

DOI: 10.13718/j.cnki.xdzk.2024.08.016

李法霖, 石军锋, 梁新成, 等. 基于改进 DeepLabV3+ 的丘陵田间道路图像分割方法研究 [J]. 西南大学学报(自然科学版), 2024, 46(8): 172-183.

基于改进 DeepLabV3+ 的丘陵田间道路 图像分割方法研究

李法霖, 石军锋, 梁新成, 李云伍, 刘鹏, 陈欣

西南大学 工程技术学院, 重庆 400715

摘要: 为解决丘陵地区智能农机装备因道路狭窄、路况复杂导致道路信息提取精度低和推理速度慢等问题, 以丘陵田间道路作为研究对象制作数据集, 提出一种基于改进 DeepLabV3+ 的丘陵田间道路图像分割方法. 首先在编码器模块中使用轻量化的主干网络 G_Ghost_RegNetX_4.0GF 提取图像特征, 保证精度并减小模型参数数量. 再采用轻量级的空洞空间金字塔池化模块, 将不同尺度特征融合. 试验结果表明, 改进模型的平均交并比和推理速度分别为 87.6% 及 116.08 f/s, 与当前主流图像分割网络 FCN、DeepLabV3 及 PSPNet 相比, MIoU 分别提升了 0.8%, 2.2%, 1%, 推理速度分别为对比网络的 1.33, 1.83, 1.76 倍. 所提模型的参数总量为 14.41×10^6 , 浮点计算量为 49.34×10^9 , 模型参数及计算量大幅减小. 改进后的算法具有较高的检测精度和推理速度, 有利于解决智能农机装备在丘陵田间道路上行驶的自主导航问题.

关键词: 丘陵道路; 机器视觉; 场景识别; 语义分割; 神经网络

中图分类号: TP183

文献标志码: A

文章编号: 1673-9868(2024)08-0172-12

开放科学(资源服务)标识码(OSID):



Research on Hilly Field Road Image Segmentation Method Based on Improved DeepLabV3+

LI Falin, SHI Junfeng, LIANG Xincheng,
LI Yunwu, LIU Peng, CHEN Xin

College of Engineering and Technology, Southwest University, Chongqing 400715, China

Abstract: To tackle the challenges of low precision in road information extraction and slow inference speeds associated with narrow roads and complex terrain in hilly areas, it is significant to initiate the creation of a dataset, with a specific focus on hilly field roads as the research subject. It is also valuable to introduce a

收稿日期: 2023-09-11

基金项目: 重庆市科技局项目(cstc2021jcyj-msxmX1062); 贵州省科技计划项目(黔科合支撑[2022]一般 168).

作者简介: 李法霖, 硕士研究生, 主要从事智能农业装备感知系统研究.

通信作者: 石军锋, 博士, 副教授, 硕士研究生导师.

hilly field road image segmentation method, leveraging an enhanced DeepLabV3+ model. In the encoder module, it is crucial to integrate a lightweight backbone network, G_Ghost_RegNetX_4.0GF, which facilitated precise feature extraction while simultaneously reducing the model's parameter count. Additionally, it is necessary to incorporate the Lite-RASPP module to fuse features of varying scales. Relevant experimental findings underscore the success of this enhanced model, achieving a remarkable Mean Intersection over Union (MIOU) of 87.6% and an impressive inference speed of 116.08 f/s. When compared to prevailing image segmentation networks like FCN, DeepLabV3, and PSPNet, our model exhibited a substantial MIOU increase of 0.8%, 2.2%, and 1% respectively, while significantly outpacing them in terms of inference speed, being 1.33, 1.83, and 1.76 times faster, respectively. Furthermore, our proposed model boasted a lean parameter count of 14.41×10^6 and a floating-point computation amount of 49.34×10^9 , substantially reducing both the model's parameters and computational demands. This optimized field image segmentation algorithm not only demonstrated enhanced detection accuracy but also delivered a superior inference speed, holding significant promise for facilitating the autonomous navigation of intelligent agricultural apparatus on hilly field roads.

Key words: hilly roads; machine vision; scene recognition; semantic segmentation; neural networks

中国总面积的 70% 左右为丘陵山地, 大约 56% 的人口居住在这些区域^[1]. 随着城镇化进程的加速, 农村劳动力向城市的持续转移导致部分耕地荒废, 这一趋势可能危及国家粮食安全. 因此, 提高丘陵地区的农业智能化水平对于保障农业正常生产和推进农业现代化具有至关重要的意义. 精准识别道路行驶区域是智能农机装备搭载视觉系统的主要目的, 由于丘陵田间道路具有狭窄蜿蜒、起伏大、路况复杂等特点, 所提取行驶区域的道路信息难以保证智能农机装备的自主导航. 因此, 精确、实时的丘陵田间道路信息提取就变得尤为重要.

在道路信息提取研究中, 传统的方法包括边缘检测和阈值分割等^[2-4], 但这些方法在处理复杂道路环境时准确率不高. 近年来, 随着神经网络技术的发展, 语义分割逐渐成为了道路信息提取的主流方法. 孟庆宽等^[5]提出一种基于通道注意力结合多尺度融合的轻量级语义分割模型, 实现对非结构化田间道路场景的识别. 但 MobileNetV2 作为主干网络采用深度可分离卷积, 易出现信息丢失和梯度消失等问题从而导致分割精度下降. 杜小强等^[6]提出一种优化 Mask R-CNN 模型的非结构化农田障碍物实例分割方法, 引入可变形卷积增大感受野并提高模型的鲁棒性. 但可变形卷积的输出结果与初始位置有关, 可能导致可变形卷积输出结果偏差较大从而影响模型的分割性能. Xu 等^[7]提出一种将 Transformer 与卷积神经网络 (CNN) 结合的 MCTNet 网络结构, 模型基于 Encoder-Decoder 在 CNN 和 Transformer 独自运行后再融合输出特征, 缺点是融合模块难以平衡各种特征的权重, 可能出现信息冲突及分割目标错误的情况. Zhang 等^[8]在 FCN 基础上提出了一种多特征全卷积网络, 通过 RGB 图像与 DEM 图像结合以提取山地道路区域. 不足之处是仅使用一次上采样将特征图恢复到原图尺寸大小, 可能影响细节和边缘信息的精确表达. Wang 等^[9]提出一种针对夜间道路场景的语义分割模型 SFNet-N, 虽然模型提高了增强后图像的真实性, 但试验结果表明其准确率和实时性还存在不足. Firkat 等^[10]提出一种用于农林业环境非结构化道路检测方法 ARDformer, 虽然试验结果表明其性能大大优于 SOTA 道路检测方法, 但计算成本过大且推理速度较慢. Liu 等^[11]基于图像级联网络 (ICNet) 架构提出一种轻量级实时语义分割网络, 通过提取不同层次特征避免了概率图模型引起的大量计算和内存消耗的问题, 但该网络易产生分割边缘模糊和不准确的问题. Tian 等^[12]提出一种多任务学习 GCS-MUL 算法, 将卷积块注意力模块 (CBAM) 作为整个模型的骨干, 设计的轻量级目标特征提取网络 Ghost CBAM (GCNet) 提高了模型分割精度并减小了模型参数量, 但多任务学习的相互作用会出现检测目标的漏检. 对于田间果园非结构化路面, 大多使用基于 Encoder-Decoder 结构的

Unet 模型提取果园道路信息^[13-14],但由于数据集的不充分及编码器卷积层的堆叠,影响提取上下文信息导致部分目标的分割精度较低。Lin 等^[15]提出一种非结构化场景的实时语义分割模型,通过融合领域泛化和注意力机制增强其在复杂无组织环境的能力,但使用轻量化的主干网络对稀有类别的感知有限可能导致分割边界模糊。Baheti 等^[16]基于非结构化道路对 DeepLabV3+语义分割网络进行改进,加入膨胀卷积的 Xception 作为主干网络,试验结果表明在流量较大的复杂环境中会出现部分重叠目标分割错误的情况。总之,上述研究基于神经网络对道路图像进行语义分割,相比于传统方法其性能有了较大提升,但也存在分割精度与模型参数量平衡的问题。

本文以丘陵田间道路作为研究对象,采集田间道路图像构造数据集,基于 DeepLabV3+网络将轻量化的 G_Ghost_RegNetX_4.0GF 作为主干网络搭建语义分割模型,引入轻量级空洞空间金字塔池化模块(Lite-RASPP)将多尺度特征融合。通过对复杂的丘陵田间道路图像进行语义分割,实时精确地提取可行驶区域和动态目标等有效信息,为智能农机装备的自主导航研究奠定基础。

1 数据集建立

1.1 试验平台和设备

丘陵田间道路狭窄、坡度起伏大、路面坑洼不平,传统轮式转运平台性能难以满足丘陵田间复杂的道路环境。鉴于履带式转运平台有较好的稳定性和路面适应能力,在丘陵田间有卓越的性能表现,故本文以履带式转运平台为试验平台。转运平台的整体布局如图 1 所示,感知硬件为无畸变工业相机,相机的像素、帧率分别为 5×10^6 和 30 f/s。

1.2 田间道路数据集制作

1.2.1 图像采集

该试验在重庆市北碚区果树重点实验室研发基地进行。研发基地中包含田地、果园、村庄等区域,基地道路类型丰富且随丘陵起伏延伸,可作为丘陵田间道路的图像采集场所。图像采集过程中手动遥控转运平台行驶路线和速度,行驶路线覆盖整个试验基地,包含了石子路、水泥路、沥青路和混合路面等多种道路类型,控制转运平台以 4 km/h 匀速行驶模拟其正常运行状态。工业相机固定于转运车正面中轴线上,数据通过 USB 端口传输并储存在手提电脑中,相机采集的图像分辨率为 640×480 像素,图像采集的帧率为 30 f/s。

1.2.2 数据集制作

从试验基地采集到的原始数据集有可能受到路面颠簸、光照及阴影等因素的影响,不能直接用于深度学习训练,需要对图片进行筛选。从原始数据集中选出 1 000 张较为清晰的图片,通过开源标注软件 EISeg 处理图像集生成标签图像,再通过人工对标注边缘进行调整。按照 8 : 1 : 1 的比例划分训练集、验证集和测试集,数据集的储存格式为 PASCAL VOC 数据集格式。

本文根据丘陵田间道路环境中对象的特征属性将环境对象类别划分为 10 种,包括道路、动态目标、障碍物、水塘、天空、标志牌、建筑、挡墙、护栏和背景,其定义见表 1,标注示例如图 2 所示。通过数据集的增强和均值化方法提高数据集的丰富程度,进而提高神经网络模型的泛化能力和鲁棒性。数据增强是指通过对原有训练数据进行一系列变换和扩充生成更多的图像数据,同时将采集到的图片格式大小转化为 520×520 像素;训练神经网络时通过程序随机翻转、裁剪及旋转图片扩充数据集并保持随机性,降低神经网

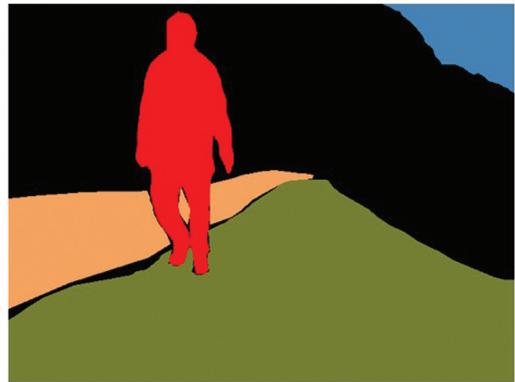


图 1 履带式转运平台

络过拟合风险, 使模型有更好的适应性和鲁棒性.

表 1 丘陵田间道路数据集标注信息

分类	颜色 RGB	定义
背景	(0, 0, 0)	指不符合以下分类的目标.
道路	(112, 128, 144)	指转运平台及动态目标移动区域, 包括水泥路、沥青路及院坝等.
动态目标	(255, 0, 0)	指行人(包含行人携带物品)、行驶的摩托车、移动的狗等.
障碍物	(255, 255, 0)	指占据路面行驶区域的石头、树干等静止物体.
水塘	(46, 139, 87)	指小型水体或人造水池.
天空	(135, 206, 250)	指未被树木大面积遮挡的天空区域.
标志牌	(0, 0, 255)	指道路信息和环境信息等的提示牌.
建筑	(128, 0, 128)	指人工修建的房屋和公共设施等.
挡墙	(244, 164, 96)	指固定道路两旁山体的墙面.
护栏	(128, 128, 128)	指果园与道路的分界物.



a. 山地道路图像

b. 标注示例

图 2 丘陵道路标注示例

2 丘陵道路图像语义分割模型构建

2.1 模型构建思路

DeepLabV3+模型是基于卷积神经网络的语义分割网络^[17], 使用了 Encoder-Decoder 的模型架构, 将空洞空间金字塔模块(ASPP)与 Decoder 模块结合提高网络的边界分割效果. 该模型选用 Xception 作为模型的主干网络并对其改进, 选用深度可分离卷积替代其最大池化层, 从而改善了主干网络的特征提取能力和网络分割精度, 不足之处是明显增加了网络的计算量并降低了推理速度.

为实现智能农机的自主导航, 语义分割模型不仅要高效准确地对丘陵田间道路信息识别分割, 还需要

在较小计算量下完成田间道路图像分割以减小对计算资源的占用。基于此,通过对 DeepLabV3+ 进行改进,搭建的语义分割模型如图 3 所示。

1) 在语义分割任务中,编码器的主干网络能够从输入图像中提取高层次的语义信息,并将其转化为适合分割任务的特征表示。这些特征使得模型能够准确地识别不同物体并进行像素级的分类和分割,从而实现图像的精确理解和分析。所构建的模型选择新型的网络模型 G_Ghost_RegNetX4_0GF(后续简称为 G_RegNetX4)作为主干网络,减小模型计算量并确保主干网络能够有效地提取丘陵田间道路目标特征。

2) 选用基于 ASPP 模块轻量化改进的 Lite-RASPP 模块作为模型的分割头(Segmentation Head),在不影响模型精度的情况下减少模型的计算量。对于每个像素,分割头计算其属于每个类别的概率,并根据最大概率值将像素归类到相应的语义类别中。通过这种方式,网络可以将输入图像的每个像素映射到相应的语义类别,从而实现像素级别的语义分割。

3) 解码器通过跳跃连接的方式结合低位特征和高级特征,融合丰富的上下文信息和空间细节信息提高分割精度,再使用双线性插值将经过编码和特征提取的信息还原至图像大小,并生成每个像素的语义标签。输出结果可以代表图像中每个像素所属的语义类别,实现了像素级别的语义分割。

模型基于 DeepLabV3+ 的 Encoder-Decoder 结构做出优化改进,通过采集设备获取丘陵道路图片,处理后传入模型的编码器,经 G_RegNetX4 主干网络提取低位特征和高级特征。低位特征经主干网络简单处理后通过跳跃连接传入解码器,再与经主干网络和分割头处理的高级特征结合,最后应用双线性插值将特征图恢复至原始尺寸并输出分割结果。

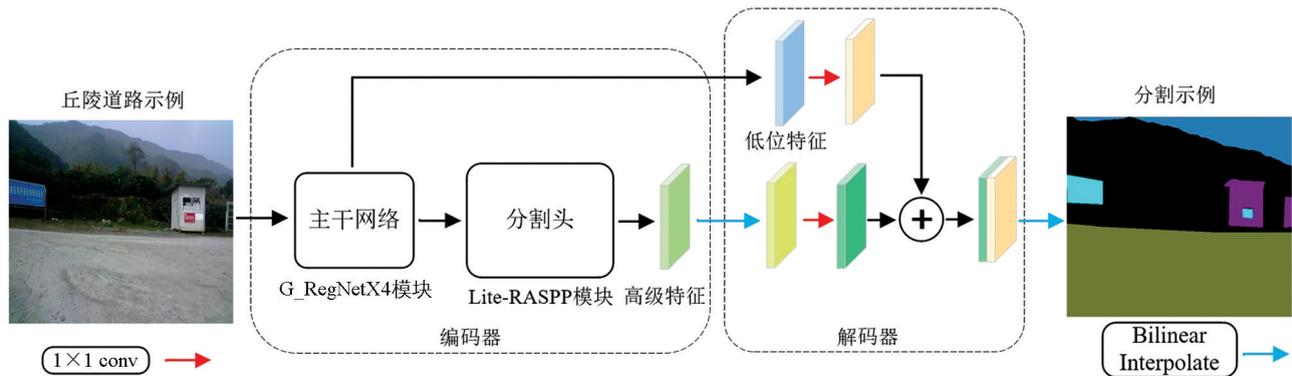


图 3 语义分割模型

2.2 主干网络

G_RegNetX4 是由华为诺亚实验室于 2022 年提出的分类模型^[18],其目的是解决由于移动设备的内存和计算资源有限,卷积神经网络(CNN)在移动设备部署困难的问题,并针对 GPU 硬件设计 G_Ghost 优化模块,进一步提高模型计算效率。

传统的卷积过程输入 X ,经过连续的卷积层 $\{B_1, B_2, \dots, B_n\}$ 得到输出特征映射 Y_n ,式(1)与式(2)分别表示第一层和最后一层的输出特征映射。G_Ghost 模块考虑到多个卷积层的堆叠会输出冗余特征,不同卷积层生成的特征存在一定的相似性,故可以在 GPU 上通过廉价变换将部分低级特征映射到高级特征,从而减小模型的计算量,提高模型的推理速度。整个 G_Ghost 模块如图 4 所示。

$$Y_1 = B_1(X) \quad (1)$$

$$Y_n = B_n(B_{n-1}(\dots B_2(Y_1))) \quad (2)$$

G_Ghost 模块中将深度特征分为复杂特征 $Y_n^c \in R^{(1-\lambda)c \times h \times w}$ 和 Ghost 特征 $Y_n^g \in R^{\lambda c \times h \times w}$,复杂特征通过一系列卷积具有更丰富的语义信息;Ghost 特征直接从第 1 层卷积层中获取,再通过 Cheap 模块(简单卷积处理,用 C 表示)使 Ghost 特征与复杂特征的通道数相同,用于后续的拼接聚合,如式(3);最后连接聚合 Y_n^c 和 Y_n^g 得到 G_Ghost 模块的最终输出,如式(4)。

$$Y_n^g = C(Y_1) \tag{3}$$

$$Y_n = [Y_n^c, Y_n^g] \tag{4}$$

不同特征的直接聚合会导致 Ghost 特征与复杂特征的信息域不符, 需要补充 Mix 模块将卷积第 2 到第 n 层的中间特征进行拼接, 将中间特征设为 $Z \in R^{c \times h \times w} = [Y_1^c, Y_2^c, \dots, Y_n^c]$, 再使用 Mix 模块(用 τ 表示)变换到与 Ghost 特征输出同域, 如式(5); 最后将聚合中间特征的 Ghost 特征与复杂特征融合得到输出. 本文所选的 G_RegNetX4 网络模型如图 5 所示.

$$Y_n^g \leftarrow Y_n^g + \tau(Z) \tag{5}$$

G_Ghost 模块一般用于 3 个及以上卷积块堆叠的结构, 在 G_RegNetX4 中当 Residuals 模块的堆积数超过 3 时, 使用 G_Ghost 模块搭建网络完成对图像复杂特征和 Ghost 特征的提取融合, 减少网络参数量并提高精度. 如图 5(a) 所示, 将 G_Ghost 模块中的 Block 块用 Residuals 模块替换.

在图 5 中, c 为卷积输出通道, k 为卷积核大小, s 为卷积步长, p 为特征图四周填充像素的行或列数, g, g_1 为分组卷积的组数, d, d_1 为膨胀卷积的膨胀因子, b_k 为 Residuals 模块堆叠次数, s_c 为捷径分支.

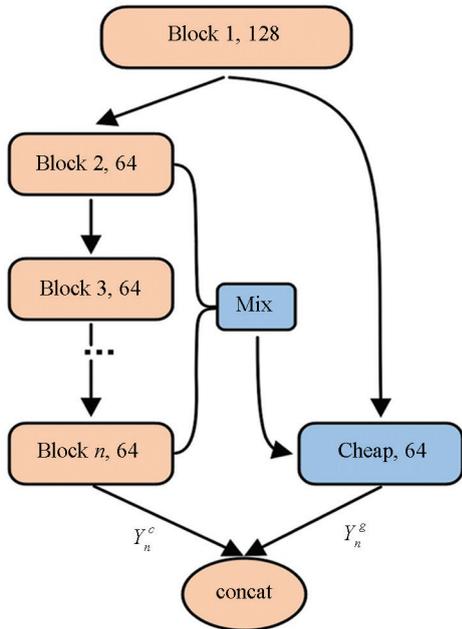


图 4 G_Ghost 模块

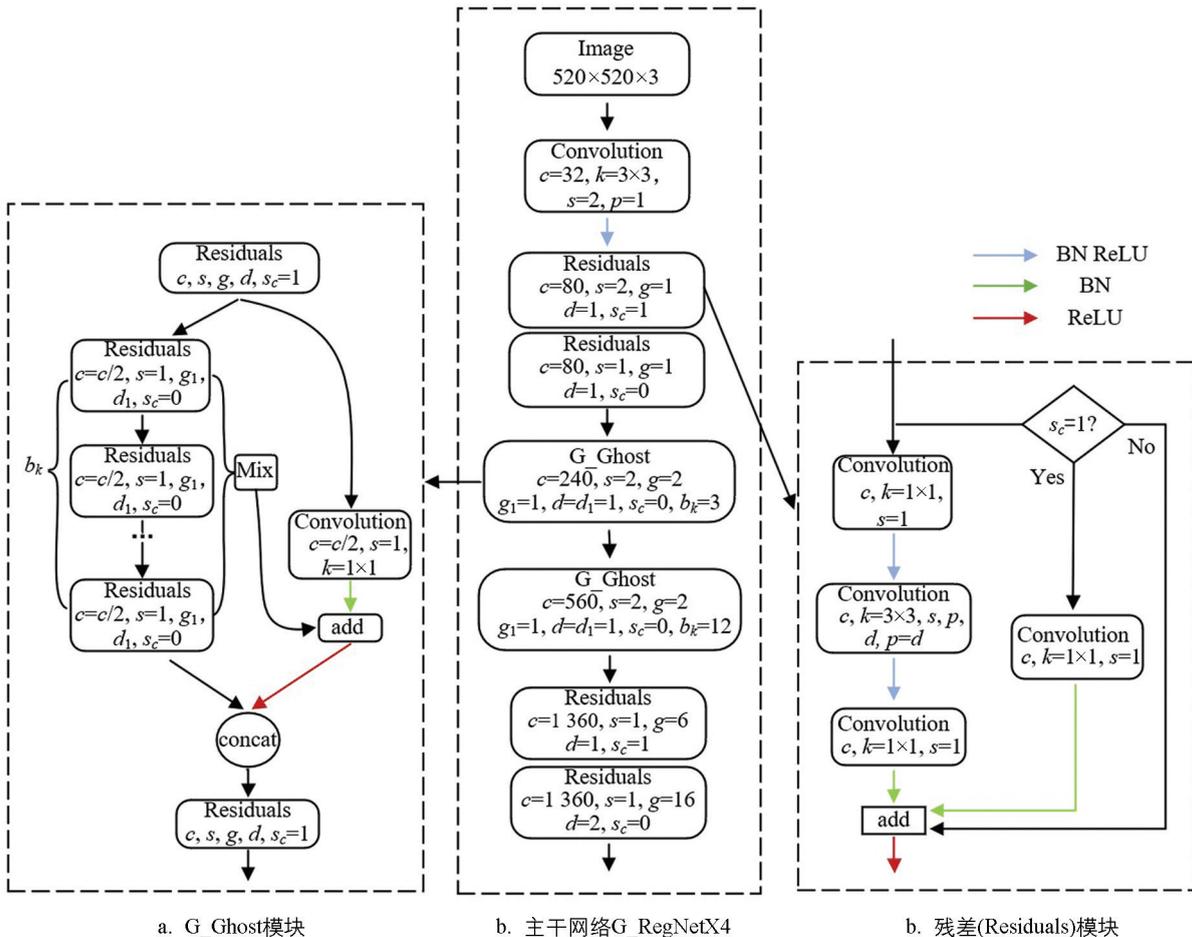


图 5 G_RegNetX4 网络模型

2.3 轻量级空洞空间金字塔池化模块

Lite-RASPP 在 MobileNetV3 中被认为是一种轻量级的分割头^[19],旨在增强模型对多尺度特征的感知能力, Lite-RASPP 结构如图 6 所示. 相比 ASPP 结构, Lite-RASPP 模块减少分支数量和膨胀卷积的膨胀率,降低了计算复杂性,使得模型更适合在资源受限的设备上使用. 在不增加参数量的情况下增大有效的感受野,通过压缩激励模块^[20]的方式部署了全局平均池化,用于主动选择和强调网络中重要的特征通道. 在主干网络的最后一个 Block 中应用空洞卷积提取更丰富的特征输入 Lite-RASPP 模块,输入特征经 2 个分支通道处理聚合获取具有更多细节信息的特征. 具体步骤为将经简单卷积、ReLU 等操作处理过的特征与全局平均池化所得特征聚合,以获取具有丰富细节信息的特征.

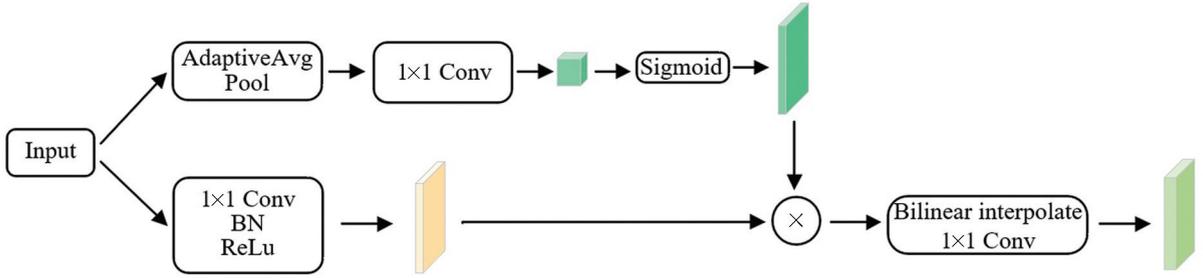


图 6 Lite-RASPP 结构

3 丘陵田间道路图像语义分割试验

3.1 训练环境

本文使用的模型训练环境为台式计算机, CPU 型号为 Intel Core i9 13900KF, 储存空间分为 32G 内存, 1T 固态硬盘, GPU 选用显存为 24G 的 NVIDIA GTX 4090 显卡; 基于 64 位 Windows 11 操作系统, 采用 Python 语言在 Pytorch 深度学习框架下进行语义分割模型搭建, 统一计算架构选择 CUDA 11.3, 神经网络加速库版本为 CUDNN 8.8.1.

3.2 模型训练

通过主干网络验证试验和模型性能试验对所提出模型进行验证, 训练网络参数如下: 学习率为 0.1, 权重衰减系数为 0.000 01, Batch_size 设置为 5, 动量指数为 0.9, 训练周期为 100. 迁移学习可以将一些标准任务上进行预训练深度神经网络模型的特征提取层重新用于其他视觉任务, 以提高模型的泛化能力^[21]. 在模型性能试验时, 对所提出模型及对比模型均载入主干网络的预训练权重, 以验证网络模型性能.

3.3 评价指标

为了验证搭建的语义分割模型在丘陵田间道路的分割效果, 使用了像素准确率 PA(Pixel Accuracy)、平均像素准确率 MPA(Mean Pixel Accuracy)、平均交并比 MIoU(Mean Intersection over Union)、浮点运算量 FLOPs(Floating Point Operations)、参数总量 Params 和推理速度 FPS(Frame Per Second) 6 个指标对语义分割模型进行性能评价. 其中评价指标 PA、MPA、MIoU 具体值的计算公式分别见式(6)~式(8), 语义分割分类数一般为 $k+1$ 类(k 类目标与 1 个背景分类).

$$A_P = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}} \quad (6)$$

$$A_{MP} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \times 100\% \quad (7)$$

$$M_{IoU} = \frac{1}{k+1} \frac{\sum_{i=0}^k p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ij} - p_{ii}} \quad (8)$$

式中: A_p 为像素准确率(%); A_{MP} 为平均像素准确率(%); M_{IoU} 为平均交并比(%); P_{ii} 表示属于 i 类且被预测为 i 类的像素数(预测正确像素数); P_{ij} 表示属于 i 类却被预测为 j 类的像素数(预测错误像素数).

3.4 主干网络验证

主干网络的特征提取能力及参数量直接影响语义分割模型的分割精度及计算效率, 对于语义分割网络而言, 合适的主干网络保证分割精度的同时还可以提升网络的计算效率. 本文对比分析了以 RegNet^[22]、MobileNetV3、ResNet 系列^[23]、G_RegNetX4 作为 DeepLabV3+ 的主干网络时的模型性能, 未载入预训练权重的情况下对加载不同主干网络的 DeepLabV3+ 模型进行训练验证.

表 2 记录了不同网络的 Params 和 FPS 以及网络在测试集中的 MPA 和 MIoU 指标. 分析表 2 可知, MobileNetV3 在 FPS 和 Params 上的表现优于其他模型, 但是它的 MPA 和 MIoU 均小于其他网络, 其 MPA 与 MIoU 比 G_RegNetX4 分别低了 0.73% 和 2.5%, MobileNetV3 使用轻量级深度可分离卷积搭建卷积模块, 该方法能够显著减少计算量和参数量, 但该方法也降低了模型的分割精度. ResNet101 的 MPA 和 MIoU 为对比主干网络中的最优量, 分别为 84.21% 和 77.1%, 比 G_RegNetX4 分别提高了 1.73% 和 1.3%, 但是其 Params 为 G_RegNetX4 的 2.30 倍, FPS 比 G_RegNetX4 慢了 30.15 f/s, ResNet101 以残差结构搭建模型, 网络层数增加提升分割精度的同时避免了网络的退化, 但是网络层数的叠加也会造成网络参数量较多以及推理速度较慢的缺点. 从表 2 中可以看出 ResNet50 和 RegNet 作为主干网络的 DeepLabV3+ 模型的 Params、FPS、MPA 和 MIoU 都不如 G_RegNetX4.

由图 7 可知, 通过观察发现各个网络能够快速稳定地达到收敛, 且损失值在 40 个 Epoch 后开始趋于稳定, 继续训练损失下降幅度微小, 此时可以停止训练. 此外, 以 RegNet、MobileNetV3 作为主干网络的 MIoU 训练曲线波动较大, 训练效果不稳定.

表 2 不同主干网络性能对比

主干网络	参数总量	推理速度 FPS/	平均像素准确率 MPA/	平均交并比 MIoU/
	Params	(f · s ⁻¹)	%	%
G_RegNetX4	25.49 × 10 ⁶	73.29	82.48	75.8
MobileNetV3	11.13 × 10 ⁶	183.93	81.75	73.3
RegNet	31.89 × 10 ⁶	65.05	81.77	73.5
ResNet101	58.75 × 10 ⁶	43.14	84.21	77.1
ResNet50	39.75 × 10 ⁶	60.21	81.94	74.9

综上所述, 结合智能农机装备计算资源有限的特点, 所建模型需保持一定分割精度及兼顾设备计算资源的限制, 而 G_RegNetX4 主干网络在具有优良的分割效果的同时, 还有参数总量少和推理速度快的优点.

3.5 模型性能试验

选用 DeepLabV3^[24]、FCN、PSPNet^[25] 模型与本文模型进行对比测试, 通过 PA、MPA、MIoU、Params、FLOPs 和 FPS 等指标验证模型性能. 试验模型均加载主干网络预训练权重, 并在丘陵田间道路训练集上训练. 在测试集上进行不同网络模型测试, 并计算相关指标. 表 3 记录了不同语义分割模型的 Params、FPS 及 FLOPs. 从表 3 可以看出本文模型的 Params 和 FLOPs 分别为 14.71 × 10⁶ 和 49.34 × 10⁹, 相比于对比模型中较优秀的 FCN 模型, 其 Params 和 FLOPs 分别减少了 18.24 × 10⁶ 和 72.74 × 10⁹, G_RegNetX4 作为主干网络, 在特征提取过程中注意到相似冗余特征的存在, 使用廉价变化将部分低级特

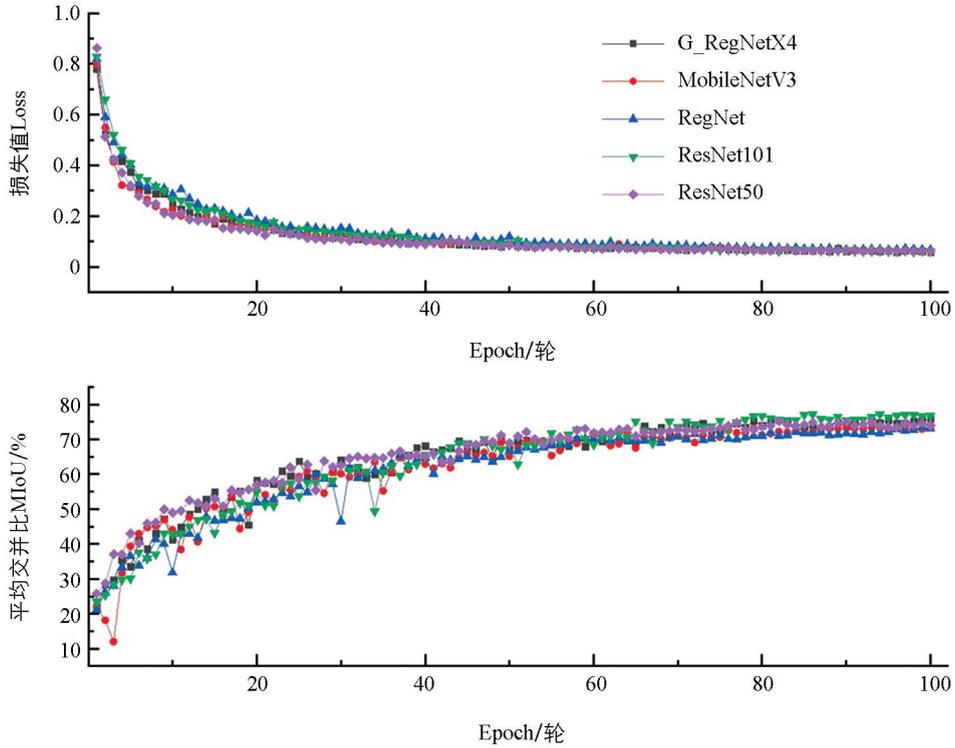


图 7 不同主干网络的训练效果曲线

征映射到高级特征从而减少模型的计算量. 在推理分割图片时, 本文模型的 FPS 为 116.08 f/s, 分别为 DeepLabV3、FCN、PSPNet 模型的 1.83, 1.33, 1.76 倍. 本文模型优势突出的原因是分割头为 Lite-RASPP, 聚合了不同尺度特征获取全局有效信息, 同时精简网络结构从而提高网络推理速度.

表 3 不同语义分割模型性能对比

网络模型	参数总量 Params	推理速度 FPS/(f · s ⁻¹)	浮点计算量 FLOPs
DeepLabV3	39.23 × 10 ⁶	63.23	144.26 × 10 ⁹
FCN	32.95 × 10 ⁶	86.71	122.08 × 10 ⁹
PSPnet	46.58 × 10 ⁶	65.93	156.13 × 10 ⁹
本文模型	14.71 × 10 ⁶	116.08	49.34 × 10 ⁹

分析图 8 可知本文模型收敛速度快于其他模型, 其中 FCN 和 DeepLabV3 在训练过程中收敛较慢且波动较大. 在 60 轮 Epoch 后各个网络均趋于稳定收敛, 但本文模型的 MPA 与 MIoU 都高于其他模型.

不同网络在测试集上的 PA、MPA 及 MIoU 由表 4 可以看出, 本文提出模型的 PA、MPA 和 MIoU 都要优于其他语义分割模型. 本文模型的 MPA 和 MIoU 分别为 93.24% 和 87.6%, 相比于 DeepLabV3 分别提高了 1.8% 和 2.2%, 比 FCN 分别提高了 1.46% 和 0.8%, 比 PSPNet 分别提高了 1.48% 和 1%. 本文模型的主干网络采用 G_RegNetX4 是分割精度提升的主要原因, 在 G_RegNetX4 的网络结构中结合残差模块部署 G_Ghost 模块, 将复杂特征与低级特征融合从而提高分割精度. 图 9 为本文模型在测试集上的分割效果对比.

表 4 不同语义分割模型的评价指标对比

网络模型	像素准确率 PA/%	平均像素准确率 MPA/%	平均交并比 MIoU/%
DeepLabV3	97.7	91.44	85.4
FCN	97.7	91.78	86.8
PSPnet	97.8	91.76	86.6
本文模型	97.9	93.24	87.6

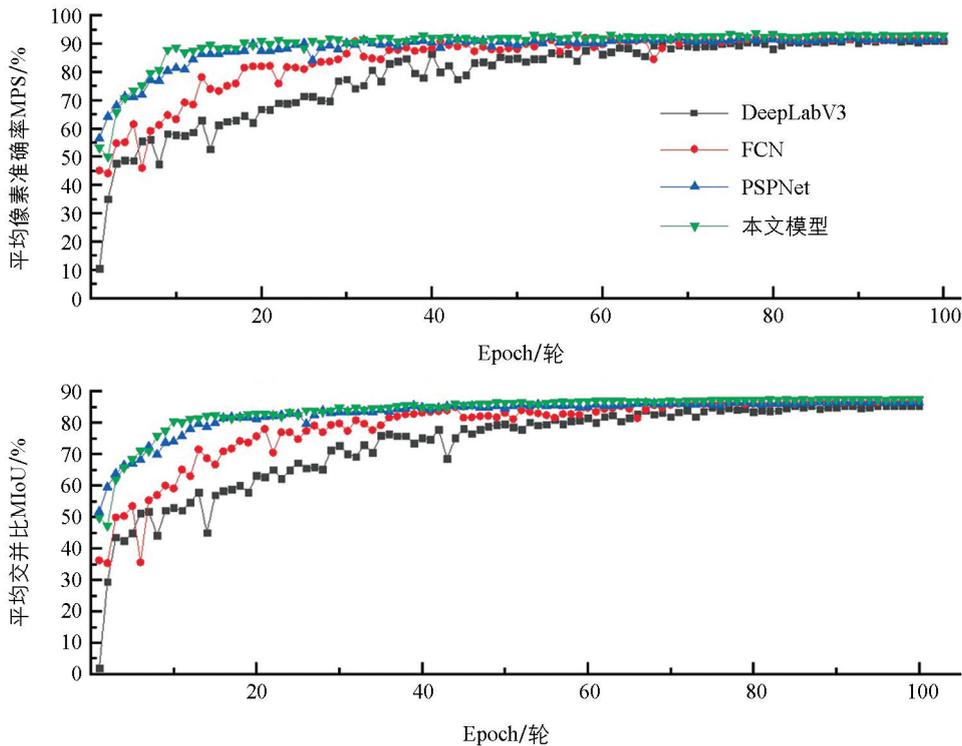


图 8 不同模型的训练效果曲线

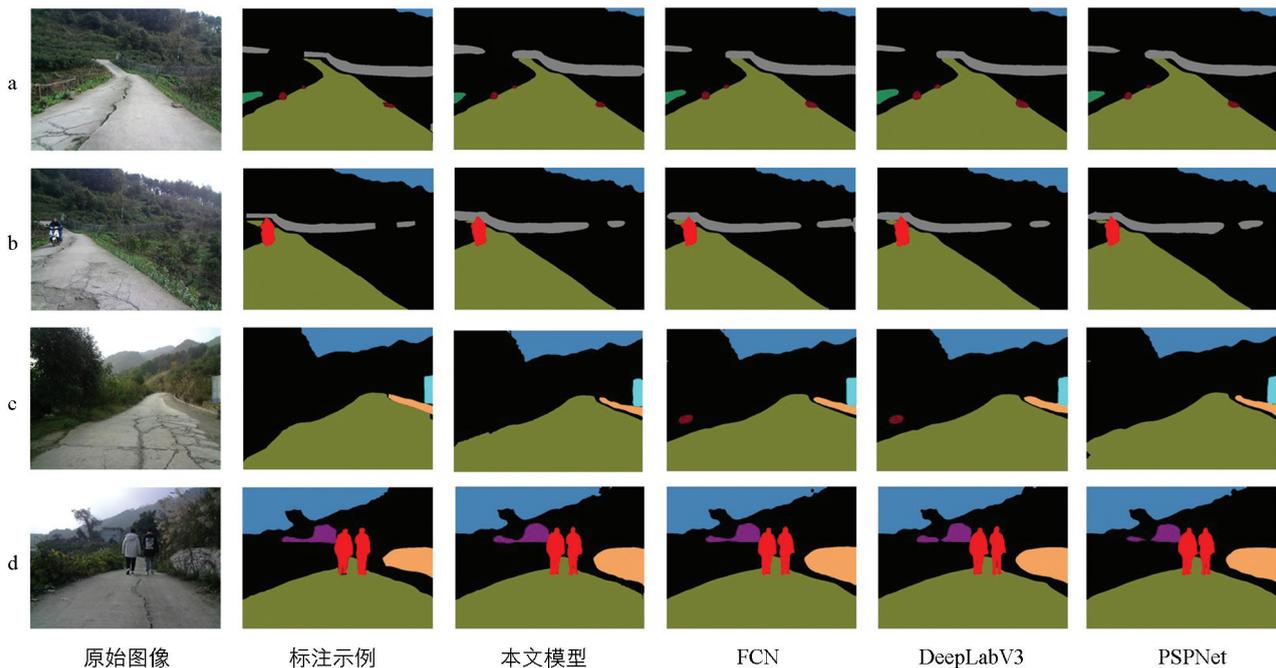


图 9 不同模型分割效果对比

从图 9 可以看出, 本文模型对丘陵田间道路中的动态目标、道路、障碍物等类别可以进行有效分割. 语义分割模型 FCN 和 DeepLabV3 在对障碍物类分割时出现分割错误, 对图 9(c) 预测分割时将道路两旁的物体预测为道路中的障碍物. PSPNet 存在过分割的情况, 处理物体边缘和细节复杂区域时易受到背景信息的干扰导致分割错误, 如将图 9(d) 中动态目标的边界分割错误. 本文模型使用的 Lite-RASPP 模块减少了模型的参数量, 但在推理分割过程中影响了部分类别的边界分割精度导致边界不平滑, 如图 9(c) 中本文

模型对挡墙边界的分割结果是波动的。

综合对比 Params、FPS、FLOPs、MIoU、MPA 及实际分割效果可知,所提出针对丘陵田间道路的语义分割模型相比于主流的语义分割模型在准确率相关指标的对比中是较优秀的;同时所提模型在 Params、FPS 及 FLOPs 的比较中也有大幅度的提升。

4 结论

提出一种基于优化 DeepLabV3+ 的丘陵田间道路图像语义分割模型,编码器环节采用轻量化的主干网络(G_RegNetX4)和轻量级的 Lite-RASPP 模块,在明显提升分割精度和推理速度的同时降低了参数总量和浮点计算量。通过采集丘陵田间道路图像建立数据集并将对象属性划分为 10 个类别,构建测试集以进行模型分割效果测试。结果表明所提出模型的平均交并比和平均像素准确率分别为 87.6%, 93.24%, 具有准确率高及泛化性能好的优点。将本文模型与 FCN、DeepLabV3 和 PSPNet 进行比较,发现其可以有效对丘陵田间道路进行分割,较好地实现了精度与速度的平衡。考虑到未对分割后的田间道路场景图像提取视觉导航线,后续将进行相关研究以期更好地实现智能农机装备的自主导航。

参考文献:

- [1] 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.
- [2] WANG G J, WU J, HE R, et al. Speed and Accuracy Tradeoff for LiDAR Data Based Road Boundary Detection [J]. IEEE/CAA Journal of Automatica Sinica, 2021, 8(6): 1210-1220.
- [3] LU X W, AI Y F, TIAN B. Real-Time Mine Road Boundary Detection and Tracking for Autonomous Truck [J]. Sensors, 2020, 20(4): 1121.
- [4] 王小娟,李云伍,刘得雄,等. 基于机器视觉的丘陵山区田间道路虚拟中线提取方法 [J]. 西南大学学报(自然科学版), 2018, 40(4): 162-169.
- [5] 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.
- [6] 杜小强,李卓林,马锶宏,等. 基于空间注意力和可变形卷积的无人机田间障碍物检测 [J]. 农业机械学报, 2023, 54(2): 275-283.
- [7] XU F L, ZHAO H K, HU F Y, et al. A Road Segmentation Model Based on Mixture of the Convolutional Neural Network and the Transformer Network [J]. Computer Modeling in Engineering & Sciences, 2023, 135(2): 1559-1570.
- [8] ZHANG Y H, XIA G H, WANG J G, et al. A Multiple Feature Fully Convolutional Network for Road Extraction from High-Resolution Remote Sensing Image Over Mountainous Areas [J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(10): 1600-1604.
- [9] WANG H, CHEN Y Y, CAI Y F, et al. SFNet-N: An Improved SFNet Algorithm for Semantic Segmentation of Low-Light Autonomous Driving Road Scenes [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11): 21405-21417.
- [10] FIRKAT E, ZHANG J L, WU D F, et al. ARDformer: Agroforestry Road Detection for Autonomous Driving Using Hierarchical Transformer [J]. Sensors, 2022, 22(13): 4696.
- [11] LIU B S, LV Y B, GU Y, et al. Implementation of a Lightweight Semantic Segmentation Algorithm in Road Obstacle Detection [J]. Sensors, 2020, 20(24): 7089.
- [12] TIAN F Y, HU G Z, YU S F, et al. An Efficient Multi-Task Convolutional Neural Network for Dairy Farm Object Detection and Segmentation [J]. Computers and Electronics in Agriculture, 2023, 211: 108000.

- [13] 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.
- [14] SHANG G G, LIU G, ZHU P, et al. A Deep Residual U-Type Network for Semantic Segmentation of Orchard Environments [J]. *Applied Sciences*, 2021, 11(1): 322.
- [15] LIN N C, ZHAO W F, LIANG S H, et al. Real-Time Segmentation of Unstructured Environments by Combining Domain Generalization and Attention Mechanisms [J]. *Sensors*, 2023, 23(13): 6008.
- [16] BAHETI B, INNANI S, GAJRE S, et al. Semantic Scene Segmentation in Unstructured Environment with Modified DeepLabV3+ [J]. *Pattern Recognition Letters*, 2020, 138: 223-229.
- [17] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation [C] //European Conference on Computer Vision(ECCV 2018). Munich, Germany. Cham: Springer, 2018: 833-851.
- [18] HAN K, WANG Y H, XU C, et al. GhostNets on Heterogeneous Devices via Cheap Operations [J]. *International Journal of Computer Vision*, 2022, 130(4): 1050-1069.
- [19] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C] //2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South). IEEE, 2019: 1314-1324.
- [20] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C] //2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [21] SHARMA S, BALL J E, TANG B, et al. Semantic Segmentation with Transfer Learning for Off-Road Autonomous Driving [J]. *Sensors*, 2019, 19(11): 2577.
- [22] RADOSAVOVIC I, KOSARAJU R P, GIRSHICK R, et al. Designing Network Design Spaces [C] //2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA. IEEE, 2020: 10425-10433.
- [23] HE K M, ZHANG X Y, REN S Q, et al. Deep Residual Learning for Image Recognition [C] //2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [24] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking Atrous Convolution for Semantic Image Segmentation [EB/OL]. (2017-12-05) [2023-08-15]. <http://arxiv.org/abs/1706.05587>.
- [25] ZHAO H S, SHI J P, QI X J, et al. Pyramid Scene Parsing Network [C] //2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 6230-6239.

责任编辑 柳剑

崔玉洁