

DOI:10.19651/j.cnki.emt.2211851

基于门控金字塔融合的遥感影像海陆分割方法^{*}

李忠伟 王晓政 罗 德 刘旭阳 郭洪升

(中国石油大学(华东)海洋与空间信息学院 青岛 266580)

摘要: 海陆分割是通过遥感影像进行海岸线变化分析、资源管理等应用的重要基础,由于遥感影像场景复杂、陆地大小形状分布不均,海陆分割面临着误分类和边界分割不清等问题。针对上述问题,提出了一种用于遥感影像海陆分割的门控金字塔融合网络。首先通过基于注意力诱导的跨层聚合模块聚合两个深层特征,捕获全局上下文,准确而粗略地获取陆地的大小和形状信息。然后将聚合的全局特征送入门控融合模块,以全局信息为指导,在多尺度特征中选择有用的上下文信息,逐层优化边界细节并突出整个陆地区域。最后对每个侧输出进行全局监督。选取两组不同数据源的遥感影像进行实验,准确率分别为 99.13% 和 98.98%, F1 分数分别为 99.03% 和 98.89%, mIoU 分别为 98.26% 和 97.97%。实验结果表明,与其他算法相比,该算法具有更好的分割效果。

关键词: 海陆分割;语义分割;遥感影像;深度学习;门控金字塔融合

中图分类号: TP2 **文献标识码:** A **国家标准学科分类代码:** 510.1050

Sea-land segmentation method of remote sensing images based on gated pyramid fusion

Li Zhongwei Wang Xiaozheng Luo Cai Liu Xuyang Guo Hongsheng

(College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China)

Abstract: Sea-land segmentation is an important basis for the application of remote sensing images, such as coastline change analysis and resource management. Due to the complex scene of remote sensing images and uneven distribution of land size and shape, sea-land segmentation is faced with problems such as misclassification and unclear boundary segmentation. Aiming at the above problems, proposes a gated pyramid fusion network for sea-land segmentation in remote sensing images. Firstly, two deep features are aggregated through the attention-induced cross-layer aggregation module to capture the global context and accurately and roughly obtain the size and shape information of the land. Then, the aggregated global features are sent to the gated fusion module, guided by the global information, useful context information is selected from the multi-scale features, optimizes boundary details layer by layer and highlights the entire land area. Finally, global supervision is performed on each side output. Two sets of remote sensing images from different data sources were selected for experiments, the accuracy was 99.13% and 98.98%, the F1 score was 99.03% and 98.89%, and the mIoU was 98.26% and 97.97%, respectively. Experimental results show that this algorithm has better segmentation effect than other algorithms.

Keywords: sea-land segmentation; semantic segmentation; deep learning; remote sensing image; gated pyramid fusion

0 引 言

海岸带是人类生活发展的“黄金地带”,受到人类活动和自然变迁的影响,海岸线处于动态变化之中。将海洋与陆地精确绘制是海岸带监测与管理的一个重要手段,并且在海岸线变化分析、城市规划和资源管理保护等领域中有着极其广泛的应用。遥感影像具有地物场景丰富、空间分

辨率高、成像周期短等特点,随着遥感技术的发展,通过遥感影像进行海陆分割已成为热点研究方向^[1-2]。但同时遥感影像具有场景变化明显、陆域情况多样、海陆边界分布复杂等问题,导致准确的海陆分割具有挑战性。

传统的海陆分割方法主要包括阈值分割法、活动轮廓模型法、面向对象法和区域生长法等^[3]。经典阈值分割法包括直方图法^[4]、最大类间方差法等,该类方法通过算法选

收稿日期:2022-10-26

^{*} 基金项目:国家自然科学基金(62071491)、中央高校基本科研业务费专项资金(22CX01004A-1)项目资助

取阈值,将图像分割为海洋与陆地两部分。阈值分割方法通常简单有效,但是阈值难以确定,分割结果通常存在椒盐噪声的问题。活动轮廓模型法(active contour mode, ACM)^[5]的基本思想是在影像中定义一条曲线,在能量函数的控制下逐渐逼近海陆边界从而完成分割任务。基于区域的 ACM 使用图像强度和纹理等信息来引导曲线向海陆边界移动,可以减轻弱边界和噪声的影响。但是 ACM 模型较复杂,计算量高,实现难度比较大。面向对象的分割方法^[6]以对象为基本处理单元,利用对象之间的光谱、空间特征进行分割,忽略了对对象内部的特征差异,不能充分的利用图像信息。区域生长法^[7]主要思想是聚集相似的像素,首先选取一个或多个种子点,依据所制定的判决准则对像素进行合并,从而完成海陆分割。区域生长法在简单的场景中分割效果较好,但是种子点需要手动选取,分割结果容易受到判决准则和噪声的影响。

随着深度学习的发展,神经网络表现出了强大的特征提取能力,在图像分割^[8-9]、图像分类^[10-11]和目标检测^[12-13]等任务中表现出成功的应用。近年来,深度学习方法逐渐在遥感影像海陆分割任务中得到应用, Li 等^[14]提出 DeepUNet,与 UNet 一样采用编码解码器结构,使用了设计的卷积块代替了 Unet 在收缩和扩展路径中的卷积层以实现更高的精度。Xia 等^[15]提出了一种残差密集 UNet,通过使用密集的残差块构建深层的网络实现海陆分割任务。Cui 等^[16]在经典编码解码器结构的基础上,设计了一个自适应多尺度特征学习模块代替传统的卷积操作,并使用 SE^[17]模块连接编解码的相应层。Kang 等^[18]基于编解码器结构设计了一个多尺度上下文提取网络用于水体分割,增强了上下文信息提取能力。海陆分割的一个关键问题是需要高精度识别海陆分界线, Cheng 等^[19]将流行的深度神经网络用于海陆分割任务,并在分割之外添加了一个边缘检测分支来获得更好的效果。梁烽等^[20]同样提出了一个结合边缘检测的多任务网络用于 SAR 图像海陆分割。

虽然现在基于深度学习的海陆分割方法能够实现较高的精度,但是仍存在如下两个问题:1)部分网络只是为单一场景提出,例如清晰的人工岸线或基岩岸线,泛化性能较差,复杂的影像场景中地物光谱复杂,容易发生误分类的问题。2)影像海陆边界分布不明显,弱边界和强边界交替分布,现有的方法在分割时容易出现边缘模糊不清的问题。为解决上述问题,本文提出了一个基于门控金字塔融合的海陆分割方法,使用跨层聚合模块聚合具有高语义信息的特征层来生成全局指导信息,通过门控融合模块逐步融合多尺度特征,有效的控制信息流并减少特征冗余。通过与主流语义分割网络的对比实验表明,本文所提网络在各项指标中均达到最高,能够适应不同场景下的海陆分割任务。

1 门控金字塔融合网络

网络结构如图 1 所示,以单个遥感影像作为输入,并以

端到端的方式工作。网络主要由用于特征提取的编码模块、跨层聚合模块(content aggregation model, CAM)以及门控融合模块(gated fusion module, GFM)组成。编码模块生成从低到高的多级特征,将最后两个高级特征输入到跨层聚合模块来集成深度特征,生成具有高级语义上下文的全局信息。然后低级特征、辅助高级特征和全局信息输入到门控融合模块,利用相邻的高层特征以及全局特征来校准。整个网络采用自上而下的策略,逐步细化,对侧输出和全局特征执行深度监督,经过多级特征融合得到最终输出。

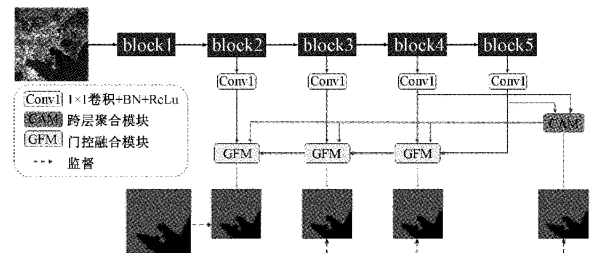


图 1 门控金字塔融合网络模型

1.1 编码模块

为了从输入图像中获取准确的语义信息,网络应该具有足够的深度和感受野来保证信息提取能力。采用预训练的 Resnet50^[21]作为主干提取特征。Resnet50 首先使用卷积层和最大池化层来降低特征空间分辨率,然后使用 4 个编码块来捕获更详细的语义特征。由于分割是像素到像素的任务,删除了最后的全局平均池化层和全连接层。5 个编码块提取不同分辨率的特征图,然后将提取的特征送入 1×1 卷积、批标准化(batch normalization, BN)和 ReLU 层将通道统一为 64。经过通道调整后得到多级特征 $F_i, i \in \{1, 2, 3, 4, 5\}$,大小分别为输入图像的 $1/2, 1/4, 1/8, 1/16$ 和 $1/32$ 。

1.2 跨层聚合模块

遥感影像具有复杂的光谱特征信息,陆地的位置和大小分布不均。为了从高级特征中捕获准确的指导信息并减少特征融合的混乱,使用 AFF^[22]来获取具有高级语义信息的跨层特征,这些特征有助于后续特征融合与优化,为整个网络的校准提供支持。AFF 通过引入多尺度通道注意力(MS-CAM)机制来有效融合跨层特征,该机制可以利用多尺度信息来获取更准确的上下文语义信息。MS-CAM 的核心思想是通过改变空间池化的大小来实现多尺度的通道注意力,为了尽可能保持轻量级,通过逐点卷积来关注通道的尺度问题,在注意力模块中将直接局部上下文添加到全局上下文中。如图 2 所示,MS-CAM 使用双分支结构,其中一个分支基于 SENet^[17],全连接层被逐点卷积代替,使用全局平均池化获取全局上下文,强调大面积分布的陆地;另一分支直接使用逐点卷积提取特征,保持原始特征大小以获得局部上下文,保证了小区域的获取。不同层次的特征融合可以相互补充,从而聚合多尺度通道上下文信息。

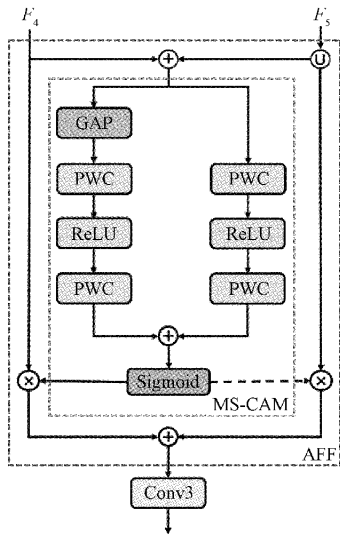


图 2 跨层聚合模块

MS-CAM 的局部通道上下文 $L(X)$ 和全局通道上下文 $G(X)$ 分别表示为:

$$L(X) = B(PWConv_2(\delta(B(PWConv_1(X)))))) \quad (1)$$

$$G(X) = B(PWConv_2(\delta(B(PWConv_1(GAP(X)))))) \quad (2)$$

其中, B 代表 BN 函数, δ 代表 ReLu 函数, $PWConv_2$ 和 $PWConv_1$ 代表内核大小不同的逐点卷积, GAP 代表全局平均池化。

将 AFF 应用于最后两层特征以获取全局信息, 基于 MS-CAM 的 AFF 过程可描述如下:

$$F_{45} = M(F_4 \oplus F_5) \otimes F_4 + (1 - M(F_4 \oplus F_5)) \otimes F_5 \quad (3)$$

其中, $M = (L(X) \oplus G(X))$ 代表 MS-CAM, F_4 和 F_5 代表两个输入特征, \oplus 表示初始特征融合, 即将上采样一倍后的 F_5 与 F_4 进行元素加法。图 2 中虚线表示 $(1 - M(F_4 \oplus F_5))$, 然后将相加后的特征送入 3×3 卷积、BN 和 ReLu 层得到最终的跨层聚合模块输出。

1.3 门控融合模块

CAM 能够大致定位陆地区域, 虽然捕获了一定的细节信息, 但是置信度较低, 缺乏岛礁地区和海陆边界的详细特征。粗略的陆地信息需要包含在高级特征图中的大尺度上下文语义信息, 而对于小规模目标的准确细节和海陆边界部分则需要使用低级特征图的空间细节。在海陆分割任务中, 一些海洋区域具有与陆地位置相似的特征信息, 传统金字塔结构直接对不同尺度的特征图进行拼接或者元素相加虽然能够获取良好的融合效果, 但是这些融合策略以同等方式对待所有特征, 受低级特征的噪声影响, 容易造成信息冗余和预测错误, 因此有必要对特征进行过滤和筛选。

门控机制是一种可以评估信息的重要性而对其进行保留、增强和抑制的机制。为了从多级特征中准确的提取详细信息并明确海陆边界, 受门控机制的启发提出了 GFM

以有效的整合多尺度特征。如图 3 所示, GFM 将全局特征和低级特征融合, 使用 CAM 输出的全局信息作为指导, 通过擦除全局特征的显著陆地区域引导网络从不同尺度探索边界细节和互补区域。

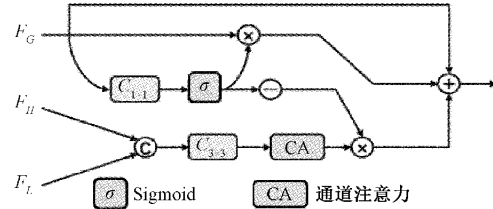


图 3 门控融合模块

首先, 对全局特征进行计算生成注意力权重:

$$g = \sigma(Conv_1(F_g)) \quad (4)$$

其中, $Conv_1$ 代表 1×1 卷积, σ 代表 Sigmoid 函数。将低级特征与辅助高级特征拼接起来, 然后采用通道注意力机制^[17]进行过滤, 以获得更具代表性的增强低级特征:

$$F_c = CA(Conv_3(Cat(F_i, F_{i+1}))) \quad (5)$$

其中, $Conv_3$ 代表 3×3 卷积、BN 与 ReLu 激活函数, CA 代表通道注意力。接下来, 将输出的低级特征和反向注意力权重组合起来以扩展对象区域, 对全局特征使用正向注意力权重。为了进一步增强特征的辨别能力, 使用跳跃连接将全局信息结合到最终输出:

$$F_{out} = F_g + F_g \otimes g + F_c \otimes (1 - g) \quad (6)$$

其中, \otimes 代表元素乘法。

2 数据集与评价指标

2.1 数据介绍

为验证所提网络的有效性, 选取两组数据集进行实验。数据集 1 采用 Landsat-8 OLI 遥感影像作为数据源, 通过人工标注的方式完成数据集的制作, 数据集 1 示例如图 4 所示。选取 2020~2022 近 3 年该卫星拍摄的中国沿海地区遥感影像, 合成 432 波段真彩图像, 接近地物真实色彩。将合成后的影像裁剪为 512×512 大小, 保证裁剪后的图像尽量全面包含海岸带周围的地物特征, 使用 Labelme 对图像进行标注、生成掩码。通过反转、旋转数据增强操作后得到网络所需数据集, 其中训练数据 1 400 张, 测试数据 410 张。数据集 2 为一组下载自 Google Earth 的公开海陆

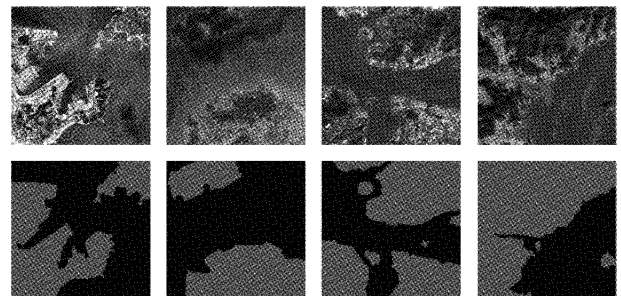


图 4 数据集 1 示例

分割数据集^[23],包含 208 幅图像,大小约为 $1\ 000 \times 1\ 000$,分辨率为 $3 \sim 5\text{ m}$ 。经过查阅相关文献数据集划分方法同时为了保持数据分布的一致性,将图像裁剪为 256×256 大小后按照 $8:2$ 的比例将数据集随机划分为训练集和测试集。

2.2 评价指标

为了全面评估所提模型与其他方法,采用 3 个指标来评价分割结果,包括平均交并比 (mIoU)、准确率 (Accuracy) 和 F1 分数。具体计算公式如下:

$$mIoU = \frac{1}{2} \left(\frac{TP}{TP + FP + FN} + \frac{TN}{TN + FN + FP} \right) \quad (7)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (8)$$

$$\begin{cases} Precision = \frac{TP}{TP + FP} \\ Recall = \frac{TP}{TP + FN} \\ F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \end{cases} \quad (9)$$

其中, TP 表示真正例,即将陆地正确分类; TN 表示真反例,即将海洋正确分类; FP 表示假正例,即将海洋错误预测为陆地; FN 表示假反例,即将陆地错误预测为海洋。

3 实验与分析

3.1 实验设置

实验环境为 GPU NVIDIA 3090, 24 GB 显存,操作系统 Ubuntu18.04。使用 pytorch 框架实现,采用交叉熵作为损失函数。选择 Adam 作为优化迭代器,初始学习率设置为 0.001,使用固定步长学习率衰减策略。令 lr_0 表示初始学习率,学习率更新如式(10)所示。

$$lr = lr_0 \times 0.94^{epoch} \quad (10)$$

其中, $epoch$ 代表当前迭代周期, lr 代表当前迭代周期学习率,即每迭代一次学习率变为原来的 0.94 倍,训练总迭代周期设为 100。

为验证本文所提方法的性能,实验选取 FCN-32s^[8]、UNet^[9]、PSPNet^[24]、Deeplabv3+^[25] 等模型进行对比实验。

3.2 对比实验

表 1 总结了各网络在两个数据集上的评估分数,所提网络同时使用注意力聚合和多尺度门控融合策略,与对比方法相比具有更高的优势。在数据集 1 中,5 种模型的准确率分别为 97.57%、98.46%、98.42%、98.89%、99.13%。在平均交并比方面本文方法比 FCN-32s 高 3.05%,比 UNet 高 1.32%,比 PSPNet 高 1.4%,比当前最先进的分割方法 Deeplabv3+ 高 0.48%。对于 F1 分数,所提网络也达到了最高的性能。

表 1 不同网络分割结果对比

方法	Data1			Data2		
	准确率	F1	mIoU	准确率	F1	mIoU
FCN-32s	97.57	97.31	95.21	97.26	97.03	94.64
UNet	98.46	98.29	96.94	98.84	98.73	97.69
PSPNet	98.42	98.24	96.86	98.54	98.41	97.11
Deeplabv3+	98.89	98.76	97.78	98.75	98.63	97.52
本文方法	99.13	99.03	98.26	98.98	98.89	97.97

在数据集 2 中,FCN-32s 分割精度最低,不及其他网络。UNet 在该数据集中准确率、F1 分数和 mIoU 均达到次优值。本文网络在数据集 2 中各项指标均达到最优,准确率、F1 分数和 mIoU 分别为 98.98%、98.89% 和 97.97%。

图 5 第 1 行是一个人工岸带,其中包含了一些养殖区,并且岸带周围泥沙含量较高。泥沙区域的光谱信息与陆地类似,FCN-32s、UNet 与 Deeplabv3+ 将部分泥沙区域认定为陆地,尤其是两个养殖区之间的海洋部分,PSPNet 虽然识别出了养殖区之间的海洋,但是边界并不明显。图 5 第 2 行是一个处于包含大量泥沙的海岸带,陆地和海洋的特征区别不明显,导致其他网络并不能准确识别出位于影像上部两个陆地之间的海域。图 5 第 3 行是一个码头区域的人工海岸,FCN 和 PSPNet 没有很好的识别出码头的特征,只能恢复码头大概的形状,边界也没有很好的对齐。

UNet 采用 U 形结构融合多尺度信息,但是这种不加区分的融合策略导致了信息冗余,将码头内部特征与海洋相似的部分识别为了海洋。Deeplabv3+ 同样发生了误检的问题,这说明多级特征的融合能够在一定程度上提升分割网络的性能。图 5 第 4 行海岸带周围漂浮了大量藻类物质,对比网络均不能很好的识别海陆区域,本文网络则较高精度的识别出海洋和陆地尤其是河流出口部分。图 5 第 5 行是一个包含自然岸线和人工岸线的大面积复杂场景,可以看出本文网络的性能也要优于其他网络。

图 6 为各个网络在数据集 2 上的分割结果,在图 6 第 1 行中,FCN 将码头附的水体错误预测为陆地并且没有识别出桥梁部分,这是因为 FCN-32s 下采样过程损失了特征信息,并且没有考虑全局上下文信息,不能很好的捕捉影像细节。UNet 和 Deeplabv3+ 将桥梁附近的水体错误预

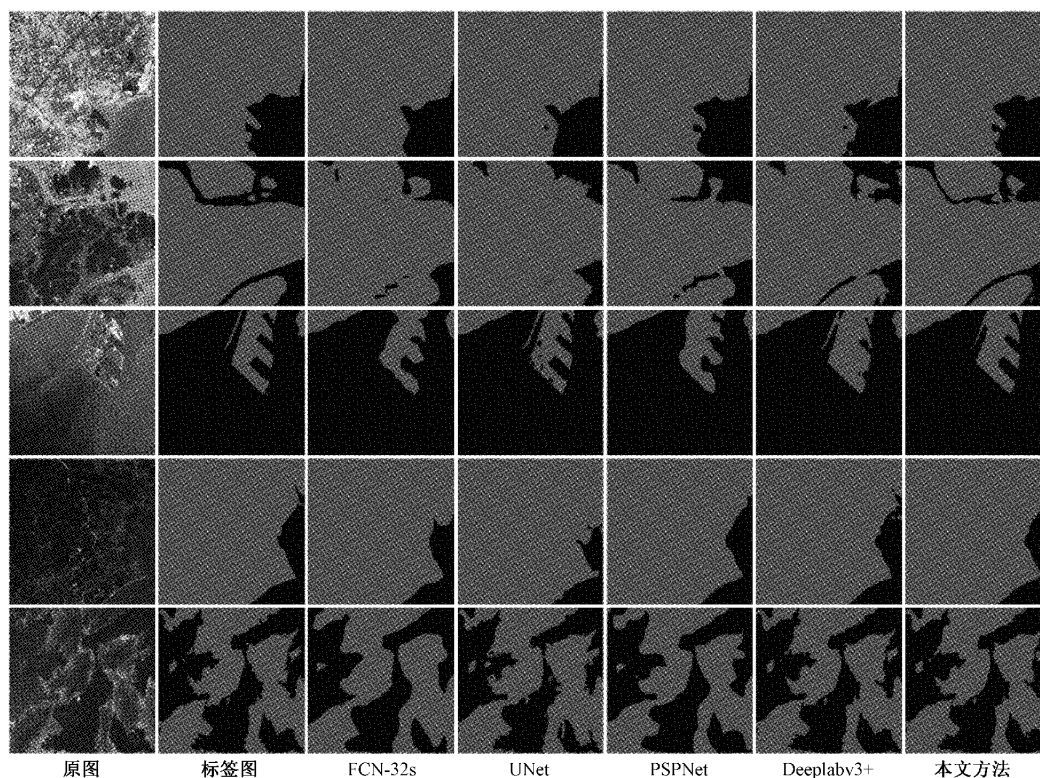


图5 各网络在数据集1的分割效果对比

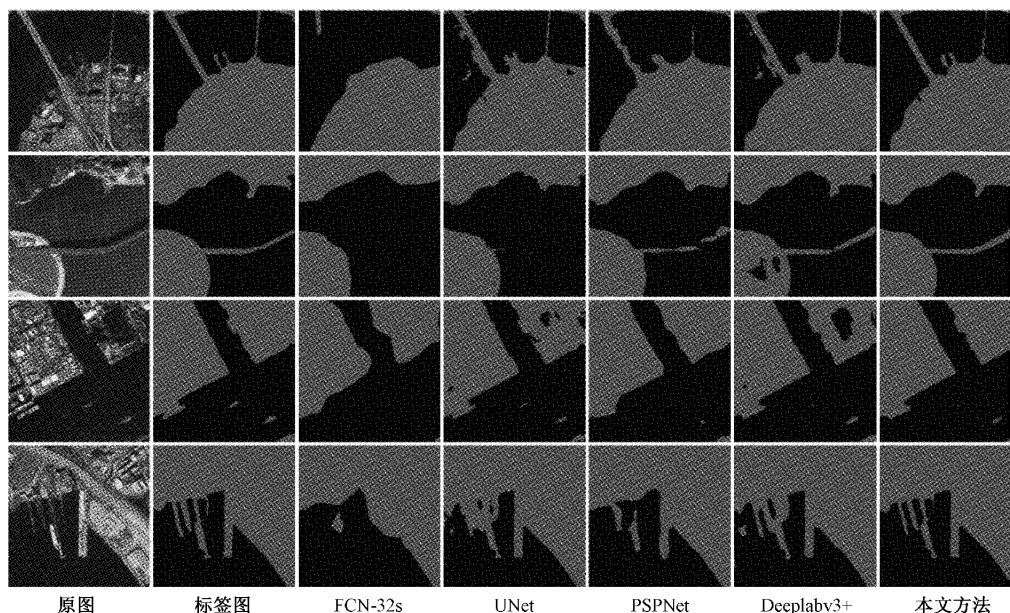


图6 各网络在数据集2的分割效果对比

测为陆地,这部分外观与陆地相似。在图6第2行中,UNet由于在低级特征提取阶段感受野较小,无法提取足够的纹理和边缘等信息,因此没有恢复出桥梁以及海岸带边界的详细特征,Deeplabv3+则出现了较大空洞现象。由于使用注意力聚合模块,本文网络更好的融合了语义尺度不一致的特征,聚合了全局上下文属性,从而提高了分割

质量。在图6第3行中,UNet和Deeplabv3+在一些特征与海洋相似的地方表现不佳,PSPNet使用了空间金字塔池化获取精确特征,不容易出现误分类的问题,但是池化层容易丢失细节信息,导致只能分割出大致的轮廓,边缘细节得不到很好的恢复。同时这些网络识别岸边码头的的能力稍显不足,而本文网络通过门控机制自适应动态调整

感受野,充分提取了图像信息以及边缘信息,精度明显优于其他网络。

3.3 消融实验

1) 模块贡献度分析

为了评估所提网络中 CAM 和 GFM 的个体贡献,在数据集 1 中进行模块贡献度分析实验。在表 2 中设计了完整的网络变体实验,即基线、基线+CAM、基线+GFM、基线+CAM+GFM。对于基线,直接移除 GFM,将 CAM 替换为上采样与元素加法融合获取预测结果。从表 2 中可以看出,CAM 将基线的准确率、F1 分数、mIoU 分别提高了 0.17%、0.19%、0.3%,GFM 将基线的准确率、F1 分数、mIoU 分别提高了 0.36%、0.41%、0.72%。在 CAM 和 GFM 的共同作用下,3 个指标分别提高了 0.44%、0.49%、0.88%。以上结果表明每个模块对于网络性能的提高都是有效的。

表 2 模块贡献度的消融实验结果 %

方法	准确率	F1	mIoU
基线	98.69	98.54	97.38
基线+CAM	98.86	98.73	97.68
基线+GFM	99.05	98.95	98.10
基线+CAM+GFM	99.13	99.03	98.26

2) 模块有效性分析

为了评估 CAM 和 GFM 的有效性,设计了 3 个消融实验在数据集 1 中进行实验。CAM¹ 代表使用卷积层代替 MS-CAM;GFM¹ 表示使用常规的卷积层代替 GFM;GFM² 表示使用逆向注意力机制代替门控融合模块。从表 3 中可以看出 MS-CAM 通过实现多尺度通道注意力,性能优于卷积层。所提出的 GFM 也比使用常规的卷积层融合实现了更高的精度。目前反向注意力的应用都是逐层递进引导网络顺序发现和补充对象细节。而所提出的 GFM 的注意力权重来自跨层的全局特征,通过擦除全局特征的显著区域引导网络从不同尺度的特征图中补充细节信息,避免相邻层特征分类错误进而对下一层产生影响。GFM² 的实验结果表明来自全局的指导信息能够更好的帮助特征融合。

表 3 模块有效性的消融实验结果 %

实验	方法	准确率	F1	mIoU
1	CAM ¹	98.95	98.83	97.91
2	GFM ¹	98.93	98.81	97.86
3	GFM ²	99.03	98.92	98.05
4	本文方法	99.13	99.03	98.26

4 结 论

本文根据遥感影像特点及分割中容易出现误分类和

边缘不清晰等问题,提出了一种新颖的门控金字塔融合网络。该网络为编码解码结构。跨层聚合模块侧重于高级特征的聚合,为下一步的特征融合提供准确的指导。门控融合模块通过相应的注意力门控机制整合了来自全局和低层的信息。通过上述两个模块,网络能够更加有效的完成特征提取和融合,有效缓解了误分类和边缘模糊等问题。此外,本文构建了一个基于 Landsat-8 数据源的海陆分割数据集,在该数据集和一组公开数据集上的实验可以看出,所提网络模型取得了更高的分割精度,具有良好的泛化性能。消融实验证明了算法各个模块合理有效,提升了海陆分割的性能。但是随着注意力门控机制的应用,使得模型变复杂,网络的计算量和训练时间也随之增加,如何研究更加轻量化且高效的特征融合策略也是后期努力的方向。

参考文献

- [1] YANG T, JAING 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.
- [2] 张静,景海涛,樊盛华. 基于海岸线数据库的遥感图像海陆分割 [J]. 电子测量技术, 2020, 43(23): 115-120.
- [3] 吴一全,刘忠林. 遥感影像的海岸线自动提取方法研究进展 [J]. 遥感学报, 2019, 23(4): 582-602.
- [4] 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.
- [5] WU Y, MA W P, GONG M G, et al. Novel fuzzy active contour model with kernel metric for image segmentation [J]. Applied Soft Computing, 2015, 34: 301-311, DOI: 10.1016/j.asoc.2015.04.058.
- [6] 陈云浩,冯通,史培军,等. 基于面向对象和规则的遥感影像分类研究 [J]. 武汉大学学报(信息科学版), 2006, 31(4): 316-320.
- [7] 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.
- [8] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [C]. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431-3440, DOI: 10.1109/cvpr.2015.7298965.
- [9] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]. Proceedings of the International Conference on Medical Image Computing and

- Computer-Assisted Intervention, 2015, 9351: 234-241, DOI: 10.1007/978-3-319-24574-4_28.
- [10] LIU Q S, HANG R L, SONG H H, et al. Learning multiscale deep features for high-resolution satellite image scene classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(1): 117-126.
- [11] PENG Y X, HE X T, ZHAO J J. Object-part attention model for fine-grained image classification[J]. IEEE Transactions on Image Processing, 2018, 27(3): 1487-1500.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2016: 779-788, DOI: 10.1109/CVPR.2016.91.
- [13] WANG C, BAI X, WANG S, et al. Multiscale visual attention networks for object detection in VHR remote sensing images [J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(2): 310-314.
- [14] LI R R, LIU W J, SUN S H, et al. DeepUNet: A deep fully convolutional network for pixel-level sea-land segmentation [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(11): 3954-3962.
- [15] XIA M, CUI Y C, ZHANG Y H, et al. DAU-Net: A novel water areas segmentation structure for remote sensing image [J]. International Journal of Remote Sensing, 2021, 42(7): 2594-2621.
- [16] CUI B G, JING W, LI Z R, 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, DOI: 10.1109/JSTARS.2020.3040176.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141, DOI: 10.1109/CVPR.2018.00745.
- [18] KANG J, GUAN H Y, PENG D F, et al. Multi-scale context extractor network for water-body extraction from high-resolution optical remotely sensed images [J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 103: 102499, DOI: 10.1016/j.jag.2021.102499.
- [19] 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.
- [20] 梁峰, 张瑞祥, 柴英特, 等. 一种结合上下文与边缘注意力的 SAR 图像海陆分割深度网络方法[J]. 武汉大学学报(信息科学版), 2023, 48(8): 1286-1295.
- [21] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778, DOI: 10.1109/CVPR.2016.90.
- [22] DAI Y M, GIESEKE F, OEHMEKE S, et al. Attentional feature fusion [C]. Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision(WACV), 2021: 3559-3568, DOI: 10.1109/WACV48630.2021.00360.
- [23] CHU Z Q, T T, FENG R Y. Sea-land segmentation with Res-UNet and fully connected CRF [C]. Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium, 2019: 3840-3843, DOI: 10.1109/IGARSS.2019.8900625.
- [24] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6230-6239, DOI: 10.1109/CVPR.2017.660.
- [25] CHEN L C, PAPANDREOU G, SCHROFF F. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]. Proceedings of the European Conference on Computer Vision, 2018, 11211: 833-851, DOI: 10.1007/978-3-030-01234-2_49.

作者简介

李忠伟, 博士, 教授, 主要研究方向为大数据智能处理及应用、光学及多手段融合与海岸带等。

E-mail: li.zhongwei@vip.163.com

王晓政, 硕士, 主要研究方向为遥感图像处理。

E-mail: wangxz029@163.com

罗德(通信作者), 博士, 副教授, 主要研究方向为人工智能、仿真无人系统等。

E-mail: luo_cai@upc.edu.cn

刘旭阳, 硕士, 主要研究方向为无人机海岸带区域覆盖路径规划。

E-mail: 851701571@qq.com

郭洪升, 硕士, 主要研究方向为无人机路径规划。

E-mail: 1272878410@qq.com