

DOI:10.19651/j.cnki.emt.2210180

基于改进 DeeplabV3+ 的遥感图像分割算法^{*}

黄聪 杨珺 刘毅 谢鸿慧

(江西农业大学软件学院 南昌 330045)

摘要: 针对高分辨率遥感图像语义分割存在地物边缘分割不连续、小目标分割精度不高的缺陷,本文提出一种基于改进 DeeplabV3+ 的遥感图像分割算法,该算法首先使用分散注意力网络 ResNeSt 替换 DeeplabV3+ 原始主干网络 Xception,以提取更丰富的深层语义信息,从而提高图像分割精度;其次引入坐标注意力机制(CA),有效获得更精确的分割目标位置信息,使得分割目标边缘更加连续;最后在解码层中采用级联特征融合方法(CFF)提高网络的语义信息表征能力。试验结果表明,该算法在中国南方某城市的高清遥感图像数据集分割任务上 mIoU 高达 97.07%,相比原始 DeepLabV3+ 模型提高了 3.39%,能够更好地利用图像语义特征信息,为解译遥感图像语义信息提供一种新的思路。

关键词: 遥感图像;DeeplabV3+;ResNeSt;注意力机制;特征融合

中图分类号: TP391 文献标识码: A 国家标准学科分类代码: 510.4050

Remote sensing image segmentation algorithm based on improved DeeplabV3+

Huang Cong Yang Jun Liu Yi Xie Honghui

(School of Software, Jiangxi Agricultural University, Nanchang 330045, China)

Abstract: Aiming at discontinuous object edge segmentation in high-resolution remote sensing image semantic segmentation and low accuracy of small object segmentation, this paper proposes a remote sensing image segmentation algorithm based on improved DeeplabV3+. The algorithm first adopts distraction network called ResNeSt instead of the DeeplabV3+ original backbone network Xception to extract richer deep semantic information, thereby improving the accuracy of image segmentation; secondly, the Coordinate Attention (CA) mechanism is introduced to effectively obtain more accurate target location information of segmentation to make the segmentation target edge more continuous; finally, the cascade feature fusion method (CFF) is adopted in the decoding layer to improve the semantic information representation ability of the network. The experimental results show that the algorithm has a high mIoU of 97.07% on the high-definition remote sensing image dataset of a city in southern China, which is 3.39% higher than that of the original model and a reflection of better utilization of image semantic feature information. This provides a new way of thinking for remote sensing image semantic information.

Keywords: remote sensing image;DeeplabV3+;ResNeSt;attention mechanism;feature fusion

0 引言

随着遥感图像空间分辨率的不断提高,高分辨率遥感图像已成为获取地理信息的最主要的现代化手段之一^[1-2],对于国土观察,土地利用,农田规划及制图具有巨大的应用价值。

遥感图像语义分割是图像处理和计算机视觉技术中关于遥感图像理解的核心任务之一,它的任务是以像素为单

位识别遥感图像,即判断出图像中每个像素所属的对象类别^[3]。早期对遥感图像分割主要通过目视解译与计算机相结合的方法,如基于阈值的分割^[4]、基于聚类的分割^[5]、基于边缘的分割^[6]、基于图论的分割^[7]等,这些传统分割方式往往只能提取到图像的低级特征,分割效率低,不能满足高分辨率遥感图像分割要求。因此近年来,国内外研究人员大都采用深度学习的方法,Maggiore 等^[8]将 FCN^[9]应用于马萨诸塞州建筑物数据集,虽然相较于传统的神经网络能

收稿日期:2022-05-31

^{*} 基金项目:江西省自然科学基金-面上项目(20212BAB205009)、江西省教育厅科技基金(GJJ13266,GJJ180374,GJJ170303)项目资助

较好的获取图像中地面特征,但是其采用直接上采样的方法忽视了图像细节信息。Cao 等^[10]提出 Res-UNet 网络,将 ResNet 网络^[11]的基本单元结构替换 U-Net^[12]网络的卷积层,通过跳跃连接将下采样获得的底层特征与上采样输入相结合,改进的 U-Net 网络在高分辨率遥感森林图像中取得了较高的分类精度,但对于少量小目标树种,分类精度会受到影响。Zhu 等^[13]基于 DeepLabV3+^[14]提出一种从高分辨率遥感图像中提取甘蔗田的算法,该算法采用分块的方法进行多时相遥感图像融合,提取精度优于传统 DeepLab 模型^[15]。Wang 等^[16]提出一种融合了类特征注意力机制的 Deeplabv3+网络,在公共数据集(GID)上的实验表明该网络可以更好的分割高分辨率遥感图像中的不同地物的边界区域,然而该模型却无法准确分割某些小场景中的地面对象。

以上研究虽然在一定程度上提高了遥感图像分割精度,但对于地物边缘和小目标仍存在分割不连续和遗漏的缺陷。因此本文基于 DeeplabV3+模型,提出一种基于改进 DeepLabV3+的遥感图像分割算法,该算法采用分散注

意力网络 ResNeSt 作为特征提取主干网络,减少信息丢失,有助于提取更丰富的深层语义信息;引入 CA 注意力机制,增强网络捕捉分割小目标对象特征信息的能力;在解码器中采用级联特征融合模块进行多尺度特征融合,避免因物边缘不连续而导致分割不准确。最后在中国南方某城市的高清遥感图像数据集上进行训练并与原始模型及常见分割模型做对比,实验验证了本文改进算法对高分辨率遥感图像具有较好的分割效果,能够准确地表达高分辨率遥感影像的地物边缘和小目标信息。

1 方法与原理

1.1 改进的 DeepLabV3+算法

在语义分割领域,传统的 DeeplabV3+算法能够很好的解决分割对象多尺度问题以及多次下采样导致的分辨率下降问题,但将其应用于高分辨率的遥感图像数据集时易出现地物边缘分割不连续、类别分割精度不高的问题。因此本文提出一种改进的 DeepLabV3+遥感图像分割算法,图 1 所示为改进的 DeeplabV3+算法采用编解码的结构。

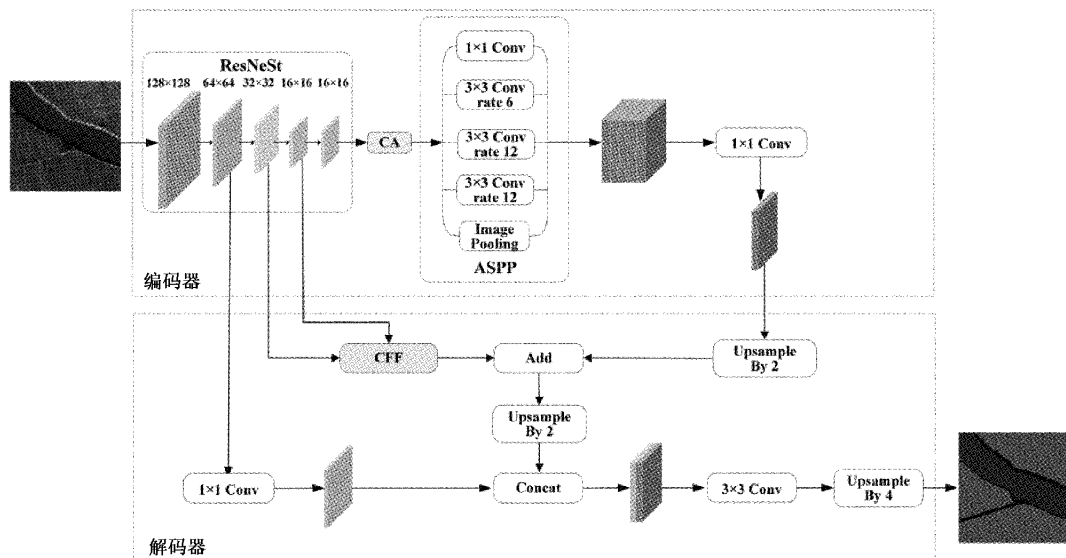


图 1 改进 DeepLabV3+算法模型结构

1)编码层首先采用分散注意力网络 ResNeSt 作为模型骨干网络对输入的图像进行特征提取,骨干网络提取特征的同时会保留原图 1/2、1/4、1/8 和 1/16 大小的特征图,本文将提取的深层语义特征(原图 1/16 大小)输入至 CA 注意力模块中,以增强图像特征的代表能力,然后 ASPP 模块对提取到的特征以不同膨胀率的空洞卷积进行并行采样,获取不同尺度的图像感受野并融合不同比例的语义信息,最后使用 1×1 卷积调整特征通道数后进行双线性插值两倍上采样操作。

2)为充分利用输入特征的细节信息,避免在提取高分辨率遥感图像特征过程中出现小目标丢失问题,通过 CFF 模块将骨干网络提取到的 1/8 和 1/16 大小的特征图进行

特征融合,然后与编码层中获取到的最终特征图进行特征叠加。为进一步降低特征信息损失,本算法同样对叠加后的特征图做双线性两倍上采样,逐渐还原图像语义信息。

3)在解码层中将骨干网络提取到的原图 1/4 大小的浅层特征图做一个 1×1 卷积操作调整特征通道数,然后与 2)中得到的多尺度特征图做 Concat 融合,最后将融合后的特征图进行 3×3 卷积以及双线性插值四倍上采样操作还原至原图大小并得到分割图像。

1.2 ResNeSt 主干网络

目前的主流图像分类网络虽然在分类方面表现出卓越的运算能力和精确度,但通常这些网络模型的感受野^[17]受限且缺乏通道之间的交互性,执行语义分割任务时存在

一定缺陷,鉴于此亚马逊团队提出了一个模块化的分散注意力块^[18](split-attention block),再通过 ResNet 网络的样式堆叠 split-attention block,从而使得网络能够跨特征图组提取语义信息。

ResNeSt 网络核心模块结构如图 2 所示,将输入的特征图沿通道维度拆分为 m 个基数组(Cardinal),每个基数组又拆分为 n 个小组,因此特征图拆分的总数为 $G = m \times n$ 。

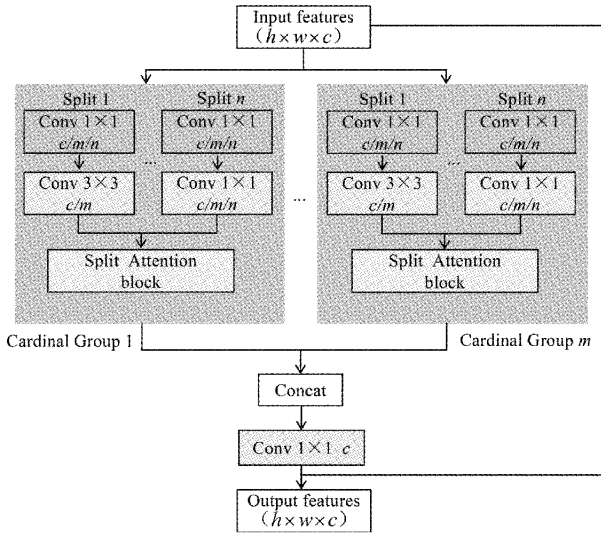


图 2 ResNeSt 核心模块结构

每个基数组中都有一个分散注意力模块,具体结构如图 3 所示。

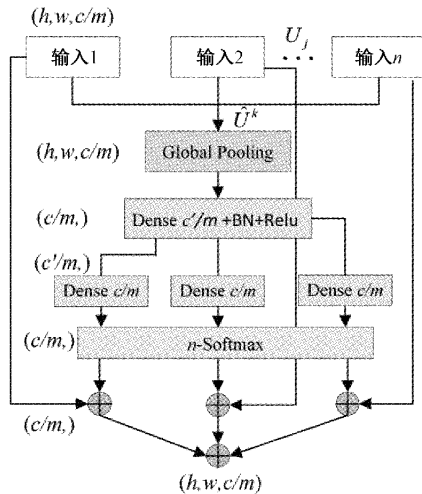


图 3 Split-Attention block 结构

首先将图 3 中的 Split 1~Split n 经 1×1 卷积和 3×3 卷积后的结果作为输入,图中 U_j 为分散注意力块中第 j 个输入特征, \hat{U}^k 为第 k 个基数组的融合特征,表达式为:

$$\hat{U}^k = \sum_{j=n(k-1)+1}^{kn} U_j (j \in \{n(k-1)+1, n(k-1)+2, \dots, nk\}) \quad (1)$$

其次使用全局平均池化^[19]来聚合跨空间通道维度的语义上下文信息,经全局池化第 k 个基数组的第 c 个通道值计算公式如下:

$$s_c^k = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W \hat{U}_c^k(i, j) \quad (2)$$

式中: $\hat{U}_c^k(i, j)$ 表示第 k 个基数组的第 c 个通道在点 (i, j) 上的值, H, W 为特征图的高和宽。

然后将第 k 个基数组的第 c 个通道值输入至 n -Softmax 层中计算第 i 个小组的权重:

$$a_i^k(c) = \begin{cases} \frac{\exp(G_i^c(s_c^k))}{\sum_{j=1}^n \exp(G_j^c(s_c^k))}, & n > 1, \\ \frac{\exp(G_i^c(s_c^k))}{1 + \exp(G_i^c(s_c^k))}, & n = 1, \end{cases} \quad (3)$$

式中: G_i^c 为根据全局上下文信息来确定权重的函数,当 $n=1$ 时, $a_i^k(c)$ 表示为 sigmoid 函数^[20]。

至此可计算每个分支的第 c 个通道加权融合值,计算公式为:

$$V_c^k = \sum_{i=1}^n a_i^k(c) U_{n(k-1)+i} \quad (4)$$

式中: $a_i^k(c)$ 表示第 i 个基数组经 softmax 后计算得到的权重。

最后对不同组的特征图通过分散注意力机制分配的不同权重融合生成输出新的特征图,如下:

$$V = \text{Concat}\{V^1, V^2, \dots, V^k\} \quad (5)$$

与 ResNet 网络一样,如果输入和输出特征图维度一致时,则使用快捷连接方式生成最终输出特征图:

$$Y = V + X \quad (6)$$

通过分组的方式对输入特征图进行加权融合成新的特征图,能够跨通道提取分割图像的语义信息,本文采用 ResNeSt 网络作为特征提取主干网络,以期精准提取高分辨率遥感图像地物特征。

1.3 CA 注意力机制

近年来,注意力机制被广泛应用于计算机视觉领域,并对于提升深度学习模型性能具有显著效果^[21-22]。注意力机制的基本思想是根据不同权重对输入特征进行关键信息的提取,在 DeeplabV3+ 中,骨干网络提取到的深层语义特征图比较小,得到的特征抽象且缺乏空间信息,因此本文引入 CA 注意力机制,在通道注意力中加入特征位置信息,增强对分割对象的关注。CA 注意力机制结构示意图如图 4 所示。

1.4 CFF 特征融合模块设计

在主干特征提取网络中往往会产生多个尺度的特征层,其中浅层特征包含一些局部、精细的特征,能够获得像素级定位精度,而深层特征是全局的特征,能够获得上下文信息和准确的语义信息。因此浅层与深层特征的融合,可以获得更加强大的特征表达。在传统的 DeeplabV3+ 模型

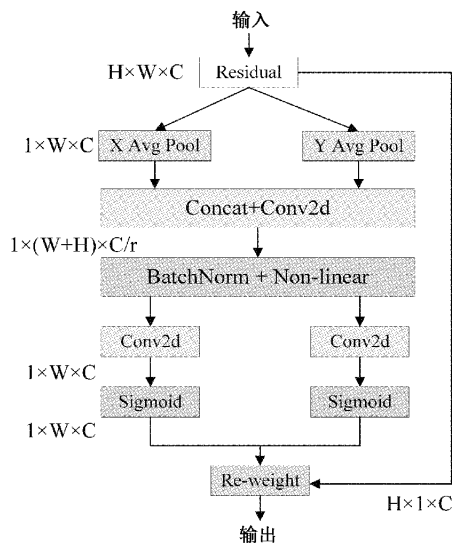


图 4 Coord Attention 结构

中,解码层只是简单地将原图尺寸 $1/4$ 大小的浅层特征与深层特征进行拼接融合,易出现分割对象细节信息丢失的问题。为提高网络分割语义信息表征能力,在本文中采用 CFF 模块对不同比例的特征图进行特征融合,如图 5 所示。

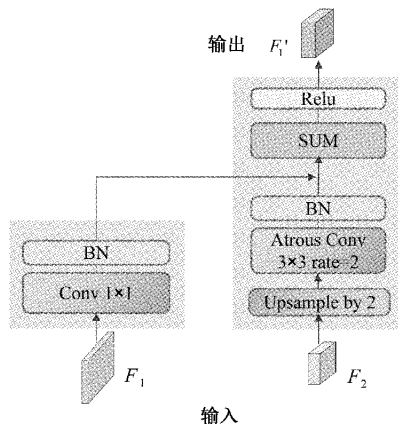


图 5 级联特征融合模块

首先将特征映射 F_1 (原图 $1/8$ 大小的特征图) 和 F_2 (原图 $1/16$ 大小的特征图) 作为 CFF 模块的输入,其次对 F_2 做双线性插值 2 倍上采样后再通过一个空洞率为 2 的 3×3 空洞卷积,使之尺寸大小和感受野与 F_1 保持一致,同时对 F_1 进行一个 1×1 卷积操作使得特征通道数与 F_2 相同,然后使用批量归一化 (Batch Normalization) 层分别对处理后的 F_1 和 F_2 做标准化,最后将得到的 F_1 和 F_2 特征相加后通过 ReLU 层得到融合特征 F'_1 。

2 实验与讨论

基于改进 DeeplabV3+ 算法的遥感图像分割主要有以下 3 个步骤: 1) 制作并划分高分辨率遥感图像数据集; 2) 将数据集加载至改进的 DeeplabV3+ 模型中进行训练; 3) 与其

他网络模型做对比和消融实验,验证并分析改进 DeeplabV3+ 算法模型对高分辨率遥感图像的分割有效性。

2.1 实验数据

本文数据集选用中国计算机学会 (China Computer Federation, CCF) 大赛提供的中国南方某地区的高分辨率遥感图像之一,其分辨率为 $5\,664 \times 5\,142$,且具有亚米级的空间分辨率,光谱为可见光波段 (R, G, B)。首先采用 Labelme 标注工具对原图进行像素级人工标注,如图 6 所示,标注分为农田、水域以及背景共 3 类,其中背景类包含建筑、道路和荒地等。

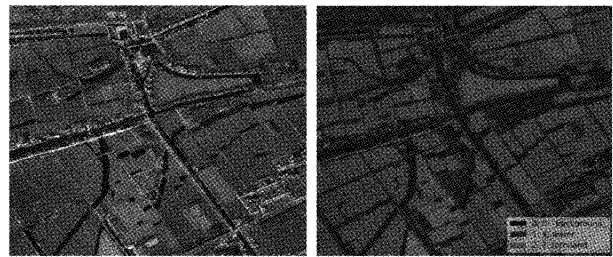


图 6 高分辨率遥感图像样本

将样本及标签随机裁剪出 6 000 张无重叠的 $256 \text{ pixel} \times 256 \text{ pixel}$ 的图像,为提高模型泛化能力,通过翻转变换、色彩抖动、加噪等处理对裁剪后的图像进行数据增强,最后按照 8 : 2 的比例划分为 4 800 张训练集和 1 200 张测试集。

2.2 实验参数及评价指标

针对本文改进的算法进行实验研究,实验软硬件环境配置情况如表 1 所示。

表 1 实验软硬件环境配置情况表

实验环境	配置版本
操作系统	Windows 10
CPU	Inter [®] Xeon [®] Silver 4112
GPU	NVIDIA Quadro P4000(8G)
Cuda	Cuda 10.2
Python	Python 3.6
框架	Pytorch 1.9.1
开发工具	Pycharm 2020.2

为保证改进的 DeeplabV3+ 算法的真实性,实验中原始模型和改进后的模型均设置相同的超参数,采用随机梯度下降 (SGD) 优化算法,动量 (momentum) 为 0.9,初试学习率 (lr) 为 0.005,选用 poly 学习策略,权重衰减率 (weight decay) 设置为,训练迭代次数 (epoch) 为 300,批处理大小 (batch size) 为 8,选用交叉熵损失函数 (Cross Entropy Loss) 计算损失。此外,实验中利用迁移学习思想在训练前先载入主干网络 ResNeSt 的预训练模型权重,将预训练参数更新至改进网络中训练,使得模型能够加速收敛。

为验证改进 DeeplabV3+ 算法在高分辨率遥感图像

中的分割性能,本文采用像素精度(PA)、平均像素精度(MPA)、交并比(IoU)、平均交并比(mIoU)、频率权重并交比(FWIoU)作为技术评价指标,计算方式如式(7)~(11)所示,其中 n_{class} 表示分割目标类别个数(含背景), n_{ij} 表示目标类别 i 被预测成类别 j 的像素的个数, $t_i = \sum_j n_{ji}$ 表示为目标类别为 i 的像素总数(真实标签)。

$$PA = \frac{\sum_i n_{ii}}{\sum_i t_i} \quad (7)$$

$$MPA = \frac{1}{n_{class}} \cdot \sum_i \frac{n_{ii}}{t_i} \quad (8)$$

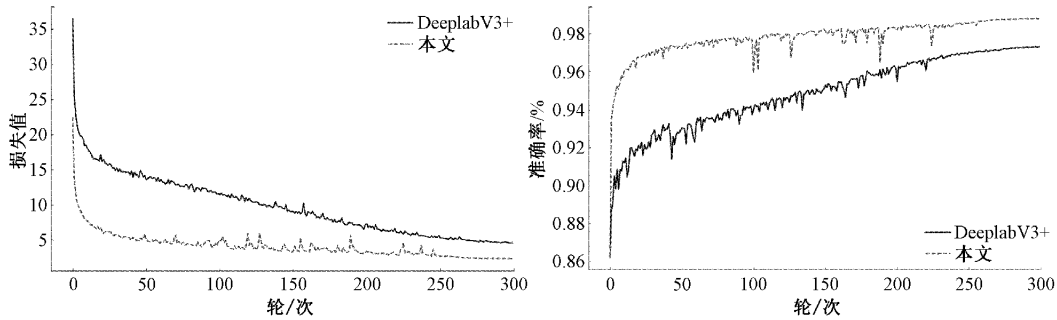


图 7 改进前后的 DeepLabV3+算法损失和准确度曲线

表 2 不同改进方案评估标准对比

模型	主干网络	CA module	CFF module	MPA/%	mIoU/%	训练时间/h
1	Xception			96.65	93.68	34.51
2	ResNeSt			98.00	96.31	27.60
3	ResNeSt	✓		98.08	96.36	28.48
4	ResNeSt		✓	98.43	97.05	40.76
5	ResNeSt	✓	✓	98.48	97.07	31.90

实线表示传统 DeepLabV3+算法,虚线为改进后的 DeepLabV3+算法,总体上二者损失值随网络训练迭代次数增加而递减,并且逐渐趋于平稳,准确度也随网络训练迭代次数增加而增加,最后趋于收敛。其中,改进的 DeepLabV3+算法的损失值曲线和准确度曲线收敛速度比传统 DeepLabV3+算法快,且损失值更低,准确度更高,后期波动小。

在表 2 中模型 1 是主干网络为 Xception 的传统 DeepLabV3+模型,其 MPA 为 96.65%,平均交并比 mIoU 为 93.68%;模型 2 是主干网络为 ResNeSt 的 DeepLabV3+模型,其 MPA 和 mIoU 较模型 1 提高了 1.35%和 2.63%,训练时间也节约了 6.91 小时,这是由于 ResNeSt 网络能够跨通道提取特征语义信息,在提高精度的同时不会增加额外的计算成本;模型 3 在模型 2 的基础上引入 CA 注意力模块,模型 4 在模型 2 上引入 CFF 模块,其中模型 3 较模型 2 在精度上仅有小幅提升,而模型 4 因引入的 CFF 模块融合了更多跨层特征信息,有助于提高图像语义分割精度,故 mIoU 有明显提高;模型 5 即本文方

$$IoU = \sum_i \frac{n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}} \quad (9)$$

$$mIoU = \frac{1}{n_{cls}} \cdot \sum_i \frac{n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}} \quad (10)$$

$$FWIoU = \frac{1}{\sum_i t_i} \cdot \sum_i \frac{n_{ii} \cdot t_i}{t_i + \sum_j n_{ji} - n_{ii}} \quad (11)$$

2.3 实验结果与分析

改进前后的 DeepLabV3+算法在高分辨率遥感图像数据集上训练的损失值曲线(Loss)和准确度曲线(Acc)以及评估指标结果分别如图 7 和表 2 所示。

法同时加入了 CA 注意力模块和 CFF 模块,相比模型 4,训练时间减少了 21%,同时精度也得到一定提高,说明了 CA 注意力模块能够集中注意力关注输入特征的重要语义信息,也验证了本文改进的 DeepLabV3+算法对于高分辨率遥感图像分割的有效性。

为进一步衡量本文提出的改进 DeepLabV3+算法对高分辨遥感图像的分割性能,选用 Unet 模型、主干网络为 Xception、MobilenetV2^[23]、Resnet 的 DeepLabV3+模型与改进的 DeepLabV3+算法做对比实验,不同网络模型在验证集上的实验结果如表 3 所示。本文方法即改进的 DeepLabV3+算法 PA 为 98.81%,MPA 为 98.48%,各项分割目标(水域、农田、背景)的 IoU 分别为 95.80%、98.63%、96.80%,mIoU 高达 97.07%、FWIoU 高达 97.65%,由表 3 可知改进的 DeepLabV3+算法在各项技术指标上均高于其他网络模型,其中,PA 和 mIoU 分别较 Unet 模型高 0.75%和 1.68%,此外,与主干网络为 Xception、MobilenetV2、Resnet 的 DeepLabV3+模型相比,本文方法的 mIoU 分别提高了 3.39%、1.48%以及

0.63%,且水域、农田和背景的 IoU 都有明显提高,说明改进的 DeepLabV3+算法整体分割效果较好。

表 3 不同网络模型评估标准对比

%

模型	PA	MPA	IoU			mIoU	FWIoU
			水域	农田	背景		
Unet	98.06	97.56	93.63	97.69	94.84	95.39	96.20
DeepLabV3+(Xception)	97.31	96.65	91.31	96.80	92.94	93.68	94.78
DeepLabV3+(Mobilenet)	98.15	97.71	93.90	97.80	95.07	95.59	96.37
DeepLabV3+(Resnet)	98.53	98.25	95.00	98.28	96.06	96.45	97.11
本文方法	98.81	98.48	95.80	98.63	96.80	97.07	97.65

不同网络模型在本文数据集上分割效果对比如图 8 所示,通过对比图中方框区域,可以发现不同主干网络的 DeepLabV3+模型在分割农田边缘时出现分割不连续的情况,这是因为传统的 DeepLabV3+模型在解码过程中直接将深层语义信息和浅层语义信息相融合,未能充分利用输入特征其他的细节信息,从而导致分割效果不佳,而改进的 DeepLabV3+算法使用了 CFF 模块加强了上下文语义信息的联系,对于地物边缘分割更加连续,分割效

果与人工标注图相似度极高。从最后一行分割效果可以看出 Unet 网络模型将黄框中农田误分割为背景,对于小目标地物分割精度不高,而本文方法相比 Unet 网络能够更加准确分割地物类别。综上所述,本文方法在高分辨率遥感图像数据集上分割更加精细,与其他常用分割模型相比效果更优,但是也存在分割边界模糊问题,如第 2 行中分割农田边界出现断裂情况,未能将田埂背景完全分割出来。

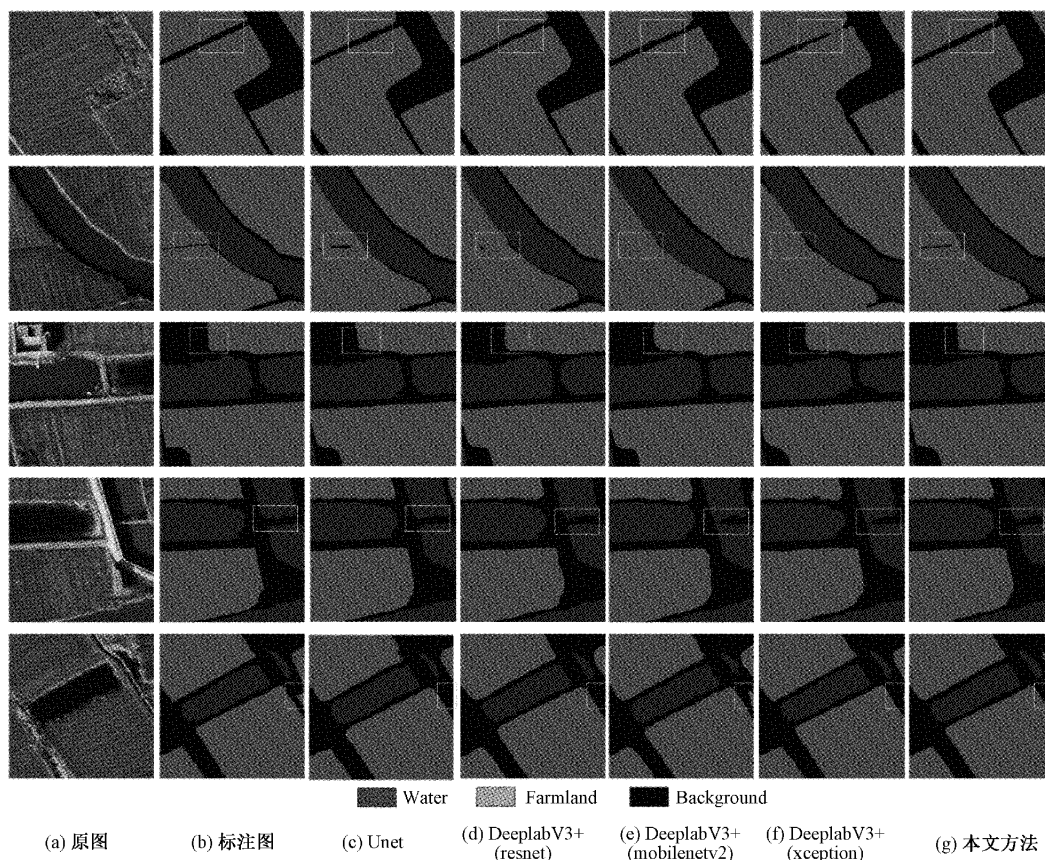


图 8 改进前后的 DeepLabV3+算法损失和准确度曲线

3 结 论

本文提出一种改进的 DeepLabV3+算法:首先在原

DeepLabV3+模型编解码的基础上使用特征提取主干网络 ResNeSt 提取更丰富的特征语义信息;然后在提取深层特征后引入坐标注意力机制获得更加精准的通道关系和

位置信息,增强对分割对象的特征表示;最后使用级联特征融合模块加强上下文语义信息的联系,提高小目标分割准确度。通过对比和消融实验,结果表明本文方法在各项评估指标上都优于其他网络模型,有效解决了地物边缘分割不连续、小目标分割精度不高问题,有助于高分辨率遥感图像的精准分割。由于本文实验数据集分割目标数较少,场景较简单,因此在后续的工作中需考虑语义分割的实用性,深入研究在复杂场景下的高分辨率遥感图像分割与应用。

参考文献

- [1] LI D, WANG M, JIANG J. China's high-resolution optical remote sensing satellites and their mapping applications [J]. *Geo-spatial Information Science*, 2021, 24(1): 85-94.
- [2] 张静,景海涛,樊盛华. 基于海岸线数据库的遥感图像海陆分割[J]. *电子测量技术*, 2020, 43(23): 115-120. DOI:10.19651/j.cnki.emt.2005133.
- [3] 田莹,王亮,丁琪. 基于深度学习的图像语义分割方法综述 [J]. *软件学报*, 2019, 30(2): 440-468. DOI: 10.13328/j.cnki.jos.005659.
- [4] AL-AMRI S S, KALYANKAR N V. Image segmentation by using threshold techniques[J]. *ArXiv Preprint*, 2010, arXiv:1005.4020.
- [5] COATES A, NG A Y. Learning feature representations with K-means[M]. *Neural Networks: Tricks of the Trade*, Springer, Berlin, Heidelberg, 2012; 561-580.
- [6] AL-AMRI S S, KALYANKAR N V, KHAMITKAR S D. Image segmentation by using edge detection[J]. *International Journal on Computer Science and Engineering*, 2010, 2(3): 804-807.
- [7] PENG B, ZHANG L, ZHANG D. A survey of graph theoretical approaches to image segmentation [J]. *Pattern Recognition*, 2013, 46(3): 1020-1038.
- [8] MAGGIORI E, TARABALKA Y, CHARPIAT G, et al. Fully convolutional neural networks for remote sensing image classification [C]. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, 2016; 5071-5074.
- [9] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015; 3431-3440.
- [10] CAO K, ZHANG X. An improved Res-UNet model for tree species classification using airborne high-resolution images[J]. *Remote Sens*, 2020, 12, 1128.
- [11] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks[C]. *European Conference on Computer Vision*, Springer, Cham, 2016; 630-645.
- [12] 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*, Springer, Cham, 2015; 234-241.
- [13] ZHU M, YAO M, HE Y, et al. Studies on high-resolution remote sensing sugarcane field extraction based on deep learning [C]. *IOP Conference Series: Earth and Environmental Science*, IOP Publishing, 2019, 237(3): 032046.
- [14] 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 (ECCV)*, 2018; 801-818.
- [15] 陈彦彤,李雨阳,陈伟楠,等. 基于深度语义分割的遥感图像海面舰船检测研究[J]. *仪器仪表学报*, 2020, 41(1): 233-240. DOI:10.19650/j.cnki.cjsi.J1905722.
- [16] 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 Deeplabv3+ [J]. *Computers & Geosciences*, 2022, 158; 104969.
- [17] LUO W, LI Y, URTASUN R, et al. Understanding the effective receptive field in deep convolutional neural networks [J]. *Advances in Neural Information Processing Systems*, 2016, 29.
- [18] ZHANG H, WU C, ZHANG Z, et al. Resnest: Split-attention networks [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022; 2736-2746.
- [19] LIN M, CHEN Q, YAN S. Network in network[J]. *ArXiv Preprint*, 2013, ArXiv:1312.4400.
- [20] EZEAFULUKWE U A, DARUS M, FADIPE-JOSEPH O. On analytic properties of a sigmoid function [J]. *Int. Journal of Mathematics and Computer Science*, 2018, 13(2): 171-178.
- [21] CHAUDHARI S, MITHAL V, POLATKAN G, et al. An attentive survey of attention models [J]. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2021, 12(5): 1-32.
- [22] 任欢,王旭光. 注意力机制综述[J]. *计算机应用*, 2021, 41(S1): 1-6.
- [23] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018; 4510-4520.

作者简介

黄聪, 硕士, 主要从事计算机视觉方面的研究。

E-mail: 2033909001@qq.com

杨珺(通信作者), 博士, 副教授, 硕士生导师, 主要从事数据挖掘、知识工程方面的研究。

E-mail: ycjun515@163.com

刘毅, 硕士, 讲师, 主要从事图像处理方面的研究。

E-mail: 308632394@qq.com

谢鸿慧, 硕士, 主要从事数据挖掘、知识图谱方面的研究。

E-mail: 1505523351@qq.com