引文格式：陈凯,肖雪,李冠达,等. 西南多源遥感影像水资源监测方法研究[J]. 地理空间信息, 2024, 22(10): 14-18. doi:10. 3969/j.issn. 1672-4623.2024.10.004

西南多源遥感影像水资源监测方法研究

陈　凯[1,2]，肖　雪[1,2*]，李冠达[1,2]，付昕乐[1,2]，饶杨莉[1,2]
（1. 自然资源部第三航测遥感院，四川 成都 61010；2. 自然资源部西南山地自然资源遥感监测工程技术创新中心，四川 成都 610101）

摘　要：为解决西南地区水资源监测的问题，结合遥感影像的近红外波段，对多源遥感影像进行全局线性拉伸，全局自适应gamma变换等预处理。提出了一种基于SegFormer水体提取方法，引入了注意力机制，提取不同分辨率下的水体语义信息，实现水体的高精度像素级分割，从而快速自动的监测西南地区水资源分布情况。以四川绵阳作为实验区，利用国家基础地理信息数据构建多源遥感影像样本数据集进行模型训练，对实验区的国产遥感卫星影像进行水体提取，实验结果显示：最高正确率为95.6%，召回率为97%，IOU为0.929。结果证明该方法的水体提取相较于其他方法具有更好地精度和鲁棒性。

关键词：水体提取；gamma变换；SegFormer；多源光学遥感影像
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2024）10-0014-05

Research on Water Resource Monitoring Method Using Multi-source Remote Sensing Images in Southwest China

CHEN Kai[1,2], XIAO Xue[1,2], LI Guanda[1,2], FU Xinle[1,2], RAO Yangli[1,2]
(1. The Third Institute of Photogrammetry and Remote Sensing, Ministry of Natural Resources, Chengdu 610101, China;
2. Technology Innovation Center for Remote Sensing Monitoring of Natural Resources in Southwest China Mountains, Chengdu 610101, China)

Abstract: In order to solve the problem of water resource monitoring in southwest areas, we proposed a water body extraction method based on SegeFormer, combining the near infrared band of remote sensing image with the pre-processing of linear stretching and gamma transformation of multi-source remote sensing images. We introduced the attention mechanism to extract the semantic information of water at different resolutions, which could achieve high-precision pixel-level segmentation of water, thus quickly and automatically monitoring the water resource distribution in the southwestern areas. Taking Mianyang, Sichuan as the experimental area, we used national basic geographical information data to build multi-source remote sensing image sample dataset for training, and carried out water extraction of the domestic remote sensing satellite images in the experimental area. The results show that the highest accuracy rate is 95.6%, the recall rate is 97%, and the IoU is 0.929, which illustrating this method has better accuracy and robustness than other methods in water extraction.

Key words: water extraction, Gamma transform, SegFormer, multi-source optical remote sensing image

2021年自然资源部发布了《自然资源调查监测体系构建总体方案》，围绕土地、矿产、森林、草原、水、湿地、海域海岛等7类自然资源的调查监测进行规划和设计，其中水资源监测是非常重要的一部分。迄今为止，国内外研究人员深入研究了遥感影像的水体提取，目前自动提取算法主要分为光谱指数法、传统的机器学习和深度神经网络。光谱指数法一般是利用遥感影像地物光谱差异，设定适当的阈值提取水体，比如水体指数NDWI[1]。传统的机器学习模拟人脑认知，利用有限的先验知识分离水体和其他地物，比如K-means分类[2]、决策树[3]、随机森林[4]等。上述方法虽然可以进行水体提取，但西南山地地区云雾多，地物复杂，土地利用景观格局极为破碎，另外，光学卫星影像获取困难，且西南地区高差变化大导致影像几何畸变大，因此，需要多源多时态的卫星影像来开展工作。

由于传感器参数存在差异，不同来源的卫星影像光谱差异大，同谱异物和同物异谱现象严重，因此使用传统的机器学习算法需要根据影像设定分类规则，导致水体自动提取难度高，泛化能力差，不适合应用[5-11]。依靠人工提取水体虽然精度高，但工作量和时间成本太高，无法进行大面积的水资源监测。深度卷

收稿日期：2023-05-22。
项目来源：兵团重点研发项目（2024AB064）。
第一作者简介：陈凯（1994—），硕士，工程师，研究方向为遥感影像AI智能解译，E-mail：271083072@qq.com。
通信作者：肖雪（1992—），硕士，工程师，研究方向为新型基础测绘，E-mail：845399731@qq.com。

积神经网络[12-14]可以对多维度信息进行更有效的特征提取、融合，提取具有高度抽象、概括性的语义信息，并且可以有效地消除光谱差异的影响，能更好地提取水体特征，区分水体与其他地物，但卷积神经网络的有效感受野较小且局部，很难提取到全局信息，所以一般通过设计解码器来增大有效感受野，这也就导致解码器越来越复杂，网络参数越来越大，增加了时间成本。2017年之后，研究者们通过将图片转换为单词的形式将Transformer从自然语言处理领域引入到计算机视觉[15-17]，发现比CNN有更好的特征提取能力。因为Transformer具备自注意力机制，比卷积神经网络的有效感受野更大，不需要复杂的操作提高感受野，因此其参数量更少，同时具有更高的并行化能力，训练和推理时间变得更短。SegFormer为其中的代表作[18]。

深度神经网络自动的从遥感影像样本数据集中学习特征，在进行水体提取过程中无需手动设置分类规则，全自动化处理，因此为了更好地服务于水资源管理、保护和修复，开展基于深度神经网络的高精度高泛化能力的多源多时态遥感影像自动水体提取研究具有非常重要的意义。

1　研究方法

多源多时态卫星影像光谱差异大，同谱异物和同物异谱现象严重，通过线性拉伸、Gamma变换等处理可降低不同光学遥感影像之间的差异，有效地提高提取精度。因此本文通过分析西南地区特有的景观特征和不同卫星影像之间的差异，设计了如图1所示的水体提取流程。

1.1　全局参数的线性拉伸

$$\mathrm{I\,dn}\left(\mathrm{where}\left(\mathrm{I_{dn}}>\mathrm{S_{max}}\right)\right)=\mathrm{S_{max}} \tag{1}$$

$$\mathrm{I\,dn}\left(\mathrm{where}\left(\mathrm{I_{dn}}<\mathrm{S_{min}}\right)\right)=\mathrm{S_{min}} \tag{2}$$

$$\mathrm{S_{dn}}=\frac{\mathrm{S_{max}}-\mathrm{S_{min}}}{\mathrm{I_{max}}-\mathrm{I_{min}}}\times\left(\mathrm{I_{dn}}-\mathrm{I_{min}}\right)+\mathrm{S_{min}} \tag{3}$$

式中，$\mathrm{S_{dn}}$为拉伸后的值；$\mathrm{S_{max}}$和$\mathrm{S_{min}}$为拉伸后的最大值和最小值；$\mathrm{I_{dn}}$为影像的像素值；where为取对应条件的位置；$\mathrm{I_{max}}$和$\mathrm{I_{min}}$为自定义的极值，一般取影像的最大最小值或者百分比范围内的极值。

当影像尺寸过小时，$\mathrm{I_{max}}$和$\mathrm{I_{min}}$将导致影像严重失真，因此本文计算整景影像的最大最小值作为参数。

1.2　全局自适应Gamma变换

$$\mathrm{Mean}=\frac{\sum S_{\mathrm{dn}}}{\mathrm{ImageCount}} \tag{4}$$

$$\mathrm{Gamma}=\frac{\log 0.5}{\log(\mathrm{Mean})}$$

$$\mathrm{G\,dn}=\left(\frac{\mathrm{S\,dn}}{255.0}\right)^{\mathrm{Gamma}}\times 255.0$$

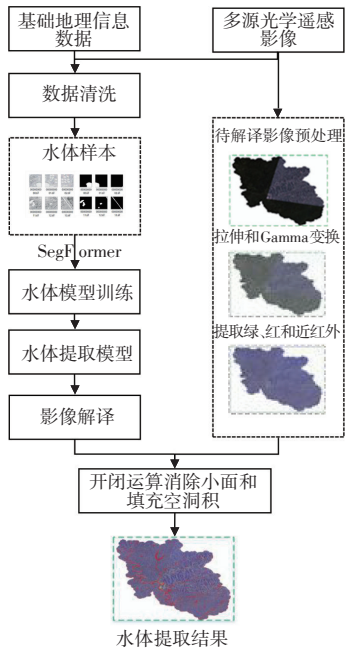


图1　水体提取流程图

其中ImageCount为影像有效值的像素数量，G_{dn}为gamma变换后的像素值。

当影像尺寸过小时，Gamma参数将失效，导致变换后影像失真，因此本文计算整景影像的Gamma作为最终的参数。

1.3　SegFormer水体提取网络

SegFormer由编码器和解码器组成，编码器通过Transform来提取特征，有4层，每一层由Patch Embeding、Transform Block和Normalize组成，解码器对不同层的特征分别进行linear保证特征维度一致，上采样到1/4分辨率并进行连接，最后用linear融合。解码器因为Transform的感受野足够大，相比于deeplab等全卷积神经网络，结构简洁。

1.4　近红外波段

水体指数证明了遥感影像近红外波段能更好地区分水体与其他地物，因此，本文选取了水体与其他地物区分度相对较高的近红外波段，结合绿、红波段来进行水体模型的训练和提取。

1.5　形态学滤波

腐蚀和膨胀是形态学中最基本的2种形态学操作，公式如下：

$$d(x,y)=\max_{(x',y'):\mathrm{element}(x',y')\neq 0}\mathrm{dn}(x+x',y+y')$$

$$e(x,y)=\min_{(x',y'):\text{element}(x',y')\neq 0}\text{dn}(x+x',y+y') \quad (6)$$

式中，element 为自定义核，可自定义尺寸，dn 为二值影像。

开运算一般用于消除小面积的目标，在纤细处分离物体，平滑大目标的边界。

$$\text{open}(x,y)=d(e(\text{dn},\text{element})) \quad (7)$$

闭运算一般用于填充小面积的空洞。

$$\text{close}(x,y)=e(d(\text{dn},\text{element})) \quad (8)$$

2 研究区与数据

2.1 研究区域概况

研究区域为绵阳市，位于四川盆地西北部，地貌复杂，西北部地势高，海拔约1 000~5 500 m，高差大，主要为山地地貌；东南部地势较低，主要为丘陵地貌，海拔约400~600 m，高差低。绵阳水系发达，包含涪江水系，白龙江水系等，水资源丰富。绵阳常年被云覆盖，可利用光学遥感影像较少，数据获取困难，水资源管理、保护和修复等应用均需要多源光学遥感影像进行分析。

2.2 数据来源

遥感影像数据源来自于2021年的国产遥感卫星，主要包括ZY1、ZY3、GF1、GF2、GF6等，空间分辨率0.8~2.5 m，时相包括春季、夏季、秋季，详细信息如表1所示。

表1 研究区域影像详细信息

传感器	时相	区域	分辨率/m
ZY1	20210329	30.96~31.50 N 105.19~105.74 E	2.5
ZY3	20210921	30.96~31.34 N 105.30~105.74 E	2.1
GF1	20210328	31.10~31.49 N 105.39~105.74 E	2
GF2	20210430	31.06~31.31 N 105.19~105.40 E	0.8
GF6	20210731	30.96~31.49 N 105.26~105.73 E	2

水体样本主要基于基础地理信息数据和国产光学遥感卫星影像生成，如表2所示。

3 实验结果与分析

3.1 评估标准

本文精度评定采用正确率，召回率和IOU，使用RTX2080 Ti显卡极限提取10 000×10 000大小的影像的水体作为时间效率评价指标，如表3，式（9）~（11）所示。

表2 水体样本统计信息

	样本量（512×512）	水体数量（像素数）
GF1	508	8 506 180
GF2	3 200	56 694 577
BJ2	1 393	84 023 461
GJ1	5 343	83 710 523
ZY3	974	23 484 659
总计	11 418	256 419 400

表3 混淆矩阵

混淆矩阵		真实值	
		水体	背景
预测值	水体	*TP*	*FN*
	背景	*FP*	*TN*

$$\text{precision}=\frac{TP}{TP+FP} \quad (9)$$

$$\text{recall}=\frac{TP}{TP+FN} \quad (10)$$

$$\text{IoU}=\frac{TP}{TP+FN+FP} \quad (11)$$

3.2 网络训练细节

实验硬件如表4所示，在训练过程中对样本进行随机数据增强，包括随机裁剪、随机缩放、随机翻转、随机校正（对比度，亮度，饱和度）、归一化，训练集和验证集比例为8∶2，初始学习率为0.01，Batch-Size设置为4，训练模型验证集精度IoU最高的为最终模型，为验证SegFormer的优越性，以同样的方式训练了DeepLabV3，HardNet，Unet等多个模型，训练正确率变化情况如图2所示，IoU变化情况如图3所示。为验证近红外波段具有更好的水体提取效果，使用SegFormer以同样方式训练了蓝、绿、红三波段样本。

表4 计算机环境

硬件	显卡	RTX2080 Ti，显存24 G
	内存	64 G
框架	Paddle Paddle	2.3

3.3 对比与分析

首先对多种国产卫星遥感影像的测试结果（表6）进行整体分析：相较于其他方法的泛化能力，本文的水体提取方法最好，整体的IOU最高，一景10 00×10 000尺寸影像使用RTX2080 Ti所需时间如表6所示。

如图7所示，针对大面积水域如湖泊，坑塘、比较宽的河流，几种深度学习提取方法效果均较好，但SegFormer在细长河流的提取效果相比其他提升很大，说明SegFormer的细节处理能力更强。

为了验证本文方法的泛化能力，在不同的遥感卫星影像上进行了测试，从表5可以看出，SegFormer的

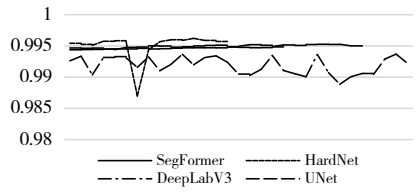


图2　各个神经网络训练验证集正确率变化

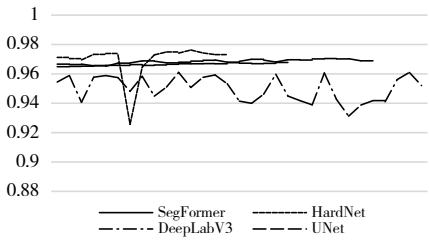


图3　各个神经网络训练验证集IOU变化

表5　实验结果精度

		SegFormer		Hard Net	Deep LabV3	UNet
		Gamma变换	无			
GF1	正确率	0.930	0.888	0.885	0.853	0.972
	召回率	0.842	0.869	0.832	0.800	0.728
	IOU	0.792	0.783	0.751	0.703	0.713
GF2	正确率	0.956	0.899	0.958	0.853	0.970
	召回率	0.970	0.990	0.964	0.800	0.959
	IOU	0.929	0.891	0.925	0.703	0.931
GF6	正确率	0.928	0.844	0.948	0.967	0.972
	召回率	0.843	0.955	0.814	0.726	0.728
	IOU	0.791	0.812	0.779	0.709	0.713
ZY1	正确率	0.944	0.895	0.969	0.947	0.979
	召回率	0.923	0.943	0.840	0.820	0.747
	IOU	0.875	0.849	0.818	0.784	0.735
ZY3	正确率	0.935	0.848	0.973	0.925	0.965
	召回率	0.909	0.944	0.827	0.845	0.857
	IOU	0.854	0.808	0.809	0.791	0.831

表6　时间效率

算法	显存/G	时间/s
SegFormer	9.6	72.86
HardNet	9.7	49.43
DeepLabV3	9.5	110.36
UNet	9.4	69.43

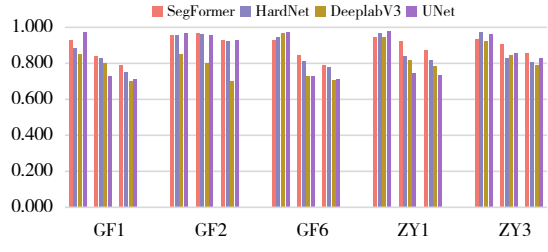


图4　水体提取精度

正确率均保持在92%以上，召回率和IOU波动较大，但召回率也保持在84%以上，IOU保持在79%以上，强力的说明了本文方法的泛化能力强，适用于不同的卫星影像，其中GF2的效果最佳，正确率95.6%，召回率97%，IOU为0.929，这是因为GF2分辨率高，影像质量较好。虽然样本未包括GF6和ZY1，但GF6、ZY1的影像水体提取均取得较好的结果，其中ZY1正确率高达94.4%，召回率为92.3%，IoU为0.875，说明本文方法仍适用于其他分辨率差异不大的卫星影像。除此之外，样本来源于四川、云南等多个省市，并未包含测试区绵阳，说明本文方法并不局限于样本所在区域，其他绝大部分省市的水体提取同样适用。

其次，实验结果表明Gamma变换在解译过程中水体提取具有更好的效果，图5为提取结果的对比图，可以看出相比于不进行Gamma变换的影像解译正确率、IOU均更高，但召回率有所降低，正确率提高了约4.2%到8.7%，IOU提高了约0.9%到4.6%。具体精度数值参考表5。原因在于不同的卫星影像因为传感器的差异，同一地区存在一定的光谱差异，经过线性拉伸降低了它们的光谱差异，提高了泛化能力，但线性拉伸同时也降低了水体与低光谱值区域的差异，导致小部分阴影、耕地、道路被误提为水，例如图6，ZY3误提了一大段的道路，ZY1误提了建筑物阴影，GF2误提了部分耕地。使用自动Gamma变换后可拉大水体与低光谱值区域的差异，因此避免了上述的问题。

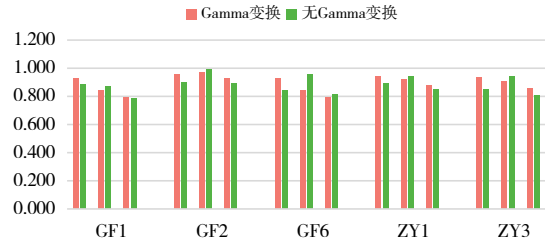
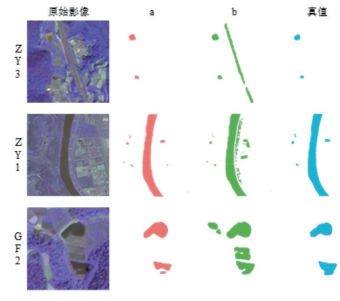


图5　SegFormer模型的水体提取精度



a Gamma变换结果　b 无Gamma变换结果

图6　SegFormer模型的水体提取

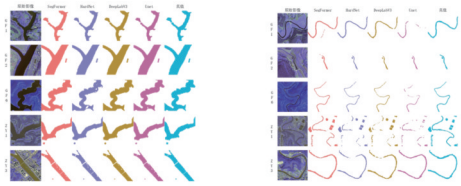


图7 不同卫星影像各方法提取结果对比

4 结 语

长期以来，遥感影像水体提取一直是自然资源监测中一个非常基础且重要的工作，本文提出了一个结合自动Gamma、SegFormer和形态学处理的多源光学遥感影像的水体提取方法，该方法综合了近红外波段在水体提取的优势，能够自动快速的提取0.8~2.5 m分辨率影像的水，最高正确率可达95.6%，召回率为97%，IoU为0.929，为水资源分析、湿地保护等提供有力的技术支撑。本文方法既可有效避免人工方法高耗时、低效率的缺点，也可避免传统方法中（如NDWI等方法）需要人工设置阈值，先验知识要求高等不足。线性拉伸、Gamma变换等影像预处理保证了本文方法泛化能力强，可适应不同地区、不同卫星影像的水体提取（图7），全程全自动化处理，无需人工干预，极大程度上提高了工作效率。由于深度学习是基于数据驱动的，而本文的样本仅有一万多个，缺少山地地区的样本，经过测试发现山地、高山地等地区召回率和绵阳地区相差不大，但正确率和IoU低。因此还需要进一步提高山地地区的水体提取精度，比如增加样本，改进影像处理模式，改善网络结构等。

参考文献

[1] Mcfeeters S K. The Use of Normalized Difference Water Index (NDWI) in the Delineation of Open Feature[J]. International Journal of Remote Sensing,1996,17(7):1 425-1 432

[2] Dhanachandra N ,Manglem K ,Chanu Y J . Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm[J]. Procedia Computer Science, 2015,54:764-771

[3] 张剑清,段艳. 一种结合分水岭与决策树C5.0的极化SAR分类方法[J]. 武汉大学学报(信息科学版),2014,39(8):891-896

[4] Chica O,M,Chica R. Machine learning Predictive Models for Mineral Prospectivity: An Evaluation of Neural Networks, Random Forest, Regression Trees and Support Vector Machines[J]. Ore Geology Reviews: Journal for Comprehensive Studies of Ore Genesis and Ore Exploration, 2015,71:804-818

[5] 杜金康,胡永胜,傅学智,等. 基于SPOT图像的Wa水体提取与分类研究spot卫星影像的水体提取方法及分类研究[J].遥感学报,2001,5(3):214-219

[6] 于欢,张树清,李晓峰,等. 基于TM影像的典型内陆淡水湿地水体提取研究[J]. 遥感技术与应用,2008(3):310-315

[7] 毕海芸,王思远,曾江源,等. 基于TM影像的几种常用水体提取方法的比较和分析[J]. 遥感信息,2012,27(5):77-82

[8] 陈文倩,丁建丽,李艳华,等. 基于国产GF-1遥感影像的水体提取方法[J]. 资源科学,2015,37(6):1 166-1 172

[9] 吉红霞,范兴旺,吴桂平,等. 离散型湖泊水体提取方法精度对比分析[J]. 湖泊科学,2015,27(2):327-334

[10] 李景刚,黄诗峰,李纪人. ENVISAT卫星先进合成孔径雷达数据水体提取研究:改进的最大类间方差阈值法[J]. 自然灾害学报,2010,19(3):139-145

[11] 刘炜,王聪华,赵尔平,等. 基于面向对象分类的细小河流水体提取方法研究[J]. 农业机械学报,2014,45(7):237-244

[12] Krizhevsky A, Sutskever I, Hinton G E. Imagenet Classification with Deep Convolutional Neural Networks[J]. Communications of the ACM,2017,60(6): 84-90

[13] Simonyan K, Zisserman A. VeryDeep Convolutional Networks for Large-scale Image Recognition[J]. arXiv preprint arXiv:1409. 1556,2014

[14] Long J, Shelhamer E ,Darrell T . Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651

[15] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale[J]. arXiv Preprint arXiv:2010. 11929,2020

[16] Liu Z, Lin Y, Cao Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision,2021

[17] Graham B, El-Nouby A , Touvron H, et al. LeViT: A Vision Transformer in ConvNet's Clothing for Faster Inference[J]. 2021.DOI:10.48550/arXiv.2104.01136

[18] Huang C H, Wu H Y, Lin Y L. Hardnet-mseg: A Simple Encoder-decoder Polyp Segmentation Neural Network That Achieves Over 0.9 Mean Dice and 86 Fps[J]. arXiv Preprint arXiv:2101,07172,2021