

## 基于 Res2Net 的多尺度遥感影像海陆分割方法

高慧, 阎晓东, 张衡\*, 牛艺婷, 王家琪

战略支援部队信息工程大学地理空间信息学院, 河南 郑州 450001

**摘要** 遥感影像海陆分割在海岸线提取、海岛礁识别、近岸目标检测等方面具有重要的应用价值。然而, 复杂多样的背景环境和海陆边界易导致出现海陆边界分割不准确的现象。针对上述问题, 提出了一种基于 Res2Net 的多尺度遥感影像海陆分割网络——MSRNet, 该网络利用深层卷积神经网络 Res2Net 提取影像的多尺度特征图, 并利用压缩和注意力模块对不同尺度的特征图进行特征增强, 以加强海陆弱边界信息; 将增强后不同尺度的特征图上采样至原始影像大小, 同时利用深度监督策略对不同尺度的预测结果进行损失增强。选取两组包含不同海岸类型的遥感影像数据集进行实验, 结果表明, 所提网络能够得到更准确的海陆分割结果和更加清晰完整的海陆边界。

**关键词** 遥感; 海陆分割; 深度学习; 多尺度特征; 深度监督策略

中图分类号 P237

文献标志码 A

DOI: 10.3788/AOS202242.1828004

## Multi-Scale Sea-Land Segmentation Method for Remote Sensing Images Based on Res2Net

Gao Hui, Yan Xiaodong, Zhang Heng\*, Niu Yiting, Wang Jiaqi

*Institute of Geospatial Information, Strategic Support Force Information Engineering University, Zhengzhou 450001, Henan, China*

**Abstract** The sea-land segmentation of remote sensing images has significance application value in applications such as coastline extraction, island reef identification, and near-shore target detection. However, it is prone to inaccurate sea-land boundary segmentation owing to the influences of complex and diverse background environment and the sea-land boundary. To address the aforementioned problem, a multi-scale sea-land segmentation network, namely MSRNet, for remote sensing images based on Res2Net is proposed in this paper. It uses a deep convolutional neural network named Res2Net to extract multi-scale features of images and applies a squeeze and attention module to enhance features at different scales to strengthen the weak sea-land boundary information. The enhanced feature maps at different scales are then upsampled to the original image size, and the prediction results at different scales are loss-enhanced using a deep supervision strategy. Two remote sensing image datasets containing complex scenes are selected for the experiments, and the results show that the proposed network can obtain more accurate sea-land segmentation results as well as clearer and more complete sea-land boundaries than other networks for various natural and artificial coastlines.

**Key words** remote sensing; sea-land segmentation; deep learning; multi-scale feature; deep supervision strategy

## 1 引言

随着遥感技术的快速发展, 遥感数据获取手段更加多样化, 遥感影像的空间分辨率、时间分辨率、光谱分辨率以及辐射分辨率也在不断提高, 为大范围海岸地区的研究提供充足的数据支撑。在遥感影像上, 将海洋区域和陆地区域进行有效的区分、实现快速高精度的海陆分割, 在海岸线提取、海岛礁识别及近岸目标

检测等方面具有重要的应用价值<sup>[1]</sup>。

传统的海陆分割方法包括阈值分割法<sup>[2]</sup>、活动轮廓模型法<sup>[3]</sup>、区域生长法<sup>[4]</sup>以及基于马尔可夫随机场的方法<sup>[5]</sup>等, 主要依靠影像中陆地和海域在灰度、纹理等方面的差异进行分割, 在水陆交界灰度差异明显、水边线形状简单的情况下能够取得较好的分割结果, 但是容易受噪声干扰, 需要人为设置参数来调控结果, 鲁棒性较差。近年来, 深度学习特别是卷积神经网络

收稿日期: 2022-01-24; 修回日期: 2022-02-22; 录用日期: 2022-03-11

基金项目: 国家自然科学基金 (41876105, 41671410)

通信作者: \*13783651715@163.com

(CNNs)在语义分割等领域取得了超越传统方法的优异成绩,为利用CNNs进行遥感影像海陆分割提供了基础<sup>[6-7]</sup>。文献[8-10]以编码-解码架构为基础,构建了更深层的语义分割网络,实现了端到端的遥感影像海陆分割,并利用全连通条件随机场(CRF)模型和形态学运算对分割结果进行后处理,进一步改善海陆分割结果。此外,由于海陆分割任务同样关注海陆边界分割的准确性,文献[11-13]针对这一问题提出了多任务框架,通过扩展网络分支提高海陆分割精度。现有研究多是针对边界清晰的岩石海岸或人工海岸的分割,然而大范围遥感影像中海岸类型多样,弱边界(淤泥岸线)和强边界(人工岸线)交替分布,语义分割网络难以同时实现多种场景下的海陆分割;此外,遥感影像中海陆交界处像素占比较低,造成了样本不平衡问题,使得网络在训练过程中对此区域的学习程度较低,容易产生边界分割不清的现象。

针对上述问题,本文设计了一种基于Res2Net的多尺度遥感影像海陆分割网络(MSRNet),用于提高网络对复杂场景影像的海陆分割能力。该网络引入深层卷积神经网络Res2Net<sup>[14]</sup>提取影像的多尺度信息,并在每层提取特征后引入注意力模块,增强弱海

陆边界的特征,同时利用深度监督策略增强网络学习目标边界信息的能力。所提网络以两组包含复杂场景的遥感数据集作为研究对象,引入边缘分割评价指标以更好地评价海陆边界处的分割结果。将该网络与目前主流的语义分割网络进行对比实验,结果表明本文网络在各评价指标中均达到最优,能够适用复杂场景下的海陆分割任务,提取到更加完整、连续的海陆水边线。

## 2 网络结构

本文提出的MSRNet结构如图1所示,网络整体为编码-解码结构,encoder<sub>m</sub>和decoder<sub>m</sub>( $m \in \{1, 2, 3, 4, 5\}$ )分别表示网络的编码层和解码层。网络编码过程使用Res2Net-50的预训练模型进行迁移学习,并在每层特征提取后引入压缩与注意力(SA)模块<sup>[15]</sup>,加强网络对弱边界的提取能力;解码过程将各特征图的上采样(upsample)作为网络的侧边输出(S1~S5),将各侧边输出进行组合以实现多层次特征融合(S6),并对各特征输出(S1~S6)使用深度监督策略进行处理,最后对网络输出的海陆分割结果进行边缘检测,从而获得水边线。

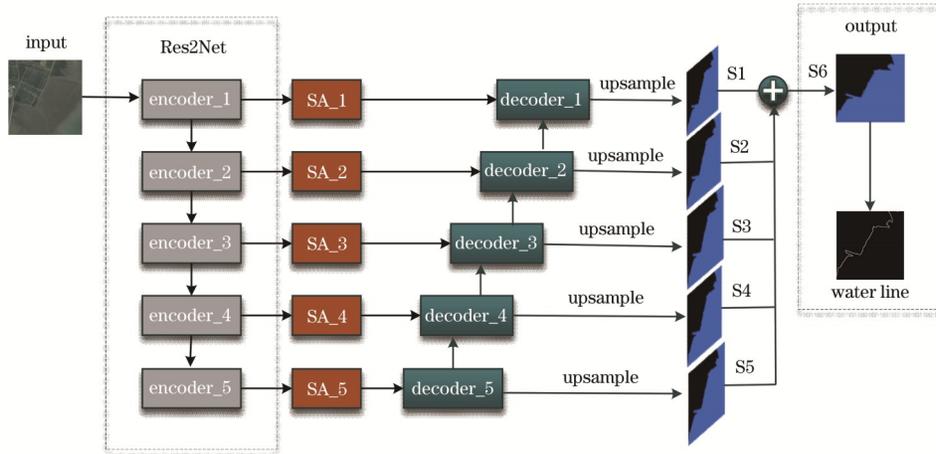


图1 网络结构图

Fig. 1 Architecture of the proposed network

### 2.1 多尺度编码结构

#### 2.1.1 Res2Net 网络

卷积神经网络中的多尺度表达能力非常重要,特别是语义分割任务需要利用来自不同尺度的上下文信息消除局部信息不足造成的歧义。Res2Net是一种新型多尺度网络,将ResNet残差块(Res\_Block)内的特征映射分成多个通道组,通过在不同的通道组之间设计一个类似残差的连接,使得网络在更细粒度的层次上提高多尺度表达能力。图2(a)为Res\_Block的结构图,通过在输入和输出之间添加一条跳跃连接,使得输入信息直接传输到后面的层,这简化了网络学习的难度。图2(b)为构成Res2Net的残差单元Res2\_Block的结构图,在 $1 \times 1$ 的卷积层之后将特征图分为 $s$ 个子集,分别用 $x_i$ 表示,其中 $i \in \{1, 2, 3, \dots, s\}$ 。每个特征图

子集 $x_i$ 具有相同的空间大小,但通道数是原输入特征图的 $1/s$ ,除 $x_1$ 外每个 $x_i$ 都有其对应的 $3 \times 3$ 卷积层,分别用 $K_i$ 表示,将 $x_i$ 经过卷积层 $K_i$ 后的输出设为 $y_i$ ,将特征图子集 $x_i$ 和经过 $K_{i-1}$ 的特征图子集 $x_{i-1}$ 相加后输入 $K_i$ ,得到输出 $y_i$ ,其中 $y_i$ 表达式定义如下:

$$y_i = \begin{cases} x_i & i = 1 \\ K_i(x_i) & i = 2 \\ K_i(x_i + y_{i-1}) & 3 \leq i \leq s \end{cases} \quad (1)$$

Res2\_Block中每一个 $3 \times 3$ 的卷积核 $K_i$ 可以接收到此前所有的特征图子集的信息,即 $\{x_j, j \leq i\}$ 的特征信息,使得特征图子集 $x_j$ 经过一个 $3 \times 3$ 的卷积后就可以得到一个比 $x_j$ 感受野更大的输出结果。为了让不同尺度的信息融合,将所有 $y_i$ 并联并采用 $1 \times 1$ 的卷积进行融合。这种分组、合并的策略使得卷积层能够更有

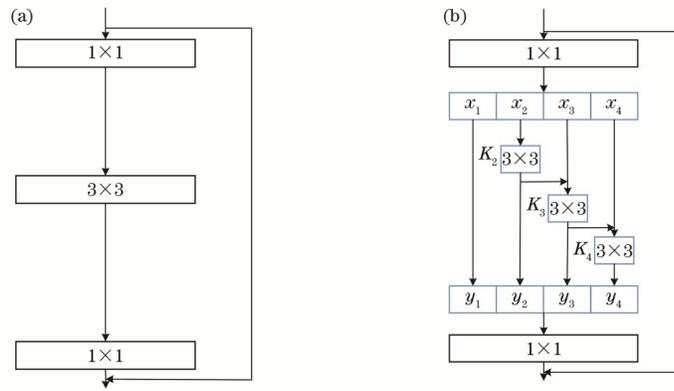


图 2 Res\_Block和Res2\_Block结构。(a) Res\_Block;(b) Res2\_Block( $s=4$ )

Fig. 2 Architectures of Res\_Block and Res2\_Block. (a) Res\_Block; (b) Res2\_Block ( $s=4$ )

效地处理特征图,使得 Res2\_Block 输出特征图包含不同感受野,有利于进行多尺度特征的提取,从而使网络可以在更佳细粒度的水平上捕捉局部或者全局的图像特征。

MSRNet 在编码层使用 Res2Net-50 网络,以输入影像大小为  $512 \times 512 \times 3$  为例,各层详细信息如表 1 所示。

表 1 编码层结构的输入和输出信息

Table 1 Input and output parameters of encoder layer

Encoder layer	Input	Convolution type	Kernel size	Stride	Output
Encoder_1	$512 \times 512 \times 3$	Conv	$7 \times 7$	2	$256 \times 256 \times 64$
	$256 \times 256 \times 64$	maxpooling	$3 \times 3$	2	$128 \times 128 \times 256$
Encoder_2	$128 \times 128 \times 256$	$3 \times$ Res2_Block	$3 \times 3$	1	$128 \times 128 \times 256$
Encoder_3	$128 \times 128 \times 256$	$4 \times$ Res2_Block	$3 \times 3$	2	$64 \times 64 \times 512$
Encoder_4	$64 \times 64 \times 512$	$6 \times$ Res2_Block	$3 \times 3$	2	$32 \times 32 \times 1\ 024$
Encoder_5	$32 \times 32 \times 1\ 024$	$3 \times$ Res2_Block	$3 \times 3$	2	$16 \times 16 \times 2\ 048$

2.1.2 压缩与注意力模块

注意力机制能够提高对有效信息的利用率,降低噪声干扰,从而被广泛应用于语义分割领域。SENet (Squeeze-and-Excitation Networks)<sup>[16]</sup>是最早提出的通道注意力机制,分为压缩和激励两部分。如图 3(a)所示,已知输入特征图  $X \in \mathbb{R}^{C \times H \times W}$ ,  $C, H, W$  分别为输入特征图的通道数、宽度以及高度,首先使用一个全局平均池化 (Avg.Pool) 操作压缩  $X$ ,然后利用两个全连接

层 (FC) 捕获各特征通道之间的相互依赖关系,从而得到激励权重  $\omega$ ,最后将其输出与  $X$  的卷积输出  $X_{res}$  逐通道相乘,以实现特征图的通道重标定。SE 模块的输出  $X_{out}$  可以表述为

$$\omega = \varphi \{ \omega_2 \times \sigma [ \omega_1 \times \text{AvP}(X) ] \}, \quad (2)$$

$$X_{out} = \omega \times X_{res} + X_{res}, \quad (3)$$

式中:  $\varphi(\cdot)$  为 sigmoid 函数;  $\sigma(\cdot)$  为 ReLU 激活函数; AvP 为平均池化操作;  $\omega_1, \omega_2$  为两个全连接层的参数。

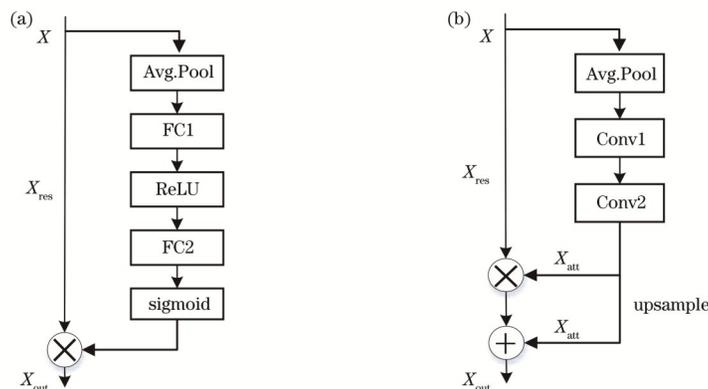


图 3 注意力模块结构图。(a) SE 模块结构;(b) SA 模块结构

Fig. 3 Architectures of attention module. (a) Architecture of SE module; (b) architecture of SA module

SA 模块扩展了 SE 模块的重新加权通道,分为压缩(Squeeze)和注意力(Attention)两部分。图 3(b)所示为 SA 模块结构图,其中 Squeeze 部分与 SE 模块一致,Attention 部分引入两个卷积核大小为  $3 \times 3$  的卷积层来聚集非局部特征,然后依据每个特征通道的重要程度,增强有用的特征并抑制对当前任务用处不大的特征。特征图  $X$  经过加权通道的两个卷积层的输出为  $\hat{X}_{\text{attn}}$ ,对其进行上采样至  $X_{\text{res}}$  大小得到  $X_{\text{attn}}$ ,将  $X_{\text{attn}}$  和  $X_{\text{res}}$  逐通道相乘,并与  $X_{\text{attn}}$  逐通道相加,得到输出  $X_{\text{out}}$  为

$$X_{\text{attn}} = \text{Up}[\sigma(\hat{X}_{\text{attn}})], \quad (4)$$

$$X_{\text{out}} = X_{\text{attn}} \times X_{\text{res}} + X_{\text{attn}}, \quad (5)$$

式中:  $\hat{X}_{\text{attn}}$  表示加权卷积层 Conv1、Conv2 的输出;  $\text{Up}(\cdot)$  表示上采样函数。MSRNet 在编码-解码对应层级引入 SA 模块,从而对弱海陆边界处的特征图赋予

更大的权重,以增强网络对弱边界特征的提取能力。

## 2.2 深度监督解码结构

MSRNet 包含 5 个解码模块,每 1 个解码模块分别由 1 个 Res2\_Block 和 1 个 upsample 组成,用于逐步恢复特征图至原始输入图像大小,特征图每经过 1 个 upsample 层,尺寸增大 1 倍。其中,将每 1 个解码层与对应层级 SA 输出的特征图融合,以更好地利用浅层特征图的细节信息和深层特征图的语义信息,从而产生更加精确的海陆分割结果。MSRNet 将解码阶段每层特征图上采样至原图大小作为侧边输出,并将各特征图进行维度上的叠加(Concat),叠加结果可作为网络的最终输出结果,用于实现多层次的特征融合,提高海陆分割的准确性。MSRNet 解码层结构的详细信息如表 2 所示。

表 2 解码层结构的输入和输出信息

Table 2 Input and output parameters of decoder layer

Layer	Input	Convolution type	Output	Side-output
Decoder_5	$16 \times 16 \times 2048$	SA_5	$16 \times 16 \times 1024$	$512 \times 512 \times 2$
	$16 \times 16 \times 1024$	Res2_Block+Conv+upsample	$32 \times 32 \times 1024$	
Decoder_4	$32 \times 32 \times 1024$	SA_4	$32 \times 32 \times 512$	$512 \times 512 \times 2$
	$32 \times 32 \times 512$	Res2_Block+Conv+upsample	$64 \times 64 \times 512$	
Decoder_3	$64 \times 64 \times 512$	SA_3	$64 \times 64 \times 256$	$512 \times 512 \times 2$
	$64 \times 64 \times 256$	Res2_Block+Conv+upsample	$128 \times 128 \times 256$	
Decoder_2	$128 \times 128 \times 256$	SA_2	$128 \times 128 \times 64$	$512 \times 512 \times 2$
	$128 \times 128 \times 64$	Res2_Block+Conv+upsample	$256 \times 256 \times 64$	
Decoder_1	$256 \times 256 \times 64$	SA_1	$256 \times 256 \times 32$	$512 \times 512 \times 2$
	$256 \times 256 \times 32$	Res2_Block+Conv+upsample	$512 \times 512 \times 2$	
Concat	$512 \times 512 \times 10$	Conv	$512 \times 512 \times 2$	

此外,网络在训练过程中采用深度监督策略(DS)<sup>[17]</sup>,分别对每个侧向输出进行一次损失计算,并将各侧向输出的损失反映在最终的损失函数中,以对网络的训练过程进行监督,计算公式为

$$L = \sum_{m=1}^M W_{\text{side}}^{(m)} l_{\text{side}}^{(m)} + W_{\text{fuse}} l_{\text{fuse}}, \quad (6)$$

式中:  $M=5$ ;  $l_{\text{side}}^{(m)}$  表示侧向输出的损失;  $l_{\text{fuse}}$  表示最终输出的总损失;  $W_{\text{side}}^{(m)}$  和  $W_{\text{fuse}}$  分别表示侧向输出损失和总损失的权重,其中损失函数选择用于语义分割二分类任务的二分类交叉熵损失函数,定义如下:

$$l = - \sum_{(r,c)}^{(H,W)} [P_{G,(r,c)} \lg P_{S,(r,c)} + (1 - P_{G,(r,c)}) \lg (1 - P_{S,(r,c)})], \quad (7)$$

式中:  $(r,c)$  表示像素点的坐标;  $H, W$  分别表示影像的高、宽;  $P_{G,(r,c)}$  和  $P_{S,(r,c)}$  分别表示该像素点的真值和预测值。训练过程中,通过不断学习网络权值参数,使得  $l$  的值趋于最低,从而达到网络收敛的目的。

## 3 实验数据与评价准则

### 3.1 实验数据

为验证所设计网络的有效性,选取两组公开的遥

感影像海陆分割数据集进行实验,分别为江苏连云港海岸带区域<sup>[18]</sup>的遥感影像记为数据集 1(Data1)、中国多个沿海地区的遥感影像<sup>[19]</sup>记为数据集 2(Data2)。由于遥感影像的覆盖范围广,语义分割等密集预测任务对计算资源的要求较高,因此在训练和预测时,两组数据集的大幅遥感影像被切分成若干片。表 3 列举了两组数据集的详细信息。

### 3.2 评价标准

本文分别从区域分割和边界检测两个方面评价各网络在海陆分割数据集上的性能。将 F1 分数(F1-score)、平均交并比(MIOU)和平均绝对误差(MAE)<sup>[20]</sup>作为区域分割的评价指标。其中:F1-score 和 MIOU 为常用的语义分割结果评价指标;MAE 是以像素点为单位计算预测结果与真值标签之间的平均绝对误差,能够更好地反映预测误差的实际情况,MAE 值越小,网络得到的预测结果与真值标签图越接近。

此外,海陆分割任务更加关注边界像素分割的准确性。为了评价网络对海陆边界的提取能力,本文采用 F1-score-b<sup>[21]</sup>作为边界分割质量的评价标准,通过对分割结果和标记真值进行边缘提取,得到像素宽度为 1 的分割边界。F1-score-b(其值用  $F_1$  表示)定义为预测边界像素与标记真值边界像素在  $\beta$  像素内的准确

表 3 数据集的详细信息  
Table 3 Details of two datasets

Parameter	Data1	Data2
Image source	Gaofen-1	Landsat-8
Bands	Red, green, blue, near-infrared	Red, green, blue
Spatial resolution /m	8	30
Image size	256×256	512×512
Number of training samples	1544	1950
Number of validation samples	178	1411
Number of test samples	192	—

率,计算公式如下所示:

$$F_1 = \frac{2P_\beta R_\beta}{P_\beta + R_\beta}, \quad (8)$$

式中: $P_\beta$ 和 $R_\beta$ 分别代表边界像素在 $\beta$ 像素内的精确率和召回率,本文 $\beta$ 设为3。

## 4 实验及结果分析

为了验证本文网络的有效性,将其与U-Net<sup>[22]</sup>、Deeplabv3+<sup>[23]</sup>、U<sup>2</sup>-Net<sup>[24]</sup>和RAUNet<sup>[25]</sup>进行对比。其中:U-Net、Deeplabv3+为经典的语义分割方法;RAUNet采用编码-解码结构,设计了增强注意力模块(AAM)以融合多层次特征并捕获全局上下文信息;U<sup>2</sup>-Net利用两级嵌套的U型结构和设计的残差U型模块(RSU),使网络能够从浅层和深层捕获到更丰富的特征信息。实验过程中,U-Net等对比模型均使用已公开的源代码实现,并将所有模型在两个海陆分割数据集上重新训练,以进行比较。

实验采用Windows下的Pytorch机器学习框架,硬件环境为CPU Inter(R) XeonE2176G, GPU GTX 2080Ti, 11G显存。各网络实验均在相同的环境中进行,各训练参数保持一致,选择Adam优化器,批大小

(batch size)设置为4,初始学习率设为0.0001,共训练50次。

### 4.1 实验结果定性分析

为了更加全面地比较各种方法对海陆的分割能力,分别对两数据集的海陆分割结果进行分析,并将本文网络的海陆分割边界与原始影像叠加,以更加直观地展示本文网络在海陆分割任务中的有效性。图4为数据集1中各网络的分割结果,该场景包含港口、防波堤等结构细长的水工建筑,用来检测网络对边界复杂的人工海岸的海陆分割能力。U-Net和RAUNet能够对部分防波堤区域进行提取,但是分割结果仍然不够准确,而U<sup>2</sup>-Net和Deeplabv3+完全将防波堤错分为海域,说明融合高低级语义信息能够在一定程度上提高网络的分割性能,但是对边界和细节的处理效果不够理想。所提网络在训练过程中使用深度监督策略,有效地保留了防波堤的边缘细节信息,在边界复杂的人工海岸区域提取到了相对完整、连续的海陆边界,分割结果明显优于其他网络。特别是图4(b)中左边矩形框区域的海水被港口包围,导致U-Net、Deeplabv3+等网络忽略此区域海域特征、将其错分为陆地,而本文网络能够区分出港口、海域类别,得到了更为准确的海陆分割结果。

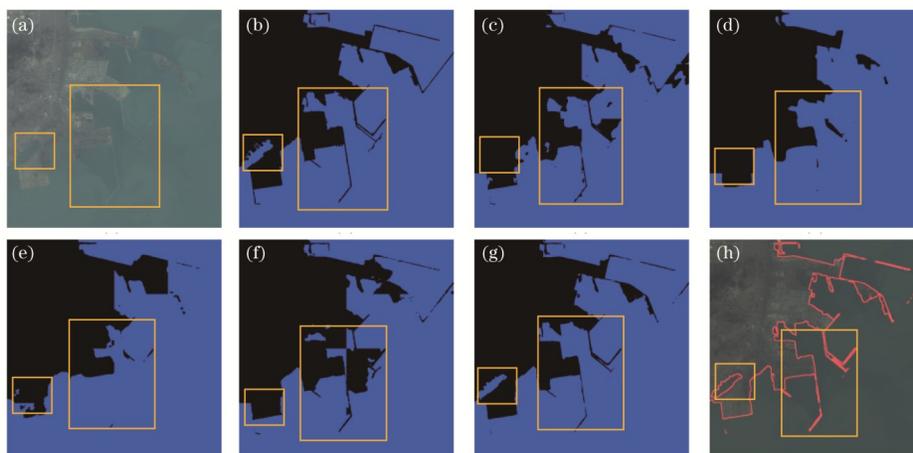


图 4 各网络在数据集 1 中的结果对比。(a)原始影像;(b)原始标签;(c)U-Net;(d)Deeplabv3+;(e)U<sup>2</sup>-Net;(f)RAUNet;(g)MSRNet;(h)水边线叠加结果

Fig. 4 Comparison of experimental results of various networks on Data1. (a) Original image; (b) original label; (c) U-Net; (d) Deeplabv3+; (e) U<sup>2</sup>-Net; (f) RAUNet; (g) MSRNet; (h) water line superposition result

图 5 为各网络在数据集 2 中的分割结果,该场景为包含多个岛屿的海岸影像,用来检测网络在水边线形状复杂下的海陆分割能力。如图 5 所示,岛屿的海陆边界通常较为明显,各网络在此类型影像的分割结果明显优于其他场景,但是岛屿形状、大小不规则,容易造成水边线出现严重凹陷边缘的现象(圆圈

区域),各对比网络难以处理此区域复杂的边界信息,均出现分类错误的现象;本文网络利用 Res2Net 提取影像的多尺度信息,弥补了局部区域信息不足造成的歧义,克服水边线凹陷带来的干扰,得到的海陆分割结果更加准确,提取的海陆边界也更加清晰、连续。

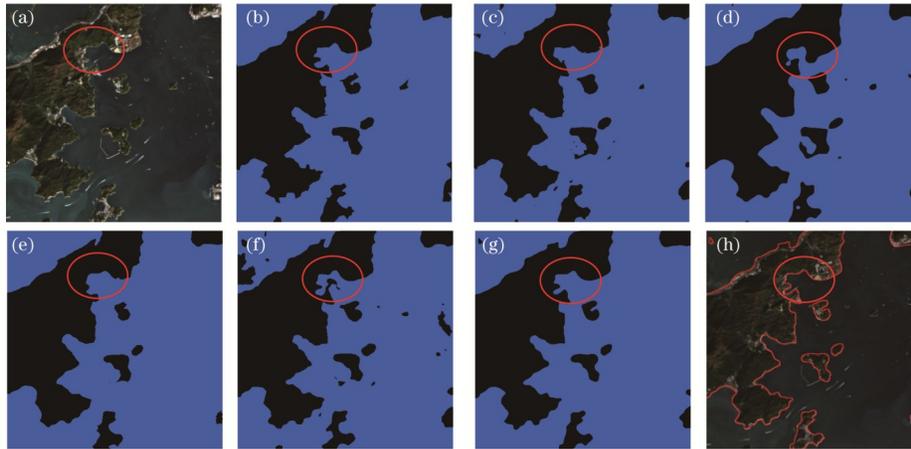


图 5 各网络在数据集 2 中的结果对比。(a)原始影像;(b)原始标签;(c) U-Net;(d) Deeplabv3+;(e) U<sup>2</sup>-Net;(f) RAUNet;(g) MSRNet;(h)水边线叠加结果

Fig. 5 Comparison of experimental results of various networks on Data2. (a) Original image; (b) original label; (c) U-Net; (d) Deeplabv3+; (e) U<sup>2</sup>-Net; (f) RAUNet; (g) MSRNet; (h) water line superposition result

#### 4.2 实验结果定量分析

表 4 统计了各网络在两个数据集上的评价结果。在数据集 1 中:U-Net 和 U<sup>2</sup>-Net 得到的 F1-score 等区域分割精度较低;而 U<sup>2</sup>-Net 采用深度监督策略,使得 F1-score-b 值高于 U-Net;Deeplabv3+ 区域分割精度均高于 U-Net 和 U<sup>2</sup>-Net,但由于 Deeplabv3+ 没有充分利用浅层特征图的细节信息,导致其 F1-score-b 较低。RAUNet 在数据集 1 中的 F1-score、MIOU 以及 MAE 分别为 98.38%、96.82%、0.016, F1-score-b 为 69.25%,达到了各网络分割结果的次优值。本文网络在数据集 1 中的各评价指标均优于其他网络, F1-score、MIOU 比 RAUNet 得到的次优值分别提高了

0.71 和 1.37 个百分点, MAE 减小为 0.009,另外本文网络在数据集 1 中 F1-score-b 值达到 80.14%,较次优值提高了 10.89 个百分点。

在数据集 2 中:U-Net 和 U<sup>2</sup>-Net 的区域分割精度较低,不及其他网络;Deeplabv3+ 在数据集 2 中 F1-score、MIOU 和 F1-score-b 均达到次优值;而 RAUNet 对该数据集的边缘分割能力较差, F1-score-b 仅为 49.77%,远低于其他网络。本文网络在数据集 2 中的各评价指标均达到最优, F1-score、MIOU 相比于次优值提高了 0.42 和 0.83 个百分点, MAE 减小了 0.187, F1-score-b 值精度达到 72.86%,较次优值提高了 17.38 个百分点。

表 4 各网络的分割结果对比

Table 4 Segmentation results of each method

Method	Data1				Data2			
	F1-score / %	MIOU / %	MAE	F1-score-b / %	F1-score / %	MIOU / %	MAE	F1-score-b / %
U-Net	97.92	95.92	0.021	51.90	98.13	96.33	0.979	52.84
Deeplabv3+	98.23	96.53	0.018	48.11	98.53	97.10	0.709	55.48
U <sup>2</sup> -Net	97.92	95.92	0.023	60.68	98.30	96.65	0.762	52.05
RAUNet	98.38	96.82	0.016	69.25	98.39	96.84	0.694	49.77
MSRNet	99.09	98.19	0.009	80.14	98.95	97.93	0.507	72.86

#### 4.3 消融实验

为了探究网络不同模块的效果,将 MSRNet 进行拆分,分别验证 Res2Net 模块、深度监督策略和 SA 模块的作用。图 6 为在数据集 1 上进行消融实验时海陆分割边界与原始影像的叠加结果。其中基本网络为编

码-解码(En\_Decoder)网络结构,为了验证 Res2Net 模块对影像多尺度特征的提取能力,将其替换为 ResNet 模块进行消融实验。如图 6(a)所示, ResNet 模块海陆分割结果中有大量错分像素, Res2Net 模块能在一定程度上改善海陆分割效果,但是对结构复杂的港口区

域分割效果仍较差。如图 6(c)所示,利用深度监督策略可提高海陆边界分割能力,本文网络在此基础上进

一步结合 SA 模块提高海陆边界特征提取能力,从而得到了更加完整、连续的海陆分割结果。

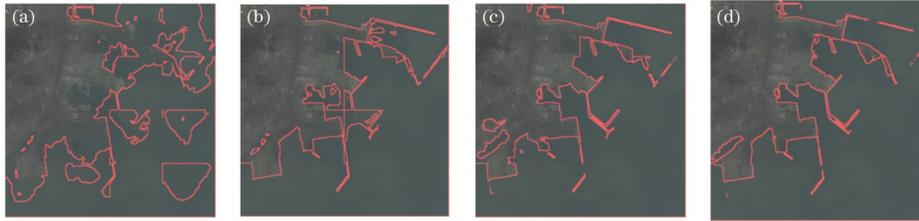


图 6 消融实验在数据集 1 中的结果对比。(a) En\_Decoder+ResNet; (b) En\_Decoder+Res2Net; (c) En\_Decoder+Res2Net+DS; (d) MSRNet

Fig. 6 Ablation experimental results on Data1. (a) En\_Decoder+ResNet; (b) En\_Decoder+Res2Net; (c) En\_Decoder+Res2Net+DS; (d) MSRNet

表 5 为在数据集 1 上进行消融实验时的 F1-score-b 值。结果表明:网络使用 Res2Net 模块时相比于使用 ResNet 模块的边界分割精度提高了 10.45 个百分点;采用深度监督策略后,相比使用 Res2Net 模块分割边界精度提高了 1.03 个百分点;加入 SA 模块后的网络 (MSRNet) 分割边界精度为 80.14%,相比使用深度监督策略提高了 5.03 个百分点,达到了各结构精度最优值。这些结果证明了 MSRNet 各模块在海陆分割任务中的必要性。

表 5 MSRNet 消融实验

Table 5 MSRNet ablation experiments

Architecture of network	F1-score-b / %
En_Decoder+ResNet	63.63
En_Decoder+Res2Net	74.08
En_Decoder+Res2Net+DS	75.11
En_Decoder+Res2Net+DS+SA (MSRNet)	80.14

通过以上分析,得出结论:1)注意力模块能够提高弱海陆边界处的特征响应,在弱海陆边界提取中具有优势;2)深度监督策略能够增强网络学习目标边界信息的能力,提高边缘分割的精度;3)本文网络能够适用于不同类型的遥感影像海陆分割,并且在区域和边缘检测结果中均能得到最优结果。

## 5 结 论

本文设计了一种基于 Res2Net 的多尺度遥感影像海陆分割网络——MSRNet,以实现复杂场景下的遥感影像海陆分割任务。网络选取 Res2Net 作为特征提取骨干网络,以获取影像的多尺度信息,并结合注意力模块加强网络对弱海陆边界的提取能力,同时利用深度监督策略改善海陆边缘分割效果。本文分别在两个数据集上进行实验,结果表明 MSRNet 在区域和边缘分割精度上均优于目前主流的语义分割网络。此外,通过消融实验验证了 MSRNet 各模块对海陆分割结果精度的必要性。

后续研究拟结合边缘检测的思想构建多任务网络结构,通过两个并行的卷积神经网络分别进行海陆区

域分割和边缘检测,以改善分割效果;通过对分割结果进行后处理,消除分割结果中的孔洞现象,以得到更加准确的海陆分割结果。

## 参 考 文 献

- [1] 张靖奇. 基于深度学习的卫星遥感图像舰船检测[D]. 哈尔滨: 哈尔滨工业大学, 2019: 1-78.  
Zhang J Q. Ship detection in satellite remote sensing images based on deep learning[D]. Harbin: Harbin Institute of Technology, 2019: 1-78.
- [2] Pardo-Pascual J E, Almonacid-Caballer J, Ruiz L A, et al. Automatic extraction of shorelines from Landsat TM and ETM+ multi-temporal images with subpixel precision[J]. Remote Sensing of Environment, 2012, 123: 1-11.
- [3] 喻金桃, 郭海涛, 李传广, 等. 二叉树与多种活动轮廓模型相结合的遥感影像水边线提取方法[J]. 测绘学报, 2016, 45(9): 1104-1114.  
Yu J T, Guo H T, Li C G, et al. A waterline extraction method from remote sensing image based on quad-tree and multiple active contour model[J]. Acta Geodaetica et Cartographica Sinica, 2016, 45(9): 1104-1114.
- [4] 詹雅婷, 朱利, 孙永华, 等. 海岸线遥感光谱角度—距离相似度生长模型自动化提取[J]. 遥感学报, 2017, 21(3): 458-469.  
Zhan Y T, Zhu L, Sun Y H, et al. Automatic extraction of coastline via spectral angle-distance similarity growth model[J]. Journal of Remote Sensing, 2017, 21(3): 458-469.
- [5] Baselice F, Ferraioli G. Unsupervised coastal line extraction from SAR images[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10(6): 1350-1354.
- [6] 张哲晗, 方薇, 杜丽丽, 等. 基于编码-解码卷积神经网络的遥感图像语义分割[J]. 光学学报, 2020, 40(3): 0310001.  
Zhang Z H, Fang W, Du L L, et al. Semantic segmentation of remote sensing image based on encoder-decoder convolutional neural network[J]. Acta Optica Sinica, 2020, 40(3): 0310001.
- [7] 李潇凡, 王胜强, 翁轩, 等. 基于 UNet 深度学习算法的东海大型漂浮藻类遥感监测[J]. 光学学报, 2021, 41(2):

- 0201002.
- Li X F, Wang S Q, Weng X, et al. Remote sensing of floating macroalgae blooms in the East China Sea based on UNet deep learning model[J]. *Acta Optica Sinica*, 2021, 41(2): 0201002.
- [8] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [9] Chu Z Q, Tian T, Feng R Y, et al. Sea-land segmentation with Res-UNet and fully connected CRF [C]//IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, July 28-August 2, 2019, Yokohama, Japan. New York: IEEE Press, 2019: 3840-3843.
- [10] Shamsolmoali P, Zareapoor M, Wang R L, et al. A novel deep structure U-net for sea-land segmentation in remote sensing images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(9): 3219-3232.
- [11] Cheng D C, Meng G F, Cheng G L, et al. SeNet: structured edge network for sea-land segmentation[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(2): 247-251.
- [12] Cheng D C, Meng G F, Xiang S M, et al. FusionNet: edge aware deep convolutional networks for semantic segmentation of remote sensing harbor images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, 10(12): 5769-5783.
- [13] 戴牧宸, 冷祥光, 熊博莅, 等. 基于改进双边网络的 SAR 图像海陆分割方法[J]. *雷达学报*, 2020, 9(5): 886-897.
- Dai M C, Leng X G, Xiong B L, et al. Sea-land segmentation method for SAR images based on improved BiSeNet[J]. *Journal of Radars*, 2020, 9(5): 886-897.
- [14] Gao S H, Cheng M M, Zhao K, et al. Res2Net: a new multi-scale backbone architecture[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(2): 652-662.
- [15] Zhong Z L, Lin Z Q, Bidart R, et al. Squeeze-and-attention networks for semantic segmentation[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 13062-13071.
- [16] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [17] Xie S N, Tu Z W. Holistically-nested edge detection [C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1395-1403.
- [18] Cui B G, Jing W, Huang L, et al. SANet: a sea-land segmentation network via adaptive multiscale feature learning[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 14: 116-126.
- [19] Yang T, Jiang S L, Hong Z H, et al. Sea-land segmentation using deep learning techniques for landsat-8 OLI imagery[J]. *Marine Geodesy*, 2020, 43(2): 105-133.
- [20] Liu J J, Hou Q B, Cheng M M, et al. A simple pooling-based design for real-time salient object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 3912-3921.
- [21] Qin X B, Zhang Z C, Huang C Y, et al. BASNet: boundary-aware salient object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 7471-7481.
- [22] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [23] 徐聪, 王丽. 基于改进DeepLabv3+网络的图像语义分割方法[J]. *激光与光电子学进展*, 2021, 58(16): 1610008.
- Xu C, Wang L. Image semantic segmentation method based on improved DeepLabv3+ network[J]. *Laser & Optoelectronics Progress*, 2021, 58(16): 1610008.
- [24] Qin X B, Zhang Z C, Huang C Y, et al. U2-Net: going deeper with nested U-structure for salient object detection [J]. *Pattern Recognition*, 2020, 106: 107404.
- [25] Ni Z L, Bian G B, Zhou X H, et al. RAUNet: residual attention U-net for semantic segmentation of cataract surgical instruments[M]//Gedeon T, Wong K W, Lee M. Neural information processing. Lecture notes in computer science. Cham: Springer, 2019, 11954: 139-149.