引文格式:虞欣,王鑫,郑肇葆,等.融合尺度注意力机制的高分辨率遥感影像水体提取网络[J].武汉大学学报(信息科学版),2025,50(7):1328-1338.DOI:10.13203/j.whugis20230079

Citation: YU Xin, WANG Xin, ZHENG Zhaobao, et al. Water Extraction of High Spatial Resolution Remote Sensing Image Fusing Scale Attention Mechanism[J]. Geomatics and Information Science of Wuhan University, 2025, 50(7): 1328-1338. DOI: 10.13203/j.whugis20230079

融合尺度注意力机制的高分辨率遥感影像水体提取网络

虞 欣[1] 王 鑫[1] 郑肇葆[2] 孟令奎[2] 李林宜[2] 尹 欣[3] 肖 洲[3]

1 北京石油化工学院,北京,102617

2 武汉大学遥感信息工程学院,湖北 武汉,430079

3 北京市水利规划设计研究院,北京,100048

摘 要:近年来,水资源短缺等一系列问题日益严重,需要对水资源的数量和变化进行及时准确的监测。从遥感影像上提取的水体信息不仅可以为水体生态分析和后续的相关风险预警提供数据支撑,还可以用于缓解日益严峻的生态环境问题,并有助于所研究流域的环境演变监测、水资源保护和开发以及生态资源的可持续发展。随着机器学习和人工智能的发展,基于机器学习的水体提取方法应运而生,特别是深度学习方法具有强大的特征表征能力,它的发展给计算机视觉和影像识别等领域带来了巨大的冲击。全卷积神经网络(fully convolutional networks, FCN)作为第一个端到端的像素级预测的全卷积网络,拉开了像素级别的语义分割序幕,其他语义分割模型都是在其基础之上发展起来的。然而,在利用FCN及其变种方法进行水体提取时仍然普遍存在水陆边界不清晰等问题。针对上述问题,提出了融合尺度注意力机制的高分辨率遥感影像水体提取加权FCN(squeeze-and-excitation weighted FCN,SE-wFCN)。该方法将注意力机制应用于不同层次或尺度的语义信息来更好地融合不同层次或尺度的特征,尽可能地减少水体边界像素的误分和漏分;考虑到FCN模型的复杂性和过拟合等问题,在FCN中引入辅助分类器,进一步降低深度学习模型在学习训练阶段过拟合发生的可能性;顾及水陆边界的像素位置信息对水体提取的重要性,在损失计算中,对水体边界像素与水体内部像素的损失赋予不同的权重,引导尺度注意力机制和FCN模型更好地关注并学习水陆边界的特征信息,进而可以进一步地提高水陆边界的提取精度。以高分辨率高分一号卫星遥感影像为研究对象,将该网络与目前主流的语义分割网络进行了对比实验,结果表明,SE-wFCN比当前主流的方法(如DeepLabV3、U-Net、U2Net、MECNet、MSResNet和SADA-Net等)在水陆边界上取得了更加满意的水体提取结果。

关键词:高分辨率遥感影像;水体提取;注意力机制;深度学习

中图分类号:P237 **文献标识码**:A **收稿日期**:2023-05-25

DOI:10.13203/j.whugis20230079 **文章编号**:1671-8860(2025)07-1328-11

Water Extraction of High Spatial Resolution Remote Sensing Image Fusing Scale Attention Mechanism

YU Xin[1] *WANG Xin*[1] *ZHENG Zhaobao*[2] *MENG Lingkui*[2] *LI Linyi*[2]
YIN Xin[3] *XIAO Zhou*[3]

1 Beijing Institute of Petrochemical Technology, Beijing 102617, China

2 School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

3 Beijing Institute of Water Resources Planning, Design and Research, Beijing 100048, China

Abstract: **Objectives**: Recently, it is gradually becoming important to timely monitor the accurate distribution and change of water resource on the earth due to the shortage of water resource globally. Water extraction from remotely sensed image not only is able to provide data support for water ecosystem analysis and

基金项目:国家重点研发计划(2021YFB3900603);北京市科技新星计划。

第一作者:虞欣,博士,教授,研究方向为影像解译、人工智能和贝叶斯统计等。china_yuxin@163.com

通信作者:李林宜,博士,副教授。lilinyi@whu.edu.cn

risk prediction, but also can alleviate the currently serious ecosystem environment and contribute to sustainable development of water resource and environment protection. Because of its powerful feature representation, deep learning has an important influence on computer vision and image recognition domains. Fully convolutional networks (FCN) is the first end-to-end fully convolutional network for image pixel-wise prediction and starts the image semantic segmentation. Hereafter, other variants of image semantic segmentation based on FCN are also put up. However, the border region between water and land is still misclassified during water extracting based on FCN or its variants. Therefore, we put up water extraction of high spatial resolution remotely sensed image fusing scale attention mechanism based on squeeze-and-excitation weighted FCN (SE-wFCN). **Methods:** We utilize GF-1 remotely sensed image to do some experiments by comparing with FCN and its variants. The proposed method has the following contribution of three aspects: (1) In order to take advantage of attention mechanism, this method applies it to features of different scales and then fuses those features more reasonably and more scientifically to reduce the possibility of misclassification of water pixels as far as possible. (2) Due to the model complexity of FCN and its possibility of overfitting, we add one auxiliary classifier module to FCN model as one branch during the training phase, which is able to help to reduce the possibility of overfitting. (3) Considering that the border pixel's position between water and land is very important to water extraction and the position information is actually ignored when computing the loss during training phase, we pay more attention to the border pixel by putting more weight in the loss function. **Results:** Experimental analysis shows that the proposed method is feasible and effective. **Conclusions:** It can achieve the more satisfying of water extraction than these current popular deep learning models (i.e. DeeplLabV3, U-Net, U2Net, MECNet, MSResNet and SADA-Net), especially in the border between water and land.

Key words: high spatial resolution remotely sensed image; water extraction; attention mechanism; deep learning

水资源作为地球上最重要的资源之一,是人类及地球上动植物赖以生存和发展的决定性物质基础[1-2]。近些年,水资源短缺等一系列问题日益严重,需要对水资源的数量和变化进行及时准确的监测。从遥感影像上进行水体提取便于及时准确地掌握现有的水资源概况,有助于对水资源更加科学、合理地规划和治理,提高水资源的利用率,对人类及社会活动具有重要影响[3-5]。遥感影像上的水体通常是指地表水,而地表水变化对生态、社会和经济的影响是当前研究的热点。从遥感影像上提取的水体信息,不仅可以为水体生态分析和后续的相关风险预警提供数据支撑,而且可以用于缓解日益严峻的生态环境问题,并有助于所研究流域的环境演变监测、水资源保护和开发以及生态资源的可持续发展。因此,水体信息对于军事侦察、环境保护、土地规划、地理测绘等领域都至关重要[6-9]。

2000年前后,由于受到卫星种类和数据获取的限制,通常是利用陆地(Landsat)卫星影像进行地表水的变化分析、提取识别和监测等研究。此外,由于Landsat卫星的空间分辨率并不高,其研究对象主要是针对大区域的河流和湖泊等。而且,考虑到当时计算机的硬件水平,其研究方法主要是针对单一像元的光谱信息提出的单波段阈值法,而且其阈值是依据某一具体波段的电磁波谱的反映特性由人工经验设定。随后,利用水体与背景地物波谱曲线变化趋势的差异,构建多个波段之间的简单逻辑关系式(即加减运算),提出了多波段谱间关系法[10]。接着,受到归一化差异植被指数的启发,依据水体在绿波段反射率较高而在近红外波段反射率较低的特性,利用Landsat卫星影像中的近红外波段和绿波段,提出了归一化差分水体指数(normalized difference water index, NDWI)[11],从此开启了水体指数法的序幕,包括改进的归一化差异水体指数[12]、修正的归一化差异水体指数[13]、WI20006[14]和WI2015[15]、新型水体指数[16]、综合权重水体指数法[17]以及自动水体提取指数[18]等。

随着卫星遥感技术和计算机技术的发展,水体指数法逐渐成为遥感影像上水体信息的一种辅助手段。许多研究者先用指数法来增强水体的光谱特征,再利用数字影像处理方法对水体进行分割提取,进一步提高水体提取的精度。由于Landsat遥感影像存在着大量的混合像元,其分割

提取水体的边界区域出现较为明显的误分现象。所以,一方面,考虑到水体光谱特征对水体提取任务的贡献是有限的,综合多种特征(如空间信息、纹理信息等)的方法可以在一定程度上弥补其缺陷,实现更加准确的水体信息提取。另一方面,已有学者开始尝试利用更高空间分辨率的卫星遥感影像(如哨兵2号[2]、SPOT[19]、高分一号[20-21]等)对水体提取算法进行研究。然而,不同卫星在相同波段上的中心波长并不完全一致,这就导致了已有的水体指数方法在用于不同的数据源时出现了阈值跳动的情况。传统水体识别方法大多数都是基于阈值,而且阈值的确定很大程度上依赖于人工经验,在大规模应用和识别精度方面也有待商榷。因此,不同的卫星应当选定不同的阈值。实际上,由于大气传输特性的影响,不同地区、不同时期的卫星影像都需要做具体实验来确定合适的阈值,这便在一定程度上制约了水体提取的大规模应用。

近些年,随着机器学习和人工智能的发展,基于机器学习的水体提取方法应运而生,避免了人工设定阈值对水体提取的影响。一类方法是将所提取的特征送入学习模型实现水体信息的提取,虽然这类方法也可以融合光谱、纹理和空间等信息,但是由于目前人工设计特征的特征描述能力仍然有较大的局限性,并不能充分利用遥感影像上的上下文信息。因此,比起传统的指数法,该类浅层学习模型的水体提取精度提升并不明显。然而,另一类方法——深度学习方法具有强大的特征表征能力,它的发展给计算机视觉和影像识别等领域带来了巨大的冲击,有时其性能可以与人类识别的精度相媲美。特别是文献[22]提出了全卷积神经网络(fully convolutional networks, FCN)结构后,卷积神经网络分类任务正式从场景级别进入到像素级别,并拓展到语义分割领域。同时,FCN被称为用于语义分割的第一个端到端的像素级预测的全卷积网络。随着影像分割精度的要求越来越高,在FCN的基础上,越来越多的网络如SegNet[23]、U-Net[24]、RefineNet[25]、DeepLab[26-28]、LRASPP[29]、U2Net[30]、MECNet[31-32]、MSResNet[33]和SADA-Net[20]等被提出,获得了广泛的研究和应用。

虽然目前主流的深度学习方法在水体提取方面取得了令人兴奋的结果,但在水陆边界处仍不能得到实用的效果。一方面,与自然影像相比,因大气传输、背景复杂等原因,遥感影像呈现结构多样、细节丰富、水体形状及大小无规则等特点,导致水体提取时在复杂的自然水体边界处更容易发生像素误分或漏分的情况。遥感影像上的水体边界信息往往因卷积和池化等操作在特征学习传递时被遗漏,因此,水体的细节保留较少。另一方面,在FCN中,不同层次或尺度的特征在融合时通常只是简单地相加或连接,并没有考虑到不同层次或尺度的特征在水体提取中的作用,导致了有的特征没有发挥出应有的效果。更为重要的是,在计算损失时完全忽略了每一个像素的位置信息。实际上,相比于水体内部区域的像素,水体边界的像素更容易被误分或漏分。

鉴于上述分析,本文提出了融合尺度注意力机制的加权FCN(squeeze-and-excitation weighted FCN, SE-wFCN),以高分辨率高分一号卫星遥感影像为研究对象,将SE-wFCN与目前主流的语义分割网络进行了对比。

1 原理与方法

1.1 总体网络结构

FCN的设计特点包括跳跃连接和多尺度融合分析,在一定程度上提升了语义分割的性能。虽然跳跃连接用于防止网络过深引起的梯度消失或梯度爆炸,但仍然存在训练时间过长和过拟合的可能性。多尺度融合分析是将浅层网络提取的信息和高层网络提取的信息融合在一起,可以更好地帮助理解图像。然而,在用FCN进行水体提取时普遍存在水陆边界不清晰等问题。针对上述问题,本文提出了SE-wFCN,更好地顾及到不同尺度特征对水体边界提取的重要性,其架构如图1所示。为了便于后续进一步描述不同模块之间的衔接,将FCN的部分网络结构细化,如图2所示。

1.2 辅助分类器

实际上,河流中的支流有两个作用:(1)为河流补充水源,因为河水也是在不断消耗的,需要补充。(2)当洪水来临时,支流会出现倒灌的现象(水从主河道流入支流,因为支流没有发洪水,水位较低),变相地延缓了洪峰。虽然这个现象不是很明显,但是支流的蓄水能力增加就能延缓更多的洪峰。受到大自然河流与支流两者之间相互作用的启发,为了更进一步防止误差梯度消失或梯度爆炸,设法将误差梯度传播到网络浅层,从而在网络浅层(即更细小的尺度上)学习到更

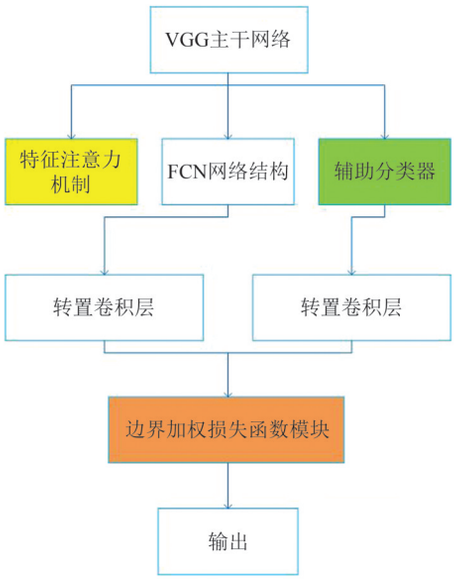


图 1　网络模块图

Fig. 1　Network Structure with Different Modules

好的特征表征能力。与此类似,在 SE-wFCN 中引入辅助分类器,作为主干网络的支流。在训练阶段,一方面它为主干网络降低误差梯度消失或梯度爆炸的可能,也可以进一步加快网络的训练速度。另一方面,它协助主干网络在细小的尺度上更好地学习特征的表征能力。而且,在尺度注意力机制和边界加权损失函数的组合引导下,网络在细小的尺度上更加关注水体边界处的特征表征与提取。具体的网络结构见图 3 中的辅助分类器部分,其输入来自图 2 中 FC6 的输出。

1.3　尺度注意力机制

本文尝试将注意力机制引入到高分辨率遥感影像水体提取中的解码部分,即在 FCN 提取的不同尺度或层次特征的基础上,借助于注意力机制的引导,在细小的尺度上更加关注水体的边界,从而有助于水体边界处的特征表示和提取。具体的网络结构见图 4。

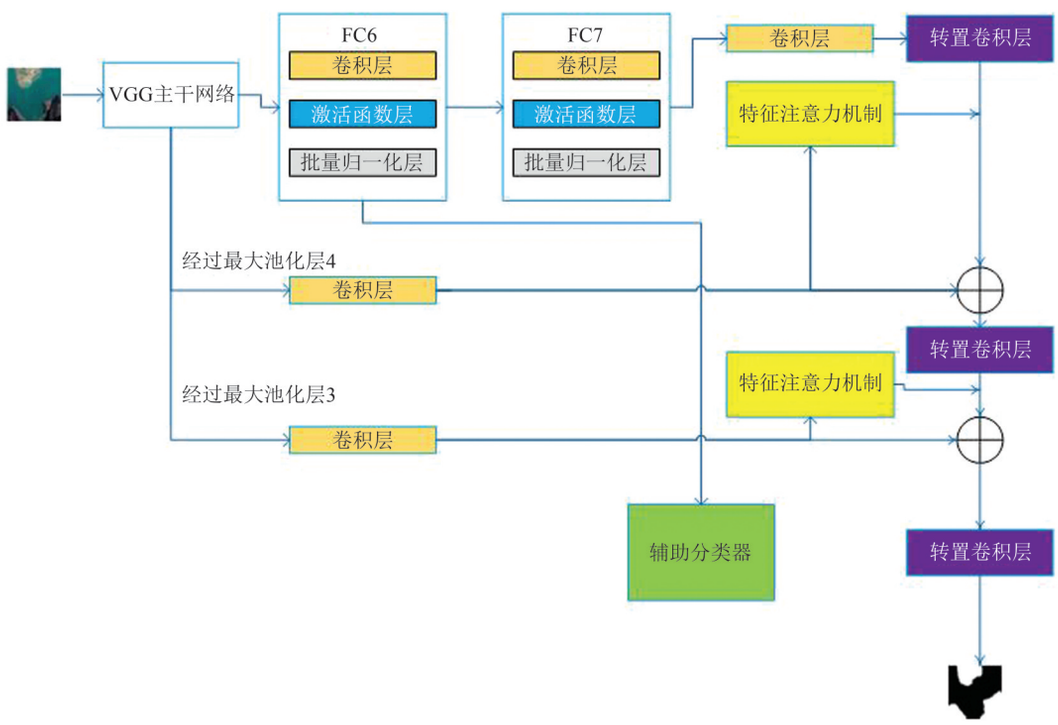


图 2　FCN 网络结构

Fig. 2　FCN Structure

注意力机制本质上是在矩阵转置相乘的基础上进行相关性的计算,依次增大依赖性较强特征的权重,降低噪声等的干扰,从而提高有效信息的最佳利用。在图 4 中,A 为输入的特征图,大小为 $C\times H\times W$(即特征数×高度×宽度)。首先对特征图 A 进行卷积,得到特征图 B、C 和 D,其宽度和高度与 A 一样。接着,将特征图 B 和 C 在特征维度进行维度重排分别得到 B_r 和 C_r,然后将重排后的 C_r 和转置后的 B_r 相乘(即 $B_r^{\mathrm{T}}\times C_r$)后,得到大小为 $(H\times W)\times(H\times W)$ 的注意力特征矩阵 E,E 的元素的计算公式为:

$$e_{ji}=\frac{\exp(B_i\cdot C_j)}{\sum_{i=1}^{N}\exp(B_i\cdot C_j)} \tag{1}$$

式中,B_i 表示矩阵 B 中的第 i 行组成的向量;C_j 表示矩阵 C 中的第 j 行组成的向量;$B_i\cdot C_j$ 表示两个

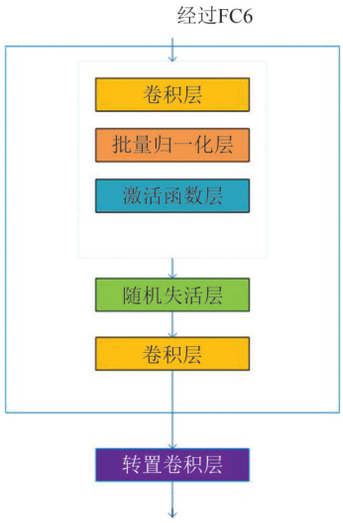


图3 辅助分类器的结构

Fig. 3 Structure of Auxiliary Classifier

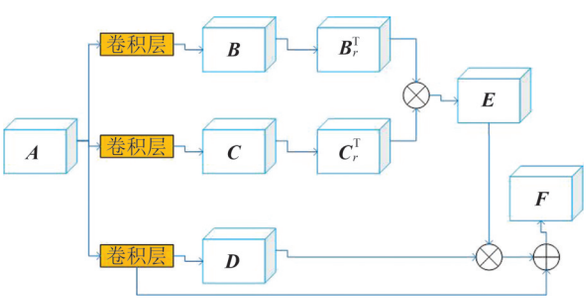


图4 尺度注意力机制结构

Fig. 4 Structure of Scale Attention Mechanism

向量的内积,即$B_iC_j^{\mathrm{T}}$;e_{ji}表示第i个像素和第j个像素之间的相关性。

最后,将E与特征图D进行矩阵相乘(即$D\times E$)后,再与原始特征图进行加和(即$D\times E+A$)得到注意力特征图F,计算公式为:

$$F_j=\alpha\sum_{i=1}^{N}(e_{ij}D_i)+A_j \tag{2}$$

式中,α为可训练的权重系数;F_j表示矩阵F中的第j行组成的向量;D_i表示矩阵D中的第i行组成的向量;A_j表示矩阵A中的第j行组成的向量。

1.4 边界加权损失函数

在水体提取中,最容易出现错分或漏分的是水陆边界,造成了水陆分界线的模糊和不确定。针对这个不足,在训练阶段,相对于水体内部的像素,本文对水陆边界处的像素赋予更高的效用权重,并借助于尺度注意力机制,在不同尺度进行有效的引导,更加准确地学习水陆边界的特征表征能力。具体的损失函数为:

$$L=w_1\times\sum_{i\in B}L(P_i,T_i)+w_2\times\sum_{i\notin B}L(P_i,T_i) \tag{3}$$

式中,B表示水陆分界线像素的集合;$L(P_i,T_i)$表示二分类的交叉熵损失函数;P_i表示第i个像素的预测值;T_i表示第i个像素的地面真实值;$w_1+w_2=1$。

2 高分辨率遥感影像水体提取实验

实验平台搭载Ubuntu 18.04.6 LTS操作系统,部署PyTorch 1.12.0(包括cuda10.2)和Python 3.9.13版本,配置Intel(R) Core™ i5-10400F@2.90 GHz 6核12线程CPU处理器,32 GB RAM和1块NAIDIA RTX 3060显卡,显存为12 GB。在深度学习模型训练过程中,均使用随机梯度下降算法对所有的模型进行优化,基础学习率设置为0.1,权重衰减因子为1×10^{-4},动量值为0.9。在训练过程中,采用文献[31]的方法进行数据增强,同时迭代次数设置为200。

2.1 实验数据集

采用的数据集是高分影像数据集(Gaofen image dataset,GID)中提取的有关水体类别的数据集,称为GID-W数据集。GID数据集是基于中国高分二号卫星数据而构建的大规模高分辨率遥感图像土地覆盖数据集。它具有两个特点:(1)土地覆盖信息分布广泛;(2)贴近真实的地物分布特性。GID数据集分为大规模分类集(GID-5)和精细土地覆盖集(GID-15)两个部分。GID-5包含建筑、农田、森林、草地和水域等5个土地覆盖类别,GID-15包含稻田、灌溉地、旱地、园地、乔木林、灌木林、自然草地、人工草地、工业用地、城市住宅、农村住宅、交通用地、河流、湖泊、池塘等15个类别。GID-W数据集是将GID-5中的水域部分和GID-15中的河流、湖泊和池塘3个类别单独提取出来,统一作为本文实验中的水体类别,包含5 490组真彩色遥感影像以及对应的地面真实标签,影像大小固定为224×224像素,按5∶1∶4比例随机划分为训练集、验证集和测试集,用以评估SE-wFCN性能。图5展示了数据集部分样本的示例(黑色表示水体,白色表示非水体)。

2.2 评价指标

本文选择总体精度(overall accuracy,OA)、Kappa系数、F-measure分数、精确率P、召回率R、平均交并比(mean intersection over union,mIoU)、绝对平均误差(mean absolute error,MAE)、筛子系数D等8个指标来评价水体提取的结果。

总体精度O_A是指提取的正确预测像素数占所有像素数的比值,计算公式为:

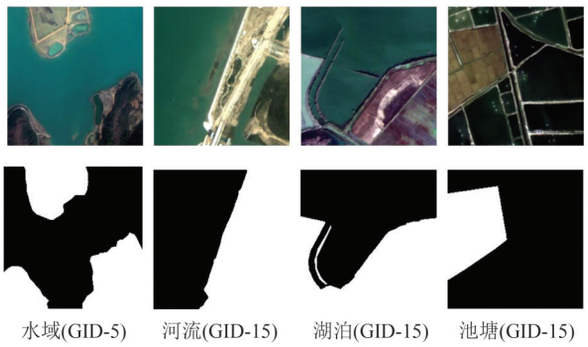


图5　实验数据集示例

Fig. 5　Examples from Experimental Datasets

精确率P_r是衡量提取结果是否有误判的指标,计算公式为:

$$P=\frac{N_{TP}}{N_{TP}+N_{FP}} \tag{7}$$

召回率R是衡量提取结果是否有遗漏的指标,计算公式为:

$$R=\frac{N_{TP}}{N_{TP}+N_{FN}} \tag{8}$$

$F1$分数是计算机领域中衡量二分类精度的评价指标,它是结合了精确率和召回率的综合指标。当$\beta=1$时,F_β便等于$F1$。计算公式为:

$$F_\beta=\frac{(1+\beta^2)\times P\times R}{\beta^2\times P+R} \tag{9}$$

交并比(intersection over union, IoU)是目标检测和语义分割中的常用指标,是指提取结果与地面真实标签交叠率,即它们的交集与并集的比值。mIoU是所有类别IoU的平均值。IoU的计算公式为:

$$P_{IoU}=\frac{N_{TP}}{N_{FP}+N_{TP}+N_{FN}} \tag{10}$$

式中,P_{IoU}表示IoU的值。

MAE是将每一个像素与地面真实标签进行比较的平均值,计算公式为:

$$E_{MAE}=\frac{1}{H\times W}\sum_{r=1}^{H}\sum_{c=1}^{W}|P(r,c)-G(r,c)| \tag{11}$$

式中,E_{MAE}表示MAE的值;P表示网络预测的概率图;G表示真实的地面标签;r、c分别表示影像中的行号和列号。

筛子系数D是医学图像比赛中使用频率最高的度量指标,它是一种集合相似度度量指标,通常用于计算两个样本的相似度,其值域为[0, 1],计算公式为:

$$O_A=\frac{N_{TP}+N_{TN}}{N_{TP}+N_{TN}+N_{FP}+N_{FN}} \tag{4}$$

式中,N_{TP}表示真实水体被预测为水体的像素个数;N_{TN}表示背景像素被预测为背景像素的数量;N_{FP}表示背景像素被预测为水体像素的数量;N_{FN}表示真实水体像素被预测为背景像素的数量。

Kappa系数K是用来平衡样本规模的一种精度指标,计算公式为:

$$K=\frac{O_A-P_e}{1-P_e} \tag{5}$$

其中,

$$P_e=\frac{(N_{TP}+N_{FN})\times(N_{TP}+N_{FP})+(N_{FN}+N_{TN})\times(N_{FP}+N_{TN})}{(N_{TP}+N_{TN}+N_{FN}+N_{FP})^2} \tag{6}$$

$$D=1-2\times\frac{|V\cap T_{val}|}{|V|+|T_{val}|} \tag{12}$$

式中,V代表预测值集合;T_{val}代表地面的真实标签集合;$|V\cap T_{val}|$表示两个集合中相同元素的数量。

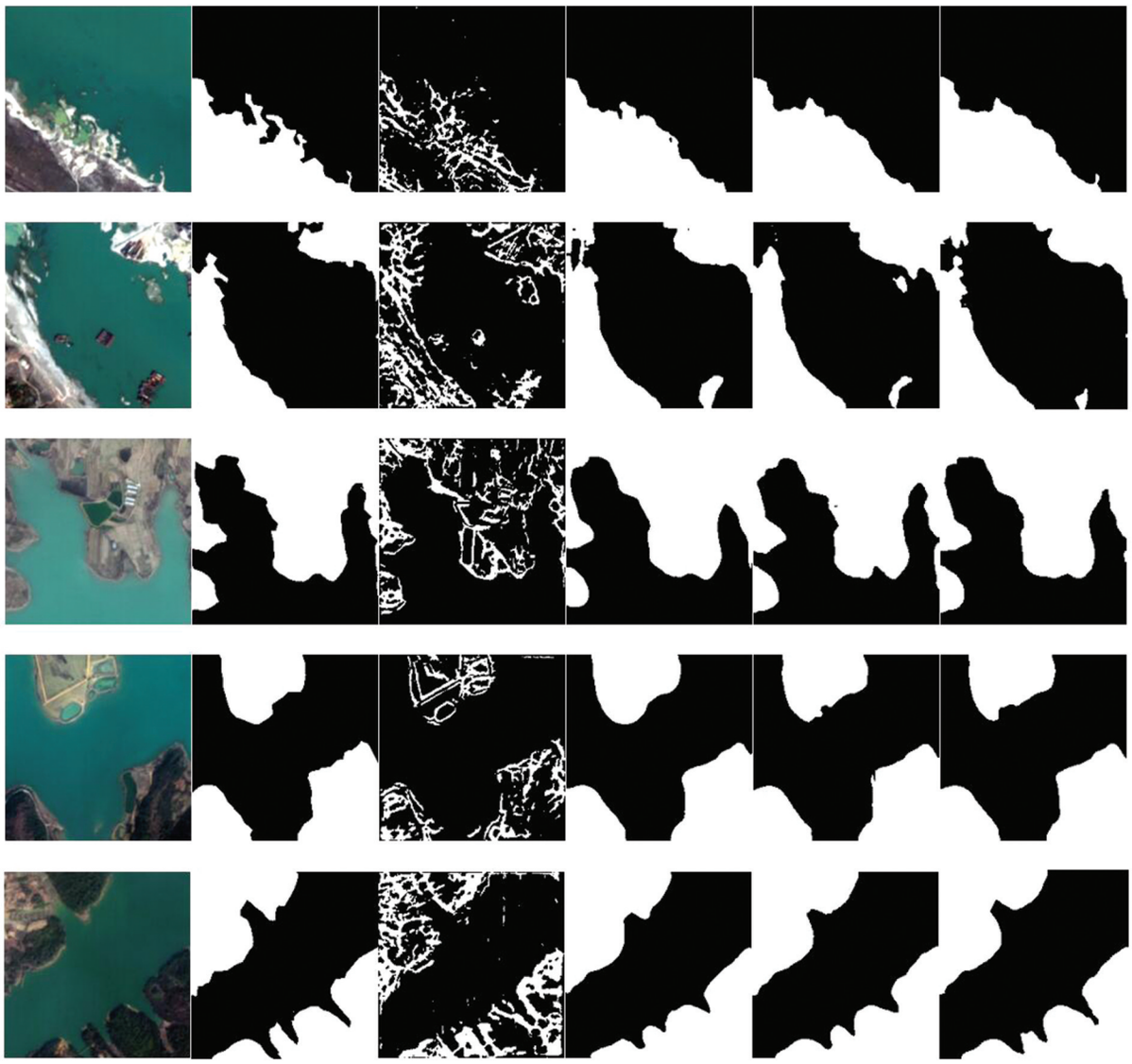
2.3　实验结果与分析

2.3.1　水体提取结果与主观分析

针对GID-W数据集,分别采用FCN、DeepLabV3、U-Net、U2Net和SE-wFCN进行水体提取实验。限于篇幅,只列出部分实验结果(见图6),一共选取了5个样本的水体提取结果(即每一行),每一列依次表示真彩色遥感影像、地面真实标签(黑色表示水体,白色表示非水体)、U-Net提取结果、U2Net提取结果、DeepLabV3提取结果和SE-wFCN提取结果。从目视效果上分析可知,SE-wFCN在水陆边界处更加清晰,与真实情况也更加接近。从图6中的水陆边界处来看,U-Net和U2Net都只能提取到水陆边界的大概位置,而SE-wFCN方法成功提取了水陆边界处的细节。此外,U-Net的提取结果存在明显的噪声和较多的误提和漏提。U2Net和DeepLabV3的边界提取效果较好,但不如SE-wFCN。

2.3.2　网络收敛速度比较分析

为了进一步为主干网络降低误差梯度消失或梯度爆炸的可能,本文提出在原始的FCN基础上引入辅助分类器,帮助主干网络进行分流,使得误差在反向传播时能够达到网络中的低层特征。实际上,它协助主干网络充分地利用了现有的训练样本,进一步降低过拟合的可能性,更好地学习网络低层(即细小尺度上)的表征能力。图7是原始FCN和SE-wFCN两种方法的训练误差和测试误差随着迭代次数变化的曲线图,可以看出FCN的训练误差和测试误差,特别是训练误



(a) 原图 (b) 标签 (c) U-Net (d) U2Net (e) Deeplabv3 (f) SE-wFCN

图6 不同方法的水体提取结果比较

Fig. 6 Comparison of Water Extract Based on Different Methods

差,明显低于SE-wFCN,其误差大小更加接近于零,呈现出了过拟合的现象,而且在测试过程中,测试误差大于训练误差,其误差的波动比较大,在图7中可以看到一些零星的红点。而相比而言,SE-wFCN表现比较稳定,测试误差小于训练误差,而且误差的波动也较小,随着迭代次数的逐步增加趋于稳定。

2.3.3 尺度注意力机制对水体提取的影响

为了进一步提高水陆边界处的提取精度,本文将尺度注意力机制引入FCN,在细小的尺度上更加关注水体的边界。限于篇幅,只列出部分实验结果(见图8),可以看出随着方法的逐步改进,从原始FCN方法(标记为FCN0)到原始FCN+辅助分类器(标记为FCN1)、原始FCN+辅助分类器+尺度注意力机制(标记为FCN2)和原始FCN+辅助分类器+尺度注意力机制+边界加权损失函数(即SE-wFCN)方法,水陆边界的提取变得更加清晰和准确,如图8中红色圆形部分。

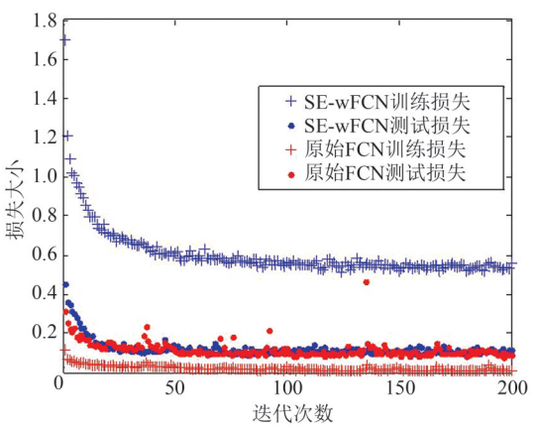
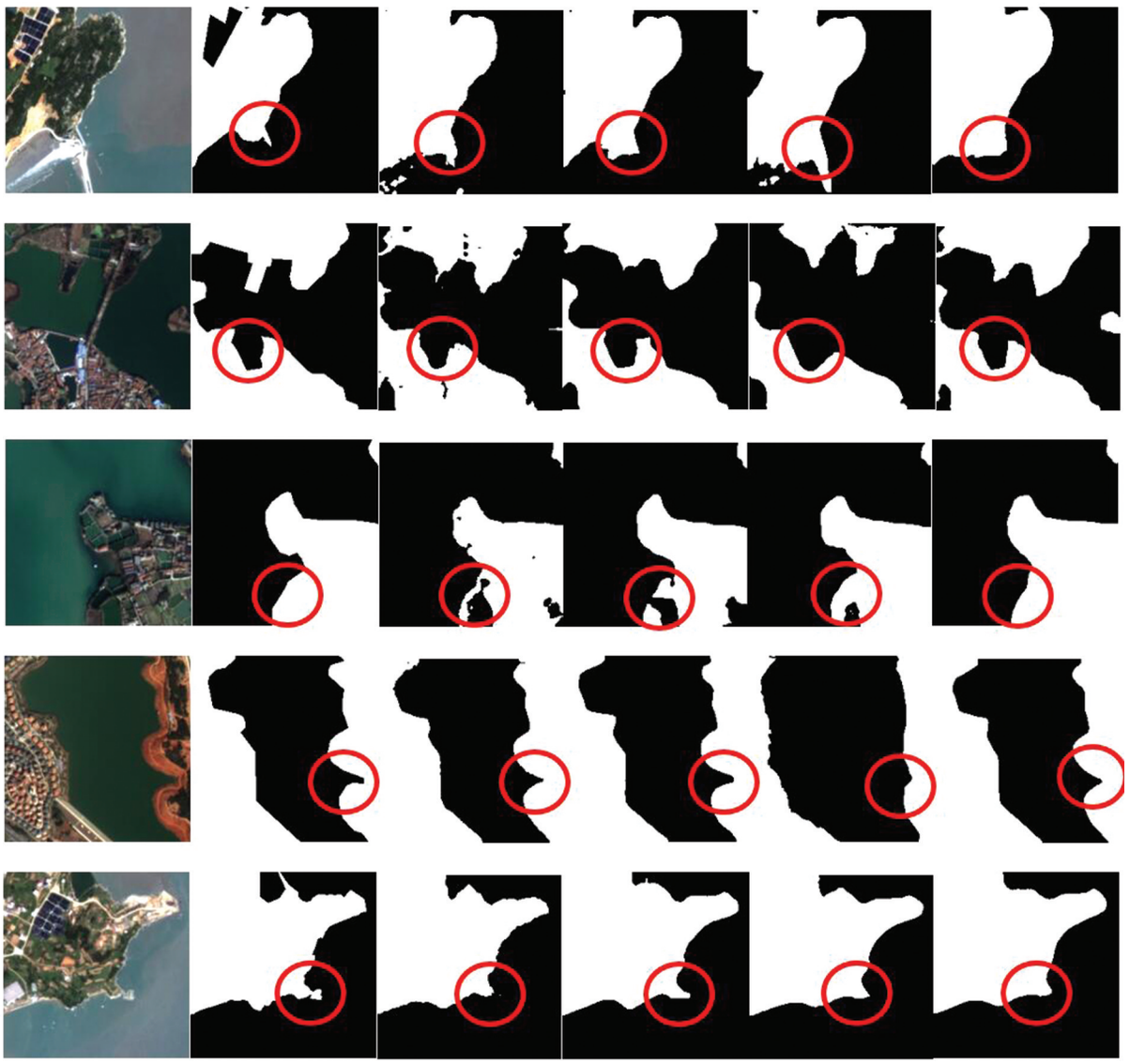


图7 两种不同方法下的误差曲线

Fig. 7 Error Curve Based on Two Different Methods



(a) 原图　(b) 标签　(c) FCN0　(d) FCN1　(e) FCN2　(f) SE-wFCN

图 8　尺度注意力机制对水体提取的影响比较

Fig. 8　Comparison of Scale Attention Affecting Water Extract

2.3.4　水体提取精度比较分析

为了更好地比较不同方法对水体提取结果的差异,本文采用 8 个评价指标进行定量分析,结果见表 1。从表 1 中可以发现,从 mIoU(其他指标类似)来看,在 FCN0 模型的基础上引入辅助分类器后,FCN1 比 FCN0 提高了 1.35%。在 FCN1 基础上增加特征注意力机制后,FCN2 比 FCN1 又提高了 1.19%。在 FCN2 基础上采用边界加权损失函数后,SE-wFCN 比 FCN2 提高了 3.27%。而且,与当前主流的深度学习模型相比,SE-wFCN 比 U-Net 方法在 mIoU 指标上提高了 17.72%,比 DeepLabV3 提高了 4.88%,比 U2Net 提高了 5.62%。此外,与 MECNet、MSResNet 和 SADA-Net 相比,在 $F1$ 指标上,SE-wFCN 比它们都提高了约 10%。

2.3.5　网络复杂度分析

在神经网络中,参数量和计算量是衡量神经网络模型的两个重要指标。本文对不同方法的模型复杂度及效率进行了比较与分析。模型的参数量用网络模型中含有的参数在内存中的大小来表示,主要影响模型运行时所需的内存和现存大小,而模型的计算量采用浮点运算数量来表示,它决定网络运行的快慢,本文使用 thop 工具包(一个轻量级的 PyThosch 工具包)进行统计。训练时间为迭代一次训练数据集所需要的时间,推理时间为在测试数据集上推理所需要的总时间。从表 2 中可以看出,SE-wFCN 的计算量较少,按模型大小来看,仅有 U2Net 的 1/7 左右,只有 DeepLabV3 的 1/4。按训练时间来看,SE-wFCN 只有 DeepLabV3 的 1/2,与 U2Net 一样。而

按测试时间来看,SE-wFCN只有DeepLabV3的1/2左右,只有U2Net的2/3。因此,SE-wFCN在精度和效率方面取得了较好的平衡点,具有较高的应用价值。

表1 不同方法的性能比较

Table 1 Performance Comparison of Different Methods

深度学习模型	总体精度	召回率	精确率	*F*1分数	Kappa系数	mIoU	MAE	筛子系数
DeepLabV3	0.960 6	0.988 6	0.966 3	0.977 3	0.826 4	0.847 8	0.039 4	0.971 2
U-Net	0.904 0	0.815 8	0.624 7	0.707 6	0.651 3	0.719 4	0.096 0	0.764 4
U2Net	0.964 2	0.768 5	0.919 9	0.837 4	0.817 4	0.840 4	0.035 8	0.039 3
FCN0	0.955 6	0.975 7	0.972 7	0.974 2	0.816 3	0.838 5	0.044 4	0.967 1
FCN1	0.961 5	0.987 7	0.968 1	0.977 8	0.833 3	0.852 0	0.038 5	0.972 4
FCN2	0.964 7	0.988 3	0.971 1	0.979 6	0.839 4	0.863 9	0.035 3	0.974 8
SE-wFCN	0.973 4	0.990 0	0.979 3	0.984 6	0.887 8	0.896 6	0.026 6	0.980 7
MECNet[32]	0.947 4	0.838 6	0.855 4	0.847 0				
MSResNet[33]	0.954 3	0.852 9	0.880 5	0.866 5				
SADA-Net[20]	0.961 8	0.853 6	0.920 3	0.885 7				

表2 网络结构复杂度对比

Table 2 Comparison of Network Structure Complexity

深度学习模型	每秒十亿次浮点运算量	参数量/(10^6个)	训练时间/s	测试时间/s	模型大小/MB
DeepLabV3	41.994 3	33.264 8	41	11	336.4
U-Net	7.755 4	4.318 0	16	1	34.6
U2Net	28.828 2	43.995 5	20	9	528.7
FCN0	19.524 1	18.643 7	19	6	74.6
FCN1	20.430 1	19.307 6	19	6	78.6
FCN2	21.994 3	20.264 8	20	6	82.9
SE-wFCN	21.994 3	20.264 8	20	6	82.9

3 结 语

本文提出了SE-wFCN用于改善高分辨率遥感影像水体提取中水陆边界模糊和不确定性等问题。首先,本文考虑到FCN模型本身的复杂性,在FCN中引入辅助分类器模块进一步降低深度学习模型在训练过程中发生过拟合的可能性。其次,本文将注意力机制应用于不同层次或尺度的语义信息中,最大限度地发挥尺度注意力机制模块的优势,更好地融合不同尺度的特征,从而尽可能地减少水陆边界像素的错分现象。然后,顾及水体边界的像素位置信息对水体提取的重要性,在损失计算中对水体边界像素与水体内部像素的损失赋予不同的权重,引导FCN模型更好地学习水陆边界的特征信息,进一步地提高水体边界的提取精度。大量的实验与分析表明,SE-wFCN不仅可以进一步降低过拟合的可能性,而且可以在水陆边界处学习到更好细节特征,使得水体提取的mIoU比DeepLabV3和U2Net分别提高了4.88%和5.62%。在网络模型复杂度方面,都比DeepLabV3和U2Net有更大的优势。因而,SE-wFCN在精度和效率方面取得了更好的平衡点,具有较高的应用价值。

参 考 文 献

[1] CHEN S T, LIU Y, ZHANG C. Water-Body Segmentation for Multi-spectral Remote Sensing Images by Feature Pyramid Enhancement and Pixel Pair Matching[J]. *International Journal of Remote Sensing*, 2021, 42(13): 5025-5043.

[2] 吴庆双, 汪明秀, 申茜, 等. Sentinel-2遥感图像的细小水体提取[J]. 遥感学报, 2022, 26(4): 781-794.
WU Qingshuang, WANG Mingxiu, SHEN Qian, et al. Small Water Body Extraction Method Based on Sentinel-2 Satellite Multi-spectral Remote Sensing Image [J]. *National Remote Sensing Bulletin*, 2022, 26(4): 781-794.

[3] 眭海刚, 陈光, 胡传文, 等. 光学遥感影像与GIS数据一体化的水体分割、配准与提取方法[J]. 武汉大学学报(信息科学版), 2016, 41(9): 1145-

1150.
SUI Haigang, CHEN Guang, HU Chuanwen, et al. Integrated Segmentation, Registration and Extraction Method for Water-Body Using Optical Remote Sensing Images and GIS Data[J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(9): 1145-1150.

[4] 王小鹏,文昊天,王伟,等. 形态学边缘检测和区域生长相结合的遥感图像水体分割[J]. 测绘科学技术学报,2019,36(2):149-154.
WANG Xiaopeng, WEN Haotian, WANG Wei, et al. Water Segmentation of Remote Sensing Image Using Morphological Edge Detection and Region Growing[J]. *Journal of Geomatics Science and Technology*,2019,36(2):149-154.

[5] LÜ Y L, TIAN S W, YU L, et al. Water Body Semantic Information Description and Recognition Based on Multimodal Models[J]. *International Journal of Computational Intelligence and Applications*, 2019, 18(2): 1950015.

[6] SUN J Y, WANG G Z, HE G J, et al. Study on the Water Body Extraction Using GF-1 Data Based on Adaboost Integrated Learning Algorithm [J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2020, XLII-3/W10: 641-648.

[7] WANG E D, JIANG Y M, LI Y, et al. MFCS-Net: Multi-scale Deep Features Fusion and Cost-Sensitive Loss Function Based Segmentation Network for Remote Sensing Images[J]. *Applied Sciences*, 2019, 9(19): 4043.

[8] DONG S, PANG L, ZHUANG Y, et al. Optical Remote Sensing Water-Land Segmentation Representation Based on Proposed SNS-CNN Network [C]//IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 2019.

[9] 王小标,谢顺平,都金康. 水体指数构建及其在复杂环境下有效性研究[J]. 遥感学报,2018,22(2):360-372.
WANG Xiaobiao, XIE Shunping, DU Jinkang. Water Index Formulation and Its Effectiveness Research on the Complicated Surface Water Surroundings[J]. *Journal of Remote Sensing*, 2018, 22(2): 360-372.

[10] 种丹,张世强,李浩杰,等. 基于光学遥感的湖泊水体信息提取研究进展[J]. 人民黄河,2018,40(3):49-53.
CHONG Dan, ZHANG Shiqiang, LI Haojie, et al. Review on the Extraction of Lake Water Body Information from Optics Remote Sensing Data[J]. *Yellow River*, 2018, 40(3): 49-53.

[11] MCFEETERS S K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features[J]. *International Journal of Remote Sensing*, 1996, 17(7): 1425-1432.

[12] XU H Q. Modification of Normalised Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery[J]. *International Journal of Remote Sensing*, 2006, 27(14): 3025-3033.

[13] 曹荣龙,李存军,刘良云,等. 基于水体指数的密云水库面积提取及变化监测[J]. 测绘科学,2008,33(2):158-160.
CAO Ronglong, LI Cunjun, LIU Liangyun, et al. Extracting Miyun Reservoir's Water Area and Monitoring Its Change Based on a Revised Normalized Different Water Index[J]. *Science of Surveying and Mapping*, 2008, 33(2): 158-160.

[14] DANAHER T,COLLETT L. Development,Optimisation and Multi-temporal Application of aA Simple Landsat Based Water Index[C]//The 13th Australasian Remote Sensing and Photogrammetry Conference,Canberra,Australia, 2024.

[15] FISHER A,FLOOD N,DANAHER T. Comparing Landsat Water Index Methods for Automated Water Classification in Eastern Australia[J]. *Remote Sensing of Environment*, 2016, 175: 167-182.

[16] 丁凤. 基于新型水体指数(NWI)进行水体信息提取的实验研究[J]. 测绘科学,2009,34(4):155-157.
DING Feng. Study on Information Extraction of Water Body with a New Water Index (NWI)[J]. *Science of Surveying and Mapping*, 2009, 34(4): 155-157.

[17] 万建鹏,官云兰,叶素倩,等. 基于综合权重水体指数的水体提取研究:以鄱阳湖为例[J]. 东华理工大学学报(自然科学版),2015,38(2):206-211.
WAN Jianpeng, GUAN Yunlan, YE Suqian, et al. Water Extraction Based on Comprehensive Weight Water Index: A Case Study in Region of Poyang Lake[J]. *Journal of East China University of Technology (Natural Science)*, 2015, 38(2): 206-211.

[18] FEYISA G L, MEILBY H, FENSHOLT R, et al. Automated Water Extraction Index: A New Technique for Surface Water Mapping Using Landsat Imagery[J]. *Remote Sensing of Environment*, 2014, 140: 23-35.

[19] 巫兆聪,喻文蕴. 不同地表覆盖下SPOT-7影像大气校正方法研究[J]. 测绘地理信息,2021,46(2):7-12.

DU Jinkang, HUANG Yongsheng, FENG Xuezhi, et al. Study on Water Bodies Extraction and Classification from SPOT Image[J]. *Journal of Remote Sensing*, 2001, 5(3): 214-219.

[20] WANG B, CHEN Z L, WU L, et al. SADA-Net: A Shape Feature Optimization and Multiscale Context Information-Based Water Body Extraction Method for High-Resolution Remote Sensing Images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 1744-1759.

[21] 洪亮, 黄雅君, 杨昆, 等. 复杂环境下高分二号遥感影像的城市地表水体提取[J]. 遥感学报, 2019, 23(5): 871-882.
HONG Liang, HUANG Yajun, YANG Kun, et al. Study on Urban Surface Water Extraction from Heterogeneous Environments Using GF-2 Remotely Sensed Images[J]. *Journal of Remote Sensing*, 2019, 23(5): 871-882.

[22] LONG J, SHELHAMER E, DARRELL T. Fully Convolutional Networks for Semantic Segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 2015.

[23] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.

[24] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]//The 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015.

[25] 张兵. 遥感大数据时代与智能信息提取[J]. 武汉大学学报(信息科学版), 2018, 43(12): 1861-1871.
ZHANG Bing. Remotely Sensed Big Data Era and Intelligent Information Extraction[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(12): 1861-1871.

[26] 陈志坤, 江俊君, 姜鑫维, 等. 一种基于改进双边滤波的鲁棒高光谱遥感图像特征提取方法[J]. 武汉大学学报(信息科学版), 2020, 45(4): 504-510.
CHEN Zhikun, JIANG Junjun, JIANG Xinwei, et al. Robust Hyperspectral Image Feature Extraction Based on Improved Bilateral Filtering[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(4): 504-510.

[27] ZHANG Z L, LU M, JI S P, et al. Rich CNN Features for Water-Body Segmentation from very High Resolution Aerial and Satellite Imagery[J]. *Remote Sensing*, 2021, 13(10): 1912.

[28] DANG B, LI Y S. MSResNet: Multiscale Residual Network via Self-Supervised Learning for Water-Body Detection in Remote Sensing Imagery[J]. *Remote Sensing*, 2021, 13(16): 3122.

[29] 陈立福, 龙凤琪, 李振洪, 等. 面向多源SAR图像的多级特征注意力水体提取网络[J]. 武汉大学学报(信息科学版), 2023, DOI: 10.13203/j.whugis20230041.
CHEN Lifu, LONG Fengqi, LI Zhenhong, et al. Multi-level Feature Attention Water Body Extraction Network for Multi-source SAR Images[J]. *Geomatics and Information Science of Wuhan University*, 2023, DOI: 10.13203/j.whugis20230041.

[30] 覃星力, 杨杰, 李平湘, 等. 迁移学习用于多时相极化SAR影像的水体提取[J]. 武汉大学学报(信息科学版), 2022, 47(7): 1093-1102.
QIN Xingli, YANG Jie, LI Pingxiang, et al. Water Body Extraction from Multi-temporal Polarimetric SAR Images Based on Transfer Learning[J]. *Geomatics and Information Science of Wuhan University*, 2022, 47(7): 1093-1102.

[31] 闫利, 周建彤. 单张无人机影像水体高光自动检测与补偿方法[J]. 武汉大学学报(信息科学版), 2018, 43(10): 1511-1517.
YAN Li, ZHOU Jiantong. A Method for Automatic Water High Light Detection and Removal in Single UAV Image[J]. *Geomatics and Information Science of Wuhan University*, 2018, 43(10): 1511-1517.

[32] 眭海刚, 陈光, 胡传文, 等. 光学遥感影像与GIS数据一体化的水体分割、配准与提取方法[J]. 武汉大学学报(信息科学版), 2016, 41(9): 1145-1150.
SUI Haigang, CHEN Guang, HU Chuanwen, et al. Integrated Segmentation, Registration and Extraction Method for Water-Body Using Optical Remote Sensing Images and GIS Data[J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(9): 1145-1150.

[33] 徐川, 华凤, 眭海刚, 等. 多尺度水平集SAR影像水体自动分割方法[J]. 武汉大学学报(信息科学版), 2014, 39(1): 27-31.
XU Chuan, HUA Feng, SUI Haigang, et al. Automatic Water Segmentation Method in SAR Images Using Multi-scale Level Set[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(1): 27-31.