

引用格式:何红术,黄晓霞,李红杏,等.基于改进U-Net网络的高分遥感影像水体提取[J].地球信息科学学报,2020,22(10):2010-2022. [ He H S, Huang X X, Li H G, et al. Water body extraction of high resolution remote sensing image based on improved U-Net network[J]. Journal of Geo-information Science, 2020,22(10):2010-2022. ] DOI:10.12082/dqxxkx.2020.190622

## 基于改进U-Net网络的高分遥感影像水体提取

何红术<sup>1,2</sup>,黄晓霞<sup>1\*</sup>,李红杏<sup>1</sup>,倪凌佳<sup>1,2</sup>,王新歌<sup>3</sup>,陈崇<sup>3</sup>,柳泽<sup>3</sup>

1. 中国科学院空天信息创新研究院,北京 100094; 2. 中国科学院大学,北京 100049;
3. 住房和城乡建设部城乡规划管理中心,北京 100835

### Water Body Extraction of High Resolution Remote Sensing Image based on Improved U-Net Network

HE Hongshu<sup>1,2</sup>, HUANG Xiaoxia<sup>1\*</sup>, LI Hongga<sup>1</sup>, NI Lingjia<sup>1,2</sup>, WANG Xingge<sup>3</sup>, CHEN Chong<sup>3</sup>, LIU Ze<sup>3</sup>

1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. Urban and Rural Planning Management Center of the Ministry of Housing and Urban-Rural Development, Beijing 100835, China

**Abstract:** There are two main methods of traditional water body extraction: a method based on spectral information and a method based on classification. Traditional water body extraction methods based on spectral information fail to take into account features such as water body shape, internal texture, water body size, and adjacent relations of water body. Also, there is a common phenomenon of “same object with different spectra and same spectrum with different objects”, which could result in low accuracy of water body extraction. Thus, the traditional methods that design features based on classification to extract water body is complex and impossible to capture the deep information of water body features. This paper proposed an improved U-Net network semantic segmentation method, which uses the de-encoding structure of the classic U-Net network to improve the network: ① Use the VGG network to shrink the path and increase the depth of the network to extract deep features of the water; ② Strengthen the low-dimensional feature information in the expansion path, fuse the feature map on the next layer of the shrinking feature pyramid with the feature map on the corresponding expansion path in the next layer, and enhance the model's low-dimensional feature information to improve the classification accuracy of the model; and ③ The Conditional Random Field (CRF) was introduced in the post-classification process to refine the segmentation results and improve the segmentation accuracy. In the study of Qingdao area, SegNet, classic U-Net network, and improved U-Net network were selected as controlled experiments while maintaining the same training set, validation set, and test set. The test results show that the improved U-Net network structure performed better than SegNet and classic U-Net networks in terms of IoU,

收稿日期:2019-10-24;修回日期:2020-02-16.

基金项目:国家重点研发计划项目(2017YFB0503905);国家自然科学基金项目(41971363)。[ **Foundation items:** National Key Research and Development Program of China, No.2017YFB0503905; National Natural Science Foundation of China, No.41971363. ]

作者简介:何红术(1993—),男,重庆人,硕士生,研究方向为深度学习遥感信息提取、地理信息系统。

E-mail: hhs10010@163.com

\*通讯作者:黄晓霞(1970—),女,广东梅县人,博士,研究员,研究方向资源环境遥感、灾害应急救援仿真。

E-mail: huangxx@aircas.ac.cn

accuracy rate and Kappa coefficient. Compared with SegNet, the three indicators increased by 10.5%, 12.3%, and 0.14, respectively. Compared with the results of the classic U-Net network, each indicator increased by 5.8%, 4.4% and 0.05, respectively. The results demonstrate the effectiveness of the improved method in this paper. In addition, the method proposed in this paper has more advantages than the other two networks in the extraction of small targets in the test area, the completeness of water body extraction, the distinction between shadows and water bodies, and the accuracy of boundary segmentation. In order to verify the space-time scalability of the model, this paper chose western Qingdao and Xining, Qinghai as the verification areas. The verification results show that the water body extraction was good for areas similar to the geographical environment of the experimental area, and the effect of water body extraction needs to be further improved in places that differ greatly from the geographical environment of the experimental area. In general, the improved U-Net network can effectively achieve the task of water extraction.

**Key words:** U-Net; water extraction; high-resolution remote sensing image; conditional random field; image segmentation; VGG16; Qingdao; Xining

**\*Corresponding author:** HUANG Xiaoxia, E-mail: huangxx@aircas.ac.cn

**摘要:**传统基于光谱信息的水体提取未能考虑水体形状、纹理、大小、相邻关系等问题,且存在同物异谱、异物同谱现象,导致水体提取精度较低。而传统基于分类提取水体方法设计特征过程较为繁琐,且不能挖掘深度信息特征。因此,本文提出改进的U-Net网络语义分割方法,借鉴经典U-Net网络的解编码结构对网络进行改进:①将VGG网络用于收缩路径以提取特征;②在扩张路径中对低维特征信息进行加强,将收缩特征金字塔上一层的特征图与下一层对应扩张路径上的特征图进行融合,以提高提取结果分割精度;③在分类后处理中引入条件随机场,以将分割结果精细化。在保持相同训练集、验证集和测试集的情况下,分别用SegNet、经典U-Net网络和改进的U-Net网络做对照试验。试验结果表明,改进的U-Net网络结构在IoU、精准率和Kappa系数指标上均高于SegNet和经典U-Net网络,与SegNet相比,3项指标分别提升了10.5%、12.3%和0.14,与经典U-Net网络结果相比,各个指标分别提升了5.8%、4.4%和0.05。改进的网络水体提取结果较为完整,对小目标水体能够准确提取。改进的U-Net网络能够有效地实现水体提取任务。

**关键词:**U-Net; 水体提取; 高分遥感影像; 条件随机场; 图像分割; VGG16; 青岛; 西宁

## 1 引言

环境变化与人类社会息息相关,地球表面74%都是水体,研究水体变化切实关乎人类利益<sup>[1-3]</sup>。遥感因具有宏观、快速和同步的优点,自20世纪70年代起,各国竞相发展遥感事业,基于此遥感技术迅猛发展,遥感越来越多地应用在资源调查、环境变化、气象监测等应用中<sup>[4]</sup>。近年来,研究人员利用遥感优势,将遥感技术应用在水体提取研究中,针对光学遥感影像特点,国内外学者提出了多种水体提取方法,主要分为2类:基于像元分类的阈值法和基于目标分类的分类法。阈值法多利用在中分辨率遥感影像上,基于目标物的光谱特征,利用光谱知识构建各种水体指数来进行水体的提取;分类法更适用于高分辨率影像,综合影像的光谱、纹理和空间特征,充分利用地物的光谱、形状、结构和纹理等特征来提取水体信息<sup>[5]</sup>。Mcfeeters<sup>[6]</sup>根据水体光谱特征,提出归一化差异水体指数(NDWI)法提取水

体。针对NDWI的不足,Xu<sup>[7]</sup>提出改进的归一化差异水体指数(MNDWI),在大型水体提取中精度达到99%。针对单一方法很难有效地提取水体,国内学者提出在各节点设计不同分类器,利用决策树方法能有效提取水体<sup>[8-9]</sup>,精度达到90%。针对雷达影像,胡德勇等<sup>[10]</sup>利用雷达影像纹理信息,使用支持向量机(SVM)分类器进行水体信息提取,较准确地提取了水体信息,水体提取精度为84%。

自2006年Geoffrey Hinton<sup>[11]</sup>提出深度学习思想后,深度学习在学界和工业界得到越来越多的研究人员关注,在自然语言处理和图像识别等方面均获得了突破性的进展。随着深度学习的发展,深度卷积神经网络在图像语义分割领域也展现出了优势,语义分割是对单个像素点进行分类,相比于其他视觉分析任务具有更大的挑战性<sup>[12]</sup>。Long等<sup>[13]</sup>提出了基于全卷积网络(Fully Convolutional Network, FCN)的语义分割方法,在单个像素上对图像进行了标注。基于全卷积神经网络的框架,Badrinarayanan

等<sup>[14]</sup>提出编码-解码结构用于道路、车辆的分割,在物体边界分割精度上有所提高。在FCN基础之上,Ronneberger等<sup>[15]</sup>提出一种U形对称网络,该结构能融合低维与高维特征,使分割精度得到较大提高。为了有效利用图像上下文信息,考虑到像素之间的局部特性,即相邻像素属于同一类别的概率应该更高,Zhen等<sup>[16]</sup>提出将条件随机场引入卷积神经网络结构中,并实现端到端的分割模型。在水体提取任务中,Chen等<sup>[17]</sup>用简单线性聚类方法(Simple Linear Iterative Clustering, SLIC)将单个像素聚类为超像素(Superpixels),再利用卷积神经网络判别超像素是否为水体。许玥<sup>[18]</sup>在研究地表水体变迁中,采用改进的U-Net网络架构提取水体,取得了90.47%的分类精度,具体改进为使用elu激活函数替换原网络的relu激活函数,在模型最后加入条件随机场进行细分割。陈前等<sup>[19]</sup>选用Deeplabv3网络提取水体,总体精度为92.14%,并比较了NDWI方法、面向对象法、SVM法、CNN方法与作者选取方法,证明了深度学习提取水体的有效性。

尽管上述研究都能实现对高分辨率影像中水体的提取,但其提取精度,以及在提取小目标、小面积水体上仍然有提升的空间。本研究相对于经典U-Net网络的改进点主要在网络结构和分类后处理方面,具体为:① 本文方法利用预训练好的VGG16作为编码器,在学习深层特征的同时,预训练好的网络能够简化模型复杂度,减少模型训练时间;② 在解码器中对低维特征信息进行加强,使模型在信息传递过程中,能更有效利用低维特征信息,使模型在小目标、小面积区域提取方面有更好的效果;③ 考虑到水体内部纹理较为相似,相邻像元属于同一类的概率较大,模型引入条件随机场,将U-Net网络结构得到的特征图作为条件随机场输入,以提高目标物边缘分割精细度。

## 2 研究方法

### 2.1 技术路线

本文水体提取技术路线如图1所示,路线分为2部分:模型训练与模型测试。模型训练中首先对遥感影像进行辐射校正、几何校正和图像融合等数据预处理操作。将数据划分为训练集、验证集和测试集,3部分之间无相互重叠区域。训练集和验证集用来训练模型,测试集在模型测试中测试训练好的模

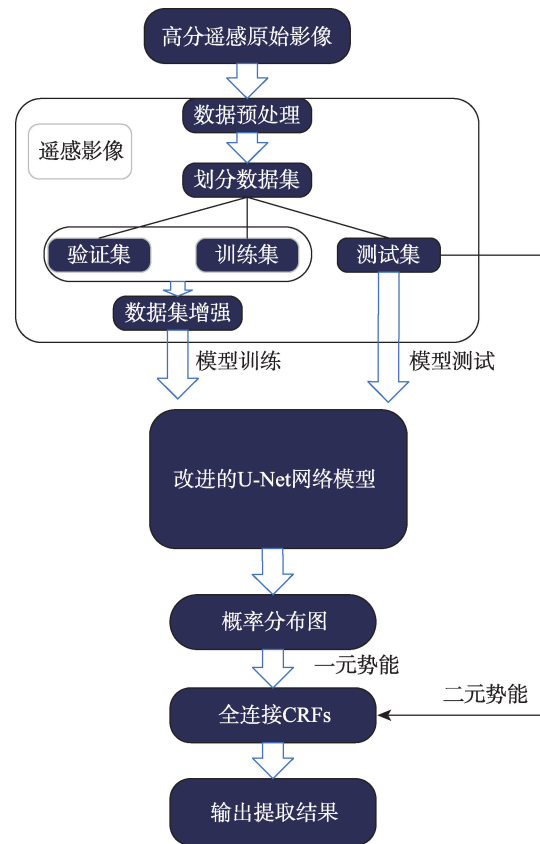


图1 水体提取技术路线

Fig. 1 Technical routes for extracting the water body

型的分割能力。训练网络模型采用改进的U-Net网络结构,该结构使用VGG16进行编码,使得网络能够提取图像高维特征信息,同时在网络解码阶段能有效增强图像低维特征信息,提高分割精度。在模型测试中,利用训练好的U-Net网络模型输出概率分布图,将其作为全连接条件随机场的一元势能输入,二元势能中的位置和颜色信息由原始影像提供,经过全连接条件随机场之后输出水体提取结果图。

### 2.2 经典U-Net网络

本文网络结构参考经典U-Net网络的对称结构。经典U-Net网络结构是基于全卷积神经网络(FCN)拓展和修改而来。网络结构清晰优雅形呈U状,具体由2部分组成:图2左半部分收缩路径用来获取上下文信息,图2右半部分对称的扩张路径用以对目标精确定位。

图2左半部分收缩路径是典型的卷积结构,包含连续2个3×3卷积核层,卷积后的特征图利用ReLU函数进行非线性激活,下采样方法采用最大池化,每次下采样操作都将特征图通道数增加2倍,同

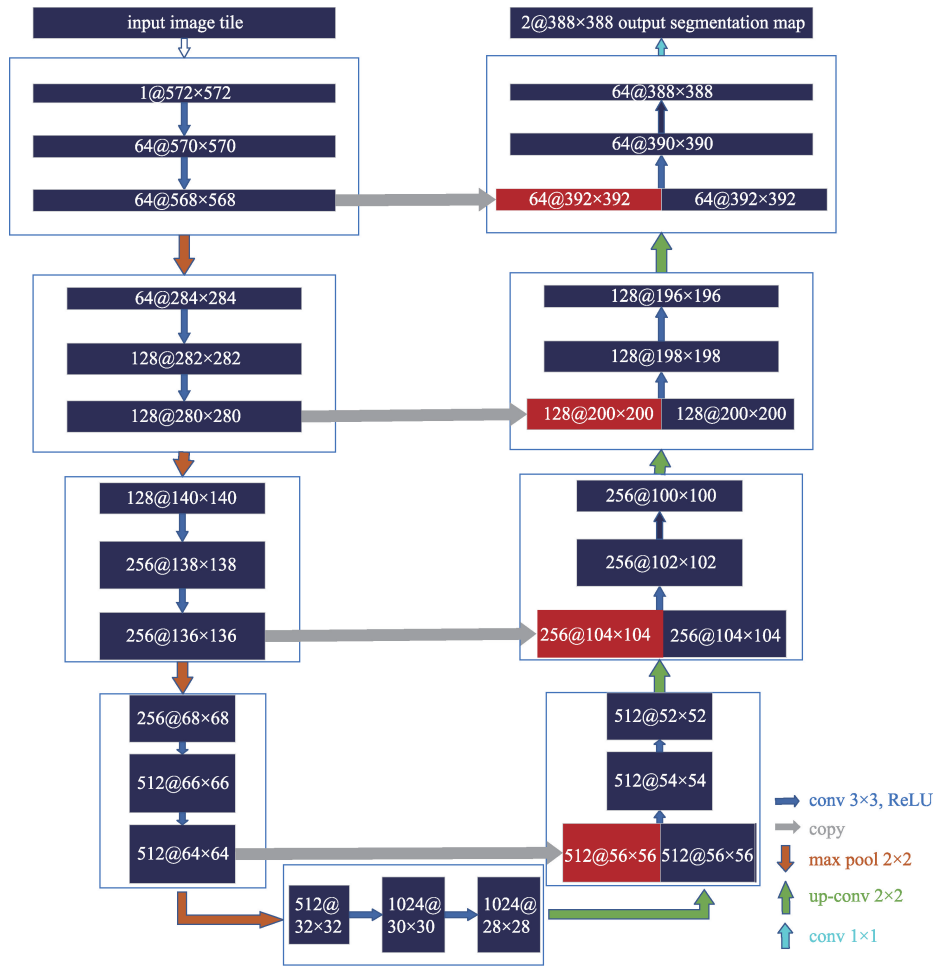


图2 经典U-Net网络结构  
Fig. 2 U-Net architecture

时将图像尺寸缩小2倍。扩张路径中,对特征图进行反卷积,使特征图尺寸扩大2倍,与对应收缩路径中的特征图进行拼接,以利用低维特征信息,对拼接后的特征图进行连续2次卷积操作,再利用ReLU函数激活。在最后的输出层中利用1×1卷积核对特征图进行卷积,生成与对应类别数相同的维数,同时还原图像的尺寸大小。在收缩路径中,每次池化操作之前,特征图都将被保存,并通过跳跃连接的方式将保存的特征图传递给对应尺度的扩张路径,从而为扩张路径提供不同尺度的特征信息。通过跳跃连接的方式,恢复由于最大池化操作降低分辨率带来的空间信息损失,为扩张路径提供不同分辨率的特征。

### 2.3 VGG16预编码与低维信息特征加强

经典U-Net网络编码阶段本质上是一个特征提取过程,经典U-Net网络被设计用于医学影像的分

割问题,医学影像背景较为简单,利用深度较浅的网络也能取得较好的效果,水体提取背景地物较为复杂,需要较深的网络学习更深层次的特征用于影像分类。VGG16网络相对于经典U-Net网络深度更深,网络已被用于多种遥感影像信息提取场景中,均取得较好的效果<sup>[20-22]</sup>。本文采用同构空间下基于特征的迁移学习,使用全卷积层替换VGG16网络最后的全连接层之后的网络作为U-Net网络编码阶段特征提取器,相较于经典U-Net特征提取结构,VGG网络结构更深,能提取水体高维特征信息,更有利于复杂背景地物下目标的分类。迁移学习在卷积神经网络中的一般过程如下:首先,利用大型数据集对网络中的参数随机初始化并进行训练,训练完毕保存训练权重,目的是增强网络的泛化能力;然后,利用训练好的卷积神经网络的卷积层和下采样层对目标应用领域的训练集进行特征提取。

VGG16网络有5个下采样层,将图像尺寸缩小为原尺寸的1/32。13个卷积层,卷积核大小为3×3,相比于大尺寸卷积核,小尺寸卷积核能将训练参数减少并将网络深度加升,卷积之后利用ReLU进行非线性激活。网络结构如表1所示。输入图像尺寸为256像素×256像素,经过13层卷积,5个下采样层之后生成大小为8像素×8像素、维度为512维的特征图。

表1 VGG16网络结构配置

Tab. 1 VGG16 network structure configuration

	感受野	步长	填充	输出大小
InputRGBImage:3@256×256				
Conv+ReLU	3×3	1	1	64@256×256
Conv+ReLU	3×3	1	1	64@256×256
Max-pooling				64@128×128
Conv+ReLU	3×3	1	1	128@128×128
Conv+ReLU	3×3	1	1	128@128×128
Max-pooling				128@64×64
Conv+ReLU	3×3	1	1	256@64×64
Conv+ReLU	3×3	1	1	256@64×64
Conv+ReLU	3×3	1	1	256@64×64
Max-pooling				256@32×32
Conv+ReLU	3×3	1	1	512@32×32
Conv+ReLU	3×3	1	1	512@32×32
Conv+ReLU	3×3	1	1	512@32×32
Max-pooling				512@16×16
Conv+ReLU	3×3	1	1	512@16×16
Conv+ReLU	3×3	1	1	512@16×16
Conv+ReLU	3×3	1	1	512@16×16
Max-pooling				512@8×8

本文所提取水体信息,背景地物多样,相邻环境复杂,随着网络深度的增加,特别是经过下采样层之后的特征图丢失了低维度的细节信息,导致提取结果的边缘轮廓效果较差,无法提取影像中较细小的水体信息。精确分割水体,需要更多地利用卷积神经网络低维特征信息。基于此,本文在对称网络结构基础之上对低维特征信息进行加强,以精确提取小目标水体信息。改进的网络结构如图3所示。

在网络右半部分扩张路径中,扩张路径特征图在与同一级低维特征连接时,加入收缩路径中VGG16网络中上一层级特征图进行融合,以进一步增强低维特征信息,在向后传递过程中能够防止细

节信息的丢失。连接后的特征图进一步利用3×3的卷积核进行卷积操作,为防止梯度消失和梯度爆炸,试验中在每次卷积操作之后都引入批量归一化(Batch Normalization)层,从而能够将前一层输出的特征归一化为均值为0、方差为1的分布,将得到的结果输入到激活函数ReLU中。经过激活函数之后对特征图进行上采样操作,每一次上采样特征图尺寸增大两倍,经过5次上采样之后,特征图恢复输入时的尺寸256像素×256像素。在网络末端采用1×1的卷积层将特征图数目缩减为2,并利用sigmoid函数对最终输出进行处理,使得网络输出中每个像素点对应值位于0到1范围内,每一个像素点对应值表示该点所属类型。这种网络结构能够在一定程度上增强水体低维特征信息,强化模型学习目标物边缘信息,提高目标分割边缘精度。

## 2.4 全连接条件随机场后处理

U-Net网络解码器中进行了上采样操作,这一步骤能够将特征图恢复至原图尺寸,但也造成了特征的损失,会产生地物边界模糊的问题。除此之外,卷积运算操作是局部连接的,它只能提供一个像素周围矩形区域内的信息,重复下采样卷积运算虽然能使矩形面积逐渐变大,但即使到最后一个卷积层,也无法获得整个图像中一个像素与其他所有像素之间的相关性。为了解决上述问题,提高分割的准确度,本文将卷积神经网络和全连接条件随机场<sup>[23-25]</sup>相结合,通过计算2个像素之间的相似性来判断他们是否属于同一类别。本文在模型测试中,将解码器最后一层的输出概率分布图作为全连接条件随机场的一元势能,二元势能中的位置和颜色信息由原始影像提供。影像后处理的结果作为最后的输出结果。全连接条件随机场后处理模型如图4所示。

全连接条件随机场的能量函数如下式:

$$E(X) = \sum_{i=1}^N \psi_u(x_i) + \sum_E \psi_p(x_i, x_j) \quad (1)$$

能量方程的第一项 $\psi_u(x_i)$ ,为一元势能函数,用于衡量当像素点*i*的颜色值为 $y_i$ 时,该像素点属于类别标签 $x_i$ 的概率。能量方程第二项成对势能函数 $\psi_p(x_i, x_j)$ ,用于衡量两事件同时发生的概率 $p(x_i, x_j)$ ,描述了每个像素与其他像素的关系,将颜色及相对距离较近的像素归为一类,计算公式如下:

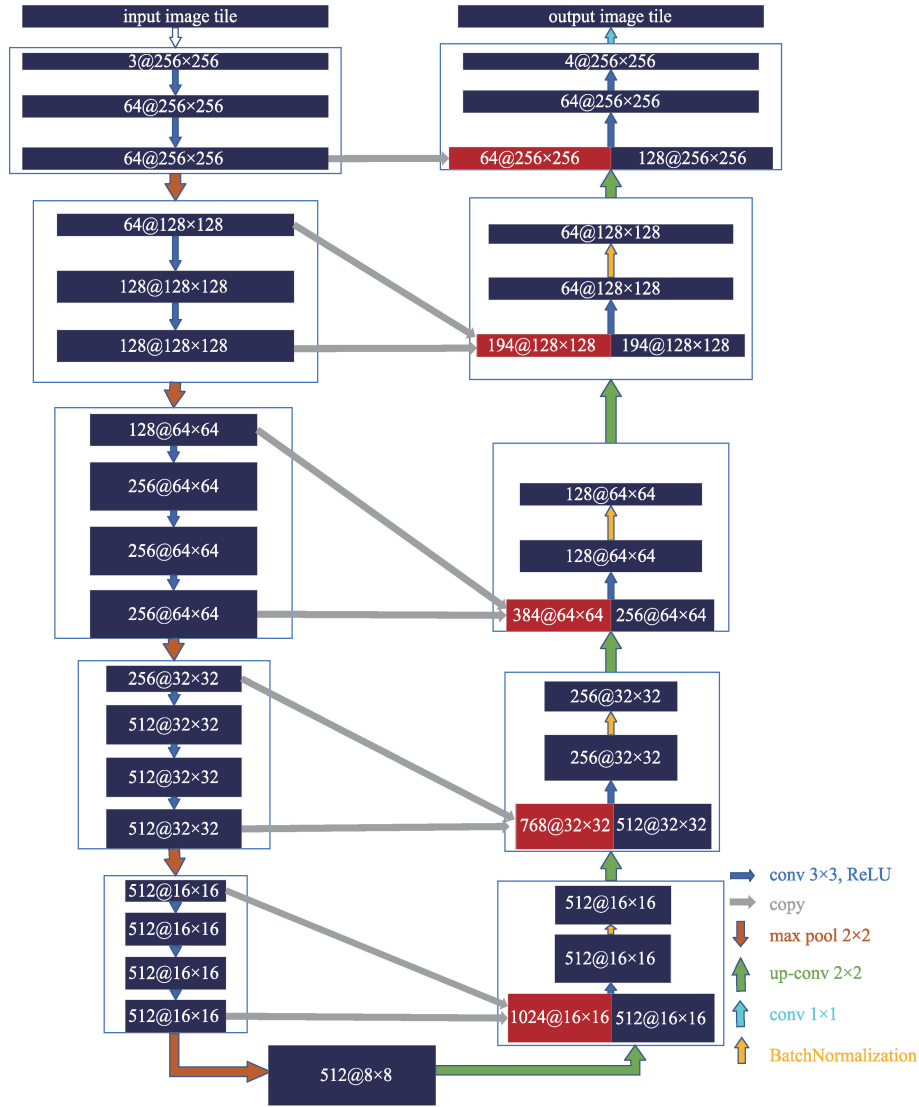


图3 改进的U-Net网络低维信息增强

Fig. 3 Improved U-Net network low-dimensional information enhancement

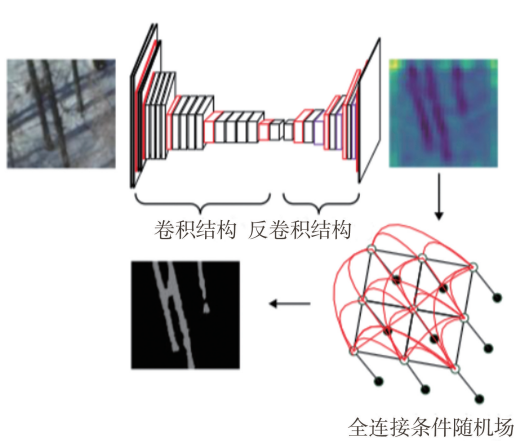


图4 全连接条件随机场后处理模型<sup>[26]</sup>

Fig. 4 Full connection condition random field post-processing model

$$\psi_P(x_i, y_j) = \frac{U(x_i, y_j) \sum_{m=1}^M \omega^m K_G^m}{K(f_i, f_j)}(f_i, f_j) \quad (2)$$

式中： $U$  为标签概率函数，计算像素  $i$  和像素  $j$  属于同一类的概率。若  $x_i \neq y_j$ ，则  $U(x_i, y_j) = 1$ ，否则为 0。 $\omega^m$  用来平衡函数， $K_G^m$  为高斯核函数， $K_G^m(f_i, f_j)$  表达式为：

$$K_G^m(f_i, f_j) = \exp\left(-\frac{1}{2}(f_i, f_j)^T \Lambda^{(m)}(f_i - f_j)\right) \quad (3)$$

式中： $f_i$  与  $f_j$  代表像素  $i$  和像素  $j$  的特征向量。

式(2)中的  $\omega^m$  是高斯  $K_G^m$  的权重，每一个高斯核  $K_G^m$  由一个对称正精度矩阵  $\Lambda^{(m)}$  表征，定义形状。对于多分类图像分割问题，式(2)中  $K(f_i, f_j)$  的

通常使用双核势能,表达式为:

$$K(f_i, f_j) = w^{(1)} \exp\left(-\frac{|p_i - p_j|^2}{2\theta_a^2} - \frac{|I_i - I_j|^2}{2\theta_\beta^2}\right) + w^{(2)} \exp\left(-\frac{|p_i - p_j|^2}{2\theta_\gamma^2}\right) \quad (4)$$

式中:  $I_i$ 、 $I_j$  代表像素位置  $p_i$ 、 $p_j$  上的颜色向量。公式右边第1项称为外观内核,第2项称为平滑内核。外观内核假设相邻且颜色相近的像素很可能属于同一类别。平滑内核的作用是消除孤立的小区域。

式(4)的作用是判别相似的像素点是否属于同一类。如果像素点属于同一类,则能量函数值相对较小。反之,若像素点不属于同一类,则能量函数相对较大。在水体影像分割中,河岸边区域往往被错误地划分为水体区域,影响后续分析。利用该能量函数,使河岸与相邻的水体分割更加精准。水体内部像素的RGB值相似,河岸与水体的RGB值存在差异。当相似区域的像素点判别为不同类时,能量函数值会变得较大。当存在差异的区域判别为同一类时,也会产生较大的能量值。通过多次迭

代,使能量函数的值最小化来获得最终结果。通过这种方式,利用整个图像的信息来细化水体边缘,提高分割的准确性。

### 3 实验数据及结果分析

#### 3.1 数据来源与预处理

本文研究区为山东青岛东部区域,地理位置如图5所示。遥感影像选用国产高分二号卫星(GF-2)影像,该卫星搭载两台1 m分辨率全色4 m分辨率多光谱相机,具有亚米级空间分辨率、高定位精度和快速姿态机动等特点。

实验选取3幅国产GF-2卫星融合影像1 m分辨率影像数据,详细信息如表2所示。影像中涵盖农村和城市区域,水体类型丰富,既有大型河流水库,又有细小水流池塘。同时,研究区内地貌类型多样,包括广袤平原与高大山体,这些地貌构成不同背景地物,为研究不同背景地物下提取水体提供数据支撑。

在遥感影像成像过程中,由于受传感器外在原因:如姿态的变化、高度、速度等因素造成遥感影像几何畸变与变形,并且遥感影像本身在空间、时间

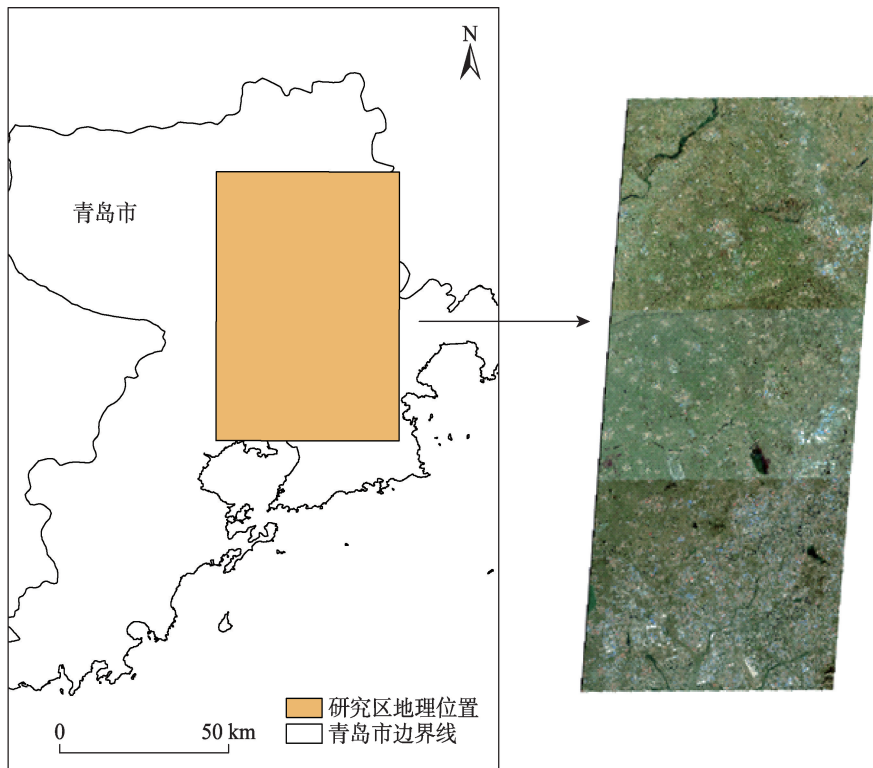


图5 青岛研究区地理位置

Fig. 5 Location of Qingdao study area

表 2 青岛研究区遥感影像信息

Tab. 2 RemotesensingimageinformationintheQingdaostudyarea

影像编号	中心经度/°E	中心纬度/°N	成像时间	影像大小/像素×像素
L1A0003593712	120.5	36.7	2018-11-12	27 620×35 273
L1A0003593719	120.4	36.3	2018-11-12	27 620×35 113
L1A0003593868	120.6	36.3	2018-11-12	27 620×35 191

以及光谱分辨率的不足,在获取数据的过程中不能够精确的进行信息的记载,在很大程度上会降低遥感数据的精度,因此,需要对遥感影像进行遥感数据的预处理<sup>[27]</sup>。辐射校正原理是运用辐射传输模型,将入瞳辐射亮度值转为地表真实反射率值。正射校正是校正了因传感器、地形的起伏不均衡等因素引起的像点元素的偏移,并利用地面控制点通过相应的数学算法模型来进行实现的过程。高分影像自带RPC文件,影像的正射校正采用自带的RPC文件和数字高程模型(DEM)来进行数据定位。图像融合能够保留多光谱的真彩色信息,同时也能保留全色波段的高空间分辨率信息。在后续应用中,影像反射率均为校正后所得。影像处理流程如图6所示。

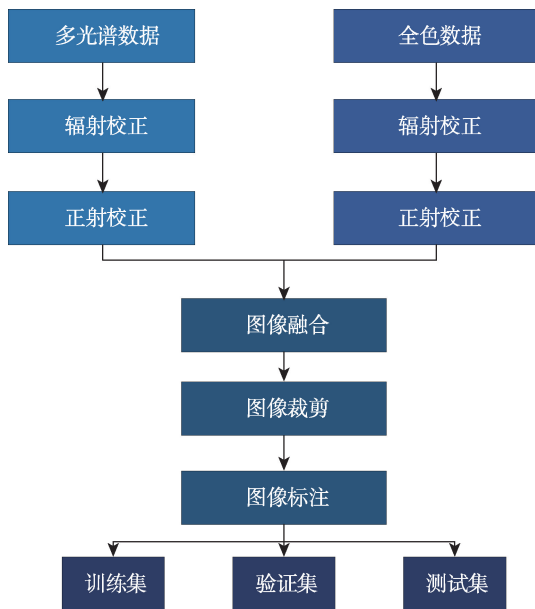


图 6 GF-2 影像处理流程

Fig. 6 GF-2 image processing flowchart

将融合后的影像裁剪为 60 幅大小为 2884 像素×2753 像素的真彩色图像,并对裁剪后的每幅图像在 labelme 软件工具中进行手工标注。打乱标记好的数据顺序,随机分配 40 幅为训练图像,15 幅为验证图像,5 幅作为测试图像。本文将地物分为水

体与背景地物两类。在反复试验下,将输入层的大小设定为 256 像素×256 像素,在图像与对应标签中随机裁剪大小为 256 像素×256 像素的图像,在裁剪过程中对图像进行随机数据增强。所做的数据增强包括:将原图像和对应的标签同时做随机旋转;将原图像与对应的标签同时做镜像处理;对图像随机增加噪声以增强模型的鲁棒性,降低训练数据集的敏感性。最后生成 30 000 个训练样本、9000 个验证样本和 2000 个测试样本。

### 3.2 训练模型

全卷积神经网络计算在训练时会进行大量计算并消耗大量内存和显存,对硬件要求较高,但受限于价格和实验环境,会在平台方面追求一个平衡。本文模型基于深度学习框架 Keras 搭建,按照目前主流配置环境搭建深度学习实验环境,基础配置如表 3 所示。

表 3 基础系统平台配置

Tab. 3 Basic system platform configuration

项目	系统	CPU	内存	硬盘	显卡
内容	Ubuntu 16.04	Intel E5-1630	8 GB	500 GB	NVIDIA GTX970

本文重要软件配置如表 4 所示。之后针对各个模型的运行情况,参考文献[28]可适当安装相应软件包以加速模型运行。

表 4 重要软件配置

Tab. 4 Important software configuration

项目	GPU-Driver	CUDA	Python	Keras	Tensorflow-gpu
内容	384	8.0	3.6	2.2.4	1.4.0

通过多次试验,综合考虑模型计算效率、结果精度以及硬件 3 个方面,实验最终将迭代次数设置为 60,批量大小设置为 14,选用 Adadelta 作为优化器,并将初始学习率设置为 1.0。经过迭代训练,网络最终收敛,在训练集上取得了 99.7% 的分类精度,在验证集中取得了 99.6% 的分类精度,训练曲线如图 7 所示。



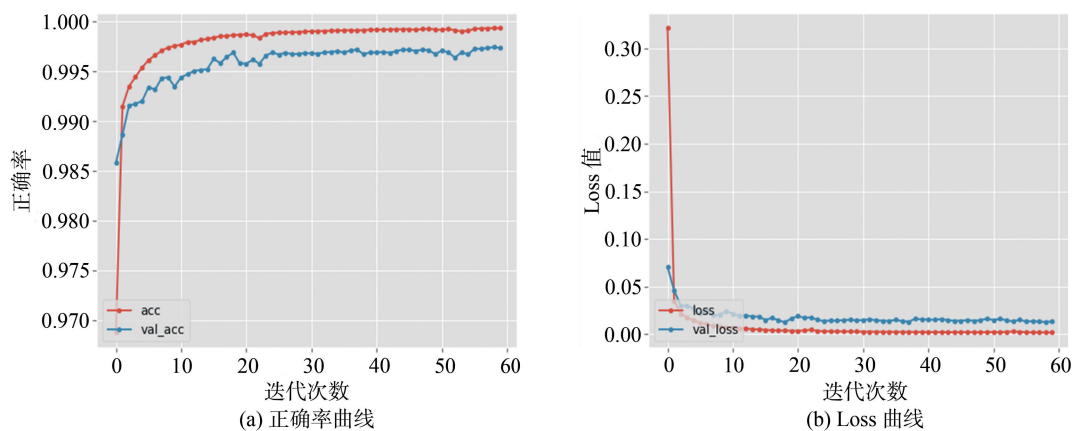


图7 改进的U-Net模型训练曲线  
Fig. 7 Training curve of improved U-Net

### 3.3 实验结果及分析

#### 3.3.1 研究区结果及对比分析

在对分割结果做定性定量评价之前,需要与其他著名网络进行对比,在保持相同训练集、验证集及测试集的情况下,本文选择SegNet与经典U-Net网络进行对比实验。选择裁剪的2000幅256像素×256像素的影像块对模型进行测试,为了直观地展示对比结果,选取典型的5块区域进行展示,如表5所示。在青岛研究区,研究结果显示3个网络都能有效识别大型且内部均一的水体(区域1、区域2),但通过实验结果对比可以发现SegNet在识别水体时,存在误提现象,将河岸边的绿色体育场误识别为水体(区域1)。在提取的水体内部存在大的空洞,导致提取的水体形态不完整,在与桥邻近区域,存在漏提现象,导致水体与桥之间存在较大的间隙(区域2)。大型山体形成阴影,SegNet在识别过程中,将阴影错误地识别为水体信息(区域3)。在小型水体识别方面,不能完整地识别出水体区域(区域5)。经典U-Net网络在分割完整度上总体要优于SegNet网络(区域1、区域2、区域5),但在内部区域会出现空洞现象(区域2)。对于较小的水体不能完整识别,存在漏提现象,边界分割不精细(区域4、区域5),也会将高山阴影错误识别为水体(区域3)。采用改进的U-Net网络能够较好地提取大面积水域,同时去除阴影和建筑物的影响,但提取的结果会存在过多地提取周围地物的现象(区域2),在对小目标、小面积水体和细长水体提取中,改进的U-Net相比较于以上2种网络有更好的发挥(区域4、区域5)。改进后的U-Net网络通过加强遥感图像

低维特征信息,能够有效结合图像低维特征信息和高维特征信息,挖掘图像深层规律的同时,保留图像低维特征信息,找到有效地区分水体和其他地物的特征。

在定量评价中,选用交并比(IoU)、精准率以及Kappa系数3个指标对2000幅影像的水体提取结果进行评定。交并比(IoU),通常被称为PASCAL VOC交叉联合度量,是图像分割定量评价系统中最常用的一种性能度量方法。在图像像素级分割任务中,可以理解为在测试集上分割像素集合P与图像标注结果的像素集合GT之间的交集比上它们的并集,公式如下:

$$IoU = \frac{TP}{TP + FP + FN} \quad (5)$$

式中:TP表示分类准确的正类像素数目;FP表示分类为正类的负类像素数目;FN表示被错分为负类的正类像素数目。分类如表6所示。

精准率就是被准确分类为正类的像元数与所有被分类为正类的像元数之比,表示预测结果为正类的像素里有多少像素是真的正类。精准率指标如式(6)所示。

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

Kappa系数的计算如式(7)所示。

$$Kappa = \frac{p_o - p_e}{1 - p_e} \quad (7)$$

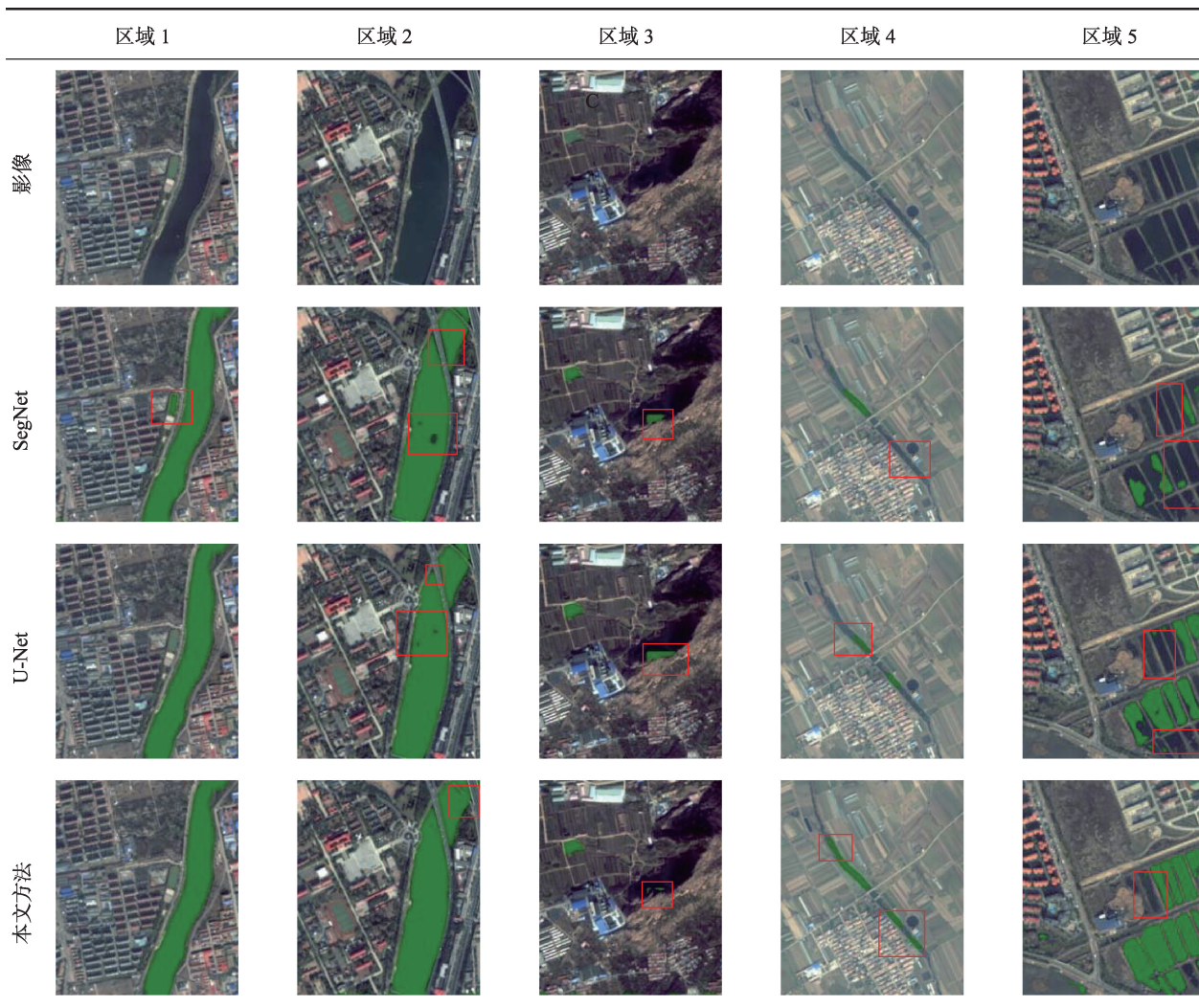
式中: $p_o$ 代表每一类中正确分类像元素; $p_e$ 则定义为:

$$p_e = \frac{a_1 \times b_1 + a_2 \times b_2 + \dots}{N \times N} \quad (8)$$

式中: $a$ 代表每一类的实际像元数; $b$ 代表每一类的

表5 研究区5个典型区域不同方法水体提取结果比较

Tab. 5 Comparison of water extraction results by different methods in 5 typical areas of the study area



注：  表示提取水体。

表6 精度评价混淆矩阵

Tab. 6 Confusion matrix for accuracy evaluation

	实际正类	实际负类
预测正类	TP	FP
预测负类	FN	TN

预测像元数。

定量评价结果如表7所示。从表中可看出,经过改进的U-Net网络分割结果指标最好,与SegNet相比,IoU、精准率和Kappa系数分别提升了10.5%、12.3%和0.14;与经典U-Net网络结果相比,3项指标分别提升了5.8%、4.4%和0.05。

### 3.3.2 模型推广应用实验

为研究模型的时空可拓展性,本文选取青岛西部区域和青海西宁地区两幅高分二号影像作为模

表7 水体提取结果精度比较

Tab. 7 Accuracy comparison of water extraction results

方法	IoU/%	精准率/%	Kappa系数
SegNet	77.6	82.5	0.79
经典U-net	82.3	90.4	0.88
改进后的U-Net网络	88.1	94.8	0.93

型应用研究数据,数据集如表8所示。对原数据做数据预处理,包括辐射校正、正射校正、图像融合和图像裁剪。从2幅影像中各裁剪2000幅256像素×256像素大小的测试样本,用训练好的改进U-Net网络模型对测试样本进行测试。

从应用区提取结果中各选取典型的5块区域进行展示,测试结果如表9所示,青岛应用区与本文研究区同属一个地理单元,地貌类型和背景地





















表8 应用区遥感影像信息

Tab. 8 Remote sensing image information in the application area

影像编号	中心经度/°E	中心纬度/°N	成像时间	影像大小/像素×像素
L1A0003553729	120.1	36.3	2018-10-28	276 20×292 00
L1A0003351642	101.5	36.8	2018-07-26	276 20×292 00

表9 应用区5个典型区域水体提取结果比较

Tab. 9 Comparison of water extraction results in 5 typical areas of the application area

		区域1	区域2	区域3	区域4	区域5
青岛	原始影像					
	水体信息					
西宁	原始影像					
	水体信息					

注:  表示提取水体。

物相似,水体类型也较为相似,从测试结果看效果较好,特别是对细长水体和小面积水体的提取,但也存在明显问题,区域2中与水体邻接的建筑阴影被误识别为水体,模型最终在测试集上测试结果IoU、精准率和Kappa系数分别为86.7%、93.2%和0.9。青海省西宁位于我国西北部,与本文研究区在地理环境上差异较大,直观地从测试结果可以看出,对于区域1类型的水体模型无法识别,可能原因是模型在训练过程中没有加入这类型的水体进行训练。模型对于大型水体识别效果较好,但也存在将淡绿色植被识别为水体的现象(区域2)。模型对形态与青岛研究区相似小面积水体的识别效果较好(区域4、区域5)。模型在西宁应用区测试结果IoU、精准率和Kappa系数分别为79.5%、84.3%、0.81。

## 4 结论与展望

本文基于U-Net网络提出一种改进的U-Net网络模型,用于提取高分二号遥感影像水体信息。研究以山东省青岛市东部高分二号遥感影像为数据源,基于改进的U-Net网络语义分割算法对遥感影像中的水体信息进行提取,并与SegNet和经典U-Net网络语义分割算法结果做对比。同时,为了验证模型的时空可扩展性,本文选择青岛西部和青海西宁高分二号遥感影像对模型进行验证。结果表明:

(1)本文改进U-Net方法、SegNet分割算法、经典U-Net网络语义分割算法3种方法比较中,本文改进的网络模型具有优异的分割性能,在精度评价指标IoU、精准率以及Kappa系数上分别达到了

88.1%、94.8%和0.93的精度。

(2)本文提出的方法在研究区小目标提取、水体提取完整度、阴影与水体的区分以及边界分割精度中更具有优势。

(3)在模型推广应用,模型对于与研究区地理环境相似区域水体提取效果较好,但与研究区环境相差较大的区域提取效果还有待进一步提高。

水体背景信息的复杂多样性为水体提取工作带来困难,从模型的推广应用中可以发现,在训练数据中应该更多地增加不同背景地物下的样本数量,以增强模型的泛化能力。因此,下一步研究工作将着重加强不同背景地物下水体的提取工作,进一步提高模型的时空可拓展性。

#### 参考文献(References):

- [ 1 ] Li K, Wu S, Dai E, et al. Flood loss analysis and quantitative risk assessment in China[J]. *Natural Hazards*, 2012,63(2):737-760.
- [ 2 ] Bond N R, Lake P, Arthington A H. The impacts of drought on freshwater ecosystems:an Australian perspective[J]. *Hydrobiologia*, 2008,600(1):3-16.
- [ 3 ] Lake P S. Ecological effects of perturbation by drought in flowing waters[J]. *Freshwater Biology*, 2003,48(7):1161-1172.
- [ 4 ] 吴赛,张秋文.基于MODIS遥感数据的水体提取方法及模型研究[J].*计算机与数字工程*,2005(7):1-4.[ Wu S, Zhang Q W. Method and model of water body extraction based on remote sensing data of MODIS[J].*Computer & Digital Engineering*, 2005(7):1-4. ]
- [ 5 ] 李丹,吴保生,陈博伟,等.基于卫星遥感的水体信息提取研究进展与展望[J].*清华大学学报(自然科学版)*,2020,60(2):147-161. [ Li D, Wu B S, Chen B W, et al. Review of waterbody information extraction based on satellite remote sensing[J]. *Journal of Tsinghua University (Science and Technology)*, 2020,60(2):147-161. ]
- [ 6 ] 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.
- [ 7 ] Xu H. 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.
- [ 8 ] 都金康,黄永胜,冯学智,等. 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. [ Du J K, Huang Y S, Feng X Z, et al. Study on water bodies extraction and classification from SPOT image[J]. *Journal of Remote Sensing*, 2001,5(3):214-219. ]
- [ 9 ] 王培培.基于 ETM 影像的水体信息自动提取与分类研究[J].*首都师范大学学报(自然科学版)*,2009,30(6):75-79. [ Wang P P. Automatic extraction and classification of water body from ETM image[J]. *Journal of Capital Normal University (Natural Science Edition)*, 2009,30(6):75-79. ]
- [ 10 ] 胡德勇,李京,陈云浩,等.单波段单极化 SAR 图像水体和居民地信息提取方法研究[J].*中国图象图形学报*,2008,13(2):257-263. [ Hu D Y, Li J, Chen Y H, et al. Water and settlement area extraction from single band single polarization SAR images based on SVM method[J]. *Journal of Images and Graphics*, 2008,13(2):257-263. ]
- [ 11 ] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006,313(5786):504-507.
- [ 12 ] 计梦予,裘肖明,于治楼.基于深度学习的语义分割方法综述[J].*信息技术与信息化*,2017(10):137-140. [ Ji M Y, Xi X M, Yu Z L. A review of semantic segmentation based on deep learning[J]. *Information Technology and Informatization*, 2017(10):137-140. ]
- [ 13 ] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015:3431-3440.
- [ 14 ] Badrinarayanan V, Handa A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling[J]. *ArXiv Preprint ArXiv: 1505.07293*, 2015:117-124.
- [ 15 ] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]. *International Conference on Medical Image Computing and Computer-assisted Intervention*, 2015:234-241.
- [ 16 ] Zheng S, Jayasumana S, Romera-Paredes B, et al. Conditional random fields as recurrent neural networks[C]. *Proceedings of the IEEE International Conference on Computer Vision*, 2015:1529-1537.
- [ 17 ] Chen Y, Fan R, Yang X, et al. Extraction of urban water bodies from high-resolution remote-sensing imagery using deep learning[J]. *Water*, 2018,10(5):585.
- [ 18 ] 许玥.基于改进Unet的遥感影像语义分割在地表水体变迁中的应用[D].重庆:重庆师范大学,2019. [ Xu Y. Application of semantic segmentation of remote sensing image based on improved Unet in surface water change[D]. Chongqing: Chongqing Normal University, 2019. ]
- [ 19 ] 陈前,郑利娟,李小娟,等.基于深度学习的高分遥感影像水体提取模型研究[J].*地理与地理信息科学*,2019,35(4):

- 43-49. [ Chen Q, Zheng L J, Li X J, et al. Water body extraction from highresolution satellite remote sensing images based on deep learning[J]. *Geography and Geo-information Science*, 2019,35(4):43-49. ]
- [20] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [21] Li Y, Guo L, Rao J, et al. Road segmentation based on hybrid convolutional network for high-resolution visible remote sensing image[J]. *IEEE Geoscience and Remote Sensing Letters*, 2018,16(4):613-617.
- [22] Hu J, Li L, Lin Y, et al. A Comparison and strategy of semantic segmentation on remote sensing images[J]. arXiv preprint arXiv:1905.10231, 2019.
- [23] Krähenbühl P, Koltun V. Efficient inference in fully connected crfs with gaussian edge potentials[C]. *Advances in neural information processing systems*, 2011:109-117.
- [24] Chen L-C, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets and fully connected crfs[J]. arXiv preprint arXiv:1412.7062, 2014: 97-105.
- [25] 刘懿兰,黄晓霞,李红春,等.基于卷积神经网络与条件随机场方法提取乡镇非正规固体废弃物[J].*地球信息科学学报*,2019,21(2):259-268. [ Liu Y L, Huang X X, Li H G, et al. Extraction of irregular solid waste in rural based on Convolutional Neural Network and conditional random field method[J]. *Journal of Geo- information Science*, 2019,21(2):259-268. ]
- [26] 黄英来,刘亚檀,任洪娥.基于全卷积神经网络的林木图像分割 [J]. *计算机工程与应用*,2019,55(4):219-224. [ Huang Y L, Liu Y T, Ren H E. Segmentation of forest image based on Fully Convolutional Neural Network[J]. *Computer Engineering and Applications*, 2019,55(4):219-224. ]
- [27] 王晨巍,王晓君.高分遥感卫星影像的预处理技术[J].*电子技术与软件工程*,2016(24):122-123. [ Wang C W, Wang X J. Pre-process technology of high precision satellite photograph[J]. *Electronic Technology & Software Engineering*. 2016(24):122-123. ]
- [28] 张新明,祝晓斌,蔡强,等.图像语义分割深度学习模型综述[J].*高技术通讯*,2017,27(Z1):808-815. [ Zhang X M, Zhu X B, Cai Q, et al. Survey of the deep learning models for image semantic segmentation[J]. *Chinese High Technology Letters*, 2017,27(Z1):808-815. ]