

方向感知金字塔聚合网络的道路中心线提取

张小青, 王青旺, 瞿信, 沈世全, 吴长义, 刘菊

引用本文

张小青, 王青旺, 瞿信, 沈世全, 吴长义, 刘菊. 方向感知金字塔聚合网络的道路中心线提取[J]. 计算机科学, 2024, 51(6A): 230400101-7.

ZHANG Xiaoqing, WANG Qingwang, QU Xin, SHEN Shiquan, WU Changyi, LIU Ju. [Direction-aware Pyramidal Aggregation Network for Road Centerline Extraction](#) [J]. Computer Science, 2024, 51(6A): 230400101-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[融合注意力机制的DeeplabV3+ 服装图像分割方法](#)

Clothing Image Segmentation Method Based on Deeplabv3+ Fused with Attention Mechanism
计算机科学, 2024, 51(6A): 230900153-7. <https://doi.org/10.11896/jsjcx.230900153>

[基于双流卷积神经网络的稻米缺陷分割](#)

Rice Defect Segmentation Based on Dual-stream Convolutional Neural Networks
计算机科学, 2024, 51(6A): 230600107-8. <https://doi.org/10.11896/jsjcx.230600107>

[基于注意力机制与密集邻域预测的轻量化图像语义分割](#)

Lightweight Image Semantic Segmentation Based on Attention Mechanism and Densely AdjacentPrediction
计算机科学, 2024, 51(6A): 230300204-8. <https://doi.org/10.11896/jsjcx.230300204>

[基于多尺度特征的地基云图分类检测算法](#)

Classification and Detection Algorithm of Ground-based Cloud Images Based on Multi-scale Features
计算机科学, 2024, 51(6A): 230400041-6. <https://doi.org/10.11896/jsjcx.230400041>

[基于图卷积神经网络的点云语义分割综述](#)

Review of Point Cloud Semantic Segmentation Based on Graph Convolutional Neural Networks
计算机科学, 2024, 51(6A): 230400196-7. <https://doi.org/10.11896/jsjcx.230400196>

方向感知金字塔聚合网络的道路中心线提取

张小青¹ 王青旺¹ 瞿信¹ 沈世全² 吴长义³ 刘菊⁴

1 昆明理工大学信息工程与自动化学院 昆明 650504

2 昆明理工大学交通工程学院 昆明 650504

3 昭通市建筑工程质量监督站 云南 昭通 657099

4 用友网络科技股份有限公司流程制造开发四部 成都 610095

(zhangxiaqing@stu.kust.edu.cn)

摘要 道路中心线作为抽象类,无明确显性特征,进而造成模型无法准确提取道路中心线。针对该问题,文中将道路中心线提取建模为语义分割任务,并根据道路中心线的线性结构特点提出了一种方向感知金字塔聚合网络(Direction-aware Pyramidal Aggregation Network, DAPANet)。首先,针对道路中心线的空间分布特性及结构特点,设计了方向感知模块(Direction-aware Module, DAM),在主干网络(ResNet18)最后输出的4个层级上分别使用4个方向感知层提取道路中心线的方向特征。然后,进一步设计融合多向性特征的金字塔聚合模块(Pyramid Aggregation Module, PAM),融合4个层级提取到的结构特征,得到更具有鲁棒性的道路中心线特征。最后,在无人机平台下采集的真实数据上进行了实验,实验结果显示所提出的 DAPANet 取得了 84.7% 的 mIoU 和 98.6% 的 Precision,道路中心线的 IoU 达到 77.28%,性能优于其他先进的对比方法,证明了所提方法对提取道路中心线的有效性。

关键词: 方向感知;金字塔聚合;语义分割;道路中心线提取

中图分类号 TP389

Direction-aware Pyramidal Aggregation Network for Road Centerline Extraction

ZHANG Xiaoqing¹, WANG Qingwang¹, QU Xin¹, SHEN Shiquan², WU Changyi³ and LIU Ju⁴

1 School of Information Engineering and Automation, Kunming University of Technology, Kunming 650504, China

2 School of Traffic Engineering, Kunming University of Technology, Kunming 650504, China

3 Zhaotong City Construction Quality Supervision Station, Zhaotong, Yunnan 657099, China

4 Process Manufacturing Development IV, Yonyou Network Technology Co., Chengdu 610095, China

Abstract As an abstract class, road centerlines have no explicit features, which in turn causes the model fail to extract road centerlines accurately. To address this problem, this paper models road centerline extraction as a semantic segmentation task, and proposes a direction-aware pyramidal aggregation network(DAPANet) based on the spatial linear structure of road centerlines. Firstly, for the spatial distribution characteristics and structural features of road centerlines, this paper designs the direction-aware module(DAM) to extract the features of road centerlines using four direction-aware layers on each of the four layers of the final output of the backbone network(ResNet18). Then, it further designs the pyramid aggregation module(PAM) to fuse the structural features extracted from the four layers to obtain a more robust road centerline feature. Experiments are conducted on real data collected under the UAV platform, and the experimental results show that the proposed DAPANet achieves 84.7% of mIoU and 98.6% of Precision, in which the IoU of road centerline reaches 77.28%, outperforming other advanced comparative methods and proving the effectiveness of the proposed method.

Keywords Direction-awareness, Pyramid aggregation, Semantic segmentation, Road centerline extraction

1 引言

传统道路巡检主要依靠专业人员通过实地考察的方式开展,效率低且耗费大量人力物力。随后,车辆半自动巡检系统出现,降低了部分人工成本,但车辆巡检易受路况影响,造成巡检效率低下^[1-2]。近年来,随着无人机技术的快速发展,

低成本、高续航、高稳定性的无人机不断涌现,无人机已在地质勘探^[3]、管道巡视^[4]、道路养护^[5]和森林巡查^[6]等各个领域广泛的应用。因此,相比传统的人工巡检和车辆巡检,采用更高效、更快速、更安全以及更低成本的无人机进行道路巡检,能提升巡检效率,推动数字化建设,实现智能自动化。

无人机道路巡检指无人机通过搭载高清摄像头、红外

基金项目:国家自然科学基金青年科学基金(62201237);云南省基础研究计划项目(202101BE070001-008)

This work was supported by the Young Scientists Fund of the National Natural Science Foundation of China(62201237) and Yunnan Fundamental Research Projects(202101BE070001-008).

通信作者:王青旺(wangqingwang@kust.edu.cn)

相机、激光雷达等设备,对道路的路面状态、交通流量、道路设施等情况进行全面的检测和记录。但是,当前无人机道路巡检严重依赖于人类飞手的控制,造成无人机无法实现自主巡检。道路中心线能为无人机自主巡检提供路线参考,因此准确提取道路中心线是实现无人机自主巡检的首要前提。

道路中心线提取的研究主要分为基于传统机器学习的方法和基于深度学习的方法。但绝大多数研究都从空间分辨率有限、成像视角存在差异的高空遥感图像中提取道路中心线,并不能为低空无人机道路巡检提供高分辨率的近地视野,使得传统方法不能准确提取道路中心线,从而导致无人机自主道路巡检任务开展困难^[7-9]。而语义分割是通过像素级语义分类,来实现高分辨率图像的准确分割,已被广泛应用于自动驾驶领域^[10-12]。因此,针对微型无人机设备,考虑将道路提取建模为语义分割任务,并设计一种轻量级、实时性算法,为无人机道路巡检提供路线参考,从而实现无人机自主道路巡检是当前研究中亟待解决的问题。

针对道路中心线是一种抽象类,无明确显性特征,导致传统语义分割模型效果不理想的问题,本文提出了方向感知金字塔聚合网络(DAPANet)。具体地,本文首先设计了方向感知模块(DAM),通过将主干网络(ResNet18)最后4个层级提取到的特征作为方向感知模块的输入,并分别使用4个方向感知层提取道路中心线特征。其中,每个方向感知层均采用4个方向(水平、垂直、左对角线、右对角线)的条带卷积操作,有效地提取到道路中心线的空间线性结构特征。随后,本文进一步设计了融合多方向性特征的金字塔聚合模块(PAM),融合了道路中心线的多尺度线性结构特征,提高了模型对道路中心线特征的提取能力,得到更具鲁棒性的道路中心线。

2 相关工作

2.1 道路中心线提取

从高空遥感图像中提取道路中心线,部分研究人员已开展大量研究。其中,部分学者通过对分类图像使用张量投票的方式,提出了一种半自动化的方法来获取道路中心线^[7]。随后有学者根据道路形状提取道路段,进一步使用局部线性核平滑回归^[13],从卫星图像中提取城市主干道路中心线。此外,部分学者提出了一种从模型驱动到样本驱动农村地区道路提取方法^[14],并通过线条序列的长度约束和平均对比度约束建立道路中心点的提取模型,从而实现道路跟踪。

随着深度学习技术的不断发展与推广,2017年,部分研究人员根据卷积神经网络(Convolutional Neural Networks, CNN)结构设计了一个级联CNN模型^[8],可以同时提取道路区域和道路中心线。2019年,部分学者引入基于深度学习的方法以实现道路中心线鲁棒性特征的自动提取,并提出了一种以多尺度Gabor滤波器和多方向非极大值抑制为基础的的道路中心线提取方法^[15]。紧接着,有学者提出了一个Deep-Window模型^[16],通过采用CNN决策函数引导的滑动窗口,解决了复杂场景下提取道路中心线偏离道路中心的问题。2021年,有研究人员注意到大多数现有研究均依赖于分别处理道路中心线提取和路面提取的工作,因此考虑了联合提取,并提出了一种具有多任务策略的MRENet^[9]来同时预测路面

和道路中心线。

虽然,目前研究人员已实现在传统高空遥感图像中提取道路中心线,但缺乏低空无人机视角图像中提取道路中心线方面的研究,导致无人机自主巡检任务开展困难。因此,面向低空无人机道路巡检设计了一种道路中心线提取算法,为无人机道路巡检提供路线参考,从而实现无人机自主巡检是当前亟待解决的问题。

2.2 语义分割方法

由于道路中心线抽象、无明确显性特征,传统目标检测方法难以生成连续性曲线,曲线拟合的方法对关键点确定过于敏感,而采用语义分割的方法能有效直接提取整条具有连续性的道路中心线。因此,本文采用语义分割的方式提取道路中心线,且为了将其应用于具有微型计算能力的无人机设备,应采用轻量级、实时性的语义分割方法。

目前,已经有许多研究人员开展大量关于轻量级、实时性语义分割方面的研究^[17-19]。2018年,面向轻量级、实时性语义分割算法研究,有学者提出双边分割网络BiSeNet^[10],保留了空间分辨率的同时提升了推理速度,满足了对实时性的需求。随后,部分研究人员在BiSeNet基础上进行了改进,提出了双分支网络BiSeNet-V2^[20],合理利用了低层次细节信息及高层次语义信息。进一步,有学者在BiSeNet的基础上提出了STDCNet^[21],去除了冗余结构,加快了推理速度。针对减少像素级标签推断的大量计算方面存在的困难,部分学者提出了一种多分支级联网络(ICNet)^[11],以快速实现高质量的分割。此后,部分学者专门为实时语义分割设计了深度双分辨率网络DDRNet^[22],减少了计算量,提升了效果。针对双分支网络中目标细节信息被淹没的问题,部分学者提出了三支网络PIDNet^[23],提升了模型的分割准确度。针对编码器-解码器结构很难为像素级分割任务提取足够背景信息的问题,一些研究人员提出了子空间金字塔融合网络SPFNet^[24]。部分研究员提出了一个将“注意力金字塔网络(APN)”作为解码器的轻量级非对称编码器-解码器结构(LEDNet)^[12]。针对轻量级网络由于减少参数导致分割准确性降低的问题,有学者提出轻量级多尺度特征融合网络LM-FFNet^[25],平衡了模型参数和分割精度。随后,部分研究人员在特征金字塔的基础上,引入了尺度感知的带状注意力和通道注意力模块,并设计了用于语义分割的架构SSFPN^[26],在自动驾驶领域实现了精度和推理速度的平衡。综上所述,虽然当前已有许多研究人员在轻量级、实时性语义分割算法方面展开大量研究,但以上绝大多数研究面向的是自动驾驶场景,缺乏低空无人机自主巡检提取道路中心线的相关研究。

因此,针对微型无人机设备,设计了一种满足轻量级、实时性需求的语义分割算法,用于提取道路中心线,从而实现无人机自主巡检,对实际的生产生活具有重要的研究价值。

3 方向感知金字塔聚合网络

针对道路中心线的线性空间分布特性及其结构特征,本文设计了一个方向感知金字塔聚合网络(DAPANet),如图1所示。

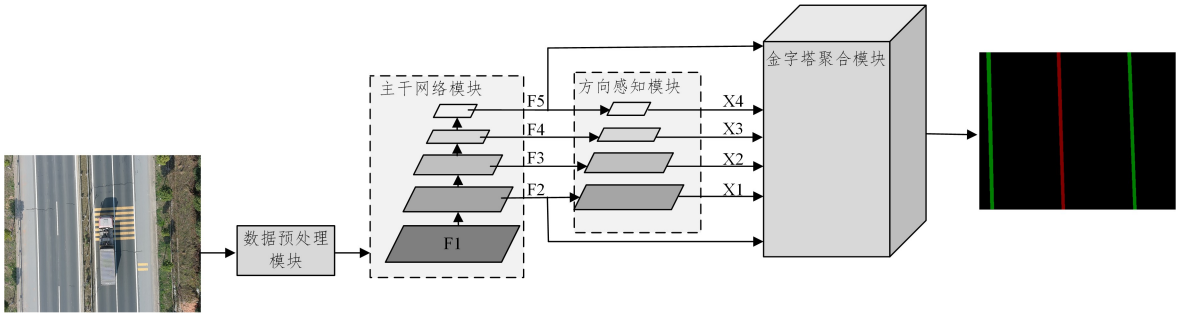


图1 DAPANet的整体结构

Fig. 1 Overall structure of DAPANet

DAPANet的总体结构主要包括:主干网络模块(Backbone)、方向感知模块(DAM)和金字塔聚合模块(PAM)。其中,以去掉自适应平均池化层和全连接层的ResNet18为主干网络,并输出最后4层提取的特征{F2,F3,F4,F5}。

3.1 方向感知模块

方向感知模块包含4个不同尺度的方向感知层。将主干网络提取的特征{F2,F3,F4,F5}分别作为4个方向感知层的输入,对应的通道数分别为{64,128,256,512},对应特征图尺寸分别为{1/4,1/8,1/16,1/32}。受一维转置卷积^[27]的启发,条状卷积相比方形网格卷积更轻量,且更有利于提取条状特征。因此,为了提取道路中心线的空间线性结构特征,每个尺度的方向感知层均通过4个方向的条带卷积操作(Strip Convolution)^[28]。其中,不同尺度方向感知层对应条带卷积核的大小不同,分别为{3,5,7,9},步长均默认为1,为了保持图像尺寸不变,每个卷积核对应的填充项值分别为{1,2,3,4}。方向感知层的结构如图2所示,具体操作步骤如下。

步骤1 降维。采用 1×1 的卷积块(包含 1×1 的卷积操作、批量归一化操作和非线性激活操作)降低输入特征 $C \times H \times W$ (C 表示通道数、 H 表示图像高度、 W 表示图像宽度)的通道数,将其压缩成 $C2 \times H \times W$ 。

步骤2 条带卷积。将步骤1后大小为 $C2 \times H \times W$ 的输出特征图同时进行4个方向(水平方向、垂直方向、正对角线方向和反对角线方向)的条带卷积操作。其中,为了便于两个对角线方向的条带卷积计算,输入特征图应保持高度和宽度相等,而输入特征图大小为 $H \times W$,高度和宽度不同。因此,在通过两个对角线方向的条带卷积操作之前,采用双线性插值的方式将大小为 $H \times W$ 的特征图转化成大小为 $W \times W$ 的特征图,并进行条带卷积操作后,同样以双线性插值的方式将特征图的尺寸恢复到原输入特征图的大小 $H \times W$ 。最后,将经过4个方向条带卷积后的特征图以拼接的方式级联起来。

设 $\mathbf{I} \in \mathbb{R}^{H \times W \times C2}$ 表示方向感知层的输入特征向量经过 1×1 的卷积块后的输出向量,其中 $H, W, C2$ 分别表示输入特征的高度、宽度和通道数。若 $K \in \mathbb{R}^{2k+1}$ 是大小为 $2k+1$ 的条带卷积滤波器, $SC_D \in \mathbb{R}^{H \times W \times C'}$ 表示滤波器 K 在方向 $D = (D_h, D_w)$ 的条带卷积输出结果,则计算式如下:

$$SC_D[i, j] = (\mathbf{I} * K)D[i, j] \\ = \sum_{u=-k}^k \mathbf{I}[i + D_h u, j + D_w u] \cdot K[k - u] \quad (1)$$

其中, $\mathbf{I} * K$ 是条带卷积运算, D 是条带卷积方向向量,对应于水平、垂直、正对角线和反对角线方向的值分别为(0,1),(1,0),(1,1)和(-1,1)。

当最后两个方向感知层经过条带卷积操作后,采用双

线性插值的方式进行上采样,以增大感受野,提升特征图的尺寸。此时,4个方向感知层得到的特征图大小为原始输入特征图尺寸的{1/4,1/8,1/8,1/16}。

步骤3 池化。将拼接后的特征图经过平均池化层(包含池化操作、批量归一化操作和非线性激活操作,下同),从而压缩数据,减小过拟合。其中,每层池化核的大小及填充项的大小与该层条带卷积核大小和填充项保持一致,分别为{3,5,7,9}和{1,2,3,4}。

步骤4 降维。池化后特征图经过 1×1 的卷积块进行降维,使得4个方向感知层的通道数保持一致,以便进行后续的融合操作。

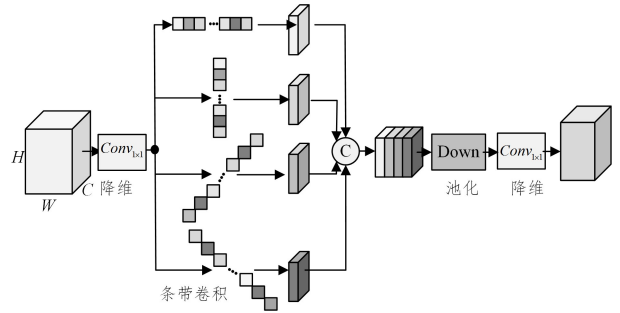


图2 方向感知层

Fig. 2 Direction-aware layer

3.2 金字塔聚合模块

为了更好地融合多层级提取到的道路中心线的空间线性结构特征,得到了更具鲁棒性的道路中心线,本节设计了金字塔聚合模块(PAM),如图3所示。输入包括经过4个方向感知层的输出特征和主干网络的两层输出特征。其中,{F2}和{F5}对应主干网络提取特征后输出的第一层与最后一层特征,{X1,X2,X3,X4}分别对应4个方向感知层的输出,{Y}对应金字塔聚合模块的输出,操作流程如下。

步骤1 降维。首先,将主干网络输出的第一层特征{F2}作为 1×1 的卷积块的输入,得到特征图{Y0}。其次,将方向感知模块输出的第一层特征{X1}经过 1×1 的卷积块后与{Y0}通道拼接得到特征{Y1}。此时,{Y0}与{Y1}具有相同的尺寸和通道数。

步骤2 特征上采样。首先,对剩下3个方向感知层的输出特征{X2,X3,X4},分别使用不同采样率的轻量级上采样算子(Content-Aware ReAssembly of FEatures, CARAFE)^[29]进行上采样。其中,特征{X4}经过上采样后扩大两倍,{X2}和{X3}经过上采样后扩大一倍,且得到的输出特征尺寸与{Y0}保持一致。最后,3个输出特征分别与特征图

{Y0}拼接,得到输出特征{Y2,Y3,Y4}。

步骤3 全局平均池化。首先,为了获得丰富的全局上下文信息,将主干网络最后一层输出的特征{F5}通过全局平均池化模块(Global Averaging Pooling,GAP)。该模块包含自适应平均池化和 1×1 的卷积块,并采用双线性插值法,得到输出特征的尺寸与{Y0}保持一致。其次,将输出特征与特征图{Y0}进行通道拼接,得到输出特征{Y5}。

设步骤1-步骤3分别与{Y0}进行通道拼接后得到的输出特征图记为{Yi}(其中, $i\in[0,5]$),其计算式如下:

$$Y_i = \begin{cases} Conv_{1\times 1}(F_2), & i=0 \\ Cat(Conv_{1\times 1}(X_1),Y_0), & i=1 \\ Cat(CARAFE(X_i),Y_0), & 2\leq i\leq 4 \\ Cat(Global(F_5),Y_0) & i=5 \end{cases} \quad (2)$$

其中, $Conv_{1\times 1}$ 表示 1×1 的卷积块;CARAFE表示特征上采样操作; $Global$ 表示全局平均池化操作; Cat 表示通道拼接操作。

步骤4 多尺度特征融合。首先,将步骤1-步骤3得到的输出特征{Y1,Y2,Y3,Y4,Y5}元素级相加;其次,经过一个 3×3 的卷积块,再与{Y0}进行通道拼接;最后,经过 1×1 的卷积块降低维度,并与{Y0}进行元素级相加,实现局部特征与全局特征的融合,得到融合后的输出特征{Y},其表达式如下:

$$Y = Conv_{1\times 1}(Cat(Conv_{3\times 3}(\sum_{i=1}^5 Y_i),Y_0)) + Y_0 \quad (3)$$

其中, $Conv_{1\times 1}$ 表示 1×1 的卷积块; $Conv_{3\times 3}$ 表示 3×3 的卷积块; Cat 表示特征拼接操作。

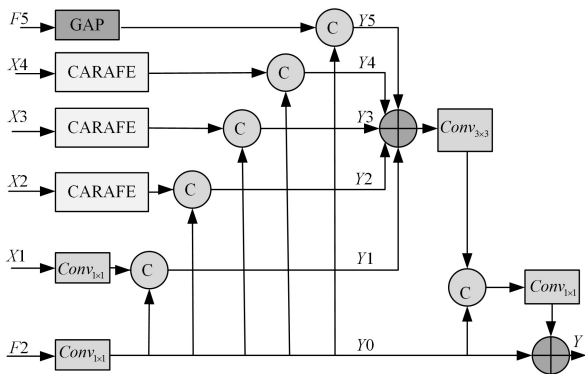


图3 金字塔聚合模块

Fig. 3 Pyramid aggregation module

3.3 损失函数

本文所采用数据集的样本表现为极度不平衡,因此本文实验的损失函数采用 focal loss^[30],能将高置信度样本的损失

降低,低置信度样本的损失提高,损失函数 \mathcal{L} 的计算式如下:

$$\mathcal{L} = -\partial_c (1 - p_c)^\gamma \log(p_c) \quad (4)$$

其中, ∂_c 表示第 c 类样本的权重; p_c 表示 softmax 输出的第 c 类的概率值; γ 表示损失衰减程度,其值越大表示对权重的调节程度越大,反之越小,通常 γ 设置为2。当 $\gamma=0$ 时,损失函数退化成原始交叉熵损失。因此,本文损失函数的参数 γ 设置为2。

4 实验

本文实验所用操作系统及其版本为 Ubuntu18.04,CPU 处理器型号为 Intel(R) Core(TM) i5-9400F CPU@2.90 GHz,显卡型号为 NVIDIA GeForce RTX 2060,显存大小是6GB,内存大小是16GB,CUDA版本是10.2,CuDNN版本是7.6.5。本文实验采用了 PyTorch 深度学习框架以及 OpenCV 图像处理库。优化器选择随机梯度下降法(Stochastic Gradient Descent,SGD),其动量参数设置为0.9,权重衰减设置为0.0005,学习率为0.001。本文实验采用“poly”的学习率策略,其学习率的计算式如下:

$$lr = lr_0 \times \left(1 - \frac{iter}{iter_{max}}\right)^{power} \quad (5)$$

其中, lr_0 是初始学习率,本文设置为0.025,指数 $power$ 是一个控制参数,设置为0.9,最大迭代次数 $iter_{max}$ 设置为300。本文实验训练轮次(epoch)设为300,批量大小(batchsize)设为12。

本文实验采用的数据集是在无人机平台下采集的真实数据,涵盖了22个不同的道路场景,包括高速柏油公路、乡村黄土小路、城市水泥道路等。本文按照7:1:2的比例将4587张大小为 512×640 的道路中心线语义分割数据集随机打乱后划分为训练集、验证集和测试集,对应有3212张、459张和916张。其中,该数据集包含背景(Background)、道路中心线(Centerline)和边缘线(Edgeline)3个类别,像素比约为60:1:2。

考虑到实验数据集中3个类别的像素差异大,因此将 ∂_c 超参数设置为0.01,0.66和0.33。

4.1 对比实验

为了进行相对公平的比较,以验证本文方法的有效性,本文在真实场景道路中心线数据集上开展对比实验,并与其他8种较先进的实时语义分割方法在参数和性能方面进行比较,包括参数量(Parameters)、浮点计算量(FLOPs)、帧率(FPS)、平均精度(Precision)以及平均交并比(mIoU)等5个方面的实验结果。

表1 不同方法的参数及性能

Table 1 Parameters and performance of different methods

方法	Parameters	FLOPs	FPS/(frame/ms)	Precision/%	mIoU/%
BiSeNet[13]	12.42×10^6	15.23×10^6	100.83	96.7	73.9
BiSeNetV2[20]	3.62×10^6	7.53×10^6	88.89	97.1	75.5
LMFFNet[25]	1.34×10^6	10.32×10^6	61.25	96.7	73.66
STDCNet[21]	12.04×10^6	10.18×10^6	92.50	97.7	79.4
PIDNet[23]	28.76×10^6	28.96×10^6	53.59	97.5	74.3
SPFNet[24]	34.52×10^6	16.79×10^6	42.02	98.3	80.5
DDNet[22]	32.36×10^6	39.75×10^6	44.14	97.8	82.7
SSFNet[26]	30.51×10^6	18.22×10^6	52.12	98.2	82.4
DAPANet(本文方法)	18.01×10^6	13.31×10^6	52.56	98.6	84.7

从表 1 可以知道,本文提出的方向感知金字塔聚合网络的 Precision 值和 mIoU 值与其他方法相比达到了当前最优,precision 值高达 98.6%,mIoU 值高达 84.7%,说明本文所提 DAPANet 能有效提取道路中心线的空间线性结构特征。

与双边分割网络 BiSeNet 及其改进版的双分支网络 BiSeNetV2 和去除了冗余结构的 STDCNet 相比,本文方法的 mIoU 值分别提升了 10.8%,9.2%和 5.3%,说明本文方法更有效。

与 PIDNet,SPFNet,DDRNet 和 SSPFN 相比,mIoU 值提升均大于 2%,而且 Parameters 值和 FLOPs 值均最小,说明本文方法所需参数数量和计算量更少,从而证明了本文所提方法满足了对算法轻量级方面的需求。当 FPS 大于 30 帧每秒时表示满足实时性需求,而本文方法的 FPS 值仅低于 PIDNet,达到 52.26 帧每秒,说明本文方法在推理速度方面满足实时性需求。

此外,为了综合评估本文方法在参数及性能方面的效果,分别与参数数量和计算量最少以及推理速度最快的 3 种方法相比,结果如下。

表 2 不同方法的分割结果对比

Table 2 Comparison of segmentation results by different methods

(%)					
方法	Background	Centerline	Edgeline	Precision	mIoU
BiSeNet ^[13]	96.56	64.43	60.84	96.7	73.9
BiSeNetV2 ^[20]	96.91	65.23	64.50	97.1	75.5
LMFFNet ^[25]	96.56	63.00	61.44	96.7	73.6
STDCNet ^[21]	97.58	71.22	69.47	97.7	79.4
PIDNet ^[23]	97.37	68.08	57.52	97.5	74.3
SPFNet ^[24]	98.22	72.11	71.30	98.3	80.5
DDRNet ^[22]	98.22	76.00	73.75	97.8	82.7
SSFPN ^[26]	98.16	73.32	75.81	98.2	82.4
DAPANet(本文方法)	98.55	77.28	78.31	98.6	84.7

由表 2 可以看出,本文方法的平均精度 Precision 为 98.6%,平均交并比 mIoU 达到 84.7%。其中,背景、中心线以及边缘线 3 个类别的 IoU 值分别为 98.55%,77.28%及 78.31%,与其他 8 种先进方法相比,本文方法在背景类和边缘线类的分割结果明显高于其他 8 种对比方法,且道路中心线的 IoU 值相比其他方法均增加 1%以上,从而证明本文方法提高了模型对道路中心线特征的提取能力,能有效地提取到道路中心线的空间线性结构特征,得到了更具鲁棒性的道路中心线,证明了本文方法的有效性。

4.2 消融实验

为了进一步验证本文方法各个模块的有效性,本节设计了消融实验,并在道路中心线语义分割数据集上进行测试。具体地,在以 RestNet18 为主干网络(BackBone)提取特征的基础上,通过控制变量法验证以下 3 个模块的有效性,包括金字塔聚合模块(PAM)、方向感知模块(DAM)及特征上采样算子(CARAFE),消融实验结果如表 3 所列。表中第一行(除表头以外)表示仅使用主干网络 RestNet18 提取道路中心线的特征,同时将最后的全连接层改为 1×1 的卷积层(包含 1×1 的卷积操作、批量归一化操作和非线性激活操作),并采用简单的双线性插值法进行上采样,将图像恢复到原始输入图像的尺寸,作为 Baseline。其中,“√”表示实验中使用到的模块,“—”表示实验中未使用到的模块。本节根据消融实验

首先,相比 Parameters 值最小的 LMFFNet 方法,虽然本文提出的方法增加了 Parameters 值,但 mIoU 值远高于 LMFFNet 方法,增加了 10%以上,说明本文 DAPANet 虽然增加了部分参数量,降低了推理速度,但显著提升了分割精度,从而证明了本文方法在道路中心线提取方面的分割结果更准确。其次,与 FLOPs 最少的 BiSeNetV2 相比,Parameters 在保持一致的情况下,本文方法的 mIoU 远高于 BiSeNetV2,且 FPS 值更高,从而证明本文推理速度更快且更准确。最后,与 FPS 值最大的 BiSeNet 方法相比,本文方法虽然在 Parameters 上略高于该方法,但 FLOPs 更低,mIoU 提升 10%左右,从而证明本文方法的浮点计算量更少,且更准确。因此,本文与其他 8 种语义分割方法相比,在满足轻量级、实时性需求的同时分割结果更准确。

为了验证本文方法能有效提取道路中心线的空间线性结构特征,表 2 详细地列出了本文方法与 BiSeNet 等 8 种对比方法,在道路中心线语义分割数据集下每个类别的分割结果包括背景(Background)、道路中心线(Centerline)和边缘线(Edgeline)的交并比值(IoU)、平均精度值(Precision)和平均交并比值(mIoU),分割结果的对比如表 2 所列。

结果展开分析,结果如表 3 所列。

表 3 消融实验

Table 3 Ablation experiments

(%)					
Baseline	PAM	DAM	CARAFE	Precision	mIoU
√	—	—	—	96.7	73.7
√	√	—	—	97.3	77.3
√	√	√	—	98.5	83.8
√	√	√	√	98.6	84.7

由表 3 可知,仅通过 Baseline 进行测试得到 96.7%的 Precision 和 73.7%的 mIoU。首先,为了验证 PAM 的有效性,表中第二行将主干网络提取 4 个层级的特征通过 PAM,并将 PAM 中的 3 个 CARAFE 采用 3 个双线性插值法代替,得到 97.3%的 Precision 和 77.3%的 mIoU,从而证明了 PAM 能有效融合道路中心线的多尺度空间线性结构特征,显著提高了模型对道路中心线特征提取的能力,从而提取到更具鲁棒性的道路中心线。然后,为了验证 DAM 的有效性,表中第三行在第二行的基础上将主干网络输出的 4 个层级的特征通过 DAM,并分别对每个方向感知层使用 4 个方向(水平、垂直、左对角线、右对角线)的条带卷积操作,最终得到了 98.5%的 Precision 和 83.3%的 mIoU,从而证明了 DAM 能有效提取到道路中心线的空间线性结构特征。为了验证

CARAFE 的有效性,表中第四行在第三行的基础上将金字塔聚合模块中的 3 个双线性插值替换为 3 个特征上采样算子,最终得到了 98.6% 的 Precision 和 84.7% 的 mIoU,从而证明了 CARAFE 的有效性。

综上所述,通过消融实验,进一步证明了本文方法的各个模块在道路中心线提取方面有效。

4.3 可视化

为了能更直观地体现出本文所提方向感知金字塔聚合网络在提取道路中心线方面的优势,本节进行了可视化实验,如图 4 所示。

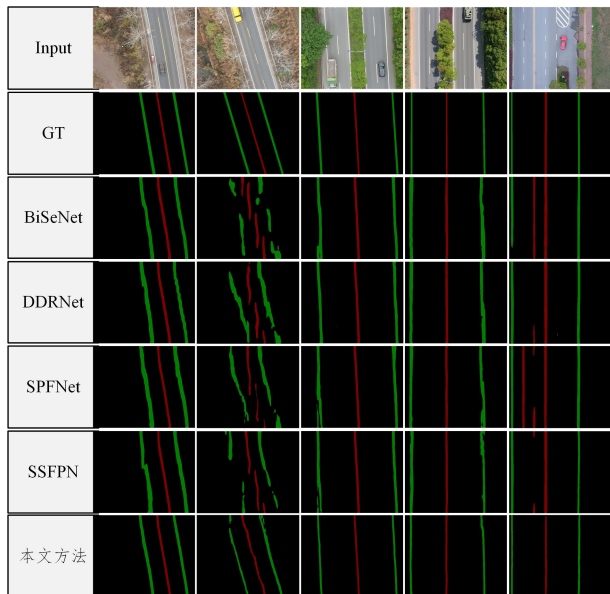


图 4 不同模型在道路中心线分割数据集上的可视化结果

Fig. 4 Visualisation results of different methods on road centerline segmentation dataset

图 4 中,第一行是原始图像,第二行是对应标签,第三行至第六行展示了对比方法的可视化结果,第七行是本文方法的可视化结果。前四列图像显示,BiSeNet 等 4 种对比方法提取到的道路中心线出现了不连续、有毛刺等现象,而本文 DAPANet 提取道路中心线更完整、更连续,说明了本文方法能更好地保持道路的空间线性结构,具有更好的鲁棒性,进而验证了本文方法的有效性。从最后一列图像的可视化结果可以看出,4 种对比方法均会出现误检的现象,而本文所提 DAPANet 能准确提取出道路中心线,进一步验证了本文方法的有效性。

结束语 本文提出了方向感知金字塔聚合网络,有效利用了道路中心线的空间线性结构,提高了模型对道路中心线特征的提取能力,得到了更具有鲁棒性的道路中心线特征。最后,本文在真实道路场景数据集上进行测试,证明了本文方法在提取道路中心线方面表现更好。然而,无人机自主巡检任务对轻量级、实时性及准确度方面的要求不断提升。因此,未来在保持高准确度分割结果的同时,思考如何进一步减少计算量和参数量,提高推理速度,是下一步需要研究的内容。

参考文献

[1] WU H B, QIU W M, WANG J, et al. Intelligent multi-track inspection robot inspection system for urban rail vehicles[J]. *Locomotive Electric Drive*, 2022(1): 45-52.

[2] ZHU J Y, LI P, LI X, et al. Intelligent inspection robot positioning system for urban rail vehicles[J]. *Railway Transport and Economy*, 2023, 45(1): 130-136.

[3] ZHAO X T, HU K, LU X P, et al. A method for fine detection of geological hazards in mines by low altitude aerial photography with unmanned aerial vehicles[J]. *Survey and Mapping Science*, 2014, 39(6): 49-52.

[4] PAN L, GUO T T. Intelligent inspection application of natural gas long-distance pipeline by drones[J]. *Chemical Management*, 2023, 653(2): 124-126.

[5] ZHU C, ZHU J, BU T, et al. Monitoring and Identification of Road Construction Safety Factors via UAV[J]. *Sensors*, 2022, 22(22): 8797.

[6] SHERSTJUK V, ZHARIKOVA M, SOKOL I. Forest fire monitoring system based on UAV team, remote sensing, and image processing[C]// *Proceedings of IEEE Conference on Data Stream Mining and Processing(DSMP)*, 2018: 590-594.

[7] MIAO Z, WANG B, SHI W, et al. A semi-automatic method for road centerline extraction from VHR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(11): 1856-1860.

[8] CHENG G, WANG Y, XU S, et al. Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(6): 3322-3337.

[9] SHAO Z, ZHOU Z, HUANG X, et al. MRENet: Simultaneous extraction of road surface and road centerline in complex urban scenes from very high-resolution images[J]. *Remote Sensing*, 2021, 13(2): 239.

[10] YU C, WANG J, PENG C, et al. Bisenet: Bilateral segmentation network for real-time semantic segmentation [C]// *European Conference on Computer Vision(ECCV)*, 2018: 325-341.

[11] ZHAO H, QI X, SHEN X, et al. Icnnet for real-time semantic segmentation on high-resolution images[C]// *European Conference on Computer Vision(ECCV)*, 2018: 405-420.

[12] WANG Y, ZHOU Q, LIU J, et al. Lednet: A lightweight encoder-decoder network for real-time semantic segmentation [C]// *International Conference on Image Processing (ICIP)*, IEEE, 2019: 1860-1864.

[13] SHI W, MIAO Z, DEBAYLE J. An integrated method for urban main-road centerline extraction from optical remotely sensed imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 52(6): 3359-3372.

[14] DAI J, MA R, GONG L, et al. A model-driven-to-sample-driven method for rural road extraction [J]. *Remote Sensing*, 2021, 13(8): 1417.

[15] LIU R, MIAO Q, SONG J, et al. Multiscale road centerlines extraction from high-resolution aerial imagery[J]. *Neurocomputing*, 2019, 329: 384-396.

[16] LIAN R, HUANG L. DeepWindow: Sliding window based on deep learning for road extraction from remote sensing images [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 1905-1916.

[17] DONG G, YAN Y, SHEN C, et al. Real-time high-performance semantic image segmentation of urban street scenes[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020,

- 22(6):3258-3274.
- [18] LI S, YAN Q, SHI W, et al. HoloParser: Holistic Visual Parsing for Real-time Semantic Segmentation in Autonomous Driving [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-15.
- [19] ZOU W, HUA G, ZHUANG Y, et al. Real-time Passable Area Segmentation with Consumer RGB-D Cameras for the Visually Impaired [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-11.
- [20] YU C, GAO C, WANG J, et al. Bisenet v2: Bilateral network with guided aggregation for real-time semantic segmentation [J]. International Journal of Computer Vision, 2021, 129: 3051-3068.
- [21] FAB M, LAI S, HUANG J, et al. Rethinkingbisenet for real-time semantic segmentation [C] // Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 9716-9725.
- [22] PAN H, HONG Y, SUN W, et al. Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes [J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(3): 3348-3460.
- [23] XU J, XIONG Z, BHATTACHARYYA S P. Pidnet: A real-time semantic segmentation network inspired from pid controller [J]. arXiv: 2206. 02066, 2022.
- [24] ELHASSAN M A M, YANG C, HUANG C, et al. SPFNet: Subspace Pyramid Fusion Network for Semantic Segmentation [J]. arXiv: 2204. 01278, 2022.
- [25] SHI M, SHEN J, YI Q, et al. LMFFNet: A well-balanced lightweight network for fast and accurate semantic segmentation [J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(6): 3205-3219.
- [26] ELHASSAN M A M, YANG C, HUANG C, et al. SS-FPN: Scale-wise Strip Attention Guided Feature Pyramid Network for Real-time Semantic Segmentation [J]. arXiv: 2206. 07298, 2022.
- [27] SUN T, DI Z, CHE P, et al. Leveraging crowdsourcedgps data for road extraction from aerial imagery [C] // Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 7509-7518.
- [28] MEI J, LI R J, GAO W, et al. CoANet: Connectivity attention network for road extraction from satellite imagery [J]. IEEE Transactions on Image Processing, 2021, 30: 8540-8552.
- [29] WANG J, CHEN K, XU R, et al. Carafe: Content-aware reassembly of features [C] // International Conference on Computer Vision (ICCV), 2019: 3007-3016.
- [30] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C] // International Conference on Computer Vision (ICCV), 2017: 2980-2988.



ZHANG Xiaoqing, born in 1997, post-graduate. Her main research interests include computer vision and deep learning.



WANG Qingwang, born in 1990, Ph.D., professor. His main research interests include the collaborative interpretation of multi-source remote sensing data and deep learning.