

毛 星, 金 晶, 张 欣, 等. 基于改进 DeepLab v3+ 模型和迁移学习的高分遥感耕地信息提取方法[J]. 江苏农业学报, 2023, 39(7): 1519-1529.

doi: 10.3969/j.issn.1000-4440.2023.07.009

基于改进 DeepLab v3+ 模型和迁移学习的高分遥感耕地信息提取方法

毛 星^{1,2}, 金 晶^{1,2}, 张 欣^{1,2}, 戴佩玉^{1,2}, 任 妮^{1,2}

(1. 江苏省农业科学院农业信息研究所, 江苏 南京 210014; 2. 农业农村部长三角智慧农业技术重点实验室, 江苏 南京 210014)

摘要: 针对实际遥感耕地信息提取工作中, 多源数据特征复杂、样本标注工作繁重等导致高空间分辨率影像解译精度不高、自动化程度不够的问题, 本研究基于 DeepLab v3+ 模型, 提出一种融合邻域边缘加权模块 (NEWM) 和轴向注意力机制模块 (CBAM-s) 的卷积网络模型 DEA-Net, 结合迁移学习方法进行高分分辨率遥感影像耕地信息提取。首先, 在浅层网络结构中加入邻域边缘加权模块, 提升高分分辨率下地物的连续性, 细化边缘分割粒度; 其次, 在深层网络结构中添加轴向注意力机制模块, 增加细小地物的关注权重, 减少深度卷积导致地物丢失的情况; 最后, 采用迁移学习的思想, 降低样本标注工作量, 提高模型学习能力。利用高分卫星土地覆盖数据集 (GID) 数据构建源域数据集进行模型预训练, 将获取的模型参数及权重信息迁移至大数据与计算智能大赛 (BDCI) 遥感影像地块分割竞赛数据集和全国人工智能大赛 (NAIC) 遥感影像数据集制作的 2 种不同目标域数据集中, 微调训练后应用于耕地信息提取研究。结果表明, 本研究构建方法能够增强模型的空间细节学习能力, 提高耕地语义分割精度的同时, 降低 2/3 以上的训练样本数量, 为遥感耕地信息提取及农业数据智能化利用提供新的思路和方法。

关键词: 耕地信息提取; 迁移学习; DEA-Net; 高分遥感

中图分类号: S126 **文献标识码:** A **文章编号:** 1000-4440(2023)07-1519-11

High-resolution remote sensing arable land information extraction method based on improved DeepLab v3+ model and transfer learning

MAO Xing^{1,2}, JIN Jing^{1,2}, ZHANG Xin^{1,2}, DAI Pei-yu^{1,2}, REN Ni^{1,2}

(1. Institute of Agricultural Information, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China; 2. Key Laboratory of Intelligent Agricultural Technology (Changjiang Delta), Ministry of Agriculture and Rural Affairs, Nanjing 210014, China)

Abstract: The complex multi-source data features and heavy sample annotation work in the practical remote sensing arable land information extraction work will lead to low accuracy and insufficient automation of high spatial resolution image interpretation. In view of the above problems, based on DeepLab v3+, we proposed a convolutional network model DEA-Net that incorporated the neighborhood edge weighting module (NEWM) and the axial attention mechanism (CBAM-s), and

combined the transfer learning method to extract arable land information of high-resolution remote sensing images. First, the NEWM was added to the shallow network structure to improve the continuity of features under high resolution and refine the granularity of edge segmentation. Then, the CBAM-s was added to the deep network structure to increase the attention weight of fine features and reduce the loss of features due to deep convolution. Finally,

收稿日期: 2022-09-15

基金项目: 高分辨率对地观测系统重大专项 (74-Y50G12-90-01-22/23)

作者简介: 毛 星 (1991-), 男, 江苏镇江人, 硕士, 工程师, 主要从事农业遥感数据分析与利用研究。(E-mail) maoxing@jaas.ac.cn

通讯作者: 任 妮, (E-mail) rn@jaas.ac.cn

the idea of transfer learning was adopted to reduce the sample annotation workload and improve the learning ability of the model. The source domain dataset was constructed using the Gaofen image dataset (GID) for model pre-training, and the acquired model parameters and weight information were migrated to two different target domain datasets produced by big data & computing intelligence contest (BDCI) and national artificial intelligence challenge (NAIC), and fine-tuned and trained for arable land information extraction. The results showed that the method constructed in this study could enhance the spatial detail learning ability of the model, improve the semantic segmentation accuracy of arable land, and reduce the number of training samples by more than 2/3. It can provide new ideas and methods for remote sensing arable land information extraction and intelligent utilization of agricultural data.

Key words: arable land information extraction; transfer learning; DEA-Net; high resolution remote sensing

及时、准确获取耕地信息对于国家粮食安全、农业生产管理、农业资源保护等具有重要意义^[1]。随着农业现代化的不断发展,中低分辨率耕地影像已无法满足精准农业对数据质量的要求。虽然遥感影像的分辨率越来越高,但随之而来的复杂空间特征信息对遥感影像的解译技术提出了更高要求。

卷积神经网络出现之前,遥感地物分类提取方法多依赖于分类特征设计和机器学习模型^[2]。光谱信息和植被指数是耕地识别中最常见的特征参数^[3],但不同波段组合、不同植被指数组成的分类特征使耕地识别结果具有很大随机性。基于支持向量机(SVM)^[4]、决策树^[5]、随机森林^[6]等分类器构建的机器学习模型虽然降低了分类特征设计过程中人为因素的影响,但模型的鲁棒性和灵活性存在一定限制,改选数据源或者切换应用区域,均需对模型的参数及输入特征做出调整。

近年来,深度学习技术的快速发展推动了遥感语义分割的广泛应用。Long 等^[7]提出了 FCN 网络,用卷积层替换最后一层全连接层,实现了像素级的图像分类,但 FCN 在上采样过程中忽略了像素之间的关系。UNet 模型采用编码器—解码器结构,并在上采样过程中充分利用下采样提取的特征信息,提高了语义分割精度^[8],但没有考虑图像的上下文信息。Chen 等^[9]引入了空洞金字塔池化模块来增加网络的上下文信息感受野,但空洞卷积会丢失部分空间信息,导致影像分割不够精细。Yao 等^[10]在 DeepLab v3+ 的基础上引入了注意力机制,增强了特征提取能力,降低了环境信息的影响。Wang 等^[11]基于 DeepLab v3+ 提出了一种名为 CFAMNet 的网络模型,使用类特征注意模块增强了类与类之间的相关性,采用多并行空洞金字塔池化结构增强了空间相关性。刘文祥等^[12]采用双注意力机制,从位置和通道 2 个角度增强特征信息的关注度,提升了模型

的训练速度。叶沅鑫等^[13]采用 R-MCN 网络结构,增大训练的感受野,增加了浅层网络特征权重,进一步细化了地物边界。虽然,基于 DeepLab v3+ 的遥感影像语义分割技术已经有了一定提高,但大多数研究针对的目标数据都是具有相同参数的少量遥感影像,模型泛化能力有限。多源遥感数据不同的成像方式、成像时间、空间分辨率、光谱波段范围等对模型的鲁棒性提出了较高要求,作物类型、灌溉方式、地形地貌、休耕轮作等也是遥感耕地自动化提取中不可避免的难点问题。

基于此,本研究拟提出一种融合邻域边缘加权模块(NEWM)和轴向注意力机制模块(CBAM-s)的高分辨率遥感耕地信息提取方法,利用邻域边缘加权模块改善浅层网络中地物的连续性,用轴向注意力机制模块提高深层网络中细小地物的关注度,基于迁移学习的方法优化模型参数,以期在增强语义分割效果的同时,能够进一步提高模型的泛化能力。

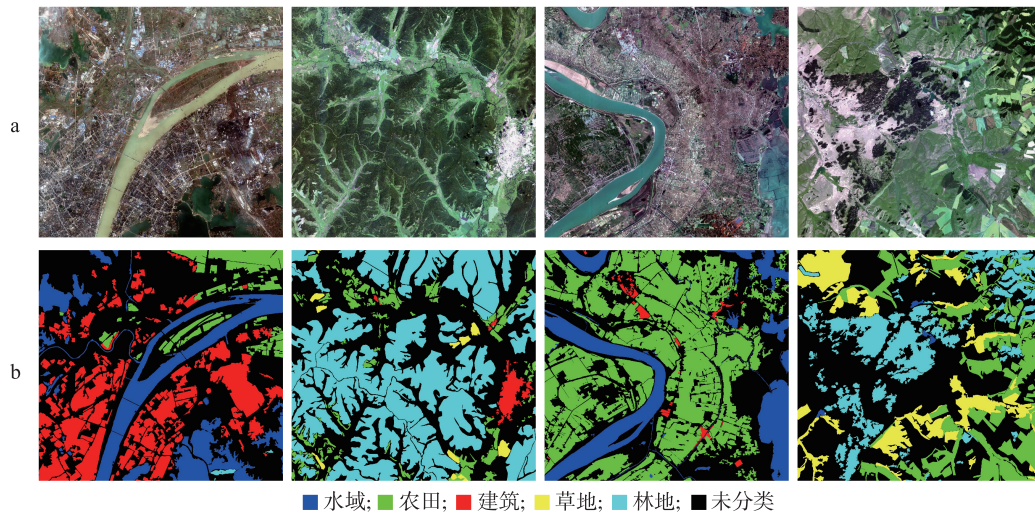
1 材料与方法

1.1 数据集制作

本研究试验数据包括两部分,第一部分为源域数据集,用于模型的预训练,该数据集来自于武汉大学公布的高分卫星土地覆盖数据集(GID),包含了 150 张像素为 7 200×6 800 的高分 2 号卫星影像,涵盖了 5 种地物类别:水域、农田、建筑、草地和林地,具体见图 1。因为研究对象是耕地,所以选择与之相似的农田地块作为模型预训练对象。为方便模型的训练,使用像素为 256×256 的随机窗口从原始影像中裁剪出 30 000 张包含农田地块的样本图像,将分类标注数据二值化为农田和其他类别,作为源域数据集,并按照 8:1:1 的比例随机划分为训练数据、验证数据和测试数据。第二部分为目标域数据集,数据来源于大数据与计算智能大赛(BDCI)遥感影像地块分割竞赛数

据集(后文统称为 BDCI 数据集)和全国人工智能大赛(NAIC)遥感影像数据集(后文统称为 NAIC 数据集),这 2 个数据集的图像大小均为 256×256 (图 2)。BDCI 数据集是分辨率为 2.0 m 的遥感图像,包含 7 类地物:建筑、耕地、林地、水体、道路、草地和其他地物。NAIC 数据集由分辨率 0.1 m 至 4.0 m 的多种卫

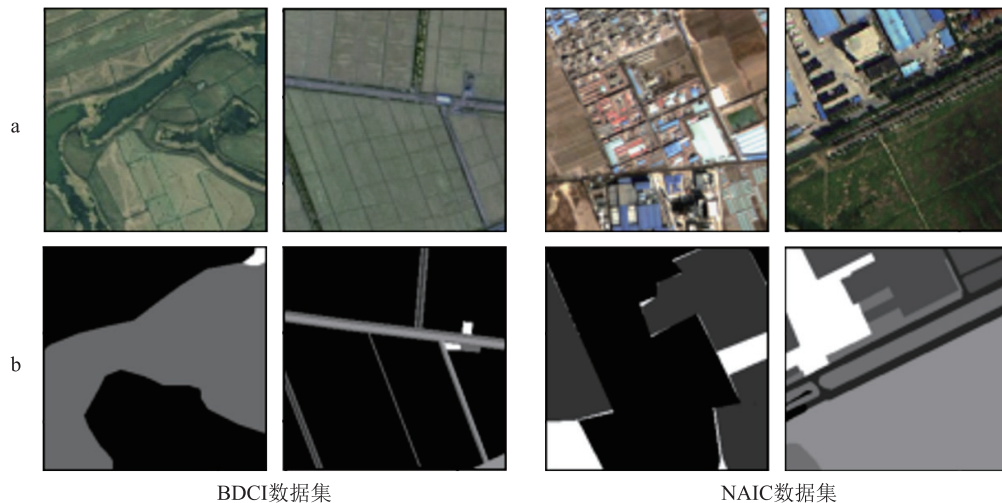
星及航空数据组成,包含 8 类地物:水体、道路、建筑、耕地、草地、林地、裸土和其他地物。分别从 2 个数据集中随机选取 8 000 张包含耕地的图像,将分类标注数据二值化为耕地和其他类别,作为目标域数据集,并按照 2 : 1 : 1 的比例随机划分为训练数据、验证数据和测试数据。



a:原始遥感影像;b:分类标注数据。

图 1 高分卫星土地覆盖数据集(GID)部分样本

Fig.1 The part of the Gaofen image dataset(GID)



a:原始遥感影像;b:分类标注数据。

图 2 BDCI 和 NAIC 数据集部分样本

Fig.2 The part of the big data & computing intelligence contest (BDCI) and national artificial intelligence challenge (NAIC) datasets

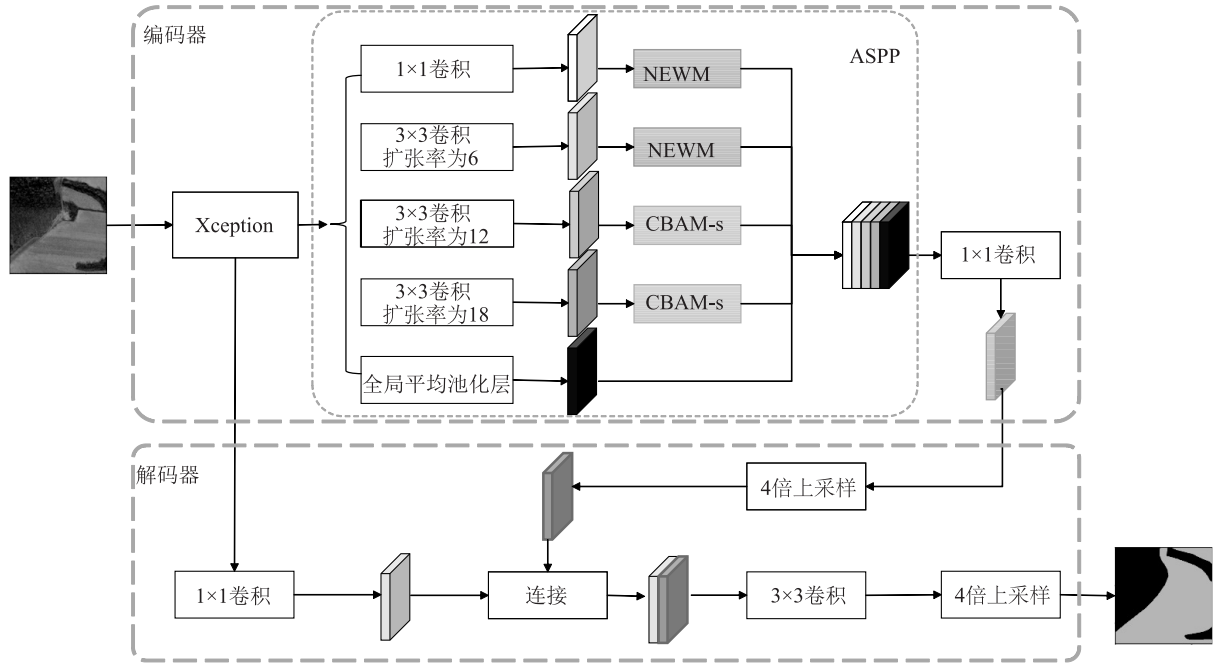
1.2 DeepLab v3+模型改进

本研究提出了一种改进的 DeepLab v3+模型,名为 DEA-Net 模型,整体采用编码器—解码器结构。图 3 显示,编码器部分主要包含 Xception 特征提取网络

和空洞金字塔池化(ASPP)模块,解码器通过 2 次 4 倍上采样得到地物的分割形状。ASPP 模块包含 1 个 1×1 卷积和 3 个扩张率分别为 6、12、18 的 3×3 卷积以及 1 个全局平均池化层。不同扩张率的空洞卷积可

以有效捕获影像上下文特征信息,提高地物不同尺度的识别效果^[14]。不同尺度影像特征包含不同的空间信息,高分辨率对应的浅层特征具有丰富的空间细节信息,有利于耕地边界细化,但是面对遥感影像同物异谱的情况,容易出现噪声点^[15]。低分辨率对应的深层特征包含丰富的全局语义信息^[16],能够有效识别耕地的大致范围,但下采样过程中容易导致沟壑、

田埂等信息的丢失。因此,本研究设计了邻域边缘加权模块和轴向注意力机制模块,对模型的 ASPP 模块进行改进,增强模型特征提取能力,提高遥感影像的耕地信息提取精度。NEWM 位于 ASPP 模块前两层卷积之后,主要用于改善高分辨率情况下地物的连续性,CBAM-s 则嵌入 ASPP 中扩张率为 12、18 的 3×3 卷积之后,以增加深层特征对周边要素的考虑。



Xception:特征提取网络;ASPP:空洞金字塔池化;NEWM:邻域边缘加权模块;CBAM-s:轴向注意力机制模块。

图3 DEA-Net 模型整体结构

Fig.3 The overall structure of DEA-Net model

1.2.1 邻域边缘加权模块 卷积神经网络中的低级特征容易忽略空间细节^[17],导致语义分割的边缘模糊不清。为此,本研究设计了邻域边缘加权模块,增强耕地边缘权重,提高空间特征关注度,改善耕地边缘识别效果。图 4 显示,首先按照栅格图像 8 邻域规则,创建 8 个不同方向的卷积核,卷积核当前方向值为 1,其余位置为 0,然后对输入特征进行 8 邻域二维卷积,卷积结果与输入特征相减后叠加,最后经过非线性激活生成邻域边缘权重特征。NEWM 的计算公式可表示为:

$$Ne(M) = \sigma \left\{ \sum_{i=0}^7 [E_i(M) - M] \right\} \otimes M \quad (1)$$

其中, M 表示输入特征; $Ne(M)$ 表示邻域边缘加权后的输出特征; E_i 表示 8 邻域二维卷积; σ 表示 Sigmoid 函数; \otimes 表示矩阵的元素乘法运算。

1.2.2 轴向注意力机制模块 注意力机制的本质就是矩阵乘法运算^[11],神经网络模型中的注意力模块可以看作是训练的加权图,通过计算像素点不同特征信息的权重,捕捉关注信息,增强网络理解能力。为提高相邻耕地之间的田埂、沟壑等细小地物的区分效果,受卷积块注意模块(CBAM)的启发^[18],本研究提出了轴向注意力机制模块,即 CBAM-s。图 5 显示,CBAM-s 通过一维通道注意力和二维空间注意力提升特征权重。通道注意力使神经网络更加关注各栅格单元哪些特征是重要信息,空间注意力则从 X 和 Y 2 个轴向考虑重要特征信息在影像中的空间位置。对于输入特征 $M \in \mathbb{R}^{C \times H \times W}$ (其中 C 表示通道数, $H \times W$ 为影像尺寸),CBAM-s 的计算公式可表示为:

$$M' = A_c(M) \otimes M \quad (2)$$

$$M'' = A_s(M') \otimes M' \quad (3)$$

其中, $A_c \in \mathbb{R}^{C \times 1 \times 1}$ 表示一维通道注意力; $A_s \in \mathbb{R}^{1 \times H \times W}$ 表示二维空间注意力; \otimes 表示矩阵的元素乘

法运算; M' 表示通过通道注意力的特征图; M'' 是最终输出的注意力特征图。

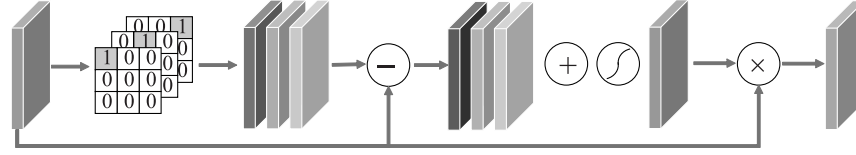


图 4 邻域边缘加权模块结构

Fig.4 Neighborhood edge weighting module structure

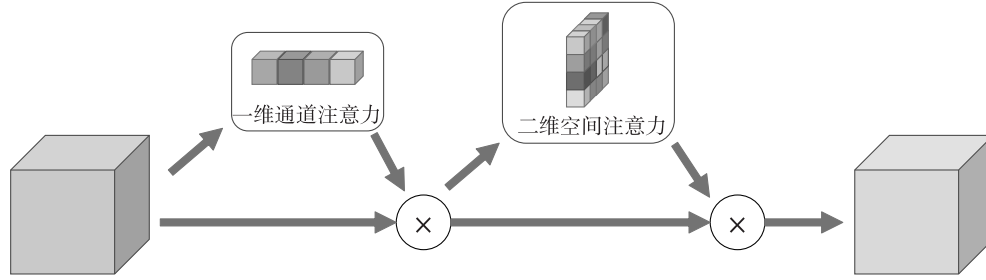


图 5 轴向注意力机制模块结构

Fig.5 The axial attention mechanism module structure

1.2.2.1 一维通道注意力 通道注意力对栅格影像的每个波段赋予一个权重,并且这些权重在训练的过程中可以自动校正^[19],这些权重特征可以被后续网络转换利用以提高卷积通道之间资源分配的粒度。图 6 显示,首先,使用全局平均池化和最大池化操作来聚合空间维度信息,生成 2 个通道描述一维矩阵,然后通过多层感知器 (MLP),最终将 2 个通道描述相加实现非线性输出,通道注意力的计算公式可以表示为:

$$A_c(M) = \sigma \{ MLP [Pool(M)_{Avg}] + MLP [Pool(M)_{Max}] \} \quad (4)$$

其中, $A_c(M)$ 表示通道注意力计算函数; M 表示输入特征; $Pool_{Avg}$ 和 $Pool_{Max}$ 分别表示全局平均池化和最大池化; σ 表示 Sigmoid 函数; MLP 表示多层感知函数,包含的单个隐藏层可以将 $A_c \in \mathbb{R}^{C \times 1 \times 1}$ 压缩成 $A_c \in \mathbb{R}^{r \times 1 \times 1}$, 本研究中 r 值为 16。

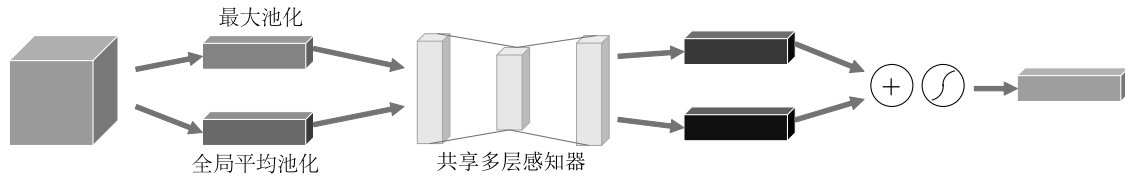


图 6 通道注意力模块结构

Fig.6 The channel attention module structure

1.2.2.2 二维空间注意力 CBAM 的空间注意力对每个像素的所有通道进行平均池化和最大池化,保留了像素的空间位置信息。本研究将注意力模块嵌入 ASPP 模块中 3×3 卷积的深层特征,增加了神经网络对周边要素的考虑。然而深层特征的低空间分辨率很容易忽略影像中的琐碎图斑。为此,本研究

从 X 和 Y 2 个轴向逐像素增强特征,空间注意力的计算公式如下:

$$A_s(M_c) = \sigma \{ Conv [XY(M_c)_{Avg} + XY(M_c)_{Max}] \} \quad (5)$$

其中, $A_s(M_c)$ 表示空间注意力计算函数; M_c 表示通道注意力特征; Avg 和 Max 分别表示平均池化和最大池化; σ 表示 Sigmoid 函数; $Conv$ 表示卷积运

算; $XY(M_C)$ 为轴向增强函数。

图 7 显示,轴向增强函数首先将输入特征 $M_C \in \mathbb{R}^{C \times H \times W}$ 拆分成 2 个轴向进行平均池化得到 $X_{Avg} \in \mathbb{R}^{C \times 1 \times W}$ 和 $Y_{Avg} \in \mathbb{R}^{C \times H \times 1}$,然后将拼接结果压缩得到特征

图 $Mer \in \mathbb{R}^{\frac{C}{r} \times 1 \times (H+W)}$, r 值为 16,接着再分割为 2 个特征图 $Xat \in \mathbb{R}^{\frac{C}{r} \times 1 \times W}$ 、 $Yat \in \mathbb{R}^{\frac{C}{r} \times H \times 1}$ 进行卷积和 Sigmoid 激活,最后通过元素乘法运算得到轴向增强的特征信息。

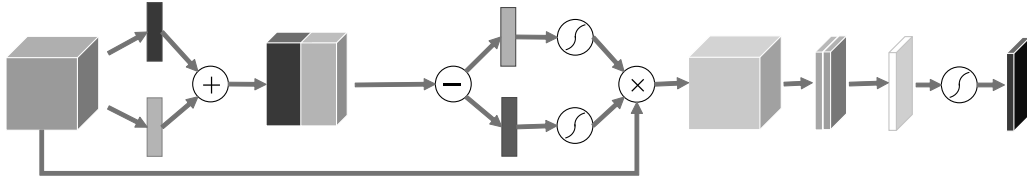
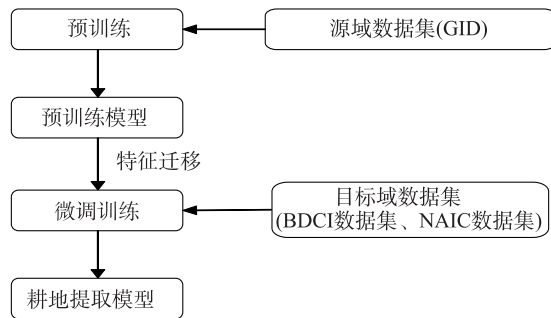


图 7 空间注意力模块结构

Fig.7 The spatial attention module structure

1.3 迁移学习

针对样本数量有限的条件下,高分辨率遥感耕地信息提取模型存在收敛缓慢、极易过拟合等问题,迁移学习可以使用少量训练样本创建特征学习能力更强的深度学习模型,最大限度减少目标领域学习成本^[20]。对于给定源域(D_S)和学习任务(T_S)以及目标域(D_T)和学习任务(T_T),迁移学习能够利用在 D_S 和 T_S 中学习的知识帮助改进目标预测函数(F_T)在 D_T 中的使用。图 8 显示,本研究将源域数据集上预训练获得的特征权重和参数信息迁移至新的网络模型,新模型最大程度复用预训练模型的网络结构和权重参数,通过目标域数据集微调训练模型参数学习隐含特征,进而提高模型性能。



GID: 高分卫星土地覆盖数据集; BDCI 数据集: 大数据与计算智能大赛遥感影像地块分割竞赛数据集; NAIC 数据集: 全国人工智能大赛遥感影像数据集。

图 8 迁移学习流程

Fig.8 The process of transfer learning

1.4 精度评价方法

采用像素准确率(PA)、类别平均像素准确率

(MPA)、平均交并比($MIoU$)和 Kappa 系数作为精度评价标准。 PA 、 MPA 和 $MIoU$ 是图像分割评价的重要指标,计算公式如下:

$$PA = \frac{\sum_{i=1}^k P_{ii}}{\sum_{i=1}^k \sum_{j=1}^k P_{ij}} \quad (6)$$

$$MPA = \frac{1}{k} \sum_{i=1}^k \frac{P_{ii}}{\sum_{j=1}^k P_{ij}} \quad (7)$$

$$MIoU = \frac{1}{k} \sum_{i=1}^k \frac{P_{ii}}{\sum_{j=1}^k P_{ij} + \sum_{j=1}^k (P_{ji} - P_{ii})} \quad (8)$$

其中, PA 表示像素准确率; MPA 表示类别平均像素准确率; $MIoU$ 表示平均交并比; k 表示类别数量; p_{ii} 表示预测正确的像素数量; p_{ij} 表示将 i 类预测为 j 类的像素数量; p_{ji} 表示将 j 类预测为 i 类的像素数量。

Kappa 系数是遥感分类中最常用的评价指标,计算公式如下:

$$Kappa = \frac{N \sum_{i=1}^k N_{ii} - \sum_{i=1}^k N_{i+} N_{+i}}{N^2 - \sum_{i=1}^k N_{i+} N_{+i}} \quad (9)$$

其中, $Kappa$ 表示 Kappa 系数; N 表示样本像素总数; k 为类别数量; N_{i+} 表示第 i 类的真实样本像素数量; N_{+i} 表示预测为第 i 类的样本像素数量; N_{ii} 表示第 i 类真实值和预测值一致的样本像素数量。

1.5 试验配置

本研究的试验在 Linux 操作系统中运行,深度学习框架采用 Pytorch 1.11,处理器为 Intel(R) Xeon(R) Gold 5218R CPU @ 2.10 GHz,显卡为 NVIDIA A100-PCIE-40 GB。Python 版本为 3.8,CUDA 版本为 11.0。使用 Adam 函数作为网络优化器,初始学习率为 0.001,单次训练的样本个数为 64。

2 结果与分析

2.1 不同网络在源域数据集上的对比

为了对比模型的特征学习能力,本研究选择 SegNet、UNet、DeepLab v3+网络模型进行对比试验。每种模型均利用源域数据集的训练数据和验证数据进行特征学习,并在测试数据上对其进行语义分割试验以及精度评价。结果(表 1)表明,DEA-Net 模型相比其他网络模型的 PA 、 MPA 、 $MIoU$ 和 Kappa 系数均有所提高, PA 、 MPA 、 $MIoU$ 分别为 88.49%、86.44%、80.11%,Kappa 系数为 0.833 7。与 DeepLab v3+模型相比,DEA-Net 模型的 PA 提高了 1.59 个百分点, MPA 提高了 0.87 个百分点, $MIoU$ 提高了 1.61 个百分点,Kappa 系数提高了 0.013 8,说明本研究提出的 DEA-Net 模型在农田遥感影像语义分割中具有一定优势。

图 9 展示了不同模型的语义分割情况,SegNet

模型在上采样和下采样过程中未曾关联上下文信息,容易丢失空间特征,造成大片农田无法识别。UNet 模型融合了编码、解码过程的特征,其精度有所提高,但对相似区域仍存在大量错分、漏分情况。DeepLab v3+模型中 ASPP 模块的不同空洞卷积增加了多尺度特征关联,其结果与本研究提出的 DEA-Net 模型最接近。DEA-Net 模型提取的农田边界信息比其他网络模型更加清晰,结果更接近真实标注。

表 1 不同模型在源域数据集上的精度结果

Table 1 Accuracy of different models on source domain dataset

模型	PA (%)	MPA (%)	$MIoU$ (%)	Kappa 系数
SegNet	85.12	83.62	73.55	0.784 0
UNet	86.20	84.89	76.08	0.798 9
DeepLab v3+	86.90	85.57	78.50	0.819 9
DEA-Net	88.49	86.44	80.11	0.833 7

PA :像素准确率; MPA :类别平均像素准确率; $MIoU$:平均交并比。

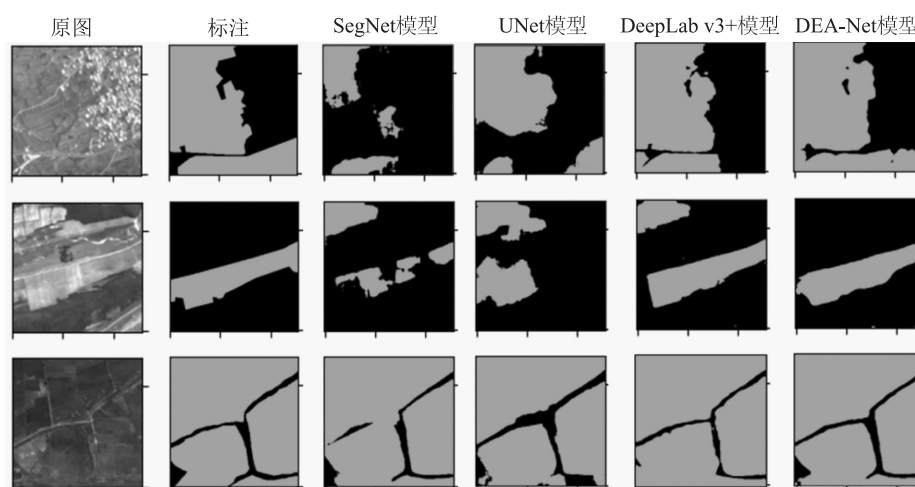


图 9 不同模型在源域数据集上的提取效果

Fig.9 Extraction effects of different models on the source domain dataset

2.2 模型在源域数据集上的消融分析

为了研究 NEWM 和 CMAM-s 对 DEA-Net 模型的影响,本研究在源域数据集上进行了消融试验。采用 DeepLab v3+模型为基准模型,逐步添加 NEWM 和 CMAM-s,验证各模块的效果。从表 2 可以看出各模块对模型精度的贡献程度,CMAM-s 对模型学习能力的提升效果优于 NEWM。从提取结果中选取了 3 张代表性影像说明各模块发挥的作用,图 10 显示,第一张影像引入 NEWM 后,细条状农田被清晰识别出,且农田的边界轮廓鲜明;第二张影像提取结果显示,

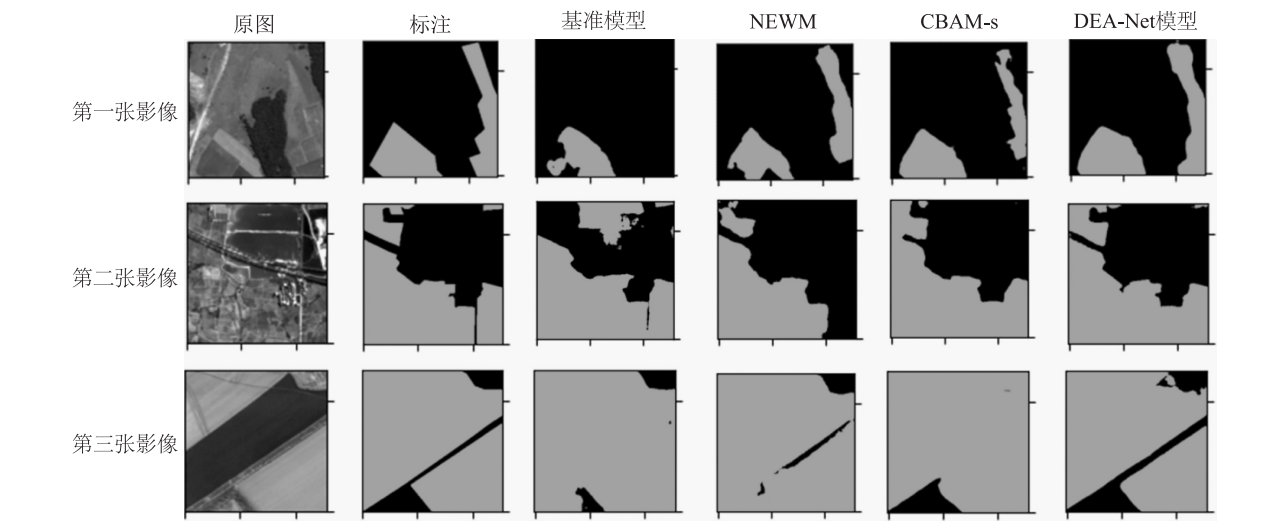
CMAM-s 能够精准确定农田的大致分布范围;第三张影像的提取结果可以看出 NEWM 提高了田埂、道路的识别精度,但有部分非农田被误认为是农田。CMAM-s 可以提高连片农田分割效果,但无法区分农田中细长的其他地物,融合 NEWM 和 CMAM-s 2 个模块的 DEA-Net 模型提取的农田信息基本接近实际标注范围。消融试验结果表明 NEWM 和 CMAM-s 均能有效提高模型的农田识别精度,但部分场景下 2 个独立模块的提取结果并不理想,而 DEA-Net 模型能够融合 2 个模块的优势,提高模型精度。

表 2 不同模块的消融结果

Table 2 Results of ablation experiments with different modules

基准模型	NEWM	CMAM-s	PA (%)	MPA (%)	MIoU (%)	Kappa 系数
√			86.90	85.57	78.50	0.819 9
√	√		87.32	85.94	79.23	0.824 6
√		√	87.86	86.15	79.37	0.828 9
√	√	√	88.49	86.44	80.11	0.833 7

基准模型:DeepLab v3+模型;NEWM:邻域边缘加权模块;CBAM-s:轴向注意力机制模块;PA:像素准确率;MPA:类别平均像素准确率;MIoU:平均交并比。



基准模型:DeepLab v3+模型;NEWM:邻域边缘加权模块;CBAM-s:轴向注意力机制模块。

图 10 不同模块的消融试验效果

Fig.10 Effects of ablation experiments with different modules

2.3 目标域数据集中的耕地信息提取结果

本研究利用 2 种目标域数据集分别进行耕地信息提取试验,将源域数据集上训练得到的模型迁移到目标域数据集上,通过 20 次迭代训练并在目标域数据集的测试数据中进行模型效果验证。本研究对比了 DEA-Net 模型和 DeepLab v3+模型在目标域数据集上的精度,结果(表 3)表明,BDCI 数据集中 DEA-Net 模型迁移学习效果优于 DeepLab v3+模型,经过迁移学习 DEA-Net 模型的 PA、MPA、MIoU、Kappa 系数分别提高了 1.55 个百分点、1.18 个百分点、2.15 个百分点和 0.031 2。NAIC 数据集中 2 种模型经过迁移学习后的精度提升效果更加明显,DeepLab v3+模型的 PA、MPA、MIoU、Kappa 系数分别提高了 2.02 个百分点、2.43 个百分点、2.12 个百分点和 0.043 5,DEA-Net 模型的 PA、MPA、MIoU、Kappa 系数分别提高了 2.26 个百分点、1.36 个百分

点、4.47 个百分点和 0.039 5,DEA-Net 模型的 PA 和 MIoU 提升效果更明显,说明 DEA-Net 模型对于目标耕地信息的提取效果更好。

为了更直观地展示不同方法的耕地信息提取效果,从目标域数据集中选取 4 个代表性场景,场景一、场景二来自 NAIC 数据集,场景三、场景四来自 BDCI 数据集。图 11 显示,场景一的耕地中间分布了一片规则房区,场景二中耕地被不规则地物隔开,场景三由耕地和林地环绕分布,场景四拥有 2 块不同形态的耕地。DEA-Net 模型的邻域边缘加权模块能够有效提高边缘精度,不论是针对规则边缘还是不规则边缘都能较好地展示细节信息。注意力机制可以加速不同特征信息学习,提高“异物同谱、同物异谱”地物的分类准确性。此外,经过迁移学习的 DEA-Net 模型和 DeepLab v3+模型的识别效果均明显提高,但 DEA-Net 模型的耕地分割结果与实际情况更为接近。

表 3 不同方法在目标域数据集上的精度结果

Table 3 Accuracy of different methods on target domain dataset

目标域数据集	网络结构	训练方式	<i>PA</i> (%)	<i>MPA</i> (%)	<i>MIoU</i> (%)	Kappa 系数
BDCI 数据集	DeepLab v3+	迁移学习	86.52	84.17	81.57	0.698 4
		无迁移学习	85.61	83.99	80.09	0.683 3
	DEA-Net	迁移学习	86.95	84.81	82.04	0.709 0
		无迁移学习	85.40	83.63	79.89	0.677 8
NAIC 数据集	DeepLab v3+	迁移学习	83.96	84.00	74.67	0.675 7
		无迁移学习	81.94	81.57	72.55	0.632 2
	DEA-Net	迁移学习	84.59	84.01	76.43	0.684 5
		无迁移学习	82.33	82.65	71.96	0.645 0

PA:像素准确率;*MPA*:类别平均像素准确率;*MIoU*:平均交并比。BDCI 数据集:大数据与计算智能大赛遥感影像地块分割竞赛数据集;NAIC 数据集:全国人工智能大赛遥感影像数据集。

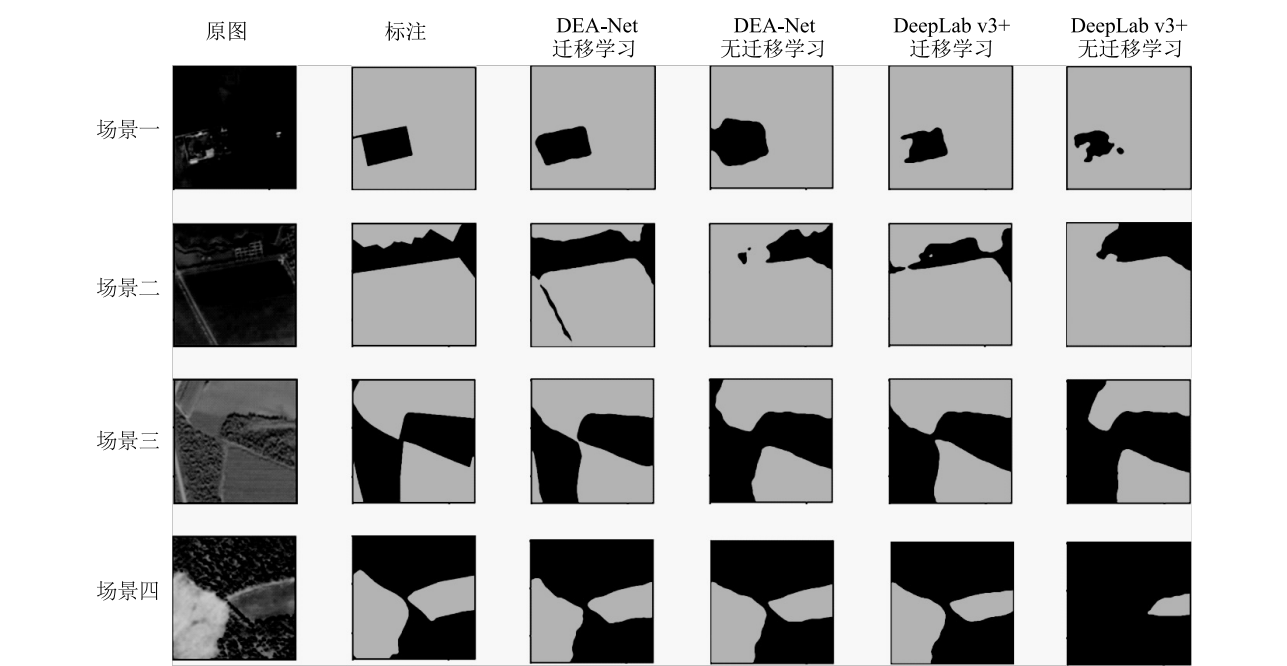


图 11 不同方法在 4 种场景上的耕地信息提取结果

Fig.11 Results of different methods for arable land information extraction on four scenarios

2.4 训练样本数量分析

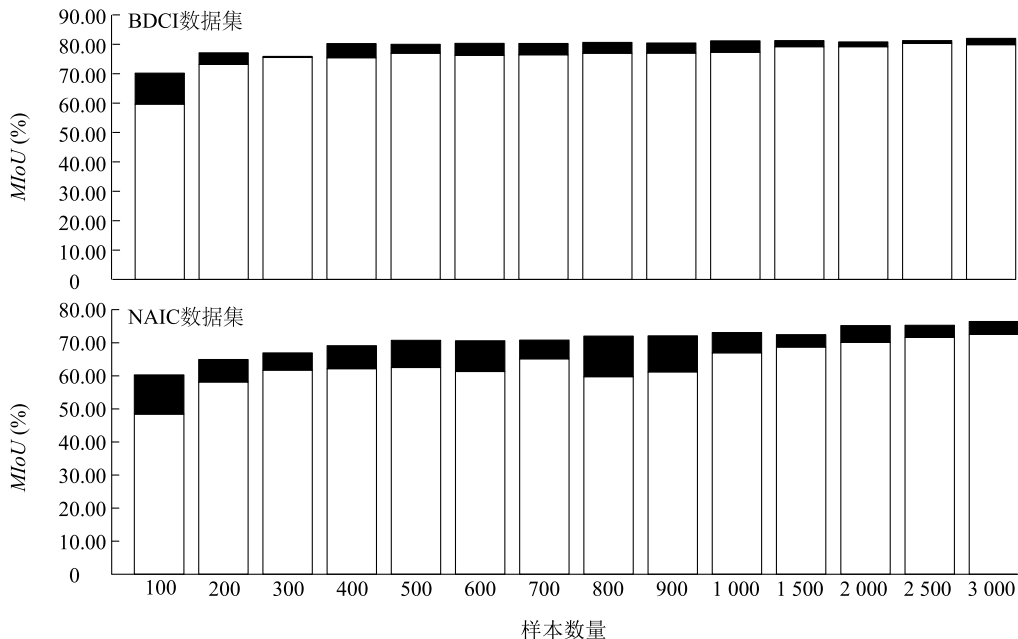
为了进一步验证 DEA-Net 模型在目标域数据集中的适用性,将目标域训练数据从 100 张逐步增加至3 000张,验证数据和测试数据不变,训练迭代次数设为 20 次。图 12 展示了 BDCI 数据集和 NAIC 数据集中 DEA-Net 模型的 *MIoU* 随训练样本数量的变化。随着样本数量的增加,模型的精度整体呈增加趋势。在 BDCI 数据集中,经过迁移学习后模型的 *MIoU* 达到 80%时,样本为 400 张影像,而未经过迁移学习的模型在1 500张样本以后 *MIoU* 才超过 79%。在 NAIC 数据集中,经过迁移学习的模型在

样本超过 500 张后 *MIoU* 达 70%以上,未经过迁移学习的模型则需要2 000张样本。因此,本研究提出的方法可以有效减少高分辨率遥感耕地分割的样本数量。

此外,在 BDCI 数据集中,样本数量超过 400 张以后,模型精度的提升幅度总体呈减小趋势。在 NAIC 数据集中,模型精度随训练样本数量的增加总体呈增加趋势。主要原因是 BDCI 数据集都是分辨率相同的遥感图像,只需少量训练样本就能覆盖大部分样本类型。NAIC 数据集包含了高分一号、高分二号、高分六号、高景二号、北京二号以及航空遥感

等多源影像,其空间分辨率、光谱载荷、成像方式等差异性较大,样本类型更加丰富,少量的样本难以覆盖全部样本类型。因此,迁移学习虽然能够降低训

练样本数量,但面对复杂多样的实际应用仍然需要一定规模的样本数据以扩大模型的适用范围。



□ 未经过迁移学习的模型平均交并比; ■ 经过迁移学习模型提高的平均交并比

MIoU: 平均交并比。BDCI 数据集: 大数据与计算智能大赛遥感影像地块分割竞赛数据集; NAIC 数据集: 全国人工智能大赛遥感影像数据集。

图 12 BDCI、NAIC 数据集中训练样本数量与 MIoU 关系

Fig.12 The relationship between the number of training samples and mean intersection over union (MIoU) in BDCI and NAIC datasets

3 结论

本研究以 DeepLab v3+ 模型为基础,引入邻域边缘加权模块和轴向注意力机制模块,提出了一种 DEA-Net 网络结构,有效提高了多尺度影像特征信息提取能力,并借助迁移学习的思想,在提高模型精度的同时,降低样本标注成本。研究结果表明,本研究提出的 DEA-Net 模型在源域数据集的语义分割中相比其他几种主流网络模型表现更好,PA、MPA、MIoU 分别为 88.49%、86.44%、80.11%,Kappa 系数为 0.833 7。NEWM 可以提高边缘细节关注度,CBAM-s 可以为深层网络提供更加丰富的空间特征,能够为耕地信息提取提供更多的语义信息。本研究提出的方法能够有效解决遥感耕地信息提取中样本数量不足的问题,减少 2/3 以上的样本标注工作。

虽然本研究提出的方法可以有效应用于遥感影像耕地信息提取工作,但仍有改进的空间。首先,本

研究仅利用红、绿、蓝 3 个波段数据,并未完全发挥遥感影像的多光谱、高光谱优势,未来可从光谱维度方面考虑如何进一步提高模型精度。其次,本研究并未对耕地种植类型开展细化分类,未来可围绕作物种植类型识别进行相关研究。

参考文献:

- [1] 唐华俊. 农业遥感研究进展与展望[J]. 农学学报, 2018, 8(1): 167-171.
- [2] 冯权波,牛博文,朱德海,等. 土地利用/覆被深度学习遥感分类研究综述[J]. 农业机械学报, 2022, 53(3): 1-17.
- [3] ARANGO R B, DÍAZ I, CAMPOS A, et al. Automatic arable land detection with supervised machine learning[J]. Earth Science Informatics, 2016, 9: 535-545.
- [4] 陈杰,陈铁桥,刘慧敏,等. 高分辨率遥感影像耕地分层提取方法[J]. 农业工程学报, 2015, 31(3): 190-198.
- [5] 朱江洪,李江风,叶菁. 利用决策树工具的土地利用类型遥感识别方法研究[J]. 武汉大学学报(信息科学版), 2011, 36(3): 301-305.
- [6] NITZE I, BARRETT B, CAWKWELL F. Temporal optimisation of

- image acquisition for land cover classification with random forest and MODIS time-series[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2015, 34: 136-146.
- [7] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. New York: IEEE Press, 2015.
- [8] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]. Munich: Springer International Publishing, 2015.
- [9] CHEN L, PAPANDREOU G, KOKKINOS I, 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.
- [10] YAO X, GUO Q, LI A. Light-weight cloud detection network for optical remote sensing images with attention-based DeepLab v3+ architecture[J]. *Remote Sensing*, 2021, 13. DOI:org/10.3390/rs13183617.
- [11] WANG Z, WANG J, YANG K, et al. Semantic segmentation of high-resolution remote sensing images based on a class feature attention mechanism fused with DeepLab v3+[J]. *Computers & Geosciences*, 2022, 158. DOI:org/10.1016/j.cageo.2021.104969.
- [12] 刘文祥,舒远仲,唐小敏,等. 采用双注意力机制 DeepLab v3+ 算法的遥感影像语义分割[J]. *热带地理*, 2020, 40(2): 303-313.
- [13] 叶沅鑫,谭 鑫,孙苗苗,等. 基于增强 DeepLab v3 网络的高分辨率遥感影像分类[J]. *测绘通报*, 2021(4): 40-44.
- [14] WANG Y, GAO L, HONG D, et al. Mask DeepLab: end-to-end image segmentation for change detection in high-resolution remote sensing images[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2021, 104. DOI:org/10.1016/j.jag.2021.102582.
- [15] GUO H, DU B, ZHANG L, et al. A coarse-to-fine boundary refinement network for building footprint extraction from remote sensing imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 183: 240-252.
- [16] HUANG J, ZHANG X, SUN Y, et al. Attention-guided label refinement network for semantic segmentation of very high resolution aerial orthoimages[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 4490-4503.
- [17] ZHU Q, LIAO C, HU H, et al. MAP-Net: multiple attending path neural network for building footprint extraction from remote sensed imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 59(7): 6169-6181.
- [18] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]. Munich: Springer International Publishing, 2018.
- [19] WANG D, CHEN X, JIANG M, et al. ADS-Net: an attention-based deeply supervised network for remote sensing image change detection[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2021, 101. DOI:org/10.1016/j.jag.2021.102348.
- [20] RIBANI R, MARENGONI M. A survey of transfer learning for convolutional neural networks[C]. Rio de Janeiro: IEEE, 2019.

(责任编辑:王 妮)