

文章编号:1671-4229(2024)03-0093-12

一种 DeepLabv3 + 结构改进的高分遥感影像红树林边界识别方法

吴耀炜, 龚建周*, 陈智勇, 袁海威, 林颖怡

(广州大学 地理科学与遥感学院, 广东 广州 510006)

摘要: 针对红树林自动监测与保护的迫切需求, 文章提出一种 DeepLabv3 + 改进模型的高分辨率遥感影像红树林的识别方案。改进方案主要包括在 DeepLabv3 + 的 ASPP (Atrous Spatial Pyramid Pooling) 结构中, 引入深度可分离卷积和 SE (Squeeze and Excitation) 注意力机制, 以及在解码端加入 CBAM (Convolutional Block Attention Module) 注意力机制和多尺度融合技术, 以提高模型对红树林关键特征的捕捉和表征能力, 从而减少漏检和误检现象。经过严格的精度评价, 改进后的 DeepLabv3 + 模型在总体精度上达到了 99.60%, 在召回率、红树林类交并比 (Mangrove-IoU) 和类 F1-score 上也分别达 96.05%、95.31% 和 97.60%。与原始 DeepLabv3 +、HRNet 和 PSPNet 模型相比, 改进模型在所有主要评价指标上表现更优, 红树林的识别准确性和边界提取能力明显提升。应用分析也进一步验证了模型的泛化能力和应用潜力。研究成果可优化红树林的实时监测技术。

关键词: 红树林边界识别; DeepLabv3 +; 注意力机制; 多尺度特征融合; 语义分割; 高分遥感影像

中图分类号: TP79; S718.5 文献标志码: A

A DeepLabv3 + architecture improved method for identifying mangrove boundaries in high-resolution remote sensing images

WU Yao-wei, GONG Jian-zhou*, CHEN Zhi-yong, YUAN Hai-wei, LIN Ying-yi

(School of Geography and Remote Sensing, Guangzhou University, Guangzhou 510006, China)

Abstract: In this study, aiming at the monitoring and protection of mangroves, an improved scheme based on the DeepLabv3 + model is proposed to enhance the recognition accuracy of mangroves in high-resolution remote sensing images. The improvement mainly involves introducing depthwise separable convolution and SE (Squeeze and Excitation) attention mechanism into the ASPP (Atrous Spatial Pyramid Pooling) structure of DeepLabv3 +, as well as incorporating a CBAM (Convolutional Block Attention Module) attention mechanism and multi-scale fusion technique into the decoding end. These innovative structural designs strengthen the model's ability to capture and represent key features of mangroves, thereby reducing missed detections and false positives. After rigorous accuracy evaluation, the improved DeepLabv3 + model achieved an overall accuracy of 99.60%, with recall, mangrove Intersection over Union (Mangrove-IoU), and mangrove F1-score reaching 96.05%, 95.31%, and 97.60%, respectively. Compared with the original DeepLabv3 + model and other popular models such as HRNet and PSPNet, the improved model demonstrated superior performance in all major eval-

收稿日期: 2024-04-03; 修回日期: 2024-04-23

基金项目: 国家自然科学基金面上资助项目(42071123)

作者简介: 吴耀炜(1999—), 男, 硕士研究生. E-mail: 18350920211@163.com

* 通信作者. E-mail: gongjzh66@126.com

引文格式: 吴耀炜, 龚建周, 陈智勇, 等. 一种 DeepLabv3 + 结构改进的高分遥感影像红树林边界识别方法[J]. 广州大学学报(自然科学版), 2024, 23(3): 93-104.

uation metrics, significantly improving the recognition accuracy and boundary extraction capability of mangroves. Furthermore, the model's generalization ability and application potential were also verified in the application analysis. The application analysis further verified the generalization ability and application potential of the model. The research results can optimize the real-time monitoring technology of mangroves.

Key words: mangrove boundary identification; DeepLabv3 +; attention mechanism; multi scale feature fusion; semantic segmentation; high resolution imaging

红树林是一类生长于热带、亚热带沿海潮滩间或河海接壤处的耐盐木本植物群落,包括红树科植物和耐盐常绿非红树科植物^[1-2]。红树林拥有诸多特殊生态功能,如:在高含盐量沙土内生长与发展,使其在抵抗海水侵蚀岸堤、防风防浪方面表现优秀^[3],并且具有净化水质^[4]、提供建筑材料、作为燃料、药用以及食用等价值^[5]。近半个世纪以来,由于人类过度砍伐、气候变化和环境污染等因素,使得红树林的面积断崖式下降^[6],进而引发海水加速侵蚀岸堤、受污染的水体流入海洋等生态环境问题^[7]。有研究表明,若红树林得不到妥善的保护与恢复,全球的红树林不出百年便全部消失殆尽,到时人类将直面诸如台风之类的自然灾害,造成无法估量的损失^[8]。

监测红树林面积的动态变化是红树林恢复与保护的重点工作,准确识别其边界则是动态监测的基础。遥感技术的进步,特别是在空间分辨率的提高和时间分辨率的缩短方面,已极大地促进了生态环境监测的效率和准确性^[9],这些技术的提升不仅使遥感数据能够提供更详细的地表信息,还能够以更高的频率更新自身数据,这对于理解和监控红树林的分布及其变化尤为重要^[10],使用遥感技术能够快速准确地识别和全面掌握红树林现状^[11]。在很长一段时间里,人工目视解译作为获取遥感地学信息的主要途径,是因为其能够有机结合遥感信息与专家知识,在提取精度上达到一个较高水准,尤其面对地物场景复杂且纹理结构多样的情景下,依旧能够保持较高水准精度。虽然相对于传统的地面测绘而言,人工目视方法在效率上占据显著优势,但依旧存在一定的局限性,主要有:①十分依赖解译人员的专业遥感及地学相关知识,用人成本大大提高;②目视解译人员自身生理和心理的原因,容易出现分类错误的情况;③非量化的标准,无法实现更为高效的解译效率^[12-13]。随着计算机技术的发展,人们逐渐将遥感解译技术与计算机相结合,在计算机高算力的支持下,人机交互解译实现了高效的数字化处理流程,主要包括:①将影像与解译结果同地理信息系统相结合;②监督分类与半监督分类实现了高效的半自动化解译;③基于光

谱特性,建立提取模型,实现对指定地物的识别提取。尽管与目视解译相比,其效率已然大幅度提升,但人为因素的参与依旧无法保证稳定结果的输出,因此,迫切需要更为智能化的影像信息提取技术^[14-15]。

随着机器学习技术的深化及其在各领域应用的推进,深度学习作为机器学习重要分支之一,其引领的卷积神经网络在计算机视觉领域大放异彩^[16]。在遥感影像处理应用方面,主要以高分辨率遥感影像作为数据源,利用卷积网络逐深度提取特征的特点,充分挖掘影像光学信息之间的联系与特征,再根据实际需求应用于图像分类、目标检测、语义分割或实例分割等领域^[17-18]。实际应用过程首先将源数据分割成切片,随后在批处理模式中提取特征,因此,它对遥感影像的尺寸没有限制,展现出优秀的鲁棒性,这一特点使其特别适用于大尺寸图像的处理^[19]。同时,当前主流的深度学习网络模型如 DeepLabv3 + 采用编码器-解码器结构,在提升分割效果的同时也会关注边缘细节。尽管如此,DeepLabv3 + 模型在编码端进行深层次的特征提取时,空间维度持续缩减,常会丢失细节特征,导致微小物体无法获得准确的分类^[20-21],还会导致空间位置信息的模糊,使得在解码端时不能实现精确的细节恢复;其空间金字塔池化模块(Atrous Spatial Pyramid Pooling, ASPP)使用的是标准空洞卷积,因其内部采用并行处理的方式,致使参数量增加,计算复杂度较高;而其解码端将ASPP模块输出的特征图进行大幅度上采样,也不利于丢失的空间位置信息的还原;除此之外,原始模型结构并未充分利用特征图特征,导致最终输出的分割结果在精度上大打折扣。进一步改进 DeepLabv3 + 模型结构,成为提升模型性能的重要途径。

基于红树林动态监测实际需求以及机器学习技术的现状,本研究提出了一种改进的模型结构,这种结构基于 DeepLabv3 + 网络框架进行了优化,探索基于高分二号遥感影像进行红树林识别与边界提取的方法。改进模型的策略涉及两个关键领域:①通过在 ASPP 模块中集成深度可分离卷积和 SE 注意力机制,旨在提升计算效率和特征的精确性;②解码阶段充分利用浅层和中

等深度的特征,通过融入 CBAM 注意力模块来增强模型对空间和通道信息的捕捉能力,并且采用多尺度融合技术,对特征图进行有效整合,从而构建了一个升级版的 DeepLabv3 + 架构。研究目的在于改进 DeepLabv3 + 模型,探索一种红树林快速识别与准确获取边界的技术方法,为红树林修护与保护的动态监测提供科技支撑。

1 实验区及数据

1.1 研究区概况

选取广东省湛江市廉江北部湾的安铺镇红树林为研究区域(图1)。湛江市位于广东西南部,同海南岛隔

海相望,介于北纬 $20^{\circ}13' \sim 21^{\circ}57'$ 之间,属于亚热带季风气候,气候适宜,年均温度为 23°C , 年均降水约为 $1\,400 \sim 1\,800\text{ mm}$ 间,海岸线长超过 $2\,000\text{ km}$, 为红树林生长提供了必要条件。湛江红树林国家自然保护区在我国国家级红树林保护区之中占地面积最大,约为全国红树林面积的 35.7% [22-23]。其中,安铺镇红树林植被覆盖度高,且分布较为连片集中,面积超过 60 万 m^2 , 主要树种有红海榄、桐花树、角果木、秋茄、白骨壤等十余个种类 [24]。九州江入海口及其沿岸延伸超过 10 km 的红树林区域,不仅面积广阔,而且形成了一片连续的生态群落,为红树林的研究提供了一个复杂且具有代表性的场所。

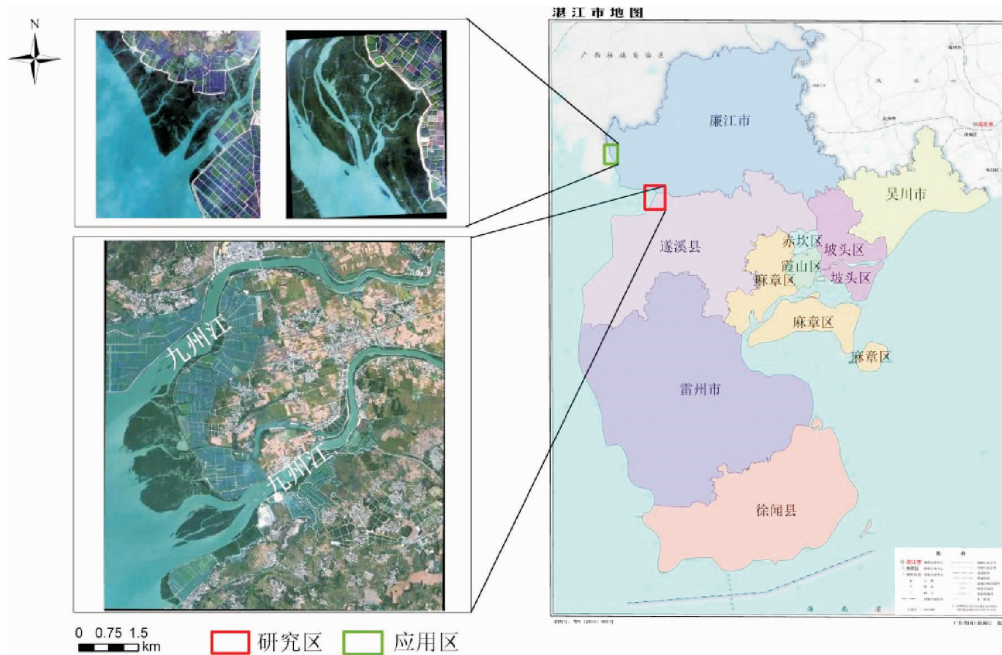


图1 研究区与应用区位置

Fig. 1 Location of research and application areas

注:该图基于广东省标准地图服务网站下载的审图号为粤 S(2018)093 号的标准地图制作,底图无修改。

1.2 数据来源与预处理

高分二号(GF-2)卫星是我国自主研发、自主发射的民用领域高空间分辨率、超宽幅光学遥感卫星,所搭载的两台高分辨率传感器能够提供 1 m 全色波段和 4 m 多光谱波段数据。高分二号卫星发射于 2014 年 8 月 19 日,并于同年 8 月 21 日首次开机工作,广泛应用于国土规划、交通运输、环境保护等领域 [25-26]。本研究使用覆盖湛江市的 GF-2 影像(2020 年 11 月 07 日),影像云层覆盖度小于 20% , 红树林区域无云。

对 GF-2 号影像进行预处理,包括辐射定标、大气校正、几何校正、全色和多光谱融合以及研究区裁切等操作,得到最终的研究区域 GF-2 遥感影像。

2 方法

运用 DeepLabv3 + 深度学习网络模型进行红树林识别,首先,需要对原始影像数据进行预处理,其次,需要制作网络模型训练的红树林训练样本数据集,然后,搭建 DeepLabv3 + 模型并进行优化,最后,运用模型对红树林进行识别与边界提取。

2.1 训练样本数据集制作

训练样本的数据规模、质量、多样性等是 DeepLabv3 + 的深度学习网络模型应用的基础与重要条件。训练样本数据集制作需要经过标签图像获取和训练数据集

生成两个流程。

(1) 标签图像获取

一方面,由于遥感影像尺寸巨大,使用常规的语义分割图像标注制作软件 LabelMe 只能先将遥感影像裁切后再进行勾绘,但裁切后的遥感影像无法保证标签的一致性。因此,利用 ENVI.5.6 中影像分类功能获得粗分类结果,随后在 ArcGIS10.8 中结合人工目视识别方法,由于红树林的冠层结构和树木分布模式在遥感图像上形成独特的纹理,这些纹理与周围地形和其他植被类型

相比较为均匀,并且通常沿海岸线或河流沿岸成带状分布,有利于对粗分类结果进行修正,勾绘出研究区内所有红树林的初步样本标签图斑。另一方面,结合野外实地调研打点、既有资料以及 Google Earth 不同时相遥感数据,对目视解译结果进行验证,以保证解译结果能够作为地面红树林真实标签。如图 2 所示,左侧图为经预处理后的 GF-2 遥感影像,右侧图为红树林真实标签图像,其中,绿色代表红树林,蓝色代表背景(非红树林地物,如水体、城镇建筑、陆地绿植等)。

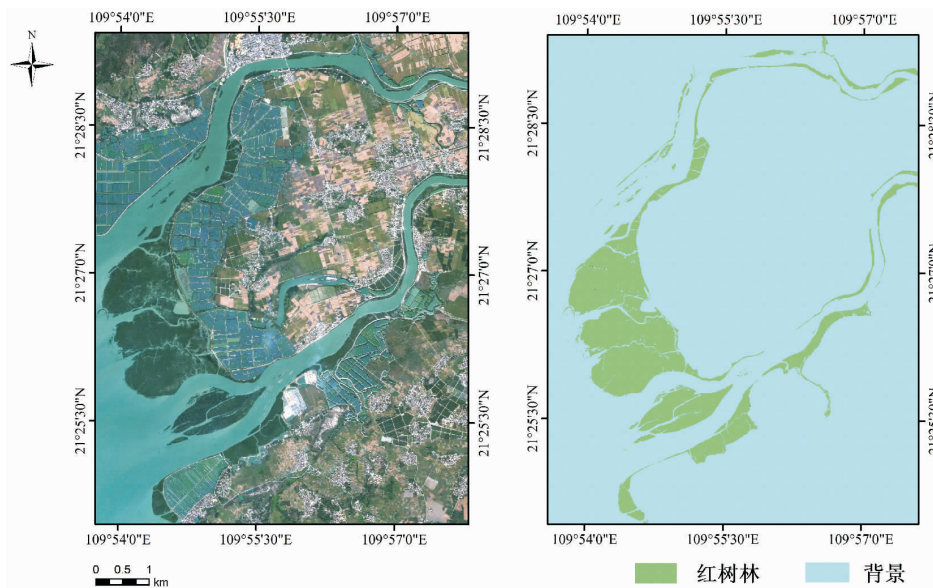


图 2 训练样本标签图

Fig. 2 Training sample label map

(2) 训练数据集生成

使用 ArcGIS10.8 的“Export Training Data For Deep Learning”工具,进行训练数据集的生成。包括采取滑动裁切的方法,尺寸为 512 × 512 像素,相邻切片 50% 重叠率,裁切方法示意如图 3 左图,共生成 1 120 对训练数据集。为防止样本过少导致模型训练出现过拟合,分别采

用色彩饱和度、亮度、对比度及锐度随机因子抖动方法结合随机角度旋转方法增强数据,最终获得 5 600 对训练数据集,如图 3 右侧的图对。使用 9:1 的比例,将训练数据集随机分组,分成 5 040 对样本作为模型训练数据,560 对样本作为验证数据。

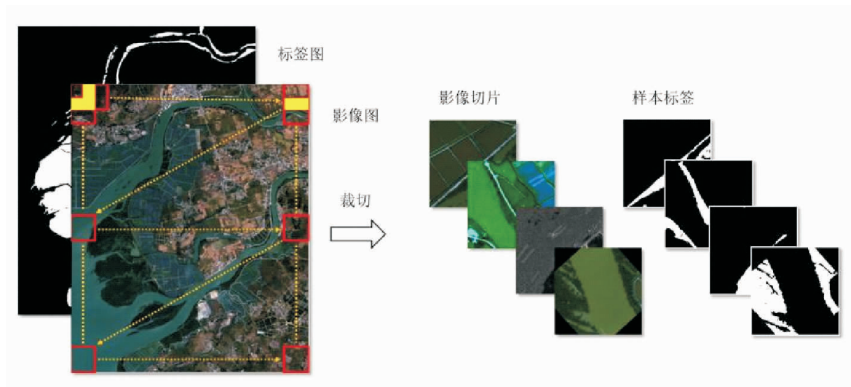


图 3 影像切片方法示意、样本对

Fig. 3 Schematic diagram of image slicing method and sample pairing

2.2 DeepLabv3 + 模型框架

DeepLabv3 + 是由谷歌研发的一种基于端到端训练方式的语义分割网络,是 DeepLab 系类中最新的分割网络,拥有更高的准确率与更快的速率^[27]。DeepLabv3 + 网络结构包括编码器(encoder)与解码器(decoder)两部分(图4):①编码器部分包括 Xception 深度卷积网络作为特征提取器,结合深度可分离卷积和残差结构,有效

解耦空间与通道相关性并避免梯度问题^[28]。ASPP 模块通过不同空洞率的卷积层及全局平均池化,捕捉不同大小目标的高层语义信息。②解码器则结合编码器输出的深层特征与浅层特征,通过上采样和卷积操作恢复空间信息,最终得到与原图尺寸相同的分割结果。整个结构充分利用了不同层次的特征信息,提高了分割的准确性和细节表现。

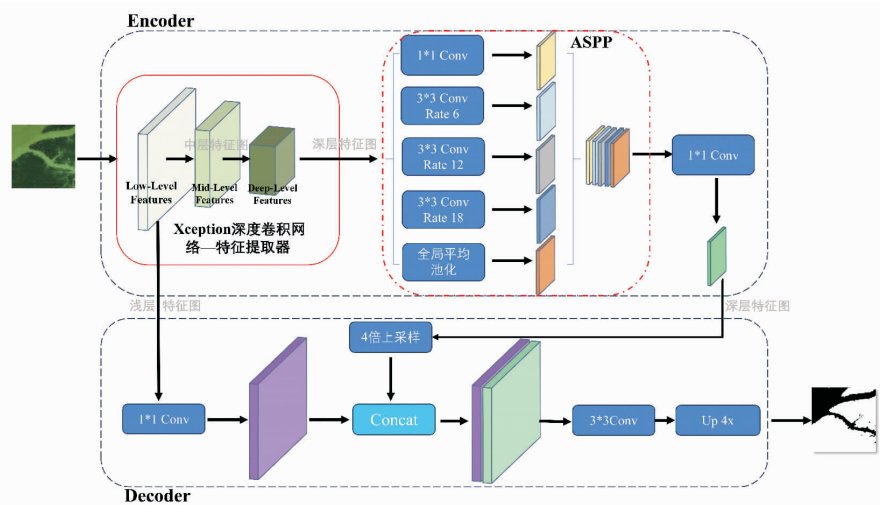


图4 DeepLabv3 + 模型结构

Fig. 4 The architecture of DeepLabv3 + model

DeepLabv3 + 语义分割网络是目前性能最为优异的分割模型之一,在众多领域已取得优秀的成绩^[29-30],但仍然存在诸多待优化之处:编码端深层特征提取易损失空间位置信息,影响解码端细节恢复;ASPP 模块采用标准空洞卷积并行处理,计算复杂度高;解码端大幅上采样 ASPP 输出,不利于恢复空间信息;此外,解码端未充分利用中层特征图,影响模型性能。因此,优化网络结构是提升 DeepLabv3 + 性能的有效途径。

2.3 DeepLabv3 + 模型结构改进

通过卷积方式、添加多种注意力机制和多尺度特征融合等网络结构改进,提高模型信息含量和模拟的空间位置准确度,从而提升 DeepLabv3 + 模型在高分遥感影像红树林边界识别上的性能。基于 DeepLabv3 + 原始模型(图4),绘制结构改进的 DeepLabv3 + 模型(图5),用灰色背景阴影突显结构改进,并以序号①、②和③区分改变卷积方式、添加多种注意力机制和多尺度特征融合等改进方案。

图5中,序号①为改变卷积方式的结构改进方案。采用深度可分离空洞卷积方式,改进编码端 ASPP 模块的各卷积层卷积方式。

序号②为添加多种注意力机制增加信息含量的结构改进方案。一是,在 ASPP 模块的 5 个分支操作后,都

分别引入 SE 注意力模块加强每个通道信息的重要性,并赋予相应的权值突出各通道的重要特征,目的在于增强深层次特征的信息。将 ASPP 改进后的模块更名为 DSASPP。二是,将 DSASPP 模块输出的特征图传入解码端,从原来的 4 倍上采样降低至 2 倍上采样,减小上采样幅度,以便于丢失空间信息的恢复。三是,增加利用特征提取器提取的中层特征图,将其输入至解码端,经 CBAM 注意力机制处理后,利用 3×3 卷积将其与 DSASPP 模块输出、经二倍上采样的深层特征图进行融合,通过丰富通道信息和空间信息来提高特征信息的含量。

序号③为多尺度特征融合提高空间位置准确度的结构改进方案。一是,使用 CBAM 注意力机制将中层特征图处理,DSASPP 模块输出的深层特征图经过两倍上采样,二者在通道维度上进行融合。二是,将浅层特征图经过 CBAM 注意力机制处理后,同上一步输出结果经两倍上采样后在通道维度上进行融合。一方面,将浅层特征图经过 CBAM 注意力机制处理,增强了特征的表达能力,更为聚焦于重要信息,如空间位置信息与通道信息;另一方面,中层特征图与深层特征图融合,增大了空间位置信息还原的力度;接下来,融合处理后的特征图逐步经 3×3 卷积、4 倍上采样恢复至输出影像尺寸,获得最终分割结果。

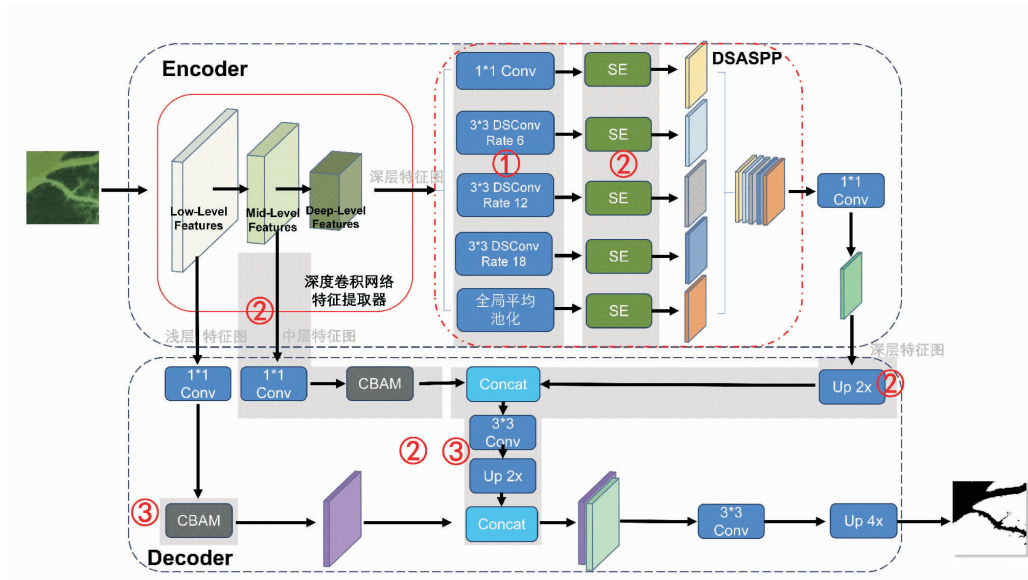


图 5 结构改进的 DeepLabv3 + 模型

Fig. 5 The architecture of improved DeepLabv3 + model

2.3.1 深度可分离空洞卷积

DeepLabv3 + 模型的 ASPP 模块采用标准卷积方式,即通过多通道卷积核对多通道特征图进行计算,其中卷积核的通道数与特征图的通道数相同,标准卷积的操作结果是生成单个特征图,这一过程计算量较大^[31]。不同空洞率的卷积核进行卷积提取运算时,由于不同大小的感受野,能够输出含有不同上下文信息的特征图。深度可分离空洞卷积是一种特殊的卷积方

式,采用两步流程的方案(图 6):第一步,采用相同空洞率但卷积核参数不同的逐深度卷积方法,即每一个通道都只对应一个空洞卷积核并都会生成一个特征图;第二步,采用 1×1 卷积核的逐点卷积方法,即对第一步提取的结果进行逐点卷积处理,以保证通道之间信息充分交流及提取特征的完整性,并且最终生成与标准空洞卷积相同尺寸的特征图^[32]。此种改进不仅降低了计算复杂度,还增强了特征信息的交流效率。

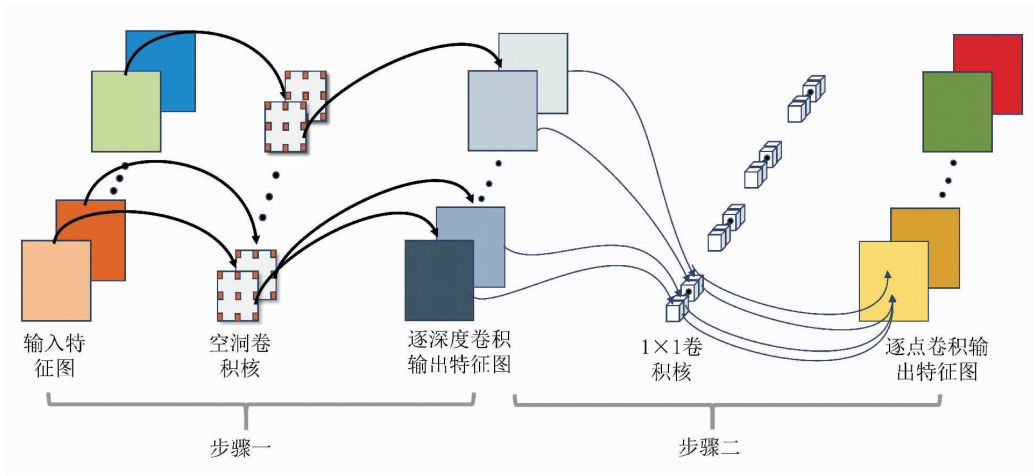


图 6 逐深度卷积图(左图)与逐点卷积图(右图)

Fig. 6 Depth wise convolutional graph (left) and point wise convolutional graph (right)

2.3.2 SE 注意力机制

DeepLabv3 + 模型在处理深层特征图时,忽略了通道间的关联信息,导致信息利用不足。注意力机制模仿人类注意力的分配,通过学习并加权关键信息来增强信息含量^[33]。其中,通道注意力机制(Squeeze and Excita-

tion,SE)是一种高效的注意力机制,通过对通道之间重要性的学习,保证通道信息的充分利用,以提升模型训练效率以及模型的准确性^[34-35]。将 ASPP 标准卷积替换为深度可分离空洞卷积,再增加 SE 注意力机制,由此形成的模块为 DSASPP(图 5)。

SE 注意力机制如图 7 所示,假设经过 ASPP 卷积处理后的深度(逐点卷积)特征图为 U,首先经过 F_{sq} 的全局平均池化处理,将特征图 U 压缩为 $1 \times 1 \times C$ 的向量,见式(1);然后,通过 F_{ex} 构建两个全连接层,配合使用 ReLU 激活函数,用于学习每个通道的重要性,即权重,见式(2);最后,将权重结果经过 sigmoid 激活函数(σ)归一化处理,得到 $1 \times 1 \times C$ 的通道权重向量,再利用 F_{scale} 将通道权重向量与特征图(U)相乘,得到通道加权后的特征图 \tilde{X} ,见式(3)。

$$z_c = F_{sq}(x_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i,j), \quad (1)$$

$$\hat{z} = F_{ex}(z_c) = T_2(ReLU(T_1(z_c))), \quad (2)$$

$$\tilde{X} = F_{scale} = X \cdot \sigma(\hat{z}), \quad (3)$$

式(1)~式(3)中, z_c 为特征图 U 压缩为 $1 \times 1 \times C$ 的向量, \hat{z} 为经过两个全连接层与 ReLU 激活函数后得到的通道权重向量, \tilde{X} 为最终得到的特征图,ReLU 为激活函数, σ 为 sigmoid 激活函数。

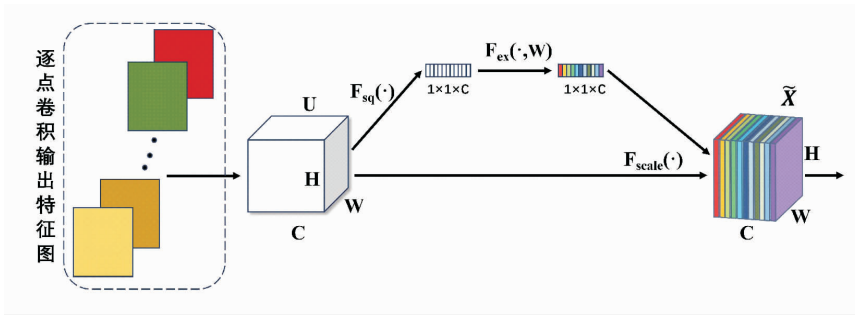


图 7 SE 注意力机制

Fig. 7 SE attention mechanism

2.3.3 CBAM 注意力机制

以往卷积神经网络注意力机制重心放置在通道维度,CBAM 注意力机制则是一种结合通道注意力与空间注意力的机制。本文对浅层特征图与中层特征图后进行了 CBAM 注意力机制处理(图 5),该机制的引入按照通道到空间的顺序依次进行。由于通道注意力能够学习通道之间的相关性,因而使通道信息得以充分利用;空间注意力则更加关注起决定作用的像素区域而忽略无关紧要的区域,这样既突出了重点又保留了像素位置信息,因此,从通道到空间将两者结合,达到特性信息融合和增强的目的^[36-37]。

CBAM 注意力机制包含 CAM(通道注意力模块)和 SAM(空间注意力模块)两个子模块(图 8)。CAM 模块通过最大池化和平均池化获取通道维度的向量,经 MLP 模块交流后相加,通过 sigmoid 激活函数得到通道注意力权重,见式(4)。

$$M_c F = \text{sigmoid}(MLP(AvgPool(F)) + MLP(MaxPool(F))). \quad (4)$$

SAM 模块则使用池化压缩通道维度,拼接特征图层后卷积融合,再经 sigmoid 激活得到空间注意力权重,见式(5)。

$$M_s F = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])). \quad (5)$$

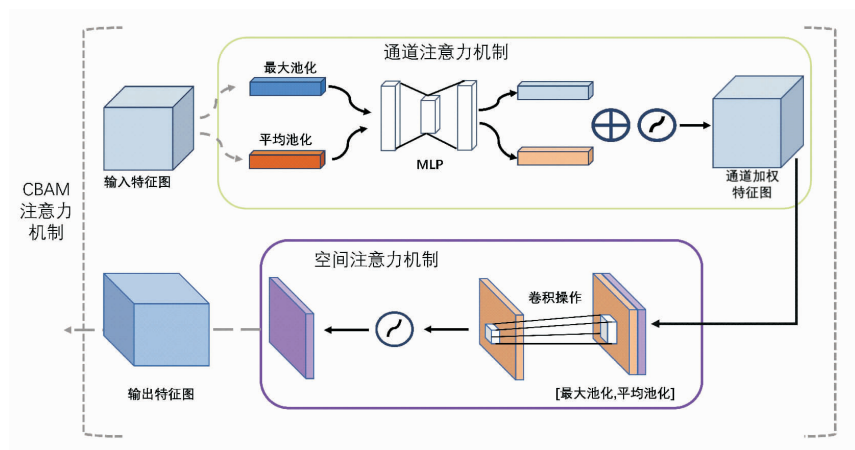


图 8 CBAM 注意力机制

Fig. 8 CBAM attention mechanism

2.4 精度评价

将 DeepLabv3 + 改进模型与 DeepLabv3 +、HRNet 和 PSPNet 3 种代表性语义分割模型进行对比实验。其中, HRNet 是由微软亚洲研究院所提出的网络模型,它因多分支并行、多尺度融合和保持全程高分辨率等特点,已在语义分割、姿态跟踪等领域取得了优异的成绩^[38-39]。PSPNet 是基于 FCN 架构所设计的一种网络模型,因其引入更多的上下文信息、融合不同区域的语境信息等特点,能够很好地理解场景中的全局语境信息,已在众多语义分割数据集上展现出优异的性能^[40-41]。

验证数据为已制作的红树林部分人工标记数据集(划分出来的 560 对验证集,图像尺寸大小为 512 × 512),标记包括红树林与非红树林两类,其中,非红树林包括水体、城镇用地、陆地绿植等土地覆盖类型。共采用精确率(P)、召回率(R)、交并比(IoU)、F1-分数(F1-Score)和总体精度(OA)共 5 个指标进行验证结果的评价。其中, P 是预测正确的样本数占所有预测样本的比值; R 是指在所有正确样本数中正样本的占比; IoU 是指模型预测所得的样本数与真实标签之间的交叠率,即预测结果与真实标签之间的交集与并集之比;F1-Score 是同时兼顾了精确率和召回率的更为全面的评价指标; OA 表示分类正确样本数占所有样本的比值,包括了红树林与非红树林的正确样本。所有指标计算公式如下^[42]:

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

$$R = \frac{TP}{TP + FN}, \quad (7)$$

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

$$IoU = \frac{TP}{TP + FN + FP}, \quad (9)$$

$$F1 - Score = 2 \times \frac{TP}{TN + FP + 2TP}, \quad (10)$$

式(6)~式(10)中, TP 表示预测为红树林,实际也是红树林; FP 表示预测为红树林,实际不是红树林; TN 表示预测为非红树林,实际也是非红树林; FN 表示预测为非红树林,实际则是红树林。

此外,为探究 DeepLabv3 + 改进模型对于未曾见过全新遥感影像识别的可行性(机器学习模型泛化能力),剪裁湛江市高桥红树林保护区的高分二号遥感影像场景作为测试数据。首先,对影像场景进行类似 2.1 小节的尺寸大小处理;然后,将两块影像场景依次传入训练好的 DeepLabv3 + 改进模型,进行红树林识别并获取识别结果;由于同样对影像场景采用重叠裁切,接下来要

对识别结果进行再拼接,但只需取切片中间最有价值部分。这种方法既能有效避免卷积操作对边缘识别效果较差的问题,又能防止拼接时产生拼接痕迹。模型应用的精度评价指标包括生产者精度(%)、用户精度(%)和 Kappa 系数^[43]。

2.5 实验平台及超参数

基于 PyTorch 网络框架来搭建 DeepLabv3 + 深度学习模型,实验具体软硬件配置如表 1 所示。模型与改进后的模型都采用相同超参数设置,以确保改进算法的真实有效性。通过查阅文献及反复实验,最终确定使用 Adam 作为优化器,初始学习率设置为 4e-4,使用 Dice Loss 和 CE Loss 作为混合损失函数, batch-size 设置为 16,不采用冻结训练方式,总训练轮数设置为 200,保证模型能够充分拟合。在训练模型阶段,加载 Xception 主干网络的预训练模型权值,预训练权值是在其他数据集上训练得到的网络权值,公开在网络上,由于特征是相通的,故使用预训练权值能够在训练过程中减少训练时间及加速网络的收敛,避免因从零开始训练而导致的提取效果不明显。

表 1 软硬件配置

Table 1 Software and hardware configuration

软/硬件名称	配置详情
CPU	Intel Xeon Platinum 8255C
GPU	V100-SXM2-32GB
CUDA	Cuda 11.3
运行内存	40G
操作系统	Ubuntu 20.04
编程语言	Python 3.8
网络框架	PyTorch 1.10.0

3 试验结果与分析

3.1 精度评价指标比较

DeepLabv3 + 改进模型与 DeepLabv3 +、HRNet 和 PSPNet 模型的对比较验证结果如表 2。除 P (精确率)外,其余 4 个精度指标都显示 DeepLabv3 + 改进模型的性能更优。其中,① DeepLabv3 + 改进模型的 R (召回率)最大,达 96.05%,高于 DeepLabv3 + (95.72%)、HRNet (94.65%)或 PSPNet (92.66%)。② 红树林 IoU 能够体现出识别红树林该类型的准确性,其中,DeepLabv3 + 改进模型达到了 95.31% 的交并比,对比其他 3 类模型有显著提升,相比于原模型提升了 0.31%,比 HRNet 提升了 1.35%,同 PSPNet 相比提升最为明显,达到了

3.51%。③在红树林的 F1-Score 评价指标下,改进模型依旧取得了最高的评价,达到了 97.60%,比较原模型提升了 0.16%,同 HRNet 相比提高了 0.71%,对比 PSPNet 提升最为明显,达到 1.88%。实验证明通过改进卷积方式、添加多种注意力集中以及多尺度特征融合的方法能

够在分割时达到更好的效果。④DeepLabv3 + 改进模型在各个评价指标上都有一定的进步,其中,总体精度 OA 达到了 99.60%,同原模型 DeepLabv3 + 相比提高了 0.03%,相较于 HRNet 与 PSPNet 都有明显的提升,分别提升了 0.12% 和 0.31%。

表 2 4 种模型运用于红树林边界识别的精度

Table 2 Comparison of recognition mangrove boundary accuracy for four models

模型	P	R	Mangrove-IoU	Mangrove-F1-Score	OA
DeepLabv3 + 改进	99.20	96.05	95.31	97.60	99.60
DeepLabv3 +	99.22	95.72	95.00	97.44	99.57
HRNet	99.23	94.65	93.96	96.89	99.48
PSPNet	99.00	92.66	91.80	95.72	99.29

3.2 模型验证结果比较

随机抽取验证数据集中的样本,根据明暗、形状、红树林斑块面积大小及破碎度等因素筛选出代表样本,图 9 为 4 种模型下红树林边界的识别结果。验证集数据未经过网络模型学习,将其用于模型红树林边界识别验证,其结果能够代表网络模型性能的真实水平。可以看出,DeepLabv3 + 改进模型对红树林边界识别的整体效果最好,特别是在细条状目标对象(图 9b),以及破碎程度较高的场景(图 9d)。其次是 DeepLabv3 + 模型,但对

一些边缘(图 9b)及破碎的斑块(图 9d)识别效果较差,这是因为 DeepLabv3 + 模型未充分利用原始场景图像内目标的空间信息,不能联系更多的上下文信息。HRNet 模型在面对较大且连续的目标物时,也能够准确地识别且保留边缘细节,但面对小目标物或细小斑块时,容易出现目标缺失现象(图 9a 和图 9b)。PSPNet 模型性能表现最差,丢失了绝大部分小目标(图 9a ~ 图 9c),且较大目标物的边缘细节被抹去,呈现光滑表现(图 9b 和图 9c)。

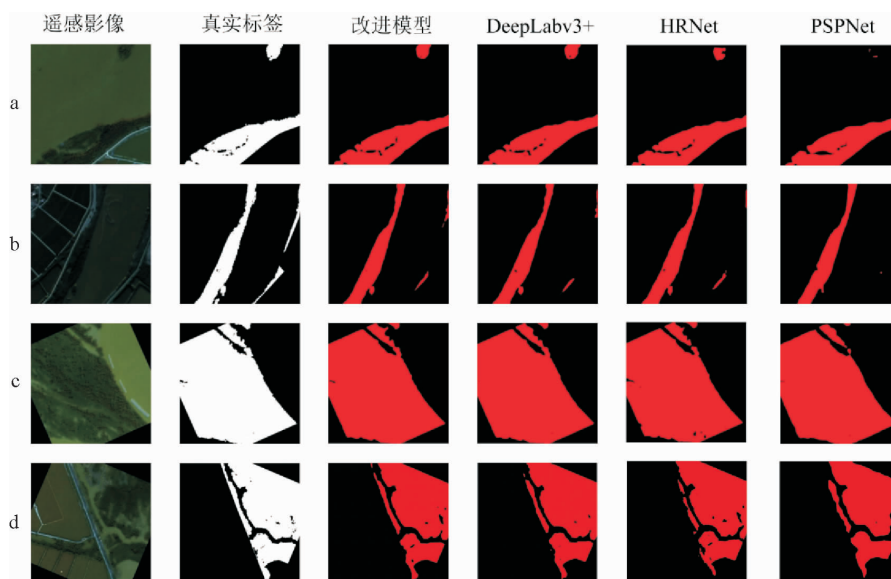


图 9 几种模型运用于红树林边界提取的结果对比

Fig. 9 Comparison results of several models used for mangrove boundary extraction

注:图中,纵列从左至右依次为遥感影像、真实标签、DeepLabv3 + 改进模型、DeepLabv3 +、HRNet 和 PSPNet;横行 a ~ d 依次为选取的 4 组样本数据。

可见,在改进卷积方法、添加多种注意力和进行多尺度空间特征融合后,DeepLabv3 + 改进模型的应用精度得以提升,实际遥感影像场景应用也能保证图像分割

的精度和细节处理的有效性。

3.3 消融实验结果比较

为验证改进方案的有效性,采用控制变量法进行消

融实验,使用 OA、Mangrove-IoU 和 Mangrove-F1-Score 作为评价指标。设计了 4 种实验方案:方案一是未改进的原始 DeepLabv3 + 模型;方案二改进 ASPP 模块,使用深度可分离卷积和 SE 注意力机制;方案三结合 CBAM 注

意力机制和多尺度特征融合;方案四是综合改进 ASPP 与多尺度特征融合的 DeepLabv3 + 改进模型。具体精度评估结果如表 3 所示。

表 3 消融实验结果对比表

Table 3 Comparison table of ablation experiments

%

顺序	模型改进方案 内容	评价指标		
		OA	Mangrove-IoU	Mangrove-F1-Score
方案一	未改进,即原始 DeepLabv3 + 模型	99.57	95.00	97.44
方案二	改进 ASPP 模块	99.59	95.29	97.59
方案三	CBAM 注意力机制及多尺度特征融合	99.58	95.13	97.51
方案四	改进 ASPP 与多尺度特征融合结合,即 DeepLabv3 + 改进模型	99.60	95.31	97.60

消融实验表明,DeepLabv3 + 模型的各改进方案均提升了性能。对比 ASPP 模块修改与 CBAM 注意力机制及多尺度特征融合,前者对 OA、Mangrove-IoU 及 Mangrove-F1-Score 提升最大,因深度可分离卷积与 SE 注意力机制结合增强了通道重要性,提高了高层语义信息精度。后者虽也有提升,但相对较小,主要实现空间位置信息恢复。最终改进模型结合两者,效果最突出,验证了 DeepLabv3 + 改进模型的有效性。

3.4 DeepLabv3 + 改进模型的应用分析

初步测试验证与指标评估都显示 DeepLabv3 + 改进模型在红树林边界识别的性能优越。为进一步验证其

应用的可行性与性能,选择了湛江市高桥红树林保护区的高分二号遥感影像作为新数据,如图 10a 和图 10b,两块遥感影像场景的面积分别为 9.5 km² 和 3.3 km²。首先,通过人机交互解译获取红树林边界作为真实标签,然后,将影像传入训练好的模型进行识别,最后,将识别结果与真实标签叠置运算,得到差分图像。结果表明,模型在测试影像上的识别结果虽有部分错分、漏分,但主要出现在红树林边界,这是因为模型在进行卷积操作时往往对识别对象边界进行了“模糊处理”,从而导致出现差分图像现状,但总体上还原了红树林真实边界,区分了大范围内的红树林与非红树林。

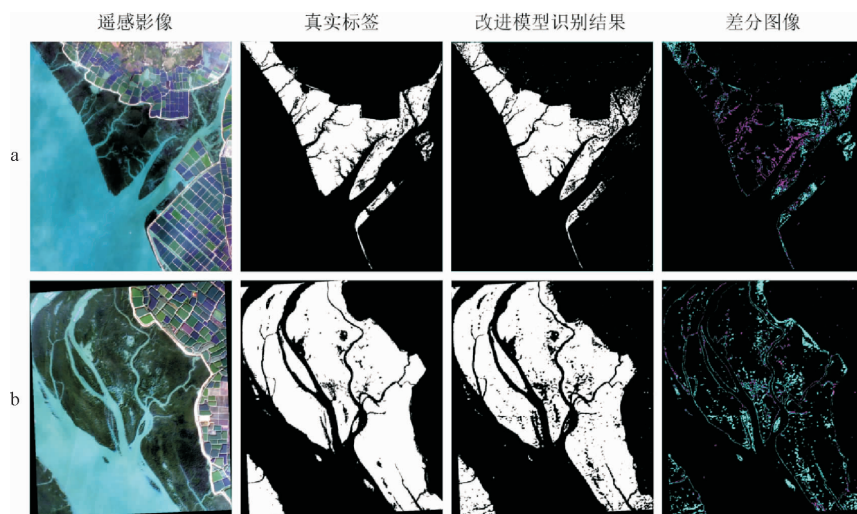


图 10 训练后的 DeepLabv3 + 改进模型在全新数据应用及结果

Fig. 10 Improved DeepLabv3 + model after training in new data applications and results

基于 DeepLabv3 + 改进模型识别结果和真实标签进行混淆矩阵计算,结果如表 4 和表 5。基于改进模型对遥感场景 a、b 进行红树林识别与边界提取的总体精度均大于 95%,Kappa 系数均大于 0.85,生产者精度与用

户精度都具有很高的精度,表明 DeepLabv3 + 改进模型在实际应用方面已经具备能够区分红树林与非红树林区域的能力,显示模型一定的泛化能力。

表4 混淆矩阵
Table 4 Confusion matrix

混淆矩阵		标签类型					
		遥感影像场景 a			遥感影像场景 b		
		红树林	非红树林	总计	红树林	非红树林	总计
预测类型	红树林	1 552 482	100 910	1 653 392	1 352 973	18 374	1 371 347
	非红树林	247 381	7 579 337	7 826 718	142 397	1 745 046	1 887 443
	总计	1 799 863	7 680 247	9 480 110	1 495 370	1 763 420	3 258 790

表5 精度评价指标
Table 5 Accuracy indices

类型	精度							
	遥感影像场景 a				遥感影像场景 b			
	生产者 精度/%	用户 精度/%	Kappa 系数	总体 精度/%	生产者 精度/%	用户 精度/%	Kappa 系数	总体 精度/%
红树林	86.26	93.90	0.876 7	96.326 1	90.48	98.66	0.900 0	95.066 5
非红树林	98.69	96.84			98.96	92.46		

4 结论与讨论

本文改进 ASPP 模块和 CBAM 注意力机制及多尺度特征融合的方法,获得一种基于 DeepLabv3 + 改进模型的红树林识别与边界提取方法;通过对原模型、广泛运用的 HRNet 和 PSPNet 模型进行控制变量对照实验,结果证明改进模型在 R、OA、Mangrove-IoU 和 Mangrove-F1-Score 4 个指标下都取得了最优成绩,其值分别达到 96.05%、99.60%、95.31% 和 97.60%;对验证集随机抽取样本进行模型验证实验,改进模型整体表现最为优异,能够识别出最多的破碎斑块及细长条目标物,也能保持目标物边缘细节。在具体应用方面,将 DeepLabv3 + 改进模型应用于湛江市高桥红树林,提取的红树林边

界结果证明改进模型已具备识别红树林与非红树林的能力。改进模型绩效的提升能够为红树林的监测、修复、管理和生态保护等提供技术支持。

尽管 DeepLabv3 + 模型得以改进,但依旧存在少许漏分、错分的情况,说明模型在红树林边界的识别与提取精度上仍有提升空间。未来可从以下 4 个方面探索提升红树林识别与提取的精度:①制作红树林的真实标准数据集并保证足以训练的样本数量;②探索海水淹没红树林边界的识别与提取精度将是未来模型改进的重点;③可尝试利用多光谱遥感数据更为丰富的影像信息;④融合更高分辨率无人机影像。不仅如此,通过无人机搭载的高光谱成像系统,获取兼顾高空间分辨率和高光谱的影像,可能会实现红树林群落或种类识别与边界提取。

参考文献:

- [1] 林鹏. 中国红树林湿地与生态工程的几个问题[J]. 中国工程科学, 2003, 5(6): 33-38.
- [2] 贾明明. 1973—2013 年中国红树林动态变化遥感分析[D]. 长春: 中国科学院研究生院(东北地理与农业生态研究所), 2014.
- [3] 林益明, 林鹏. 中国红树林生态系统的植物种类、多样性、功能及其保护[J]. 海洋湖沼通报, 2001(3): 8-16.
- [4] 张乔民, 隋淑珍. 中国红树林湿地资源及其保护[J]. 自然资源学报, 2001, 16(1): 28-36.
- [5] 黎夏, 刘凯, 王树功. 珠江口红树林湿地演变的遥感分析[J]. 地理学报, 2006, 61(1): 26-34.
- [6] 郭伟, 朱大奎. 深圳围海造地对海洋环境影响的分析[J]. 南京大学学报(自然科学版), 2005, 41(3): 286-296.
- [7] 孟宪民. 湿地与全球环境变化[J]. 地理科学, 1999, 19(5): 385-391.
- [8] Duke N C, Meynecke J O, Dittmann S, et al. A world without mangroves? [J]. Science, 2007, 317(5834): 41-42.
- [9] 马明国, 王建, 王雪梅. 基于遥感的植被年际变化及其与气候关系研究进展[J]. 遥感学报, 2006, 10(3): 421-431.
- [10] 张树文, 颜凤芹, 于灵雪, 等. 湿地遥感研究进展[J]. 地理科学, 2013, 33(11): 1406-1412.
- [11] 赵英时. 遥感应用分析原理与方法[M]. 北京: 科学出版社, 2003.
- [12] 张继贤, 顾海燕, 杨懿, 等. 高分辨率遥感影像智能解译研究进展与趋势[J]. 遥感学报, 2021, 25(11): 2198-2210.

- [13] 杨桃, 刘湘南. 遥感影像解译的研究现状和发展趋势[J]. 国土资源遥感, 2004, 16(2): 7-10, 15.
- [14] 赵春霞, 钱乐祥. 遥感影像监督分类与非监督分类的比较[J]. 河南大学学报(自然科学版), 2004, 34(3): 90-93.
- [15] 潘建刚, 赵文吉, 宫辉力. 遥感图像分类方法的研究[J]. 首都师范大学学报(自然科学版), 2004, 25(3): 86-91, 80.
- [16] 郑远攀, 李广阳, 李晔. 深度学习在图像识别中的应用研究综述[J]. 计算机工程与应用, 2019, 55(12): 20-36.
- [17] 周培诚, 程焱, 姚西文, 等. 高分辨率遥感影像解译中的机器学习范式[J]. 遥感学报, 2021, 25(1): 182-197.
- [18] 张康, 黑保琴, 周壮, 等. 变异系数降维的 CNN 高光谱遥感图像分类[J]. 遥感学报, 2018, 22(1): 87-96.
- [19] 曲景影, 孙显, 高鑫. 基于 CNN 模型的高分辨率遥感图像目标识别[J]. 国外电子测量技术, 2016, 35(8): 45-50.
- [20] 唐璐, 万良, 王婷婷, 等. DECANet: 基于改进 DeepLabv3 + 的图像语义分割方法[J]. 激光与光电子学进展, 2023, 60(4): 92-100.
- [21] 赵为平, 陈雨, 项松, 等. 基于改进的 DeepLabv3 + 图像语义分割算法研究[J]. 系统仿真学报, 2023, 35(11): 2333-2344.
- [22] 杨加志, 胡喻华, 罗勇, 等. 广东省红树林分布现状与动态变化研究[J]. 林业与环境科学, 2018, 34(5): 24-27.
- [23] 毛丽君. 基于遥感的广东湛江红树林湿地动态变化研究[D]. 南京: 南京林业大学, 2011.
- [24] 唐秋霞, 王友绍. 雷州半岛红树林群落特征及其分布格局[J]. 生态科学, 2021, 40(5): 23-32.
- [25] 王芳, 杨武年, 邓晓宇, 等. 高分二号数据的城市生态用地分类方法探讨[J]. 测绘科学, 2018, 43(3): 71-76.
- [26] 宋明辉. 基于高分二号数据的面向对象城市土地利用分类研究[J]. 遥感技术与应用, 2019, 34(3): 547-552, 629.
- [27] Chen L C, Zhu Y K, Papandreou G, et al. European Conference on Computer Vision, October 06, 2018 [C]. Cham: Springer, 2018.
- [28] Chollet F. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017 [C]. Piscataway: IEEE, 2017.
- [29] 刘文祥, 舒远仲, 唐小敏, 等. 采用双注意力机制 Deeplabv3 + 算法的遥感影像语义分割[J]. 热带地理, 2020, 40(2): 303-313.
- [30] 刘文雅, 岳安志, 季珏, 等. 基于 DeepLabv3 + 语义分割模型的 GF-2 影像城市绿地提取[J]. 国土资源遥感, 2020, 32(2): 120-129.
- [31] 邵伟平, 王兴, 曹昭睿, 等. 基于 MobileNet 与 YOLOv3 的轻量化卷积神经网络设计[J]. 计算机应用, 2020, 40(S1): 8-13.
- [32] Howard A G, Zhu M L, Chen B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [EB/OL]. (2017-04-17) [2023-10-21]. <http://arXiv.org/abs/1704.04861>.
- [33] Mnih V, Heess N, Graves A, et al. Recurrent models of visual attention [EB/OL]. (2014-06-24) [2023-10-21]. <http://arXiv.org/abs/1406.6247>.
- [34] 徐沁, 梁玉莲, 王冬越, 等. 基于 SE-Res2Net 与多尺度空谱融合注意力机制的高光谱图像分类[J]. 计算机辅助设计与图形学学报, 2021, 33(11): 1726-1734.
- [35] Hu J, Shen L, Sun G, et al. 2018 IEEE Transactions on Pattern Analysis and Machine Intelligence, June 18-23, 2018 [C]. Piscataway: IEEE, 2018.
- [36] 李永上, 马荣贵, 张美月. 改进 YOLOv5s + DeepSORT 的监控视频车流量统计[J]. 计算机工程与应用, 2022, 58(5): 271-279.
- [37] Woo S, Park J, Lee J Y, et al. Proceedings of the European Conference on Computer Vision (ECCV), September 08-14, 2018 [C]. Cham: Springer, 2018.
- [38] Sun K, Xiao B, Liu D, et al. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019 [C]. Piscataway: IEEE, 2019.
- [39] 赵紫旋, 吴谨, 朱磊. 基于 GLNet 和 HRNet 的高分辨率遥感影像语义分割[J]. 红外技术, 2021, 43(5): 437-442.
- [40] Zhao H S, Shi J P, Qi X J, et al. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017 [C]. Piscataway: IEEE, 2017.
- [41] 刘钊, 廖斐凡, 赵桐. 基于 PSPNet 的遥感影像城市建成区提取及其优化方法[J]. 国土资源遥感, 2020, 32(4): 84-89.
- [42] 范吉延, 柯长青, 姚国慧, 等. 基于深度学习的全极化 SAR 影像冰川边界识别[J]. 遥感学报, 2023, 27(9): 2098-2113.
- [43] 韩文霆, 郭聪聪, 张立元, 等. 基于无人机遥感的灌区土地利用与覆被分类方法[J]. 农业机械学报, 2016, 47(11): 270-277.