

结合全国第三次国土调查资料的 DeepLabV3+ 林地、草地、湿地资源监测研究

詹雅婷¹,戴宁江²,戎 欣¹,徐 钊³,朱叶飞¹,郑 刚³,祝亚云⁴,戚阳艳³

(1. 江苏省地质调查研究院/自然资源江苏省卫星应用技术中心,江苏 南京 210018;2. 江苏省林业局,江苏 南京 210036;
3. 江苏省森林资源监测中心,江苏 南京 210036;4. 江苏省林业科学研究院,江苏 南京 211153)

摘要:森林、草地、湿地资源的消长动态和变化趋势的及时掌握是评估生态状况和供给能力的基础。为高效利用遥感卫星数据开展林地、草地、湿地资源监测,该研究以 DeepLabV3+模型为基础,并作出以下改进:(1)骨干网络的改进,以 Xception 作为模型的骨干结构;(2)空洞空间金字塔池化模块的改进,首先在并联的每个分支的前一层增加一个空洞率相对较小的空洞卷积,其次在分支后层加入串联的空洞率逐渐减小的空洞卷积层。以江苏省无锡市江阴市为研究区,利用高分二号卫星影像,并结合全国第三次国土调查及其国土变更调查数据实现林地、草地、湿地资源信息的获取。结果表明:顾及全国第三次国土调查资料的 DeepLabV3+算法能有效利用遥感图像中的信息,提升林地、草地、湿地资源监测的效率和准确性。

关键词:高分二号;林地;草地;湿地;资源;DeeplabV3+;国土调查

中图分类号:TP79 **文献标志码:**A **doi:**10.3969/j.issn.1001-7380.2023.03.007

森林、湿地、草地资源是“山水林田湖草沙”的重要组成部分,是陆地生态空间的主体和生态系统最大的碳库。因此,对林地、草地、湿地等林业资源的实时高效监测管理至关重要。传统的监测方法以野外调研为主,费时费力,成本高且效率低下,无法满足对自然资源长期实时高效监测管理的需要^[1]。卫星遥感数据是地球表面信息的重要来源,通过遥感数据,能够快速且准确获取大规模的地面信息^[2]。

利用遥感技术对植被信息的提取是开展传统林业资源遥感监测的关键。该技术主要是获取相关植被遥感指标,例如比值植被指数、归一化植被指数等植被监测遥感指标^[3]。董立新^[4]等利用 Landsat TM 计算的植被指数通过多元逐步回归模型和 PCA 模型等估算得到三峡库区区域尺度的森林 LAI,但需要根据季节不断选择模型。陈艺华^[5]等利用高分卫星影像通过光谱特征、形状特征、纹理特征等通过贝叶斯分类法等面向对象方法提取草地,获得了较高的精度,但颜色和纹理特征会对人工草地的提取产生一定影像。Han^[6]等利用 Landsat TM 数据通过 SVM 方法对珠三角的湿地进行提取,

精度较高,但是 SVM 等机器学习方法模型结果简单,无法表征复杂的地物特征。

近年来,随着深度学习的快速发展,在遥感图像分类和提取任务中的应用也越来越广泛^[7]。卷积神经网络作为深度学习中最常用的模型^[8],它从数据中自适应地学习到层次化的特征表达,实现对数据基本特征的表达,可以胜任高分辨率遥感影像的分类^[9]。江涛等^[10]基于 GF-2 影像采用 CNN 对林地进行提取,结果表明 CNN 能够挖掘深层特征,从而提高提取精度,但是在提取的过程中会出现噪声和边界不确定性问题。马建^[11]等利用 CNN 对高光谱遥感数据进行草地识别,证明了 CNN 在草地提取方面的可行性,但是采用的模型结果过于简单,相较于传统的 SVM 并没有提升效果。Liu^[12]等提出一种融合多光谱影像和高分影像的 CNN 模型,较好地提取了黄河口湿地和盐城湿地,但是在精细尺度上提取效果较差。

除了上述的模型外,DeepLabV3+^[13]网络作为最典型的语义分割网络由谷歌提出。采用编码器解码器结构,使用空洞空间金字塔模块和空洞卷积,能够

收稿日期:2023-05-11;修回日期:2023-06-05

作者简介:詹雅婷(1985-),女,高级工程师,博士研究生,遥感科学与技术专业,主要从事自然资源遥感调查监测研究工作,E-mail:zytayx@sina.com

提取丰富的多尺度空间上下文信息,在提升分割效果的同时还关注了地物边界信息。但是模型存在对小尺度地物错分漏分和边界划分粗糙的问题。

基于上述研究背景,本研究结合深度学习方法,以 DeepLabV3+网络为基础进行改进,并结合全国第三次国土调查(以下简称“三调”)及其国土变更调查数据,实现对遥感影像的林地、草地、湿地等资源的智能提取。本研究的优势在于:(1)充分利用遥感、模型、大数据、RTK 等先进技术手段,以“三调”及其国土变更调查数据为本底,通过深度学习识别结合现地调查方式,协同开展江苏省内林地、草地、湿地等自然资源的实时监测管理需要。(2)本文通过对 DeepLabV3+模型的改进,特别是对其空洞空间金字塔模块的改进,有效提升了小尺度地物和地物边界的分类效果。

1 研究区概况与数据预处理

1.1 研究区概况

本研究以江苏省无锡市江阴市为研究区,地理位置为东经 119°59′—120°34′,北纬 31°40′—31°57′之间,总面积 987.50 km²。

江阴市自然植被为北亚热带落叶常绿阔叶混交林,林地面积 1.64 万 hm²,草地 0.51 万 hm²,自然湿地面积 660 hm²。

1.2 数据预处理

高分二号卫星于 2014 年 8 月成功发射,2015 年 3 月投入使用,广泛应用于土地利用动态监测、城乡规划管理和评价、交通运输管理和林业生态工程监测等领域,其有效技术载荷指标如表 1。

对获取的高分二号卫星影像进行辐射校正、正射校正、影像融合、拼接镶嵌、匀光匀色等预处理后

按照研究区的矢量范围进行剪裁,得到实验样本数据,并通过目视解译方法对测试区域制作真实地物的样本标签。

表 1 高分二号有效技术载荷指标

参数	1 m 分辨率全色/ 4 m 分辨率多光谱相机	波段范围
光谱范围	全色	0.45–0.90 μm
		0.45–0.52 μm
	多光谱	0.52–0.59 μm
		0.63–0.69 μm
		0.77–0.89 μm
空间分辨率	全色	1 m
	多光谱	4 m
幅宽	45 km(2 台相机结合)	

2 研究方法

2.1 模型整体架构

改进的 DeepLabV3+的网络结构如图 1 所示。模型总体上采用了 DeepLabV3+的架构。模型总体采用编码器-解码器结构。编码器主要包括 2 个部分:带空洞的卷积神经网络(Dilated Convolution Neural Network,DCNN)和增强空洞金字塔池化模块(Atrous Spatial Pyramid Pooling Plus,ASPP+),以带空洞卷积的 Xception 架构为骨干网络,进行特征提取;通过 ASPP+模块提取不同尺度的空间上下文信息。在编码器中,骨干网络提取的高级特征进行并联的 ASPP+模块后输出多通道的特征图,进入 1×1 卷积改变通道进入解码器。解码器中,先使用 1×1 卷积将多个低层级特征图融合,之后经过 4 倍上采样的编码器输出特征图合并,经过 3×3 卷积细化后,使用双线性差值上采样 4 倍得到预测结果。

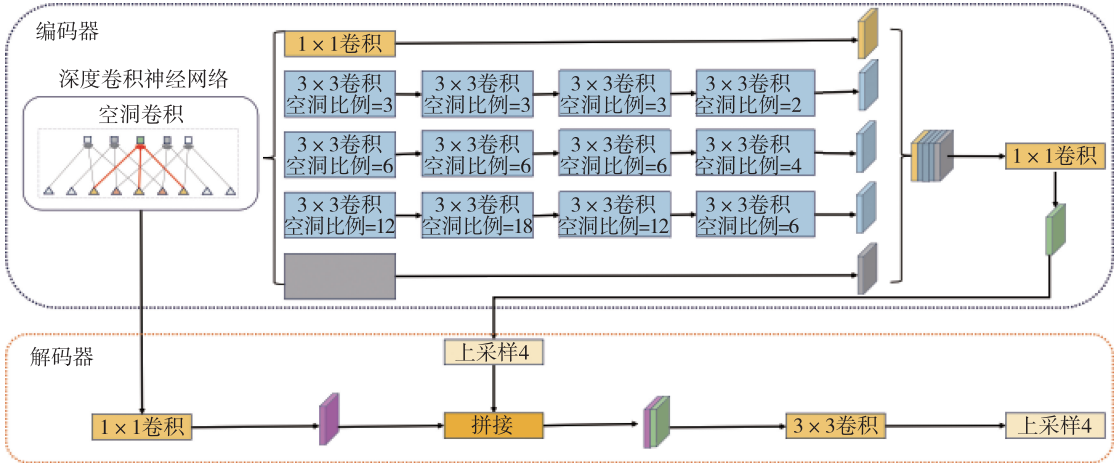


图 1 改进的 DeepLabV3+网络结构

2.2 ASPP+模块

在遥感图像分类的过程中,空洞空间金字塔池化层(ASPP)对小尺度地物的特征表达能力欠缺,且过大的空洞率会使得遥感图像的局部空间信息丢失,削弱相邻空间的一致性,导致网络对小尺度地物特征提取能力的不足^[14]。为了进一步加强模型对小尺度地物的提取能力,首先在并联的每个分支的前一层增加一个空洞率相对较小的空洞卷积,提取更多小尺度地物的特征,加强特征表达能力;其次在分支后层加入空洞率逐渐减小的空洞卷积层,在更深的层次中逐步提取多尺度的局部空间信息,提高空间一致性,改进后的模块结构如图2所示。

2.3 Xception 模块

DeepLabV3+网络使用的主体网络是 ResNet 结

构,其通过残差连接解决了深层神经网络中的梯度消失问题,使得训练更深的网络变得更加稳定和有效。但是会导致网络变得很冗大,拥有大量的参数。这导致训练和推理都会带来较高的计算成本,且标准卷积操作会在空间和通道维度上同时执行,导致计算效率较低。Xception^[15]将深度可分离卷积将卷积操作分解成逐通道的逐点卷积和深度卷积,也就是通过使用深度可分离卷积^[16](Separable Convolution)。深度可分离卷积的使用有助于减少参数数量,提高计算效率,同时还可以更好地捕获通道间的相关性,提高模型的表示能力,且 Xception 模块对小尺寸特征的处理更为优秀,能更好地处理小尺寸的地物(如图3)。

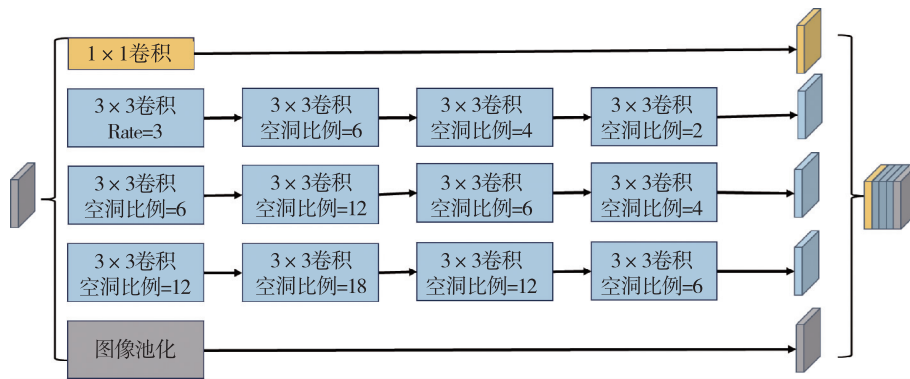


图2 ASPP+模型结构

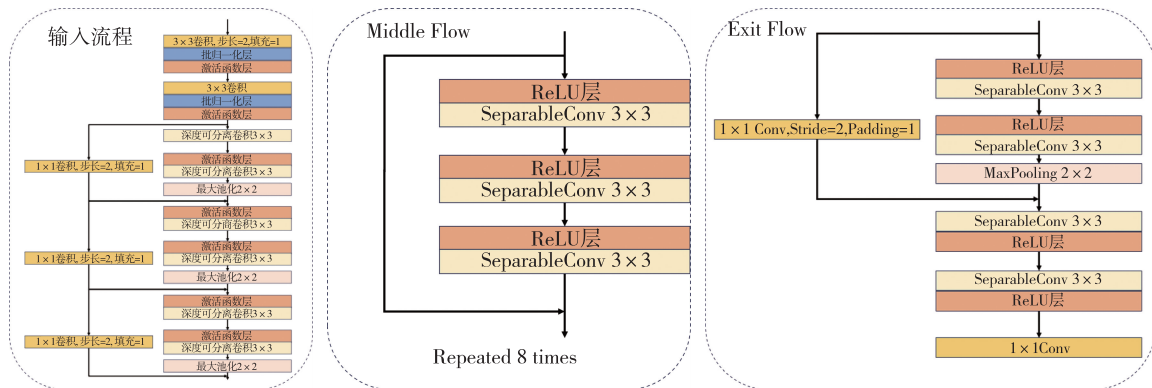


图3 Xception 结构图

3 实验与分析

3.1 深度学习样本制作

使用 ArcMap 工具对影像中部分典型区域的林地目视解译标注,制作深度学习样本数据库。由于设备和网络输入尺寸的限制,将标注好的

标签数据与对应的原始影像数据相匹配,利用 Python 采用滑动窗口采样的方式将 2 者按位置匹配裁剪为 256×256 大小的子图。将数据集按照 8:2 的比例随机划分为训练集和验证集。训练集用于训练模型,验证集用于对模型进行调参。

3.2 模型训练与优化

在训练过程中,使用 python 语言中的 PIL, skimage, simplecv 等多种第 3 方库对遥感图像影响及标签进行数据增强处理,使用 numpy 等第 3 方库实现矩阵运算;相关模型在 Pytorch 框架下搭建完成,在 2 块 Nvidia Quadro P4000 高算力的支持下完成模型训练优化。

本文采用分类问题中最常用的交叉熵函数作为损失函数。交叉熵(Cross Entropy, CE)是给定随机变量或事件集的 2 个概率分布之间差异的度量,它描述 2 个概率分布之间的差异,和模型的优化方向一致,因此常用它来衡量模型的损失。与其他损失函数相比,交叉熵损失函数收敛快,利于网络的优化,在大部分语义分割场景中均有应用。交叉熵误差的计算公式如

$$CE = -\frac{1}{N} \sum_{i=1}^n \sum_{j=1}^k t_{ij} \lg y_{ij}$$

式中, N 为 1 次训练的样本总数, y_{ij} 为第 i 个训练样本的第 j 个元素的预测值, t_{ij} 为第 i 个样本独热编码的第 j 个元素。

在训练过程中,将训练集进行数据翻转、随机裁剪等增广方式以增强模型的泛化能力,设置 epochs 为 50,初始学习率为 0.001。为了防止训练后期由于学习率过大可能导致的模型不收敛问题,设置学习率衰减因子为 0.8,每隔 5 个 epoch 衰减 1 次。

作为反向传播的核心内容,梯度下降算法是最基本的一类优化器。本文采用 Adam 优化器,其计算高效,对内存需求少,且参数的更新不受梯度伸缩变化的影响。

3.3 精度评价

本文使用以下 3 个指标进行提取结果的定量评价,分别为总体精度, Kappa 系数,综合精度得分 ($F1-score$)。

(1) 总体精度 (Overall Accuracy, OA) 总体精度为每个类别被正确分类的样本之和与提取的总样本数的比值。公式如:

$$OA = \frac{1}{N} \sum_{i=1}^m C_i$$

式中, N 为总体像元数, m 为分类数, C_i 为被正确分为第 i 类的像元数量。

(2) Kappa 系数 Kappa 系数主要用于分类一致性,其可以考虑模型对样本的偏向性,当分类样

本数量不平衡时,使用 Kappa 系数比总体精度更具说服力。Kappa 系数公式如:

$$Kappa = \frac{OA - \frac{\sum_{i=1}^m (a_i \times b_i)}{N \times N}}{1 - \frac{\sum_{i=1}^m (a_i \times b_i)}{N \times N}}$$

式中, a_i 为第 i 类像元的实际个数; b_i 为被预测为第 i 类的像元个数。

(3) $F1-score$ $F1-score$ 是适合于样本不平衡情况的精度评价指标,将真实值定为负例 (Negative, TN); 预测值也为负例的情况定位真反例 (True Negative, TN); 真实值为正例,预测值也为正例的情况定义为真正例 (True Positive, TP); 真实值为负例,预测值却为正例的情况定义为假正例 (False Positive, FP); 真实值为正例,预测值却为负例的情况定义为假反例 (False Negative, FN)。 $F1-score$ 的计算公式如:

$$F1 - score = \frac{2 \times \frac{TP}{TP + FP} \times \frac{TP}{TP + FN}}{\frac{TP}{TP + FP} + \frac{TP}{TP + FN}}$$

4 结果与讨论

4.1 结果

为了探究改进的 DeepLabV3+网络对林地、草地、湿地等提取的效果,本文选取无锡市江阴市作为精度验证研究区。改进的 DeepLabV3+网络对江阴市的卫星影像提取结果如图 4 所示。可以看到,模型对于影像中的林地、草地、湿地资源有着良好的提取效果。

为了量化评价模型的提取效果,将用模型提取到的林地、草地、湿地资源结果进行精度评价,采用总体精度和 Kappa 系数,结果如表 2 所示。可以看出,利用深度学习搭建的改进的 DeepLabV3+网络对于林地、草地、湿地等自然资源有着较好的提取效果。

表 2 深度学习林地提取精度评价

评价指标	林地	草地	湿地
总体精度	86.73%	87.76%	88.08%
Kappa 系数	84.47%	85.67%	86.04%

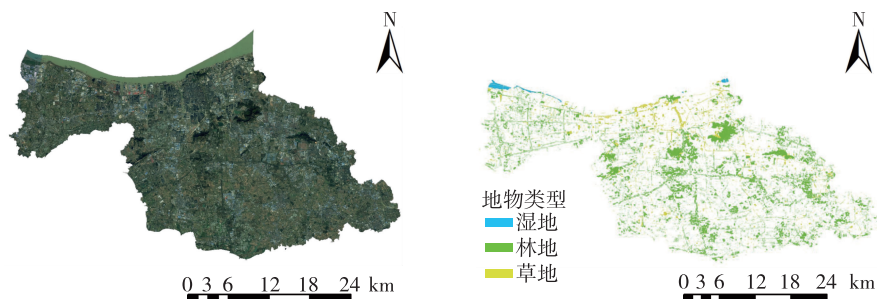


图4 江阴市高分二号影像及模型提取结果

4.2 讨论

本文基于高分二号数据,以 DeepLabV3+网络为基础进行改进,使用了以带空洞卷积的 Xception 模块为改进模型的主体模块,提取更多更充分的细节信息以提升模型的分类效果,并对原模型的 ASPP 模块进行改进,以获得更加细节的多尺度空间上下文信息,并结合“三调”及其国土变更调查数据,实现对遥感影像的林地、草地、湿地等资源的智能提取。

随后研究对模型的效果进行评价和分析。其提取林地的总体精度为 86.73%, kappa 系数为 84.47%;其提取草地的总体精度为 87.76%, kappa 系数为 85.67%;其提取的湿地总体精度为 88.08%, kappa 系数为 86.04%。总体来说,提取的林地、草地、湿地精度较高,能够满足林地、草地、湿地监管的需要。实验表明,本研究的改进模型在“三调”及其国土变更调查数据的基础上,能够有效地提取卫星影像中的林地、草地、湿地等自然资源,能够实现对这些自然资源长期实时监测管理。

但是,本研究仍然存在着不足。(1)本文使用了 Xception 作为主体架构用以增强特征提取的效果,但 Xception 模块会引入额外的参数,带来额外的计算量,需要额外的算力去运行,日后可以尝试其他的轻量化但精度也较高的模型,例如 MobileNetV2 等架构。(2)本研究虽然结合了“三调”及其国土变更调查数据,但只是用作于深度学习提取林地、草地、湿地等自然资源的后处理工作上,未充分挖掘利用数据,日后可以将“三调”及其国土变更调查数据转化为深度学习样本中的一些特征,更加充分的利用数据,以得到更好的效果。

参考文献:

[1] 赵宪文,李崇贵.基于信息技术的森林资源调查新体系[J].北京林业大学学报,2002,24(5):9.
[2] GONG P, WANG J, YU L, et al. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat

TM and ETM+ data[J]. International Journal of Remote Sensing, 2013, 34(7): 2607-2654.

- [3] 方臣,胡飞,陈曦,等.自然资源遥感应用研究进展[J].资源环境与工程,2019,33(4):563-569.
[4] 董立新.三峡库区森林叶面积指数多模型遥感估算[J].国土资源遥感,2019,31(02):73-81.
[5] 陈艺华,陈振杰.一种面向对象的人工草地遥感监测方法[J].江苏农业学报,2021,37(6):1545-1553.
[6] HAN X, PAN J, DEVLIN A T. Remote sensing study of wetlands in the Pearl River Delta during 1995 - 2015 with the support vector machine method [J]. Frontiers of Earth Science, 2018, 12: 521-531.
[7] ZHANG S, CHAU K W. Dimension reduction using semi-supervised locally linear embedding for plant leaf classification [C]//Emerging Intelligent Computing Technology and Applications: 5th International Conference on Intelligent Computing, ICIC 2009, Ulsan, South Korea, September 16-19, 2009. Proceedings 5. Springer Berlin Heidelberg, 2009: 948-955.
[8] 史加荣,马媛媛.深度学习的研究进展与发展[J].计算机工程与应用,2018,54(10):1-10.
[9] KRIZHEVSKY A, SUTSKEVER I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25.
[10] 江涛,王新杰.基于卷积神经网络的高分二号影像林分类型分类[J].北京林业大学学报,2019(9):20-29.
[11] 马建,刘文昊,靳瑰丽,等.基于CNN和SVM的地面高光谱遥感草地植物识别[J].草业科学,2023,40(2):394-404.
[12] LIU C, TAO R, LI W, et al. Joint classification of hyperspectral and multispectral images for mapping coastal wetlands[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 14: 982-996.
[13] 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.
[14] 衡雪彪,许捍卫,唐璐,等.基于改进全卷积神经网络模型的土地覆盖分类方法研究[J].地球信息科学学报,2023,25(3):495-509.
[15] Chollet F. EXCEPTION: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 1251-1258.
[16] SIFRE L, MALLAT S. Rigid-motion scattering for texture classification[J]. International Journal of Computer Vision, 2014.