

引用格式: 贾克斌, 何岩, 魏之皓. 融合多尺度特征的高分辨率森林遥感图像分割[J]. 北京工业大学学报, 2024, 50(9): 1089-1099.

JIA K B, HE Y, WEI Z H. Fusing multi-scale features for segmentation of high resolution forest remote sensing images [J]. Journal of Beijing University of Technology, 2024, 50(9): 1089-1099. (in Chinese)

# 融合多尺度特征的高分辨率森林遥感图像分割

贾克斌<sup>1,2,3</sup>, 何岩<sup>1,2,3</sup>, 魏之皓<sup>4</sup>

(1. 北京工业大学信息学部, 北京 100124; 2. 先进信息网络北京实验室, 北京 100124;  
3. 计算智能与智能系统北京市重点实验室, 北京 100124; 4. 北京大学地球与空间科学学院, 北京 100871)

**摘要:** 为实现对青海三江源国家级自然保护区高原森林的有效监测, 基于深度学习技术提出一种融合多尺度特征的遥感图像分割算法。首先, 构建了该地区首个 2 m 空间分辨率的高原森林数据集; 其次, 为解决遥感图像真值标签不足影响网络模型训练的问题, 针对森林遥感图像分割的特点提出一种将图像打乱重组的数据增强方法, 将训练数据扩充至 1 600 张; 然后, 为解决主流分割网络处理大范围遥感图像存在无法聚焦细节的缺陷, 基于编解码结构, 提出一种融合多尺度特征的高分辨率森林遥感图像分割网络模型, 该模型融合了所设计的卷积模块、多尺度特征融合模块和特征放大提取模块。实验结果表明, 所提数据增强方法提升了模型的分割精度, 同时该模型经数据增强训练, 交并比(intersection over union, IoU)高达 89.64%, 结果优于当前主流图像分割模型。

**关键词:** 深度学习; 遥感; 图像分割; 多尺度特征融合; 数据增强; 数据集构建

中图分类号: TP 75

文献标志码: A

文章编号: 0254-0037(2024)09-1089-11

doi: 10.11936/bjtxb2023010021

## Fusing Multi-scale Features for Segmentation of High Resolution Forest Remote Sensing Images

JIA Kebin<sup>1,2,3</sup>, HE Yan<sup>1,2,3</sup>, WEI Zhihao<sup>4</sup>

(1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China;

2. Beijing Laboratory of Advanced Information Network, Beijing 100124, China;

3. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China;

4. School of Earth and Space Sciences, Peking University, Beijing 100871, China)

**Abstract:** To achieve the effective monitoring of plateau forests in the Sanjiangyuan National Nature Reserve in Qinghai, a fusing multi-scale features remote sensing image segmentation algorithm based on deep learning technology was proposed. First, the first 2 m spatial resolution plateau forest dataset in the region was constructed. Second, to solve the problem of insufficient ground-truth label of remote sensing images which affects the training of network models, a data augmentation method of shuffling and reorganizing images was proposed according to the characteristics of forest remote sensing image segmentation, and the training data was expanded to 1 600 images. To address the problem of mainstream segmentation networks that cannot focus on details in processing large-scale remote sensing images, a fusing multi-scale features high-resolution forest remote sensing image segmentation network model based

收稿日期: 2023-01-30; 修回日期: 2023-04-24

基金项目: 青海省应用基础重点基金资助项目(2020-ZJ-709)

作者简介: 贾克斌(1962—), 男, 教授, 主要从事智能化信息与信号处理技术方面的研究, E-mail: kebinj@bjut.edu.cn

on encoding and decoding structures was proposed. The model incorporated the designed convolution block, multi-scale feature fusion block and feature amplification extraction block. Results show that the data augmentation algorithm proposed improves the segmentation accuracy of the model, while the proposed model trained by the proposed data augmentation achieves an intersection over union (IoU) of 89.64%, and the results are better than that of the current mainstream image segmentation models.

**Key words:** deep learning; remote sensing; image segmentation; multi-scale features fusion; data augmentation; dataset construction

作为陆地生态系统的重要组成部分,森林生态系统在全球碳平衡、全球气候变化减缓、气候调节、环境改善中发挥着重要作用<sup>[1]</sup>。监测高原森林能够帮助人们更好地了解气候变化在地方、区域和全球层面的影响<sup>[2]</sup>。

传统的森林监测手段主要依靠人工判读大比例尺航片或人工实地考察。野外考察工作环境恶劣、强度大、考察周期长会造成人力、财力成本上升。人工判读大比例尺航片与人工实地考察均面临森林面积的监测精度低,无法对森林变化进行高时效、大范围的监测等问题<sup>[3-4]</sup>。

随着遥感卫星技术的发展,高分辨率卫星图像可用于大规模高海拔高原森林监测。其中,许多由卫星数据驱动的目标检测研究都是利用基于阈值的植被指数<sup>[5]</sup>,如20世纪90年代首次提出的归一化植被指数(normalized difference vegetation index, NDVI)<sup>[6]</sup>和比值植被指数(ratio vegetation index, RVI)<sup>[7]</sup>。同时,传统的基于机器学习的目标检测方法,如支持向量机(support vector machine, SVM)<sup>[8]</sup>、随机森林<sup>[9]</sup>等也被用于卫星数据分析。例如,王二丽等<sup>[10]</sup>在基于多时相遥感影像的北京平原人工林树种分类任务中,对比使用了基于SVM的方法。Wickramarathna等<sup>[11]</sup>通过计算NDVI与RVI等参数,利用随机森林算法对单个树冠进行分割,提供了一种估计单个树木位置和森林覆盖面积变化的方法。

伴随深度学习技术的发展,卷积神经网络(convolutional neural network, CNN)在图像分割领域展现了显著的效果。Shelhamer等<sup>[12]</sup>提出了全卷积网络(fully convolutional networks, FCN),将传统的全连接层替换为卷积层,从端到端训练,首次实现了对图像的像素级分类。Ronneberger等<sup>[13]</sup>提出U-Net网络模型用于医学影像分割,该模型将在编码层中提取到的特征信息通过跳跃连接拼接到对应的解码层中,实现了较好的分割效果。此外,Zhao等<sup>[14]</sup>提出的金字塔场景解析网络(pyramid scene parsing

network, PSPNet)、Badrinarayanan等<sup>[15]</sup>提出的SegNet、Chen等<sup>[16]</sup>提出的DeepLab V3+、Fu等<sup>[17]</sup>提出的双注意力网络(dual attention network, DANet)等模型都得到了广泛应用。

近年来,基于CNN的算法模型在遥感图像分析领域中得到了广泛的应用。Wagner等<sup>[18]</sup>使用U-Net网络和来自GeoEye卫星的高分辨率多光谱图像实现识别、分割,并绘制超过3 000 km<sup>2</sup>的亚马逊森林树冠棕榈树。贾莉等<sup>[19]</sup>在GlobeLand30和全国30 m分辨率森林分类结果的基础上,利用U-Net网络模型对山区地表覆盖分类提取。

Wei等<sup>[20]</sup>将小样本学习的方法用于三江源地区大比例尺高原森林制图,对比实验结果表明,传统的阈值方法与传统机器学习方法在该地区的高原森林遥感图像分割效果表现较差。与此同时,基于深度学习的图像分割在遥感领域展示出很好的应用前景,但由于其在遥感领域的应用起步较晚,针对遥感图像分割任务的研究主要还是采用处理自然界图像或医学图像的主流分割算法。遥感图像覆盖范围广泛、数据规模大、景物类型多样、目标相对密集且涵盖大量复杂多样的地理地貌种类,相同尺寸下的遥感图像相较于自然界图像或医学图像包含了更多的语义信息。现有主流的深度学习分割算法在处理遥感影像分割任务时,虽然能达到相对满意的效果,但无法较好地兼顾图像整体性与图像细节的处理,特别是对细节特征信息的提取能力不足,最终导致分割结果的图像中部分细节丢失。同时,在处理遥感图像分割任务时,由于缺乏足够的手工标注真值标签对分割模型进行训练,无法使模型训练到相对较好的状态,最终影响了图像分割效果。

本文以青海省三江源地区为研究区,为实现使用深度学习的方法对该地区高原森林覆盖面积的精确监测,构建了该地区首个2 m空间分辨率的高原森林数据集。首先,面对缺乏大量手工标注真值标签进行训练的问题,针对森林遥感图像分割的特点,提出一种数据增强方法对训练集数据进行扩充。该

方法将原图像平均切割,在保持尺寸不变的前提下随机排列组合成新图像,将训练数据数量扩充至1 600张。其次,面对主流深度学习分割模型对遥感图像细节特征信息提取能力不足的问题,针对遥感图像的特点提出一种融合多尺度特征的深度神经网络分割模型。该模型基于编解码结构,融合了所提的卷积模块、多尺度特征融合模块和特征放大提取模块。卷积模块通过增加 $1 \times 1$ 卷积,从小尺度提取特征信息,加强模型对图像细节的提取能力。多尺度特征融合模块基于空洞空间金字塔池化(atrous spatial pyramid pooling, ASPP)设计,用于编解码结构的底层,融合多尺度进行全局上下文信息的提取。特征放大提取模块融入编解码结构之间连接的部分,从不同尺度放大提取遥感图像的细节特征信息。与深度学习领域的主流图像分割模型对比的实验结果表明,本文所提网络模型达到了更好的效果。同时,通过所提出的数据增强方法扩充数据训练,所有模型效果均有不同程度的提升,并且本文所提模型在使用所提出的数据增强方法训练后,效果进一步提升,交并比(intersection over union, IoU)高达89.64%。

## 1 研究区及数据

### 1.1 研究区概况

本文研究区位于中国青海省、地处青藏高原腹地的三江源国家级自然保护区,其地理位置位于北纬 $31^{\circ}39' - 36^{\circ}16'$ ,东经 $89^{\circ}24' - 102^{\circ}23'$ 。该区域气候类型为高原大陆性气候,湿地广布,沼泽密集,具有丰富的水资源,是长江、黄河、澜沧江的发源地。三江源地区平均海拔约为4 500 m,总面积约为39.5万 $\text{km}^2$ ,其中林地面积约为3.0万 $\text{km}^2$ 。森林植被种类繁多、分布复杂广泛,主要以寒温性的针叶林为主。因此,针对该区域高原森林覆盖面积变化的监测,对我国西部地区气候变化分析以及该地区水资源保护具有重要意义。

### 1.2 数据集构建

为实现应用深度学习的方法开展青海省三江源国家级自然保护区高原森林监测任务的研究,本文采用资源三号遥感卫星(ZY-3)构建了该区域首个2 m空间分辨率的高原森林遥感数据集,其中ZY-3卫星参数信息如表1所示。

本文将来自ZY-3卫星的高原森林遥感图像切割为38 708张 $128 \times 128$ 像素的样本。此外,为了生成合理的森林分割标签,花费几个月的时间,基于可视化Labelme工具开展像素级别的高精度手工真值

表1 资源三号遥感卫星参数信息

Table 1 Parameter information of ZY-3 remote sensing satellite

关键参数	数值
时间分辨率/d	5
空间分辨率/m	2
光谱范围/ $\mu\text{m}$	0.45 ~ 0.89
轨道高度/km	505.984

样本标注。在当前版本的数据集中共手工标记了1 187个2 m空间分辨率的精确森林分割人工地面真实数据。因为ZY-3卫星遥感影像具有4个光谱通道,比传统RGB三光谱通道图像多一个近红外(near-infrared, NIR)光谱通道,所以在遥感图像可视化过程中,只采用RGB通道进行符合人眼常见规律的可视化展示。其中部分遥感影像样本及其手工标注真值标签数据的可视化图像如图1所示。

从图1中可以看出,该地区的森林呈现地貌、种类、分布多样化的特点。同时,Han等<sup>[21]</sup>对多种主流深度学习图像分割算法进行了对比,在实现对中国中部河南省新乡市地区的大尺度土地覆盖制图任务中,利用高分2号遥感卫星分别制作仅有RGB的三通道数据集和拥有RGB+NIR的四通道数据集进行实验。通过对比实验效果发现,训练网络模型时,采用RGB+NIR的四通道数据集展示出了更好的效果。因此,在数据集的构建中,本文构建了通道数为RGB+NIR的四通道数据集。其中所构建的高原森林遥感数据集的相关信息如表2所示。

### 1.3 数据增强

在当前情况下,受有效采集数量、人工标注真值标签成本等多因素的限制,所使用的训练数据集数据量相对有限。在深度学习模型训练中,为增加训练模型的鲁棒性,普遍采用数据增强来扩充训练集的数据量。现有数据增强方法主要是一些对图像进行简单的转换,例如:几何变换,包括将图像水平或垂直翻转、随机旋转、随机缩放等;颜色与空间变换,包括对图像随机剪裁、随机变形,改变图像的颜色及亮度、对比度、饱和度等参数;增加噪声方法,包括对图像添加随机的椒盐噪声、高斯噪声等<sup>[22]</sup>。这些数据增强方法对数据的处理相对简单,只能提升数据集的数量,但无法保证数据质量的提升,进而无法有效提升训练结果,甚至在某些情况下会导致模型训练效果下降。有效的数据增强方法不仅要数据集的数据量进行扩充,同时还要兼顾所扩充数据的质量。



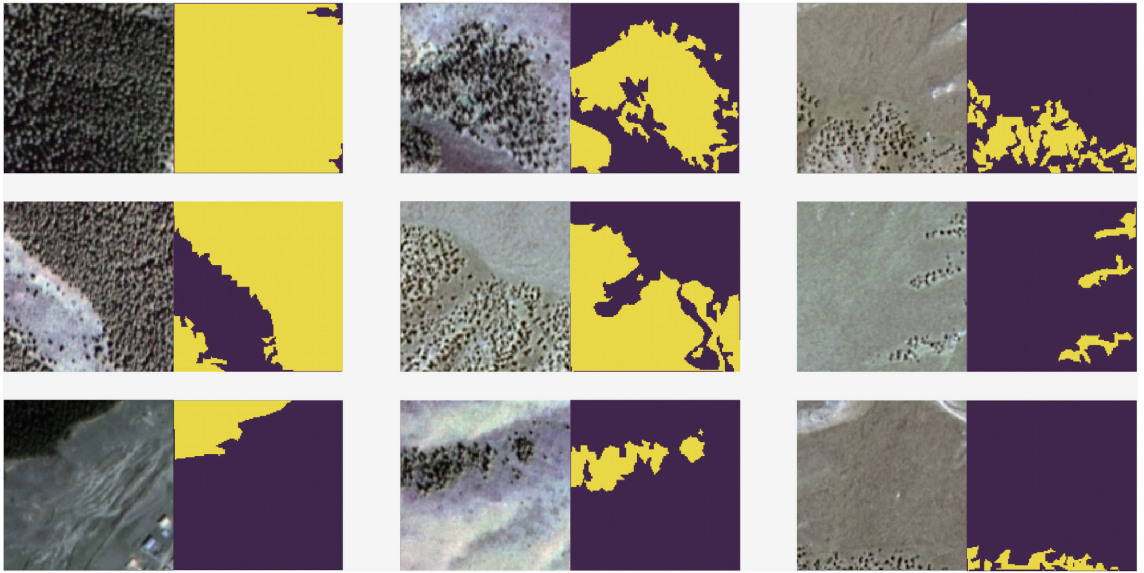


图1 部分遥感影像样本及其真值标签可视化图像

Fig. 1 Visualization of some remote sensing image samples and their ground-truth label

表2 森林遥感数据集相关信息

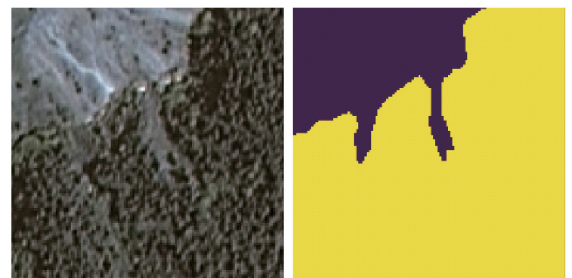
Table 2 Relevant information of the forest remote sensing dataset

参数名称	参数信息
影像数据源	资源三号遥感卫星
数据样本数	38 708
人工标注真值标签数	1 187
影像样本尺寸/像素	128 × 128
影像样本通道	四通道
人工标注真值标签尺寸/像素	128 × 128
人工标注真值标签通道	一通道
样本像素点分辨率/m	2
影像样本时间分布	2017-01—12
人工标注真值时间分布	2017-05—06

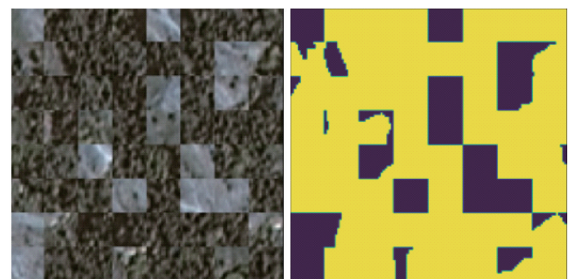
本文所构建的数据集的图像仅为森林和非森林二分类图像,并且森林种类繁多,地理地貌分布复杂多样。为解决遥感图像分割模型训练时缺乏大量手工标注真值标签的问题,进一步提升 CNN 模型的鲁棒性,本文综合高原森林地貌的特点及大尺度遥感图像二分类分割任务的特点,提出一种特殊数据增强方法。该方法将原始图像剪裁并重新随机排列组合成一张森林分布更为随机的新图像。其具体操作为:将原始  $128 \times 128$  像素的图像按  $8 \times 8$  平均切割为 64 个大小相同的小块,并将这些小块重新随机排列组合成一张新的  $128 \times 128$  像素的图像。其可视

化效果如图 2 所示。

经过所提数据增强方法操作,扩充后的训练模型图像中森林树木的分布更加随机,虽然牺牲了图像的部分整体性,但极大地提升了数据复杂度。通过本文所提数据增强方法,训练数据扩充至 1 600 张,同时强化了训练模型对图像细节的感知能力,增加了训练模型的鲁棒性。实验结果表明,通过本文所提出的数据增强方法训练,能有效提升分割模型的准确率。



(a) 原始图像及其真值



(b) 数据增强后图像及其真值

图2 数据增强可视化图像

Fig. 2 Visualization of data augmentation image

## 2 研究方法

用于处理医学影像分割任务的 U-Net 算法在遥感领域得到了广泛的应用。为提升遥感图像分割精度, 本文基于 U-Net 网络的编解码结构, 针对大尺度遥感影像的特点, 提出融合多尺度特征的高分辨率森林遥感图像分割网络模型。该模型融合了本文所提出的卷积模块、多尺度特征融合模块与特征放大提取模块。卷积模块在 U-Net 的 2 层  $3 \times 3$  卷积的基础上增加 2 层  $1 \times 1$  卷积, 加强网络模型对细节特征信息的提取能力。多尺度特征融合模块基于 ASPP 设计, 在编解码结构的底层融合不同尺度以提取全局上下文信息。特征放大提取模块应用于编解码结构的跳跃连接部分, 通过  $2 \times 2$  与  $4 \times 4$  尺度的上采样放大提取图像的细节特征信息。本文所提网络模型的主体结构如图 3 所示。其流程为: 首先, 在编码器部分对输入的遥感图像连续逐层编码。每一

层编码均采用所提卷积模块卷积, 并使用尺寸为  $2 \times 2$  的最大池化层进行下采样, 将特征信息尺寸缩小为上一层的一半, 同时在下一层卷积模块中将特征信息的通道数扩大为上一层的 1 倍。其次, 在编解码结构底层对编码后的特征信息使用多尺度特征融合模块, 融合不同尺度以提取全局上下文语义信息。然后, 在解码器部分对提取到的特征信息连续逐层解码。每一层解码均先使用  $2 \times 2$  大小的反卷积对特征信息上采样, 将特征尺寸扩大为上一层的 1 倍, 特征通道数缩小为上一层的一半。与此同时, 本文使用所提出的特征放大提取模块在编解码结构的连接部分放大提取来自编码器的细节特征信息, 并与解码器上采样后的同尺寸特征信息相拼接, 随后使用所提卷积模块对拼接过的特征信息进行卷积操作。最后, 将解码器解码后的特征信息通过  $1 \times 1$  卷积与 Sigmoid 激活函数得到最终的图像分割结果。

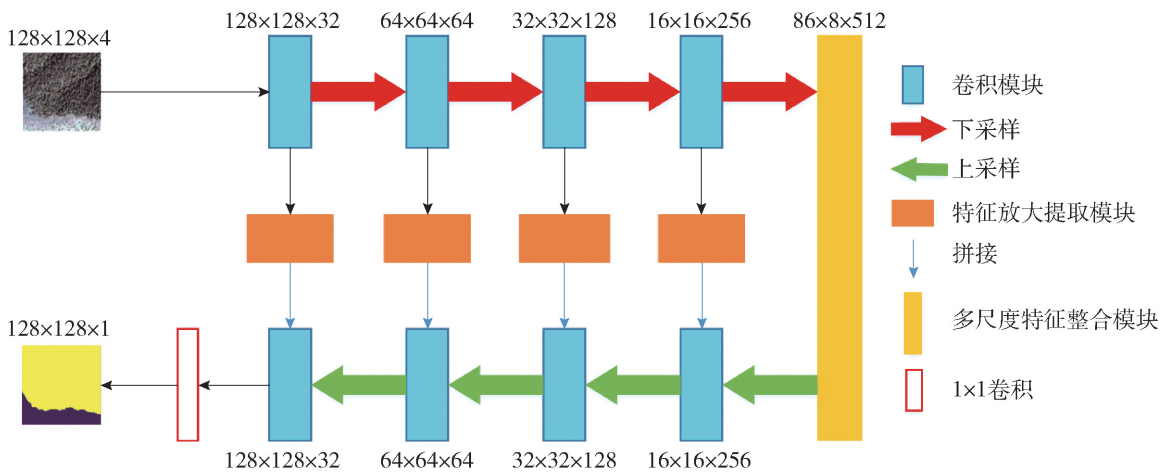


图3 网络模型主体结构

Fig. 3 Main architecture of the network model

### 2.1 卷积模块

在编解码结构的网络中处理特征信息时, 以 U-Net 为主流的分割模型均采用 2 层  $3 \times 3$  卷积进行图像特征信息提取。遥感图像覆盖范围广、数据规模大, 同尺寸下的遥感图像与自然界图像或医学图像相比包含了更多的语义信息。因此, 在处理遥感图像分割任务中, 需要相对较小尺度的卷积进行特征信息的提取。为提高模型对遥感图像细节的提取能力, 本文提出卷积模块, 先使用 1 层  $1 \times 1$  卷积提取小尺度细节特征信息, 再使用 2 层  $3 \times 3$  卷积对特征信息进一步提取, 最后使用 1 层  $1 \times 1$  卷积进一步提取小尺度细节特征信息。具体结构如图 4 所示。其流程为: 首先, 通过  $1 \times 1$  卷积操作缓解编解码层之

间特征尺寸与通道变化时导致的特征信息丢失问题; 然后, 连续进行 2 次  $3 \times 3$  卷积操作以提取特征信息; 最后, 将特征信息以  $1 \times 1$  卷积进行小尺度特征的进一步提取。其中每次卷积操作都伴随着使用 Relu 激活函数, 在最后一次  $3 \times 3$  卷积后通过批处理归一化 (batch normalization) 加速网络的收敛。

### 2.2 多尺度特征融合模块

空洞卷积可以通过改变膨胀率控制卷积层的接收野, 通过收集不同分辨率的图像特征获取多尺度的全局上下文信息, 达到兼顾处理全局与细节特征信息的目的, 进而提高分割性能。本文受 DeepLab 网络模型中的 ASPP 模块启发, 设计了多尺度特征融合模块并融入所提网络的编解码结构底层部分。

具体结构如图5所示。

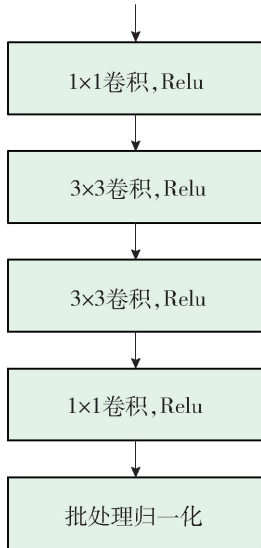


图4 卷积模块

Fig. 4 Convolution block

外,还有一个分支为全局平均池化,用来合并全局上下文信息生成全局特征。随后将5个分支得到的不同尺度特征信息拼接在一起,最后通过 $1 \times 1$ 卷积降低通道数,使输出结果与输入特征图的尺寸保持一致,达到了提取多尺度上下文信息的目的,较好地兼顾了遥感图像整体性与细节的处理。

### 2.3 特征放大提取模块

由U-Net网络所提出的跳跃连接只是将网络中编解码两部分所对应层的特征相拼接,并没有对来自编码器的特征信息处理。对此,很多文献针对这一部分进行了改进,例如, Oktay等<sup>[23]</sup>提出的Attention U-Net网络将跳跃连接部分改为所提出的注意力模块; Ibtehaz等<sup>[24]</sup>提出的MultiResUNet网络将跳跃连接部分改为所提出的卷积模块。但是,针对遥感影像分割任务的处理,由于网络模型缺乏对图像细节信息的提取,这些改进后的模型效果并不突出。为加强深度学习网络对遥感图像细节特征的处理能力,本文提出了特征放大提取模块。该模块用于编解码结构的跳跃连接部分,对来自编码器的特征信息进行放大卷积操作,提取从不同尺度下放大的图像细节特征信息。具体结构如图6所示。其流程为:首先,分别使用 $2 \times 2$ 与 $4 \times 4$ 尺寸的双线性插值对输入特征进行上采样;然后,分别对2种尺度放大后的特征使用卷积模块进行细节特征信息的提取;最后,分别使用 $2 \times 2$ 与 $4 \times 4$ 尺寸的最大池化层将放大提取到的特征的尺寸恢复为输入时的大小,接着使用 $1 \times 1$ 卷积与Relu激活函数减少尺寸变化导致的特征信息丢失问题。为了保留来自编码器中原有的特征信息,将2种尺度下放大提取到的细节特征与来自编码器的原始特征拼接,然后使用 $3 \times 3$ 卷积与Relu激活函数在通道数不变的情况下卷积,进一步提取特征信息。最后,使用卷积模块将特征通道数还原为输入时的大小,进一步分配提取到的细节特征。

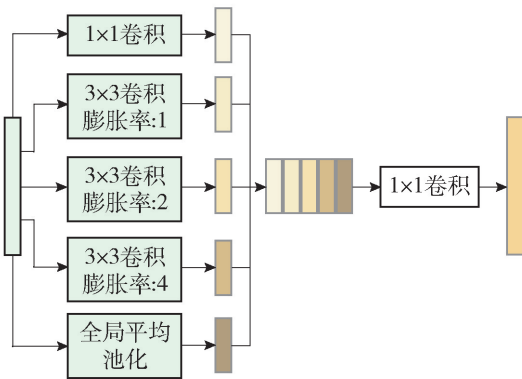


图5 多尺度特征融合模块

Fig. 5 Multi-scale feature fusion block

多尺度特征融合模块采用5个分支并行处理输入的特征信息。其中第1个分支为 $1 \times 1$ 卷积,另外3个分支均采用卷积核大小为 $3 \times 3$ ,膨胀率分别为1、2、4的空洞卷积以提取不同尺寸的图像特征。此

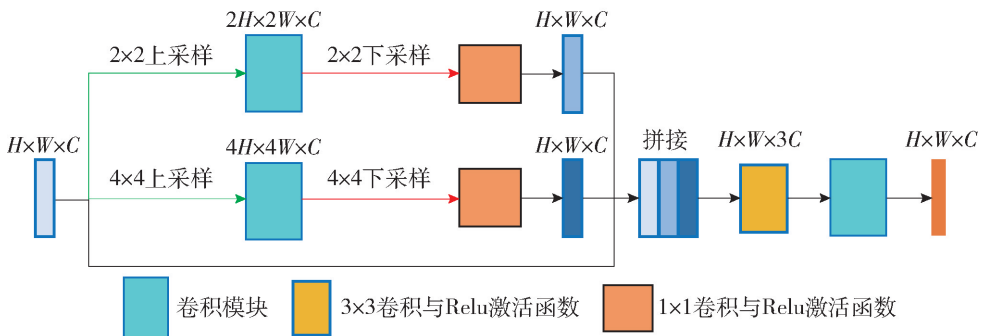


图6 特征放大提取模块

Fig. 6 Feature amplification extraction block



### 3 实验与分析

#### 3.1 实验环境及数据集

本文的实验环境为 Ubuntu 20.04 系统, Intel Xeon Gold 5218R CPU, 显卡为 Nvidia GeForce RTX 3090 24 GB。本文的图像分割网络模型由 Keras 2.7.0 框架搭配 CUDA 11.7 和 CUDNN 8.2.4 实现, 开发语言为 Python 3.8。所有模型训练的优化器使用收敛相对较快的 Adam 优化器, 学习率设置为 0.0001。使用的损失函数为图像分割任务所常用的二元交叉熵函数。每个训练批次大小设置为 16, 共进行 50 轮训练。

实验数据集为本文所构建的基于青海省三江源地区的高原森林遥感影像数据集。由于数据集中所标注的相邻图像具有环境类型相似的特点, 先将 1187 张人工标注真值随机打乱顺序, 然后选取打乱顺序后的前 800 张人工标注真值作为训练集, 其余的 387 张人工标注真值作为测试集。在打乱顺序后的数据中, 森林类型和地貌环境分布更为平均, 这样的数据分布可以在训练中增强训练模型的鲁棒性, 在测试集验证模型效果中也更具有科学性与泛化性。在验证本文所提数据增强方法效果的实验中, 训练集为打乱顺序后的前 800 张人工标注真值标签及其所对应的数据增强后的数据, 共 1600 张。测试集使用的数据同样为其余 387 张人工标注真值标签, 与未使用数据增强的实验中所使用的测试集数据保持一致。

#### 3.2 评价标准

图像分割准确性的评价标准有多种。本文使用精确率 (precision)  $P$ 、召回率 (recall)  $R$ 、 $P$  和  $R$  的调和平均数  $F_1$  值与交并比  $R_{IoU}$  这 4 种精度指标作为算法性能的评价标准, 其中以在图像分割领域应用最为广泛的  $R_{IoU}$  作为衡量精度的主要指标。计算公式分别为

$$P = \frac{T_p}{T_p + F_p} \quad (1)$$

$$R = \frac{T_p}{T_p + F_n} \quad (2)$$

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (3)$$

$$R_{IoU} = \frac{T_p}{T_p + F_p + F_n} \quad (4)$$

式中: 真阳性  $T_p$  表示被正确分割的森林特征像素数量; 假阳性  $F_p$  表示被错误分割的森林特征像素数

量; 假阴性  $F_n$  表示被错误分割的非森林特征像素数量。

#### 3.3 实验结果及分析

在实验设计上, 本文使用多种主流的图像分割算法与本文所提出的算法进行对比实验, 从客观与主观 2 个方面进行了比较。与本文所提算法相比较的主流算法包括 FCN、DeepLab 系列、PSPNet、DANet、SegNet、U-Net 以及对其改进的网络 MultiResUNet 和 Attention U-Net。其中 FCN 有 3 种架构, 包括 FCN-8s、FCN-16s、FCN-32s, 本文使用的是经大量实验证明在语义分割中表现最好的 FCN-8s。在 DeepLab 系列中, 本文使用的是其最新版本 DeepLab V3+。

首先在客观评价方面, 本文以  $R_{IoU}$  为主要评价标准, 同时参考  $P$ 、 $R$ 、 $F_1$  进行算法性能评价。不同网络的分割性能对比如表 3 所示。由表可知: 在本文所提出的数据集上所训练的主流深度学习图像分割网络模型中, DeepLab V3+ 所展现的效果低于平均水平; 相较于其他主流网络, U-Net 与基于 U-Net 改进的 Attention U-Net 网络模型在分割效果上表现较好, 但是基于 U-Net 改进的 MultiResUNet 的效果并不如 U-Net 网络; 本文所提网络模型均优于主流分割模型, 其  $R_{IoU}$  高达 89.54%。同时, 综合参考  $P$ 、 $R$  与  $F_1$ , 所提网络模型同样达到最优水平。通过使用本文所提网络模型结合本文所提数据增强方法训练后, 得到的模型效果进一步提升,  $R_{IoU}$  达到 89.64%, 虽然提升效果相对有限, 但在主流的分割网络中使用本文所提的数据增强方法训练, 效果均有明显提升。本文所提出的数据增强方法主要用于强化网络对图像细节的感知能力, 主流分割网络对遥感图像细节感知能力较差, 提升效果明显, 而本文所提网络模型针对遥感图像特点设计, 所设计的模块注重提取细节特征信息的能力。

为进一步展示本文所提算法的优越性, 基于人眼主观对本文所提算法与主流图像分割算法的分割结果进行对比, 其中 2 张特点较为突出的遥感图像分割结果对比如图 7 所示。图中算法含 \* 的为使用经本文所提数据增强方法扩充后的 1600 张数据训练的结果, 不含 \* 的为仅使用原始 800 张数据训练的结果。综合图 7 中 2 张图像的森林分割结果对比可以看出, 由 FCN-8s 算法得到的分割结果呈现方块状且边缘部分含有大量噪声点。与之相似的是, 由 PSPNet 算法得到的分割结果中, 图像边界同样存在大量噪声点, 两者均未能很好地刻画图像中森林的

表3 不同网络的分割性能对比

Table 3 Comparison of segmentation performance of different networks

算法	不含数据增强实验				含数据增强实验				%
	$P$	$R$	$F_1$	$R_{IoU}$	$P$	$R$	$F_1$	$R_{IoU}$	
	FCN-8s	93.05	89.92	91.06	84.58	93.63	92.39	92.77	
DeepLab V3 +	88.26	87.67	86.92	78.28	91.79	90.42	90.53	83.60	
PSPNet	92.74	93.73	92.98	87.44	92.51	95.05	93.48	88.29	
DANet	92.87	93.93	93.04	87.65	92.14	94.84	93.18	87.90	
SegNet	92.38	93.74	92.76	87.11	92.51	94.17	93.09	87.61	
U-Net	93.31	93.84	93.30	87.95	92.60	94.86	93.41	88.20	
MultiResUNet	92.45	94.09	93.00	87.57	92.58	94.47	93.35	88.11	
Attention U-Net	94.09	93.98	93.72	88.74	93.82	94.70	94.00	89.14	
本文算法	94.17	94.80	94.26	89.54	94.46	94.63	94.30	89.64	

边缘部分。相比之下,由其他主流图像分割算法得到的分割结果中未含有噪声点,森林的边缘部分相对平滑。主流图像分割算法在遥感图像局部细节信息的提取方面,展现了不同程度的效果,但均未能将大尺度遥感图像中的小范围非森林部分较好地提取出来,在不同程度上错误地把非森林部分分割为森林部分。本文所提出的分割模型是基于大尺度森林遥感图像的特点而设计,融合了不同尺度的特征信息,既保证森林图像分割的整体性,又加强了对图像中细节特征信息的提取能力。从图中可以看出,本文所提算法的分割结果中,边缘平滑清晰且能较为精细地分割出遥感图像中的细节部分,并能将图中非森林部分更精确地提取出来。同时,为验证本文所提数据增强方法的有效性,本文通过对边缘存在噪声点的 PSPNet 算法、在遥感图像分割任务中应用最为广泛的 U-Net 算法、主流算法中  $R_{IoU}$  表现最好的 Attention U-Net 算法以及本文所提出的遥感图像分割算法,使用本文所提数据增强方法扩充后的 1 600 张训练数据进行训练,与未使用数据增强的原始 800 张数据进行训练得到的效果进行对比。通过数据量更多且森林分布更为复杂的训练集数据训练,分割模型对图像细节的提取能力得到了加强,不同的分割算法模型在不同程度上提高了对遥感图像中非森林部分的识别准确度,其中本文所提分割算法的效果最好。同时,PSPNet 算法由于获得了足够的训练数据,分割结果中噪声点消失,得到了平滑清晰的森林边缘部分。从图 7(a) 的图像 1 森林分割结果对比可以看出,本文所提算法的细节提取能力优于主流算法,能够较好地提取图像中间部分的非

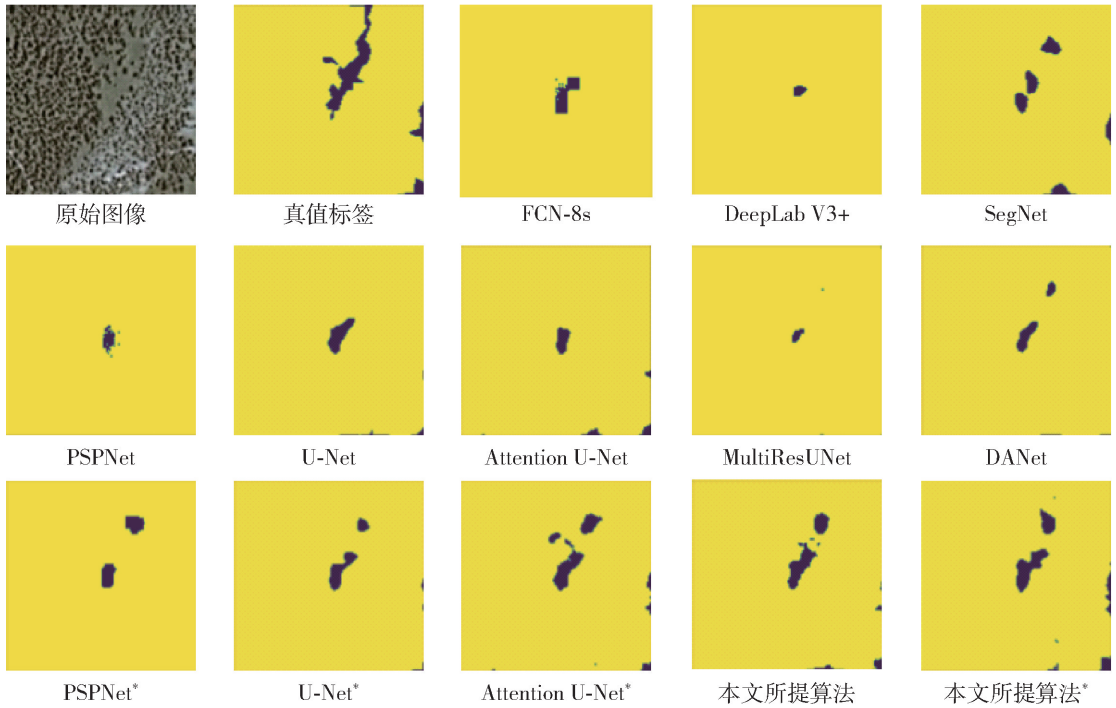
森林区域,甚至右下部分极小的细节也能展现出来。本文所提出的数据增强方法也进一步加强了分割算法对这些部分的提取能力。从图 7(b) 的图像 2 森林分割结果对比可以看出,即使人工标记的真值图中将左下部分非森林错误标记为森林,由于本文所提算法对细节的强大处理能力,也能较为准确地进行分割,而主流分割算法很难对这部分正确识别。经本文所提数据增强方法训练后,主流算法在一定程度上正确提取了部分非森林区域,而本文提出的算法的分割效果进一步得到了提升。因此,本文认为,由于所提数据增强方法能将数据中被错误划分部分准确分割出来,本身正确的分割被判定为错误分割,同样导致所提数据增强方法在客观评价方面表现并不突出,但可以从主观人眼可视化的部分展示出其优越性。

### 3.4 消融实验

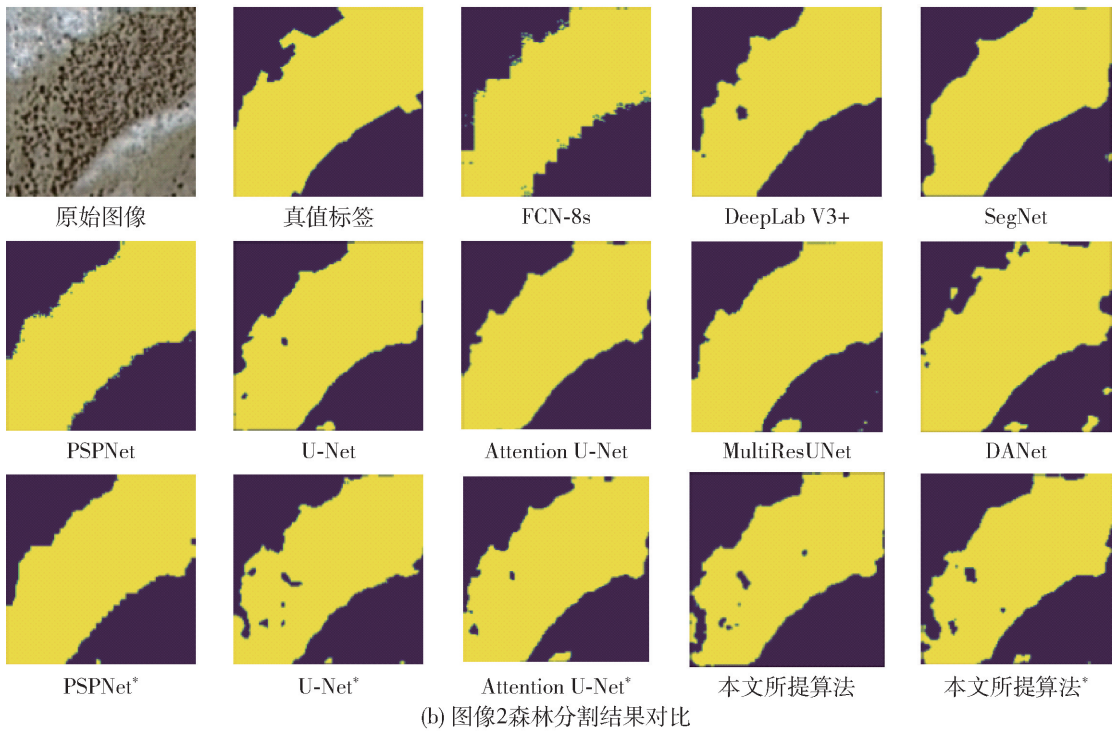
本文所提出的融合多尺度特征的高分辨率森林遥感图像分割算法是基于具有编解码结构的 U-Net 而设计,融合了本文所提卷积模块、多尺度特征融合模块以及特征放大提取模块。为验证所提模块在模型中的有效性,本文设计了消融实验,使用 800 张大尺度高原森林遥感图像数据进行实验,对比了基于 U-Net 添加所提出的不同模块的分割性能,实验结果如表 4 所示。

从表中可以看出:本文所提出的特征放大提取模块对分割效果的提升最为明显,在 U-Net 中仅添加该模块, $R_{IoU}$  高达 88.75%。其次为多尺度特征融合模块,在 U-Net 中仅添加该模块, $R_{IoU}$  也达到了 88.56%。卷积模块同样对分割效果的提升有促进





(a) 图像1森林分割结果对比



(b) 图像2森林分割结果对比

图 7 不同算法的森林分割结果对比

Fig. 7 Comparison of forest segmentation results of different networks

作用,在 U-Net 中将卷积部分替换为所提出的卷积模块,  $R_{10U}$  达到了 88.11%, 相较于原来的 87.95% 提高了 0.16 个百分点,但在 U-Net 中同时添加卷积模块和多尺度特征融合模块或特征放大提取模块时,效果得到了更好的提升。二者经添加卷积模块后效

果分别达到了 88.92% 和 89.05%, 相较于不添加卷积模块,效果分别提升了 0.36 个百分点和 0.30 个百分点。最后,在 U-Net 中同时添加多尺度特征融合模块和特征放大提取模块达到了更好的效果,  $R_{10U}$  达到了 89.23%。消融实验通过在 U-Net 算法

模型中增添本文所提出的3个不同模块,效果均有不同程度的提升,展现了本文所提卷积模块、多

尺度特征融合模块以及特征放大提取模块的有效性。

表4 基于U-Net的不同模块分割性能对比

Table 4 Comparison of segmentation performance of different blocks based on U-Net

方法	$P$	$R$	$F_1$	$R_{IoU}$
仅卷积模块	93.06	94.29	93.36	88.11
仅多尺度特征融合模块	93.44	94.32	93.66	88.56
仅放大特征提取模块	93.98	94.08	93.79	88.75
卷积模块和多尺度特征融合模块	93.42	94.86	93.88	88.92
卷积模块和放大特征提取模块	93.51	94.87	93.93	89.05
多尺度特征融合模块和放大特征提取模块	93.61	95.30	94.06	89.23

## 4 结论

1) 本文以青海省三江源自然保护区为研究区,建立了一个用于训练深度学习网络模型的高原森林遥感图像数据集。针对训练深度模型数据不足的问题,提出一种将图像切割并重新组合的数据增强方法。在使用数据增强与不使用数据增强的条件下分别进行分割网络模型的训练,实验结果表明,扩充训练数据的同时强化了网络对图像细节的感知能力,针对不同模型不同程度地提高了遥感图像分割的精确度。

2) 本文针对主流分割网络处理大范围遥感图像的分割任务时存在无法聚焦图像细节的问题,提出一种融合多尺度特征的高分辨率森林遥感图像分割模型。该模型融合了所提出的卷积模块、多尺度特征融合模块与特征放大提取模块。卷积模块使用 $1 \times 1$ 卷积提取小尺度细节特征信息,基于ASPP的多尺度特征融合模块融合多尺度提取全局上下文信息,特征放大提取模块用于编解码器结构的连接部分,从不同尺度放大提取遥感图像的细节特征信息。与多种主流图像分割对比的实验结果证明本文所提出的分割模型能够对高分辨率遥感影像进行更准确的森林区域提取。

3) 本文所提出的基于数据增强与多尺度特征融合的高分辨率森林遥感图像分割方法,能够更好地服务于我国固碳减排、全球气候变化及水资源循环的相关研究。

## 参考文献:

[1] 郑冬梅. 三峡库区森林生物量和碳储量的遥感估测研究[J]. 遥感信息, 2013, 28(5): 95-98, 108.

ZHENG D M. Forest biomass and carbon estimation in Three Gorges Region based on remote sensing [J]. Remote Sensing Information, 2013, 28(5): 95-98, 108. (in Chinese)

[2] BELGIU M, DRĂGUȚ L. Random forest in remote sensing: a review of applications and future directions [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 114: 24-31.

[3] 吕杰, 郝宁燕, 李崇贵, 等. 利用随机森林和纹理特征的森林类型识别[J]. 遥感信息, 2017, 32(6): 109-114.

LV J, HAO N Y, LI C G, et al. Identification of forest type based on random forest and texture characteristics [J]. Remote Sensing Information, 2017, 32(6): 109-114. (in Chinese)

[4] 李增元, 陈尔学. 中国林业遥感发展历程[J]. 遥感学报, 2021, 25(1): 292-301.

LI Z Y, CHEN E X. Development course of forestry remote sensing in China [J]. National Remote Sensing Bulletin, 2021, 25(1): 292-301. (in Chinese)

[5] DYMOND C C, MLADENOFF D J, RADELOFF V C. Phenological differences in Tasseled Cap indices improve deciduous forest classification [J]. Remote Sensing of Environment, 2002, 80(3): 460-472.

[6] DEFRIES R S, TOWNSHEND J R G. NDVI-derived land cover classifications at a global scale [J]. International Journal of Remote Sensing, 1994, 15(17): 3567-3586.

[7] ARII M, VAN ZYL J J, KIM Y. A general characterization for polarimetric scattering from vegetation canopies [J]. IEEE Transactions on Geoscience and Remote Sensing, 2010, 48(9): 3349-3357.

[8] MELGANI F, BRUZZONE L. Classification of hyperspectral remote sensing images with support vector machines [J]. IEEE Transactions on Geoscience and

- Remote Sensing, 2004, 42(8): 1778-1790.
- [9] PAL M. Random forest classifier for remote sensing classification [J]. International Journal of Remote Sensing, 2005, 26(1): 217-222.
- [10] 王二丽,李存军,周静平,等. 基于多时相遥感影像的北京平原人工林树种分类 [J]. 北京工业大学学报, 2017, 43(5): 710-718.  
WANG E L, LI C J, ZHOU J P, et al. Classification of Beijing afforestation species based on multi-temporal images [J]. Journal of Beijing University of Technology, 2017, 43(5): 710-718. (in Chinese)
- [11] WICKRAMARATHNA S, VAN DEN HOEK J, STRIMBU B M. Automated detection of individual juniper tree location and forest cover changes using Google Earth Engine [J]. Annals of Forest Research, 2021, 64(1): 61-72.
- [12] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [13] 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.
- [14] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network [C] // Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 6230-6239.
- [15] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [16] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C] // Proceedings of the European Conference on Computer Vision. Cham: Springer, 2018: 801-818.
- [17] FU J, LIU J, TIAN H J, et al. Dual attention network for scene segmentation [EB/OL]. [2023-01-26]. <https://arxiv.org/abs/1809.02983>.
- [18] WAGNER F H, DALAGNOL R, CASAPIA X T, et al. Regional mapping and spatial distribution analysis of canopy palms in an Amazon forest using deep learning and VHR images [J]. Remote Sensing, 2020, 12(14): 2225.
- [19] 贾莉,郑柯,唐娉,等. 地形校正对 U-Net 深度神经网络分类器分类精度的影响 [J]. 遥感学报, 2022, 26(4): 698-710.  
JIA L, ZHENG K, TANG P, et al. Topographic correction effect on classification accuracy for deep neural network classifier—a case study of the U-Net model [J]. National Remote Sensing Bulletin, 2022, 26(4): 698-710. (in Chinese)
- [20] WEI Z H, JIA K B, JIA X W, et al. Mapping large-scale plateau forest in Sanjiangyuan using high-resolution satellite imagery and few-shot learning [J]. Remote Sensing, 2022, 14(2): 388.
- [21] HAN Z M, DIAN Y Y, XIA H, et al. Comparing fully deep convolutional neural networks for land cover classification with high-spatial-resolution Gaofen-2 images [J]. International Journal of Geo-Information, 2020, 9(8): 478.
- [22] OUBARA A, WU F L, AMAMRA A, et al. Survey on remote sensing data augmentation: advances, challenges, and future perspectives [C] // Advances in Computing Systems and Applications. Cham: Springer, 2022: 95-104.
- [23] OKTAY O, SCHLEMPER J, LE FOLGOC L, et al. Attention U-Net: learning where to look for the pancreas [EB/OL]. [2023-01-26]. <https://arxiv.org/abs/1804.03999>.
- [24] IBTEHAZ N, RAHMAN M S. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation [J]. Neural Networks, 2020, 121: 74-87.

(责任编辑 梁洁)