

基于改进 DeepLabV3+ 的遥感影像语义分割方法

苏志鹏¹, 李景文^{1,2*}, 姜建武^{1,2}, 陆妍玲^{1,2}, 朱明³¹桂林理工大学测绘地理信息学院, 广西 桂林 541004;²广西空间信息与测绘重点实验室, 广西 桂林 541004;³广西壮族自治区自然资源信息中心, 广西 南宁 510023

摘要 针对普通卷积神经网络在遥感图像分割中小目标识别度不高、分割精度低的问题,提出了一种结合特征图切分模块和注意力机制模块的遥感影像分割网络 AFSSM-Net。首先在编码阶段引入特征图切分模块,对每个切分的特征图进行放大,通过参数共享的方式进行特征提取;然后,将提取的特征与网络原输出图像进行融合;最后,在网络模型中引入注意力机制模块,使其更关注图像中有效的特征信息,忽略无关的背景信息,从而提高模型对小目标物体的特征提取能力。实验结果表明,所提方法的平均交并比达到 86.42%,相比于 DeepLabV3+ 模型提升了 3.94 个百分点。所提方法充分考虑图像分割中小目标的关注度,提升了遥感图像的分割精度。

关键词 遥感; 遥感图像; DeepLabV3+; 特征图切分; 注意力机制; 语义分割

中图分类号 TP751

文献标志码 A

DOI: 10.3788/LOP213268

Semantic Segmentation Method for Remote Sensing Images Based on Improved DeepLabV3+

Su Zhipeng¹, Li Jingwen^{1,2*}, Jiang Jianwu^{1,2}, Lu Yanling^{1,2}, Zhu Ming³¹College of Geomatics and Geoinformation, Guilin University of Technology, Guilin 541004, Guangxi, China;²Guangxi Key Laboratory of Spatial Information and Geomatics, Guilin 541004, Guangxi, China;³Natural Resources Information Center of Guangxi Zhuang Autonomous Region, Nanning 510023, Guangxi, China

Abstract A remote sensing image segmentation network called AFSSM-Net, which combines a feature map segmentation module and an attention mechanism module, is proposed to address the issues of low recognition and low segmentation accuracy of small targets in remote sensing image segmentation using conventional convolutional neural networks. First, the feature map segmentation module is introduced in the coding stage to enlarge each segmented feature map and extract features by sharing parameters; then, the extracted features are fused with the original output image of the network; and finally, the attention mechanism module is introduced into the network model to make it pay more attention to the effective feature information in the image and ignore the irrelevant background information, to improve the feature extraction ability of the model for small target objects. The experimental results show that the average intersection ratio of the proposed method is 86.42%, which is 3.94 percentage points higher than that of the DeepLabV3+ model. The proposed method fully considers the attention of small and medium targets in image segmentation, and improves the segmentation accuracy of remote sensing images.

Key words remote sensing; remote sensing image; DeepLabV3+; feature image cut; attention mechanism; semantic segmentation

1 引言

遥感图像在城市规划、土地利用、农作物监测、灾

害监测、气候预测等多个领域应用广泛^[1-3]。传统的遥感图像语义分割主要通过纹理、边缘、光谱、长度等信息解译出遥感地物信息^[4]。随着深度学习的快速发

收稿日期: 2021-12-17; 修回日期: 2022-01-06; 录用日期: 2022-01-13; 网络首发日期: 2022-01-23

基金项目: 国家自然科学基金(41961063)、广西自然科学基金创新研究团队项目(2019GXNSFGA245001)、国家文化和旅游科技

技术创新工程项目(2019-011)

通信作者: *lijw@glut.edu.cn

展,卷积神经网络(CNN)被广泛应用到遥感影像语义分割领域。2015年,Long等^[5]提出了第1个图像语义分割的全卷积神经网络(FCN)。之后,语义分割模型U-Net^[6]、SegNet^[7]、PSPNet^[8]、DeepLab系列等均是FCN的经典变体,都在遥感图像语义分割领域取得不错的分类效果。DeepLab系列网络可以有效解决传统CNN下采样空间分辨率下降的问题,是图像语义分割的常用模型,可用于图像逐像素的分类。文献[9]首次提出DeepLabV1网络,结合深度卷积网络、空洞卷积和全连接条件随机场,分类的精度和速度都得以提升。文献[10]又提出DeepLabV2网络,提出空间金字塔池化扩大感受野,大大降低了模型的参数量。文献[11]又提出了DeepLabV3网络,设计并行的空洞卷积模块,通过不同的膨胀率捕获上下文信息。文献[12]再次提出了DeepLabV3+网络,引入编码-解码结构,用深度可分离卷积代替原来的普通卷积,大大减小了模型的参数量,使图像分割效果更好。文献[13]提出了一种多注意力融合U型网络,通过多注意力机制处理不同层次的语义信息,强化不同特征图的相关性,提高了模型在复杂条件下的分类性能。文献[14]在网络模型中引入双注意力模块,可以有效地改善DeepLabV3+的缺点,但仍存在边界信息模糊、大目标分割带空洞、小目标分割模糊、拟合速度慢等问题。文献[15]提出了一种对称的编码-解码结构网络模型,通过减少池化、卷积操作,减少模型的参数量,使网络模

型提取到更多细节的图像特征,且大大降低网络的训练时间,与经典的语义分割方法相比,该模型能够在复杂的地物类型中分类出地物信息。但是,利用DeepLabV3+网络提取遥感图像仍存在拟合速度慢、小目标分割不精确、边界信息较粗糙等问题^[16]。

基于以上分析,本文提出了一种结合特征图切分模块和注意力模块遥感影像语义分割网络(AFSM-Net)。首先引入通道、空间注意力模块,使模型关注遥感图像中显著的特征,忽略无关的背景信息;其次引入特征图切分模块,扩大图像中小目标,有效提高遥感语义分割的精度;最后结合膨胀卷积,通过不同的膨胀率,扩大感受野来捕获到更多图像的多尺度特征信息。以2015年中国南方某地区的高分辨遥感影像(CCF卫星遥感数据集)作为研究对象,将AFSM-Net模型与经典的语义分割网络进行对比实验。

2 AFSM-Net模型设计

本研究并没有采用传统的U型网络,而是将注意力机制与DeepLabV3+网络相融合,提出了一种基于改进的DeepLabV3+网络遥感影像语义分割方法,如图1所示。主要的工作包括以下几个方面:在编码器部分,基于Resnext50_32x4d残差网络对遥感图像高维特征进行训练,且添加注意力机制模块(CBAM)^[17];当遥感图像分辨率为1/16时,对特征图进行均等切分,并放大到切分前的大小,使得模型更好

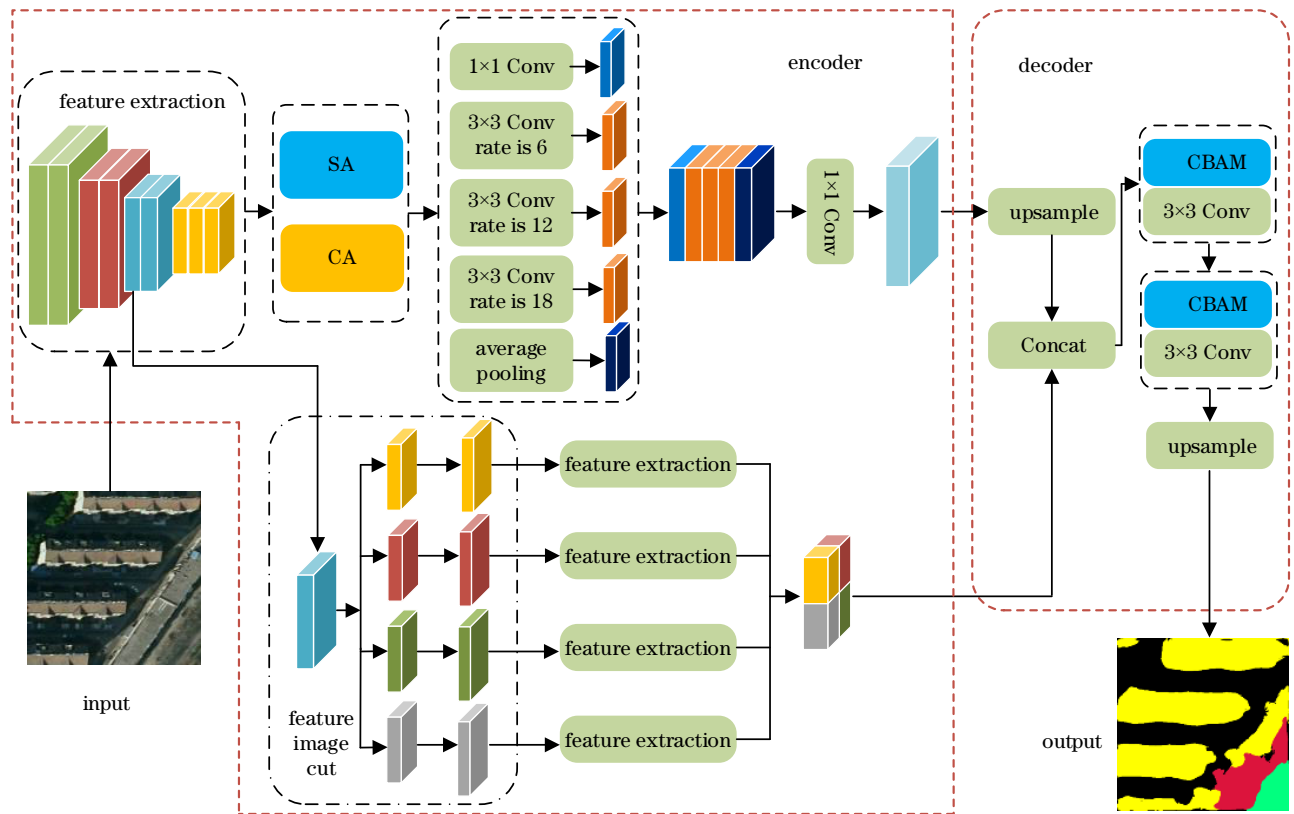


图1 AFSM-Net总体结构图

Fig. 1 Overall structure diagram of AFSM-Net

地关注局部区域内小目标的物体;将提取到的特征图输入空间金字塔模块中,使模型捕获到更多的多尺度信息,且进行1次平均池化操作和1次 1×1 卷积操作,膨胀卷积中的膨胀率分别设置为6、12、18;最后利用 1×1 卷积输出特征图。在解码部分,利用双线性插值对图像进行4倍上采样;设计特征融合模块将不同感受野的特征信息融合到一起;在解码器模块的2个 3×3 卷积层后添加注意力机制模块CBAM;再次利用双线性插值进行4倍上采样,输出预测图像。

2.1 特征图切分模块

遥感影像存在某些小目标的特征,目前大多遥感

语义分割方法没有很好地考虑到这些小特征信息,特征图的切分模块可以有效地考虑遥感影像小物体的空间信息。当遥感图像分辨率为 $1/16$ 时,引入特征图切分模块对特征图进行均等切分,可以有效地提高模型对小物体信息的关注。假设 k 为切分比率,且 $1/16$ 分辨率的图像空间维度为 (w, h) ,此时 k 的取值为1、2、3、4。切分后的图像包含不同位置、空间特征信息,然后对切分后的图像进行上采样操作,放大到切分前的大小,这样有利于将遥感图像中的小目标放大成大目标,最后对大目标的特征进行提取,一定程度上可以提高遥感语义分割的精度。具体切割方法如图2所示。

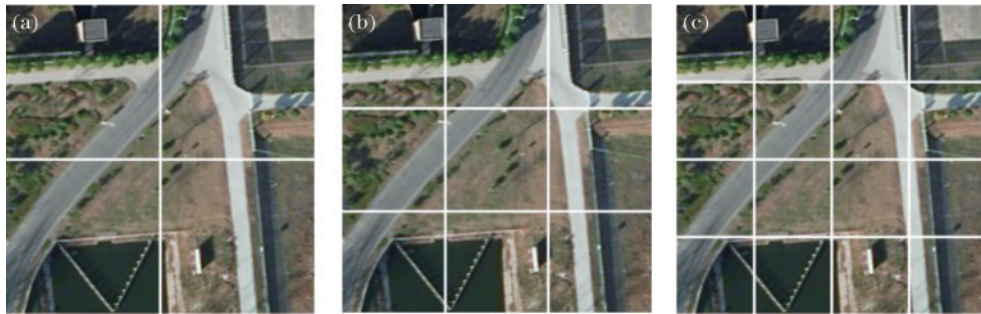


图2 特征图切割。(a) $k=2$; (b) $k=3$; (c) $k=4$
Fig. 2 Feature map cut. (a) $k=2$; (b) $k=3$; (c) $k=4$

将特征图切分模块中 $1/16$ 分辨率的图像还原成切分前的大小,此时应该设计合理的特征提取网络才能有效地提取出图像中小目标的物体。目前对中间层特征的提取一般采用跳跃结构融合特征信息或者通过简单的卷积操作提取出特征,然而,这些操作对遥感语

义分割中间层特征信息的提取并不能达到一个理想的效果。为了更好地获取遥感图像小物体的特征信息,受Inception V3网络结构的启发^[18],所提模型的中间层特征提取网络采用 1×1 、 3×3 大小的滤波器,分别捕获不同尺度的特征信息,如图3所示。

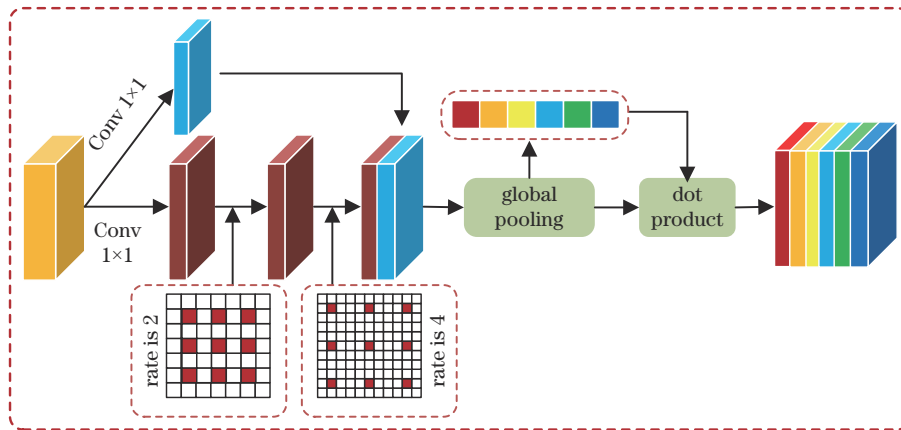


图3 特征提取网络结构
Fig. 3 Structure of feature extraction network

由图3可以看出,特征提取网络主要通过2条支路对中间层的特征图进行输出。该特征提取网络首先采用 1×1 大小的卷积核将通道数降到原来的 $1/2$,同时下面 1×1 卷积的分支中,串联2次膨胀卷积,膨胀卷积的rate分别设置为2、4,可以有效地扩大网络的感受野,减少下采样过程中导致的信息丢失问题;然后将2个支路提取的特征在通道维度上进行拼接,对拼接后

图像特征进行全局池化操作,再与拼接后的特征图进行点乘操作。通过这种操作可以获得到不同切分区域的语义类别信息,为模型提供更加有效的中间层特征信息,从而提高模型对切分放大后的小目标物体特征提取能力。

2.2 空间注意力机制

注意力机制的本质是让模型更加关注图像中的重

要特征信息,忽略无关的背景信息,并且生成不同通道和不同区域的重要性权重,减弱对背景区域的关注,帮助网络模型融合重要的特征信息,可以有效提高特征表示的鉴别能力。每一个卷积层在通道注意力模块可以提取出特征图中的通道特征信息,在空间注意力模块可以提取出特征图中的空间特征信息。通过注意力机制模块的特征学习可以有效地帮助网络模型特征信息的扩散与流动。所提网络引入通道注意力模块和空间注意力模块。

通道注意力模块。首先将输入的特征图 F 经过全局池化与平均池化处理,聚合遥感影像的特征信息,然后再进入多层感知器,即共享参数,将多层感知器输入的特征向量进行相加;接着特征向量经过 Sigmoid 激活函数获得通道注意力权重向量,得到通道注意特征图 $F(a)$;最后将权重矩阵与输入特征图点乘,生成空间注意力模块所需要的输入特征。通道注意力具体结构如图 4 所示。

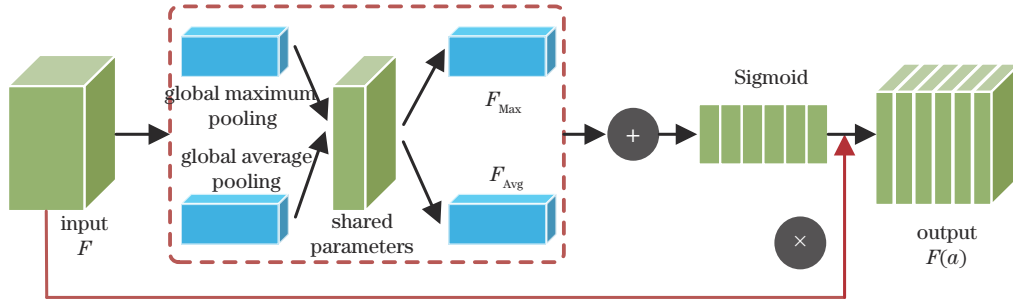


图 4 通道注意力示意图

Fig. 4 Schematic diagram of channel attention

通过注意力的具体计算公式为

$$F(a) = S\{M[\text{MaxPool}(F)] + M[\text{AvgPool}(F)]\} = S\{\omega_1[\omega_0(F_{\max}^c)] + \omega_1[\omega_0(F_{\text{avg}}^c)]\}, \quad (1)$$

式中: S 代表的是 Sigmoid 激活函数; M 表示的是共享参数的全连接层; ω_0, ω_1 表示的是输入的共享参数。

空间注意力模块。首先沿着通道进行最大池化和平均池化操作,然后将它们连接成一个有效的特征信息,最后应用卷积操作生成空间注意力图。具体的计

算过程为

$$F(b) = S\{f^{7 \times 7}[\text{AvgPool}(F(a)); \text{MaxPool}(F(a))]\} = S\{f^{7 \times 7}(F_{\text{avg}}^s; F_{\max}^s)\}, \quad (2)$$

式中: $f^{7 \times 7}$ 代表的是滤波器大小为 7×7 的卷积运算。图 5 为空间注意力运算示意图。前面卷积得到的 $F(a)$ 经过 Sigmoid 函数激活函数后,空间权重信息与输入特征相乘得到输出特征 $F(b)$ 。

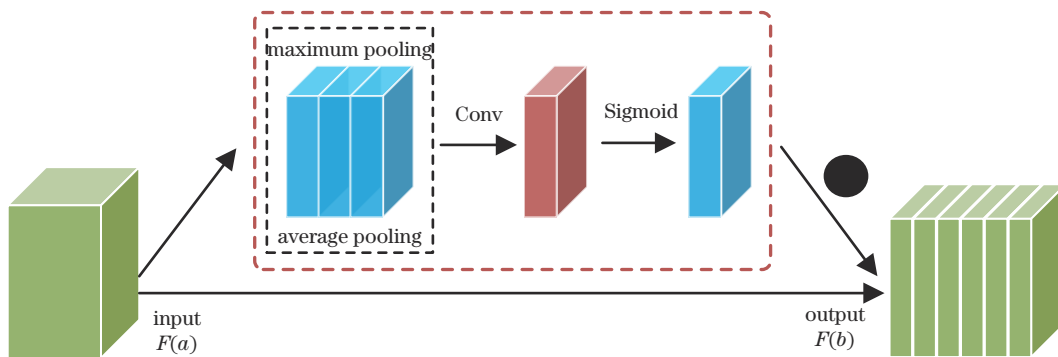


图 5 空间注意力示意图

Fig. 5 Spatial attention diagram

3 遥感图像分割方法的技术流程

基于 AFSM-Net 模型的遥感影像语义分割方法的技术流程思路如下:首先对遥感图像及对应的标注图像进行训练;其次针对遥感图像类别分布不均匀的问题,构建一个带有注意力机制及特征切割模块的

DeepLabV3+ 网络模型,并通过膨胀卷积增大卷积核的感受野;最后通过上采样计算出原始遥感图像中每一个像素的类别值,并将预测值与预测值进行对比分析得出遥感图像的分类结果。

3.1 数据源及预处理

训练数据集为 CCF 卫星数据集,包含 5 张高分辨

率卫星影像,提供的训练数据集一共划分为5类:标记0(其他)、标记1(植被)、标记2(建筑物)、标记3(水体)及标记4(道路),植被包括草地、林地、耕地。由于数据集为5幅尺寸大且不统一的遥感影像,不适宜直接输入卷积神经网络中进行训练。因此,将图像与标注

图像滑动切割为 352×352 的子图,并且对 CCF 卫星影像数据集进行增强操作,采用 OpenCV 对切割后的图像进行扩充,主要经过平移、翻转、旋转、裁剪等操作,最终得到 10000 张子图,通过数据增强操作有效地解决了数据类别不均衡的问题,部分数据如图 6 所示。

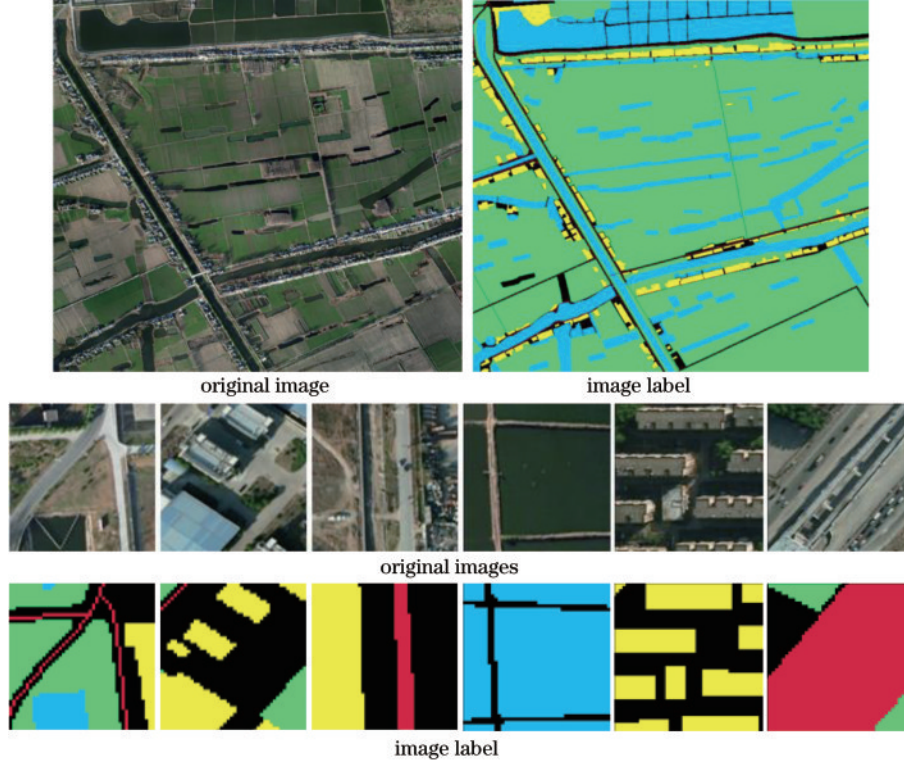


图 6 真实图像与标注图像
Fig. 6 Real images and marked images

3.2 实验参数设置及实验平台

将 CCF 卫星影像数据集按 8:2 的比例进行随机划分,8000 张图像作为训练数据集,2000 张图像作为测试集。选择 SGD 优化器进行网络梯度的更新,采用 ReLU 激活函数及交叉熵损失函数对网络进行训练,经过多次的调参,令学习率为 0.01,批尺寸设置为 32,实验总共迭代 50 次。实验在 Windows 10 操作系统上进行,采用 Pytorch 深度学习框架、Cuda 7.0、Cudnn 9.0,并通过 OpenCV 对数据进行处理。实验通过 NVIDIA 1080TI 8 GB GPU 进行加速。

3.3 实验评价指标

在遥感图像语义分割中,通常使用分类总体精度(OA)、各类别精度(PA)及平均交并比(MIoU)等 3 个指标来评估模型的好坏。OA 表示预测正确的类别占总类别的比例:

$$A_{OA} = \frac{\sum_{n=0}^k p_{nn}}{\sum_{n=0}^k \sum_{m=0}^k p_{nm}}, \quad (3)$$

式中: k 表示的是类别的数量; p_{nn} 表示真正例的像素数量; p_{nm} 表示假正例的像素数量。

PA 表示各个类别像素准确率的平均值:

$$A_{PA} = \frac{1}{k+1} \sum_{n=0}^k \frac{p_{nn}}{\sum_{m=0}^k p_{nm}}, \quad (4)$$

式中: $k+1$ 表示类别数量中含有一类背景。

MIoU 表示的是遥感图像语义分割的标准量度,它是模型预测得到的标准图像与人工标记的图像交集与并集之比:

$$R_{MIoU} = \frac{1}{k+1} \sum_{n=0}^k \frac{p_{nn}}{\sum_{m=0}^k p_{nm} + \sum_{m=0}^k p_{mn} - p_{nn}}, \quad (5)$$

式中: p_{mn} 表示的是假负例的数量。

4 实验结果与分析

4.1 AFSM-Net 消融实验结果与分析

为了验证 AFSM-Net 模型的有效性,分别对通道注意力模块、空间注意力模块和特征图切分模块进行消融实验。选取 DeepLabV3+ 网络模型作为实验的基线模型。

在注意力机制模块消融实验中以 MIoU 作为评价指标。消融实验在 CCF 卫星数据集的评价结果如表 1

所示。从表 1 可以看出:在原始 DeepLabV3+网络上,MIoU 仅有 82.58%;引入通道注意力模块后,MIoU 比基线模型提高了 1.68 个百分点;添加空间注意力模块后,MIoU 比基线模型提高了 2.44 个百分点;同时添加空间注意力模块、通道注意力模块后,MIoU 比基线模型提高了 4.20 个百分点。即空间注意力、通道注意力模块对遥感图像的分割精度都有较好的提升。

表 1 消融实验在 CFF 测试集上评价结果对比
Table 1 Comparison of evaluation results of ablation experiment on CFF test set

ID	Baseline	CA	SA	MIoU / %
a	✓	×	×	82.58
b	✓	✓	×	84.26
c	✓	×	✓	85.02
d	✓	✓	✓	86.78

在特征图切分模块消融实验中,评价指标也是 MIoU,且模型训练参数也相同。在遥感图像特征分辨率为 1/16 时,引入特征切分模块进行均等切分。但在进行特征图的切分时,需要考虑切分比率 k 取何值时,算法的分割精度最高。当切分比率 k 不同时,MIoU 指

标如表 2 所示。当切分比率 $k=1$ 时,图像不进行特征图切分,直接将提取的特征与上采样相同的图像进行融合。从表 2 可以看出: $k=2,3,4$ 时,MIoU 都比 $k=1$ 时高,表明特征切分模块引入后,模型的分割精度会提高;但 $k=3,4$ 时,MIoU 都比 $k=2$ 时低,这是因为随着 k 的增大,遥感图像被切分成过多的小区域,像素的之间关联性大大降低,最终影响了分割的精度。

表 2 不同切分比率的实验结果
Table 2 Experimental results of different cutting ratios

ID	k	MIoU / %
a	1	79.65
b	2	86.40
c	3	83.64
d	4	80.02

4.2 AFSM-Net 模型与其他模型比较

为了验证所提 AFSM-Net 模型的有效性,将其与 FCN、U-Net、SegNet、DeeplabV3+ 等 4 种经典的语义分割模型进行对比实验,4 种经典的语义分割模型和所提改进遥感图像语义分割模型的结果如图 7 所示。从图 7 可以看出,所提模型的分割准确率大大提升,对

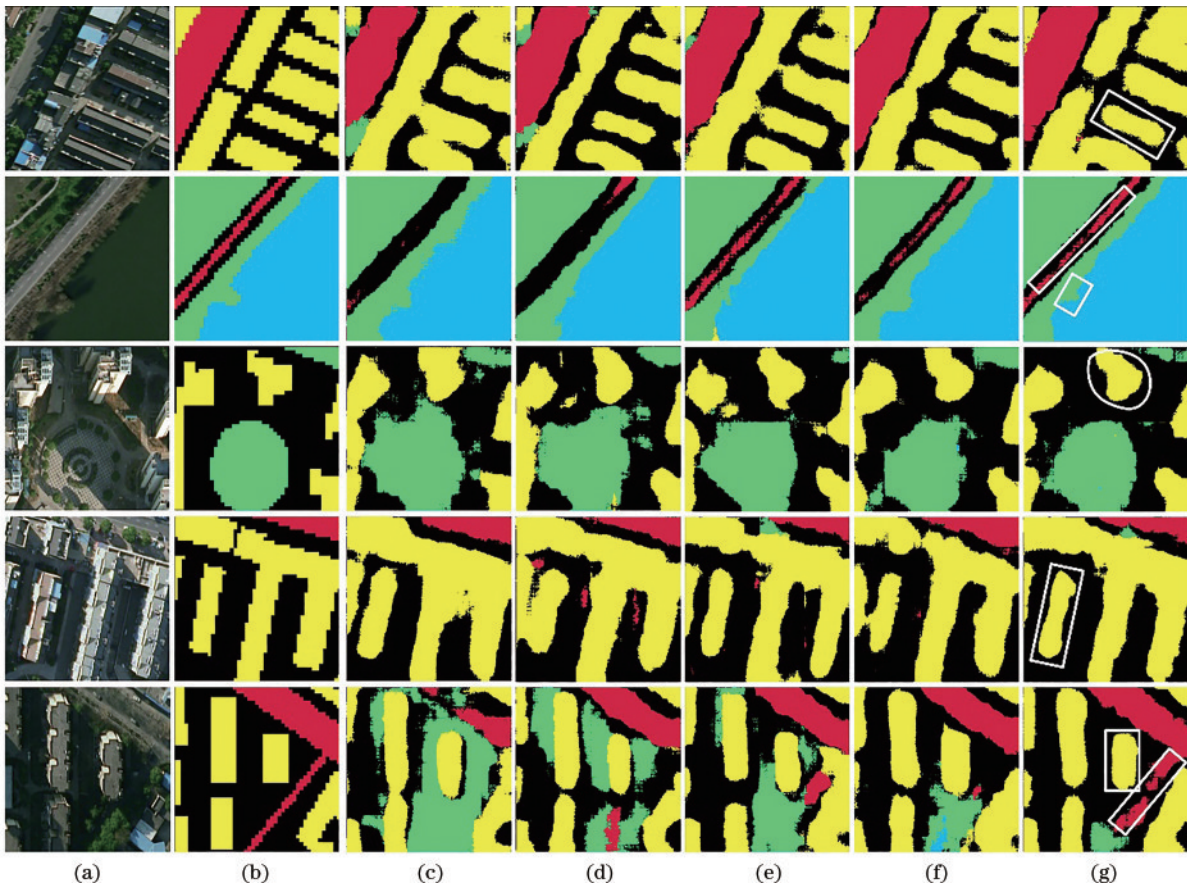


图 7 实验结果对比。(a)原始图像;(b)标注图像;(c)FCN 提取结果;(d)U-Net 提取结果;(e)SegNet 提取结果;(f) DeepLabV3+ 提取结果;(g)AFSM-Net 提取结果

Fig. 7 Comparison of experimental results. (a) Original images; (b) image labels; (c) FCN processing results; (d) U-Net processing results; (e) SegNet processing results; (f) DeepLabV3+ processing results; (g) AFSM-Net processing results

小目标物体的分割效果更好,与标签图像更加接近,且分割的效果更加直观。相比于 FCN、U-Net、SegNet 和 DeepLabV3+ 等经典语义分割算法,所提算法可以清晰分割出遥感影像中较小的建筑物、道路。FCN 算法整体的分割能力较差,出现较多错分、漏分的情况,如图 7(c) 所示;U-Net 算法分割能力比 FCN 算法有了一定的提升,但对建筑物、道路分割效果并不理想,如图 7(d) 所示;SegNet 算法在植被、建筑物、道路等类别也出现了错分、漏分的现象,但整体分割效果比 FCN、U-Net 算法有了一定的提升,如图 7(d) 所示;DeepLabV3+ 算法分割能力比经典的语义分割算法有了较大的提升,但还是无法准确分割出较小的建筑物、道路及植被等类别的边缘特征信息,如图 7(f) 所示;所提算法可以准确分割出遥感影像中小目标,分割效果较其他经典语义分割算法有较大的提升,对图像中小目标边缘特征的提取有了明显的改善,并没有出现较多的错分、漏分现象,如图 7(g) 所示。实验结果表明,所提方法可以有效地提升模型对图像中小目标物体的分割能力。

为了进一步分析不同的语义分割模型的性能,计算了每一个语义分割模型中植被、建筑物、水体、道路的分类准确率及总体分类的准确率,具体如表 3 所示。

表 3 不同算法的分类结果

Table 3 Classification accuracy of different algorithms unit: %

Algorithm	Vegetation	Building	Water	Road	OA	MIoU
FCN	70.22	71.55	69.25	70.22	70.31	70.84
U-Net	81.85	80.26	80.32	79.84	80.57	75.26
SegNet	83.46	83.28	82.84	82.78	83.09	79.45
DeepLabV3+	88.26	88.53	87.00	88.45	88.56	82.48
AFSM-Net	90.46	90.56	89.84	89.64	90.13	86.42

从表 3 可以看出,所提 AFSM-Net 模型各类精度均达到 89% 以上,分类的总体精度达到 90.13%。与经典的语义分割模型 FCN 相比,植被分割的准确率提高了 20.24 个百分点,建筑物的分割准确率提高了 19.01 个百分点,水体的分割准确率提高了 19.42 个百分点,道路分割的准确率提高了 19.82 个百分点。与 U-Net、SegNet、DeepLabV3+ 相比,分割精度也有一定的提升。AFSM-Net 的 MIoU 达到了 86.42%,相比 FCN 的 MIoU 高 15.58 个百分点,相比 U-Net 高 11.16 个百分点,相比 SegNet 高 6.97 个百分点,相比 DeepLabV3+ 高 3.94 个百分点。即所提方法是可行的,且可有效地提高遥感图像的分割结果。

5 结 论

提出了一种结合特征图切分模块和注意力模块的遥感图像语义分割网络模型 AFSM-Net。在编码器中引入特征图切分模块,使得模型可更好地提取局部区

域的小物体的特征;同时引入注意力模块,提升网络模型对遥感图像特征的提取能力,保持空间信息不变的情况下,有效地提取遥感图像的局部细节信息。AFSM-Net 模型提升了对遥感图像的感知能力,改善了遥感图像分割中漏分、错分等问题。实验结果表明,与其他经典语义分割方法相比,AFSM-Net 模型的分类精度更高,可以获得更优的遥感图像分割结果。

参 考 文 献

- [1] 汪亚妮,汪西莉.基于注意力和特征融合的遥感图像目标检测模型[J].激光与光电子学进展,2021,58(2):0228003.
Wang Y N, Wang X L. Remote sensing image target detection model based on attention and feature fusion[J]. Laser & Optoelectronics Progress, 2021, 58(2): 0228003.
- [2] 刘金香,班伟,陈宇,等.融合多维度 CNN 的高光谱遥感图像分类算法[J].中国激光,2021,48(16):1610003.
Liu J X, Ban W, Chen Y, et al. Multi-dimensional CNN fused algorithm for hyperspectral remote sensing image classification[J]. Chinese Journal of Lasers, 2021, 48(16): 1610003.
- [3] Kemker R, Salvaggio C, Kanan C. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 145: 60-77.
- [4] 李树涛,李聪好,康旭东.多源遥感图像融合发展现状与未来展望[J].遥感学报,2021,25(1):148-166.
Li S T, Li C Y, Kang X D. Development status and future prospects of multi-source remote sensing image fusion[J]. National Remote Sensing Bulletin, 2021, 25(1): 148-166.
- [5] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [6] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [7] 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.
- [8] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6230-6239.
- [9] Chen L C, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[EB/OL]. (2014-12-22) [2021-02-05]. <https://arxiv.org/abs/1412.7062>.
- [10] Chen L C, 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.
- [11] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-06-17)[2021-02-05]. <https://arxiv.org/abs/1706.05587>.
- [12] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[EB/OL]. (2018-02-07)[2021-02-05]. <https://arxiv.org/abs/1802.02611>.
- [13] 李道纪, 郭海涛, 卢俊, 等. 遥感影像地物分类多注意力融和U型网络法[J]. 测绘学报, 2020, 49(8): 1051-1064.
Li D J, Guo H T, Lu J, et al. A remote sensing image classification procedure based on multilevel attention fusion U-Net[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(8): 1051-1064.
- [14] 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.
Liu W X, Shu Y Z, Tang X M, et al. Remote sensing image segmentation using dual attention mechanism Deeplabv3+ algorithm[J]. Tropical Geography, 2020, 40(2): 303-313.
- [15] 张哲晗, 方薇, 杜丽丽, 等. 基于编码-解码卷积神经网络的遥感图像语义分割[J]. 光学学报, 2020, 40(3): 0310001.
Zhang Z H, Fang W, Du L L, et al. Semantic segmentation of remote sensing image based on encoder-decoder convolutional neural network[J]. Acta Optica Sinica, 2020, 40(3): 0310001.
- [16] 郭梦利, 阮顺领, 卢才武, 等. 基于改进 Deeplabv3+ 网络的露天矿路网提取方法[J]. 激光与光电子学进展, 2021, 58(22): 2228005.
Guo M L, Ruan S L, Lu C W, et al. Road extraction method of open-pit mine based on improved Deeplabv3+ network[J]. Laser & Optoelectronics Progress, 2021, 58(22): 2228005.
- [17] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer Vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [18] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2818-2826.