doi: 10.6046/zrzyyg.2021357

引用格式:沈骏翱,马梦婷,宋致远,等. 基于深度学习语义分割模型的高分辨率遥感图像水体提取[J]. 自然资源遥感,2022, 34(4):129-135. (Shen J A, Ma M T, Song Z Y, et al. Water information extraction from high - resolution remote sensing images using the deep - learning based semantic segmentation model[J]. Remote Sensing for Natural Resources, 2022, 34(4):129-135.)

# 基于深度学习语义分割模型的 高分辨率遥感图像水体提取

沈骏翱<sup>1</sup>,马梦婷<sup>2</sup>,宋致远<sup>1</sup>,柳汀洲<sup>1</sup>,张 微<sup>1,2</sup> (1.浙江大学软件学院,宁波 315048;2.浙江大学计算机科学与技术学院,杭州 310027)

**摘要:**水体提取是高空间分辨率遥感影像应用中重要研究方向之一。传统识别方法仅利用水体的浅层特征,为了 更好地挖掘遥感影像的深度信息,从而提升水体提取算法的鲁棒性,提高分割精度,提出了一种基于深度学习语义 分割模型的水体提取方法。利用深度神经网络挖掘高分辨率遥感影像信息,同时引入注意力模块,整合深层信息 与浅层地物的形状、结构、纹理和色调等信息,拟建立比现有模型具有更高准确率、更快预测速度的全新深度语义 分割模型。最后,和传统识别方法以及常见语义分割模型进行对比消融实验。实验证明所提出算法模型的总体精 度和效率均优于现有方法,且算法参数设置简单,受人工干预少。文章证明了深度学习以及注意力机制在高分辨 率遥感影像水体提取任务上的准确性和高效性,提供了一种使用深度学习方法解决高分辨率遥感影像分割任务的 可能,并对未来进行了展望。

关键词:语义分割;多尺度;遥感影像;全卷积网络;注意力机制 中图法分类号:TP 75 文献标志码:A 文章编号:2097-034X(2022)04-0129-07

### 0 引言

自1999年美国洛克希德·马丁公司发射空间 分辨率为1m的IKONOS卫星遥感影像之后,国际 上正式开始进入了高分辨率遥感影像处理、分析与 应用的时代。如何依据获取的高分辨率遥感影像中 包含的光谱信息以及空间信息提取出各种需要的地 物信息,即如何高效、准确地利用原始遥感影像获取 需要的地理信息,是高分辨率遥感影像应用的核心 环节之一<sup>[1-2]</sup>。同时,水资源作为地球上最重要的 资源之一,是一切人类与生物得以生存和发展的重 要物质基础,从高分辨率遥感影像中进行水体提取 便于了解现有的水资源概况,有助于对水资源更加 合理的规划和治理,提高水资源的利用效率,对人类 生活及社会活动具有重大影响。

近 20 a 来,遥感影像处理专家也发展出了许多 基于像元的传统提取方法,Work 等<sup>[3]</sup>通过研究 MSS 影像发现水体在近红外波段上的反射率较低,利用 这一特性人为的设定一个较小的阈值来进行水陆分 离,被称为单波段方法。多波段方法是 Frazier 等<sup>[4]</sup> 在研究澳大利亚湖泊时使用 TM 影像对比分析了 TM4, TM5, TM7 这3个波段的提取效果, 并利用多 波段融合的方法通过设置不同阈值提高了分割精确 度,但依旧受到噪声的干扰。近期,被广泛认可的水 体指数方法,是由 McFeeters<sup>[5]</sup>受到归一化差异植被 指数(normalized difference vegetation index, NDVI) 的启发,总结并提出了基于绿光波段与近红外波段 的归一化差异水体指数(normalized difference water index,NDWI)<sup>[6]</sup>,可有效减少噪声和阴影。其后由 Xu<sup>[7]</sup>在考量建筑物与土壤的波段特性后进行了优 化,将近红外波段改为中红外波段,并命名为 MND-WI. 消除了大部分噪声的影响, 但人为观测设定的 阈值依旧存在缺陷,因此在地物信息复杂的实际应 用中提取精确度并不高。

近些年,随着人工智能算法的发展,深度学习为 计算机视觉和图像识别等领域带来了猛烈冲击,其 目标分类与识别和追踪技术甚至能与人类识别性能

通信作者:张 微(1980-),男,博士,教授,博士生导师,研究方向为时空大数据。Email: cstzhangwei@zju.edu.cn。

收稿日期: 2021-10-25;修订日期: 2022-06-17

基金项目:浙江省重点研发计划项目"基于大数据的时空信息平台系统建设"(编号:2021C01031)和宁波市自然科学基金项目"基于时空大数据和 AloT 技术的污泥专运溯源管理系统研发与应用"(编号:2022S125)共同资助。

第一作者:沈骏翱(1997-),男,硕士研究生,研究方向为遥感影像深度学习分析。Email: 22051094@ zju. edu. cn。

相媲美<sup>[8-10]</sup>。在 Long 等<sup>[11]</sup>提出了全卷积神经网络 (fully convolutional networks, FCN)结构后,卷积神经 网络分类任务正式进入到像素级别,拓展到语义分 割领域。当然就高分辨率遥感影像分割而言,由于 其地表目标具有尺度效应,综合高分辨率遥感影像 具有丰富的纹理结构信息,需要多尺度图像的分割 方法,但从低层次特征到高层次语义鸿沟,实现高分 辨率遥感影像语义分割依旧是具有挑战的问题。近 期,基于机器学习的遥感影像水体提取任务也开始 受到了关注,许明<sup>[12]</sup>在传统 U – Net 深度学习网络 上加入条件随机场对水体变迁进行预测。但是目前 的方法仅仅使用深度学习进行深层数据的挖掘而无 法解释深层特征对于最终结果判断的影响,引入注 意力机制可以在结合深层和浅层特征的同时提供更 好的可解释性。

注意力机制作为最简便有效的即插即用模块, 在传统计算机视觉任务中广泛使用,近段时间,遥感 影像领域的各种任务中也开始广泛融入注意力机 制,Ying等<sup>[13]</sup>在2019年提出了在对地目标检测任 务中引入多头注意力机制与特征金字塔结合,有效 提升了深度学习网络关注目标浅层特征中提取出重 要的空间位置信息;He等<sup>[14]</sup>在2020年也为城市建 筑物语义分割任务设计了一种多阶注意力机制,充 分利用中间特征之间的相关性,从而获得更强大、更 具有代表性的特征。

因此,本文将研究基于深度学习的高分辨率遥 感影像中水体的语义分割。利用深度学习强大的特 征表达能力,尝试构造从底层特征到对象语义的映 射,同时考虑多尺度特征的监测问题,并适时地引人 注意力机制,提升对于中间特征的通道方向以及空 间区域上重要区域的关注度,在重要的通道与空间 区域块上将会提高响应值,通过激活模块后重要部 分对于最终水体分类的影响将提高,弥补纯卷积神 经网络无法关注到细节区域分类的缺陷,提高语义 分割时关注原始信息中重要部分的可能性。

1 数据集预处理与标注

本实验数据集使用高分六号(GF-6)遥感卫星 获取长江三角洲某地区包含水体的4个波段(红光 波段、绿光波段、蓝光波段、红外波段)影像,由于目 前深度学习卷积网络模型一般只使用三通道图像进 行语义分割,因此首先使用 eCognition 软件对原始 遥感影像进行合并,然后根据遥感影像特点,以及后 续使用传统指数方法提取水体信息(NDWI)时需要 使用红外波段(IR),生成 22 953 像素×17 605 像素 大小的 IRRG(红外波段、红光波段、绿光波段) PNG 三通道图像进行存储。由于遥感影像传感器原因, 使用 Python 对影像进行预处理,减少噪声对语义分 割模型识别造成的影响,最终得到完成几何纠正的 影像。



图1 最终遥感影像标注图局部



2 研究方法

### 2.1 注意力模块

2.1.1 通道级联空间注意力模型概述

注意力模块目前已经成为了深度学习处理中重要的概念,最初被用于机器翻译,最终在深度学习领域中的自然语言处理、计算机视觉以及图神经网络被广泛使用<sup>[15]</sup>。

本文尝试将注意力机制引入到遥感影像语义处 理中的解码部分,应用于卷积神经网络的注意力模 块主要包含通道注意力(channel attention)以及空间 注意力(spatial attention)2部分,将这2类注意力处 理方法级联形成新的注意力模块,并将其命名为 S&CMNet。遥感影像通过主干网络特征提取映射后 输入到通道注意力模块,在通道上提升网络对特定 通道关注度,对通道特征进行细化。将细化后的通 道特征图输入到空间注意模块中,同时空间域上学 习关注点,提取重要局部信息。最后通过大量卷积 层及池化层后最终连接 SoftMax 分类器得到语义分 割结果。

### 2.1.2 特征提取结构

研究过程中对于编码器部分的特征提取骨干网

络主要是使用了目前在计算机视觉领域以及语义分 割领域取得较好效果的 ResNet101 和 U – Net。本次 研究对于 ResNet101 进行了一定改进,使用了空洞 卷积操作,最终得到原图像 1/8 大小的特征映射图。 在特征提取最后一层中并不直接进行语义推理,而 是将最后一层特征图输入注意模块进行特征细化, 从而更好地学习特征表达。图 2 为 S&CMNet 概览 图,图中 C,H,W 分别代表特征图通道数、高度与宽 度。



图 2 S&CMNet 概览图 Fig. 2 S&CMNet overview map

### 2.2 混合注意力模块

### 2.2.1 通道注意力模块

高分遥感影像输入特征提取骨干网络通过 C核卷积操作之后得到多通道特征图  $F \in R^{C \times H \times W}$  作为 通道注意力模块的输入数据。通道的特征表示该图 像在不同卷积核上的分量,通道信息表示关键信息 的贡献多少。

通道注意力模块的目的是利用每个信道之间的 关系特征映射学习一个权值  $W_{c} \in R^{C\times1\times1}$ 。整体过 程可以分为压缩(Squeeze)、激活(Excitation)以及注 意(Attention)3 部分。

1)Squeeze 模块目的为降低每个通道的信息维度,聚合空间维度信息。池化操作,见式(1)和(2), 为每个通道生成2个特征描述符,输入含有隐层的 1×1卷积核组成的多层感知器,生成更具有代表性 的特征向量,即

$$Z_{c1} = F_{sq}(u_c) = \frac{1}{H} \frac{W}{W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i,j) , \quad (1)$$

$$Z_{c2} = F_{sq}(u_c) = \max\left[\sum_{i=1}^{H}\sum_{j=1}^{W}u_c(i,j)\right], \quad (2)$$

式中: Z<sub>e1</sub>为平均上采样之后生成的更具代表性的 通道特征向量; Z<sub>e2</sub>为最大上采样之后生成的更具 代表性的通道特征向量; F<sub>sq</sub>为上采样函数; u<sub>e</sub>为输 入通道注意力模块的特征图。 2) Excitation 模块主要是通过利用 Sigmoid 函数,对于 Squeeze 模块得到的特征进行激活,得到最终的通道注意图。公式为:

 $W_{c} = \text{Sigmoid}[\text{MLP}(Z_{c1}) + \text{MLP}(Z_{c2})], \quad (3)$ 

式中:W。为生成的特征权重;Sigmoid 为激活函数; MLP 为多层感知机组成的线性分类层。

3) Attention 模块将最后生成的特征权重与输入 的特征图向量相乘,本质上是缩放的过程,从而增强 对关键通道的注意力。计算公式为:

$$\widetilde{u}_{c} = F_{scale}(u_{c}, W_{c}) = u_{c} \cdot W_{c}$$
, (4)

式中: $F_{\text{scale}}$ 为缩放及特征权重融合函数; $\hat{u}_{c}$ 为最终输出特征图。

2.2.2 空间注意力模块

空间域注意力模块关注的是对当前任务更具有 价值的空间区域,级联在通道注意力模块之后,是对 通道注意力的补充。以遥感影像为例,使用空间注 意力模块有助于聚合空间信息,特别是对小型地物 信息进行处理。

空间注意力模块以通道注意力模块输出的特征

图  $\tilde{u}_{e}$  为输入,利用不同空间位置之间的关系来学习 一个二维空间权重  $W_{s}$ ,将学习得到的权重值乘以 相应的空间位置学习更具代表性的特征。空间权重 学习过程类似于通道学习过程,由3部分组成。其中,Squeeze模块公式为:

$$Z_{\rm sl} = F_{\rm sq}(\tilde{u}_{\rm c}) = \frac{1}{C} \sum_{i=1}^{C} \tilde{u}_{\rm c}(i) ,$$
 (5)

$$Z_{s2} = F_{sq}(\tilde{u}_{c}) = \max\left[\sum_{i=1}^{C} \tilde{u}_{c}(i)\right]$$
, (6)

式中: Z<sub>st</sub>为平均上采样之后生成的更具代表性的 空间特征向量; Z<sub>s2</sub>为最大上采样之后生成的更具 代表性的空间特征向量。

Excitation 模块公式为:

$$M_{\rm s} = f^{7\times7} [(Z_{\rm s1}); (Z_{\rm s2})] , \qquad (7)$$

式中: *M*<sub>s</sub>为经过卷积融合后的通道特征; *f*<sup>7×7</sup>(;) 为合并特征向量后经过7×7卷积网络。

Attention 模块公式为:

 $U = F_{\text{scale}}(\widetilde{u}_{c}, W_{s}) = \widetilde{u}_{c} \cdot \text{sigmoid}(M_{s})$ , (8)

式中U为通过所有注意力模块后最终得到的特征图。

S&CMNet 模型的混合注意力模块部分参数量 非常少,可以忽略,空间注意力模块的参数个数甚至 只有98个。除此之外本次实验只在特征提取主干 网络的末端层增加了注意模块,因此仅为网络少量 增加了计算的复杂度<sup>[16]</sup>。

### 2.3 实验流程

为了评估所使用的不同模型方法之间的性能, 引入语义分割较为常见的语义分割准确率指标,包 括平均交叉合并比(mean intersection over union, mI-oU)、F1 – Score 和像素准确率(pixel accuracy, PA) 以及一个图像预测速度指标 TIME<sup>[17]</sup>。

考虑到实验过程中,同时将标记数据集按照7: 2:1 的比例分为3部分,其中70%的数据用于模型 训练,20%的数据用于观察模型效果,10%数据用于 最终模型图像预测,验证模型效果。

使用 NDWI 作为传统方法的对比基线,实验过 程使用 ENVI 5.3 软件在 i7 - 9750H CPU 以及 NVIDIA 1660 TI GPU 设备上进行, 阈值设定范围为 0.2~1.0之间为水体。基于深度学习模型学习过 程,为防止 U-Net 模型编码器结构以及 ResNet101 主干网络对于实验的影响,从头开始训练 U-Net, SegNet, PSPNet 以及 DeepLabV3+ 模型<sup>[18-20]</sup>, 用以进 行消融对比试验。S&CMNet 模块实验部分,通过实 验优选,分别选择 ResNet101 和 U - Net 编码器为特 征提取主干训练,并分别命名为 S&CMNet V1 和 S&CMNet V2,模型架构如图3所示。网络均未单独 设置学习率,使用默认的 1e-3,采用 Adam 作为优 化器,进行自适应性的学习率更新方法,交叉熵作为 损失函数,考虑到二分类语义分割任务训练轮次过 多会造成过拟合现象,因此模型训练轮次初始设定 均为100轮。同时将 U-Net 相关网络批次、大小 (Batch Size)设置为4,将其他Batch设置为8。所有 深度网络模型均在 Pytorch 平台上实现,用 NVIDIA 2080Ti GPU 来进行实验<sup>[8]</sup>。



Fig. 3 S&CMNet V2 network model

#### 实验结果分析 3

表1列出了本文模型在所设计数据集上的分割 结果,其中100<sup>-1</sup>表示预测每一百张512×512大小 的图需要花费的时间。从表1中实验结果可以看 出,与传统遥感图像水体检测方法 NDWI 相比,使用 神经网络的方法在3个评价指标上都取得了更高的 检测精度,在3个评价指标上平均增加了23.98, 46.06 和 41.73 百分点,证明了深度学习算法在遥感 影像的语义分割方面具有可靠性和有效性;另外, 对于 S&CMNet V2,在 mIoU 和 F1 - Score 上比同时 将 VGG 网络作为主干网络的 U-Net 和 SegNet 分 别增加了1.49,0.93 和 1.97,1.21 百分点,这说明 混合注意力模块能够捕获有意义的通道特征信息以 及聚合更多的位置信息;最后,对于 S&CMNet V1, 预测时间为190.66 s,是所有语义分割模型中耗时 最少的,这在需要应用于实时语义分割场景任务有显 著优势。

表1 S&CMNet 模型语义精度 Tab. 1 S&CMNet model semantic accuracy

模型名称	PA/%	mIoU/%	F1 – Score/%	<i>TIME/</i> (s • 100 <sup>-1</sup> )
NDWI	67.52	41.48	51.30	_
U – Net	90.43	88.33	93.50	252.25
SegNet	93.04	87.85	93.22	419.03
S&CMNet V2	92.66	89.82	94.43	291.31
PSPNet	88.64	86.07	92.16	223.62
DeeplabV3+	92.83	88.09	93.38	239.88
S&CMNet V1	91.37	85.09	91.51	190.66

为了进一步展示分割结果,分别进行了3次消 融实验,同时选取了3张不同区域的图像进行可视 化处理,表2展示了与传统方法 NDWI 相比,深度学 习分割模型如 U - Net 等能够得到具有更为连贯和 准确的分割图,同时可以看出对于红外波段反射不 敏感的区域 NDWI 并不能做出良好判断,而通过深 度神经网络进行特征提取后则对不同空间域大小的 水体都有较好的表现。

表 2 NDWI 分割方法与 U – Net 语义分割方法的分割结果 Tab. 2 Segmentation results of NDWI and U – Net semantic segmentation methods



为了从现有深度语义分割模型中寻找能够更好 适应水体提取任务的基线网络,本文进行了大量实 验,表3展示了现有深度分割模型之间的对比结果, 从网络模型的角度出发,SegNet 模型中使用的解码 器结构使最终的分割影像中出现了大量噪声,而 U-Net 更好地避免了这种情况;另外 DeeplabV3+和 PSPNet 的结果是对相当于原图 1/8 的特征图进行 下采样从而使结果会存在少许边缘模糊的情况,但 水体

是由于是以 ResNet101 为主干网络,其在多层语义 特征表述上有着更好的效果。结合表3得出的结 论,将 ResNet101 与 U - Net 的特征提取网络作为主 干网络,在此基础上引入混合注意力模块得到模型 S&CMNet V1 以及 S&CMNet V2,表4 展示了2 个模 型的可视化结果,可以看出 S&CMNet 模型均有较好 的表现,同时在细小河流区域展现出了更为准确的 分割结果。

### 表 3 基于 ResNet101 的语义分割网络的分割结果与先前实验结果之间对比

Tab.3 Segmentation results of based ResNet101 segmentation network compared with the previous experimental results



表4 S&CMNet 网络模型的分割结果

Tab. 4 Segmentation results of the S&CMNet network model



### 4 结论

本文提出了一种新的深度分割模型 S&CMNet 用于高分辨率遥感图像水体提取任务。通过实验评 估,证明了本文模型能够显著提高水体检测性能,并 具有高效的推理速度,且在 GPU 上实时运行速度快。

本文方法依旧存在一些需要解决的问题,本文 只引入了2个较为简单的通道注意力和空间注意 力,性能提升有限,但同时也证明了注意力机制在遥 感影像语义分割中的可行性。目前兴起于自然语言 处理技术的 Transformer 注意力模块快速发展,为语 义分割领域提供更多可行性。高分辨率遥感影像包 含了丰富地物信息以及影像普遍连续性等特点,似 乎更加适合将传统的端到端训练过程用序列到序列 的预测任务来提供替代视角。如何针对高分遥感影 像设计更好的深度学习网络结构,从而更好地解决 遥感应用领域中的诸多问题,需要计算机视觉、遥感

### 科学等众多领域研究者的共同努力。

### 参考文献(References):

[1] 方涛,霍宏,马贺平.高分辨率遥感影像智能解译[M].北京:科学出版社,2016:18-25.
 Fang T,Huo H,Ma H P. Intelligent interpretation of high resolution

remote sensing image[M]. Beijing: Science Press, 2016:18 – 25.

- [2] Deng F L. Research on multi level segmentation method and application of high resolution remote sensing image [D]. Beijing: University of Chinese Academy of Sciences, 2013.
- [3] Work E A, Gilmer D S. Utilization of satellite data for inventorying prairie ponds and lakes[J]. Photogrammetric Engineering and Remote Sensing, 1976, 42(5):685-694.
- [4] Frazier P S, Page K J. Water body detection and delineation with Landsat TM data [J]. Photogrammetric Engineering and Remote Sensing, 2000,66(12):1461-1468.
- [5] 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.
- [6] Kaufman Y J, Tanre D. Atmospherically resistant vegetation index

(ARVI) for EOS - MODIS[J]. IEEE Transactions on Geoscience and Remote Sensing, 1992, 30(2):261 - 270.

- [7] Xu H Q. A study on information extraction of water body with the modified normalized difference water index (MNDWI) [J]. Journal of Remote Sensing, 2005, 9(5):589-595.
- [8] Lecun Y L, Bottou L, Bengio Y, et al. Gradient based learning applied to document recognition [J]. Proceedings of the IEEE, 1998,86(11):2278 – 2324.
- [9] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25 (2):1097-1105.
- [10] He K, Gkioxari G, Dollár P, et al. Mask R CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017:2961 – 2969.
- [11] 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.
- [12] 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 changes [D]. Chongqing: Chongqing Normal University,2019.
- [13] Ying X, Wang Q, Li X, et al. Multi attention object detection model in remote sensing images based on multi – scale [J]. IEEE

Access, 2019, 7:94508 - 94519.

- [14] He N, Fang L, Plaza A. Hybrid first and second order attention U net for building segmentation in remote sensing images [J]. Science China Information Sciences, 2020, 63 (4):1-12.
- [15] Cortes C, Vapnik V. Support vector networks [J]. Machine Learning, 1995, 20(3):273 – 297.
- [16] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313 (5786):504 – 507.
- [17] Zhang X, Xiao P, Feng X, et al. Toward evaluating multiscale segmentations of high spatial resolution remote sensing images [J].
   IEEE Transactions on Geoscience and Remote Sensing, 2015, 53 (7):3694 3706.
- [18] Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4):834 – 848.
- [19] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:2881 – 2890.
- [20] 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. Cham; Springer, 2015;234 – 241.

## Water information extraction from high – resolution remote sensing images using the deep – learning based semantic segmentation model

SHEN Jun'ao<sup>1</sup>, MA Mengting<sup>2</sup>, SONG Zhiyuan<sup>1</sup>, LIU Tingzhou<sup>1</sup>, ZHANG Wei<sup>1,2</sup>

(1. School of Software Technology, Zhejiang University, Ningbo 315048, China; 2. School of Computer

Science and Technology, Zhejiang University, Hangzhou 310027, China)

**Abstract**: Water information extraction is an important study direction in the application of high spatial resolution remote sensing images. Conventional recognition methods only focus on the shallow features of water. Therefore, to further improve the robustness of water information extraction algorithms and increase the segmentation precision by extracting more deep information from remote sensing images, this study proposed a water classification method using the semantic segmentation model based on deep learning. First, deep neural networks were used to mine the information from high – resolution remote sensing images. Then, attention modules were used to integrate the deep information, a new deep semantic segmentation model with higher precision and prediction efficiency than existent models was built. Finally, the ablation experiment was conducted to compare with conventional recognition methods and common semantic segmentation models. The experiment demonstrates that the proposed algorithm model yields higher overall precision and efficiency than previous methods and that the algorithm parameters are easy to set and less human intervention is required in the model. This study proved the accuracy and efficiency of deep learning and attention mechanism on water information extraction from high – resolution remote sensing images. Moreover, this study provided a possible solution for the segmentation of high – resolution remote sensing images using the deep learning method and explored the future prospect of the solution.

Keywords: semantic segmentation; multi - scale; remote sensing image; full convolutional network; attention mechanism