



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2020
12
VOL.25

ISSN1006-8961
CN11-3758/TB



数字浮雕建模 P2647



第25卷第12期 (总第296期)
2020年12月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿, 均视为同意在本刊网站及CNKI等全文数据库出版, 所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用, 所有内容, 未经许可, 不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院空天信息创新研究院

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 吴一戎

编辑出版 《中国图象图形学报》编辑出版委员会

通信地址 北京市海淀区北四环西路19号

邮 编 100190

电子信箱 jig@radi.ac.cn

电 话 010-58887035

网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

(邮政信箱: 北京399信箱 邮编: 100048)

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Aerospace Information Research Institute, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong

Editor, Publisher Editorial and Publishing Board of Journal of
Image and Graphics

Address No. 19, North 4th Ring Road West, Haidian District,
Beijing, P. R. China

Zip code 100190

E-mail jig@radi.ac.cn

Telephone 010-58887035

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China))

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

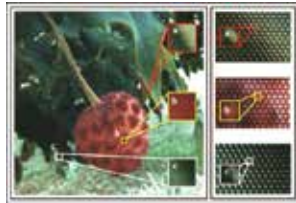
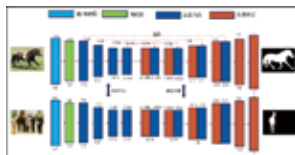
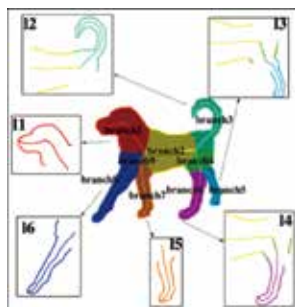
ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

国内定价 60.00元

光场显著性检测研究综述
(第2465页)融合注意力机制与知识蒸馏
的孪生网络压缩(第2563页)面向可解释性的物体拓扑结
构骨架表征方法(第2587页)

综述

光场显著性检测研究综述

刘亚美, 张骏, 张旭东, 孙锐, 高隼 2465

图像处理和编码

自适应卷积的残差修正单幅图像去雨

王美华, 何海君, 李超 2484

局部特定空间关系统计特征的RVIN噪声检测器

于海雯, 易昕炜, 徐少平, 林珍玉, 刘蕊蕊 2494

全局与局部属性一致的图像修复模型

孙劲光, 杨忠伟, 黄胜 2505

图像分析和识别

关键语义区域链提取的视频人体行为识别

马淼, 李贻斌, 武宪青, 高金凤, 潘海鹏 2517

针对形变与遮挡问题的行人再识别

史维东, 张云洲, 刘双伟, 朱尚栋, 暴吉宁 2530

多特征融合的行为识别模型

谭等泰, 李世超, 常文文, 李登楼 2541

结构特征下的可撤销人脸识别

孙浩浩, 邵珠宏, 尚媛园, 陈滨, 赵晓旭 2553

融合注意力机制与知识蒸馏的孪生网络压缩

耿增民, 余梦巧, 刘峡壁, 吕超 2563

联合聚焦度和传播机制的光场图像显著性检测

李爽, 邓慧萍, 朱磊, 张龙 2578

图像理解和计算机视觉

面向可解释性的物体拓扑结构骨架表征方法

危辉, 余莉萍 2587

自监督深度残差函数映射网络的3维模型对应关系计算

杨军, 马中昌 2603

融合自编码器和one-class SVM的异常事件检测

胡海洋, 张力, 李忠金 2614

抗高光的光场深度估计方法

王程, 张骏, 高隼 2630

计算机图形学

不同类型高度图控制浅浮雕模型生成方法

尚菁, 余文杰, 王美丽 2647

遥感图像处理

残差密集空间金字塔网络的城市遥感图像分割

韩彬彬, 张月婷, 潘宗序, 台宪青, 李芳芳 2656

结合判别相关分析与特征融合的遥感图像检索

葛芸, 马琳, 储珺 2665

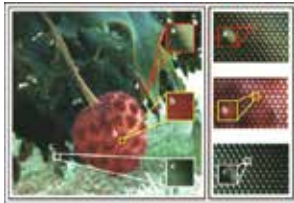
改进卷积网络的高分遥感图像城镇建成区提取

侯博文, 闫冬梅, 郝伟, 黄青青, 苏秀琴, 李青雯 2677

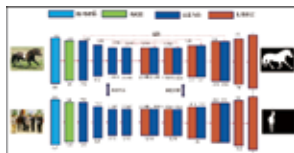
总目次 I

CONTENTS

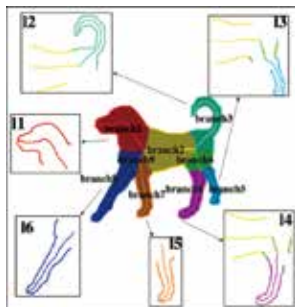
JOURNAL OF IMAGE AND GRAPHICS



Review of saliency detection on light fields(P2465)



Combining attention mechanism and knowledge distillation for Siamese network compression (P2563)



Skeleton characterization of object topology toward explainability (P2587)

Review

Review of saliency detection on light fields

Liu Yamei, Zhang Jun, Zhang Xudong, Sun Rui, Gao Jun 2465

Image Processing and Coding

Single image rain removal based on selective kernel convolution using a residual refine factor

Wang Meihua, He Haijun, Li Chao 2484

Random-valued impulse noise detector using local spatial structure statistics

Yu Haiwen, Yi Xinwei, Xu Shaoping, Lin Zhenyu, Liu Ruirui 2494

Image inpainting model with consistent global and local attributes

Sun Jinguang, Yang Zhongwei, Huang Sheng 2505

Image Analysis and Recognition

Human action recognition in videos utilizing key semantic region extraction and concatenation

Ma Miao, Li Yibin, Wu Xianqing, Gao Jinfeng, Pan Haipeng 2517

Person re-identification based on deformation and occlusion mechanisms

Shi Weidong, Zhang Yunzhou, Liu Shuangwei, Zhu Shangdong, Bao Jining 2530

Multi-feature fusion behavior recognition model

Tan Dengtai, Li Shichao, Chang Wenwen, Li Denglou 2541

Cancelable face recognition with fusion of structural features

Sun Haohao, Shao Zhuhong, Shang Yuanyuan, Chen Bin, Zhao Xiaoxu 2553

Combining attention mechanism and knowledge distillation for Siamese network compression

Geng Zengmin, Yu Mengqiao, Liu Xiabi, Lyu Chao 2563

Saliency detection on a light field via the focusness and propagation mechanism

Li Shuang, Deng Huiping, Zhu Lei, Zhang Long 2578

Image Understanding and Computer Vision

Skeleton characterization of object topology toward explainability

Wei Hui, Yu Liping 2587

Correspondence Calculation of 3D Models by a Self-Supervised Deep Residual Functional Maps Network

Yang Jun, Ma Zhongchang 2603

Anomaly detection with autoencoder and one-class SVM

Hu Haiyang, Zhang Li, Li Zhongjin 2614

Anti-specular light-field depth estimation algorithm

Wang Cheng, Zhang Jun, Gao Jun 2630

Computer Graphics

Bas-relief generation controlled by different height maps

Shang Jing, Yu Wenjie, Wang Meili 2647

Remote Sensing Image Processing

Residual dense spatial pyramid network for urbanremote sensing image segmentation

Han Binbin, Zhang Yueting, Pan Zongxu, Tai Xianqing, Li Fangfang 2656

Remote sensing image retrieval combining discriminant correlation analysis and feature fusion

Ge Yun, Ma Lin, Chu Jun 2665

Urban built-up area extraction using high-resolution remote sensing images with an improved convolutional neural network

Hou Bowen, Yan Dongmei, Hao Wei, Huang Qingqing, Su Xiuqin, Li Qingwen 2677

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2020)12-2677-13

论文引用格式: Hou B W, Yan D M, Hao W, Huang Q Q, Su X Q and Li Q W. 2020. Urban built-up area extraction using high-resolution remote sensing images with an improved convolutional neural network. Journal of Image and Graphics, 25(12): 2677-2689 (侯博文, 闫冬梅, 郝伟, 黄青青, 苏秀琴, 李青雯. 2020. 改进卷积网络的高分遥感图像城镇建成区提取. 中国图象图形学报, 25(12): 2677-2689) [DOI:10.11834/jig.190539]

改进卷积网络的高分遥感图像城镇建成区提取

侯博文^{1,4}, 闫冬梅^{2,3}, 郝伟¹, 黄青青², 苏秀琴¹, 李青雯^{2,4}

1. 中国科学院西安光学精密机械研究所, 西安 710119; 2. 中国科学院空天信息创新研究院, 北京 100094;

3. 海南省地球观测重点实验室, 三亚中科遥感研究所, 三亚 572029; 4. 中国科学院大学, 北京 100049

摘要: 目的 城镇建成区是城镇研究重要的基础信息,也是实施区域规划、落实城镇功能空间布局的前提。但是遥感影像中城镇建成区的环境复杂,同时不同城镇建成区在坐落位置、发展规模等方面存在许多差异,导致其信息提取存在一定困难。**方法** 本文基于面向图像语义分割的深度卷积神经网络,使用针对特征图的强化模块和通道的注意力模块,对原始 DeepLab 网络进行改进,并通过滑动窗口预测、全连接条件随机场处理方法,更准确地实现城镇建成区提取。同时,针对使用深度学习算法容易出现过拟合和鲁棒性不强的问题,采用数据扩充增强技术进一步提升模型能力。**结果** 实验数据是三亚和海口部分地区的高分二号遥感影像。结果表明,本文方法的正确率高于 93%, Kappa 系数大于 0.837,可以有效地提取出大尺度高分辨率遥感影像中的城镇建成区,且提取结果最为接近实际情况。**结论** 针对高分辨率遥感卫星影像中城镇建成区的光谱信息多样化、纹理结构复杂化等特点,本文算法能在特征提取网络中获取更多特征信息。本文使用改进的深度学习方法,提出两种处理方法,显著提高了模型的精度,在实际大幅遥感影像的使用中表现优秀,具有重要的实用价值和广阔的应用前景。

关键词: 卷积神经网络;注意力机制;遥感图像;城镇建成区;信息提取

Urban built-up area extraction using high-resolution remote sensing images with an improved convolutional neural network

Hou Bowen^{1,4}, Yan Dongmei^{2,3}, Hao Wei¹, Huang Qingqing², Su Xiuqin¹, Li Qingwen^{2,4}

1. Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China;

2. Institute of Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;

3. Key Laboratory of Earth Observation, Institute of Remote Sensing of Sanya, Sanya 572029, China;

4. University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: **Objective** The urban built-up area is an important source of basic information for urban research and serves as a prerequisite for the regional planning and implementation of the spatial layout of urban functions. Given the recent developments in Earth observational technologies and improvements in the resolution of remote sensing images, accurately and efficiently extracting information on urban built-up areas has become possible. However, due to the complex environment of urban built-up areas in high-resolution remote sensing images and the variations in their locations and development scales, various forms of remote sensing image representations increase the difficulty of using traditional information extraction methods for urban built-up areas. Recent studies show that deep learning algorithms have significant advantages in processing of

收稿日期: 2019-10-22; 修回日期: 2020-02-19; 预印本日期: 2020-02-26

基金项目: 海南省重点研发计划项目 (ZDYF2018001)

Supported by: Hainan Key Research and Development Program Project (ZDYF2018001)

large-scale images. This paper then examines these deep learning algorithms and reviews previous research that apply deep convolutional neural network methods, which have been widely used in computer vision to extract information on urban built-up areas from high-resolution satellite images. This article also improves the application of computer image processing technology in the field of remote sensing. **Method** Semantic image segmentation is crucial in image processing and computer vision. This process recognizes an image at the pixel level and then labels the object category to which each pixel in the image belongs. Based on the deep convolutional neural network oriented to semantic image segmentation, this paper uses the refinement module for the feature map and the attention module of the channel domain to improve the original DeepLab v3 network. The feature refinement module accurately obtains relevant information between pixels and reduces the grid effect. Afterward, the network model processes the feature map through atrous spatial pyramid pooling. The decoding part of the network extracts the attention information of the channel domain and then weighs the low-level features to achieve a better representation and to restore the detailed information. Afterward, the urban built-up area is extracted via the sliding window prediction and full connection conditional random fields methods, both of which can be applied to extract urban built-up areas with better accuracy. However, the use of deep learning algorithms is prone to overfitting and poor robustness. Accordingly, data augmentation and extension are used to enhance the capabilities of the model. Specifically, we use rotation and filter operations while cutting the original training and verification data into 256×256 samples. **Result** Extracting information from remote sensing images involves an effective mining and category judgment of such information. The experimental data are taken from Gaofen-2 remote sensing images of Sanya and Haikou cities in Hainan Province, China. These images are specifically taken at the Qiongsan District of Haikou City and at the Tianya District, Jiyang District, and the sea surrounding the Jiaotouding Island of Sanya City. Given their weak sample processing ability, traditional classification algorithms have achieved an accuracy rate of no higher than 85% in the experiments. Meanwhile, deep learning methods, such as SegNet and DeepLab v3, have relatively high accuracy and better performance in extracting urban built-up area information from remote sensing satellite images. By using the refinement module for the feature map and the attention module of the channel domain, this paper improves the accuracy rate of the original DeepLab v3 network by 1.95%. Meanwhile, the proposed method has an accuracy rate of above 93%, a Kappa coefficient of greater than 0.837, a missed detection rate of less than 4.9%, and a false alarm rate of below 2.1%. This method can effectively extract urban built-up areas from large-scale high spatial resolution remote sensing images, and its extraction results are the closest to the actual situation. **Conclusion** The comparative experiment shows that the proposed method outperforms others in extracting urban built-up area information from high-resolution remote sensing satellite imagers with diverse spectral information and complex texture structure. Two processing methods are also proposed to significantly improve the accuracy of the model. Both the sliding window method and conditional random fields processing demonstrate an excellent performance in extracting information from high-resolution remote sensing images and show high application value for large-scale remote sensing images.

Key words: convolutional neural network (CNN); attention mechanism; remote sensing image; urban built-up area; information extraction

0 引言

自 20 世纪 90 年代以来,随着中国的经济增长,城镇的空间实现了大规模的扩展。为了有效地对城镇进行动态监测,做好规划管理和发展预测,传统方法已经很难满足客观需要(Chen 等,2003)。由于城镇建成区类内差异大,难以准确地确定图像中某一部分是否属于城镇建成区。因此,高效地从遥感影像中进行城镇建成区信息提取仍然是一个具有挑战性的研究课题。

遥感影像能够快速准确地获取地球表面信息,特别是随着高分辨率遥感影像的应用,使其成为地理信息系统(geographic information system, GIS)数据库以及地图更新的重要信息来源(李德仁,2003)。卫星遥感具有大尺度周期性获取地面信息的能力,其获取的影像可以作为城镇建成区信息提取的数据源(邓刘洋等,2018)。高分辨率遥感卫星获取的高空间分辨率影像中的地物目标空间结构和表层纹理特征十分清晰,同时地物的可辨析内部结构和边缘信息也同样清楚(胡晓曦等,2010),为准确有效地解译分析和信息提取提供良好的条件与基础。

遥感图像的解译分类和信息提取技术的发展经历了从最初目视人工解译到目前面向对象的方法。Welch(1980)突破性地使用 Landsat TM 影像进行假彩色合成,采用目视人工解译的方法,实现了城市建筑信息的提取,并简单分析了建成区面积与人口之间的关系。Klonowski 和 Koch(1997)实现了基于马尔可夫随机场理论对遥感图像中建筑物的提取。Lin 和 Nevatia(1998)提出感知分组,利用边缘检测提取出建筑物的矩形轮廓,进而确定城镇建筑物外形和位置的新方法。杨存建和周成虎(2000)分析了 TM 影像上各地物谱间结构,利用居住用地与其他地物不同的谱间特征,实现了居住用地的自动提取。赵萍等人(2003)基于决策树算法,在 SPOT 卫星影像上实现了居民地信息的自动提取。Unsalan 和 Boyer(2004)利用一定区域内的直线统计特征,在高分辨率遥感影像中识别、提取居民区信息。徐涵秋(2005)基于谱间特征和归一化指数分析,进一步实现了城市建筑用地的信息提取。Pesaresi 等人(2008)通过构造建筑区因子提出基于模糊准则的灰度共生矩阵提取方法,实现了在全色遥感影像中提取城市区域。Rodriguez-Galiano 等人(2012)使用随机森林算法实现了遥感图像的地表覆盖分类。过林等人(2012)基于 Radon 和树型小波变换提取出了遥感影像中的居民用地。陶超等人(2014)利用角点等局部不变特征,提出了高分辨率遥感影像居民地的检测方法。Rougier 等人(2016)基于主动学习算法完成了面向对象的高分辨率卫星图像城市植被类别信息提取。

基于传统方法的遥感图像分类和信息提取算法在不断发展的同时,研究人员对基于深度学习的相关算法也展开了广泛研究。李石华等人(2005)相继提出了基于神经网络和知识发现的遥感图像分类方法。Mnih 和 Hinton(2010)首次使用深度学习技术提取遥感影像中的道路信息,此后,深度学习技术逐渐应用于高分辨率遥感影像的分类、信息提取、变化监测等任务。高常鑫和桑农(2014)通过分层方法建立深度学习模型,完成了对影像的高精度分类。陈磊士等人(2018)使用卷积神经网络和 BP 神经网络两种分类器对 Landsat 多光谱影像进行城市建设用地信息提取,实现了较高的精度。

为进一步提高高分辨率遥感影像分类精度,使高分辨率遥感影像中丰富的信息得到充分利用,面

向对象的高空间分辨率影像分类方法应运而生,而面向对象的分类技术主要包括影像分割和分类两个核心步骤(杨惠雯,2018)。图像语义分割算法对光谱及空间特征较强的提取能力使更多的学者将其引入到遥感影像分类和信息提取中(Hu 等,2018)。Zhang 和 Du(2016)提出高分辨率遥感影像自适应分割并应用了多级分类器,使建筑物提取精度进一步提高。Pu 等人(2011)使用面向对象的图像分割算法对遥感图像进行分割,再用人工神经网络对分割后的图像进行分类,分类效果优于像元分类方法。

针对以上研究现状,结合城镇建成区的特点和实际应用需求,本文将探索应用深度学习技术实现于高分辨率遥感影像上准确提取城镇建成区信息,从而便于相关政府部门对城镇建成区的发展规模进行统计和分析。

1 深度学习技术

1.1 卷积神经网络

卷积神经网络(convolutional neural network, CNN)作为深度学习技术的一种经典模型,是受到生物视觉认知机制启发而专门设计的一种针对2维结构图像的多层感知器。近年来,在信号处理、模式识别、生物医学等领域都有着非常广泛并且成功的应用。

卷积神经网络的结构主要是卷积层、非线性映射层、池化层(也称下采样层)和全连接层。本世纪以来,数值计算设备能力的不断提高使得卷积神经网络在计算机视觉等领域取得快速发展,同时,海量数据的获取也推进了卷积神经网络结构的不断创新和优化。基于卷积神经网络算法对高分辨率遥感图像进行分类和信息提取,可有效避免特征提取和判断过程中数据重建的复杂度。相比传统的非监督分类方法,基于卷积神经网络的遥感图像分类算法在分类精度方面有所提高;与此同时,卷积神经网络利用卷积层和池化层分别降低噪声影响、聚合遥感影像中不同空间位置特征特性,降低特征维度的同时,特征的尺度不变,并且在一定程度上避免了过拟合的现象。

卷积神经网络的早期应用是 LeCun 等人(1998)于20世纪90年代设计的 LeNet-5 网络模

型,该网络实现了对手写数字的识别分类。但由于数据处理设备能力的限制,同时缺乏大规模的数据对网络进行训练,该网络模型对复杂问题的处理结果并不理想。Krizhevsky 等人(2012)提出了一个经典的卷积神经网络结构,名为 AlexNet,在当时的 ImageNet 图像竞赛中取得了分类任务的冠军,其性能大幅超过了第 2 名。AlexNet 的成功让人们意识到卷积神经网络在计算机视觉等领域具有广阔的应用前景。

1.2 ResNet 网络结构

在卷积神经网络的训练过程中,由于梯度反向传播,网络深度的增加会导致网络的梯度消失问题。为解决这个问题,He 等人(2016)提出了 ResNet,即深度残差网络,在 ImageNet 图像竞赛的分类、检测和定位任务中均取得第 1 名。

ResNet 对卷积神经网络结构的优化效果明显,在标准卷积网络基础上,每隔几层进行直接绕道,每绕过一段就会产生一个瓶颈块结构(bottleneck),如图 1 所示。通过这种使用 1×1 卷积核巧妙改变特征映射维度的结构,网络能够有效地保护训练过程中信息的完整性,即使网络层逐渐加深,仍然可以保持较好的性能。

1.3 全卷积神经网络

全卷积神经网络(fully convolutional networks, FCN)是基于卷积神经网络结构改进而得到的,广泛应用于图像语义分割领域,即实现对图像的像素级分类。Long 等人(2015)首次使用全卷积网络对自然图像进行端到端分割和分类,实现了深度学习技术在图像语义分割方向的应用。

经典的卷积神经网络结构在卷积层之后通过若干个全连接层进行降维,得到一定长度的特征向量

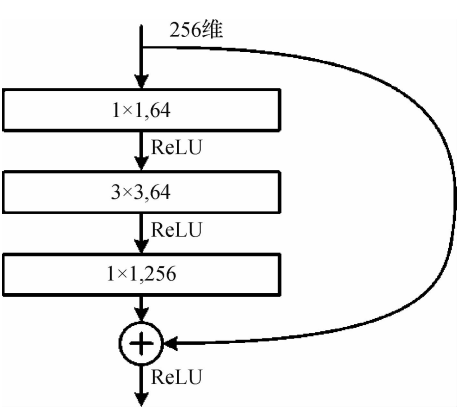


图 1 瓶颈块结构

Fig. 1 Bottleneck block

进行分类。而全卷积网络丢弃全连接层,使用反卷积层对最后一个卷积层的特征映射进行上采样,使图像恢复到与输入相同的尺寸,进而对每个像素都产生一个预测结果,同时保留了原始图像中的空间信息,在最后一个上采样层的特征图上进行逐像素类别信息判断。

1.4 DeepLab 架构

在图像语义分割任务中,为保证卷积特征分辨率不变的基础上实现感受野的扩大,Chen 等人(2014)提出了 DeepLab 系列语义分割模型,通过全部使用空洞卷积(atrous convolution)和空间维度金字塔结构空洞池化(atrous spatial pyramid pooling, ASPP),提取出更紧凑的特征信息。在此基础上,Chen 等人(2017)改进并提出了 DeepLab v3 架构。该架构以 ResNet 部分网络为编码器。对于空间金字塔的空洞池化,DeepLab v3 采用 4 个并行的不同尺度空洞卷积对特征映射进行处理,如图 2 所示,通过挖掘卷积特征并编码全局内容特征,实现不同尺度的特征采样,提升分类效果。

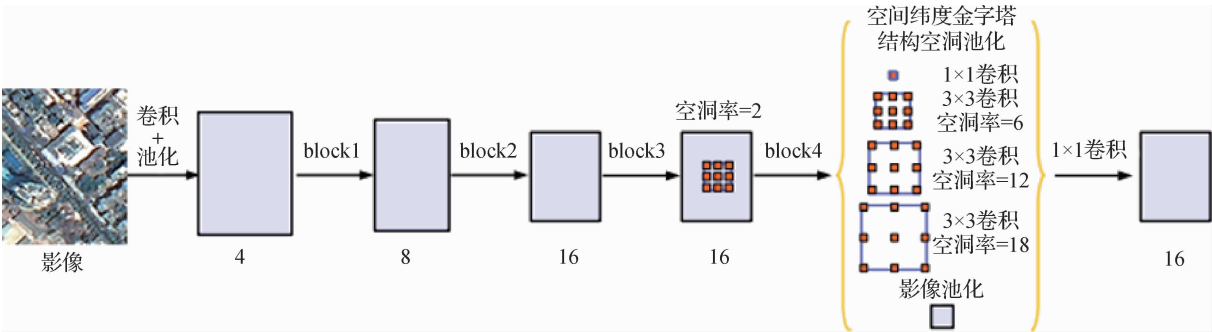


图 2 DeepLab v3 空间金字塔的空洞池化结构

Fig. 2 Atrous spatial pyramid pooling structure of DeepLab v3

2 改进的卷积网络模型

2.1 特征强化模块

原始网络采用空洞卷积扩大感受野,进而提取特征信息,但空洞卷积是通过在卷积核的中间补零操作实现的,实际上只获取了区域中非零部分信息,这种稀疏采样方式很可能使得特征图出现网格效应,进而导致网络中的一些提取结果来自上一层的

互相独立部分,丢失遥感影像的局部信息,网络模型学习到的特征也会缺少相关性。虽然原始网络并行使用了扩张率各异的空间卷积以提高特征信息的获取能力,但扩张率的大幅增加会导致有效采样点的减少,并且无法避免网格效应。网格效应会导致网络的训练不充分,学习到局部偏差严重的特征映射信息。针对这类问题,本实验设计了特征图的强化模块(refinement module, RM),如图3所示。

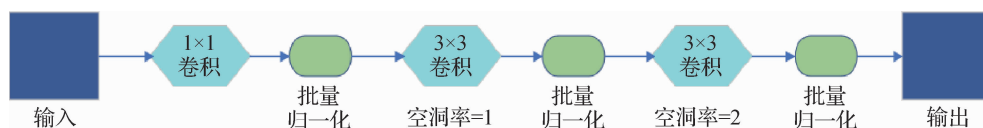


图3 强化模块

Fig.3 Refinement module

在原始网络的ASPP结构之前,增加了一个卷积单元和两个扩张率很小的空洞卷积,使得采样点更为密集、更加全面地提取特征图信息。为了提高网络对特征图的学习效率,强化模块先对图像降维,然后输入至扩张率为1的空洞卷积,具体为

$$(C * k)(p) = \sum_{a+b=p} C(a)k(b) \quad (1)$$

之后,输入到扩张率为2的空洞卷积单元,具体为

$$(C * k)(p) = \sum_{a+2b=p} C(a)k(b) \quad (2)$$

式中, $C(a)$ 表示卷积向量, $k(b)$ 为卷积核, a 为步长, b 为位置偏移量, $*$ 为卷积运算。

除此之外,模块通过批量归一化(batch normalization, BN)层对卷积单元的数据进行标准化处理,每层使用可训练的参数对批量数据进行归一化处理,进而提高运算效率,增强模型的泛化能力,还能加速网络的收敛,控制过拟合,在一定程度上避免因网络层过深而出现的梯度消失。

加入强化模块的算法模型对比结果如表1所示。可以看出,对于特征图的强化模块使得网络模型提取的特征更具有相关性,提高了遥感影像信息提取的准确率。

2.2 通道域注意力模块

实际上,卷积神经网络从浅层到深层,感受野不断增大,学习的特征信息也有着变化。在浅层网络上,模型的感受野较小,特征图分辨率较高,提取到的多是细节特征信息;而深层网络上,模型的感受野较大,特征图分辨率较低,获取的更多是语义特征信

表1 原始DeepLab v3与加入特征图强化模块的结果对比

Table 1 Compared results of DeepLab v3 with and without refinement module

算法	Kappa	正确率/%	漏检率/%	虚警率/%
DeepLab v3	0.769 1	89.94	7.40	2.66
加入RM的DeepLab v3	0.792 3	90.95	6.52	2.53

注:加粗字体表示各列最优结果。

息。在预测判别阶段,原始网络将最终的特征图进行16倍解码上采样,这一操作忽视了浅层网络提取的部分细节特征,模型提取边界信息的能力较差。为了避免简单的大倍率一次性上采样产生的误差,更好地提取遥感地物边界,本实验使用多次上采样的方法,结合不同尺度的特征图实现特征融合预测。在不同分辨率的特征图合并之前,本实验引入通道域的注意力模块(attention module, AM),对特征图的通道附加权重(Hu等,2018)。先通过池化对特征图进行挤压,具体为

$$z_c = F_{\text{squeeze}}(x_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (3)$$

式中, x 表示大小为 $C \times H \times W$ 的特征图,总通道数为 C 。 x_c 是此特征图中通道为 c 的2维矩阵。

再进行激活和归一化,具体为

$$s_c = F_{\text{excitation}}(z_c, W) = \delta(g(z_c, W)) = \delta(W_2 \sigma(W_1 z_c)) \quad (4)$$

式中,采用权重为 W_1 的全连接层将通道降维,中间通过softmax层 σ 进行激活,再使用权重为 W_2 的全

连接层将通道升维,接着用 ReLU 函数 δ 归一化输出权重。

这些权重通过 scale 函数加权到每个特征通道上,进而完成特征通道的权重分配,具体为

$$y_c = F_{\text{scale}}(x_c, s_c) = x_c \cdot s_c \quad (5)$$

通道域的注意力模块如图 4 所示。网络卷积层的加深和遥感影像特征信息的丰富使得每个通道都

提取出了多种信息,这些特征信息与目标地物的相关程度也各有不同。原始网络对每个通道信息的权重相同,本文使用注意力模块对特征图通道加权,算法模型对比结果如表 2 所示。可以看到,采用通道域注意力模块的算法可以学习并选出对目标信息提取贡献较大的特征,实现更好的信息提取效果。

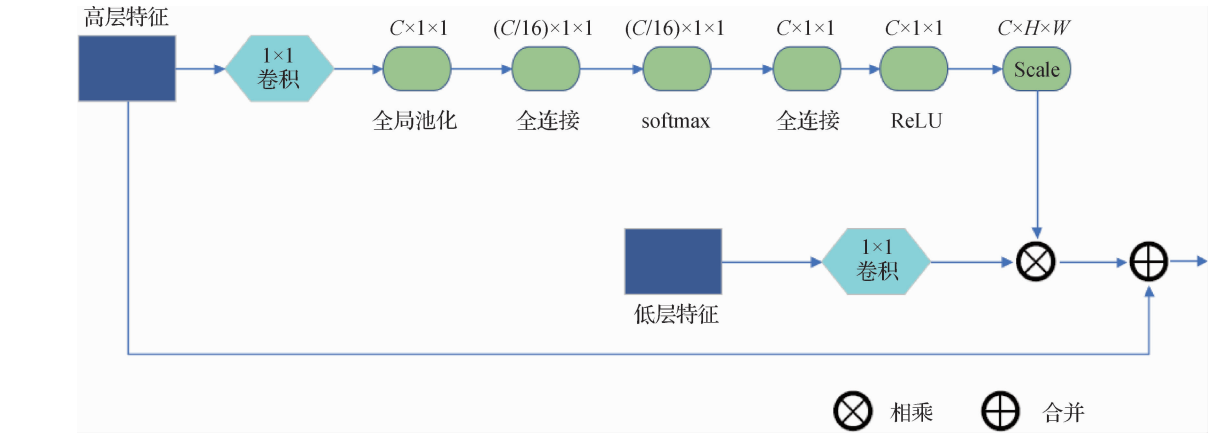


图 4 通道域的注意力模块

Fig. 4 Attention module of channel domain

表 2 原始 DeepLab v3 与引入通道域
注意力模块的结果对比
Table 2 Compared results of DeepLab v3
with and without attention module

算法	Kappa	正确率/%	漏检率/%	虚警率/%
DeepLab v3	0. 769 1	89. 94	7. 40	2. 66
引入 AM 的 DeepLab v3	0. 804 4	91. 17	6. 34	2. 49

注:加粗字体表示各列最优结果。

2.3 改进的卷积网络结构

输入的遥感影像经过原始网络的编码器提取后,特征强化模块更准确地获取了像素间的相关信息,减少了网格效应。接着,网络模型通过空间维度金字塔结构空洞池化,使用不同尺度的并行空洞卷积对特征映射进行处理,提取特征信息。网络的解码部分将深层语义信息进行了与编码器相对应的 3 次上采样操作,并且把深层特征输入注意力模块,提取通道域的注意力信息,对低层特征加权,实现更强的表征性,并且有效地恢复细节信息。通道域注意力模块的训练过程,本质上就是利用网络对数据的学习和通道间的相关性来产生最佳的 scale 。随后,高层特征和筛选出的低层特征相融合,判别提取出较

为详细的地物边界,进一步提升了网络的信息提取精度。本文的网络模型如图 5 所示,实际应用中,能有效地学习并选出对城镇建成区信息贡献较大的特征,一定程度上减轻卷积网络处理大量数据的负担,更好地学习、拟合输入图像至输出图像的映射关系,实现更高的预测精度和更强的泛化能力。

2.4 滑动窗口预测法

一般来说,高分辨率遥感图像的尺寸较大,本文使用的研究区域内高分二号遥感影像也是如此。卷积神经网络结构在应用过程中,如果将大幅高分辨率遥感影像直接输入其中,很可能会导致显卡内存不足。未使用滑动窗口预测法时,实验在大尺寸遥感影像输入网络架构之前,将影像依次裁剪为固定大小的多个较小图像并分别进行预测,得到每个较小图像的结果后,通过拼接取得网络架构预测的影像。但是由于网络模型中的大量卷积处理,此种常规预测方式会使得每幅小图像边界像素的预测效果一般,将多幅小图像得到的结果拼接成原始影像大小后,部分位置会出现较为明显的拼接痕迹。

为了解决上述问题,本实验在网络架构预测之前使用滑动窗口方法。本实验将大尺寸的遥感影像

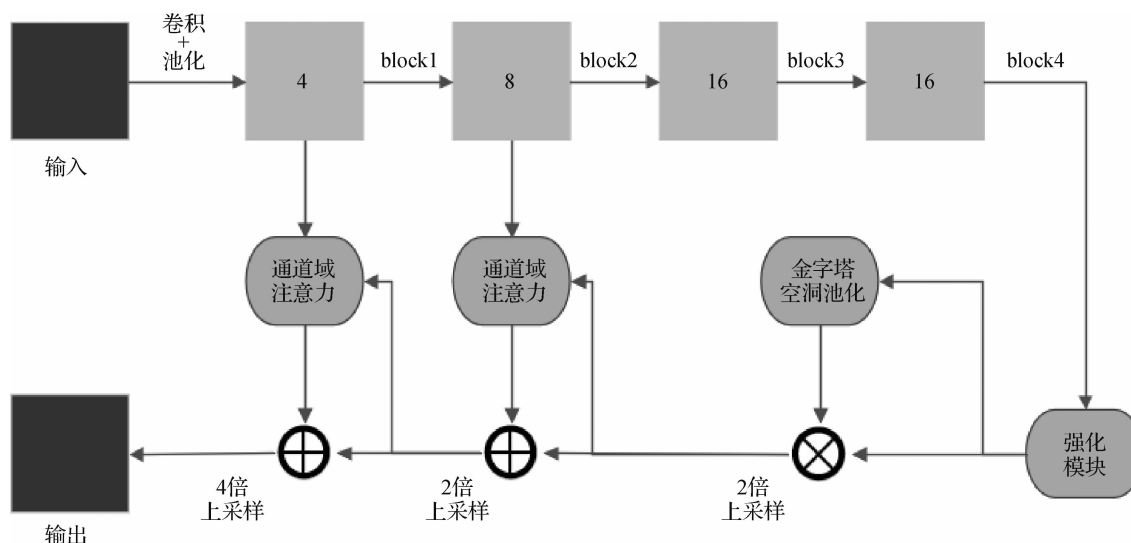


图5 改进的DeepLab v3结构

Fig. 5 Proposed DeepLab v3 structure

裁剪成多个 256×256 像素的小图像,但在裁剪过程中,设置小于图像大小的固定数值区域以便控制每次输入网络框架预测区域范围。本实验中窗口滑动参数为 64,滑动窗口预测的处理过程如图 6 所示。显然,原来的大尺寸遥感影像中大量像素会被裁剪窗口多次处理,网络框架对每个像素进行 1~16 次预测,得到结果并进行判断,进而得出它的类别信息。

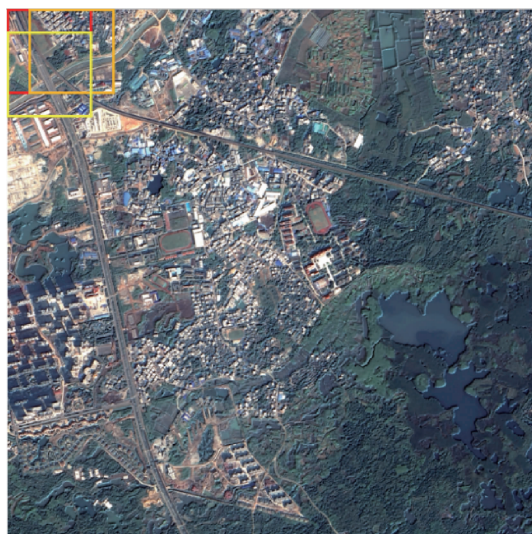


图6 滑动窗口法

Fig. 6 Sliding window method

2.5 全连接条件随机场

本文采用的全卷积神经网络架构使用非线性卷积层进行上采样操作,虽然能够将特征图恢复至原

始图像的尺寸,但也造成了特征映射的损失,产生提取目标边界模糊的问题。为了对提取地物的边界进行优化,得到更精确的分类结果,本实验进行了图像后处理。人们通常使用概率图模型解决机器学习相关算法产生的此类问题。概率图模型是用图论方法以表现数个独立随机变量之间概率相关关系的建模总称。其中,隐马尔可夫模型(hidden Markov model, HMM)和条件随机场(conditional random fields, CRFs)应用最为广泛。条件随机场能考虑原始输入图像的 RGB 颜色特征与空间位置关系,估计出网络模型预测结果的最大后验分布,进一步改善结果。条件随机场通过结合单像素的一元势能和相邻像素之间的二元势能,使得在空间上相邻的像素具有相同标签,进而平滑预测结果中物体的边缘。但对于高分辨率遥感影像,网络架构得到的结果中地物边缘还有模糊,需要进一步恢复细节信息。夏梦等人(2017)使用全连接条件随机场(DenseCRF)结合深度学习实现遥感图像分类,通过像素之间的预测值和实际空间距离衡量像素之间的相似性,对所有像素建立点对势能,从而最大程度地实现地物边缘处分割。

全连接条件随机场对网络模型的判断结果进行处理流程如图 7 所示。一元势能是每个像素的类别判断概率分布图,二元势能的颜色和相对位置信息可在原始影像中获得。本文使用平均场近似法(mean field approximation)迭代更新能量项,迭代

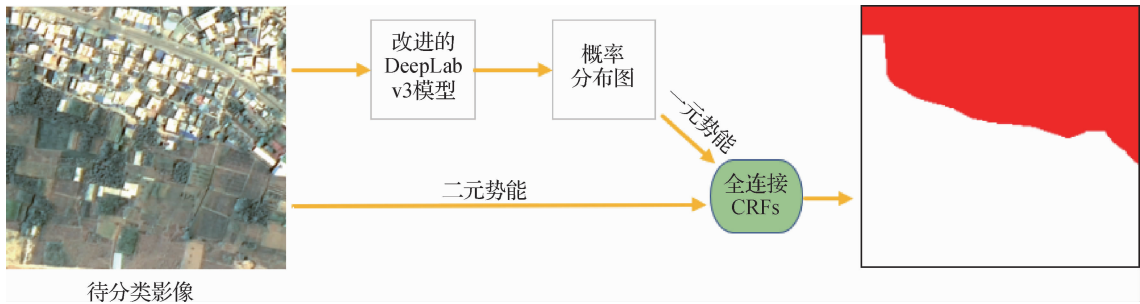


图 7 全连接条件随机场处理流程图

Fig. 7 Diagram of DenseCRFs processing

10 次,得到最终的结果。

作为后阶段的处理流程,全连接条件随机场使本文模型变为不是端到端的算法框架。原始的 DeepLab v3 网络框架(Chen 等,2017)并没有说明使用 DenseCRFs 对模型能力的影响。全连接条件随机场通过连接全局信息,融合了图像级特征,可以更好地提取出高分辨率遥感地物目标的边界,如图 8 所示,最终对比结果见表 3。实际上,针对遥感影像的信息提取,将全连接条件随机场引入到 DeepLab v3 + 网络框架,也能提升网络模型能力(王俊强 等,2019)。

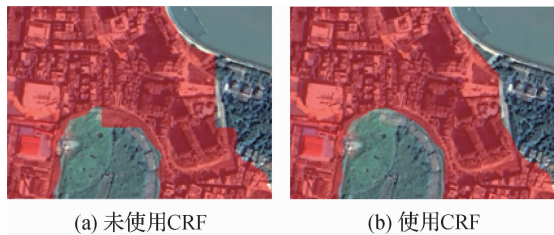


图 8 使用 CRF 与未使用 CRF 的对比图

Fig. 8 Compared images of this experience with and without CRF((a) without CRF; (b) with CRF)

表 3 使用 CRF 与未使用 CRF 的结果对比

Table 3 Compared results of this experience with and without CRF

算法	Kappa	正确率/%	漏检率/%	虚警率/%
未使用 CRF	0. 831 2	92. 63	5. 19	2. 18
使用 CRF	0. 837 5	93. 06	4. 85	2. 09

注:加粗字体表示各列最优结果。

3 城镇建成区信息提取

3.1 研究区域数据集制作

海南省是我国最南端的省份,也是我国唯一的

热带海岛省份。海南岛中部高,四周低平;属海洋性热带季风气候,热带雨林较多。本实验采用同时期的高分二号遥感数据,以海南省三亚市部分地区的遥感影像作为训练和验证数据,使用海南省海口市部分地区进行测试。

原始高分二号遥感影像包括全色波段和 4 个多光谱波段,其中多光谱波段有蓝(0. 45 ~ 0. 52 μm)、绿(0. 52 ~ 0. 59 μm)、红(0. 63 ~ 0. 69 μm)、近红外(0. 77 ~ 0. 89 μm)共 4 个波段。影像预处理依次进行了辐射定标、FLAASH(fast line-of-sight atmospheric analysis of spectral hypercubes)大气校正、正射校正、图像融合、波段选取与处理。

本实验是提取城镇建成区的信息,然而城镇建成区在高分遥感影像中的细节丰富,光谱特征复杂多样。因此,本实验选用具有代表性的 4 幅经过预处理的 1 m 分辨率 RGB 三波段 8 位遥感影像进行网络模型的训练,大小均为 6 400 × 6 400 像素,中心点坐标分别为东经 109°41'11"、北纬 18°32'03";东经 109°42'65"、北纬 18°23'83";东经 109°50'78"、北纬 18°31'45";东经 109°52'10"、北纬 18°25'57"。训练数据的地理位置如图 9 所示,以三亚市的天涯区、吉阳区和角头顶附近岛屿和海域为主。数据集真值由人工标注制作。由于样本过大和计算设备的限制,未经切割的大图样本不能直接输入到模型中训练和验证,所以在分类前需对样本进行切割,切割大小为 256 × 256 像素。

3.2 样本增强与扩充

深度学习算法的原理决定了其实现需要大量数据进行学习训练。由于卷积神经网络的参数极多,在数据量不足的情况下很容易出现过拟合问题,即随着神经网络对数据的学习和训练迭代次数的增加,其泛化能力反而会降低。



图9 训练数据的地理位置

Fig. 9 Geographic location of training data



图10 训练样本示例

Fig. 10 Examples of training sample

像素的百分比。虚警率(F_i)是指在整个图像中被预测为 i ,但预测错误的像素和图像所有像素的百分比。Kappa系数(Kappa)是一个统计量,用来衡量预测和地面真相之间的一致性。

各评价指标的具体定义为

$$Ac = \frac{\sum_{i=1}^K P_{ii}}{\sum_{i=1}^K t_i} \quad (6)$$

$$M_i = \frac{\sum_{j=1}^K P_{ij} - P_{ii}}{\sum_{j=1}^K t_j} \quad (7)$$

$$F_i = \frac{\sum_{j=1}^K P_{ji} - P_{ii}}{\sum_{j=1}^K t_j} \quad (8)$$

式中, P_{ab} 表示真实类别 a 被预测到类别 b 的像素点数, $t_a = \sum_b P_{ab}$ 表示实际属于 a 类别的像素点

针对此类问题,本次实验使用了数据扩充和增强技术。将原始的训练和验证数据切割为 256×256 像素样本的过程中,对数据进行旋转、滤波操作,如图10所示,参数均是合理范围内的随机数值。同时,实验使用有重叠的切割策略达到样本的扩充。最终,将样本数据扩充为80 000幅图像,大大增加了数据的多样性和算法的鲁棒性。

3.3 实验评价参数

为了更好地评估算法的准确性和有效性,本实验采用正确率、漏检率、虚警率和Kappa系数对实验结果进行精度评价。

正确率(Ac)是指在整个图像中正确分类像素和图像总像素数的百分比。漏检率(M_i)是指在整幅影像中实际上为 i ,但未被预测正确的像素和所有

总数。

$$Kappa = \frac{P_0 - P_c}{1 - P_c} \quad (9)$$

$$P_0 = \frac{\sum_{i=1}^K P_{ii}}{\sum_{i=1}^K t_i} \quad (10)$$

$$P_c = \frac{\sum_{i=1}^K (\sum_b P_{ib} \times \sum_a P_{ai})}{\sum_{i=1}^K t_i \times \sum_{i=1}^K t_i} \quad (11)$$

正确率、Kappa系数、漏检率和虚警率的值在0~1之间。前两者值越高、后两者值越低,模型的分类精度越好。

3.4 实验环境与模型训练

实验在Win10系统下,采用32 GB内存Intel Xeon E5-2620@2.00 GHz六核CPU处理器,内存为8 GB的NVIDIA GeForce GTX 1080-显卡(GPU处理单元)。实验使用Tensorflow1.10深度学习框架、

Nvidia 公司的 CUDA (compute unified device architecture) 9.2 GPU 运算平台和 cuDNN 7.2.1 深度学习加速库。实验通过 Scikit-learn 开源库,实现浅层机器学习算法。

模型训练过程中,本实验把数据集以 4 : 1 的比例随机生成训练集和验证集,使用随机梯度下降 (stochastic gradient descent, SGD) 算法更新网络参数值。将初始学习率设为 0.01,动量为 0.9,权值衰减系数为 0.001。每经过 20 000 次迭代便把学习率下降为 1/10,迭代 80 000 次结束训练。

4 实验结果与分析

高分辨率遥感影像的信息提取本质上是对遥感影像中信息的有效挖掘和类别判断。本实验选取以海南省海口市琼山区为主的 10 000 × 10 000 高分辨率遥感影像。实验改进了以 ResNet 为特征提取网络的 DeepLab v3 深度学习架构,并以传统的分类算法最大似然法 (maximum likelihood, ML)、浅层的机器学习方法随机森林 (random forest, RF) 和深度学习方法 SegNet 作为对比实验。由于图像非常大,所有的对比实验均采用相同的常规切割、处理和拼

接流程。采用随机森林方法处理,树的最大数量设为 100,最大深度设为 25。实验结果如图 11 所示。

值得注意的是,海南岛上的城镇建成区大多分布于农田、绿地、水域之间,实验选取的大幅遥感影像中城镇建成区也是如此。影像包含了体育场、穿梭于城镇建成区和非城镇建成区的公路、零星村落等典型地物,符合海南省绝大多数地区的实际情况。每组实验的分类准确率和评价指标如表 4 所示。可以看出最大似然法的城镇建成区信息提取精度最低,其虚警率高达 15.56%。受制于对数据的学习能力,最大似然法对于水域的判断能力也非常差。浅层学习算法随机森林的信息提取精度有所提高,但由于计算参数有限,并且大尺度高分辨率遥感图像数据量大、地物特征复杂多样,提取结果和性能指标都没有深度学习算法好。相对于以 VGG (visual geometry group) 为特征提取网络的 SegNet 架构,DeepLab v3 可以实现更好的结果。实验使用特征强化模块和通道域的注意力模块,增强了 DeepLab v3 模型对遥感图像信息的提取能力,提高了 1.95% 的正确率。滑动窗口预测能够减少网络模型产生的拼接痕迹,提高模型泛化能力,验证了本文针对大尺度图像在模型预测阶段的改进效果。全连接条件随机

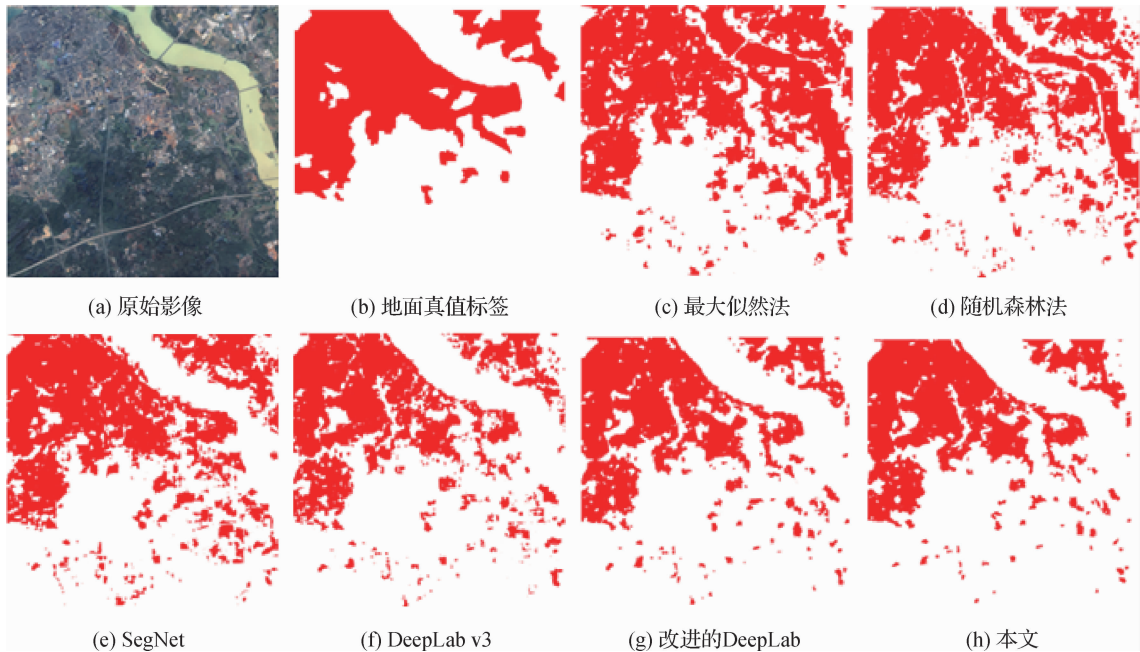


图 11 城镇建成区的信息提取结果

Fig. 11 Results of urban built-up area information extraction ((a) original image; (b) label of ground truth; (c) maximum likelihood; (d) random forest; (e) SegNet; (f) DeepLab v3; (g) proposed DeepLab; (h) ours)

表4 城镇建成区提取结果评价
Table 4 Evaluation of urban built-up area
extraction results

算法	Kappa	正确率/%	漏检率/%	虚警率/%
ML	0.592 8	80.2	4.24	15.56
RF	0.663 8	84.35	5.48	10.16
SegNet	0.741 3	88.21	7.96	3.84
DeepLab v3	0.769 1	89.94	7.4	2.66
改进的 DeepLab	0.816 2	91.89	5.74	2.37
本文	0.837 5	93.06	4.85	2.09

注:加粗字体表示各列最优结果。

场虽然增加了算法的时间复杂度,但它对预测的概率分布图进行处理,结合一元势能和二元势能,可以增强模型能力。本文方法的正确率高于93%,Kappa系数大于0.837。上述的结果和分析,充分验证了本文提出的算法处理流程在大幅高空间分辨率遥感影像上的有效性。

5 结 论

针对大尺度高分辨率遥感影像中地物信息提取困难的问题,本文提出了基于深度学习的城镇建成区信息提取方法。该方法充分考虑了像素之间的空间几何关系,相比传统算法,在实际应用中具有更高的准确率。

由于遥感图像中原始样本数据较少,本文对数据进行增强,并使用有重叠的切割策略进行样本扩充。本实验使用特征强化模块和通道域的注意力模块,对DeepLab网络模型进行大幅改进,更好地提取出遥感影像中特定地物的特征信息。后处理在深度学习技术应用中也非常重要。本文采用滑动窗口预测和全连接条件随机场处理方法,更准确地实现了城镇建成区信息提取,提升了网络模型在工程应用中的能力。

本文是对海南省部分区域的遥感影像进行信息提取,后续还可以扩大范围,比如高分辨率遥感影像中提取出整个地级市甚至更大范围内的城镇建成区。实验结果充分证明了本文算法的有效性,对我国城镇建成区的规划、监测和管理具有重要的辅助作用,能够有效地分析一个地区的城镇化发展水平。

虽然使用深度学习技术进行遥感影像信息提取

取得了好的实验结果,但受制于数据规模和计算设备的处理能力,本文仍有不足之处。对于未来的研究,本文使用的方法可以进行以下改进优化:1)丰富训练样本。在原始高空间分辨率遥感影像的基础上,增加归一化植被指数(normalized difference vegetation index, NDVI)等作为额外的光谱信息。使用生成对抗网络进行样本数据的增强扩充,进而提高算法的鲁棒性。2)多种遥感影像特征提取方法相结合。将高层深度特征与中层稀疏特征进行联合建模,更高效地判断和提取出遥感影像中的各类信息。

参考文献 (References)

- Chen L C, Papandreou G, Kokkinos I, Murphy K and Yuille A L. 2014. Semantic image segmentation with deep convolutional nets and fully connected CRFs [EB/OL]. [2019-09-20]. <https://arxiv.org/pdf/1412.7062.pdf>
- Chen L C, Papandreou G, Schroff F and Adam H. 2017. Rethinking Atrous convolution for semantic image segmentation [EB/OL]. [2019-09-20]. <https://arxiv.org/pdf/1706.05587.pdf>
- Chen L S, Zhao J S, Dong Z W and Zhu Q F. 2018. Urban construction land information extraction based on deep learning by multi-spectral remote sensing imagery of Yunnan central urban agglomeration area. Software Guide, 17(11): 177-180, 186 (陈磊士, 赵俊三, 董智文, 朱祺夫. 基于深度学习的滇中城市多光谱影像建设用地信息提取. 软件导刊, 17(11): 177-180, 186) [DOI: 10.11907/rjdk.181447]
- Chen Z J, Chen J, Shi P J and Tamura M. 2003. An IHS-based change detection approach for assessment of urban expansion impact on arable land loss in China. International Journal of Remote Sensing, 24(6): 1353-1360 [DOI: 10.1080/0143116021000047910]
- Deng L Y, Shen Z F and Ke Y M. 2018. Built-up area extraction and urban expansion analysis based on remote sensing images. Journal of Geo-information Science, 20(7): 996-1003 (邓刘洋, 沈占锋, 柯映明. 2018. 城市建成区遥感影像边界提取与扩张分析. 地球信息科学学报, 20(7): 996-1003) [DOI: 10.12082/dqxxkx.2018.170571]
- Gao C X and Sang N. 2014. Deep learning for object detection in remote sensing image. Bulletin of Surveying and Mapping, (S1): 108-111 (高常鑫, 桑农. 2014. 基于深度学习的高分辨率遥感影像目标检测. 测绘通报, (S1): 108-111) [DOI: 10.13474/j.cnki.11-2246.2014.0625]
- Guo L, Ma L and Ye S C. 2012. Remote sensing image extraction of residential areas based on radon and tree-structured wavelet transformation. Geomatics and Spatial Information Technology, 35(1): 150-153 (过林, 马莉, 叶诗淳. 2012. 基于Radon和树型小波变换的遥感影像居民地提取. 测绘与空间地理信息, 35(1): 150-

- 153) [DOI: 1672-5867(2012)01-0150-04]
- He K M, Zhang X Y, Ren S Q and Jian S. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hu F, Xia G S, Hu J W and Zhang L P. 2015. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sensing*, 7 (11): 14680-14707 [DOI: 10.3390/rs71114680]
- Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 7132-7141 [DOI: 10.1109/CVPR.2018.00745]
- Hu X X, Li Y S, Li H C and Xu Y N. 2010. The comparative analysis about spatial information access technologies based on light UAV low-altitude digital aerophotogrammetry and high-resolution satellite remote sensing. *Engineering of Surveying and Mapping*, 19 (4): 68-70, 74 (胡晓曦, 李永树, 李何超, 许懿娜. 2010. 无人机低空数码航测与高分辨率卫星遥感测图精度试验分析. *测绘工程*, 19(4): 68-70, 74) [DOI: 10.19349/j.cnki.issn1006-7949.2010.04.018]
- Klonowski J and Koch K R. 1997. Two level image interpretation based on Markov random fields. *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps*, Germany: Gustav Stresemann Institut in Bonn Bad Godesberg, 37-55
- Krizhevsky A, Sutskever I and Hinton G E. 2012. ImageNet classification with deep convolutional neural networks//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: ACM: 1097-1105
- LeCun Y, Bottou L, Bengio Y and Haffner P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11): 2278-2324 [DOI: 10.1109/5.726791]
- Li D R. 2003. Towards the development of remote sensing and GIS in the 21st century. *Geomatics and Information Science of Wuhan University*, 28(2): 127-131 (李德仁. 2003. 论21世纪遥感与GIS的发展. *武汉大学学报·信息科学版*, 28(2): 127-131) [DOI: 10.13203/j.whugis2003.02.001]
- Li S H, Wang J L, Bi Y, Chen Y, Zhu M Y, Yang S and Zhu J. 2005. A review of methods for classification of remote sensing images. *Remote Sensing for Land and Resources*, 17(2): 1-6 (李石华, 王金亮, 毕艳, 陈姚, 朱妙园, 杨帅, 朱佳. 2005. 遥感图像分类方法研究综述. *国土资源遥感*, 17(2): 1-6) [DOI: 10.6046/gtzyyg.2005.02.01]
- Lin C G and Nevatia R. 1998. Building detection and description from a single intensity image. *Computer Vision and Image Understanding*, 72(2): 101-121 [DOI: 10.1006/cviu.1998.0724]
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 3431-3440 [DOI: 10.1109/cvpr.2015.7298965]
- Mnih V and Hinton G E. 2010. Learning to detect roads in high-resolution aerial images//Proceedings of the 11th European Conference on Computer Vision. Heraklion, Greece: Springer: 210-223 [DOI: 10.1007/978-3-642-15567-3_16]
- Pesaresi M, Gerhardinger A and Kayitakire F Ç. 2008. A robust built-up area presence index by anisotropic rotation-invariant textural measure. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1 (3): 180-192 [DOI: 10.1109/jstars.2008.2002869]
- Pu R L, Landry S and Yu Q. 2011. Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery. *International Journal of Remote Sensing*, 32 (12): 3285-3308 [DOI: 10.1080/01431161003745657]
- Rodriguez-Galiano V F, Ghimire B, Rogan J, Chica-Olmo M and Rigol-Sanchez J P. 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67: 93-104 [DOI: 10.1016/j.isprsjprs.2011.11.002]
- Rougier S, Puissant A, Stumpf A and Lachiche N. 2016. Comparison of sampling strategies for object-based classification of urban vegetation from very high resolution satellite images. *International Journal of Applied Earth Observation and Geoinformation*, 51: 60-73 [DOI: 10.1016/j.jag.2016.04.005]
- Tao C, Zou Z R and Ding X L. 2014. Residential area detection from high-resolution remote sensing imagery using corner distribution. *Acta Geodaetica et Cartographica Sinica*, 43 (2): 164-169, 192 (陶超, 邹峥嵘, 丁晓利. 2014. 利用角点进行高分辨率遥感影像居民地检测方法. *测绘学报*, 43(2): 164-169, 192) [DOI: 10.13485/j.cnki.11-2089.2014.0024]
- Unsalan C and Boyer K L. 2004. Classifying land development in high-resolution panchromatic satellite images using straight-line statistics. *IEEE Transactions on Geoscience and Remote Sensing*, 42 (4): 907-919 [DOI: 10.1109/tgrs.2003.818835]
- Wang J Q, Li J S, Zhou H C and Zhang X. 2019. Typical element extraction method of remote sensing image based on Deeplabv3 + and CRF. *Computer Engineering*, 45(10): 260-265, 271 (王俊强, 李建胜, 周华春, 张旭. 2019. 基于Deeplabv3+与CRF的遥感影像典型要素提取方法. *计算机工程*, 45(10): 260-265, 271) [DOI: 10.19678/j.issn.1000-3428.0053359]
- Welch R. 1980. Monitoring urban population and energy utilization patterns from satellite data. *Remote Sensing of Environment*, 9 (1): 1-9 [DOI: 10.1016/0034-4257(80)90043-7]
- Xia M, Cao G, Wang G Y and Shang Y F. 2017. Remote sensing image classification based on deep learning and conditional random fields. *Journal of Image and Graphics*, 22(9): 1289-1301 (夏梦, 曹国, 汪光亚, 尚岩峰. 2017. 结合深度学习与条件随机场的遥感图像分类. *中国图象图形学报*, 22(9): 1289-1301) [DOI: 10.11834/jig.170122]

- Xu H Q. 2005. Fast information extraction of urban built-up land based on the analysis of spectral signature and normalized difference index. *Geographical Research*, 24(2): 311-320 (徐涵秋. 2005. 基于谱间特征和归一化指数分析的城市建筑用地信息提取. *地理研究*, 24(2): 311-320) [DOI: 10.11821/yj2005020018]
- Yang C J and Zhuo C H. 2000. Extracting residential areas on the TM imagery. *Journal of Remote Sensing*, 4(2): 146-150 (杨存建, 周成虎. 2000. TM 影像的居民地信息提取方法研究. *遥感学报*, 4(2): 146-150) [DOI: 10.3321/j.issn:1007-4619.2000.02.012]
- Yang H W. 2018. *High Resolution Remote Sensing Imagery Classification Based on Deep Learning*. Beijing: University of Chinese Academy of Sciences (Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences) (杨惠雯. 2018. 基于深度特征学习的高分辨率遥感影像分类. 北京: 中国科学院大学(中国科学院遥感与数字地球研究所))
- Zhang X Y and Du S H. 2016. Learning selfhood scales for urban land cover mapping with very-high-resolution satellite images. *Remote Sensing of Environment*, 178: 172-190 [DOI: 10.1016/j.rse.2016.03.015]
- Zhao P, Feng X Z and Lin G F. 2003. The decision tree algorithm of automatically extracting residential information from SPOT images. *Journal of Remote Sensing*, 7(4): 309-315 (赵萍, 冯学智, 林广发. 2003. SPOT 卫星影像居民地信息自动提取的决策树方法研究. *遥感学报*, 7(4): 309-315) [DOI: 10.11834/jrs.20030413]

作者简介



侯博文, 1995 年生, 男, 硕士研究生, 主要研究方向为遥感信息提取与应用、计算机视觉。

E-mail: hbwer@foxmail.com



闫冬梅, 通信作者, 女, 研究员, 主要研究方向为遥感图像处理、光学遥感影像信息提取与应用、计算机视觉。

E-mail: yandm@radi.ac.cn



郝伟, 通信作者, 男, 研究员, 主要研究方向为图像处理、高速成像信号处理。

E-mail: haowei@opt.ac.cn

黄青青, 女, 助理研究员, 主要研究方向为遥感信息变化检测。E-mail: huangqq@radi.ac.cn

苏秀琴, 女, 研究员, 博士生导师, 主要研究方向为光电经纬仪伺服控制系统、监控、红外及可见光跟踪测量和视频记录。

E-mail: suxiuqin@opt.ac.cn

李青雯, 女, 硕士研究生, 主要研究方向为遥感图像处理。

E-mail: lqwncut@163.com