动化奠定更好的基础。

## 参考文献 References

- [1] 郑永军,江世界,陈炳太,等.丘陵山区果园机械化技术与装备研究进展[J].农业机械学报,2020,51(11):1-20.ZHENG Y J,JIANG S J,CHEN B T,et al.Review on technology and equipment of mechanization in hilly orchard[J].Transactions of the CSAM,2020,51(11):1-20(in Chinese with English abstract).
- [2] 鲍秀兰,孙培林,胡成,等.一种并联缆绳牵引式山地果园运输机的建模与分析[J].华中农业大学学报,2020,39(6):121-127.BAO X L,SUN P L,HU C,et al.Modeling and analysis of a parallel cable-towed transporter for hilly orchard[J].Journal of Huazhong Agricultural University,2020,39(6):121-127(in Chinese with English abstract).
- [3] 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 M,JI Y H,LI S C,et al.Research progress of agricultural machinery navigation technology[J].Transactions of the CSAM,2020,51(4):1-18(in Chinese with English abstract).
- [4] 邹宝玲,刘佛良,张震邦,等.山地果园机械化:发展瓶颈与国外经验借鉴[J].农机化研究,2019,41(9):254-260.ZOU B L,LIU F L,ZHANG Z B,et al.Mechanization of mountain orchards:development bottleneck and foreign experiences[J].Journal of agricultural mechanization research,2019,41(9):254-260(in Chinese with English abstract).
- [5] 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 Z H,LI J,YUAN Y W,et al.Path recognition of orchard visual navigation based on U-Net[J].Transactions of the CSAM,2021,52(1):30-39(in Chinese with English abstract).
- [6] 郭成洋,范雨杭,张硕,等.果园车辆自动导航技术研究进展[J].东北农业大学学报,2019,50(8):87-96.GUO C Y,FAN Y H,ZHANG S,et al.Progress on vehicle automatic navigation in orchard[J].Journal of Northeast Agricultural University,2019,50(8):87-96(in Chinese with English abstract).
- [7] 魏建胜,潘树国,田光兆,等.农业车辆双目视觉障碍物感知系统设计与试验[J].农业工程学报,2021,37(9):55-63.WEI J S,PAN S G,DIAN G Z,et al.Design and experiments of the binocular visual obstacle perception system for agricultural vehicles[J].Transactions of the CSAE,2021,37(9):55-63(in Chinese with English abstract).
- [8] 龚章鹏,王国业,于是.基于体素网络的道路场景多类目标识别算法[J].汽车工程,2021,43(4):469-477.GONG Z P,WANG G Y,YU S.The algorithm of multi-category object recognition in road scene based on voxel network[J].Automotive engineering,2021,43(4):469-477(in Chinese with English abstract).
- [9] 邝辉宇,吴俊君.基于深度学习的图像语义分割技术研究综述[J].计算机工程与应用,2019,55(19):12-21,42.KUANG H Y,WU J J.Survey of image semantic segmentation based on deep learning[J].Computer engineering and applications,2019,55(19):12-21,42(in Chinese with English abstract).
- [10] 冯成浩.基于深度学习的智能车辆环境语义分割算法研究[D].长春:吉林大学,2019.FENG C H.Research on outdoor scene semantic segmentation algorithm based on deep learning for intelligent vehicle[D].Changchun:Jilin University,2019(in Chinese with English abstract).
- [11] 王嫣然,陈清亮,吴俊君.面向复杂环境的图像语义分割方法综述[J].计算机科学,2019,46(9):36-46.WANG Y R,CHEN Q L,WU J J.Research on image semantic segmentation for complex environments[J].Computer science,2019,46(9):36-46(in Chinese with English abstract).
- [12] 李云伍,徐俊杰,刘得雄,等.基于改进空洞卷积神经网络的丘陵山区田间道路场景识别[J].农业工程学报,2019,35(7):150-159.LI Y W,XU J J,LIU D X,et al.Field road scene recognition in hilly regions based on improved dilated convolutional networks[J].Transactions of the CSAE,2019,35(7):150-159(in Chinese with English abstract).
- [13] PENG C,ZHANG X Y,YU G,et al.Large kernel matters-improve semantic segmentation by global convolutional network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).July 21-26,2017.Honolulu,HI:IEEE,2017:1743-1751.
- [14] FU J,LIU J,WANG Y H,et al.Stacked deconvolutional network for semantic segmentation[J/OL].IEEE transactions on

- image processing, 2019, 1[2021-19-12].<http://ieeexplore.ieee.org/document/8626414dol;10.1109/TIP.2019.2895460>.
- [15] 王康,梁秀英,周风燃,等.基于磁导引的履带式小车作物图像自动采集系统设计与试验[J].华中农业大学学报,2020,39(2):141-149. WANG K, LIANG X Y, ZHOU F R, et al. Automatic crop image acquisition system based on crawler car by magnetic guidance[J]. Journal of Huazhong Agricultural University, 2020, 39 (2): 141-149 (in Chinese with English abstract).
- [16] 刘洋,冯全,王书志.基于轻量级 CNN 的植物病害识别方法及移动端应用[J].农业工程学报,2019,35(17):194-204.LIU Y, FENG Q, WANG S Z. Plant disease identification method based on lightweight CNN and mobile application[J]. Transactions of the CSAE, 2019, 35(17): 194-204 (in Chinese with English abstract).
- [17] 刘洋,周宁宁.基于 SLIC 的图像分割方法研究[J].计算机技术与发展,2019,29(1):75-79.LIU Y, ZHOU N N. Research on image segmentation method based on SLIC[J]. Computer technology and development, 2019, 29 (1): 75-79 (in Chinese with English abstract).
- [18] 张小凤,刘向阳.基于图像超像素分析的图像分割方法[J].计算机技术与发展,2018,28(7):25-28,47.ZHANG X F, LIU X Y. Image segmentation based on image superpixel analysis[J]. Computer technology and development, 2018, 28(7): 25-28, 47 (in Chinese with English abstract).

## Semantic segmentation based road recognition technology of hilly fruit and tea garden

WU Weibin<sup>1,2</sup>, TANG Ting<sup>1</sup>, LIU Qiang<sup>1</sup>, ZHAO Xin<sup>1,2</sup>, HAN Chongyang<sup>1</sup>, LI Jie<sup>1</sup>

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

2. Ministry of Education Key Laboratory of Key Technology on Agricultural Machine and Equipment/Division of Citrus Machinery, China Agriculture Research System/Guangdong Engineering Technology Research Center for Creative Hilly Orchard Machinery, Guangzhou 510642, China

**Abstract** Aiming at the trend of continuous expansion of fruit and tea gardens, the gradual development of intelligent agricultural mechanization and the lack of fruit and tea garden road scenes in common road semantic segmentation data sets, the semantic segmentation technology was used to some fruit and tea garden roads in Meizhou City, Guangdong Province to realize pixel-level segmentation of roads in fruit and tea gardens. Roads, people, and cars were used as classification objects to establish a scene image data set including 6 032 images of fruit and tea garden road. The data set was randomly divided into a training set including 5 429 images and a test set including 603 images according to a 9 : 1 ratio. The MS-PSPNet semantic segmentation model was established based on the PSPNet (pyramid scene parsing network) segmentation model for optimization. The results of training showed that MS-PSPNet model mean intersection over union (mean intersection over union, MIoU) was 83.41%. The number of frames per second (frames per second, FPS) was 22.31. The MS-PSPNet model was applied to fruit and tea gardens under different road conditions and light intensity to conduct field tests and evaluate the accuracy. The results showed that the category pixel accuracy (mean pixel accuracy, MPA) of MS-PSPNet model exceeded 92%. MIoU exceeded 91% in all cases of non-hardened road conditions. It is indicated that the MS-PSPNet model has good validity and applicability in road recognition of fruit and tea gardens.

**Keywords** hilly fruit and tea garden; intelligent orchard machinery; road recognition; semantic segmentation; superpixel segmentation; precision agriculture; autonomous navigation

(责任编辑:陆文昌)