引文格式：丁书培,张文辉,王腾飞. 利用复合注意力网络提取遥感影像道路[J]. 地理空间信息, 2026, 24(3): 25-28. doi:10. 3969/j.issn. 1672-4623.2026.03.006

利用复合注意力网络提取遥感影像道路

丁书培，张文辉，王腾飞
（中科星图智源科技（安徽）有限公司，安徽 合肥 230031）

摘　要：道路检测在城市规划、地图更新、车辆导航等场景应用中具有重要作用。基于语义分割的数据驱动方法通常仅考虑了局部空间特征，导致道路信息提取不连续、边界呈现锯齿状，因此提出了一种改进条带注意力和复合注意力的增强架构（SA-UNet），以实现高分辨率遥感影像的精细化道路提取。首先采用ResNet进行特征提取，并在网络的中心区域引入复合注意力机制，获得道路信息的全局特征；然后在网络解码过程中采用条带注意力机制，获得道路局部信息。结果表明，SA-UNet方法在道路提取的精度和连续性方面均优于其他语义分割网络，可服务于不同的道路信息应用场景。

关键词：道路提取；条带注意力机制；复合注意力机制；高分辨率影像；深度学习
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2026）03-0025-04

Road Extraction from Remote Sensing Images Based on Composite Attention Network

DING Shupei, ZHANG Wenhui, WANG Tengfei
(Zhongke Xingtu Zhiyuan Technology (Anhui) Co., Ltd., Hefei 230031, China)

Abstract: Road detection plays a vital role in various scenarios such as urban planning, map update and vehicle navigation. The data-driven methods based on semantic segmentation usually only consider local spatial features, resulting in discontinuous road information extraction and jagged boundaries. In order to solve these problems, we proposed an improved strip attention and composite attention enhancement architecture (SA-UNet) to achieve refined road extraction from high-resolution remote sensing images. Firstly, we used ResNet to extract features, and introduced a composite attention mechanism in the central area of network to obtain the global features of road information. Then, in the network decoding process, we used a strip attention mechanism to obtain local road information. The results show that SA-UNet method is superior to other semantic segmentation networks in terms of the accuracy and continuity of road extraction and can serve different road information application scenarios.

Key words: road extraction, strip attention mechanism, composite attention mechanism, high-resolution image, deep learning

道路是现代交通基础设施建设中不可或缺的组成部分，对于引导最佳交通路线和正确的驾驶行为具有重要意义[1]。详细的路网信息为高清地图绘制和高精度地图导航提供了重要依据，因此及时更新道路网络细节以提高完整性和精度非常重要；但对于测绘人员实地调查，快速绘制复杂且大规模的道路网络地图并不是一项轻松任务。与一般分割任务相比，遥感图像道路分割具有独特的难度和挑战性[2]。随着计算机视觉图像分割技术的发展，道路提取已成为一个重要的应用领域。图像分割方法可分为传统算法和深度学习算法[3]。传统算法包括Gabor滤波器、Sobel算子、分水岭算法以及支持向量机、随机森林[4]等更先进的机器学习方法，通过提取遥感影像的纹理、边缘、形状等特征进行影像分割[5]，实现目标提取[6-7]。由于需要提取的特征十分复杂且存在较多干扰，因此传统方法难以满足道路提取需求。近年来深度学习技术在计算机视觉领域快速发展，能自动提取图像中的非线性和层次特征，更好地解决道路提取中面临的问题[8-12]。尽管现有语义分割网络取得了一定的效果，但由于道路场景的复杂性，现有道路提取方法的准确性有待进一步提高。鉴于此，本文提出了一种针对遥感影像道路特征的SA-UNet网络，在U-Net网络的基础上引入了条带注意力模块，以提高网络对道路空间特征的关注度，可有效提升网络对道路信息的提取能力，同时保证道路的连通性。

1　方法原理

1.1　整体流程

按照行政等级可将道路分为国家公路、省公路、

第一作者简介：丁书培（1996—），硕士，工程师，研究方向为遥感深度学习算法开发，E-mail：dingsp@geovis.com.cn。

县公路、乡公路、村公路（以下简称国道、省道、县道、乡道、村道）以及专用公路6个等级，等级越低的道路在影像上的特征越不显著。因此，为提取村道及以上的道路，本文提出了SA-UNet模型，以U-Net网络为基础进行修改添加模块[13]。SA-UNet模型主要由编码器、跳跃连接和解码器3部分构成。解码器以ResNet18为特征提取主干网络，由堆叠4阶段的ResBlock组成，每个ResBlock包含两个3×3卷积层、批量归一化层和一个ReLU激活层以及跳跃连接；每个阶段以特征图尺寸减半的方式进行下采样；最后一层ResBlock的特征图即CBAM_ASPP模块的输入。空洞空间金字塔池化（ASPP）模块通过不同膨胀率下的卷积操作来捕获多尺度上下文信息。在ASPP模块后添加卷积块注意力机制（CBAM）模块，可在通道和空间上提高模型对重要特征的关注度，自适应保留重要特征。编码器与解码器之间采用SAM作为跳跃连接（输入与输出维度不变），帮助二者特征的融合。解码器的4个模块均包含上采样和卷积层，上采样层将特征图分辨率放大一倍，与SAM输出结果进行拼接以实现信息的融合和重建；卷积层则对拼接后的特征图进行进一步处理，逐层恢复到输入图像大小。

1.2 复合注意力机制模块

注意力机制是一种使模型关注图像重要区域并忽略不相关区域的方法。目前CBAM[14]和ASPP[15]能很好地应用于道路提取任务中，特别是在复杂场景下；但由于道路的形状和纹理变化多样，易出现误分类和漏检问题，因此本文对二者进行有效结合，形成CBAM_ASPP模块（图1），通过引入通道注意力机制和空间注意力机制来增强特征表达能力，进一步提高道路提取的准确性和鲁棒性。

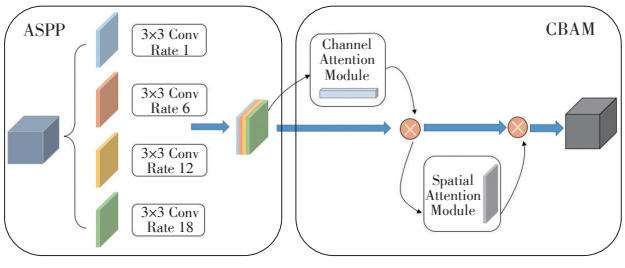


图1 CBAM_ASPP模块网络结构

1.3 条带注意力模块

道路具有不同的跨度、宽度和连续性。城市道路形状较规则，而乡村道路则弯曲无规律，普通注意力模块不能很好地捕捉各种道路特征。为更好地分割道路，网络应更加关注不同方位（行和列）的像素分布，同时关注宏观尺度的空间信息。因此，本文通过SAM模块[16]来关注每一行和每一列的像素分布，将更精细的水平、垂直信息编码为局部特征，在空间维度上沿高度或宽度维将图像的特征图分成一组条带，可有效捕捉局部上下文信息并避免其他非相关区域干扰，能捕捉道路长距离相关性，以帮助更准确的网络分割道路，提高道路分割精度。

条带注意力模块首先提取图像中的列、行像素特征，为研究列像素的分布，对输入特征映射 X 的每一行 $H\times C$ 数据进行平均池化，得到 $W\times 1$ 维列特征向量 $\boldsymbol{v}_W$，第 m 个数据点的 $\boldsymbol{v}_W$ 为：

$$\boldsymbol{v}_{W_m}=F_a\left(X_{W_m}\right)=\frac{1}{H\times C}\sum_{i=1}^{H}\sum_{j=1}^{C}X_{W_m}(i,j) \tag{1}$$

式中，F_a 为自适应平均池化；X_{W_m} 为第 m 列中所有 $H\times C$ 数据。

然后将 $\boldsymbol{v}_W$ 经过激活层和全连接层，得到新的特征向量 $\boldsymbol{v}_{W_1}$，即

$$\boldsymbol{v}_{W_1}=\sigma(W_2\delta(W_1\boldsymbol{v}_W)) \tag{2}$$

式中，δ 为ReLU函数；σ 为Sigmoid函数；$W_1\in R^{\frac{W}{r}\times W}$、$W_2\in R^{W\times\frac{W}{r}}$ 为两个全连通层操作。

最后对 $\boldsymbol{v}_{W_1}$ 进行扩维，恢复到 $H\times W\times C$ 的大小，再与输入的特征图 X 相乘，得到新的特征图 X_W。

为评估行像素特征的分布，采用相同的方法得到行注意力特征图 X_H。将 X_W 和 X_H 进行矩阵求和，得到最终的底层特征输出 X^l，有助于提高道路提取的精度。

2 实验数据与评价指标

本文在Windows 11系统通过PyCharm编译程序进行实验，采用主流的Pytorch框架完成模型构建。平台硬件主要包括Intel I9-9900KF CPU、64 GB内存和NVIDIA 2080Ti（11G）显卡。模型训练优化器均采用Adam；设置Batch_size为8、初始学习率为0.001；递减方式为等差方法；所有对比模型Epoch均为200。

2.1 数据集

为验证本文模型在不同环境中道路提取的有效性，实验采用两个公开可用的数据集，即CHN6-CUG数据集[17]是一种新型的中国代表性城市大比例尺卫星影像数据集，包含4 511幅512×512大小的标记影像，分为3 608个训练集实例和903个测试集实例，遥感

影像底图来源于谷歌地球，标注内容采用手工标注生成，包括有遮盖道路和无遮盖道路；2018年DeepGlobe机器视觉挑战赛中道路提取数据集[18]，包含6 226幅遥感影像，每幅影像大小为1 500像素×1 500像素，标签主要包括道路和背景，按照512×512大小对该数据集进行裁剪，得到24 905个样本，按照4∶1的比例随机选取训练集和测试集。两个数据集的真彩色图像和标签见图2。

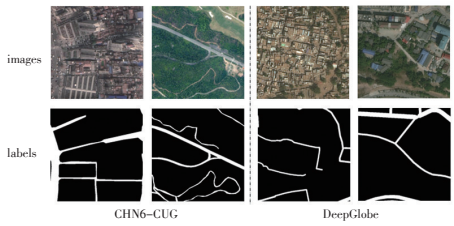


图2　CHN6-CUG数据集和DeepGlobe道路提取数据集

2.2　评价指标

道路语义分割中常用F_1-score和交并比（IoU）评估算法性能。IoU和F_1-score综合考虑了正确检测和错误检测，提供了两个综合评价指标。

3　实验与结果分析

3.1　消融实验

本文通过引入基于注意力的级联残差块来增强多尺度空间细节和语义特征。为评估本文模型的效果，在DeepGlobe数据集上设计了一个消融实验。以U-Net为基线网络，分别引入CBAM和SAM两种注意力机制模块，以探究不同模块对道路提取精度的影响。实验结果表明（表1），引入SAM模块可显著提高网络性能，相对于基线网络IoU提升了4.71%，说明该模块能使网络更加关注道路的空间位置信息，对道路的定位更加准确；引入CBAM模块后，相对于基线网络IoU提高了1.84%，说明复合注意力方法比基线网络更能关注全局特征；本文模型相较于基线网络，IoU提高了6.43%。

表1　消融实验结果

消融模型	F_1-score	IoU
U-Net	0.723 5	0.622 4
U-Net+CBAM	0.731 6	0.640 8
U-Net+SAM	0.752 4	0.669 5
本文模型	0.781 9	0.686 7

3.2　经典网络模型对比分析

为了评估SA-UNet模型在道路提取中的性能，本文通过U-Net、DeepLabV3和D-LinkNet进行比较分析。U-Net是一种典型的编码器—解码器结构；DeepLabV3采用空洞卷积池化来提取多尺度信息；D-LinKNet采用编码器—解码器结构、空洞卷积和预训练编码器进行道路提取任务，通过LinkNet架构构建，并在中心部分具有空洞卷积层。空洞卷积可在不降低特征图分辨率的情况下扩大感受野。不同网络在CHN6-CUG数据集和DeepGlobe数据集上的结果见表2，可以看出，SA-UNet模型在两个数据集上都表现出最佳性能，明显优于D-LinkNet、DeepLabV3和U-Net，证明了本文改进的网络在道路分割方面的优越性，其中用于道路设计的SAM可以更好地聚焦于道路区域，CBAM模块则能更好地提取通道信息并融合低层和高层特征以提高分割精度。

表2　不同模型结果对比

网络模型	CHN6-CUG数据集		DeepGlobe数据集	
	F_1-score	IoU	F_1-score	IoU
U-Net	0.714 6	0.628 7	0.723 5	0.622 4
DeepLabV3	0.729 3	0.629 3	0.749 2	0.643 8
D-LinkNet	0.748 2	0.650 8	0.765 0	0.661 6
SA-UNet	0.770 6	0.670 2	0.781 9	0.689 3

两个数据集下不同网络的预测结果见图3、4，可以看出，SA-UNet模型在两个数据集中的分割效果均优于其他模型，道路连通性最好，未识别道路数量最少，道路清晰度最高，与表2的数据结果相吻合。

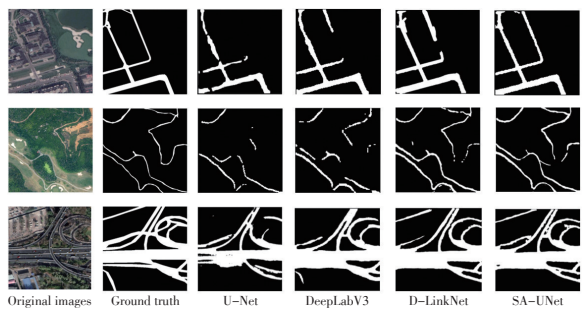


图3　CHN6-CUG数据集各模型对比图

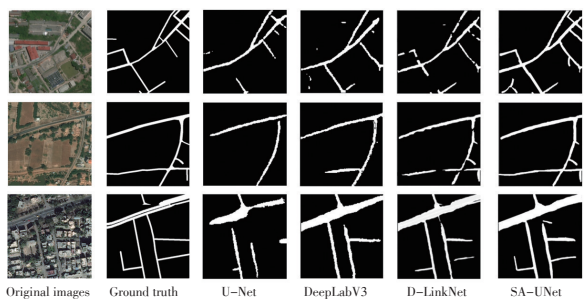


图4　DeepGlobe数据集各模型对比图

3.3　应用分析

为验证本文模型在大范围遥感影像中的适用性，

本文选取合肥市肥西县2022年空间分辨率0.8 m的高分2号卫星影像数据进行大范围遥感影像道路提取。利用滑窗推理方法进行大范围影像预测，结果见图5，可以看出，本文模型能完整提取国道、省道、县道、乡道和村道。为验证提取精度，对肥西县提取的道路影像进行人工标注，并与提取结果进行精度对比。结果表明，本文模型的F_1-score为0.76，IoU为0.67，精度与开源数据集精度相差较小，说明该模型在道路提取中具有较好的模型迁移性，能满足实际应用需求。

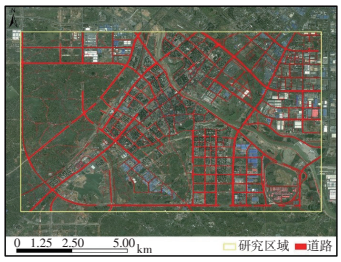


图5 肥西县道路提取结果图

4 结 语

本文提出了一种新的图像分割网络SA-UNet用于从遥感影像中提取道路，旨在解决语义分割网络在分割道路时连续性较差的问题。结果表明，SA-UNet模型引入的SAM能准确提取道路的上下文信息并对道路进行定位，从而提高道路提取的准确率；同时CBAM_ASPP模块可有效提取道路特征并融合低层和高层信息，降低错误分割率，提高分割精度。分割效果受低层特征中的位置信息和高层特征中的类别信息共同影响，二者缺一不可。因此，未来将研究位置信息和类别信息的提取和融合方法，进一步提高分割的准确性和稳定性。

参考文献

[1] 赫晓慧,陈明扬,李盼乐,等. 结合DCNN与短距条件随机场的遥感影像道路提取[J]. 武汉大学学报(信息科学版),2024,49(3):333-342

[2] 林娜,张小青,王岚,等. 空洞卷积U-Net的遥感影像道路提取方法[J]. 测绘科学,2021,46(9):109-114

[3] Mei J,Li R,Gao W,et al. CoANet: Connectivity Attention Network for Road Extraction from Satellite Imagery[J]. IEEE Transactions on Image Processing,2021,30:8540-8552

[4] 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

[5] 张丰凡,吴昊,刘明波. 基于sDNA的路网演变及其对景观格局的影响[J]. 地理空间信息,2024,22(9):67-71

[6] 王建华,秦其明,高中灵,等. 加入空间纹理信息的遥感图像道路提取[J]. 湖南大学学报(自然科学版),2016,43(4):153-156

[7] 任建平,刘勇. 基于高分辨率影像的多尺度最优分割层次模型的城市道路网提取方法[J]. 兰州大学学报(自然科学版),2019,55(1):42-50

[8] 高建文,管海燕,彭代锋,等. 基于局部—全局语义特征增强的遥感影像变化检测网络模型[J]. 地球信息科学学报,2023,25(3):625-637

[9] Xin J, Zhang X, Zhang Z, et al. Road Extraction of High-resolution Remote Sensing Images Derived from DenseUNet[J]. Remote Sensing (Basel,Switzerland),2019,11(21):2499

[10] 赵凌虎,袁希平,甘淑,等. 改进Deeplabv3+的高分辨率遥感影像道路提取模型[J]. 自然资源遥感,2023,35(1):107-114

[11] Wu Q,Luo F,Wu P,et al. Automatic Road Extraction from High-resolution Remote Sensing Images Using a Method Based on Densely Connected Spatial Feature-enhanced Pyramid[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,2021,14:3-17

[12] 龙伊娜,谷玉海,吴文昊,等. 基于改进D-Linknet的高分遥感影像道路提取方法[J]. 激光杂志,2023,44(5):162-168

[13] 申传庆,王凯,王文杰. 基于ResNet-UNet的地表覆盖自动分类技术研究[J]. 地理空间信息,2023,21(6):21-23

[14] Wang Y,Peng Y,Liu W,et al. DDU-Net: Dual-decoder-U-Net for Road Extraction Using High-resolution Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing,2022,60:1-12

[15] Liu R,Tao F,Liu X,et al. RAANet: A Residual ASPP with Attention Framework for Semantic Segmentation of High-resolution Remote Sensing Images[J]. Remote Sensing (Basel,Switzerland),2022,14(13):3109

[16] Huan H,Sheng Y,Zhang Y,et al. Strip Attention Networks for Road Extraction[J]. Remote Sensing,2022,14(18):1-14

[17] 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

[18] Xie Y,Miao F,Zhou K,et al. HsgNet: A Road Extraction Network Based on Global Perception of High-order Spatial Information[J]. ISPRS International Journal of Geo-information,2019,8(12):571-579