

融合空洞空间金字塔池化和注意力的轻量化 遥感影像道路提取

刘志恒¹ 岳子腾^{2,*} 周绥平¹ 江澄³ 节永师³ 陈雪梅⁴

(1 西安电子科技大学空间科学与技术学院, 西安 710126)

(2 北京航空航天大学电子信息工程学院, 北京 100191)

(3 北京空间机电研究所先进光学遥感技术北京市重点实验室, 北京 100094)

(4 西安航天天绘数据技术有限公司, 西安 710100)

摘要 针对高分辨率遥感影像中道路形状结构错综复杂, 出现窄小型道路提取错误或漏分的问题, 提出一种基于空洞空间金字塔池化和注意力机制的轻量化遥感影像道路提取方法。首先, 在原始高分辨率网络 (HRNet) 基础上, 通过引入空洞空间金字塔池化模块, 实现多尺度道路信息融合; 再引入挤压激励通道注意力机制, 增强网络特征表征质量; 最后使用深度可分离卷积方法改进网络残差模块实现模型轻量化, 以降低模型计算复杂度。在公开数据集上进行了模型性能测试, 实验结果表明, 文章所提算法的准确率、精确率、召回率、F1 分数和平均交并比, 相比原始 HRNet 分别提升了 5.35%、2.15%、4.1%、3.15% 和 14.34%, 且减少了 36.1% 的参数数量; 相比其他网络, 该算法突出了细小道路的特征, 道路预测结果连续性、完整性好, 并且模型小易于部署在实时检测设备中, 有效改善了道路提取任务中错分和缺失的情况, 是一种适应性更强、分割精度更高、更轻量化的多尺度道路提取算法。

关键词 道路提取 空间金字塔池化 通道注意力机制 可分离卷积 高分辨率网络 遥感影像

中图分类号: TP751

文献标志码: A

文章编号: 1009-8518(2024)01-0111-12

DOI: 10.3969/j.issn.1009-8518.2024.01.010

Lightweight Remote Sensing Image Road Extraction Combining Atrous Spatial Pyramid Pooling and Attention Mechanism

LIU Zhiheng¹ YUE Ziteng^{2,*} ZHOU Suiping¹ JIANG Cheng³ JIE Yongshi³ CHEN Xuemei⁴

(1 School of Aerospace Science and Technology, Xidian University, Xi'an 710126, China)

(2 School of Electronics and Information Engineering, Beihang University, Beijing 100191, China)

(3 Beijing Key Laboratory of Advanced Optical Remote Sensing Technology,

Beijing Institute of Space Mechanics & Electricity, Beijing 100094, China)

(4 Xi'an Aerospace Remote Sensing Data Technology Co., Ltd., Xi'an 710100, China)

收稿日期: 2023-08-08

基金项目: 陕西省自然科学基金基础研究计划资助项目 (2023-JC-QN-0299); 先进光学遥感技术北京市重点实验室开放基金项目 (AORS20238); 自然资源部矿山地质灾害成灾机理与防控重点实验室项目 (2022-08); 中央高校基本科研业务费专项资金资助项目 (300102353502); 自然资源部国土卫星遥感应用重点实验室开放基金项目 (KLSMNR-G202303)

引用格式: 刘志恒, 岳子腾, 周绥平, 等. 融合空洞空间金字塔池化和注意力的轻量化遥感影像道路提取[J]. 航天返回与遥感, 2024, 45(1): 111-122.

LIU Zhiheng, YUE Ziteng, ZHOU Suiping, et al. Lightweight Remote Sensing Image Road Extraction Combining Atrous Spatial Pyramid Pooling and Attention Mechanism[J]. Spacecraft Recovery & Remote Sensing, 2024, 45(1): 111-122. (in Chinese)

Abstract Aiming at the problem of intricate road shape and structure in high-resolution remote sensing images, where narrow and small roads are extracted incorrectly or omitted, a lightweight remote sensing image road extraction method based on Atrous Space Pyramid Pooling and Attention Mechanism is proposed. Firstly, based on the original HRNet network, multi-scale road information fusion is realized by introducing the ASPP. Secondly, the Squeeze and Excitation channel attention mechanism (SE-networks) is introduced to enhance the quality of network feature representation. Finally, using deep separable convolution to improve the network residual module to realize the model lightweight and reduce the complexity of model calculation. Experimental results on the publicly available dataset show that the accuracy, precision, recall, F1 score and the MIOU of the proposed algorithm was improved respectively by 5.35%, 2.15%, 4.1%, 3.15% and 14.34%, compared with the original HRNet network, and reduce the number of parameters by 35.6%. Compared with other networks, the algorithm highlights the characteristics of small roads, and the prediction results have good continuity and integrity. As the small size, the proposed model is easier to deploy in real-time detection equipment. The proposed model effectively reduces the road extraction fault and missing, implements a stronger adaptability, higher segmentation accuracy, more lightweight multi-scale road semantic segmentation algorithm.

Keywords road extraction; ASPP; channel attention mechanism; separable convolution; High-Resolution Network; remote sensing images

0 引言

在城市发展和规划过程中,道路是不可或缺的元素之一。道路信息是城市地理信息系统研究与分析的重要组成部分,在车辆导航、智能交通、地图绘制、城市规划等应用领域发挥着广泛的作用^[1-3]。随着遥感技术的快速发展,遥感图像的空间分辨率在逐步提高,目前已经达到分米级别,且呈上升发展趋势。从遥感图像中提取道路信息是城市规划与社会发展等领域的一个热点问题。

目前,大多数道路提取研究仍采用传统的半自动化方式^[4]。此类道路提取方法大致可划分为三类:模板匹配法、知识驱动法、面向对象法^[5]。这种传统提取方法需要人工参与标定,按照特定的规则和逻辑提取道路信息和非道路信息,且依赖于研究人员经验,存在算法迁移能力不足、分割精度受限、流程繁琐耗时长等问题,不适用于大量数据的处理,给道路提取带来诸多困难。

以卷积神经网络(Convolution Neural Networks, CNNs)^[6]为代表的深度学习图像处理理论与技术近年来发展迅速。基于CNNs的图像分析方法,在各种计算机视觉任务中取得了良好的性能,如目标检测、语义分割和时间序列预测等。相较于传统方法,基于深度学习的方法在遥感影像目标识别与图像语义分割等方面具有显著优势,利用此类方法解决遥感影像数据中道路信息提取问题,已成为当前研究的主要趋势。该类算法利用深度学习的强大的学习能力、更高效的计算能力和更精准的特征表达能力,准确地区分道路和背景信息。Minh等^[7]将深度学习应用于遥感图像道路提取任务中,证明了利用深度学习提取遥感影像上的道路信息是可靠的。刘笑等^[8]利用全卷积神经网络和迁移学习方法,细化抽离模型的中间层信息,实现道路信息的快速提取。该算法提取的道路中心线位置较为准确,但道路细节处理不佳(如密集建筑区的细小道路),或与其忽略纹理特征有关。贺浩等^[9]通过改进二分类交叉熵损失函数解决了正负样本不平衡的问题,提出的编码-解码结构道路提取网络,有效突出了模型对道路信息的细节表达能力。王卓^[10]利用VGG16改进Unet网络^[11]的编码结构,提取道路特征语义信息,并使用指数激活函数(Exponential Linear Units, ELU)^[12],提升了网络训练速度,但在阴影和树木等遮挡的道路区域,提取效果欠佳,需进一步改进。李朝奎等^[13]基于全卷积网络的特点,利用Deeplabv1网络^[14]中的空洞卷积

来突出道路的多尺度特征信息,在不牺牲特征空间分辨率的同时扩大了特征感受野(Field-of-View, FOV),实验结果表明该算法对细窄道路提取有明显帮助。此外,韩玲等^[15]以 Deeplabv3 网络^[16]为主干框架,利用空洞空间金字塔池化(ASPP)突出了高分辨率遥感影像中的道路边缘特征,实验表明空洞空间金字塔池化可有效提取遥感影像上道路的多尺度特征,更有利于细小道路的准确提取。王勇和曾祥强^[17]以 Unet 网络为基础,引入残差模块(Residual Learning Unit, RLU)降低训练复杂度,并使用混合损失函数(Binary cross entropy 和 Dice)解决类不平衡问题,有效提取了被阴影等遮挡的道路范围。张亚宁等^[18]以 Unet 网络为主干框架,提出了一种结合改进的残差模块和卷积注意力机制(Convolutional Block Attention Module, CBAM)^[19]的道路提取网络,相较原始的 Unet 网络,所提算法有效提高了网络对道路特征的表达能力,抑制了“同谱异物”的非道路背景信息。以上研究表明:随着遥感技术的进步与革新,遥感影像的空间、光谱、时间等分辨率不断提高;高分辨率遥感图像中的地物背景等细节信息更加丰富,非道路信息如植被阴影、车辆流动、高楼建筑遮挡、人流流动等作为干扰信息十分繁杂,一方面周围环境可能与道路十分相似,如城市中楼宇的纹理、狭长的河道、沙漠中由风吹形成的痕迹等,另一方面,遥感图像中的道路成分可能被周围障碍物遮挡,如形成视觉遮挡或阴影遮挡等,这些问题给道路信息的提取造成了困难,已有道路分割模型在高分辨率遥感影像道路提取任务中仍存在模型适应性不足、模型分割精度欠佳等问题。

本文针对高分辨率遥感影像中道路形状结构错综复杂,出现窄小型道路提取错误或漏分的问题,提出了一种基于空洞空间金字塔池化和注意力机制的轻量化遥感影像道路提取方法。该方法以原始 HRNet 网络为主干框架,以公开的 CHN6-CUG 道路遥感影像数据集为基础数据源(https://grzy.cug.edu.cn/zhuqiqi/zh_CN/yjgk/32368/list/index.htm)^[20],结合空洞空间金字塔池化结构扩大道路提取模型感受野,优化模型提取多尺度道路特征信息能力;引入通道注意力机制构建特征图不同通道间的依赖关系,强化特征图中道路特征信息表征,从而提高遥感影像道路提取精度;利用深度可分离卷积方法改进残差模块,减少模型参数量,有效降低模型计算复杂度。

1 网络结构

1.1 HRNet 网络结构

原始的 HRNet 是 2019 年由微软亚洲研究院^[21]提出的神经网络结构,被广泛应用于图像语义分割、姿态估计和目标检测等任务中。原始 HRNet 网络结构的特征是能够在保持数据图像高分辨率的同时,完成特征提取,其融合结构如图 1 所示。在同一个 Stage 中,前一层的特征图,分别通过图 1 中所示 3 种不同的特征融合方式,得到后一层的特征图。3 种方式融合前的输入,由上而下分别为低级、中级、高级特征图分辨率大小的前一层特征图,融合后的输出分别为低级、中级、高级特征图分辨率大小的后一层特征图。图 1 中方式(1)表示将中级和高级特征图分辨率大小的前一层特征图,通过上采样,并与低级特征图进行特征融合,恢复成前一层低级特征图分辨率大小的后一层特征图;方式(2)表示将前一层低级特征图经过卷积得到的特征图和高级特征图经过上采样得到的特征图,并与前一层中级特征图进行特征融合,恢复成前一层中级特征图分辨率大小的后一层特征图;方式(3)表示将前一层低级特征图和中级特征图经过卷积得到的特征图,与前一层高级特征图进行特征融合,恢复成前一层高级特征图分辨率大小的后一层特征图。原始的 HRNet 网络通过逐步添加、互联并行的低分辨率子网络,不断对并行的图像特征信息进行多尺度融合,使模型预测获得较好的准确性。

现有基于原始 HRNet 的语义分割模型,仍存在一定的缺陷与不足,主要包括:1)在道路密集的城市中,道路结构错综复杂,已有模型在分割多样的道路结构时存在适应性不足的问题,对多尺度道路目标分割效果不佳;2)在高分辨率遥感图像中,一方面道路周围环境可能与道路十分相似,另一方面图像中的道路成分可能被周围障碍物遮挡,这些问题导致深度学习模型难以准确提取出图像中的道路信息,

使模型提取精度受限；3) 原始 HRNet 对并行的子网络进行不断地信息交换与特征融合，此过程存在许多重复计算，导致模型训练产生大量参数，增大了模型计算复杂度。

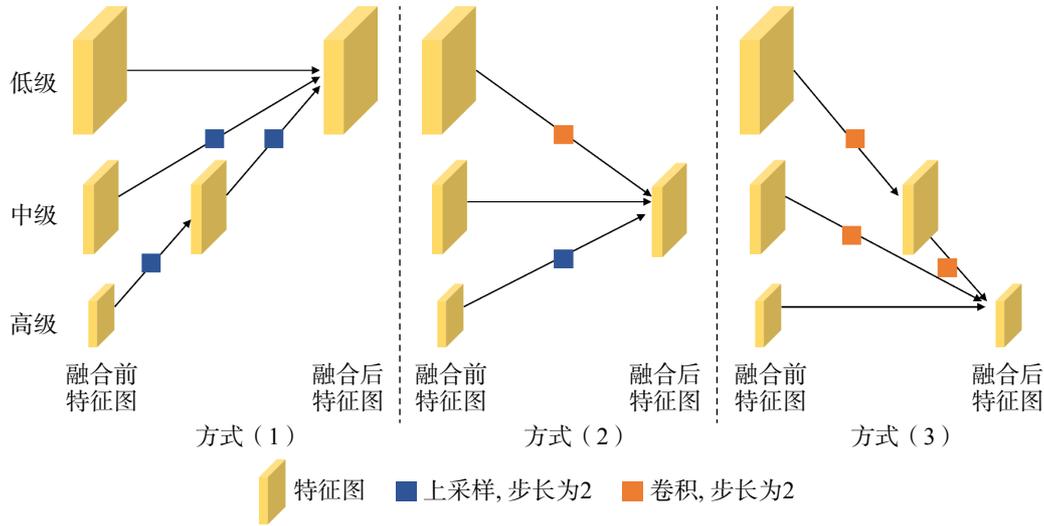


图1 原始的 HRNet 多尺度特征融合结构示意图
Fig.1 Original HRNet multiscale feature fusion structure map

为解决上述问题，本文以原始 HRNet 网络为主干框架，提出一种融合空洞空间金字塔池化和注意力机制的轻量化遥感图像道路提取方法。该方法利用空洞空间金字塔池化模块，提取遥感数据中的多尺度道路特征信息，提高原始 HRNet 网络提取多尺度道路目标的能力；引入 SE 通道注意力机制，构建特征图通道间的相互依赖关系，加强特征图中重要特征表征，改善原始 HRNet 网络的特征提取质量；利用深度可分离卷积方法改进网络残差模块，减少原始 HRNet 网络模型训练过程中产生的参数量，降低模型计算复杂度。本文提出的改进后的 HRNet 网络模型结构见图 2。

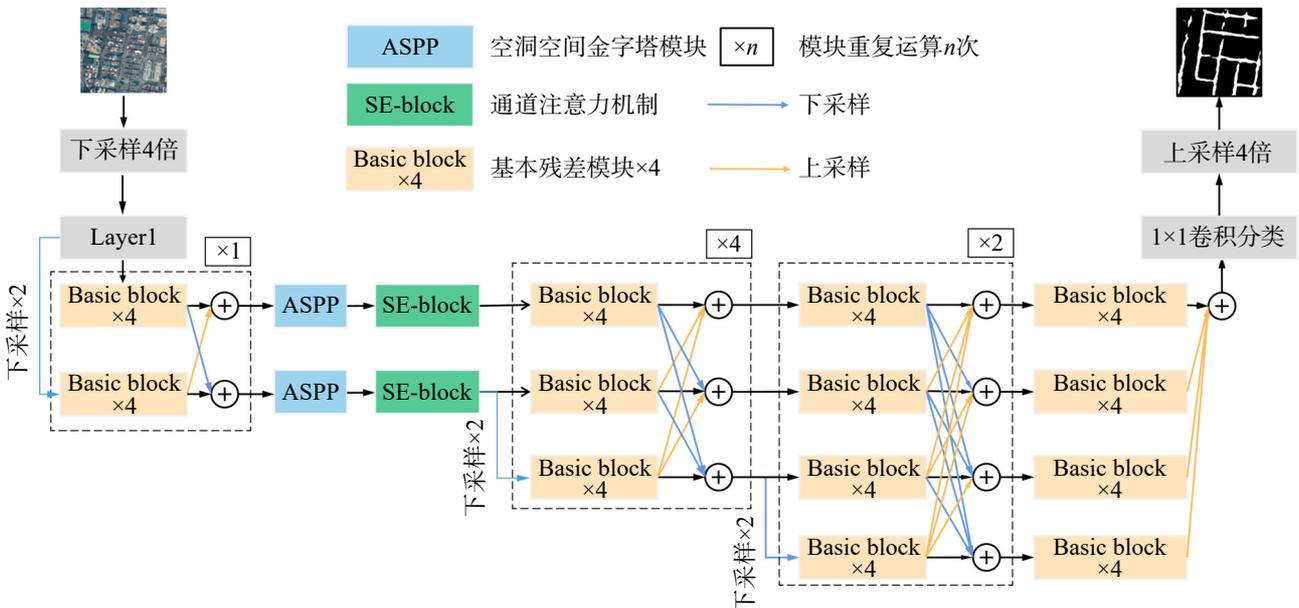


图2 网络结构示意图
Fig.2 Schematic diagram of the network structure

1.2 空洞空间金字塔池化模块

高分辨率遥感影像中的道路形状结构错综复杂、变化多样，道路的上下文信息复杂。仅使用简单的

池化方式处理卷积后的特征图,降低特征图的维度,可能造成细小狭窄道路或者多路并行的宽阔道路的部分区域被误判为背景成分。提取道路的多尺度特征,有助于减少空间信息损失。结合道路区域的上下文信息,学习不同尺度道路的融合特征,能够提升对多尺度道路目标的分割效果。

本文提出的空洞空间金字塔池化模块共包括 6 个分支,其结构如图 3 所示。其中第一层为原特征图像,用于补充特征图中的空间信息;最后一层是平均池化层,作用是获得图像级别的全局特征;第二层至第五层为采用不同膨胀率的空洞卷积层,用于扩大特征图感受野,综合提取多尺度道路特征信息。根据 HDC 准则^[22]对 4 个分支的膨胀率 r_i (r_i 为第 i 个膨胀卷积分支的膨胀率, $i \in \{1,2,3,4\}$) 进行选择,设置的膨胀率分别为 2、3、7、13。这组膨胀率间不存在大于 1 的公约数,且:

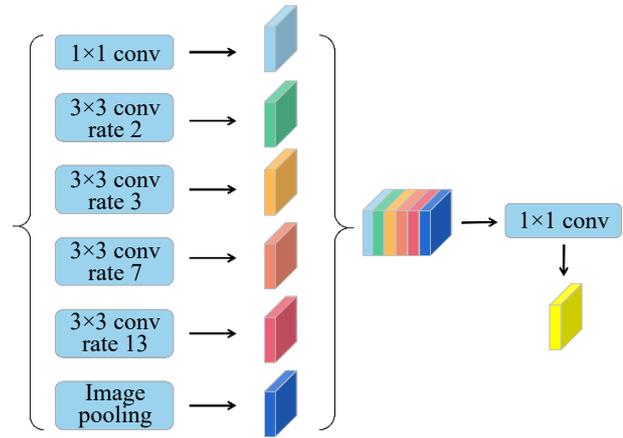


图 3 空洞空间金字塔池化模块

Fig.3 Atrous spatial pyramid pooling module

根据 HDC 准则^[22]对 4 个分支的膨胀率 r_i (r_i 为第 i 个膨胀卷积分支的膨胀率, $i \in \{1,2,3,4\}$) 进行选择,设置的膨胀率分别为 2、3、7、13。这组膨胀率间不存在大于 1 的公约数,且:

$$M_2 = \max(M_3 - 2r_2, 2r_2 - M_3, r_2) = 3 \tag{1}$$

$$M_2 \leq C \tag{2}$$

式中 C 为卷积核大小; M_2 、 M_3 为定义的第 2、3 层膨胀卷积中两个非零点间的最大距离。该组膨胀率满足 HDC 准则,证明 4 个膨胀率组成的空洞卷积金字塔其感受野可以覆盖底层特征图的整个区域。这一方法可有效改善网络问题,增强原始 HRNet 网络提取多尺度道路信息的能力。

1.3 挤压激励通道注意力机制

针对高分辨率遥感图像中地物及背景信息丰富、非道路干扰信息繁杂、道路特征不突出等问题,本文引入挤压激励通道注意力机制 (Squeeze and Excitation Networks, SE) 改进原始的 HRNet 网络,该机制的结构如图 4 所示。采用 SE 通道注意力机制,构建特征图中不同通道间的相互依赖关系,改善道路提取模型中细小道路特征不突出问题,增强原始 HRNet 网络的特征表征质量。

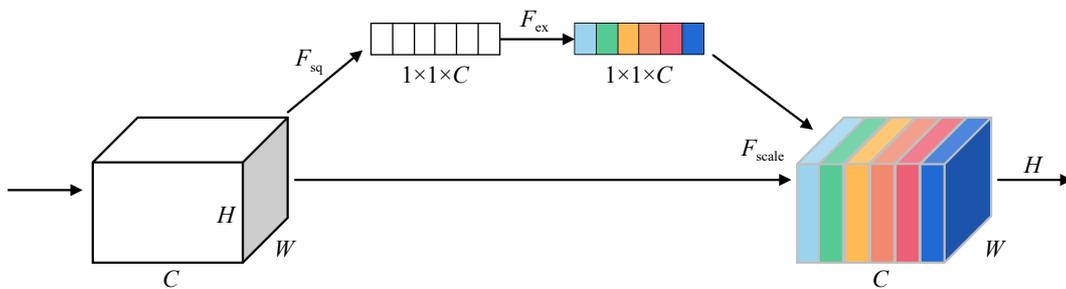


图 4 通道注意力机制结构

Fig.4 Squeeze and Excitation Networks structure

1.4 模型轻量化

由于原始 HRNet 网络使用并行结构对子网络不断地进行信息交换与特征融合,在网络计算过程中不可避免会产生许多重复计算,占用大量内存且导致过拟合。本文采用深度可分离卷积方法改进原始 HRNet 中的残差模块,实现模型轻量化。改进后的残差模块如图 5 所示,将原始 HRNet 残差结构中的一层 3×3 普通卷积替换为 3×3 深度可分离卷积,保留一层 3×3 普通卷积。模型轻量化能够有效减少神经网络计算过程中产生的参数量,降低计算复杂度,同时对模型分割精度不产生显著影响。

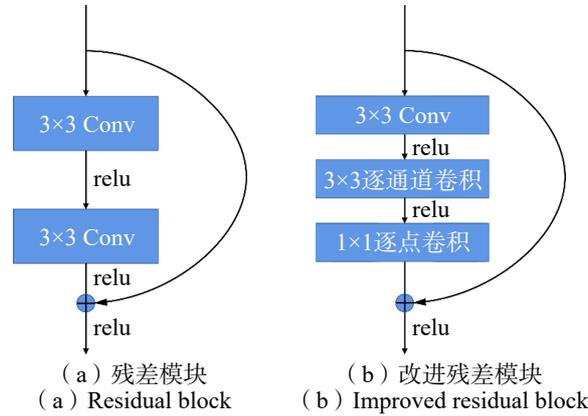


图5 深度可分离方法改进结构
Fig.5 Depthwise convolution methods for improved structure

2 实验与评价指标

2.1 数据集简介

随着高分辨率遥感技术的迅速发展，传感器空间分辨率得到显著提升，人们获得了越来越多的高质量遥感影像，出现了大量高分辨率遥感影像数据集，为开展遥感影像道路提取研究提供了丰富的数据来源。在深度学习模型的训练实验中，数据集的大小、数量及质量的高低，最终都会影响实验的准确度和召回率等结果。本文采用公开的中国区域道路数据集（CHN6-CUG 道路数据集，该数据集由中国地质大学（武汉）高性能空间智能计算实验室（HPSCIL）团队制作），选取北京市朝阳区、上海市杨浦区、武汉市中心、深圳南山区、香港沙田、澳门等6个地区作为实验区。CHN6-CUG 道路数据集的影像来自谷歌地球，共包含3681组标记图像，图像大小为512像元×512像元，图像分辨率为0.5 m。CHN6-CUG 数据集中样本及其标签图像如图6所示，其中第一行为原始遥感图像，第二行为对应的道路标签图像。

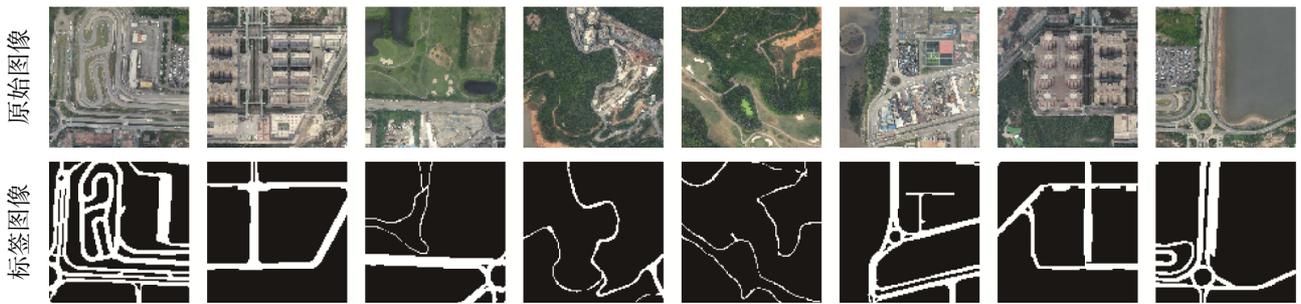


图6 CHN6-CUG 数据集标注图像样本示意
Fig.6 Schematic diagram of annotated image samples of CHN6-CUG dataset

2.2 数据集预处理

由于单张样本图像所包含的特征信息有限，因此需使用大量样本对深度学习模型进行训练，才能使模型学习到更多更深层次的特征，从而实现精细化的语义分割，提升模型的泛化能力。深度学习算法的实现依靠大量数据样本的迭代计算，实际实验中会存在样本数量不足的情况，因此需要采取一定方法对已有数据进行扩增^[23]。研究表明，在已确定的数据集上进行合适的数据样本扩增，能很好地提高数据样本复杂度，改善训练所得模型性能。本文采用的遥感影像数据集中图像是正射投影后的高分辨率遥感影像，图像的角度特征变化不明显，可以通过旋转及镜像等操作增加样本数据的复杂性，本文分别采用旋转90°、180°、270°及水平镜像操作对数据集进行扩充，扩充后的样本示例如图7所示。

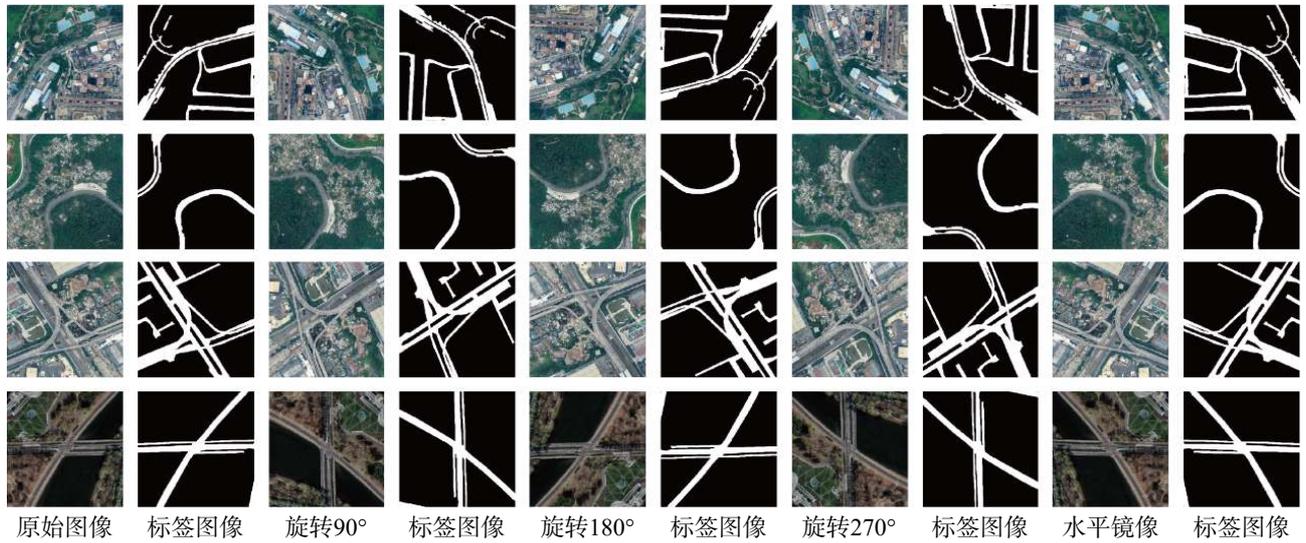


图 7 经扩充操作后的样本数据实例

Fig.7 Example of sample data after expansion operation

对数据集的预处理包括数据归一化与图像标准化。数据归一化处理将输入数据的特征值大小调整至相近范围,避免得到的梯度值过大或过小,使模型训练能够收敛平稳。图像标准化处理可以实现数据去均值中心化。数据中心化使输入数据符合分布规律,更易于取得训练后的泛化效果,加速训练和拟合过程。

归一化计算和标准化计算分别如下:

$$I_{\text{output}} = (I_{\text{input}} - \min(I_{\text{input}})) / (\max(I_{\text{input}}) - \min(I_{\text{input}})) \quad (3)$$

$$I_{\text{output}} = (I_{\text{input}} - \text{mean}(I_{\text{input}})) / \text{std}(I_{\text{input}}) \quad (4)$$

式中 I_{output} 、 I_{input} 分别为输出像素值与输入像素值; std 、 mean 、 max 和 min 分别为标准差、平均值、最大值和最小值。样本经归一化、标准化处理后所得结果如图 8 所示。

将经过预处理后的遥感影像数据集按照 6 : 2 : 2 的比例划分为训练集、验证集、测试集,将训练集与验证集作为模型训练输入,取 DiceLoss 函数与 CrossEntropyLoss 函数的均值作为模型损失函数,用来解决类不平衡的问题,即不同类别的训练样例数目差别很大的情况,特别是当正例个数远小于反例个数,则训练出的模型难以检测正例。

实验平台采用 Intel(R) Xeon(R) Platinum 8358P, GPU 选取使用 GeForce RTX 3090 24 G 显存容量,能够保证网络的训练运行。

2.3 评价指标

为检测模型道路提取结果的精度,本文采用语义分割领域常用的评价指标——精确率 (Precision)、准确率 (Accuracy)、召回率 (Recall)、F1 分数和均交并比 (MIoU),来定量评估模型性能以及标签图像与提取结果间具体差异,上述指标的具体计算公式为:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (5)$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (6)$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (7)$$

$$F_1 = 2\text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall}) \quad (8)$$

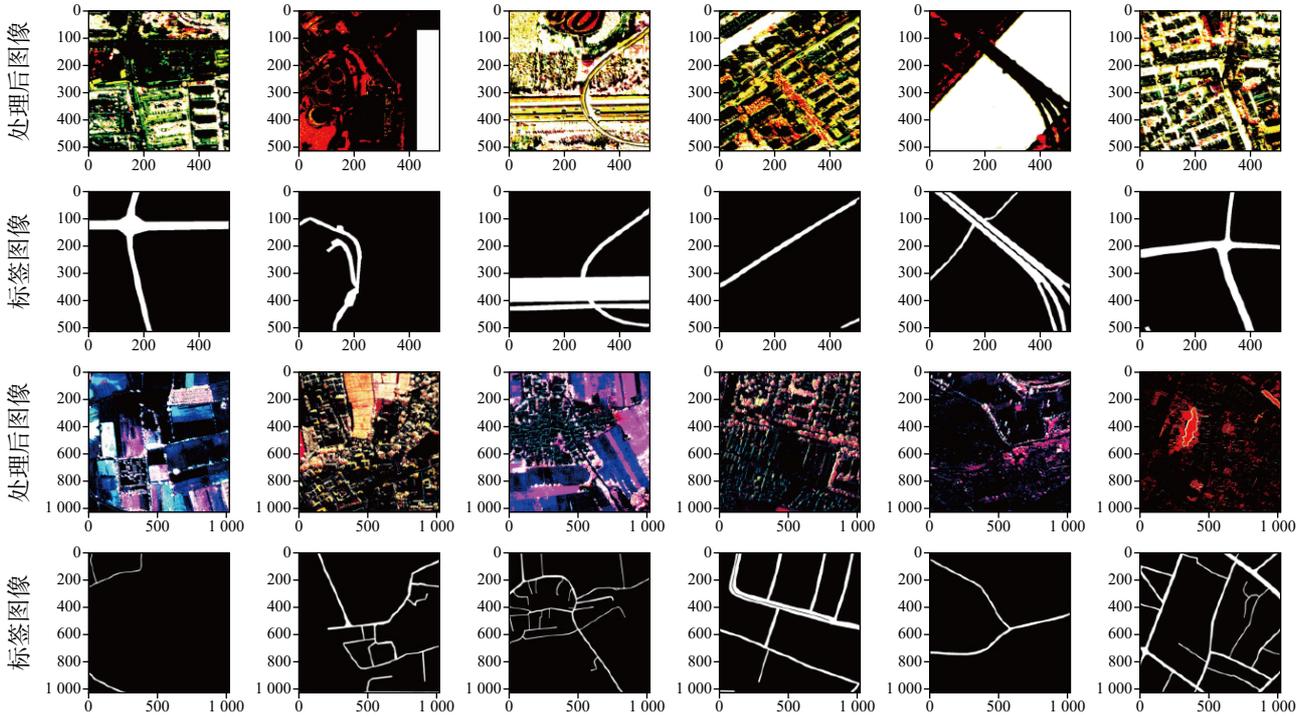


图8 CHN6-CUG数据集标准化、归一化处理图像

Fig.8 CHN6-CUG dataset standardized, normalized processing images

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \left\{ p_{ii} / \left(\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii} \right) \right\} \quad (9)$$

式中 TP 代表真阳性，表示实际道路像素被预测为道路像素；FP 代表假阳性，表示非道路像素被预测为道路像素；FN 代表假阴性，表示道路信息被预测成非道路信息；TN 代表真阴性，表示非道路信息被预测为非道路信息； k 表示目标类总数，式 (9) 中 $k+1$ 表示类别数目 (k 个目标类和 1 个背景类)； p_{ij} 表示将实际类别为第 i 类预测为第 j 类的像素数目； p_{ji} 表示将实际类别为第 j 类预测为第 i 类的像素数目； p_{ii} 表示将实际类别为第 i 类预测为第 i 类的像素数目。

3 实验结果对比与分析

为验证本文所提出的改进后的 HRNet 各模块后的有效性，在相同实验条件下，比较了原始 HRNet、原始 HRNet+ASPP 模块、原始 HRNet+SE-Block 模块及本文方法的性能。如表 1 所示，加入 SE-Block 的 HRNet 其 Accuracy、Precision、Recall 及 F-Score 等指标结果，相较原始 HRNet 分别提高了 4.85%、1.8%、3.85% 和 2.85%，表明在原始 HRNet 结构上增加通道注意力机制，提高了网络特征提取质量。加入 ASPP 模块的 HRNet 相较原始 HRNet，上述各指标分别提高了 4.96%、1.34%、4.41% 和 2.90%，表明空洞空间金字塔池化模块有效扩大了网络感受野范围，增强了模型提取全局特征的能力。

表 1 消融实验的评价指标对比

Tab.1 Comparison of evaluation metrics for ablation experiment

算法名称	Accuracy	Precision	Recall	F1-Score	%
原始 HRNet	89.41	95.05	92.40	93.70	
原始 HRNet+SE-Block	94.26	96.85	96.25	96.55	
原始 HRNet+ASPP	94.37	96.39	96.81	96.60	
本文方法 (原始 HRNet+SE-Block+ASPP)	94.76	97.20	96.50	96.85	

为验证本文所提算法在高分辨率遥感影像道路提取任务中的性能, 使用不同的深度学习网络 (UNet、FCN、PSPNet^[24]、DeepLabV3+^[25]、D-LinkNet、原始 HRNet 及本文算法), 在相同条件下对选取的 CHN6-CUG 数据集图像进行预测和验证。利用准确率、召回率等参数对验证集进行评估对比, 场景主要覆盖了城市中多种复杂路网。训练得出的具体模型评估结果如表 2 所示。

表 2 不同模型的评估结果
Tab.2 Assessment results from different models

名称	Accuracy	Precision	Recall	F1-Score	MIoU
UNet	88.76	96.90	90.26	93.46	65.32
FCN	87.70	94.02	91.39	92.68	65.32
PSPNet	89.26	96.84	90.81	93.73	66.90
DeepLabV3+	90.88	96.12	93.11	94.59	72.46
D-LinkNet	87.79	95.58	90.26	92.85	63.95
原始 HRNet	89.41	95.05	92.40	93.70	69.14
本文算法	94.76	97.20	96.50	96.85	83.48

不同模型各评估指标对比如图 9 所示, 可以看出, 原始 HRNet 网络的总体性能优于 UNet、FCN、PSPNet、D-LinkNet, 在准确率、召回率、F1 分数、MIoU 等方面也不同程度的优于 UNet、FCN、PSPNet 和 D-LinkNet, 证明了利用原始 HRNet 提取高分辨率遥感图像中道路信息的可行性。但原始 HRNet 相较其他网络精确率较低, 说明原始 HRNet 易将背景成分误判为道路信息。DeepLabV3+网络评估数据相较原始 HRNet、UNet、FCN、PSPNet、D-LinkNet 网络有明显提升, 其中召回率达到 93.11%, MIoU 指数达到 72.46%。由于 DeepLabV3+网络使用了空洞空间金字塔池化结构, 有效扩大了网络感受野范围, 增强了网络提取全局道路特征信息的能力, 表明在图像语义分割网络中引入空洞空间金字塔池化结构, 可以有效提高模型对道路全局特征信息提取的能力。因此, 在 DeepLabV3+网络启发下, 本文在原始 HRNet 网络特征融合结构之后添加 ASPP 模块, 以扩大原始 HRNet 网络的感受野范围。同时, 在 ASPP 模块后添加 SE 通道注意力机制, 优化原始 HRNet 网络的特征提取质量。通过引入 ASPP 模块和 SE 通道注意力机制, 共同改进原始的 HRNet 网络结构, 构成本文所提出的改进后的 HRNet 网络。根据表 2 数据分析, 本文所提算法在准确率、精确率、召回率、F1 分数、MIoU 指数方面的评估结果相较原始 HRNet 分别提高了 5.35%、2.15%、4.1%、3.15% 和 14.34%, 性能提升明显。

进一步分析各模型预测图像与标签图的差异, 结果如图 10 所示。图 10 所选示例图像包括城市密集道路网络、有阴影遮挡的道路、有植被覆盖的道路、密集高楼间的道路、河道旁道路、不规则曲线道路等多种道路情况。UNet、FCN、PSPNet、D-LinkNet 预测结果与标签图相比, 出现漏检情况较多, 主要表现为预测结果中道路不连续; 原始 HRNet 网络预测结果中出现较多漏检情况, 无法完整预测遥感图像中的道路信息, 但细节信息相较其他网络更丰富; DeepLabV3+网络预测结果较为准确, 能较为全面的预测出遥感图像中的道路信息, 但在道路细节处, 预测结果较为模糊, 道路边界不明晰; 本文所提算法继承了原始 HRNet 高分辨率与分割细节优良的特点, 利用 ASPP 模块提高模型综合多尺度语义信息能力,

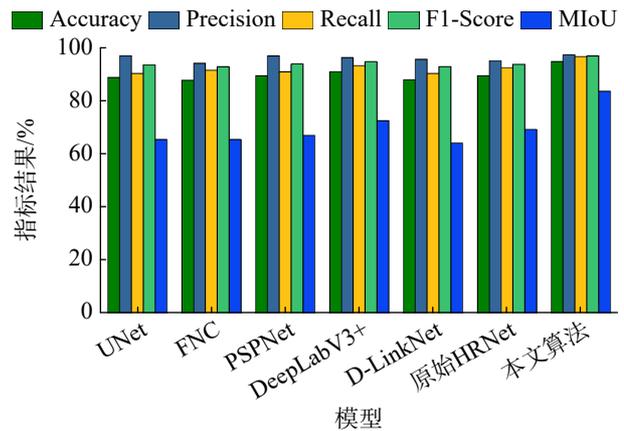


图 9 不同指标评估模型性能
Fig.9 Model performance evaluation with different

同时增加 SE 通道注意力机制，网络预测结果与标签图像重合程度高，预测结果中道路细节丰富清晰，道路完整性、连续性好，能有效地从遥感图像中提取完整的道路信息。

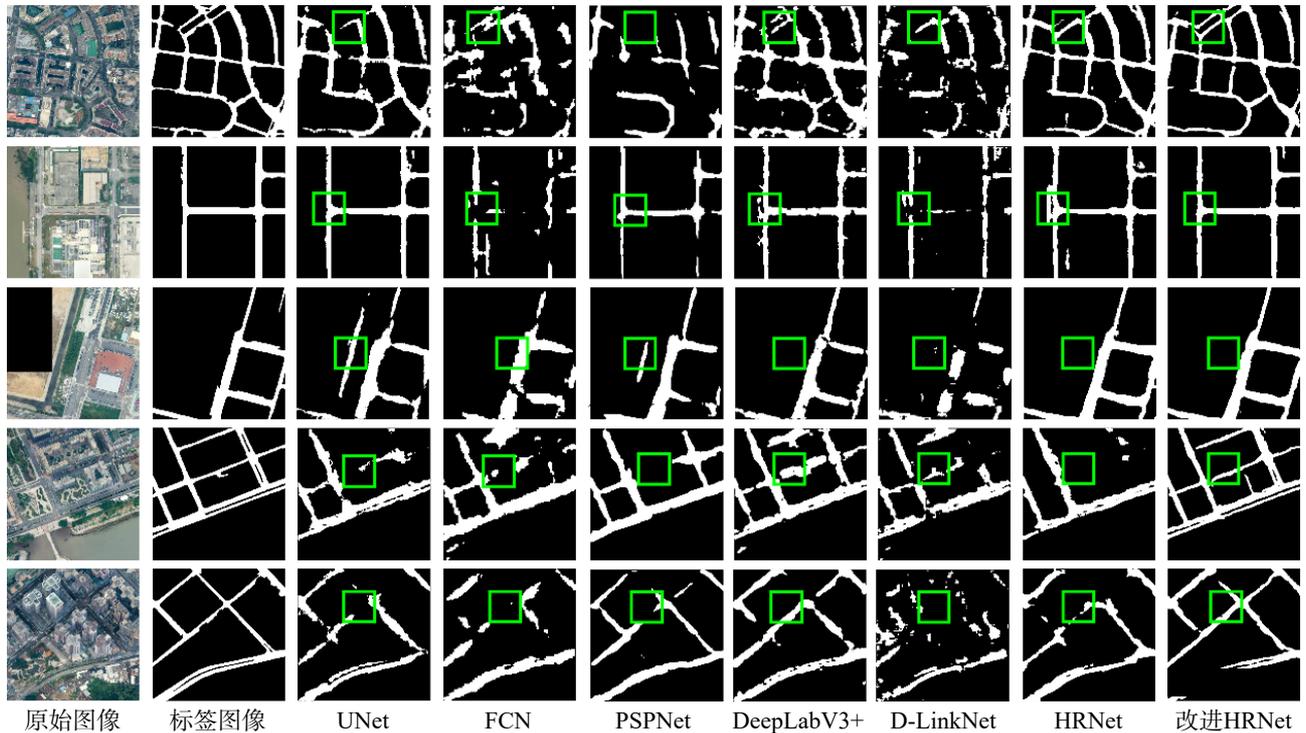


图 10 不同分割模型的道路提取结果对比

Fig.10 Comparison of road extraction results with different segmentation models

为实现轻量化部署，本文所提算法中引入了深度可分离卷积方法，引入前后算法的网络参数量结果如表 2 所示。本文方法剔除深度可分离卷积模块后网络参数量约为 9.71×10^6 个，引入深度可分离卷积的本文算法网络参数量约为 6.21×10^6 个，网络参数量减少了约 36.1%，有效减少了网络参数量，降低了模型训练的计算复杂度与运算成本。

4 结束语

针对高分辨率遥感影像中道路形状结构错综复杂，出现窄小型道路提取错误或漏分的问题，本文以原始 HRNet 网络为主干框架，引入空洞空间金字塔池化模块，有效扩大了网络感受野，提高了网络提取多尺度道路信息的能力；引入挤压激励通道注意力机制，改善网络特征提取质量，突出了模型中细小道路特征；使用深度可分离卷积方法，改进网络的残差模块，在不损失模型精确度条件下，有效减少了网络参数量，提高了网络训练速度。在 CHN6-CUG 道路数据集上，通过不同区域的提取结果，本文所提算法对细窄道路的提取效果更明显，同时道路预测结果连续性、完整性良好。后续改进的重点是在已有模型的基础上，设计实现针对线性地物复杂、噪声明显的遥感影像道路提取算法。

参考文献 (References)

- [1] 张彩丽, 向隆刚, 李雅丽, 等. 顾及路网与轨迹多模特征的道路等级分类研究[J]. 地球信息科学学报, 2022, 24(10): 1925-1940.
ZHANG Caili, XIANG Longgang, LI Yali, et al. A Research on Road Type Classification Considering the Multi-mode

- Features of Road Network and Trajectories[J]. *Journal of Geo-information Science*, 2022, 24(10): 1925-1940. (in Chinese)
- [2] 谢波, 肖扬谋. 城市道路网络特征对交通事故作用机理的研究进展[J]. *地理科学进展*, 2022, 41(10): 1956-1968.
XIE Bo, XIAO Yangmou. Progress of Research on the Mechanism of Impact of Urban Road Network Characteristics on Traffic Accidents[J]. *Progress in Geography*, 2022, 41(10): 1956-1968. (in Chinese)
- [3] 李倩雯, 靳甜甜, 蒋爱萍, 等. 道路建设对西南地区景观格局的影响[J]. *生态学报*, 2023, 43(6): 2310-2322.
LI Qianwen, JIN Tiantian, JIANG Aiping, et al. Effects of Road Construction on Landscape Pattern in Southwest China[J]. *Acta Ecologica Sinica*, 2023, 43(6): 2310-2322. (in Chinese)
- [4] 岳贵杰, 杜黎明, 刘凤德, 等. 利用局部梯度最大值的半自动道路提取[J]. *测绘科学*, 2015, 40(5): 89-91+105.
YUE Guijie, DU Liming, LIU Fengde, et al. An Algorithm for Semi-automatic Road Extraction Based on Local Gradient Maximum[J]. *Science of Surveying and Mapping*, 2015, 40(5): 89-91+105. (in Chinese)
- [5] 戴激光, 王杨, 杜阳, 等. 光学遥感影像道路提取的方法综述[J]. *遥感学报*, 2020, 24(7): 804-823.
DAI Jiguang, WANG Yang, DU Yang, et al. Development and Prospect of Road Extraction Method for Optical Remote Sensing Image[J]. *National Remote Sensing Bulletin*, 2020, 24(7): 804-823. (in Chinese)
- [6] LECUN Y, BENGIO Y. Convolutional Networks for Images, Speech, and Time Series[M]. [S.l.]: MIT Press, 1998.
- [7] MNIH V, HINTON G E. Learning to Detect Roads in High-Resolution Aerial Images[C]//Computer Vision-ECCV 2010-11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part VI. Berlin, Heidelberg: Springer, 2010.
- [8] 刘笑, 王光辉, 杨化超, 等. 全卷积神经网络遥感影像道路提取方法[J]. *遥感信息*, 2018, 33(1): 69-75.
LIU Xiao, WANG Guanghui, YANG Huachao, et al. Road Extraction from Remote Sensing Image Based on Fully Convolutional Networks[J]. *Remote Sensing Information*, 2018, 33(1): 69-75. (in Chinese)
- [9] 贺浩, 王仕成, 杨东方, 等. 基于Encoder-Decoder网络的遥感影像道路提取方法[J]. *测绘学报*, 2019, 48(3): 330-338.
HE Hao, WANG Shicheng, YANG Dongfang, et al. An Road Extraction Method for Remote Sensing Image Based on Encoder-Decoder Network[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(3): 330-338. (in Chinese)
- [10] 王卓, 闫浩文, 禄小敏, 等. 一种改进U-Net的高分辨率遥感影像道路提取方法[J]. *遥感技术与应用*, 2020, 35(4): 741-748.
WANG Zhuo, YAN Haowen, LU Xiaomin, et al. High-Resolution Remote Sensing Image Road Extraction Method for Improving U-Net[J]. *Remote Sensing Technology and Application*, 2020, 35(4): 741-748. (in Chinese)
- [11] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]//NAS-SIR N, WILLIAM M W, JOACHIM H, et al. Medical Image Computing and Computer-assisted Intervention - MICCAI 2015. Cham: Springer International Publishing, 2015.
- [12] CLEVERT D A, UNTERTHINER T, HOCHREITER S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUS)[EB/OL]. (2016-02-15)[2023-08-08]. <https://arxiv.org/pdf/1511.07289.pdf>.
- [13] 李朝奎, 曾强国, 方军, 等. 改进全卷积网络方法的高分二号影像农村道路提取[J]. *遥感学报*, 2021, 25(9): 1978-1988.
LI Chaokui, ZENG Qiangguo, FANG Jun, et al. Road Extraction in Rural Areas from High Resolution Remote Sensing Image Using a Improved Full Convolution Network[J]. *National Remote Sensing Bulletin*, 2021, 25(9): 1978-1988. (in Chinese)
- [14] CHEN L, MURPHY K, KOKKINOS I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 40(4): 834-848.
- [15] 韩玲, 杨朝辉, 李良志, 等. 利用Deeplab V3提取高分辨率遥感影像道路[J]. *遥感信息*, 2021, 36(1): 22-28.
HAN Ling, YANG Zhaohui, LI Liangzhi, et al. Road Extraction of High Resolution Remote Sensing Imagery Based on Deeplab V3[J]. *Remote Sensing Information*, 2021, 36(1): 22-28. (in Chinese)
- [16] CHEN L, PAPANDREOU G, SCHROFF F, et al. Rethinking Atrous Convolution for Semantic Image Segmentation[EB/OL]. (2022-12-20)[2023-08-08]. <https://arxiv.org/abs/1706.05587>.
- [17] 王勇, 曾祥强. 集成注意力机制和扩张卷积的道路提取模型[J]. *中国图象图形学报*, 2022, 27(10): 3102-3115.
WANG Yong, ZENG Xiangqiang. Road Extraction Model Derived from Integrated Attention Mechanism and Dilated Convolution[J]. *Journal of Image and Graphics*, 2022, 27(10): 3102-3115. (in Chinese)
- [18] 张亚宁, 张春亢, 王朝, 等. 融合残差和卷积注意力机制的U-Net网络高分影像道路提取[J]. *航天返回与遥感*, 2023, 44(3): 119-132.
ZHANG Yaning, ZHANG Chunkang, WANG Chao, et al. Road Extraction Method of High-Resolution Image Based on U-Net

- Network Combining Residual and Convolution Attention Mechanism[J]. *Spacecraft Recovery & Remote Sensing*, 2023, 44(3): 119-132. (in Chinese)
- [19] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module[C]//*Computer Vision - ECCV 2018*. Berlin: Springer, 2018.
- [20] ZHU Q, ZHANG Y, WANG L, et al. A Global Context-aware and Batch-independent Network for Road Extraction from VHR Satellite Imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 175: 353-365.
- [21] SUN K, XIAO B, LIU D, et al. Deep High-Resolution Representation Learning for Human Pose Estimation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020.
- [22] WANG P, CHEN P, YUAN Y, et al. Understanding Convolution for Semantic Segmentation[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE, 2018.
- [23] 杨军, 于茜子. 结合空洞卷积的FuseNet变体网络高分辨率遥感影像语义分割[J]. *武汉大学学报(信息科学版)*, 2022, 47(7): 1071-1080.
- YANG Jun, YU Xizi. Semantic Segmentation of High-Resolution Remote Sensing Images Based on Improved FuseNet Combined with Atrous Convolution[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(7): 1071-1080. (in Chinese)
- [24] ZHAO H, SHI J, QI X, et al. Pyramid Scene Parsing Network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017.
- [25] CHEN L, ZHU Y, PAPANDREOU G, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation[C]//*Computer Vision - ECCV 2018*. Berlin: Springer, 2018.

作者简介

刘志恒, 男, 1990年生, 2021年获长安大学摄影测量与遥感专业工学博士学位, 讲师。主要研究方向为遥感图像处理与大数据。E-mail: liuzhiheng@xidian.edu.cn。

通讯作者

岳子腾, 男, 2001年生, 现为北京航空航天大学电子信息工程学院在读硕士研究生。主要研究方向为遥感图像处理应用。E-mail: 1248208525@qq.com。

(编辑: 夏淑密)