

doi:10.6041/j.issn.1000-1298.2022.11.025

基于改进 DeepLab V3 + 的果园场景多类别分割方法

刘 慧¹ 姜建滨¹ 沈 跃¹ 贾卫东² 曾 潇¹ 庄珍珍¹

(1. 江苏大学电气信息工程学院, 镇江 212013; 2. 江苏大学农业工程学院, 镇江 212013)

摘要: 果园环境实时检测是保证果园喷雾机器人精准作业的重要前提。本文提出了一种基于改进 DeepLab V3 + 语义分割模型的果园场景多类别分割方法。为了在果园喷雾机器人上部署,使用轻量化 MobileNet V2 网络替代原有的 Xception 网络以减少网络参数,并在空洞空间金字塔池化 (Atrous spatial pyramid pooling, ASPP) 模块中运用 ReLU6 激活函数减少部署在移动设备的精度损失,此外结合混合扩张卷积 (Hybrid dilated convolution, HDC),以混合扩张卷积替代原有网络中的空洞卷积,将 ASPP 中的扩张率设为互质以减少空洞卷积的网格效应。使用视觉传感器采集果园场景 RGB 图像,选取果树、人、天空等 8 类常见的目标制作了数据集,并在该数据集上基于 Pytorch 对改进前后的 DeepLab V3 + 进行训练、验证和测试。结果表明,改进后 DeepLab V3 + 模型的平均像素精度、平均交并比分别达到 62.81% 和 56.64%,比改进前分别提升 5.52、8.75 个百分点。模型参数量较改进前压缩 88.67%,单幅图像分割时间为 0.08 s,与原模型相比减少 0.09 s。尤其是对树的分割精度达到 95.61%,比改进前提高 1.31 个百分点。该方法可为喷雾机器人精准施药和安全作业提供有效决策,具有实用性。

关键词: 果园; 喷雾机器人; 语义分割; DeepLab V3 +; 混合扩张卷积; 感受野

中图分类号: TP391; S24 文献标识码: A 文章编号: 1000-1298(2022)11-0255-07

OSID:



Multi-category Segmentation of Orchard Scene Based on Improved DeepLab V3 +

LIU Hui¹ JIANG Jianbin¹ SHEN Yue¹ JIA Weidong² ZENG Xiao¹ ZHUANG Zhenzhen¹

(1. School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China

2. School of Agricultural Engineering, Jiangsu University, Zhenjiang 212013, China)

Abstract: Real-time detection of orchard environment is an important prerequisite to ensure the accurate operation of orchard spray robot. An improved DeepLab V3 + semantic segmentation model was proposed for multi-category segmentation in orchard scene. For deployment on the orchard spray robot, the lightweight MobileNet V2 network was used to replace the original Xception network to reduce the network parameters, and ReLU6 activation function was applied in atrous spatial pyramid pooling (ASPP) module to reduce the loss of accuracy when deployed in mobile devices. In addition, hybrid dilated convolution (HDC) was combined to replace the void convolution in the original network. The dilated rates in ASPP were prime to each other to reduce the grid effect of dilated convolution. The RGB images of orchard scene were collected by using visual sensor, and eight common targets were selected to make the dataset, such as fruit trees, pedestrians and sky. On this dataset, DeepLab V3 + before and after improvement was trained, verified and tested based on Pytorch. The results showed that the mean pixel accuracy and mean intersection over union of the improved Deeplab V3 + model were 62.81% and 56.64%, respectively, which were 5.52 percentage points and 8.75 percentage points higher than before improvement. Compared with the original model, the parameters were reduced by 88.67%. The segmentation time of a single image was 0.08 s, which was 0.09 s less than the original model. In particular, the accuracy of tree segmentation reached 95.61%, which was 1.31 percentage points higher than before improvement. This method can provide an effective decision for precision spraying and safe operation of the spraying robot, and it was practical.

Key words: orchard; spray robot; semantic segmentation; DeepLab V3 +; hybrid dilated convolution; receptive field

收稿日期: 2022-01-13 修回日期: 2022-03-20

基金项目: 国家自然科学基金项目(32171908)、江苏省现代农机装备与技术示范推广项目(NJ2021-14)、宁夏回族自治区重点研发计划重大项目(2018BBF02020)、江苏省重点研发计划项目(BE2018372)和江苏高校优势学科建设工程项目(PAPD)

作者简介: 刘慧(1980—),女,教授,博士生导师,主要从事农业电气化与自动化、智能控制与信号处理研究,E-mail: amity@ujs.edu.cn

0 引言

我国是果业大国,露天果园环境的病虫害防治主要途径是农药喷洒。随着我国农业机械化水平的提高和现代化果园的规模化种植,果园喷雾机器人成为林果园植保机械的重要研究内容和发展方向^[1-2]。果园喷雾机器人的喷雾决策离不开果园环境感知,尤其是靶标识别,目前常用的果园喷雾靶标探测技术包括超声波探测、红外线探测、激光雷达探测等^[3-4]。红外传感器易受光照强度影响且只能探测靶标的有无,无法探测靶标的形态特征;超声波探测精度低,范围小,需多传感器配合;激光雷达成本较高^[5]。果园场景的实时感知和多目标分割,能够为喷雾机器人的变量施药和安全作业提供有效决策,是果园喷雾机器人精准作业的重要前提和保障^[3]。近些年来,计算机视觉、图像处理技术在农业生产中得到了广泛的应用^[6-9]。基于深度学习的语义分割^[10-13]和目标检测^[14-15]已成为大量学者的研究方向,果园喷雾机器人使用计算机视觉技术对周围场景实时检测,对靶标精准识别,可以提高农药利用率,从而减少农药污染。

近些年来,学者们提出了一系列基于深度学习

的图像语义分割方法^[16],其中 DeepLab V3 + 在 PASCAL VOC 2012 和 Cityscapes 数据集上分别达到 89.0% 和 82.1% 的测试集性能^[17],其精度可以满足在果园环境下分割达到较高准确率的要求。但其主干网络 Xception^[18] 计算量大,消耗大量显存,且模型占用计算机内存大,尚未满足在机器人移动设备上部署的要求。

为了保证果园喷雾机器人精准作业,本文旨在实现精准识别靶标,基于 DeepLab V3 + 算法进行改进和优化,采用 MobileNet V2 作为主干网络,并对卷积进行改进,同时减少网络计算量,设计更为轻量化的网络结构,提高网络的运行速度和分割精度,以满足在移动端部署的要求,从而达到在复杂果园环境中,喷雾机器人进行精准喷雾和安全作业的目的。

1 果园场景语义分割模型

1.1 DeepLab V3 + 网络

DeepLab V3 + 是在 DeepLab V3 的基础上进行了改进,它是一种新的编码器-解码器结构。DeepLab V3 作为其编码器部分,添加了一个简单有效的解码器模块,DeepLab V3 + 总体原理如图 1 所示。

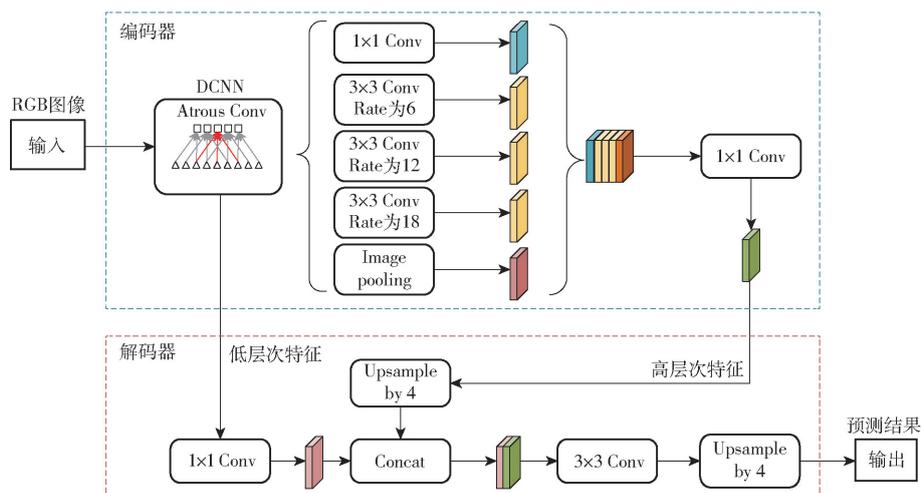


图1 DeepLab V3 + 总体原理图

Fig. 1 DeepLab V3 + overall principle diagram

编码器使用空洞空间金字塔池化,以 Xception 网络作为特征提取网络,并引入深度可分离卷积进行改进应用到 ASPP 中,采用多网格法,3 个卷积采用不同的扩张率,将得到的图像级特征输入 256 个滤波器的 1×1 卷积并进行批量归一化处理。

Xception 由具有残差连接的深度可分离卷积线性堆叠,分为 Entry flow、重复 16 次的 Middle flow 以及 Exit flow。在经过批量归一化处理后进行 1×1 卷积,对通道降维以得到编码器部分的输出。

解码器部分旨在连接低层次特征与高层次特

征,解码器部分同样应用深度可分离卷积,将编码器部分输出(高层次特征, output_stride 为 16)经过 4 倍双线性上采样得到的特征层与编码器中对应着相同分辨率的特征层(低层次特征, output_stride 为 4)进行 1×1 卷积降维后融合,再经过一个 3×3 卷积细化,最终再经过 4 倍双线性上采样得到原图分辨率的预测结果。

1.2 改进的 DeepLab V3 + 果园场景语义分割模型

1.2.1 改进 DeepLab V3 + 主干网络的轻量化设计

DeepLab V3 + 在语义分割方面已经能达到较高

精度,但其主干网络 Xception 参数量庞大,消耗大量的显存,在 Xception 组成部分中,虽然网络层数最多的是 Middle flow,但显存消耗最大的却是 Entry flow,这是由于在 Entry flow 里的特征还未被充分下采样,分辨率相当高,同时 Entry flow 里的通道数很多,这也是导致 Xception 计算量大的一个重要原因,因而需要耗费较长的时间。

农业机器人需快速获取周围环境信息,形成决策并控制执行器做出反应。为了在果园喷雾机器人上部署,引入轻量化网络 MobileNet V2^[19], MobileNet 系列网络以深度可分离卷积^[20] (Depthwise separable convolution) 来减少运算量及参数量,适合在移动设备上部署。将 MobileNet V2 作为基础网络,为了满足精度与速度之间的平衡,且对于一般情况下的果园喷雾机器人而言,达到较为精准和快速的作业即可,故而在 MobileNet V2 的基础上,将 ASPP 中的 ReLU^[21] 激活函数替换成 ClippedReLU^[22] 激活函数,ClippedReLU 激活函数是 ReLU 激活函数的变形,在 ReLU 的基础上增加了上限 z ,即

$$f(x) = \min(\text{relu}(x), z) = \min(\max(0, x), z) \quad (1)$$

式中 $f(x)$ ——ClippedReLU 激活函数的输出

x ——输入特征值

z ——ClippedReLU 激活函数上限

令 $z = 6$, 则得到 ReLU6 激活函数,对于果园喷雾机器人移动设备部署,主要是对果树进行识别,在果园喷雾存在误差的低精度场景下,ReLU6 较 ReLU 具有更好的鲁棒性^[14]。

1.2.2 混合扩张卷积

自 DeepLab V2 起,空洞(扩张)卷积被运用在 ASPP 中,空洞卷积是在标准的卷积图注入空洞,以此增加感受野,与标准的卷积相比,空洞卷积多了一个超参:扩张率,即卷积图上卷积核之间的间隔数。在语义分割领域,将原始图像输入到卷积网络中,传统的方法会对图像卷积后再做池化处理,降低图像尺寸的同时增大感受野,但语义分割是像素级的预测,在最后会进行上采样以得到与原图图像尺寸一样的预测结果,但在先减小后增大尺寸的过程中,伴随着信息的损失,而空洞卷积的引入一定程度上可解决这一问题,因空洞卷积在非池化操作卷积的情况下,增大了感受野。但空洞卷积存在网格效应问题,网格效应指多次叠加多个具有相同扩张率的卷积核会造成网格中有一些像素自始至终都没有参与运算。当空洞卷积在高层使用的扩张率变大时,对输入的采样会变得很稀疏,不利于学习,因为一些局部信息完全丢失,而长距离上的一些信息可能并不

相关;并且网格效应可能会打断局部信息之间的连续性,虽然 ASPP 叠加扩张率不同的卷积,但这些扩张率并不互质,仍然会存在网格效应。

语义分割不同于目标检测,不仅仅识别目标,还需了解像素分布,尤其是针对果园喷雾机器人,果树形态分布对喷雾的控制决策十分重要,即果树的特征信息很重要,而 DeepLab V3+ 的 ASPP 采用的空洞卷积存在信息丢失问题。故本文针对空洞卷积网格效应的问题,引入混合扩张卷积^[23],混合扩张卷积将卷积的扩张率设置为锯齿状,如 $[r_1, r_2, \dots, r_i, \dots, r_n], \dots, [r_1, r_2, \dots, r_i, \dots, r_n]$ 的循环结构,并满足

$$M_i = \max(M_{i+1} - 2r_i, M_{i+1} - 2(M_{i+1} - r_i), r_i) \quad (2)$$

其中

$$M_n = r_n$$

式中 M_i ——第 i 层卷积最大扩张率

r_i ——第 i 层卷积的扩张率

式(2)应满足 $M_2 \leq K$ (卷积核大小为 $K \times K$),以此来保证至少能够以扩张率 1 来弥补所有空洞,即至少以标准卷积计算。

在此基础之上,将 ASPP 中的卷积扩张率设置为 1、2、5、9,以此实现对不同像素占比大小目标的特征捕获,大的扩张率对应像素占比大的目标,小扩张率对应像素占比小的目标。HDC 的优点是使用任意满足式(2)的扩张率,可以在不增加额外模块的情况下扩大网络的感受野。

改进后的 ASPP 模块如图 2 所示,图中 DW 表示深度卷积 (Depthwise convolution),PW 表示逐点卷积 (Pointwise convolution),DW + PW 则组成深度可分离卷积, BN 表示批量归一化 (Batch normalization),Dilated 表示该卷积为空洞(扩张)卷积。

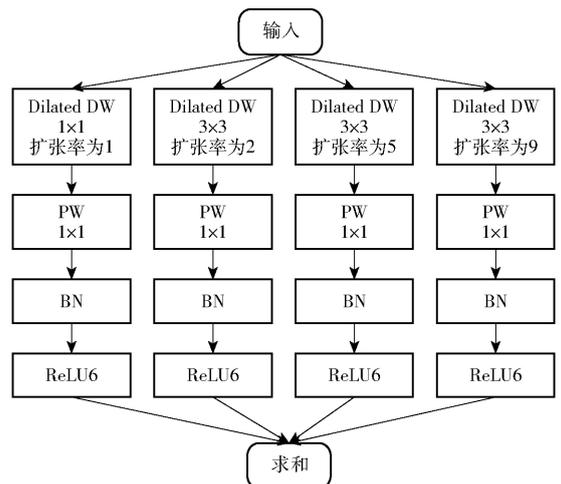


图 2 改进后的 ASPP 模块

Fig. 2 Improved ASPP module

1.2.3 改进后的 DeepLab V3+ 网络结构

将 Mobilenet V2 替代 DeepLab V3+ 的 Xception

主干网络,将 HDC 的思想应用到 DeepLab V3 + 的 ASPP 模块中,把 ASPP 模块中的卷积扩张率设置成

锯齿状结构,这些扩张率互质,改进后的网络简图如图 3 所示。

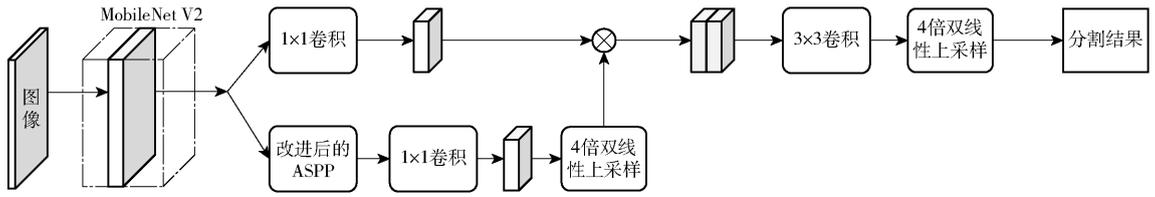


图 3 改进后的 DeepLab V3 + 网络

Fig. 3 Improved DeepLab V3 + network

2 果园场景分割试验

2.1 试验平台

深度学习硬件平台为一台配有 Intel(R) Core (TM) i9-10900K CPU 64 GB 内存、NVIDIA Quadro RTX 4000 GPU 的计算机,安装有 CUDA10.0 版本的并行计算框架和 CUDNN7.0 版本的深度神经网络加速库。在 Linux 操作平台基于 Pytorch 框架实现本文的 DeepLab V3 + 语义分割模型并完成对模型的训练、验证及测试。

2.2 试验数据采集及标注

由于目前还有很多未包含多类别的果园场景公开数据集,所以本试验中使用的是实景拍摄并制作的数据集。用车载 Realsense D435 型相机(图 4,相机距离地面 1.34 m)在果园内共拍摄了 360 幅 RGB 图像,图像分辨率为 960 像素 × 720 像素。包含 8 类果园常见场景目标:天空、树、人、路、车(农用车)、建筑、地面及标志物(灌溉管、支撑杆和标志牌等)。



图 4 喷雾机器人试验平台

Fig. 4 Spray robot experimental platform

采用 GamVid 数据集格式^[24],使用其官方标注工具 InteractLabeler 对每幅图像进行像素级数据标注,得到真实框(Ground truth)用于训练,将标注图像保存为 PNG 格式,如图 5 所示。除 8 类常见果园

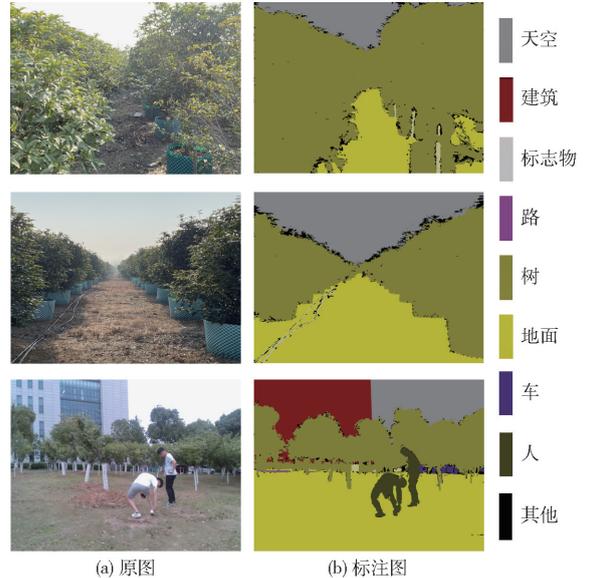


图 5 数据集图像示意图

Fig. 5 Schematic of dataset images

场景目标外,在对采集图像进行标注时,像素占比小的目标标注较为困难且存在误差,将其新增为一类:其他类,即未标注类。将数据集按照比例 4:1 分为训练集和测试集。

2.3 模型训练

为防止训练过拟合、提高网络准确率、减少图像中光照影响,同时扩增图像的数量,首先采用直方图均衡化增强图像的对比度,再通过随机翻转、随机平移、随机裁剪、随机反射进行数据增强,图像数量扩充为 1 440 幅,图像分辨率为 960 像素 × 720 像素,在训练前将 DeepLab V3 + 在 GamVid 数据集上的公开预训练模型的参数值除分类层之外都赋给本文模型的初始化参数,分类层采用均值为 0、标准差为 0.01 的高斯分布随机初始化。用于训练的优化算法是动量随机梯度下降算法。损失函数 L 为交叉熵损失函数;训练迭代次数设为 30 000,冲量设为 0.9,学习率设为 0.01,采用 Poly 策略,批量大小设为 8,每 100 次迭代验证一次并保存为最新训练模型,最终选取保存精度最高的模型。计算式为

$$L = \frac{1}{N} \sum_i L_i = \frac{1}{N} \sum_{i=1}^N \sum_{c=1}^M -y_{ic} \lg p_{ic} \quad (3)$$

式中 N ——当前 batch 样本量
 M ——类别
 y_{ic} ——指示变量, 如果 i 类别和样本 c 类别相同则为 1, 否则是 0
 p_{ic} ——观测样本 i 属于类别 c 的预测概率

3 试验结果与分析

3.1 分割评价指标

分割评价的指标包括像素精度 (Pixel accuracy, PA)、平均像素精度 (Mean pixel accuracy, MPA)、交并比 (Intersection over union, IoU)、平均交并比 (Mean intersection over union, MIoU)。

3.2 改进 DeepLab V3 + 模型对果园场景的分割结果

分别对改进前后的模型进行测试, 8 类主要场景目标分割结果对比如表 1 所示。

表 1 改进前后 DeepLab V3 + 模型的 8 类果园场景目标分割结果对比

Tab. 1 Comparison of target segmentation results of eight types of orchard scenes of DeepLab V3 + model before and after improvement %

类别	模型	PA	IoU
树	改进前 DeepLab V3 +	94.30	87.13
	改进后 DeepLab V3 +	95.61	89.61
人	改进前 DeepLab V3 +	71.23	44.37
	改进后 DeepLab V3 +	92.87	79.44
车	改进前 DeepLab V3 +	0.71	0.71
	改进后 DeepLab V3 +	0.08	0.08
地面	改进前 DeepLab V3 +	89.98	82.06
	改进后 DeepLab V3 +	94.75	86.81
路	改进前 DeepLab V3 +	87.61	82.06
	改进后 DeepLab V3 +	89.98	82.25
标志物	改进前 DeepLab V3 +	1.03	1.03
	改进后 DeepLab V3 +	3.58	3.52
建筑	改进前 DeepLab V3 +	62.55	40.21
	改进后 DeepLab V3 +	76.68	66.92
天空	改进前 DeepLab V3 +	91.99	85.76
	改进后 DeepLab V3 +	94.24	86.74

由表 1 可知, 除车外, 改进后模型各类别的分割精度和交并比均比改进前有不同程度的提高。由于在果园中, 车出现的频率较小, 且在行间行走时, 一般不会同时有多辆喷雾机器人同时作业, 所以对车类的识别不会影响喷雾机器人作业安全和效率。除了车以外, 标志物也是出现频率很低的目标, 所以这两者精度和平均交并比都较低, 但对于果园喷雾机器人而言, 果树是主要的识别目标, 改进后的模型对树的 PA 比改进前提高 1.31 个百分点, IoU 提高 2.48 个百分点, 这说明改进后的模型更利于果园喷雾机器人的靶标识别, 对精准喷雾具有指导决策的

作用。

3.3 不同模型的分割结果对比分析

针对不同主干网络, 基于 Pytorch 分割结果如图 6 所示。图 6c ~ 6f 分别为 4 种模型的分割结果, 其中图 6e 为改进前以 Xception 为主干网络的 DeepLab V3 + 分割结果, 图 6f 为本文 DeepLab V3 + 分割结果。本文方法以 MobileNet V2 为主干网络的 DeepLab V3 + 模型总体优于主干网络为 Xception、ResNet50 以及 ResNet101 的模型, 对像素占比大的目标, 分割效果更为明显; 改进前以 Xception 为主干网络的 DeepLab V3 + 模型识别错误率相对较高 (尤其是对天空出现大片的错误识别), 本文改进的 DeepLab V3 + 未出现大片的误识别, 说明改进的 DeepLab V3 + 模型其分割准确率有了一定的提升。除天空外, 4 种模型对像素占比大的目标识别相差不大, 但对距离较远, 像素占比小的目标识别却有差别, 如改进前以 Xception 为主干网络的 DeepLab V3 + 模型对远处像素占比小的建筑并不敏感, 并未识别出远处像素占比小的建筑; 对于灌溉管的识别, 改进后的模型表现明显优于其他 3 种; 对人的识别, 改进前以 Xception 为主干网络的 DeepLab V3 + 模型表现也不如其他 3 种, 改进后模型对人的识别, 轮廓的分割更为清晰, 更与标注贴近; 改进前以 Xception 为主干网络的 DeepLab V3 + 模型对支撑杆并不敏感, 虽然与树干的颜色相近, 但其他 3 种对支撑杆都有一定的识别。

使用不同主干网络, 不同模型的分割效果及模型参数量对比如表 2 所示。由表 2 可知, 本文改进后的 DeepLab V3 + 模型的像素精度达到 91.84%, 比改进前以 Xception 为主干网络的 DeepLab V3 + 模型提高 2.75 个百分点, 高出以 ResNet50 为主干网络的 DeepLab V3 + 模型 0.7 个百分点, 比以 ResNet101 为主干网络的 DeepLab V3 + 低 0.26 个百分点。本文改进后 DeepLab V3 + 模型的 MPA 和 MIoU 分别达到 62.81% 和 56.64%, 比改进前以 Xception 为主干网络的 DeepLab V3 + 分别提升 5.52 个百分点和 8.75 个百分点, 且单幅图像分割时间也缩短 0.09 s。MPA 和 MIoU 略低于 DeepLab V3 + (ResNet50) 和 DeepLab V3 + (ResNet101), 但其分割速度快于 DeepLab V3 + (ResNet50) 和 DeepLab V3 + (ResNet101), 尤其是参数量较以 ResNet50、ResNet101 和 Xception 为主干网络的模型分别压缩 83.73%、93.30% 和 88.67%。试验结果表明改进后模型表现较好, 且速度有一定的提升, 更利于语义分割模型在果园农业设备上的部署和应用。

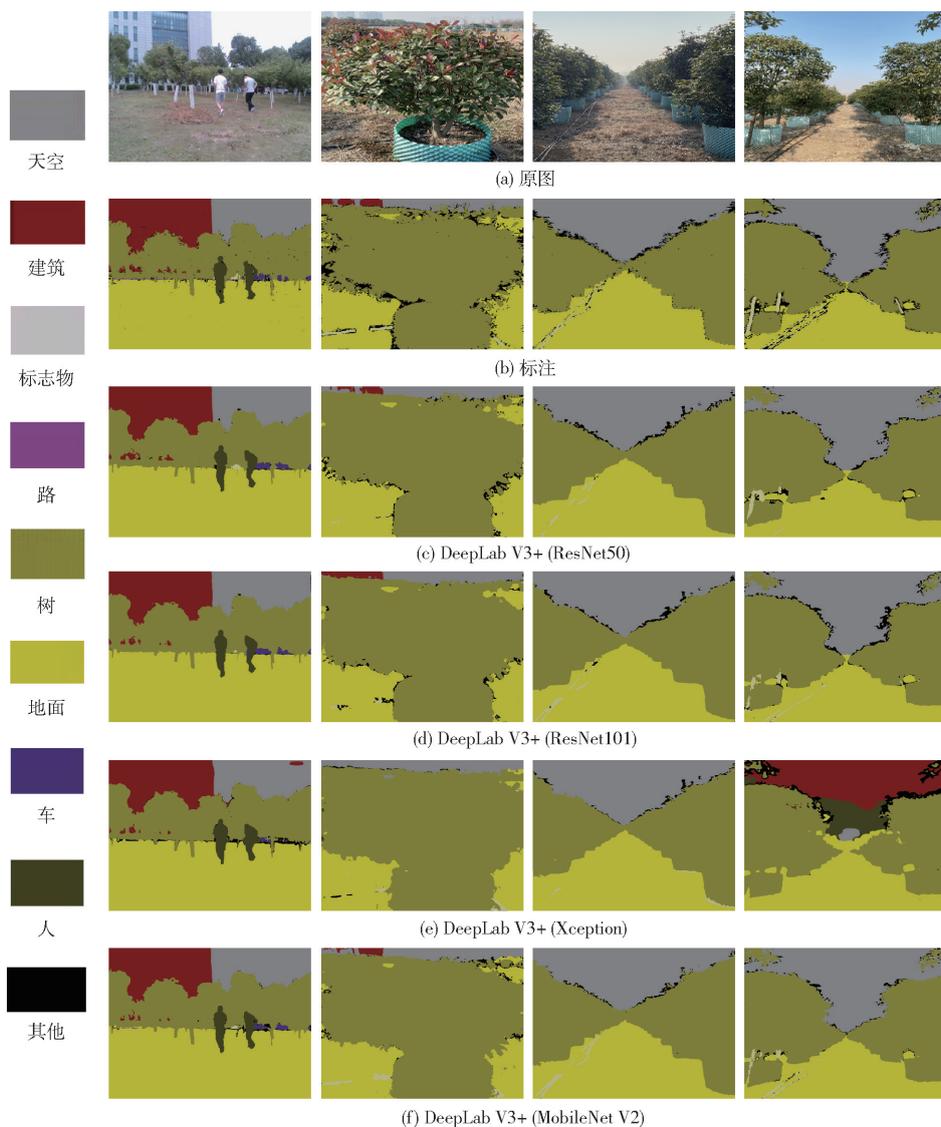


图6 不同模型的分割结果

Fig. 6 Segmentation results of different models

表2 DeepLab V3+不同主干网络的分割结果对比

Tab. 2 Comparison of segmentation results of DeepLab V3+ with different backbone networks

模型(主干网络)	PA/%	MPA/%	MIOU/%	分割时间/(s·幅 ⁻¹)	参数量/MB
DeepLab V3+ (ResNet50)	91.14	64.03	58.16	0.11	58.1
DeepLab V3+ (ResNet101)	92.10	64.63	58.85	0.14	141.0
DeepLab V3+ (Xception)	89.09	57.29	47.89	0.17	83.4
DeepLab V3+ (MobileNet V2)	91.84	62.81	56.64	0.08	9.45

4 结论

(1) 提出了改进的 DeepLab V3+ 语义分割模型,并将其用于果园场景多类别的分割。引入 MobileNet V2,替代原有 Xception 主干网络,并在 ASPP 模块中用 ReLU6 激活函数替代原有 ReLU 激活函数,将 HDC 思想应用到 DeepLab V3+ 的 ASPP 中,把 ASPP 中的卷积扩张率设置成锯齿状结构,且扩张率互质。

(2) 根据果园中常见的场景目标制作了数据

集,包含果树、行人、天空等 8 类常见场景目标,分别用于改进前后 DeepLab V3+ 语义分割模型的训练和测试。结果表明,改进后的 Deeplab V3+ 模型具有较高的准确性和实时性,平均像素精度、平均交并比分别达到 62.81% 和 56.64%,比改进前以 Xception 为主干网络的 DeepLab V3+ 模型分别提高 5.52、8.75 个百分点;尤其是对树的识别在平均像素精度、平均交并比分别达到 95.61% 和 89.61%,比改进前提高 1.31、2.48 个百分点;改进后模型参数量压缩 88.67%,单幅图像分割时间为 0.08 s。但

改进前后模型共同存在的问题是对远处像素占比小的目标,其识别分割效果一般,但对于果园内行驶的喷雾机器人只需对近处的物体进行快速、准确识别和分割即可,从而实施精准喷雾作业。

参 考 文 献

- [1] 边永亮, 李建平, 薛春林, 等. 果园风送式喷雾机智能化发展现状与前景分析[J]. 东北农业大学学报, 2020, 51(2): 86-94. BIAN Yongliang, LI Jianping, XUE Chunlin, et al. Analysis on status quo and prospects of intelligent development of air-feed sprayer in orchard[J]. Journal of Northeast Agricultural University, 2020, 51(2): 86-94. (in Chinese)
- [2] JIN Y C, LIU J Z, XU Z J, et al. Development status and trend of agricultural robot technology[J]. International Journal of Agricultural and Biological Engineering, 2021, 14(4): 1-19.
- [3] 翟长远, 赵春江, WANG Ning, 等. 果园风送喷雾精准控制方法研究进展[J]. 农业工程学报, 2018, 34(10): 1-15. ZHAI Changyuan, ZHAO Chunjiang, WANG Ning, et al. Research progress on precision control methods of air-assisted spraying in orchards[J]. Transactions of the CSAE, 2018, 34(10): 1-15. (in Chinese)
- [4] 江世界, 马恒涛, 杨圣慧, 等. 果园喷雾机器人靶标探测与追踪系统[J]. 农业工程学报, 2021, 37(9): 31-39. JIANG Shijie, MA Hengtao, YANG Shenghui, et al. Target detection and tracking system for orchard spraying robots[J]. Transactions of the CSAE, 2021, 37(9): 31-39. (in Chinese)
- [5] 南玉龙, 张慧春, 徐幼林, 等. 农林仿形对靶喷雾及其控制技术研究进展[J]. 世界林业研究, 2018, 31(4): 54-58. NAN Yulong, ZHANG Huichun, XU Youlin, et al. Research progress on profiling target spray and its control technology in agriculture and forestry[J]. World Forestry Research, 2018, 31(4): 54-58. (in Chinese)
- [6] 郑小南, 杨凡, 李富忠. 图像处理在农业应用的研究进展[J]. 物联网技术, 2020, 10(9): 80-82.
- [7] 孙俊, 张跃春, 毛罕平, 等. 基于计算机视觉的土壤镉胁迫生菜叶片污染响应分析[J]. 农业机械学报, 2018, 49(3): 166-172. SUN Jun, ZHANG Yuechun, MAO Hanping, et al. Responses analysis of lettuce leaf pollution in cadmium stress based on computer vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(3): 166-172. (in Chinese)
- [8] 钱亮. 基于 DSP 和机器视觉的精准施药平台研究[J]. 农机化研究, 2021, 43(11): 120-124. QIAN Liang. Research on precise pesticide platform based on DSP and machine vision [J]. Journal of Agricultural Mechanization Research, 2021, 43(11): 120-124. (in Chinese)
- [9] 袁洪波, 赵努东, 程曼. 基于图像处理的田间杂草识别研究进展与展望[J]. 农业机械学报, 2020, 51(增刊 2): 323-334. YUAN Hongbo, ZHAO Nudong, CHENG Man. Review of weeds recognition based on image processing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(Supp. 2): 323-334. (in Chinese)
- [10] 熊俊涛, 刘柏林, 钟灼, 等. 基于深度语义分割网络的荔枝花叶分割与识别[J]. 农业机械学报, 2021, 52(6): 252-258. XIONG Juntao, LIU Bolin, ZHONG Zhuo, et al. Litchi flower and leaf segmentation and recognition based on deep semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(6): 252-258. (in Chinese)
- [11] 郑艳梅, 黄光耀, 芦碧波, 等. 结合深度学习和引导滤波的苹果叶片图像分割[J]. 中国农业大学学报, 2020, 25(9): 112-120. ZHENG Yanmei, HUANG Guangyao, LU Bibo, et al. An apple leaf image segmentation algorithm using deep learning and guide filtering[J]. Journal of China Agricultural University, 2020, 25(9): 112-120. (in Chinese)
- [12] 王璨, 武新慧, 张燕青, 等. 基于双注意力语义分割网络的田间苗期玉米识别与分割[J]. 农业工程学报, 2021, 37(9): 211-221. WANG Can, WU Xinhui, ZHANG Yanqing, et al. Recognition and segmentation of maize seedlings in field based on dual attention semantic segmentation network[J]. Transactions of the CSAE, 2021, 37(9): 211-221. (in Chinese)
- [13] 李余康, 翟长远, 王秀, 等. 基于 DeepLab v3+ 的葡萄叶片分割算法[J]. 农机化研究, 2022, 44(2): 149-155. LI Yukang, ZHAI Zhanguan, WANG Xiu, et al. Grape leaf segmentation based on DeepLab v3+ [J]. Journal of Agricultural Mechanization Research, 2022, 44(2): 149-155. (in Chinese)
- [14] 刘慧, 张礼帅, 沈跃, 等. 基于改进 SSD 的果园行人实时检测方法[J]. 农业机械学报, 2019, 50(4): 29-35, 101. LIU Hui, ZHANG Lishuai, SHEN Yue, et al. Real-time person detection in orchard based on improved SSD[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4): 29-35, 101. (in Chinese)
- [15] 景亮, 王瑞, 刘慧, 等. 基于双目相机与改进 YOLOv3 算法的果园行人检测与定位[J]. 农业机械学报, 2020, 51(9): 34-39, 25. JING Liang, WANG Rui, LIU Hui, et al. Orchard pedestrian detection and location based on binocular camera and improved YOLOv3 algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(9): 34-39, 25. (in Chinese)
- [16] 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 Huiyu, WU Junjun. Survey of image semantic segmentation based on deep learning[J]. Computer Engineering and Applications, 2019, 55(19): 12-21, 42. (in Chinese)
- [17] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 801-818.
- [18] FRANCOIS C. Xception: deep learning with depthwise separable convolutions[J]. arXiv preprint arXiv:1610.02357v3, 2017.
- [19] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [20] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications [J]. arXiv preprint arXiv:1704.04861, 2017.
- [21] ALEX K, I SUTSKEVER, GE H. ImageNet classification with deep convolutional neural networks[C]//NIPS, 2012: 1-9.
- [22] AWNI H, CASE C, CASPER J, et al. Deep speech: scaling up end-to-end speech recognition[J]. arXiv preprint arXiv: 1412.5567v2, 2014.
- [23] WANG P, CHEN P, YUAN Y, et al. Understanding convolution for semantic segmentation[J]. arXiv preprint arXiv:1702.08502v3, 2018.
- [24] BROSTOW G J, FAUQUEUR J, CIPOLLA R. Semantic object classes in video: a high-definition ground truth database[J]. Pattern Recognition Letters, 2009, 30(2): 88-97.