

基于高分辨率遥感影像的城市水体语义分割和制图 ——以中国成都为例

王婷婷, 程 熙, 李 斌, 王滢晨, 宋泽毅, 孟丽康

(成都理工大学地球物理学院, 成都 610059)

摘要: 城市水体对城市规划和发展极为重要, 快速且准确地获取城市水体信息具有重要意义。大规模提取水体范围通常使用中低分辨率遥感影像来实现, 但水体边缘检测和小水体的提取需要使用高分辨率遥感影像。针对成都市城市水体, 笔者构建了一份高分辨率遥感影像城市水体数据集, 并采用了一种基于注意力机制的 SegFormer 模型的语义分割方法。实验结果表明, 构造的数据集和该模型结合进行城市大规模水体信息提取是可行的, 提取精度指标均达 96% 以上。该研究可应用于其他大区域的水体提取和制图, 为其他利用高分辨率遥感影像提取大规模水体信息的方法提供参考。

关键词: 城市水体; 高分辨率遥感; SegFormer; 大规模

中图分类号: P237 **文献标志码:** A **DOI:** 10.12474/wthtjs.20240314-0007

Semantic segmentation and mapping of urban water bodies based on high-resolution remote sensing images: a case study in Chengdu, China

WANG Tingting, CHENG Xi, LI Bin, WANG Yingchen, SONG Zeyi, MENG Likang

(Chengdu University of Technology College of Geophysics, Chengdu 610059, China)

Abstract: Urban water bodies are critically crucial for urban planning and development, underscoring the need for rapid and precise acquisition of urban water body information. Typically, the large-scale water body extraction is achieved using low- to medium-resolution remote sensing images. However, detecting edges and extracting smaller water bodies necessitate high-resolution remote sensing images. For this purpose, a high-resolution remote sensing image dataset has been developed for urban water bodies in Chengdu. The dataset employs a semantic segmentation approach using the SegFormer model, which integrates an attention mechanism. Experimental results validate the efficacy of combining this dataset with the model for large-scale extraction of urban water body information, achieving extraction accuracy rates exceeding 96%. This study demonstrates potential applications in extracting and mapping water bodies across other extensive regions, serving as a benchmark for different methods that utilize high-resolution remote sensing images to extract large-scale water body information.

收稿日期: 2024-03-14

基金项目: 火灾感知软件集成及预警模型功能测试项目(80303-AH20230088)

第一作者: 王婷婷(1999—), 女, 硕士, 主要从事神经网络、机器学习和遥感方面的研究, E-mail: wting9849@163.com。

通信作者: 程熙(1982—), 男, 博士, 主要研究应用人工智能从高分辨率遥感影像中自动提取城市要素, E-mail: chengxi13@cdut.edu.cn。

引用格式: 王婷婷, 程熙, 李斌, 等, 2025. 基于高分辨率遥感影像的城市水体语义分割和制图——以中国成都为例[J]. 物探化探计算技术, 47(3): 466-474.

Wang T T, Cheng X, Li B, et al., 2025. Semantic segmentation and mapping of urban water bodies based on high-resolution remote sensing images: a case study in Chengdu, China[J]. Computing Techniques for Geophysical and Geochemical Exploration, 47(3): 466-474.

Keywords: urban water body; high-resolution remote sensing; SegFormer; large-scale.

0 引言

城市水体对监测城市环境和维护城市生态稳定具有重要意义(Fletcher et al., 2013),它们的形态和空间分布在城市规划和发展中扮演着重要角色。因此,快速、准确地获取城市水体信息十分必要(Chen et al., 2020; Huang et al., 2015)。遥感影像包含丰富的地物光谱信息,其观测范围广、获取成本低,被广泛应用于城市水体测绘(Wu et al., 2019; Zhou et al., 2014)。常用基于光谱指数的方法来获取城市水体的位置、面积、形状等基本信息(Chang et al., 2022; Kaplan and Avdan, 2017)。然而,这些方法主要集中在 Landsat 和 Sentinel 系列的中低分辨率遥感影像,这限制了对城市区域内小型和狭窄水体的提取能力(张霖晨等, 2018; 蒋丹丹等, 2019)。随着遥感影像分辨率的不断提高,高分辨率遥感影像为此提供了更精确的数据支持。已有在高分辨率遥感影像上进行城市水体提取工作的结果表明,与中低分辨率遥感影像相比,更高分辨率的影像在小型和狭窄水体提取方面具有优势(Liu et al., 2022; Senaras et al., 2014)。然而,这些方法通常需要对每幅影像单独进行阈值计算及参数选择,限制了在大规模应用中的普适性和便捷性。

近年来,随着深度学习技术的快速发展,其在地物精细识别和语义分割等方面展现出了优越的性能(姜懿珈等, 2025),并在遥感地物信息提取领域得到了广泛的应用(雷雨竺等, 2025)。基于此,许多学者利用语义分割模型成功从高分辨率遥感影像中准确识别水体。Li et al.(2019)利用 FCN 模型从 GF-2 影像中提取中国北京南部地区的水体;张永洪等(2023)利用 DeepLabV3+网络探索在 GF-2 影像中城市水体的提取方法;Wang et al.(2021)利用高分辨率航空影像,引入 Areal-BiSeNet 语义分割网络,对水体和其他各类地物的语义分割效果表现出色;还有学者通过改进 U-Net 和 DeepLabV3+等模型,在高分辨率遥感影像上进行城市水体的提取工作(Li et al., 2021; Li et al., 2019)。这些方法显示了基于卷积神经网络(Convolutional Neural Networks, CNN)的语义分割模型从遥感图像中提取水体的优势。2020 年底, Vision Transformer(ViT)将在自然语言处理领域

取得巨大成功的 Transformer 架构引入计算机视觉任务领域(Chen et al., 2021),其中基于注意力机制的 Swin Transformer(Liu et al., 2021)和 SegFormer(Xie et al., 2021)模型展现了卓越的性能。随后有研究者们基于上述模型,在 Sentinel 系列影像上进行水体提取。结果表明注意力机制模型能够准确地从遥感图像中提取水体信息(Li et al., 2023; Yang et al., 2023)。然而,目前的研究聚焦在提高小区域水体信息提取的准确性,很少有研究关注使用高分辨率遥感影像来检测大规模区域的水体。因此,有必要探索高分辨率遥感影像在大规模水体检测中的应用。

笔者制作了一份高分辨率遥感影像成都市城市水体数据集,采用一种基于注意力机制的 SegFormer 模型的水体语义分割方法,通过结合构建的数据集和 SegFormer 模型,形成了一个基于注意力机制语义分割模型的水体提取框架,实现了对城市水体的高精度提取,并以成都城市区域为例,绘制详细的地表水体信息。本研究可为其他利用高分辨率遥感影像进行大规模水体信息提取的相关研究或应用提供参考。

1 研究区域和数据来源

1.1 研究区域

本文选择四川省的省会城市成都作为研究区域(图 1),全市河网密度高达 1.22 km/km^2 ,为四川首位,因此对成都市进行城市水体提取具有很高的研究价值。考虑到中国城市的行政区划面积大于其城市发展的实际区域,选择了成都市中心城市化的区域而非传统的行政区域作为研究区域,更符合对城市水体的研究。

1.2 数据来源

从谷歌地球获取到成都市 2020 年夏季的 Level 20 的高分辨率遥感影像,空间分辨率为 0.27 m ,包含红绿蓝三个波段。研究区域由 29 幅 1:25 000 的国家标准比例尺影像覆盖,如图 2 所示,单张影像尺寸约为 $93\,000 \times 62\,000$ 像素。其中图幅编号根据《国家基本比例尺地形图分幅和编号》(GB_T 13989-2012)文件的命名规则得到。

已有的城市水体研究涉及湖泊、大河流、小河流、城市边缘农田水域和池塘五类水体(Song et al., 2020;

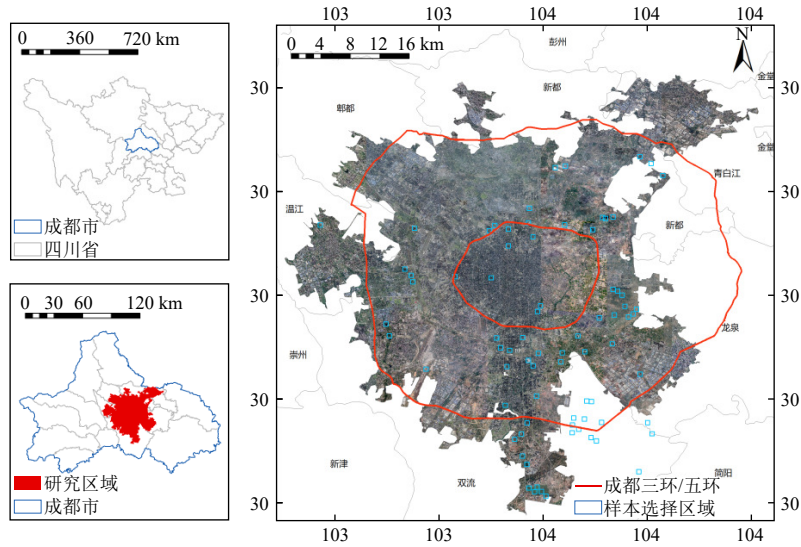


图 1 研究区域

Fig. 1 The study area

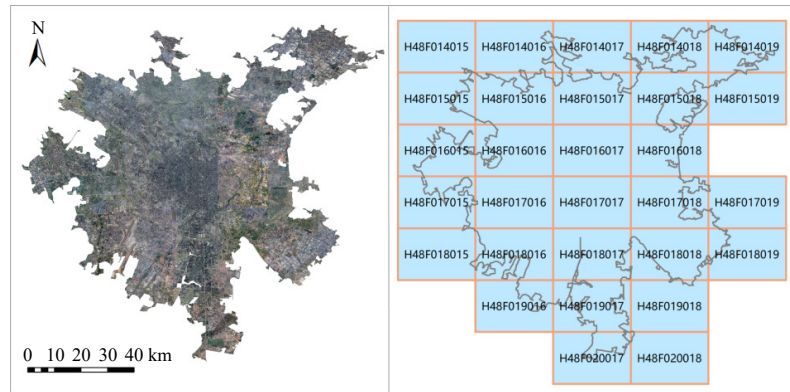


图 2 高分辨率遥感影像的覆盖范围

Fig. 2 The coverage of high resolution remote sensing images

Yang et al., 2017)。然而,为了更符合研究区域水体的特点,本研究将成都城市水体分为主流、支流、湖泊、小水体、其他水体五类,它们承担着城市的不同功能,对城市经济和自然生态的协调发展至关重要。

2 方法

本文受深度学习的启发,利用注意力机制模型 SegFormer 对成都市城市区域的高分辨率遥感影像进行了大规模的水体提取,研究工作流程如图 3 所示。

2.1 高分辨率遥感数据集制作

在研究区域内选择 77 个具有典型和代表性的城市水体样本点,以这些样本点为中心,建立了 4 000 * 4 000 像素大小的样本区域。对样本中的水体进行人工标注,考虑到硬件限制和更好的匹配深度学习模型的输入要求,采用公共区域为 32 像素的重叠裁剪

方法将影像和标注裁剪为 1024 * 1024 像素的样本小块,以提供更多的数据增强,减少裁剪导致的信息丢失和边界效应。此外还增加了 63 个靠近城市区域的以外的 1024 * 1024 像素样本,使得数据集不仅限于城市区域,增加数据集的普适性。最终制作了 1295 个 1024 * 1024 像素的样本数据集,包括 1036 个训练集、259 个验证集。数据集中的正负样本和各种水体类型比例均衡,正负样本图幅比例保持为近似 1:1。数据集制作过程如图 4 所示。

数据集的构建基于地理学的三个基本原则(相关性、异质性和相似性)和制图理论。样本涵盖了成都市区内的主流、支流、湖泊、小水体和其他水体。除水体样本外,样本数据中还有阴影、绿地、道路、农田和其他易与水体混淆的非水体样本。数据集中部分水体样本示例和具体情况如表 1 所示,其中白色为水体,黑色为非水体。

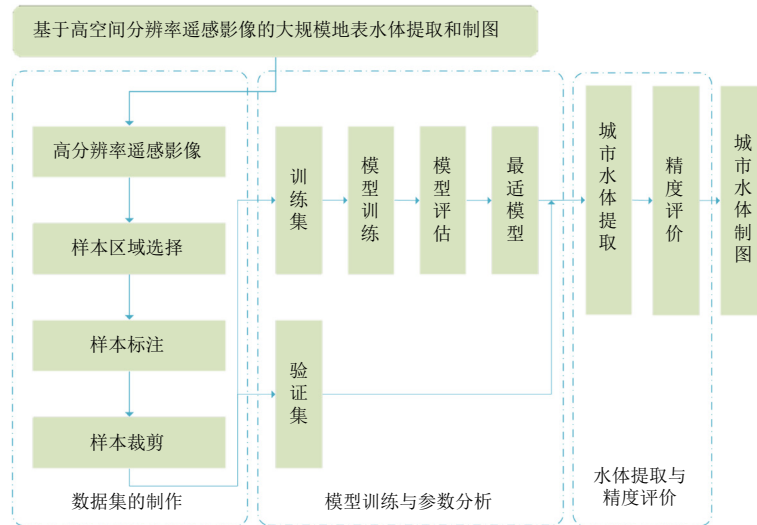


图3 研究流程图

Fig. 3 Flowchart of this study

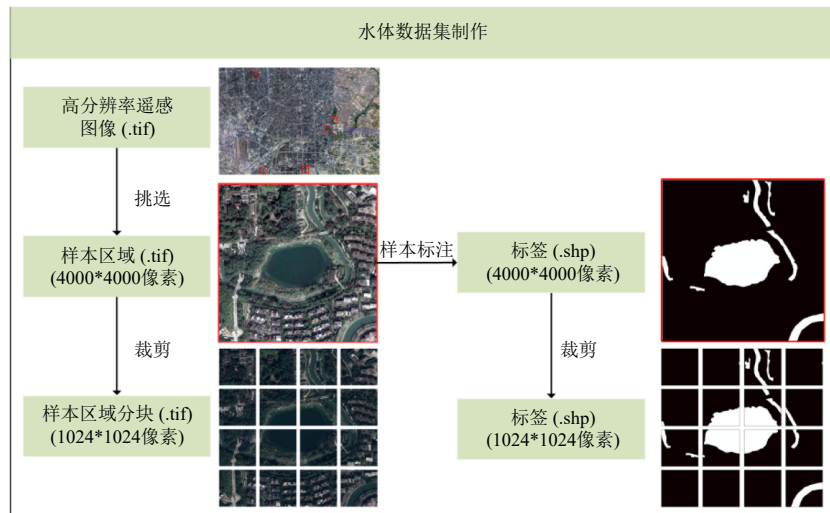


图4 水体数据集制作过程

Fig. 4 Process for the production of water body datasets

表1 数据集中水体样本示例

Tab. 1 Examples of water samples in datasets

水体类型	主流	支流	湖泊	小水体	其他水体	非水体
样本						
标签						
数量	168个	136个	225个	66个	50个	650个
占比	13.0%	10.5%	17.4%	5.1%	3.8%	50.2%

2.2 SegFormer 模型多尺度提取城市水体

高分辨率遥感影像数据由于其空间分辨率高,每个像素都包含丰富的信息,能够提供更加详细

的地物信息。注意力机制模型通过关注全局上下文信息而非仅限于局部区域,有助于获取更加详尽的信息。特别是基于注意力机制的模型 SegFormer,该

模型采用了四个不同尺度的 Transformer 编码器, 每个编码器都设计来捕获从粗糙到精细的特征, 这种层级化的特征提取策略有助于模型在保持高分辨率细节的同时, 增强对全局信息的理解。例如, 较低层次的编码器专注于捕捉大范围的水体特征, 而高层次的编码器则精细化处理水边界和小水体的细节, 从而提高模型对水体边界的检测精度。此外,

SegFormer 利用其广阔的有效感受野, 完整地捕获上下文语义信息, 其内置的注意力机制能够动态地调整模型对各个区域的关注程度, 特别是在城市环境中, 各种地物密集且相互交织的情况下。这一机制确保模型不仅能识别大片水域, 还能准确地标识出被建筑物和绿地环绕的小型水体。SegFormer 的整体架构如图 5 所示。

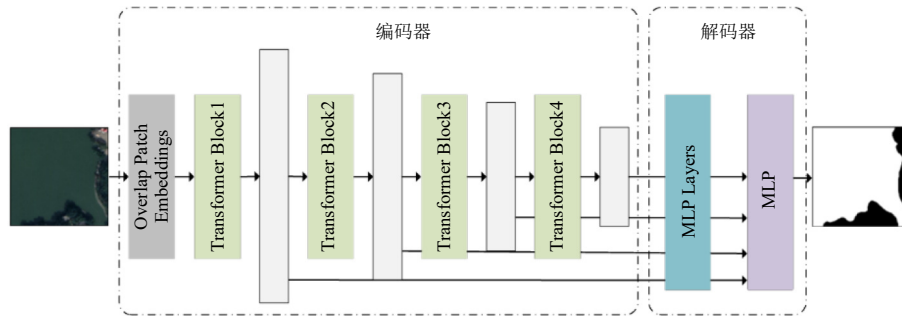


图 5 SegFormer 的整体结构

Fig. 5 The overall architecture of the SegFormer

2.3 数据增强

为确保构建的数据集能在城市水体提取任务中得到充分利用, 对数据集进行数据增强。数据增强旨在增加数据量, 提高模型的泛化能力。包括使用 Resize(调整图像和分割图的大小)、Random Crop(随机裁剪图像和分割图)、Random Flip(翻转图像和分割图)和 Photometric Distortion(依次对图像进行光度畸变)。基于前期实验, 选择 0.5 到 2.0 的范围设置数据大小调整比例, 旨在模拟图像尺度变化。图像的随机裁剪和填充是为了匹配模型输入的尺寸, 同时保持数据的完整性, 并以 0.5 的概率翻转图像, 以确保数据在水平方向的均匀性, 最终生成多样性的训练样本。

2.4 评估指数

为全面评估各模型基于超高分辨率遥感图像的城市水体提取效果, 采用精确度(P)、召回率(R)、联合交叉(IoU)、F1 分数进行评估。计算公式如下。

$$P = TP / (TP + FP) \quad (1)$$

$$R = TP / (TP + FN) \quad (2)$$

$$IoU = TP / (TP + FP + FN) \quad (3)$$

$$F1score = 2 \times (P \times R) / (P + R) \quad (4)$$

其中 TP 为正确提取的水体、TN 为正确提取的非水体、FP 为错误归类为水体的区域、FN 为错误预测为非水体的水体区域。精确度表征预测正确水体的准

确性。召回率反映模型识别水体的能力, 召回率越高, 模型识别水体的能力越强。

3 实验结果和分析

3.1 不同模型提取结果比较

在当前水体提取任务中, FCN、DeepLabV3、BiSeNet、Swin Transformer 以及 SegFormer 等语义分割模型均有较好的精度表现, 因此选取上述模型进行城市水体提取实验。为了评估不同模型在水体提取中的效果, 实验采用相同的数据集、训练参数、实验设置和预训练权重。训练数据选择 $1024 * 1024$ 像素, 实验环境由 Nvidia GeForce GTX 3060 显卡和英特尔酷睿 i7-11700F CPU 组成, 每台机器共迭代 16 万次, 优化器选择 AdamW, 损失函数为 Cross Entropy, 学习策略采用预热学习率后衰减, 初始学习率为 0.0001, Batch Size 为 2。考虑到实验计算资源和硬件条件的限制, 采用 ResNet50 作为卷积模型的骨干, 选择 tiny 版本作为 Swin Transformer 模型的骨干, 选择 mit-b3 版本作为 SegFormer 模型的骨干。具体来说, ResNet50 具有优秀的深度残差学习框架, 不仅能有效解决梯度消失问题, 同时计算复杂度低; Swin Transformer 则可以通过窗口化自注意力机制优化计算效率, 其中 tiny 版本较为轻量; SegFormer 具有高效的 Transformer 编码器和多尺度特征融合

策略,mit-b3 版本在模型大小和性能之间提供了良好的平衡,三者都十分适用于计算资源有限的环境。所有模型的预训练权重均来自 ImageNet 数据集。各模型的评估结果如表 2 所示。

上述模型的评估结果表明,在相同条件下,各模型在构建的数据集上都有不错的表现。最低的是 FCN 模型的 IoU 为 80.2%,最高的是 SegFormer 模

型的 Precision 为 98.47%。其中基于注意力机制的 SegFormer 和 Swin Transformer 模型的评价指标普遍优于基于卷积的 FCN、DeepLabV3、BiSeNet 模型,为探索城市水体信息提取的方法提供了思路。图 6 显示了上述模型在测试集上的提取结果。

结果表明,在城市水体提取任务中,基于 CNN 模型的 FCN、DeepLabV3 以及 BiSeNet 能够提取出大部分的城市水体,但是存在小型水体漏提和水体边界不够明确的情况。基于注意力机制的模型 Swin Transformer 和 SegFormer 在保有 CNN 模型提取大型水体的优势的同时,能够提取到更多的小型水体,获取到更明确的水体边缘。对主流(图 6(a)中第一行)、支流(图 6(a)中第二行、第四行)、湖泊(图 6(a)中第三行、第四行)、小水体(图 6(a)中第五行、第六行)、其他水体(图 6(a)中第七行)都有很好

表 2 各模型的评估结果

Tab. 2 Evaluation results of each model

	Backbone	Precision(%)	Recall(%)	IoU(%)	F1-Score(%)
FCN	Resnet50	87.22	90.88	80.2	89.01
DeepLabV3	Resnet50	89.72	93.96	84.83	91.79
BiSeNet	Resnet50	86.99	94.04	82.44	90.37
SwinTransformer	tiny	97.92	98.40	96.39	98.16
SegFormer	mit-b3	98.47	98.11	96.12	98.02

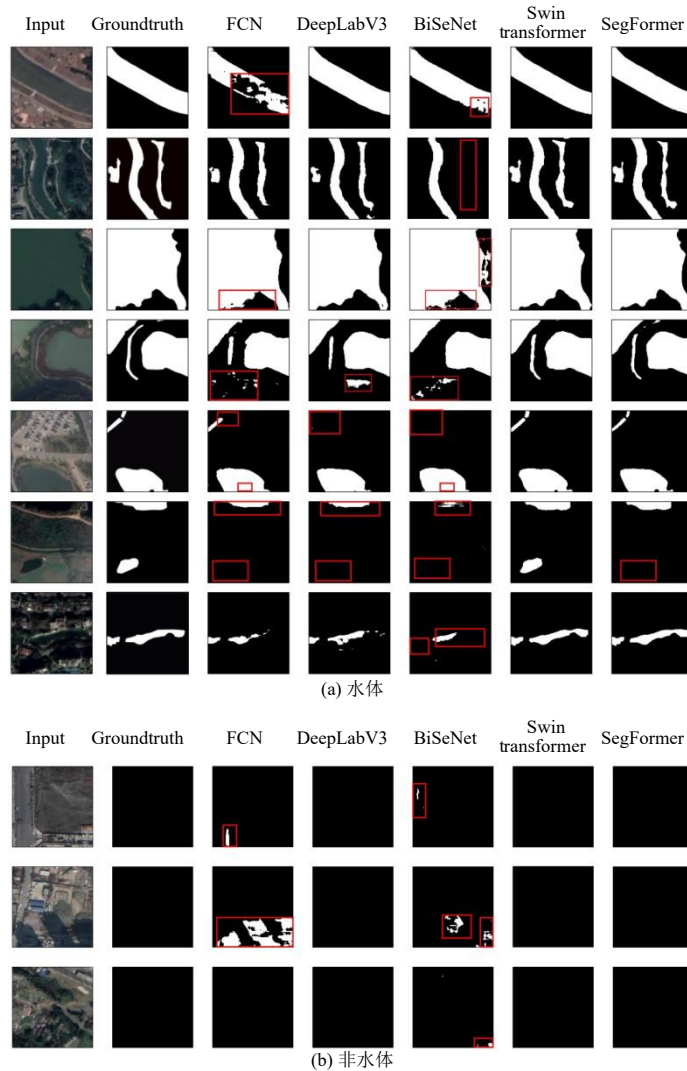


图 6 不同模型在测试集上的提取结果

Fig. 6 Different models' extraction results on the test set

的表现。这是由于 Swin Transformer 的层级设计,很好地结合了局部和全局的语义信息,能有效地处理图像的多尺度信息,从而获得更完整的水体提取结果。而 SegFormer 的多层编码器,可以生成高分辨率粗特征和低分辨率细特征,对影像中提取目标的边界有更好的语义分割效果。除此之外,注意力机制模型在带状道路、建筑物阴影、绿色植被等复杂场景中少有误提现象(图 6(b))。表明了注意力机制模型在高分辨率遥感影像上提取城市水体的优势。

除了对模型精度进行定量评估外,还使用 Flops(浮点运算)和参数来定量评估不同模型的性能。Flops 作为评估算法或模型复杂性的指标,反映计算工作量并评估模型效率。Flops 越小,说明性能越好。结果如表 3 所示。

上述模型的性能评估结果表明,在相同条件下,SegFormer 在城市水体提取任务中的 Flops 为 117.88 GFLOPs,参数为 4 460 万,均低于其他模型,在城市水体提取时具有更高的效率。

综合上述评价指标,在相同条件下,基于注意力机制模型的 Swin Transformer 和 SegFormer 在城市水体提取任务中能够获得更好的精度。其中,

表 3 各模型的 Flops 和参数

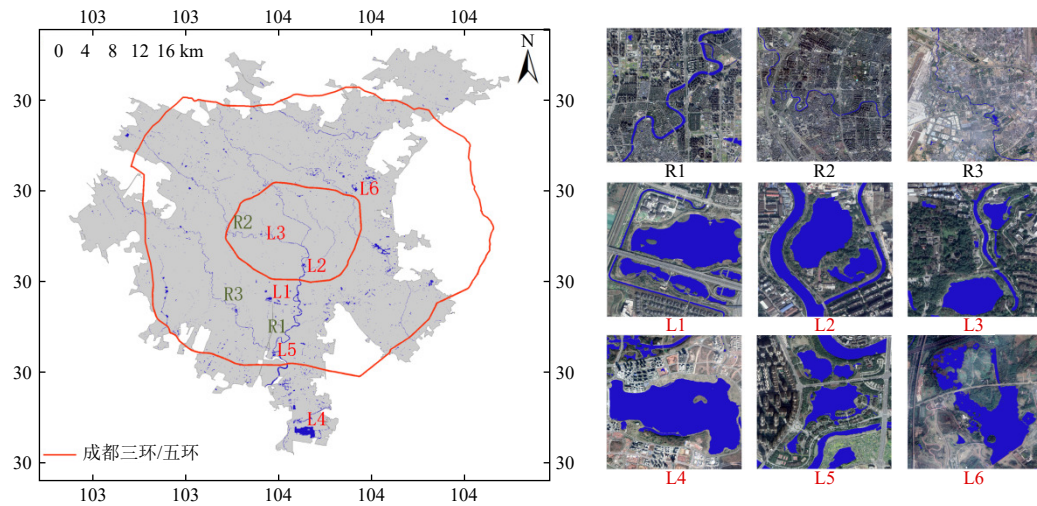
Tab. 3 Flops and Parameters of each model

	FCN	DeepLabV3	BiSeNet	SwinTransformer	SegFormer
Flops (GFLOPs)	791.9	1079.74	396.31	798.73	117.88
Parameter (M)	49.48	68.1	59.24	59.83	44.6

SegFormer 的 Flops 和参数均低于其他模型,在保证精度的同时,效率更高,所需的计算资源更少,具有精度高、效率好、成本低的优点。在大规模应用中,效率和成本是必须考虑的重要因素。因此,本研究选择 SegFormer 模型来完成成都市城区大规模的水体信息提取任务。

3.2 成都市城市水体制图

通过对上述实验结果的分析,在本文构建的数据集上,SegFormer 模型能够提取出更完整的城市水体,并且贴合水体实际边缘,在拥有高精度的同时还兼顾效率,适合应用于成都市大规模的城市水体提取和制图。将其模型训练得到最优权重文件用于水体预测,得到成都市地表水体绘制图,如图 7 所示。



R1-R3 为城市河流, L1-L6 为城市湖泊

图 7 成都市城市水体制图

Fig. 7 Mapping of urban water bodies in Chengdu City

结果表明,本文构造的数据集和 SegFormer 模型结合进行城市大规模水体信息提取是可行的,对流经成都的宽阔河流和细小的河流支流(如府河、清水河、江安河,图 7 中的 R1-R3)有着良好的表现,对成都市城市中心区域内的典型湖泊(如锦城湖、东湖、浣花溪,图 7 中的 L1-L3)和城市边缘区域的典

型湖泊(如兴隆湖、南湖、北湖,图 7 中的 L4-L6)也有着较好的提取结果。

4 讨论和结论

本文针对城市水体提取任务,成功构建并应用

了一份高分辨率遥感影像水体数据集。研究选取了多个模型进行水体提取实验,结果显示,在构建的数据集上,各模型均展现了良好的精度,特别是 SegFormer 模型。相较于其他模型, SegFormer 模型先进的多尺度特征提取及分层融合策略,大幅提高了对复杂城市场景中水体的完整捕捉能力,尤其在精确描绘水体边缘和减少邻近非水体地物的误识别方面表现突出。此外,其轻量级的多层感知机设计帮助它在拥有高精度的同时具有高效率。然而,该模型在提取极小型水体时仍面临漏提和边缘分割不清晰的挑战。未来工作将致力于以下三个方面:

1)数据集扩充:受限于人工标注的成本,本文构建的数据集规模对于大型的数据集来说还是不够的,后续将不断增加样本数量并丰富样本多样性,以提升数据集的全面性。

2)可扩展性:开发更加高效的大规模图像处理深度学习模型,将 SegFormer 模型与其他先进技术结合,包括引入更精细的数据预处理和后处理技术,以及与其他类型的遥感数据融合,进一步提高水体的提取精度和效率。

3)丰富水体提取细节:提取不同类型、不同深度等更加细节的水体,得到更加精细的水体提取结果。解决这些问题将有助于在实际应用中更好地利用高分辨率遥感影像和深度学习技术进行水体提取与管理。

通过构建的高分辨率遥感影像城市水体数据集,结合 SegFormer 模型能实现高精度和高效率的城市水体提取任务,证明本文构建的数据集和该模型在大规模应用中的适用性,最终实现了成都市大规模城市水体提取与制图,为城市水体的精细化管理和规划提供参考依据。在成都市城市区域内成功提取水体得到的经验可以扩展到其他背景复杂且尺度差异大的地物提取应用中。

参考文献:

蒋丹丹,原娟,武文娟,等,2019.基于 Sentinel-2 卫星影像的面向对象城市水体提取[J].地理空间信息,17(5):10—13+4.

Jiang D D, Yuan J, Wu W J, et al., 2019. Object-oriented Urban Water Body Extraction Based on Sentinel-2 Satellite Imagery[J]. Geospatial Information, 17(5): 10—13+4. (In Chinese)

姜懿珈,周怀来,陈姝颖,等,2025.基于 HS-DnCNN 网络的地震数据随机噪声去除[J].物探化探计算技术,47(2):

169—178.

Jiang Y J, Zhou H L, Chen S Y, et al., 2025. Removal of random noise in seismic data based on HS-DnCNN network [J]. Computing Techniques For Geophysical and Geochemical Exploration, 47(2): 169—178. (In Chinese)

雷雨竺,刘瑞,曹礼刚,等,2025.基于变化检测的北川羌族自治县震后重建监测与动态分析[J].物探化探计算技术,47(2):300—310.

Lei Y Z, Liu R, Cao L G, et al., 2025. Post-earthquake reconstruction monitoring and dynamic analysis of Beichuan Qiang Autonomous County based on change detection[J]. Computing Techniques For Geophysical and Geochemical Exploration, 47(2): 300—310. (In Chinese)

张霖晨,赵煦,钟何平,2018.利用 LandSat 8 卫星影像进行城市水体提取研究[J].计算机与数字工程,46(12):2556—2559+2564.

Zhang J C, Zhao X, Zhong H P, 2018. Research on Urban Water Bodies Extraction Using LandSat 8 Satellite Image[J]. Computer & Digital Engineering, 46(12): 2556—2559+2564. (In Chinese)

张永洪,孙路遥,席梦丹,2023. DeepLab V3+网络在遥感影像中城市水体的提取方法[J].西安文理学院学报(自然科学版),26(2):93—98+104.

Zhang Y H, Sun L Y, Xi M D, 2023. The Extraction Method of Urban Water Body in Remote Sensing Images by DeepLab V3+ Network[J]. Journal of Xi'an University (Natural Science Edition), 26(2): 93—98+104. (In Chinese)

Chang L, Cheng L, Huang C, et al., 2022. Extracting Urban Water Bodies from Landsat Imagery Based on mNDWI and HSV Transformation[J/OL]. Remote Sensing, 14(22): 5785.

Chen F, Chen X, Van De Voorde T, et al., 2020. Open water detection in urban environments using high spatial resolution remote sensing imagery[J/OL]. Remote Sensing of Environment, 242: 111706.

Chen Z S, Xie L X, Niu J W, et al., 2021. Visformer: the vision-friendly transformer[EB/OL]. 2021: arXiv: 2104.12533.

Fletcher T D, Andrieu H, Hamel P, 2013. Understanding, management and modelling of urban hydrology and its consequences for receiving waters: A state of the art[J/OL]. Advances in Water Resources, 51: 261—279.

Huang X, Xie C, Fang X, et al., 2015. Combining Pixel- and Object-Based Machine Learning for Identification of Water-Body Types From Urban High-Resolution Remote-Sensing Imagery[J/OL]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(5): 2097—2110.

Kaplan G, Avdan U, 2017. Object-based water body extraction

- model using Sentinel-2 satellite imagery[J/OL]. *European Journal of Remote Sensing*, 50(1): 137—143.
- Li J, Li G, Xie T, et al., 2023. MST-UNet: a modified Swin Transformer for water bodies' mapping using Sentinel-2 images[J/OL]. *Journal of Applied Remote Sensing*, 17(2): 026507.
- Li L, Yan Z, Shen Q, et al., 2019. Water Body Extraction from Very High Spatial Resolution Remote Sensing Data Based on Fully Convolutional Networks[J/OL]. *Remote Sensing*, 11(10): 1162.
- Li W, Li Y, Gong J, et al., 2021. Urban Water Extraction with UAV High-Resolution Remote Sensing Data Based on an Improved U-Net Model[J/OL]. *Remote Sensing*, 13(16): 3165.
- Li Z, Wang R, Zhang W, et al., 2019. Multiscale Features Supported DeepLabV3+ Optimization Scheme for Accurate Water Semantic Segmentation[J/OL]. *IEEE Access*, 7: 155787—155804.
- Liu Q, Tian Y, Zhang L, et al., 2022. Urban Surface Water Mapping from VHR Images Based on Superpixel Segmentation and Target Detection[J/OL]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15: 5339—5356.
- Liu Z, Lin Y T, Cao Y, et al., 2021. Swin transformer: hierarchical vision transformer using shifted windows[EB/OL]. 2021: arXiv: 2103.14030.
- Senaras C, Gedik E, Yardimci Y, 2014. A novel dynamic thresholding and categorizing approach to extract water objects from VHR satellite images[C]//2014 IEEE Geoscience and Remote Sensing Symposium. July 13-18, 2014, Quebec City, QC, Canada. Piscataway, NJ: IEEE: 4934—4937.
- Song S, Liu J, Liu Y, et al., 2020. Intelligent Object Recognition of Urban Water Bodies Based on Deep Learning for Multi-Source and Multi-Temporal High Spatial Resolution Remote Sensing Imagery[J/OL]. *Sensors*, 20(2): 397.
- Wang F, Luo X, Wang Q, et al., 2021. Aerial-BiSeNet: A real-time semantic segmentation network for high resolution aerial imagery[J/OL]. *Chinese Journal of Aeronautics*, 34(9): 47—59.
- Wu Y, Li M, Guo L, et al., 2019. Investigating Water Variation of Lakes in Tibetan Plateau Using Remote Sensed Data Over the Past 20 Years[J/OL]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7): 2557—2564.
- Xie E, Wang W, Yu Z, et al., 2021. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers[M/OL]. arXiv: 2105.15203.
- Yang F, Guo J, Tan H, et al., 2017. Automated Extraction of Urban Water Bodies from ZY - 3 Multi - Spectral Imagery[J/OL]. *Water*, 9(2): 144.
- Yang X, Chen M, Yu C, et al., 2023. WaterSegformer: A lightweight model for water body information extraction from remote sensing images[J/OL]. *IET Image Processing*, 17(3): 862—871.
- Zhou Y, Luo J, Shen Z, et al., 2014. Multiscale Water Body Extraction in Urban Environments From Satellite Images[J/OL]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(10): 4301—4312.