

基于改进 DeepLabV3+ 的石漠化地区 裸岩信息提取

吴永俊¹ 汪泓^{2,*} 杨晨²

(1 黔西南州自然资源管理服务中心, 兴义 562400)

(2 贵州大学矿业学院, 贵阳 550025)

摘要 针对传统喀斯特地区裸岩提取方法成本高、精度低的问题, 文章构建了一种基于改进 DeepLabV3+ 的裸岩提取方法。该方法首先在编码器中用 CA-DC-MobileNetV3 替换 DeepLabV3+ 骨干网络 Xception 进行特征提取, 很大程度上减少了模型的参数量; 其次, 将编码器提取的特征通过特征金字塔网络和坐标注意力机制进行加强特征提取, 以获取更多小目标信息并减少图像细节损失; 最后在空洞空间金字塔池化模块将不同空洞率的卷积层进行特征融合, 提高信息的利用率。研究表明: 文章方法在不同场景裸岩提取任务中表现最好, 模型参数量约为 DeepLabV3+ 的 1/13, 交并比、 F_1 分数分别为 72.46%、84.03%, 上述 2 个指标相比于 DeepLabV3+ 模型分别提高了 4.62 和 3.19 个百分点, 并优于其余常用语义分割模型, 提高了裸岩提取精度。

关键词 裸岩提取 深度学习 语义分割 坐标注意力机制

中图分类号: P237

文献标志码: A

文章编号: 1009-8518(2024)01-0123-13

DOI: 10.3969/j.issn.1009-8518.2024.01.011

Extraction of Bare Rock Information in Rocky Desertification Area Based on Improved DeepLabV3+

WU Yongjun¹ WANG Hong^{2,*} YANG Chen²

(1 Qianxi'nan Prefecture Natural Resource Management Service Center, Xingyi 562400, China)

(2 College of Mining, Guizhou University, Guiyang 550025, China)

Abstract Aiming at the problems of high cost and low precision of traditional bare rock extraction methods in karst areas, this paper constructs a bare rock extraction method based on improved DeepLabV3+. This method first uses CA-DC-MobileNetV3 to replace DeepLabV3+ backbone network Xception in the encoder for feature extraction, which greatly reduces the amount of model parameters. Secondly, the features extracted by the encoder are enhanced through the feature pyramid network and the coordinate attention mechanism to obtain more small target information and reduce the loss of image details. Finally, in the atrous spatial pyramid pooling module, the features of the convolutional layers with different dilation rates are fused to improve the utilization of information. The results show that the method in this paper performs best in the bare rock extraction tasks in

收稿日期: 2023-03-20

基金项目: 国家自然科学基金 (41901225)

引用格式: 吴永俊, 汪泓, 杨晨. 基于改进 DeepLabV3+ 的石漠化地区裸岩信息提取[J]. 航天返回与遥感, 2024, 45(1): 123-135.

WU Yongjun, WANG Hong, YANG Chen. Extraction of Bare Rock Information in Rocky Desertification Area Based on Improved DeepLabV3+[J]. Spacecraft Recovery & Remote Sensing, 2024, 45(1): 123-135. (in Chinese)

different scenarios, the number of model parameters is about 1/13 of that of DeepLabV3+, and the intersection ratio and F_1 -Score are 72.46% and 84.04% respectively. Compared with the DeepLabV3+ model, the above two indicators have improved by 4.62 and 3.19 percentage points, respectively, and are superior to other commonly used semantic segmentation models, improving the accuracy of bare rock extraction.

Keywords bare rock extraction; deep learning; semantic segmentation; coordinate attention mechanism

0 引言

石漠化是我国西南地区最严重的生态环境问题之一,与沙漠化、水土流失并称为我国的三大生态灾害^[1-2]。石漠化是指在喀斯特地区,不合理的人类活动,自然植被遭到破坏,造成土壤流失、基岩裸露等类似荒漠景观的土地退化过程。喀斯特石漠化严重的影响着区域生态环境,会导致水土流失、自然灾害频发、生物多样性下降等生态危害^[3];其带来的危害已成为制约我国西南地区可持续发展的重要因素。因此,石漠化的治理对我国西南地区改善生态环境、保障经济社会可持续发展具有重要战略意义。而石漠化调查又是石漠化治理过程中必不可少的工作环节^[4];在我国现行岩溶地区水土流失综合治理技术标准^[5]中,将基岩裸露率作为石漠化等级划分的直接指标。因此,准确地提取喀斯特地区的岩石裸露信息对石漠化调查、治理工作具有重要意义。

传统的岩石裸露信息主要通过目测和人工机械布点来获取,对人力、物力造成了较大的损耗,且当区域地物细碎化程度较高时,测量误差较大^[6]。随着遥感技术的发展,学者们通过遥感影像开展了诸多石漠化地区裸岩信息的提取工作。常用的方法有指数法、人机交互解译法、面向对象法等^[7]。文献[8]分析 Landsat8 裸岩和其他土地覆盖类型的光谱特征,通过蓝光、近红外波段差和比,构建归一化裸岩指数(Carbonate Rock Indice, CRI)对裸岩信息的提取,但遥感影像易受气候等因素的影响,且受影像空间分辨率的制约,提取精度往往较低;文献[9]通过人工勾绘获取无人机影像裸岩率,并用其矫正相同研究区 Landsat8 影像,通过指数法获取的裸岩率,使 Landsat8 影像提取裸岩的精度得到提高,但该方法工作量大,效率低;文献[10]基于无人机影像,通过面向对象的方法对裸岩信息进行提取,提取精度进一步提高,但喀斯特地区地表覆盖极其不规则,地物通常呈斑块状,破碎度较高,空间分布不均匀,岩石之间大小形状不尽相同,分割的尺度和参数较难确定,需要反复调整相关阈值参数,人为划定的主观性强。

近年来,深度学习在图像分割、目标检测等领域得到了广泛的应用。文献[11]基于“简译”软件与无人机影像实现全卷积神经网络对石漠化地区裸岩信息的提取,提取精度优于传统提取方法,但其方法封装于软件中,提取精度仍有提升空间;文献[12]将改进的残差块作为 SegNet 网络^[13]的编码器,解决了网络退化问题,并引入深度可分离卷积来减少网络参数量,用膨胀卷积来扩大感受野、提高网络特征提取能力,提高了水体提取精度;文献[14]在 U-Net 网络^[15]跳跃连接层中嵌入密集连接块和空间注意力,使得边界块引入底层空间信息,解决了提取建筑物模糊边界的问题;文献[16]将自导向注意整合到特征金字塔(Feature Pyramid Networks, FPN)^[17]结构中,利用 FPN 获取不同网络层的多尺度特征,并将其集成到 DeepLabV3+^[18]、U-Net、PANet^[19]中来提取地表裂缝信息,结果表明该模块的引入增强了纹理的细节并抑制了噪声;文献[20]将垂直和空间注意力机制分别嵌入 DeepLabV3+骨干特征提取网络和空洞空间金字塔池化(Atrous Spatial Pyramid Pooling, ASPP)模块^[21]中,加强了网络获取上下文信息的能力,结果表明改进的 DeepLabV3+算法有效细化了特征细节,分割效果得到提升。上述研究中,学者们针对不同提取任务,在不同模型的基础上进行改进,使得提取精度得以提高,但是针对喀斯特石漠化地区裸岩信息进行提取的研究,却鲜有报道。

目前,DeepLabV3+在语义分割领域得到越来越多的关注,并在多个公开数据集上取得较好的成绩,DeepLabV3+的 ASPP 模块能获取更多的上下文信息和边缘信息;但其骨干网络结构复杂且参数量庞大,

带来了庞大的计算量，甚至可能降低模型的泛化能力。针对该缺点，同时为进一步探索深度学习在石漠化方面的应用，突破传统方法在复杂地物条件下裸岩提取精度低的问题，本文构建了一种基于改进 DeepLabV3+的裸岩信息提取方法。首先使用改进的轻量化网络替换 DeepLabV3+特征提取网络；其次在特征提取网络中结合 FPN、坐标注意力机制（Coordinate Attention, CA）模块^[22]提高模型对裸岩特征上下文的联系能力，使模型更关注细碎裸岩区域并获取浅层网络中更多的细节信息；最后在 ASPP 模块中添加异感受野特征融合特征层。

1 模型构建

1.1 DeepLabV3+模型

喀斯特地区地物破碎化严重，石漠化区域裸岩通常以斑块的形态分散在植被、裸土中，部分成片分布，因此其边界轮廓、形态大小极其不规律。为了尽可能地提取不同大小形态的裸岩目标，采用 DeepLabV3+模型进行裸岩提取。DeepLabV3+引入了编码-解码结构，编码部分包括骨干网络深度卷积神经网络（Deep Convolutional Neural Networks, DCNN）和 ASPP 两个部分。其中骨干网络为 Xception，ASPP 模块由 1×1 卷积，空洞率为 6、12、18 的空洞卷积以及全局平均池化并行而成，该结构在保证空间分辨率的情况下，扩大了网络的感受野，能获得不同尺度的石漠化裸岩特征信息，从而更好地对不同形态大小的裸岩目标进行特征提取。

输入的影像经骨干网络得到两个特征层，浅层特征直接传入解码器中，深层特征经 ASPP 获取多尺度特征后进行堆叠，再经 1×1 卷积调整通道数传入解码器中。解码器对传入的深层特征进行 4 倍上采样，并对浅层特征进行 1×1 卷积通道数调整，将调整后的两个特征堆叠，最后经 3×3 卷积细化特征与 4 倍上采样得到最终预测结果。整体网络结构如图 1 所示。

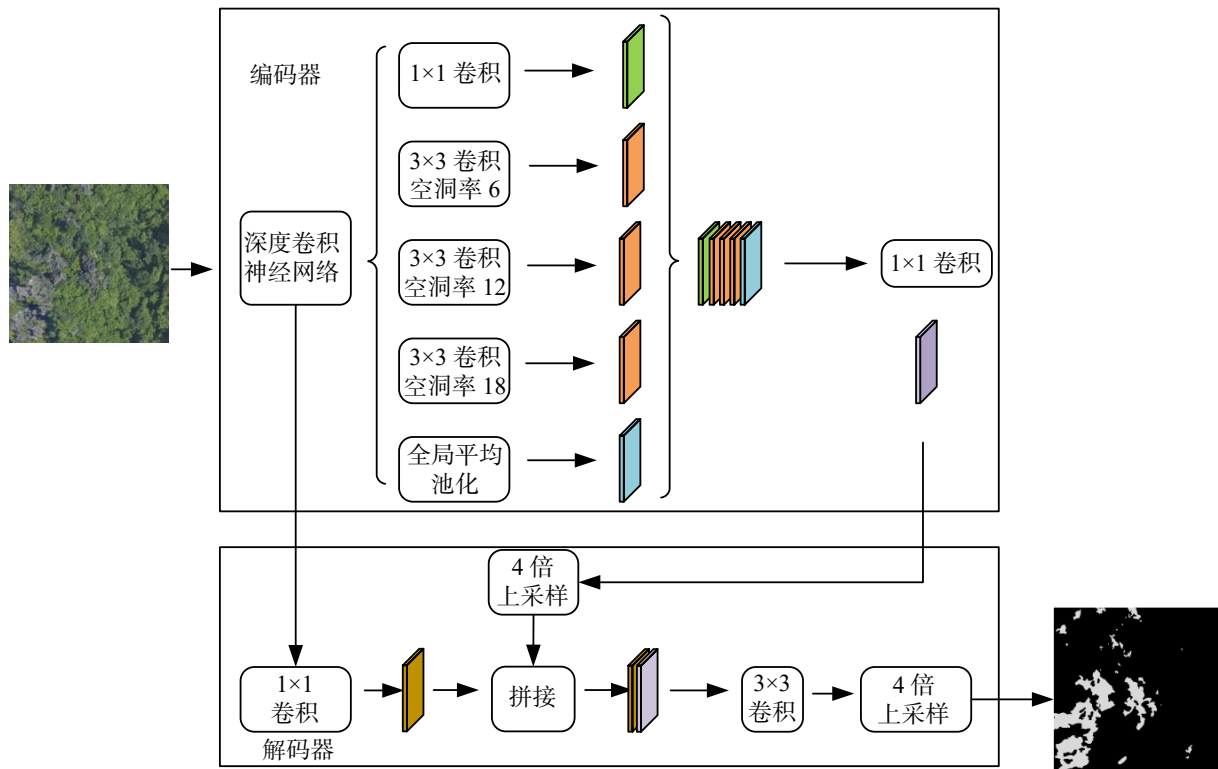


图 1 DeepLabV3+模型结构
Fig.1 DeepLabV3+ model structure

1.2 坐标注意力机制

注意力机制通过增强网络对目标特征的关注和忽略无关信息来改善网络性能，可以有效地提升网络特征提取的能力，被广泛应用于各种深度学习的任务中。将其引入本文的网络，使得模型提升对裸岩目标的关注度，忽略地物背景的干扰。CA将位置信息嵌入通道注意力中，使网络获取更大区域的信息。坐标注意力机制的实现主要有两个步骤：坐标信息的嵌入和坐标注意力生成。其模块结构如图2所示。

首先是坐标信息的嵌入。全局池化通常用于通道注意力机制中对空间信息进行编码，但其将全局信息压缩至一个通道内进行描述，会使位置信息难以保留。因此，该部分将全局池化层分解成两个一维特征编码操作，对于输入的特征图 $\mathbf{X} \in \mathbb{R}^{C \times H \times W}$ ， C 为特征图的通道数， H 为特征图的高， W 为特征图的宽，使用大小为 $(H, 1)$ 和 $(1, W)$ 的池化核沿水平方向和垂直方向坐标进行通道编码，使模块能获取长范围依赖的信息，其表达式如式(1)、(2)所示：

$$\mathbf{Z}_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} \mathbf{x}_c(h, i) \quad (1)$$

$$\mathbf{Z}_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} \mathbf{x}_c(j, w) \quad (2)$$

式中 c 、 h 、 w 分别表示当前注意力模块所对应特征图的通道、高度和宽度； $\mathbf{x}_c \in \mathbb{R}^{H \times W}$ 表示输入的特征； $\mathbf{Z}_c^h \in \mathbb{R}^{H \times 1}$ 表示第 c 通道在高 h 处的输出； $\mathbf{Z}_c^w \in \mathbb{R}^{1 \times W}$ 表示第 c 通道在宽度 w 处的输出。

其次是坐标注意力生成。该部分将具有编码信息的特征进行拼接，并使用卷积变换和非线性激活获得特征映射 \mathbf{f} ，其表达式如(3)所示

$$\mathbf{f} = \delta\{F_{1 \times 1}[\mathbf{Z}^h, \mathbf{Z}^w]\} \quad (3)$$

式中 δ 为非线性激活函数； $F_{1 \times 1}$ 为 1×1 卷积。

在得到特征图 \mathbf{f} 后沿空间方向将其分解为两个张量，并分别对两个张量进行卷积变换和 Sigmoid 激活，将通道数调整至与输入特征一致，得到张量 \mathbf{g}^h 、 \mathbf{g}^w

$$\mathbf{g}^h = \sigma(F_h(\mathbf{f}^h)) \quad (4)$$

$$\mathbf{g}^w = \sigma(F_w(\mathbf{f}^w)) \quad (5)$$

式中 F_h 、 F_w 为 1×1 卷积； \mathbf{f}^h 、 \mathbf{f}^w 为 \mathbf{f} 沿着空间维度分解的两个张量； σ 为 Sigmoid 激活函数。

最后对 \mathbf{g}^h 、 \mathbf{g}^w 进行拓展，作为注意力权重，最终得到坐标注意力机制输出 $\mathbf{y}_c(i, j)$ ，其表达式如式(6)

$$\mathbf{y}_c(i, j) = \mathbf{x}_c(i, j) \times \mathbf{g}_c^h(i) \times \mathbf{g}_c^w(j) \quad (6)$$

1.3 CA-DC-MobileNetV3 模型

考虑到石漠化区域裸岩提取仅涉及裸岩和背景两个标签，语义信息相对简单，原始 DeepLabV3+ 主干网络 Xception 结构复杂且参数量大，对算力要求高，且当数据集的量级与参数量不匹配时，网络的泛化效果较差^[23]，因此选择轻量化网络 MobileNetV3 作为骨干网络。MobileNetV3 于 2019 年提出，其在 MobileNetV2 的基础上在线性瓶颈的逆残差结构块 (The Inverted Residual with Linear Bottleneck Block,

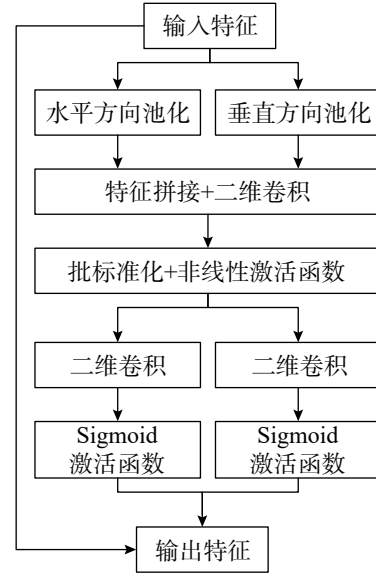


图2 坐标注意力机制

Fig.2 Coordinate attention mechanism

Bneck) 引入基于压缩奖惩 (Squeeze and Excitation, SE) 的轻量级注意力机制结构和 Hard-Swish 激活函数。MobileNetV3 的基本单元 Bneck 如图 3 所示, 输入特征经 1×1 卷积进行通道数调整, 再使用 3×3 的深度可分离卷积进行特征提取, 对提取的结果进行全局平均池化并施加 SE 注意力机制, 经 Relu 激活函数和 Hard- σ 激活函数加入非线性因素后与特征图相乘, 最后经 1×1 卷积进行通道数调整并与输入特征相加完成残差边模块^[24]。

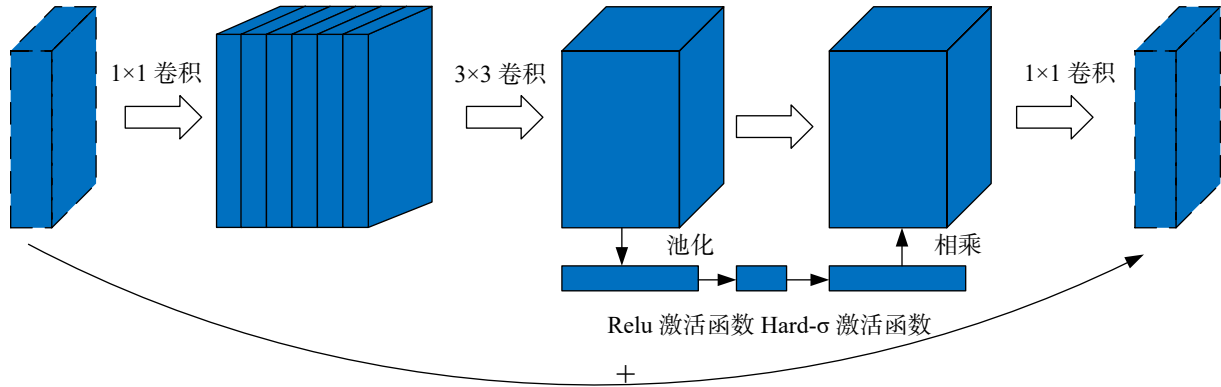


图 3 MobileNetV3 基本单元
Fig.3 MobileNetV3 block

其整体结构如表 1 所示, 表中 3×512^2 表示输入特征为 3 通道, 大小为 512 像素 \times 512 像素; pool, 7×7 表示池化及对应的池化核大小; Con2d, 1×1 , NBN 表示卷积操作、对应的卷积核大小、不使用批归一化层 (Batch Normalization, BN) 结构; YES 表示施加相应的注意力机制; HS、RE 分别表示使用的激活函数为 Hard-Swish、Relu 激活函数; 步距为 1 表示卷积过程中的步长为 1。原始 MobileNetV3 共进行 5 次下采样; 为深层特征进入 ASPP 时保留更多的像素点特征, 本文对原始输入特征进行 3 次下采样, 选

表 1 MobileNetV3 网络结构
Tab.1 MobileNetV3 network structure

特征层	输入特征	操作	SE	激活函数	步距
1	3×512^2	Conv2d	—	HS	2
2	16×256^2	Bneck, 3×3	—	RE	1
	16×256^2	Bneck, 3×3	—	RE	2
3	24×128^2	Bneck, 3×3	—	RE	1
	24×128^2	Bneck, 5×5	YES	RE	2
4	40×64^2	Bneck, 5×5	YES	RE	1
	40×64^2	Bneck, 5×5	YES	RE	1
	40×64^2	Bneck, 3×3	—	HS	2
5	80×32^2	Bneck, 3×3	—	HS	1
	80×32^2	Bneck, 3×3	—	HS	1
	80×32^2	Bneck, 3×3	—	HS	1
	80×32^2	Bneck, 3×3	YES	HS	1
6	112×32^2	Bneck, 3×3	YES	HS	1
	112×32^2	Bneck, 5×5	YES	HS	2
7	160×16^2	Bneck, 5×5	YES	HS	1
	160×16^2	Bneck, 5×5	YES	HS	1
	160×16^2	Con2d, 1×1	—	HS	1
8	960×16^2	pool, 7×7	—	—	1
	960×1^2	Con2d, 1×1 , NBN	—	HS	1
9	$1\ 280 \times 1^2$	Con2d, 1×1 , NBN	—	—	1

用 MobileNetV3 的前 7 层，同时将第 4 层之后的卷积修改为空洞率为 2 的深度可分离空洞卷积（Dilated Convolutions, DC），以扩大感受野；原始网络中的 SE 注意力只考虑内部通道信息，忽略了位置信息的重要性，而视觉中目标的空间结构是很重要的，因此使用 CA 代替。改进后的 MobileNetV3 网络结构如表 2 所示。

表 2 CA-DC-MobileNetV3 网络结构
Tab.2 CA-DC-MobileNetV3 network structure

特征层	输入	操作	CA	激活函数	步距	空洞率
1	3×512^2	Conv2d	—	HS	2	1
2	16×256^2	Bneck, 3×3	—	RE	1	1
	16×256^2	Bneck, 3×3	—	RE	2	1
3	24×128^2	Bneck, 3×3	—	RE	1	1
	24×128^2	Bneck, 5×5	YES	RE	2	1
4	40×64^2	Bneck, 5×5	YES	RE	1	1
	40×64^2	Bneck, 5×5	YES	RE	1	1
	40×64^2	Bneck, 3×3	—	HS	1	1
5	80×64^2	Bneck, 3×3	—	HS	1	2
	80×64^2	Bneck, 3×3	—	HS	1	2
	80×64^2	Bneck, 3×3	—	HS	1	2
	80×64^2	Bneck, 3×3	YES	HS	1	2
6	112×64^2	Bneck, 3×3	YES	HS	1	2
	112×64^2	Bneck, 5×5	YES	HS	1	2
7	160×64^2	Bneck, 5×5	YES	HS	1	2
	160×64^2	Bneck, 5×5	YES	HS	1	2
	160×64^2	Con2d, 1×1	—	HS	1	1

1.4 特征金字塔

考虑到喀斯特地区地物分布较为细碎，石漠化区域裸岩目标尺度不尽相同，为进一步加强网络对不同尺度目标提取的能力，获取浅层特征所包含的局部精细特征，将 FPN 引入本文网络。FPN 通过自深而浅、自浅而深、横向连接网络将深层高语义特征与浅层特征融合，以获取不同尺度语义信息，得到更丰富的空间细节，达到提高分割精度的目的。本文涉及 FPN 结构如图 4 所示，其中特征层 3、特征层 4、

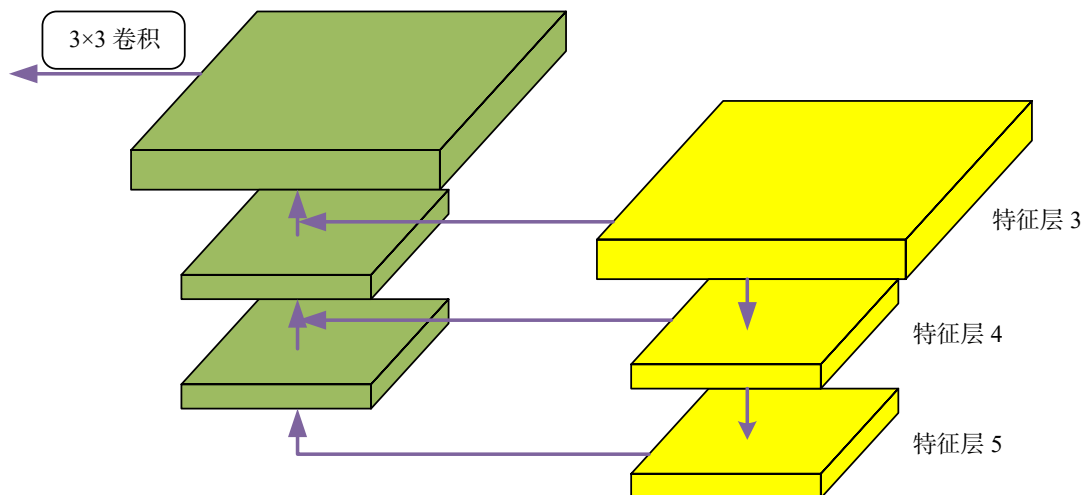


图 4 特征金字塔

Fig.4 Feature pyramid network structure

特征层 5 分别为 CA-DC-MobileNetV3 的特征层 3、特征层 4、特征层 5 输出的结果。FPN 通过上采样及通道数调整将不同尺度特征进行融合，最后经 3×3 卷积以消除上采样的混叠效应，作为浅层特征传入解码器中。

1.5 改进的 DeepLabV3+模型

图 5 为改进的 DeepLabV3+模型，使用 CA-DC-MobileNetV3 作为骨干网络，以减少模型参数；引入 FPN 模块进行加强特征提取，将深层高语义特征与浅层特征融合，获取更多细节特征，并通过 CA 对特征进行加强提取获取关键信息；为增加不同空洞卷积层之间的联系，提高信息利用率，在 ASPP 结构的基础上，将空洞率为 6、12、18 的空洞卷积层通过 Add 操作进行异感受野特征融合，并作为新的特征层与其余特征进行堆叠。

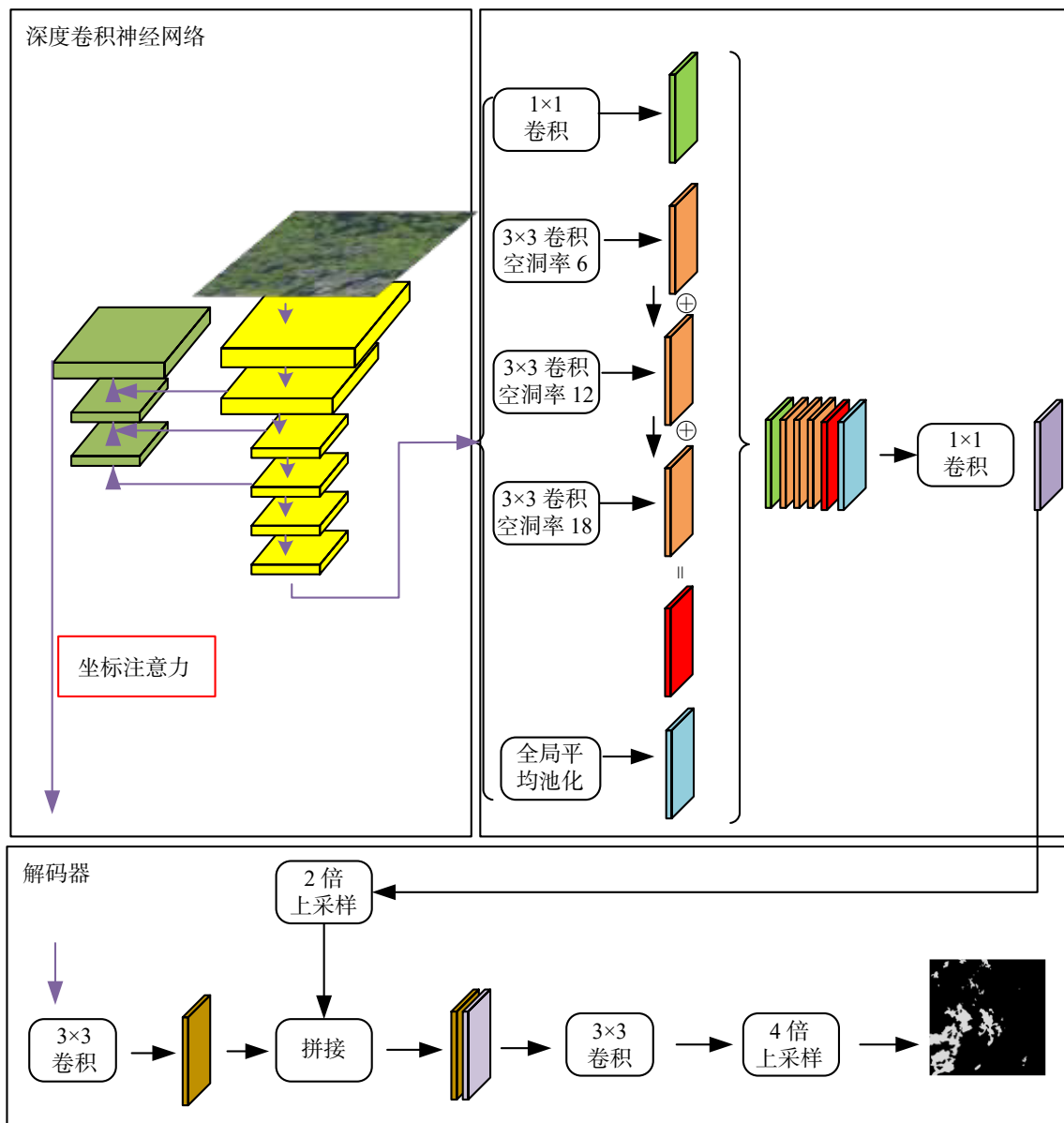


图 5 改进的 DeepLabV3+模型
Fig.5 Improved DeepLabV3+ model

2 实验与分析

2.1 实验数据与环境

实验数据来源于罗宾逊 R44 直升机搭载的哈苏 A6D-100C 航摄影采集得到的影像，区域位于盘江流域，流域内岩溶发育广泛，水土流失严重，喀斯特集中分布。所获取影像有红、绿、蓝三个波段，空间分辨率为 0.1 m。选取其中一景经预处理的影像，通过 ArcGIS 进行人工标注制作标签。将制作好的标签与原始影像进行同步处理，通过滑动窗口裁剪为 512 像素×512 像素的影像，使用水平翻转、旋转、镜像操作进行数据增广，最后得到 2 000 张影像，按照 9 : 1 的比例随机划分训练集和验证集。

实验采用的操作系统为 Ubuntu18.04，深度学习框架为 PyTorch，使用的 GPU 为 NVIDIA RTX 2080 Ti。训练超参数设置如下：迭代轮次（Epoch）为 80，批大小（Batch Size）为 4，动量为 0.9，使用 Adam（Adaptive Moment Estimation）优化器，初始学习率为 0.000 5，并通过余弦退火下降调整策略自适应调整学习率。

2.2 评价指标

选择交并比 IOU（Intersection Over Union）、 F_1 分数（ F_1 -Score）、模型参数量作为本文的评价指标。IOU 表示真实值和预测值两个集合交集和并集之比，表示预测结果与标签像素的交叠率，是衡量图像分割精度的重要指标，具体见公式（7）。 F_1 是精确率 P （Precision）和召回率 R （Recall）的调和平均值， F_1 同时考虑了正类别和负类别的性能，避免了精确率和召回率可能出现牺牲其中一个而提升另一个的情况， F_1 综合考虑二者，达到平衡。具体见公式（8）~（10）

$$\text{IOU} = \frac{\text{TP}}{\text{TP} + \text{FN} + \text{FP}} \quad (7)$$

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (8)$$

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (9)$$

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (10)$$

式中 TP 为被正确预测为裸岩的像素数；FP 为非裸岩被错误预测为裸岩的像素数；FN 为被错误预测为非裸岩的像素数。

模型参数量指模型中需要进行训练的可学习参数的数量。在不考虑偏置的情况下，对于输入特征图 $X \in \mathbb{R}^{C \times H \times W}$ ；卷积核 $k \times k \times n$ ，其中 k 为卷积核的高和宽， n 为卷积核的个数。标准卷积层参数数量为 $C \times k \times k \times n$ ；深度可分离卷积可分为深度卷积和逐点卷积，其参数数量为这两个步骤中的参数数量之和，参数数量为 $C \times k \times k + C \times 1 \times 1 \times n$ ；由此可知，深度可分离卷积与标准卷积的参数数量之比为 $\frac{1}{n} + \frac{1}{k^2}$ 。计算语义分割模型的参数量则是模型中的各个层的参数数量之和。

2.3 消融实验

为验证 CA-DC-MobileNetV3、FPN、CA、ASPP 模块异感受野特征融合方案的有效性，在相同数据集和实验环境上进行逐层的消融实验，以 IOU、 F_1 、模型参数量作为评价指标，实验结果如表 3 所示，实验 1~6 为调整性能指标的不同改进方案。为直观表达注意力机制作用效果，同时获取网络关注区域，使用 Grad-CAM^[25] 对施加注意力机制作用效果进行可视化，如图 6 所示。

对比实验 1、2 可知，改进的 CA-DC-MobileNetV3 网络相较于原始 MobileNetV3 网络 IOU 提高了 0.92 个百分点， F_1 提高了 0.64 个百分点，同时参数量得到减少，该方法减少了下采样的次数，为传入解码器的特征保留并获取了更丰富的特征，并使用 CA 使模型更准确地定位和识别感兴趣的对象，模型预

测精度得到提高。对比实验 2、3 可知, FPN 融合了深层与浅层特征, 获取了更丰富的细节, 提高了网络的预测精度, IOU 与 F_1 分别提高了 1.5、1.03 个百分点, 与此同时也增加了参数量。对比实验 3、4 可知, CA 与 FPN 的搭配使用使得模型的预测精度进一步提高, IOU 与 F_1 分别提高了 0.36、0.35 个百分点, 参数量保持平衡, 再次证明注意力机制的引入对网络预测起到正向作用。根据图 6 中高亮区域表示模型关注显著区域, 注意力机制的引入使得模型对特征不明显目标对象和阴影处的目标给予了更多的关注。对比实验 3、5 可知, 异感受野特征融合通过提高信息的利用率, 提升了模型的预测精度, IOU 与 F_1 分别提高 0.53、0.34 个百分点, 但同时也增加了模型的参数量。综合表 3, 实验 6 通过 FPN、CA、Add 操作进行特征融合, 相较于实验 1 模型 IOU、 F_1 分别提高了 3.39、2.32 个百分点, 同时参数量减少了 1.12×10^6 , 较好地平衡了参数量以及模型精度。

表 3 不同改进方案性能指标对比

Tab.3 Performance comparison of different improvement strategies

实验	特征提取网络		FPN	CA	异感受野特征融合	IOU/%	F_1 /%	参数量
	MobileNetV3	CA-DC-MobileNetV3						
1	√	×	×	×	×	69.07	81.71	5.63×10^6
2	×	√	×	×	×	69.99	82.35	4.41×10^6
3	×	√	√	×	×	71.49	83.38	4.44×10^6
4	×	√	√	√	×	71.85	83.62	4.44×10^6
5	×	√	√	×	√	72.02	83.73	4.50×10^6
6	×	√	√	√	√	72.46	84.03	4.51×10^6

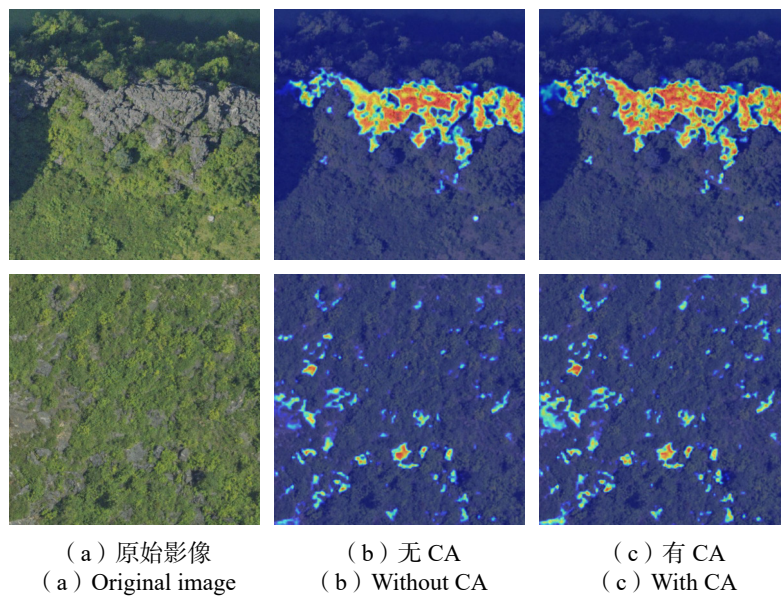


图 6 可视化的注意力图
Fig.6 Visual attention maps

2.4 不同模型对比

为验证本文提出模型的有效性, 选择基于面向对象的方法、PSPNet^[26]、SegNet、DeepLabV3+、SegFormer MiT-b0^[27] 语义分割模型与本文模型进行对比试验, 结果如表 4; 几种提取方法在四种不同的场景下的提取结果如表 5 所示。

由表 4 可知, 本文提出的模型在裸岩信息提取任务中表现最佳, 提取精度最高, IOU 与 F_1 分别达到了 72.46%、84.03%。比基于面向对象的方法、PSPNet、SegNet、DeepLabV3+、SegFormer MiT-b0 的





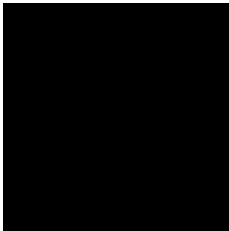
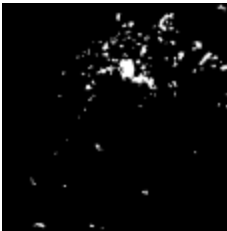
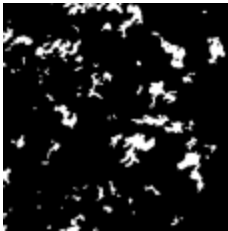

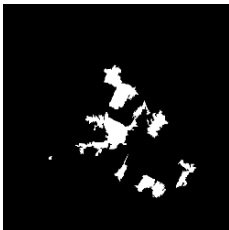
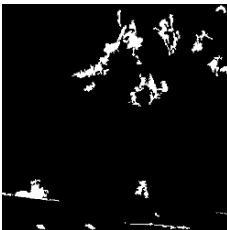


表4 不同算法性能对比
Tab.4 Performance comparison of different algorithms

提取方法	IOU/%	F_1 /%	参数量
面向对象	33.55	50.25	—
PSPNet	52.54	68.88	4.67×10^7
SegNet	67.23	80.40	2.94×10^7
DeepLabV3+	67.84	80.84	5.69×10^7
SegFormer MiT-b0	68.45	81.27	0.37×10^7
本文提取方法	72.46	84.03	0.45×10^7

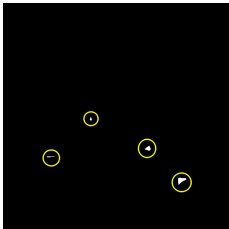
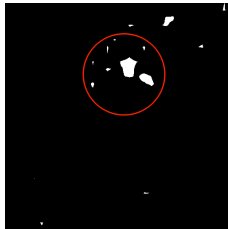


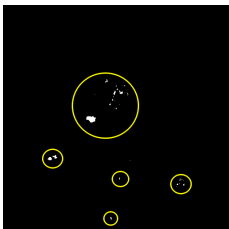
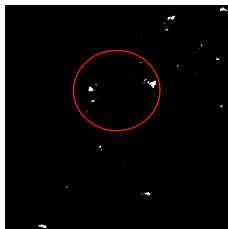


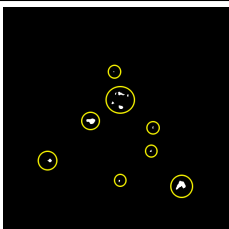
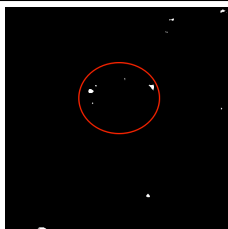


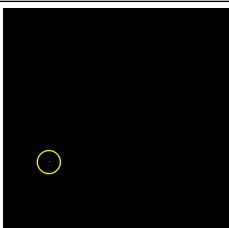


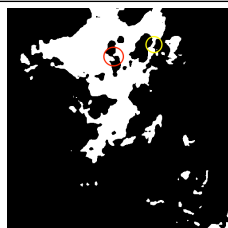

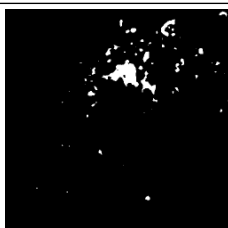


IOU 分别提高了 38.91、19.92、5.23、4.62 和 4.01 个百分点， F_1 分别提高了 33.78、15.15、3.63、3.19 和 2.76 个百分点。从模型参数规模上来看，SegFormer MiT-b0、本文模型参数量远低于其余模型，其中本文改进模型参数量约为原始模型的 1/13，略高于 SegFormer MiT-b0，但在精度上得到了提升，综合来看，本文模型优于 SegFormer MiT-b0。

表5 更直观地展示了不同语义分割模型在不同场景的预测结果，表5 中红色线圈为部分漏提区域、黄色为部分误提区域。表5 中场景1 为无裸岩区域，其中基于面向对象的方法误提现象最为严重，颜色上与裸岩相近的瓦块屋顶建筑物被误提为裸岩；除本文提出模型外，其余模型均出现了不同程度误提现象，存在噪声，其中 SegFormer MiT-b0 在场景1 误提现象最为轻微。场景2 中裸岩分布细碎，存在裸土背景及阴影区域，除本文模型外，其余模型均出现了较大区域的漏提现象，对于阴影目标敏感度较低，本文模型对细碎裸岩目标、阴影区域裸岩目标提取效果最佳，结果更为精确。场景3 中存在大量分布破碎目标，几种模型均出现了漏提的现象，其中本文模型提取轮廓边界更加清晰。场景4 中裸岩目标范围较大、边界轮廓清晰，除 PSPNet 模型提取结果外，几种模型在该场景表现较好，但仍然存在部分细节信息提取错误的现象。

表5 实验结果对比
Tab.5 Comparison of experimental results

模型	场景1	场景2	场景3	场景4
原始影像				
标签				
面向对象				

续表

模型	场景 1	场景 2	场景 3	场景 4
PSPNet				
SegNet				
DeepLabV3+				
SegFormer MiT-b0				
本文提取方法				

总体来看, 基于面向对象的提取方法、PSPNet 在裸岩提取任务中表现较差, 提取结果较为粗糙, 误提、漏提、边界信息丢失现象严重; SegNet、DeepLabV3+、SegFormer MiT-b0 相较于面向对象法、PSPNet 提取效果有明显的提高, 但仍存在大量误提、漏提现象; 本文提出模型的泛化性、细节表现力相较于其他模型具有更好的表现。

3 结束语

本文构建了一种结合 FPN、CA、添加异感受野特征融合改进的 DeepLabV3+ 喀斯特地区裸岩信息提取模型。该模型首先将 DeepLabV3+ 骨干提取网络替换为 CA-DC-MobileNetV3, 在很大程度上降低了模型的参数量, 同时提升了特征提取效果。其次, 引入 FPN 模块获取影像更丰富的细节, 并通过 CA 模块强化特征, 进一步加强特征提取能力。最后在 ASPP 模块新增异感受野特征融合特征层, 提高了特征信

息的利用率。结果表明,改进的 DeepLabV3+网络提升了模型预测精度,较好地改善了裸岩提取任务中误提、漏提的问题。在后续的工作中,将继续增加裸岩提取任务数据集,进一步提高模型的泛化能力;并继续优化模型的性能,提高模型在更复杂区域的提取精度,为石漠化调查工作提供更高效率的方案。

参考文献 (References)

- [1] 马一琳. 南宁市岩溶地区石漠化动态分析[D]. 南宁: 广西大学, 2015.
MA Yilin. Dynamic Analysis of Rocky Desertification of Karst Area in Nanning[D]. Nanning: Guangxi University, 2015. (in Chinese)
- [2] 滕永青, 孙娟, 王永雄. 北方沙漠化和水土流失防治成熟有效方案对南方石漠化防治的借鉴[J]. 农业环境与发展, 2007, 24(1): 61-63.
TENG Yongqing, SUN Juan, YU Yongxiong. The Reference of Mature and Effective Schemes for Desertification and Water and Soil Loss Control in the North to Rocky Desertification Control in the South[J]. Journal of Agricultural Resources and Environment, 2007, 24(1): 61-63. (in Chinese)
- [3] 王世杰. 喀斯特石漠化——中国西南最严重的生态地质环境问题[J]. 矿物岩石地球化学通报, 2003, 22(2): 120-126.
WANG Shijie. The Most Serious Eco-geologically Environmental Problem in Southwestern China— Karst Rocky Desertification[J]. Bulletin of Mineralogy, Petrology and Geochemistry, 2003, 22(2): 120-126. (in Chinese)
- [4] 陈希, 兰安军, 熊康宁, 等. 基于多源遥感数据的石漠化信息提取研究[J]. 信阳师范学院学报(自然科学版), 2018, 31(3): 437-442.
CHEN Xi, LAN Anjun, XIONG Kangning, et al. Research on the Information Extraction of Rocky Desertification Based on Multi-Source Remote Sensing Image Data[J]. Journal of Xinyang Normal University(Natural Science Edition), 2018, 31(3): 437-442. (in Chinese)
- [5] 岩溶地区水土流失综合治理技术标准: SL 461-2009[S]. 中华人民共和国水利部, 2009.
Techniques Standard for Comprehensive Control of Soil Erosion and Water Loss in Karst Region: SL 461-2009[S]. Ministry of Water Resources of the People's Republic of China. (in Chinese)
- [6] 张建军, 朱金兆. 水土保持监测指标的观测方法[M]. 北京: 中国林业出版社, 2013: 87-89.
ZHANG Jianjun, ZHU Jinzhao. Observation Method of Water and Soil Conservation Monitoring Indicators[M]. Beijing: China Forestry Publishing House, 2013: 87-89. (in Chinese)
- [7] 杨董琳, 王忠诚, 胡文敏, 等. 石漠化地区遥感影像信息提取方法综述[J]. 安全与环境工程, 2020, 27(3): 133-141.
YANG Donglin, WANG Zhongcheng, HU Wenmin, et al. Survey of Remote Sensing Image Information Extraction Methods in Rocky Desertification Areas[J]. Safety and Environmental Engineering, 2020, 27(3): 133-141. (in Chinese)
- [8] XIE X, DU P, XIA J, et al. Spectral Indices for Estimating Exposed Carbonate Rock Fraction in Karst Areas of Southwest China[J]. IEEE Geoscience & Remote Sensing Letters, 2015, 12(9): 1988-1992.
- [9] 杨晶. 基于无人机影像与Landsat8影像关联的重庆市石漠化信息提取与现状评价[D]. 武汉: 长江科学院, 2016.
YANG Jing. The Information Extraction and Evaluation of Rocky Desertification in Chongqing City Based on the Association of Unmanned Aerial Vehicle Photo and Landsat8 Image[D]. Wuhan: Changjiang River Scientific Research Institute, 2016. (in Chinese)
- [10] 张志慧, 刘雯, 李笑含, 等. 基于无人机航摄影像的喀斯特地区裸岩信息提取及景观格局分析[J]. 地球信息科学学报, 2020, 22(12): 2436-2444.
ZHANG Zhihui, LIU Wen, LI Xiaohan, et al. The Spatial Distribution Pattern of Rock in Rocky Desertification Area Based on Unmanned Aerial Vehicle Imagery and Object-oriented Classification Method[J]. Journal of Geo-information Science, 2020, 22(12): 2436-2444. (in Chinese)
- [11] 王梦娟. 基于多源遥感数据的岩溶石漠化信息提取研究[D]. 兰州: 兰州理工大学, 2019.
WANG Mengjuan. Research on Information Extraction of Karst Rocky Desertification Based on Multi-source Remote Sensing Data[D]. Lanzhou: Lanzhou University of Technology, 2019. (in Chinese)
- [12] WENG L, XU Y, XIA M, et al. Water Areas Segmentation from Remote Sensing Images Using a Separable Residual SegNet Network[J]. International Journal of Geo-Information, 2020, 9(4): 256.
- [13] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Im-

- age Segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [14] YE H, LIU S, JIN K, et al. CT-UNet: An Improved Neural Network Based on U-Net for Building Segmentation in Remote Sensing Images[C]//2020 25th International Conference on Pattern Recognition (ICPR), January 10-15, 2021, Milan, Italy. IEEE, 2021: 116-172.
- [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, October 5-9, 2015, Munich, Germany. Cham: Springer, 2015: 234-241.
- [16] ONG J C, LAU S L, ISMADI M, et al. Feature Pyramid Network with Self-guided Attention Refinement Module for Crack Segmentation[J]. *Structural Health Monitoring*, 2022, 22(1): 672-688.
- [17] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection[C]//30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 936-944.
- [18] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with Atrous Separable Convolution for Semantic Image Segmentation[C]//15th European Conference on Computer Vision (ECCV), September 8-14, 2018, Munich, Germany. Cham: Springer International Publishing AG, 2018: 833-851.
- [19] LIU S, QI L, QING H F, et al. Path Aggregation Network for Instance Segmentation[C]//31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 8759-8768.
- [20] LIU R, HE D. Semantic Segmentation Based on Deeplabv3+ and Attention Mechanism[C]//2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), June 18-20, 2021, Chongqing, China. IEEE, 2021: 255-259.
- [21] CHEN L C, PAPANDREOU G, KOKKINOS L, 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*, 2018, 40(4): 834-848.
- [22] HOU Q B, ZHOU D Q, FENG J S. Coordinate Attention for Efficient Mobile Network Design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 21-25, 2021, Nashville, TN, USA. IEEE, 2021: 13780-13717.
- [23] 葛小三, 曹伟. 一种改进DeepLabV3+网络的高分辨率遥感影像道路提取方法[J]. *遥感信息*, 2022, 37(1): 40-46.
GE Xiaosan, CAO Wei. A Road Extraction Method for High Resolution Remote Sensing Imagery Based on Improved DeepLabV3+ Model[J]. *Remote Sensing Information*, 2022, 37(1): 40-46. (in Chinese)
- [24] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, South Korea. IEEE, 2019: 1314-1324.
- [25] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. IEEE, 2017: 618-626.
- [26] ZHAO H, SHI J P, QI X J, et al. Pyramid Scene Parsing Network[C]//IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 2881-2890.
- [27] XIE E, WANG W, YU Z, et al. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers [EB/OL]. (2021-10-28) [2023-01-20]. <https://arxiv.org/abs/2105.15203>.

作者简介

吴永俊, 男, 1998年生, 2020年毕业于南京师范大学地理信息科学专业, 2023年获贵州大学测绘科学与技术专业硕士学位, 助理工程师。主要研究方向为喀斯特地区遥感应用的研究。E-mail: wuyongjnew@163.com。

通信作者

汪泓, 男, 1979年生, 2019年获东北大学博士学位, 副教授。主要研究方向为摄影测量与遥感。E-mail: 7653606@qq.com。

(编辑: 庞冰)