

文章编号:1007-2780(2025)10-1520-12

基于细节自适应和浅层引导的遥感影像道路提取

翟保明^{1,2}, 邵攀^{1,2*}, 熊彪^{1,2}, 漆琛巍^{1,2}, 高玉倩^{1,2}, 李净漪²

(1. 水电工程智能视觉监测湖北省重点实验室, 湖北 宜昌 443002;

2. 三峡大学 计算机与信息学院, 湖北 宜昌 443002)

摘要:针对道路具有多尺度多方向特性和细节信息容易丢失两个道路提取难题,提出一种基于细节自适应增强和浅层特征引导的遥感影像道路提取网络。首先,借助条形卷积、注意力机制和浅层特征引导策略,提出一种浅层特征引导注意力模块,增强网络对空间细节信息重构能力;然后,基于Gabor卷积、局部池化和全局方向池化,设计一种多尺度多方向特征融合模块,提升网络对不同尺度和不同方向道路特征的表征能力;最后,基于高斯滤波、差值运算、蛇形卷积和标准卷积,构建一种细节自适应增强模块,实现对细节信息的自适应增强。所提网络在两个常用公开道路提取数据集CHN6-CUG和Massachusetts上的F1分数和IoU达到74.62%、62.38%和77.86%、65.17%,相比于8个对比方法,在CHN6-CUG和Massachusetts上均取得了最高的F1分数和IoU,F1分数和IoU分别至少提高了1.41%、1.32%和1.04%、0.88%。与其他先进方法相比,本文模型在道路提取任务中表现更优,在保持较高提取精度的同时具备较低的参数量,展现出良好的性能与效率平衡。

关键词:遥感影像;道路提取;多尺度;注意力;多向卷积

中图分类号:TP751.1; TP391 **文献标识码:**A **doi:**10.37188/CJLCD.2025-0131 **CSTR:**32172.14.CJLCD.2025-0131

Road extraction from remote sensing images based on shallow guidance and detail-adaptive

ZHAI Baoming^{1,2}, SHAO Pan^{1,2*}, XIONG Biao^{1,2}, QI Chenwei^{1,2}, GAO Yuqian^{1,2}, LI Jingyi²

(1. Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, Yichang 443002, China;

2. College of Computer & Information Technology, Yichang 443002, China)

Abstract: To address two key challenges—namely, the multi-scale and multi-directional nature of roads and the frequent loss of fine-grained details, we propose a novel remote sensing road extraction network based on detail-adaptive and shallow feature guidance. Firstly, we design a shallow feature-guided attention module, which integrates strip convolutions, attention mechanisms, and shallow feature guidance to enhance the reconstruction of detailed information. Then, a multi-scale and multi-directional feature fusion

收稿日期:2025-06-23;**修订日期:**2025-08-02.

基金项目:国家自然科学基金青年基金(No.41901341);湖北省自然科学基金面上项目(No.2024AFB867;No.2024AFB217);自然资源部地理国情监测重点实验室开放基金项目(No.2025NGCM03)

Supported by National Natural Science Foundation of China (No.41901341); Natural Science Foundation of Hubei Province (No.2024AFB867; No.2024AFB217); Open Fund Project of the Key Laboratory of Geo-Environmental Monitoring, Ministry of Natural Resources (No.2025NGCM03)

*通信联系人, E-mail: panshao@whu.edu.cn

module is developed, leveraging Gabor convolutions, local pooling, and global directional pooling to improve the representation of road features at various scales and directions. Finally, a detail-adaptive enhancement module is constructed using Gaussian filtering, interpolation, serpentine convolution, and standard convolution, which adaptively enhances the detailed information. The proposed network achieves F1-scores and IoU values of 74.62% and 62.38% on the CHN6-CUG dataset and 77.86% and 65.17% on the Massachusetts dataset, outperforming eight existing methods by at least 1.41% and 1.32% and 1.04% and 0.88%, respectively. Compared with other advanced methods, the proposed model demonstrates superior performance in road extraction tasks, maintaining high extraction accuracy while having a lower parameter count, showing a good balance between performance and efficiency.

Key words: remote sensing image; road extraction; multi-scale; attention; multi-directional convolution

1 引言

随着高分辨率卫星遥感技术的快速发展,道路提取已成为遥感影像智能解译的关键任务,在智慧交通^[1](如无人驾驶^[2]、车辆导航^[3])、城市规划^[4]等领域具有重要应用价值^[5]。当前道路提取方法主要可分为传统道路提取和基于深度学习的道路提取两类^[6]。传统的道路提取方法主要依赖于遥感影像的光谱、纹理等信息,并通过手工设计特征来实现。然而手工设计特征对影像的表征能力较弱且鲁棒性较差,在某些情况下道路提取效果不理想。

鉴于深度学习在计算机视觉领域的巨大成功,其在遥感领域受到越来越多的关注,已有许多研究人员将深度学习中的语义分割技术应用到遥感影像道路提取^[7]。李朝奎^[8]等人将 FCN 应用到遥感影像道路提取,充分利用像素的位置信息,从而显著提高了道路提取的准确性和鲁棒性。Li^[9]等人提出将两个 U-Net 结构进行串联,通过强化跳跃连接以提升特征传递效率,从而更有效地保留边缘信息。Liu^[10]等人提出了一种基于生成方法的遥感图像增强策略,充分利用道路像素间的一致性提升整体分割精度。受深度残差学习和 U-Net 的启发,Zhang^[11]等人提出了一种深度残差 U-Net 结构用于道路提取。Qu^[12]等人将空洞空间池化金字塔引入到 LinkNet 的瓶颈结构中,提取道路的多尺度特征并扩大特征感受野。Zhou^[13]等人将串并联的扩张卷积引入到 LinkNet 的瓶颈结构中,提出 D-LinkNet 网络。Wang^[14]等人和 Yang^[15]等人分别利用条形卷积和多尺度条形卷积来提升 D-LinkNet 对道路的重构能力,分

别提出了 MSMDFFNet 和 OARENet。

针对 CNN 感受野受限问题,遥感学者将注意力机制引入 CNN 道路提取建模影像全局信息,提升道路提取网络的性能。Wang^[16]等人通过在编码器和解码器之间引入卷积块注意力模块(Convolutional Block Attention Module, CBAM),提出 DDU-Net 道路提取网络,实现了在空间与通道维度的双重特征校准。Lu^[17]等人通过在编码器之间引入全局感知注意力模块,提出了 GAMSNet 网络,实现了在空间与通道维度建模上下文信息。Gao^[18]等人在 D-LinkNet 的基础上引入 DANet 的双注意力模块来强化特征学习,提出了 DAD-LinkNet 网络。Wang^[19]等人在 D-LinkNet 的基础上引入非局部注意力模块来增强全局上下文建模能力,提出了 NL-LinkNet 道路提取网络。

最近,一些遥感学者将 Transformer 集成到 CNN 道路提取方向中。Dai^[20]等人将 Transformer 模块插入到编码器和解码器之间,结合 Transformer 模块和可变形卷积实现了局部几何细节与全局上下文信息协同学习,提出道路提取模型 RADANet。Wang^[21]等人先利用残差网络提取特征,然后通过并行的条形 Transformer 模块在高和宽的方向上建模全局上下文信息。Meng^[22]等人分别将条形 Transformer 模块和 Transformer 模块串联到 D-LinkNet 的瓶颈结构中,从不同角度建模全局相关性。

尽管现有深度学习方法在道路提取精度上取得了显著进展,但由于道路提取面临诸多挑战,它仍然是一个未完全解决的问题。所面临的挑战包括:(1)在编码器的特征传递过程中,多采

用简单的级联或相加方式进行特征融合,或直接迁移在其他任务中表现优异的注意力机制,而未充分考虑道路的独特性质及不同层级特征之间的指导或交互;(2)道路在图像中表现出不规则性,如宽度、形态和方向的变化,现有方法仅通过多尺度特征提取应对尺度变化,而未充分考虑道路的多方向特征;(3)在编码提取过程中经过多次的下采样导致一些边界模糊和细节信息的丢失。

为克服上述挑战,本文提出一种细节自适应增强(Detail-Adaptive Enhancement, DAE)和浅层特征引导(Shallow Feature Guidance, SFG)道路提取网络 DSNet。首先,浅层特征引导注意力模块(Shallow Features Guidance Attention Module, SFGAM)通过主辅引导机制,将浅层特征作为辅助,引导深层特征的恢复,从而弥补深层特征在细节表达上的不足。其次,设计一种多尺度多方向特征融合模块(Multi-Scale and Multi-Directional

Feature Fusion Module, MSMDFFM),该模块通过提取局部多尺度多方向细节信息和全局多方向结构信息,增强了模型的多样性和鲁棒性。最后,设计了细节自适应增强模块(Detail-Adaptive Enhancement Module, DAEM),利用高斯滤波、重采样和差值运算等实现细节区域的自适应定位,并通过多种卷积操作实现细节信息的增强,使得网络能够精准恢复道路边界和小尺度结构特征。

2 本文方法

如图 1 所示,所提出的 DSNet 网络采用 D-LinkNet 作为基础架构,主要改进包括浅层特征引导注意力模块(SFGAM)、多尺度多方向特征融合模块(MSMDFFM)和细节自适应增强模块(DAEM),其余结构参数与 D-LinkNet 保持一致。

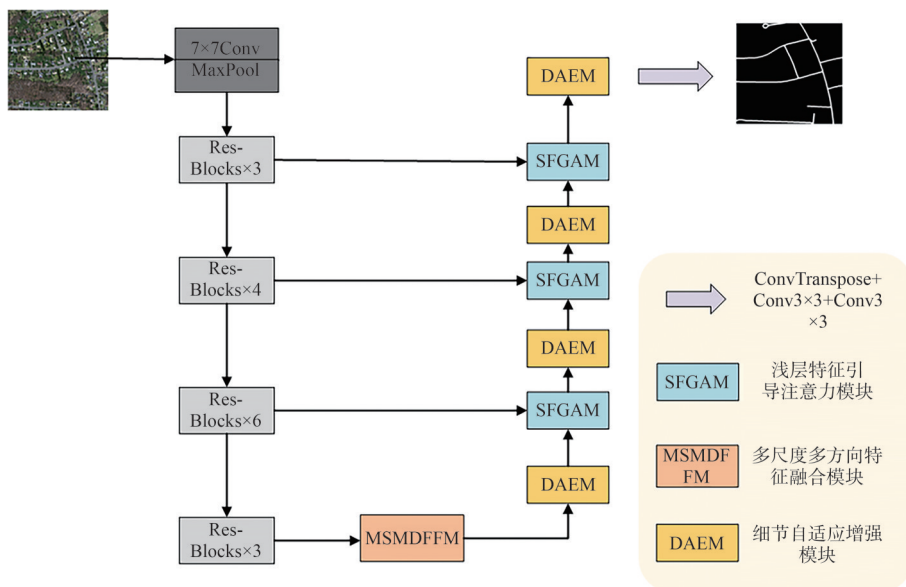


图 1 DSNet 的总体网络结构

Fig. 1 Overall structure of DSNet

2.1 浅层特征引导注意力模块

现有道路提取方法(如 D-LinkNet)通常在解码阶段通过相加、级联或引入注意力机制融合不同层级特征,以增强空间细节信息。然而,这些融合方式往往将不同层级特征同等对待,忽视了其在语义层次和空间细节表达上的差异,导致其有时效果欠佳。

为解决上述问题,本研究提出一种浅层特征引导注意力模块(SFGAM)来增强解码过程中的空间细节信息。SFGAM 具有如下优势和创新:(1)SFGAM 采用主辅引导机制,将浅层特征作为辅助,引导深层特征的恢复。该机制使浅层特征为深层特征提供精细的空间细节信息,从而弥补深层特征在细节表达上的不足,提升特征融合

效果。(2)通道注意力引导在通道维度上对浅层与深层特征进行级联,以保留多层级语义信息,并引入双分支结构,相较于传统的单分支结构,更有利于细节信息的保持。(3)空间注意力引导的目标是突出关键位置,通过对浅层与深层特征在空间维度进行逐像素相加,增强在两者中均具有较强响应的位置,从而突出关键空间区域。(4)为了更好地适应道路的几何特性,SFGAM采用条形卷积核替代传统的方形卷积核,使得卷积操

作更自然地符合道路的细长结构。

如图 2 所示,在通道引导过程中,首先将 E_{i-1} 和 D_i 进行级联,并分别进行全局平均池化和全局最大池化,以提取全局信息。随后,这两组特征分别经过一个 1×1 卷积将通道数减少,Relu 函数以及第二个 1×1 卷积将通道数再次恢复。将两个分支进行相加,采用一个 1×1 卷积调整通道数,再经过 Sigmoid 激活,以获得通道特征关系的权重系数。

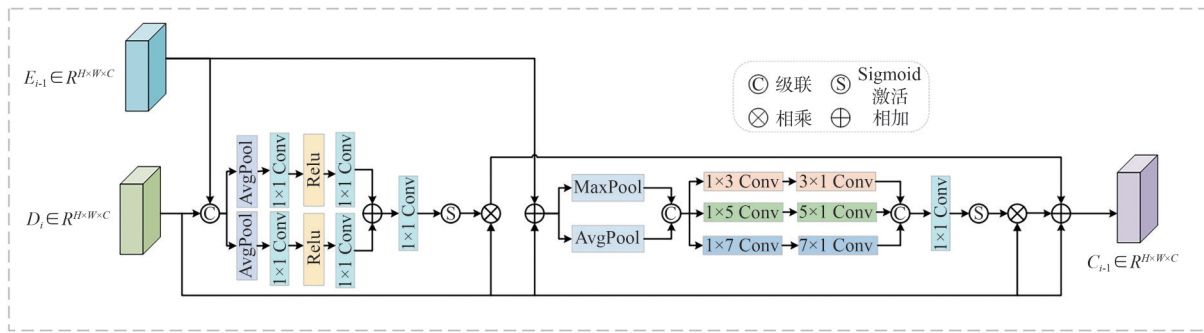


图 2 浅层特征引导注意力模块

Fig. 2 Shallow features guidance attention module

最终,将该权重与 D_i 相乘,得到增强后的特征 F_c 。这种方式能够强化网络对道路相关通道特征的注意,同时有效过滤背景中的干扰信息。其数学表达如式(1)~式(3)所示:

$$F_a = \text{Conv}_{1 \times 1}(\gamma(\text{Conv}_{1 \times 1}(\text{GAP} \times (\text{Cat}(E_{i-1}, D_i))))), \quad (1)$$

$$F_m = \text{Conv}_{1 \times 1}(\gamma(\text{Conv}_{1 \times 1}(\text{GMP} \times (\text{Cat}(E_{i-1}, D_i))))), \quad (2)$$

$$F_c = \sigma(\text{Conv}_{1 \times 1}(F_a + F_m)) \times D_i, \quad (3)$$

其中: Cat 是级联操作, GAP 是全局平均池化, GMP 是全局最大池化, Conv 是卷积操作, γ 是 Relu 激活函数, σ 是 Sigmoid 激活函数。

在空间引导过程中,首先将 E_{i-1} 和 D_i 进行相加,分别通过平均池化和最大池化并在通道上进行级联,并通过多尺度条形卷积提取道路的线性特征,以更有效地拟合道路形态。最后,经过一个 1×1 卷积调整特征分布,并通过 Sigmoid 计算注意力权重,最终与 D_i 相乘,以强化空间特征表达。最后再将两个注意增强和 D_i 相加得到最后的输出结果 C_{i-1} ,其数学表达如式(4)~式(8)所示:

$$F_i = \text{Cat}(\text{MaxPool}(E_{i-1} + D_i), \text{AvgPool}(E_{i-1} + D_i)), \quad (4)$$

$$F_1 = \text{Conv}_{1 \times 3}(\text{Conv}_{3 \times 1}(F_i)), \quad (5)$$

$$F_2 = \text{Conv}_{1 \times 5}(\text{Conv}_{5 \times 1}(F_i)), \quad (6)$$

$$F_3 = \text{Conv}_{1 \times 7}(\text{Conv}_{7 \times 1}(F_i)), \quad (7)$$

$$F_s = D_i \times \sigma(\text{Conv}_{1 \times 1}(\text{Cat}(F_1, F_2, F_3))), \quad (8)$$

其中: Cat 是级联操作, MaxPool 是最大池化, AvgPool 是平均池化, Conv 是卷积操作, σ 是 Sigmoid 激活函数。

2.2 多尺度多方向特征融合模块

在遥感图像中,道路通常同时具有多尺度与多方向的结构特性。然而,现有方法通常通过级联空洞卷积或者空间金字塔池化只考虑多尺度特性,未考虑道路的多方向特性。为此,本文提出了多尺度多方向特征融合模块(MSMDFFM),旨在同时捕捉局部多尺度多方向细节与全局多方向结构信息。

如图 3 所示,MSMDFFM 具有如下创新和贡献:(1)同时考虑多尺度多方向:首先通过不同参数的平均池化操作引入多尺度信息,随后在不同尺度分支中结合 Gabor 滤波器执行多方向卷积,精细提取具有方向性的局部结构特征。(2)多方向全局平均池化:通过在不同方向上执行全局平均池化实现全局特征压缩,从而获取道路在不同

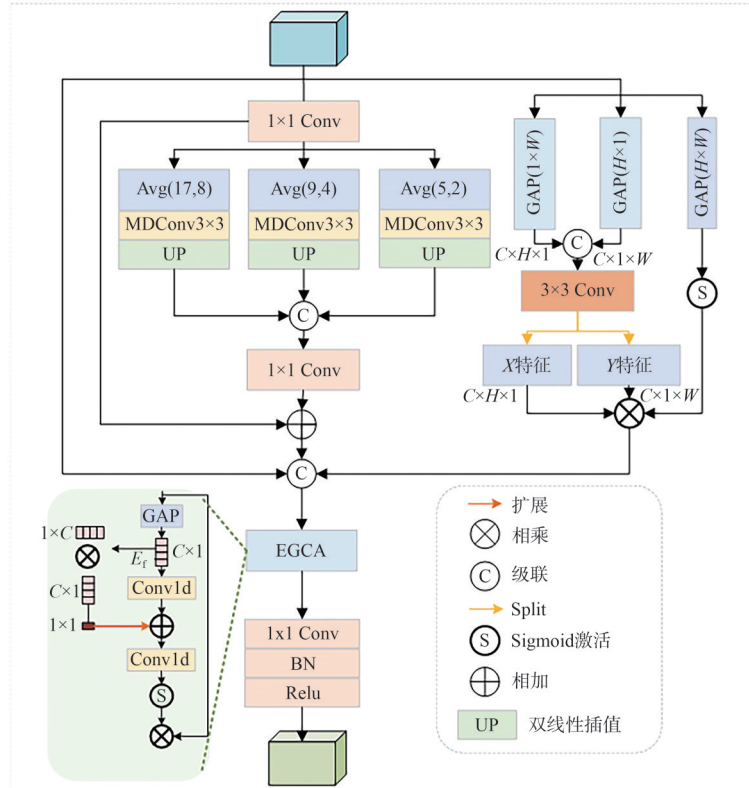


图3 多尺度多方向特征融合模块

Fig. 3 Multi-scale and multi-directional feature fusion module

方向上的整体结构与全局延展方向。(3)局部与全局信息的融合:为更充分融合局部多方向细节与全局结构特征,本文在ECA^[23]基础上提出了改进的高效全局通道注意力模块(EGCA)。原始ECA模块通过自适应调整通道权重,增强通道间的相关性,有效提升通道维度的特征表达能力。然而,由于其仅依赖一维卷积建模局部通道关系,缺乏通道维度的全局建模能力,难以充分表征遥感图像中道路所具有的长距离依赖语义结构特征。为弥补上述局限,EGCA引入特征转置与矩阵乘操作,以显式建模通道间的全局相关性,从而显著增强模型对通道维度长距离依赖的表达能力。

多方向性是道路提取的一个重要特征。为了获取道路中复杂的方向特征,本文设计了多向卷积,在单个感受野内捕捉多方向道路信息。多向卷积由一组Gabor滤波器组成,广泛应用于2D图像处理,其数学表达如式(9)和式(10)所示:

$$G(x, y; \lambda, \theta, \varphi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \times \exp\left(i\left(2\pi \frac{x'}{\lambda}\right) + \varphi\right), \quad (9)$$

$$x' = x \cos \theta + y \sin \theta,$$

$$y' = -x \sin \theta + y \cos \theta, \quad (10)$$

式中: x 和 y 分别代表像素的坐标; λ 是波长, $1/\lambda$ 表示余弦函数的空间频率; γ 是空间长宽比; φ 是相位偏移; σ 是标准差。根据不同的 θ ,可以获取不同方向的Gabor滤波器。

相较于传统将多方向Gabor卷积核通过相加合并为一个统一卷积核的方式,本文遵循Chen^[24]等人的工作,通过使用注意力来学习每个定向Gabor内核的权重系数来集成更加适合道路方向的卷积算子。

$$\text{MDConv} = \sum_{i=1}^k w_i (K \times G(\theta_i)), \quad (11)$$

式中: w_i 是注意力模块学习的权重系数, K 是普通卷积核, $G(\theta_i)$ 是Gabor内核。

在全局线性分支中,首先分别沿水平方向和垂直方向全局平均池化进行方向感知来提取道路的线性结构特征,再对输入全局平均池化生成上下文信息并进行Sigmoid激活得到权重 E_c 。随后,将水平和垂直池化后的特征进行级联并使用 3×3 卷积进一步增强局部空间细节,然后分离成

水平特征 E_w 和垂直特征 E_h , 最后将 E_c 、 E_h 和 E_w 进行相乘得到线性特征 E_x 。

$$E_s = \text{Conv}_{3 \times 3} \left(\text{Cat} \left(\frac{1}{W} \sum_{0 \leq i < W} E_i(h, i), \frac{1}{H} \sum_{0 \leq j < H} E_i(j, W) \right) \right), \quad (12)$$

$$E_c = \sigma(\text{GAP}(x)), E_h = \text{Sp}(E_s), E_w = \text{Sp}(E_s), \quad (13)$$

其中: Cat 是级联操作, Sp 是 Split。

在局部多尺度多方向分支中, 首先对输入特征图 E_i 进行 1×1 卷积降维来减少计算量。随后, 采用 3 种不同尺度的池化和 3×3 多向卷积来提取不同尺度下的局部信息, 这种方式能够增强模型对复杂道路结构的感知能力。在此基础上, 对各个分支的特征图进行双线性插值以恢复到原始分辨率, 并通过 1×1 卷积调整通道数。最终, 将处理后的特征图与输入特征图进行逐元素相加, 得到融合多尺度和多方向信息的特征图 E_s 。

$$E_1 = \text{up}(\text{MDConv}_{3 \times 3}(\text{Avgpool}_{5,2} \times (\text{Conv}_{1 \times 1}(E_i))), \quad (14)$$

$$E_2 = \text{up}(\text{MDConv}_{3 \times 3}(\text{Avgpool}_{9,4} \times (\text{Conv}_{1 \times 1}(E_i))), \quad (15)$$

$$E_3 = \text{up}(\text{MDConv}_{3 \times 3}(\text{Avgpool}_{17,8} \times (\text{Conv}_{1 \times 1}(E_i))), \quad (16)$$

$$E_s = \text{Conv}_{1 \times 1}(E_i) + \text{Conv}_{1 \times 1}(\text{Cat}(E_1, E_2, E_3)), \quad (17)$$

其中: Cat 是级联操作; AvgPool _{i,j} 是平均池化, i 是卷积核大小, j 是步距; up 是双线性插值。

EGCA 模块旨在增强通道维度的全局建模能力, 从而挖掘更多与道路结构相关的判别性通道信息。首先, EGCA 通过全局平均池化操作提取空间维度上的全局上下文信息。随后, 利用一维卷积实现相邻通道间的局部交互, 以建模通道维度的局部依赖关系。为进一步建模通道间的全局依赖性, 对特征 E_i 进行转置并执行矩阵乘操作, 得到通道间全局相关性矩阵, 用于显式捕捉各通道之间的长距离依赖结构。最终, 将局部一维卷积输出与全局相关性结果进行融合, 并通过卷积层和激活函数生成通道注意力权重, 进而对输入特征图进行通道加权。

$$E_i = \text{Gap}(\text{Cat}(E_x, E_s, E_i)), \quad (18)$$

$$E_{\text{out}} = E_f \times \text{Conv}_{1D}(E_f \times E_f^T + \text{Conv}_{1D}(E_f)), \quad (19)$$

其中: Cat 是级联操作, GAP 是全局平均池化, Conv 是卷积操作, γ 是 Relu 激活函数, σ 是 Sigmoid 激活函数。

2.3 细节自适应增强模块

在道路提取任务中, 解码器需逐步将低分辨率的特征恢复为高分辨率特征, 以准确重建道路结构。然而, 在编码阶段, 由于连续的下采样操作, 一些空间细节信息(尤其是道路边界和小尺度结构的特征)被压缩或丢失, 容易出现边界模糊和信息遗漏问题。传统道路提取方法在解码阶段缺乏针对细节的显式处理机制, 一定程度上影响了模型对道路细节信息的重构能力。为缓解这一问题, 本文提出细节自适应增强模块(DAEM), 该模块能够自适应地定位并强化解码过程中的细节区域特征, 有效增强网络对道路边界和小尺度区域的感知能力。

如图 4 所示, DAEM 包括细节区域自适应定位与细节区域增强两部分。为有效定位细节区域, DAEM 首先通过高斯滤波和重采样操作生成去除细节信息的平滑特征, 然后基于原始特征与平滑特征通过做差操作和阈值函数自适应定位细节区域。步骤如下: (1) 采用高斯滤波对输入特征进行平滑处理, 以抑制边缘与纹理等高频信息, 生成初步平滑特征; (2) 通过最大池化下采样和双线性插值上采样对初步平滑特征进行重构, 进一步去除特征细节信息, 生成最终的平滑特征; (3) 通过对平滑特征与原始特征进行做差运算生成差值特征, 初步定位细节区域; (4) 通过对差值特征进行阈值运算生成二值特征 $\text{Sign}(F_x)$ 来去除噪声区域, 生成更准确的细节区域。定位细节区域具体计算公式如式(20)~式(22)所示:

$$F_u = \text{up}(\text{MaxPool}(\text{Gauss}(F_1))), \quad (20)$$

$$F_x = |F_1 - F_u|, \quad (21)$$

$$F_s = F_1 \times \text{Sign}(F_x),$$

$$F_n = F_1 \times (1 - \text{Sign}(F_x)), \quad (22)$$

式中: MaxPool 为最大池化, up 为双线性插值。

自适应定位细节区域后, DAEM 通过在标准卷积、蛇形卷积和多方向卷积三支卷积结构增强细节区域: 标准卷积提取基础局部纹理, 蛇形卷积凭借非对称感受野提升对细长、弯曲结构的

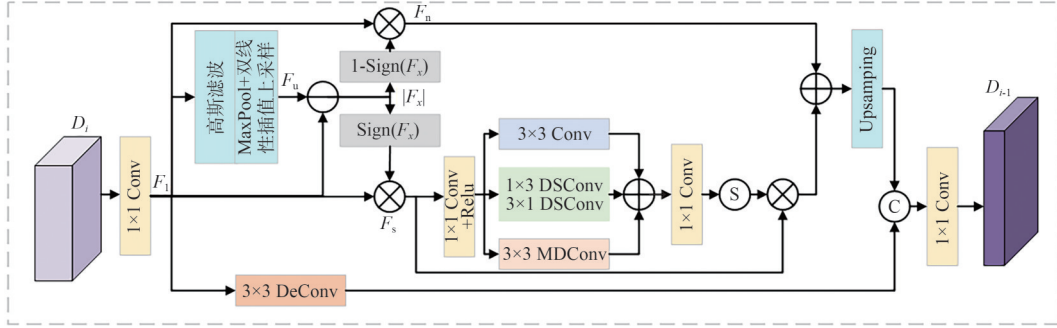


图4 细节自适应增强模块

Fig. 4 Detail-adaptive enhancement module

建模能力,多方向卷积则通过多方向卷积核捕捉道路方向的多样性。三者互为补充,使网络能够自适应感知并增强不同形态的道路细节特征。3个分支的输出通过加权融合后,再经 1×1 卷积调整通道数和Sigmoid激活与 F_s 相乘得到细节增强后的特征图 F_s' 。

$$D_1 = \text{Conv}_{3 \times 3}(\gamma(\text{Conv}_{1 \times 1}(F_s))), \quad (23)$$

$$D_2 = \text{DSCConv}_{3 \times 1}(\text{DSCConv}_{1 \times 3}(\gamma \times (\text{Conv}_{1 \times 1}(F_s)))), \quad (24)$$

$$D_3 = \text{MDConv}_{3 \times 3}(\gamma(\text{Conv}_{1 \times 1}(F_s))), \quad (25)$$

$$F_s = D_i \times \sigma(\text{Conv}_{1 \times 1}(\text{Cat}(D_1, D_2, D_3))), \quad (26)$$

式中:DSCConv是蛇形卷积, $\times$ 是矩阵相乘, γ 是Relu激活函数, σ 是Sigmoid激活函数。

最后将 F_s' 和 F_n 相加上采样和常规上采样分支进行级联,并经过 1×1 卷积得到最后的输出 D_{i-1} 。

3 实验结果及分析

3.1 实验数据与实验设计

CHN6-CUG是中国具有代表性的大型城市卫星遥感影像道路数据集,包含了6个中国具有代表性的城市。数据集收录了4 511张遥感图像,尺寸为 $512 \text{ 像素} \times 512 \text{ 像素}$,空间分辨率 0.5 m 。3 608用于模型训练,903张用于测试和评估。

Massachusetts道路数据集涵盖了城市、郊区和农村地区多种场景遥感道路数据集,数据集收录了1 171张遥感图像,尺寸为 $1 500 \text{ 像素} \times 1 500 \text{ 像素}$,空间分辨率 1 m 。数据集包含1 108个训练影像,14个验证影像和49个测试影像。

为确保实验的公正性与数据的完整性,本研

究严格采用数据集官方提供的划分方案,且未对数据样本进行任何筛选或修改。考虑到设备的GPU资源限制,采用非重叠裁剪方式将原始影像裁剪为 $512 \text{ 像素} \times 512 \text{ 像素}$ 大小,对于不足该尺寸的边缘区域,通过边缘像素填充的方式进行补齐。为提升模型的泛化能力与鲁棒性,对训练集样本进行了多种数据增强操作,包括水平翻转、垂直翻转及尺度缩放等。

模型训练采用Adam优化器,初始学习率设为 $0.000 2$,批量大小(Batch Size)为4,最大训练轮数(Epoch)为200。在训练过程中,若损失在连续6个epoch内未下降,则提前终止训练。本研究所提方法已在GitHub上开源,链接如下:<https://github.com/mingasx/DSNet>。

为了验证本文所提出网络的有效性,采用精准率(Precision, P)、召回率(Recall, R)、F1分数(F1)和交并比(Intersection over Union, IoU)作为评价指标。

3.2 实验结果与分析

图5和图6分别给出CHN6-CUG和Massachusetts数据集上5组道路提取结果图:(a)和(b)分别代表原图和真实标签;(c)~(k)分别表示CARNet^[25]、RCFSNet^[26]、MSMDFFNet^[14]、DBRANet^[27]、D-LinkNet^[13]、TransRoadNet^[28]、HSN-Net^[1]、SwinMSNet^[4]和所提出的方法的道路提取图,黑色区域是背景区域,白色区域是道路区域,红色区域是漏检区域,绿色区域是虚检区域。

从图5可知,对于CHN6-CUG数据集,本文方法在多数场景中都获得了与真实标签最接近的道路提取结果,在不同尺度的道路识别方面表现出较强的稳定性与完整性,且边界提取更为连

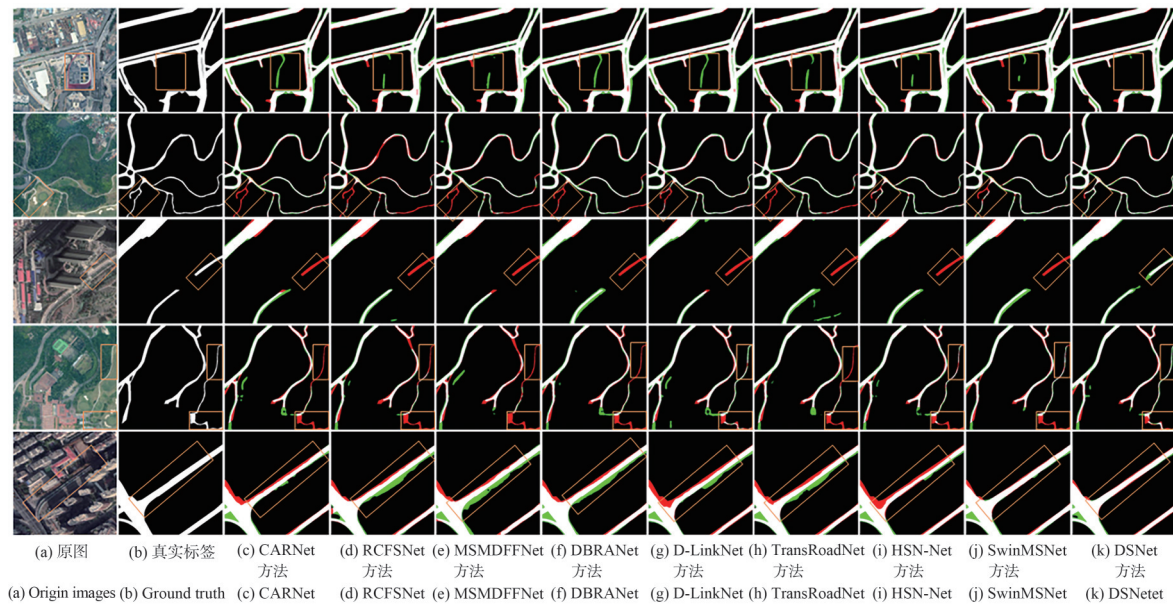


图 5 不同方法在 CHN6-CUG 数据集上的道路提取结果

Fig. 5 Road extraction results from different methods on CHN6-CUG dataset

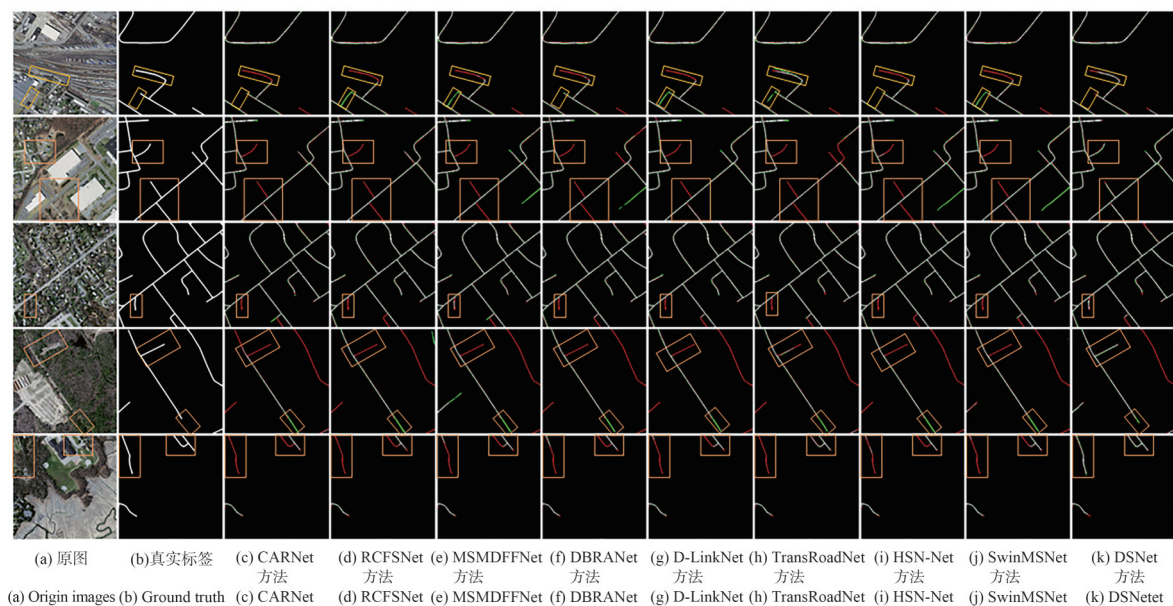


图 6 不同方法在 Massachusetts 数据集上的道路提取结果

Fig. 6 Road extraction results from different methods on Massachusetts dataset

续、准确,表明其在适应复杂道路结构方面具有显著优势。相比之下,其余8种对比方法存在明显不足。例如在第1行所示区域,大多数对比方法存在明显虚检现象,而本文方法仅出现少量误判,显著提升了提取精度;在第2行和第4行中,面对多方向、交叉密集的道路结构,其他方法普遍出现断裂提取,而本文方法能有效保持道路结构的连贯性;在第3行中,除本文方法外,其他

方法均存在不同程度的断裂和虚检;在第5行区域,本文方法在道路边界处的表现更为完整,显著减少了误检与漏检。这主要得益于本文引入的MSMDFFM模块在特征学习过程中强化了多尺度和多方向上下文的建模能力。从图6可知,对于Massachusetts数据集,本文方法在视觉效果上依然优于其他方法。在第1行中,虽然多数方法能够提取出道路主干结构,但本文方法在

轮廓完整性和边界精度方面表现更佳;在第 2 行中,MSMDFFNet 与 DBRANet 出现明显虚检;RCFSNet 和 TransRoadNet 漏检了较细的支路;CARNet 虽取得与真实标签较为接近的效果,但在边界细节处理上略逊一筹。在第 3 行中,尽管背景信息复杂多变,本文方法仍准确还原了大部分道路区域,展现出对噪声干扰的良好鲁棒性,这得益于 SFGAM 模块增强了对空间细节信息的建模能力;在第 4 行和第 5 行中,针对被植被覆盖的区域,本文方法所提取的道路结构更为连续完整,充分说明 DAEM 模块在边界和细小道路区域提升了特征重建能力,从而增强了整体提取效果。

为更加客观地评价不同道路提取方法的提取性能,表 1 给出不同方法在 CHN6-CUG 和 Massachusetts 数据集上道路提取结果的定量评价指标。从表 1 可知,本文方法 DSNet 在两个数据集上的 F1 分数和 IoU 均得到最优定量指标,显著优于现有主流方法。具体而言,在 Massachusetts 数据集上,所提出方法的 F1 分数和 IoU 分别达到 77.86% 和 65.17%,相比其他对比方法,至少提升了 1.04% 和 0.88%。在 CHN6-CUG 数据集上,所提出方法的 F1 分数和 IoU 分别达到 74.62% 和 62.38%,相比其他对比方法,至少提升了 1.41% 和 1.32%。

表 1 不同模型道路提取结果的定量分析指标

Tab. 1 Quantitative indices of the road extraction results of different models %

Models	CHN6-CUG				Massachusetts			
	Precision	Recall	F1	IoU	Precision	Recall	F1	IoU
CARNet	76.15	72.56	72.24	60.12	80.37	74.65	75.77	63.33
RCFSNet	77.98	67.21	69.37	56.37	80.25	73.63	75.54	62.98
MSMDFFNet	78.24	68.15	71.27	58.85	80.04	73.45	75.66	63.18
DBRANet	74.64	73.46	71.63	59.91	79.97	74.15	75.74	63.28
D-LinkNet	74.93	73.22	71.83	59.72	79.60	74.40	75.74	63.15
TransRoadNet	75.60	72.85	72.04	59.85	79.65	74.09	75.41	62.94
HSN-Net	77.01	72.37	72.73	60.54	81.11	75.05	76.82	64.29
SwinMSNet	77.38	73.52	73.21	61.06	81.13	74.95	76.74	64.08
Ours	78.51	74.73	74.62	62.38	81.95	75.72	77.86	65.17

注: 粗体数字代表最优结果

3.3 消融实验

为了验证本文方法中多方向特征融合模块(MSMDFFM)、细节自适应增强模块(DAEM)和浅层特征引导注意力模块(SFGAM)的有效性,在 CHN6-CUG 和 Massachusetts 数据集上对 3 个模块进行消融实验,实验结果如表 2 所示,其中“√”表示模块保留,“×”表示模块去掉。

从表 2 实验结果可知,移除本文中的任一模

块,都会导致道路提取指标明显下降,这进一步验证了 SFGAM、MSMDFFM 和 DAEM 在提升道路提取精度方面的有效性。在 3 组消融实验中,去掉 SFGAM,两个数据集的 F1 分数分别下降了 0.81% 和 0.48%,IoU 分别下降了 0.65% 和 0.7%,说明浅层引导机制对细节信息的增强具有显著贡献。去掉 MSMDFFM,两个数据集的 F1 分数分别下降了 0.93% 和 0.71%,IoU 分别下

表 2 网络结构消融实验结果

Tab. 2 Ablation experiment results of DSNet structure %

SFGAM	MSMDFFM	DAEM	CHN6-CUG		Massachusetts	
			F1	IoU	F1	IoU
×	√	√	73.81	61.73	77.38	64.47
√	×	√	73.69	61.27	77.15	64.56
√	√	×	73.57	61.45	77.21	64.62
√	√	√	74.62	62.38	77.86	65.17

降了 1.11% 和 0.61%,说明多尺度多方向结构在捕捉复杂道路形态方面具有关键作用。去掉 DAEM,两个数据集的 F1 分数分别下降了 1.07% 和 0.65%,IoU 分别下降了 0.93% 和 0.55%,说明细节自适应增强模块对道路边缘结构和细节信息的重建具有明显优势。

3.4 复杂度分析

本节分析不同模型的模型复杂度,包括模型的参数量(Parameters,Params)和浮点运算次数(Floating-point Operations,FLOPs),具体数值如表 3 所示。从表 3 可知,所提模型的 Params 介于多个现有方法之间,略高于 D-LinkNet,低于 RCFSNet、MSMDFFNet 等模型。在 FLOPs 指标上,所提模型明显低于 RCFSNet 和 MSMDFFNet,也优于计算开销较大的 TransRoadNet,仅略高于部分轻量化模型如 D-LinkNet 和 CARNet,展现

出较优的计算效率。
尽管在模型复杂度上本文方法并非最优,但与 8 种不同方法相比总体偏低。并且在道路提取精度上,本文方法明显优于所有对比方法。

4 结 论

针对高分辨率遥感影像中的道路提取任务,本文提出了一种细节自适应增强与浅层特征引导相结合的道路提取网络——DSNet。该网络通过集成 3 个关键模块:浅层特征引导注意力模块(SFGAM)、多尺度多方向特征融合模块(MSMDFFM)以及细节自适应增强模块(DAEM),显著提升了对复杂道路结构的建模能力和细节还原效果。具体而言,SFGAM 引入了浅层特征引导策略,该策略使浅层特征为深层特征提供精细的空间细节信息,从而弥补深层特征在细节表达上的不足;MSMDFFM 通过结合多方向池化与 Gabor 卷积机制,全面提升了网络对多尺度多方向道路特征的表征能力,特别适用于处理复杂交汇点和曲折道路等结构;DAEM 则利用高斯滤波、重采样与差值操作实现对细节区域的自适应感知,并结合标准卷积、蛇形卷积和多方向卷积,进一步增强对边界细节与细小道路结构的提取精度。

在 CHN6-CUG 和 Massachusetts 两组数据集的实验结果表明,DSNet 的提取结果均优于 8 种主流对比方法,在 F1/IoU 指标上相较于次优方法分别提升了 1.41%、1.32% 与 1.04%、0.88%,为遥感影像道路提取提供了新的方法。

致谢:本文得到三峡大学先进计算中心支持。

表 3 不同模型的 FLOPs 和参数量
Tab.3 FLOPs and number of parameters for different models

模型	FLOPs/ Params/		IoU/%	
	G	M	CHN6	Massachusetts
CARNet	196.21	42.35	60.12	63.33
RCFSNet	729.23	58.23	56.37	62.98
MSMDFFNet	298.41	39.26	58.85	63.18
DBRANet	212.63	47.68	59.91	63.28
D-LinkNet	134.35	31.1	59.72	63.15
TransRoadNet	321.49	36.64	59.85	62.94
HSN-Net	313.31	59.40	60.54	64.29
SwinMSNet	375.53	55.45	61.06	64.08
Ours	228.43	35.41	62.38	65.17

参 考 文 献:

[1] HUANG B, LU Y W, YANG R P, *et al.* HSN-Net: a hybrid segmentation neural network for high-resolution road extraction [J]. *IEEE Geoscience and Remote Sensing Letters*, 2025, 22: 2502105.

[2] 王龙春,方巍,张丽娟,等. 基于 YOLOv8s 改进的自动驾驶目标检测[J]. 液晶与显示,2025,40(5):773-784.

WANG L C, FANG W, ZHANG L J, *et al.* Improved autonomous driving object detection based on YOLOv8s [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(5): 773-784. (in Chinese)

[3] 李智杰,惠爱婷,李昌华,等. 面向遥感图像道路提取的多尺度上下文感知网络[J]. 光学精密工程,2025,33(4): 610-623.

LI Z J, HUI A T, LI C H, *et al.* Multi-scale context-aware network for road extraction in remote sensing images [J]. *Optics and Precision Engineering*, 2025, 33(4): 610-623. (in Chinese)

[4] WANG Y C, TONG L, YANG J X, *et al.* Swin transformer embedding MSMDFFNet for road extraction from

- remote sensing images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2025, 22: 6006105.
- [5] 郭昊虎,高若谦,葛明锋,等. 基于特征图金字塔的冠脉造影图像血管分割方法[J]. 中国光学(中英文),2024,17(4): 971-981.
GUO H H, GAO R Q, GE M F, *et al.* Coronary artery angiography image vessel segmentation method based on feature pyramid network [J]. *Chinese Optics*, 2024, 17(4): 971-981. (in Chinese)
- [6] 戴激光,常嘉骏,李宛潼,等. 一种基于密集多尺度 Unet 的山区道路提取方法[J]. 测绘科学,2025,50(4):92-102.
DAI J G, CHANG J J, LI W T, *et al.* A mountain road extraction method based on dense multi-scale Unet [J]. *Science of Surveying and Mapping*, 2025, 50(4): 92-102. (in Chinese)
- [7] MA X Z, ZHANG X K, ZHOU D X, *et al.* StripUnet: a method for dense road extraction from remote sensing images [J]. *IEEE Transactions on Intelligent Vehicles*, 2024, 9(11): 7097-7109.
- [8] 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 C K, ZENG Q G, FANG J, *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)
- [9] LI A, LI X H, MA X S. Residual dual U-shape networks with improved skip connections for cloud detection [J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 21: 5000205.
- [10] LIU R Q, LI F T, JIANG W P, *et al.* Generating pixel enhancement for road extraction in high-resolution aerial images [J]. *IEEE Transactions on Intelligent Vehicles*, 2024, 9(10): 6313-6325.
- [11] ZHANG Z X, LIU Q J, WANG Y H. Road extraction by deep residual U-Net [J]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(5): 749-753.
- [12] QU S M, ZHOU H F, YANG X Y, *et al.* CEBRNet: connectivity enhancement and boundary refinement network for road extraction from satellite imagery [J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 21: 6010005.
- [13] ZHOU L C, ZHANG C, WU M. D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018: 192-1924.
- [14] WANG Y C, TONG L, LUO S Y, *et al.* A multiscale and multidirection feature fusion network for road detection from satellite imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5615718.
- [15] YANG R Y, ZHONG Y F, LIU Y H, *et al.* Occlusion-aware road extraction network for high-resolution remote sensing imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5619316.
- [16] WANG Y, PENG Y X, LI 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: 4412612.
- [17] LU X Y, ZHONG Y F, ZHENG Z, *et al.* GAMSNet: globally aware road detection network with multi-scale residual learning [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 175: 340-352.
- [18] GAO L P, WANG J Y, WANG Q X, *et al.* Road extraction using a dual attention dilated-LinkNet based on satellite images and floating vehicle trajectory data [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 10428-10438.
- [19] WANG Y, SEO J, JEON T. NL-LinkNet: toward lighter but more accurate road extraction with nonlocal operations [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 3000105.
- [20] DAI L, ZHANG G Y, ZHANG R T. RADANet: road augmented deformable attention network for road extraction from complex high-resolution remote-sensing images [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5602213.
- [21] WANG C W, XU R T, XU S B, *et al.* Toward accurate and efficient road extraction by leveraging the characteristics of road shapes [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 4404616.
- [22] MENG Q L, ZHOU D X, ZHANG X K, *et al.* Road extraction from remote sensing images via channel attention and multilayer axial transformer [J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 21: 5504705.
- [23] WANG Q L, WU B G, ZHU P F, *et al.* ECA-Net: efficient channel attention for deep convolutional neural networks [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11531-11539.

- [24] CHEN Y P, DAI X Y, LIU M C, *et al.* Dynamic convolution: attention over convolution kernels [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11027-11036.
- [25] QU Z L, LI M Z, CHEN Z H. CARENet: satellite imagery road extraction via context-aware and road enhancement [J]. *IEEE Geoscience and Remote Sensing Letters*, 2025, 22: 6005005.
- [26] YANG Z G, ZHOU D X, YANG Y, *et al.* Road extraction from satellite imagery by road context and full-stage feature [J]. *IEEE Geoscience and Remote Sensing Letters*, 2023, 20: 8000405.
- [27] CHEN S B, JI Y X, TANG J, *et al.* DBRANet: road extraction by dual-branch encoder and regional attention decoder [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 3002905.
- [28] YANG Z G, ZHOU D X, YANG Y, *et al.* TransRoadNet: a novel road extraction method for remote sensing images via combining high-level semantic feature and context [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 6509505.

作者简介:



翟保明,男,硕士研究生,2022年于郑州轻工业大学获得学士学位,主要从事遥感影像道路提取的研究。E-mail: 1649111726@qq.com



邵攀,男,博士,副教授,2016年于武汉大学获得博士学位,主要从事遥感影像处理、变化检测、人工智能等研究。E-mail: panshao@whu.edu.cn