



JOURNAL OF IMAGE AND GRAPHICS

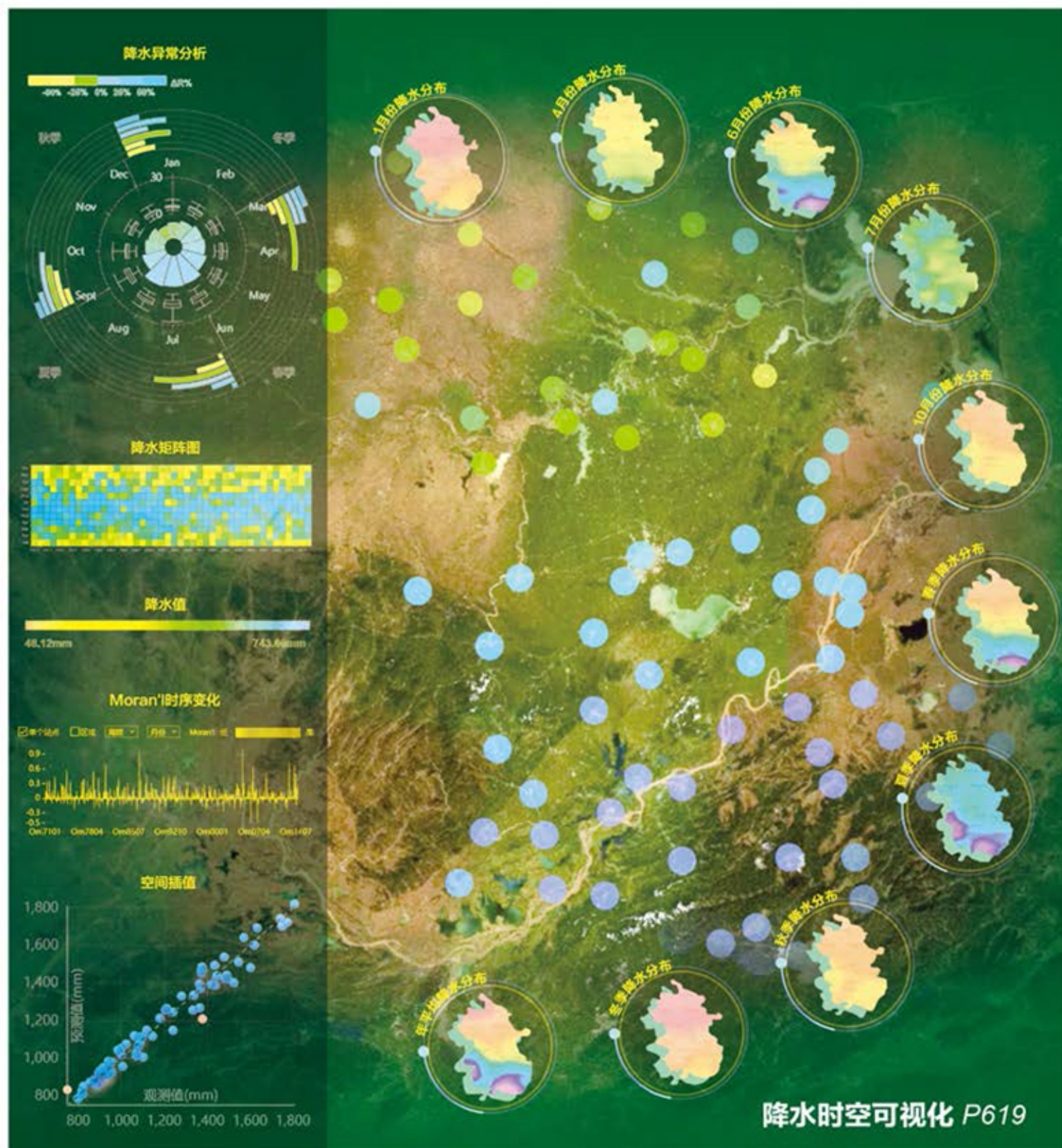
主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象学报

图形学

2021
03
VOL.26

ISSN1006-8961
CN11-3758/TB



中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）



第26卷第3期（总第299期）

2021年3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》编辑部
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

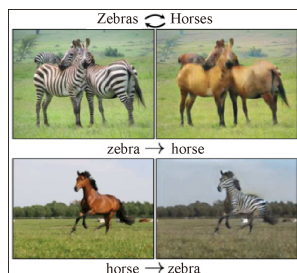
Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial Office 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

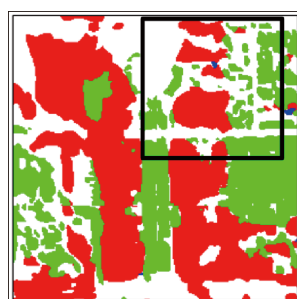
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



深度学习中的图像数据增广方法研究(第0487页)



多通道递归残差网络的图像超分辨重建(第0605页)



面向GF-2遥感影像的U-Net城市绿地分类(第0700页)

学者观点

深度学习与生物医学图像分析2020年综述

陈弘扬, 高敬阳, 赵地, 汪红志, 宋红, 苏庆华 0475

综述

深度学习中的图像数据增广方法研究综述

马崇鼻, 唐娉, 赵理君, 张正 0487

图像处理和编码

自适应B样条小波函数模糊距离测量

梁锐, 魏阳杰 0503

图像分析和识别

增强二阶网络调制的目标跟踪

王献海, 宋慧慧, 张开华, 刘青山 0516

背景与方向感知的相关滤波跟踪

姜文涛, 涂湖, 刘万军 0527

SSD与时空特征融合的视频目标检测

尉婉青, 禹晶, 柏鳗晏, 肖创柏 0542

自然场景下变形车牌检测模型DLPD-Net

余焯, 付源梓, 陈维笑, 刘海涛 0556

两阶段特征提取策略的图像去雾

袁非牛, 李志强, 史劲亭, 夏雪, 李雅 0568

图像理解和计算机视觉

由粗到精的多尺度散焦模糊检测

衡红军, 叶何斌, 周末, 黄睿 0581

Re-GAN: 残差生成式对抗网络算法

史彩娟, 涂冬景, 刘靖祯 0594

多通道递归残差网络的图像超分辨重建

程德强, 郭昕, 陈亮亮, 寇旗旗, 赵凯, 高蕊 0605

计算机图形学

降水时空关联特征可视分析

周志光, 谢琬滢, 郑微桦, 陈圆圆 0619

医学图像处理

结合残差路径及密集连接的乳腺超声肿瘤分割

陈杨怀, 陈胜, 姚莉萍 0633

遥感图像处理

增强小目标特征的航空遥感目标检测

赵文清, 孔子旭, 周震东, 赵振兵 0644

密集子区域切割的任意方向舰船快速检测

陈华杰, 吴栋, 谷雨 0654

CGAN样本生成的遥感图像飞机识别

王耀领, 王宏琦, 许滔 0663

高分辨率遥感影像的边缘损失增强地物分割

陈琴, 朱磊, 吕燧栋, 吴谨 0674

含多级通道注意力机制的CGAN遥感图像建筑物分割

余帅, 汪西莉 0686

面向GF-2遥感影像的U-Net城市绿地分类

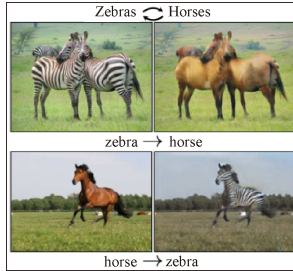
徐知宇, 周艺, 王世新, 王丽涛, 王振庆 0700

条件生成对抗遥感图像时空融合

李昌洁, 宋慧慧, 张开华, 张晓露, 刘青山 0714

CONTENTS

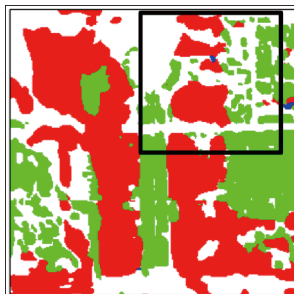
JOURNAL OF IMAGE AND GRAPHICS



Review of data augmentation for image in deep learning(P0487)



Image super-resolution reconstruction from multi-channel recursive residual network (P0605)



U-Net for urban green space classification in Gaofen-2 remote sensing images(P0700)

Scholar View

Review of the research progress in deep learning and biomedical image analysis till 2020
Chen Hongyang, Gao Jingyang, Zhao Di, Wang Hongzhi, Song Hong, Su Qinghua 0475

Review

Review of data augmentation for image in deep learning
Ma Dongao, Tang Ping, Zhao Lijun, Zhang Zheng 0487

Image Processing and Coding

Adaptive distance measurement method with blur of B-spline wavelet function
Liang Rui, Wei Yangjie 0503

Image Analysis and Recognition

Object tracking using enhanced second-order network modulation
Wang Xianhai, Song Huihui, Zhang Kaihua, Liu Qingshan 0516

Background and direction-aware correlation filter tracking
Jiang Wentao, Tu Chao, Liu Wanjun 0527

Video object detection using fusion of SSD and spatiotemporal features
Yu Wanqing, Yu Jing, Bai Manyan, Xiao Chuangbai 0542

DLPD-Net: distorted license plate detection model in natural scenarios
Yu Ye, Fu Yuanzi, Chen Weixiao, Liu Haitao 0556

Image defogging algorithm using a two-phase feature extraction strategy
Yuan Feiniu, Li Zhiqiang, Shi Jinting, Xia Xue, Li Ya 0568

Image Understanding and Computer Vision

Coarse-to-fine multiscale defocus blur detection
Heng Hongjun, Ye Hebin, Zhou Mo, Huang Rui 0581

Re-GAN: residual generative adversarial network algorithm
Shi Caijuan, Tu Dongjing, Liu Jingyi 0594

Image super-resolution reconstruction from multi-channel recursive residual network
Cheng Deqiang, Guo Xin, Chen Liangliang, Kou Qiqi, Zhao Kai, Gao Rui 0605

Computer Graphics

Visual analytics of spatial and temporal correlation features of rainfall
Zhou Zhiguang, Xie Wanying, Zheng Weihua, Chen Yuanyuan 0619

Medical Image Processing

Tumor segmentation in breast ultrasound combined with Res paths and a dense connection
Chen Yanghuai, Chen Sheng, Yao Liping 0633

Remote Sensing Image Processing

Target detection algorithm of aerial remote sensing based on feature enhancement technology
Zhao Wenqing, Kong Zixu, Zhou Zhendong, Zhao Zhenbing 0644

Fast detection algorithm for ship in arbitrary direction with dense subregion cutting
Chen Huajie, Wu Dong, Gu Yu 0654

Aircraft recognition of remote sensing image based on sample generated by CGAN
Wang Yaoling, Wang Hongqi, Xu Tao 0663

Segmentation of high-resolution remote sensing image by collaborating with edge loss enhancement
Chen Qin, Zhu Lei, Lyu Suidong, Wu Jin 0674

Remote sensing building segmentation by CGAN with multilevel channel attention mechanism
Yu Shuai, Wang Xili 0686

U-Net for urban green space classification in Gaofen-2 remote sensing images
Xu Zhiyu, Zhou Yi, Wang Shixin, Wang Litao, Wang Zhenqing 0700

Spatiotemporal fusion of satellite images via conditional generative adversarial learning
Li Changjie, Song Huihui, Zhang Kaihua, Zhang Xiaolu, Liu Qingshan 0714

中图法分类号: TP79 文献标识码: A 文章编号: 1006-8961(2021)03-0700-14

论文引用格式: Xu Z Y, Zhou Y, Wang S X, Wang L T and Wang Z Q. 2021. U-Net for urban green space classification in Gaofen-2 remote sensing images. *Journal of Image and Graphics*, 26(03):0700-0713 (徐知宇, 周艺, 王世新, 王丽涛, 王振庆. 2021. 面向 GF-2 遥感影像的 U-Net 城市绿地分类. *中国图象图形学报*, 26(03):0700-0713 [DOI:10.11834/jig.200052])

面向 GF-2 遥感影像的 U-Net 城市绿地分类

徐知宇^{1,2}, 周艺¹, 王世新¹, 王丽涛¹, 王振庆^{1,2}

1. 中国科学院空天信息创新研究院, 北京 100094; 2. 中国科学院大学, 北京 100049

摘要: **目的** 高分 2 号卫星 (GF-2) 是首颗民用高空间分辨率光学卫星, 具有亚米级高空间分辨率与宽覆盖结合的显著特点, 为城市绿地信息提取等多领域提供了重要的数据支撑。本文利用 GF-2 卫星多光谱遥感影像, 将一种改进的 U-Net 卷积神经网络首次应用于城市绿地分类, 提出一种面向高分遥感影像的城市绿地自动分类提取技术。**方法** 先针对小样本训练集容易产生的过拟合问题对 U-Net 网络进行改进, 添加批标准化 (batch normalization, BN) 和 dropout 层获得 U-Net + 模型; 再采用随机裁剪和随机数据增强的方式扩充数据集, 使得在充分利用影像信息的同时保证样本随机性, 增强模型稳定性。**结果** 将 U-Net + 模型与最大似然法 (maximum likelihood estimation, MLE)、神经网络 (neural networks, NNs) 和支持向量机 (support vector machine, SVM) 3 种传统分类方法以及 U-Net、SegNet 和 DeepLabv3 + 这 3 种深度学习语义分割模型进行分类结果精度对比。改进后的 U-Net + 模型能有效防止过拟合, 模型总体分类精度比改进前提高了 1.06%。基于改进的 U-Net + 模型的城市绿地总体分类精度为 92.73%, 平均 F_1 分数为 91.85%。各分类方法按照总体分类精度从大到小依次为 U-Net + (92.73%)、U-Net (91.67%)、SegNet (88.98%)、DeepLabv3 + (87.41%)、SVM (81.32%)、NNs (79.92%) 和 MLE (77.21%)。深度学习城市绿地分类方法能充分挖掘数据的光谱、纹理及潜在特征信息, 有效降低分类过程中产生的“椒盐噪声”, 具有较好的样本容错能力, 比传统遥感分类方法更适用于城市绿地信息提取。**结论** 改进后的 U-Net + 卷积神经网络模型能够有效提升高分遥感影像城市绿地自动分类提取精度, 为城市绿地分类提供了一种新的智能解译方法。

关键词: 城市绿地; 卷积神经网络; U-Net; 高分遥感; 语义分割

U-Net for urban green space classification in Gaofen-2 remote sensing images

Xu Zhiyu^{1,2}, Zhou Yi¹, Wang Shixin¹, Wang Litao¹, Wang Zhenqing^{1,2}

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

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

Abstract: Objective High-precision monitoring of the spatial distribution of urban green space has important social, economic, and ecological benefits for optimizing the spatial structure of such space, maintaining urban ecological balance, and developing green city construction. As the first civilian optical satellite with high spatial resolution, Gaofen2 (GF-2) exhibits the remarkable characteristics of sub-meter high spatial resolution and wide coverage. GF2 provides important data sup-

收稿日期: 2020-02-25; 修回日期: 2020-04-17; 预印本日期: 2020-04-24

基金项目: 国家重点研发计划项目 (2017YFB0504000)

Supported by: National Key Research and Development Program of China (2017YFB0504000)

port to multiple fields, such as urban environmental monitoring and urban green space information extraction. However, traditional classification methods still encounter many problems. For example, training a method to be an effective classifier for massive data is difficult, and the accuracy of classification results is generally low. The use of massive high-resolution remote sensing images to achieve large-scale rapid and accurate urban green space distribution extraction is an urgent task for urban planning managers. With the rapid development of deep learning technology, full convolutional networks (FCN) provide novel creative possibilities for semantic segmentation and realize pixel-level classification of images in the field of deep learning for the first time. Inspired by the U-Net network structure, we applied an improved U-Net to urban green space classification for the first time and proposed an automatic classification technique for urban green space by using high-resolution remote sensing images.

Method First, we improved the U-Net model to obtain the U-Net + model. The main structure of U-Net + is composed of an encoder and a decoder that can achieve end-to-end training. The encoding channel realizes the multi-scale feature recognition of an image through four-time maximum pooling, and the decoding channel restores the position and detailed information of an image through upsampling. The network uses skip connection to realize the fusion of feature information with the same scale at different levels, overcoming accuracy loss caused by upsampling. In addition, we improved the model by adding batch normalization (BN) after each layer of network convolution operation, effectively regulating the input of the network layer and improving model training speed and network generalization capability. To solve the overfitting problem, which is easily produced by the limited sample training set, we added the dropout layer with a 50% probability of dropping neurons after the convolution operation of the fourth and fifth layers of the network. Second, deep learning requires a large amount of label data related to the classification objectives for training. However, existing open-source datasets cannot meet the requirements of the urban green space classification task. Manually establishing an urban green space tag dataset is necessary. We selected three typical urban green space sample areas in Beijing (urban parks, residential areas, and golf courses) as study areas. By combining GF-2 images and Google Earth remote sensing images in summer and winter, we drew all types of urban green space in the study areas through visual interpretation by using ArcGIS. The visual interpretation results are corrected with actual field investigation. Third, random cropping and data augmentation techniques are adopted to expand the dataset, ensuring the randomness of the samples and enhancing the stability of the model while fully utilizing image information. We adopt the Adam optimizer with an initial learning rate of 0.000 1.

Result 1) The overall classification accuracy of the U-Net + model is improved by 1.06% compared with that of the original U-Net. After 40 training epochs, the accuracy of the U-Net + model reaches a high level, and the loss function realizes rapid convergence. The U-Net + model effectively prevents overfitting and improves generalization capability. 2) To verify the effectiveness of our method, the classification accuracy of the U-Net + results is compared with those of three traditional classification methods, namely, maximum likelihood estimation (MLE), neural networks (NNs), and support vector machine (SVM), and three semantic segmentation models, i. e., U-Net, SegNet, and DeepLabv3+. Among the seven classification methods, the U-Net + model achieves the highest overall classification accuracy for urban green space. The seven classification methods are arranged in order of classification accuracy from large to small: U-Net + (92.73%) > U-Net (91.67%) > SegNet (88.98%) > DeepLabv3+ (87.41%) > SVM (81.32%) > NNs (79.92%) > MLE (77.21%). 3) In the three types of urban green space, evergreen trees have the highest classification accuracy ($F_1 = 93.65%$), followed by grassland ($F_1 = 92.55%$) and deciduous trees ($F_1 = 86.55%$). 4) Deep learning exhibits strong fault-tolerant capability for training samples. By training and learning a large number of label data, it can effectively reduce the impact of errors and improve recognition capability, making it more suitable for urban green space information extraction than traditional remote sensing classification methods.

Conclusion Deep learning urban green space classification methods can fully mine the spectral, textural, and potential feature information of data. Meanwhile, the U-Net + model proposed in this study can also effectively reduce the salt-and-pepper noise the classification process and realize high-precision pixel-level classification of urban green space. The improved U-Net + can effectively improve the accuracy of automatic classification of urban green space in high-resolution remote sensing images and provide a new intelligent interpretation method for urban green space classification in the future.

Key words: urban green space; convolutional neural network (CNN); U-Net; high-resolution remote sensing; semantic segmentation

0 引言

城市绿地作为改善城市自然生态和生活环境的“净化器”,是现代城市生活和生产不可或缺的一部分。城市绿地现状信息的实时、高效、精细化监测,对城市生态环境建设,提高城市生态效益具有重要意义(Qian等,2015)。然而城市绿地遥感影像分类是一个复杂的数据处理过程,尤其是高分辨率遥感影像的城市绿地分类存在很多困难:1)城市绿地的植被类型多样,多光谱特征差异显著;2)绿地斑块具有高度异质性和分散性,形状破碎且不规则;3)高分辨率遥感影像提供地物的高清细节信息的同时,也增大了类内差异(刘大伟等,2016)。现有影像分类方法对数据本身的要求严格,难以从海量数据中训练出有效分类器以充分挖掘数据之间的关联,泛化能力较弱,过程耗时耗力且难以得到更好的分类效果。

深度学习(deep learning)逐渐成为神经网络研究的热点和前沿。深度学习是Hinton等人(2006)基于人脑学习的思想,提出的一种深度神经网络的机器学习方法。相对BP(back propagation)神经网络、支持向量机等“浅层”神经网络方法而言,深度神经网络的优势在于能够更充分挖掘海量数据中的光谱、纹理及潜在特征信息,智能地进行特征选择与特征工程构建,在影像分类中取得较好的分类结果。全卷积神经网络(fully convolutional network, FCN)(Long等,2015)的提出,打开了深度学习语义分割的大门,首次实现了深度学习图像的像素级分类。U-Net模型(Ronneberger等,2015)在FCN基础上进行改进,结合了反卷积网络与跳跃结构,在遥感影像分类中取得了一定的研究成果(Zhang等,2018)。深度学习图像分类技术已经广泛用于建筑物提取(Liu等,2019)、土地分类(Helber等,2018)、作物分类(Kussul等,2017)和医学图像分类(Hirasawa等,2018)等领域,均取得了不错效果。

利用海量高分辨率遥感影像,开展大范围快速准确的城市绿地分布提取,是实现精细化城市绿地

空间规划管理的当务之急。针对传统城市绿地影像分类存在的问题,本文将U-Net卷积神经网络首次应用于城市绿地分类,提出一种面向高分遥感影像的城市绿地自动分类提取技术。通过将改进的U-Net+模型与传统的分类模型和当前流行的深度学习语义分割模型进行对比,证明本文方法在城市绿地遥感分类中有着显著的优势,为今后城市绿地分类提供了一种新的智能解译方法。

1 研究区域与数据

1.1 研究区概况

北京作为中国的首都,是全国政治、文化和科技创新中心,是著名的现代化国际城市。北京市总面积16 410.54 km²,是典型的北温带半湿润大陆性季风气候,夏季高温多雨,冬季寒冷干燥,城市绿地主要集中于五环以内,主要种植暖温带落叶阔叶林和温性针叶林,形成了以青山相拥,三环环绕,十字绿轴,十条楔形绿地穿插,百余个城市公园和小微绿地为主的,由绿色通道串联成点、线、面、带、环相结合的城市绿地格局系统(北京市城市规划设计研究院,2010)。

本文以北京市城市绿地为研究对象,根据城市绿地类型分布特点选取了3个典型研究区域,分别为城市公园(天坛公园)、住宅区(东城)和高尔夫球场(万柳高尔夫球场),如图1所示。面积分别是2.56 km²、2.56 km²和2.42 km²。城市公园绿地作为城市绿地系统中的主要场所,分布相对集中独立,植被类型多样,具有观赏游憩功能;住宅区绿地主要包括住宅小区庭院绿地、花园绿地和交通运输用地中的附属绿地,分布较破碎零散;高尔夫球场主要以人工草地为主,绿地种植较为规则,具有辨识度高的特点。这3个样本区可以基本覆盖北京市城市绿地典型类型,减少数据量冗余带来的工作量,提高分析效率。

1.2 绿地分类

北京市常用绿地植被可归纳为3种绿地类型:常绿树,落叶树和草地。其中常绿树主要包括雪松、侧柏、圆柏、白皮松、樟子松、华山松、油松、乔松、白

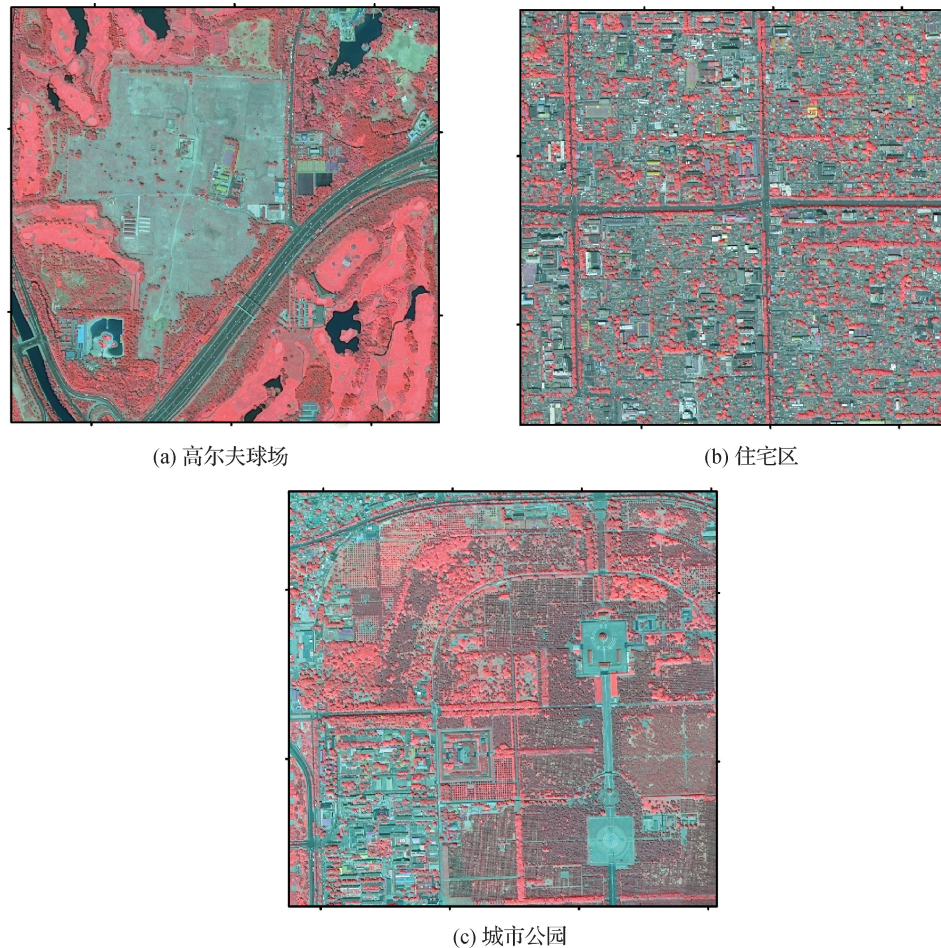


图 1 研究区

Fig. 1 Study areas((a) golf courses; (b) residential areas; (c) urban parks)

杆、冬青卫矛、叉子圆柏和锦熟黄杨等松柏类乔木和冬青卫矛等灌木;落叶树种类繁多,覆盖范围较广,主要包括梧桐、银杏、白蜡树、玉兰、水杉、海棠、合欢、椿树、红枫、白桦、黄金树、梓树、山楂、杜仲、樱桃、杏树、槐树、丁香、柳树、榆树、枣树、竹、蜡梅、流苏、卫矛、连翘、圆锥绣球、木槿、迎春花、枸杞和金叶女贞等;草地包括人工草地和天然草地(吴丽娟,2006)。各植物类型及对应的遥感解译特征如表 1 所示。

1.3 数据源与数据预处理

本文以国产高分 2 号(GF-2)卫星多光谱影像为主要研究数据(数据来源于中国资源卫星应用中心,获取时间为 2019 年 5 月)。高分 2 号 PMS(panchromatic multispectral sensor)传感器具有 1 个 0.8 m 空间分辨率的全色波段和 4 个 3.2 m 空间分辨率的多光谱波段,通过正射校正、影像融合、影像镶嵌与裁剪等预处理,得到研究区空间分辨率为 1 m、位

表 1 北京市常用植物类型及对应遥感解译特征

Table 1 Common plant types and corresponding remote sensing interpretation features in Beijing

类别	种类	解译特征	标准假彩色
常绿树	雪松、侧柏等	色调呈暗红色,纹理排列清晰,多分布于公园	
落叶树	梧桐、银杏等	色调呈红色,纹理粗糙,多分布于公园、居民区、道路两旁	
草地	人工草地、天然草地	色调呈浅红色,纹理均匀,分布于高尔夫球场、公园	

深 8 bit 的 GF-2 影像。为了更好地表征城市绿地的植被特征,本文采用 432 波段组合的标准假彩色影像进行城市绿地分类提取。

深度学习需要大量与分类目标有关的标签数据进行训练,而现有的开源数据集无法满足本文分类任务需求,因此需要人工建立城市绿地标签数据集。结合 GF-2 与 Google earth 夏季和冬季高分辨率遥感影像,通过人工目视解译的方式勾绘出实验区影像中所有的绿地类型轮廓,通过实地调查对 3 种绿地目视解译结果进行校正,结果存为 8 bit 的栅格图像,即为地表真实标签数据,其中,1 代表常绿树,2 代表落叶树,3 代表草地,0 代表其他。将处理好的影像数据和对应标签数据按照面积比例 3 : 1 : 1 划分成训练区、测试区和验证区,保证训练、测试、验证不存在交叉重叠,通过随机裁剪扩充训练集和测试集样本量,弱化背景噪声,最终获得训练集样本图像 1 500 幅,测试集样本图像 500 幅,如图 2 所示。

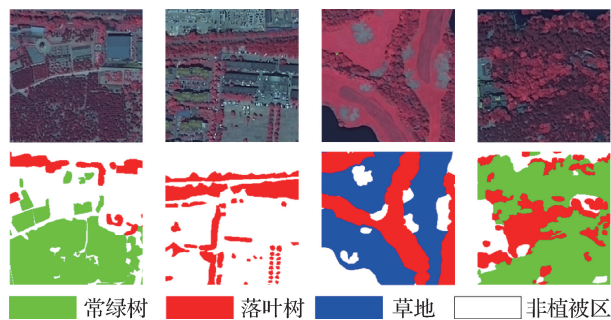


图 2 数据集影像及对应标签

Fig. 2 Remote sensing images and corresponding label images

2 基于 U-Net + 模型的城市绿地分类方法

本文在 U-Net 深度学习模型的基础上加以改进,构建了基于高分 2 号遥感影像的城市绿地自动分类提取方法。首先在制作好的数据集上对改进后的 U-Net + 模型进行训练,得到最佳参数组合。然后预测得到北京市五环区域的城市绿地分类图。为了验证本文方法的有效性,选取最大似然法 (maximum likelihood estimate, MLE)、神经网络 (neural networks, NNs)、支持向量机 (support vector machine, SVM)、SegNet、DeepLabv3+ 和原始 U-Net 模型等 6 种分类方法进行对比分析。

2.1 U-Net + 卷积神经网络

U-Net 卷积神经网络是 Ronneberger 等人 (2015) 提出的一种改进的 FCN 结构,是目前扩展性

较好的全连接神经网络。其结构清晰呈现字母 U 形,由左半边的压缩通道和右半边扩展通道组成。与其他深度神经网络相比,U-Net 模型的最大优势在于结合了编码—解码结构和多尺度特征融合的特点。压缩通道是一个编码器,通过 4 次最大池化实现网络对影像特征的多尺度特征识别。扩展通道是一个解码器,通过上采样还原影像的位置信息和细节信息。通过特征拼接 (skip-connection) 实现不同层次同尺度特征信息的融合,克服因上采样造成的精度损失问题,以优化分类结果。

目前,对 U-Net 模型的改进主要包括:1) 通过增加高性能语义分割模块改进 U-Net 模型网络框架。Diakogiannis 等人 (2020) 将 U-Net 编码解码结构与残差连接、金字塔池化等体系结合,提出一种深度学习架构 ResUNet-a,用于高分辨遥感领域语义分割任务。Li 等人 (2018) 使用 DownBlock 和 UpBlock 模块构建 DeepUNet 模型,用于海陆影像分类。2) 通过改进损失函数提高模型性能。何凯明团队提出 focal loss 损失函数以解决样本不均衡问题 (Lin 等, 2017)。3) 结合多源特征数据提高深度学习模型分类精度。许慧敏等人 (2019) 结合 nDSM 数据可以有效提高 U-Net 模型的地形分类精度。

本文总结深度学习常用且有效的改进方法,对 U-Net 网络进行改进,提出了 U-Net + 模型,解决了 U-Net 模型对小样本训练中存在的过拟合问题,提高了模型分类效果和泛化能力。

激活函数对神经网络学习和理解复杂分类问题具有重要作用,常用的有 sigmoid 函数和 ReLU 函数。当输入数据较大或较小时,sigmoid 函数的导数趋近于零,存在饱和效应,故在反向传播阶段可能会因为梯度消失导致无法收敛。ReLU 函数在提升网络稀疏性的同时减少了参数之间的依赖性,可以有效改善梯度消失,加快梯度下降收敛速度。本文提出的 U-Net + 模型沿用 U-Net 模型的激活函数策略,采用 ReLU 激活函数提高模型训练速度。

批标准化 (batch normalization, BN) 是 Ioffe 和 Szegedy (2015) 提出的一种改善神经网络性能和稳定性的设计模式,将数据分成小批次 (batch size) 进行随机梯度下降,计算机内部自动保存读取的每批数据均值和方差,在每批数据进行前向传递时,对每一层神经网络输入值都进行标准化处理,服从均值为 0、方差为 1 的分布。本文提出的 U-Net + 模型在

U-Net 模型每层网络卷积操作后增加 BN 处理,可以有效规范网络层的输入,克服了因网络层数加深导致难以训练的问题,有助于梯度传播。

随机失活 (dropout) 是神经网络正则化的一种方式 (Srivastava 等, 2014), 在训练过程中通过将隐含层的部分权重或输出随机归零, 降低节点间的相互依赖性, 可以有效解决网络过拟合和梯度消失问题。本文提出的 U-Net + 模型在 U-Net 网络第 4 层和第 5 层卷积操作后添加概率为 0.5 的 dropout 层, 通过丢弃 50% 概率的神经元减少数据过拟合现象。

针对遥感影像人工构建的标签数据集数量有限的问题, 本文提出的 U-Net + 模型采取随机裁剪 (random crop) 和随机数据增强 (random data aug-

mentation) 的方式扩充样本量。随机裁剪可以通过控制裁剪数量保证 3 个样本区 3 种绿地训练样本的比例相近, 能够高效大量扩充训练样本量。与规则网格裁剪和滑动窗口裁剪相比, 随机裁剪能获取更多的图像数量, 且不会造成过度的数据冗余, 保证样本随机性的同时可以有效学习数据信息。随机数据增强也是一种降低过拟合的有效方法, 实验采取在线随机增强方式, 对网络中批次抽取的样本进行旋转、图像水平偏移、垂直偏移、缩放和水平翻转等随机增强处理, 这样的方式保证了每次训练过程的样本不完全相同, 使训练得到的模型具有更好的泛化能力, 防止过拟合。

本文提出的 U-Net + 模型结构如图 3 所示 (其中, D 为 dropout)。

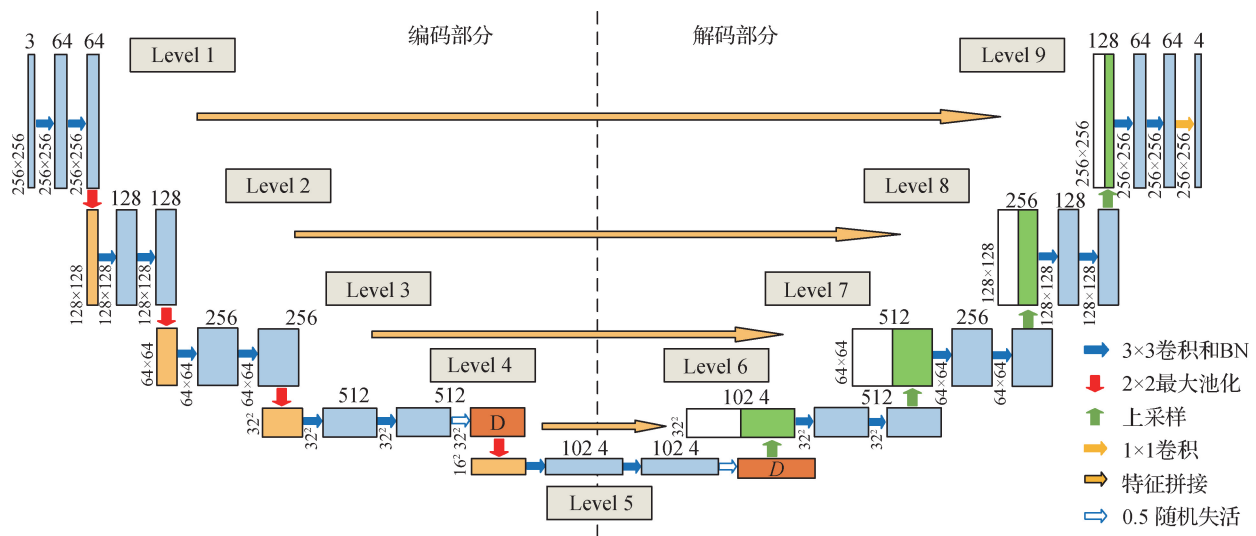


图 3 U-Net + 模型结构

Fig. 3 Structure of U-Net + model

2.2 模型实验

实验硬件 CPU 为英特尔 Core i7-8700@3.20 GHz, GPU 为 NVIDIA GeForce RTX 2080 Ti 11 GB, 采用 TensorFlow 后端的 Keras 深度学习框架实现 U-Net + 模型的搭建与实验。

在模型训练中, 共有 1 500 幅 256×256 像素的训练图像、500 幅测试图像和对应标签输入模型, 初始学习率为 0.001, 最大迭代次数 100 次, 每次训练的主要步骤如下:

1) 在训练集中每批次 (batch size) 随机抽取 2 组训练样本 x 和对应标签 y , 经随机增强后输入网络, 其中城市绿地分类先验知识通过标签 y 输入网络。

2) 训练样本输入网络后进行前向传播 (forward

pass), 在编码部分提取输入影像的特征, 在解码部分恢复其空间位置与分辨率, 在像素分类层对每一像素进行绿地分类获得 4 种类别信息, 得到一组预测值 y_{pred} 。

3) 采用交叉熵函数计算这批数据的损失值, 用于衡量标签 y 与预测值 y_{pred} 的距离。

4) 通过 Adam 优化器, 计算损失值相对于网络参数的梯度, 将参数进行反向传播 (backward pass), 从而降低损失值。

按照上述步骤反复循环, 经过 100 次迭代, 计算机逐层优化模型中的参数达到最优, 损失函数降到 0.02, 此时训练结束。将待分类影像输入网络进行预测, 获得最终分类结果图。

2.3 分类后精度评价

为定量评估影像分类精度,采用混淆矩阵(郑明国等,2006)对分类后影像全部像元进行类别统计,并将总体分类精度(overall accuracy, OA)、kappa系数和 F1 分数作为评价指标。

OA 表示全部像素分类结果与标签数据对应区域的实际类型一致的概率。kappa 系数(K)表示分类与完全随机分类产生错误减少的比例。具体为

$$OA = \sum_{i=1}^n \frac{x_{ii}}{N} \quad (1)$$

$$K = \frac{N \sum_{i=1}^n x_{ii} - \sum_{i=1}^n (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^n (x_{i+} x_{+i})} \quad (2)$$

式中, x_{ii} 为图像中像元正确分类的个数; x_{i+} 、 x_{+i} 为混淆矩阵各行、各列之和, N 为图像中像元总数。

F1 分数(F_1)是表示精确度(precision)与召回率(recall)的综合评价指标,反映模型对正负样本的识别和区分能力。

$$F_1 = \frac{(1 + \beta^2)P \times R}{\beta^2(P + R)} \quad (3)$$

式中, P 表示精确率,又称制图精度,指正确分为某类的像元数与该类真实参考总数的比率; R 表示召回率,又称用户精度,指正确分为某类的像元总数与分类器将整个影像的像元分为该类的像元总数的比率; β 为权重系数,反映精确度和召回率对 F_1 的重要程度,本文设 β 为 1。

U-Net + 模型的技术流程如图 4 所示。

3 结果与分析

3.1 模型训练结果

将 1 500 幅 256×256 像素的图像和对应标签输入原始 U-Net 和改进后的 U-Net + 模型中迭代 100 次,得到两组随迭代次数变化的精确度曲线和损失值曲线,如图 5 和图 6 所示。通过图 5 和图 6 对 U-Net 模型改进前后的训练情况和过拟合现象进行分析,可以看出在训练过程中,U-Net 和 U-Net + 模型训练集和验证集的精确度均随迭代次数的增多先快速上升,迭代 40 次后逐渐升高趋于平稳,二者的整体变化趋势比较一致;U-Net 和 U-Net + 模型训练集的损失值随迭代次数的增加先快速下降,迭代

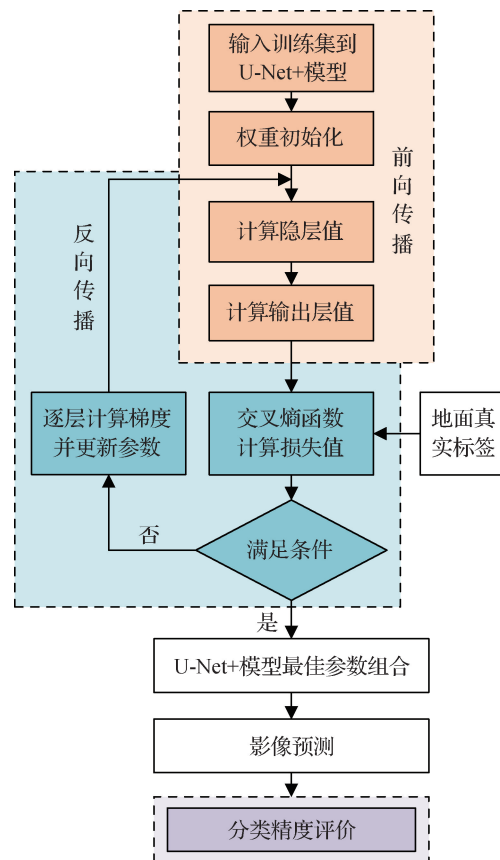


图 4 U-Net + 模型的技术流程

Fig. 4 Technical process of U-Net + model

40 次后逐渐降低趋近于 0;U-Net 模型验证集的损失值先保持稳步下降,迭代 60 次后 loss 出现上升趋势,说明模型存在过拟合现象;U-Net + 模型验证集的损失值在前 40 次迭代时有显著下降,随后趋于平缓,损失值在 0.16 左右波动。训练结果表明,增加了 BN、dropout 和数据随机增强的 U-Net + 模型可以有效防止过拟合现象,提高了模型的泛化能力和鲁棒性。改进后的模型经过 40 次,迭代准确率已经达到较高水平,曲线相对平滑,波动性较小,说明学习率设置较为合理,损失函数实现快速收敛,可判定网络对数据集进行了有效学习。

将 U-Net + 模型预测的测试影像分类图进行混淆矩阵计算,结果如表 2 所示。共有 1 888 818 个像元参与混淆矩阵计算,1 751 518 个像元分类正确,kappa 系数为 0.89,总体分类精度为 92.73%,相比 U-Net 模型提高了 1.06%。在分类结果中,常绿树的分类精度最高($F_1 = 93.65\%$),草地($F_1 = 92.55\%$)和落叶树($F_1 = 86.55\%$)次之,主要在于常绿树与落叶、草地的光谱、纹理等特征差异明显,

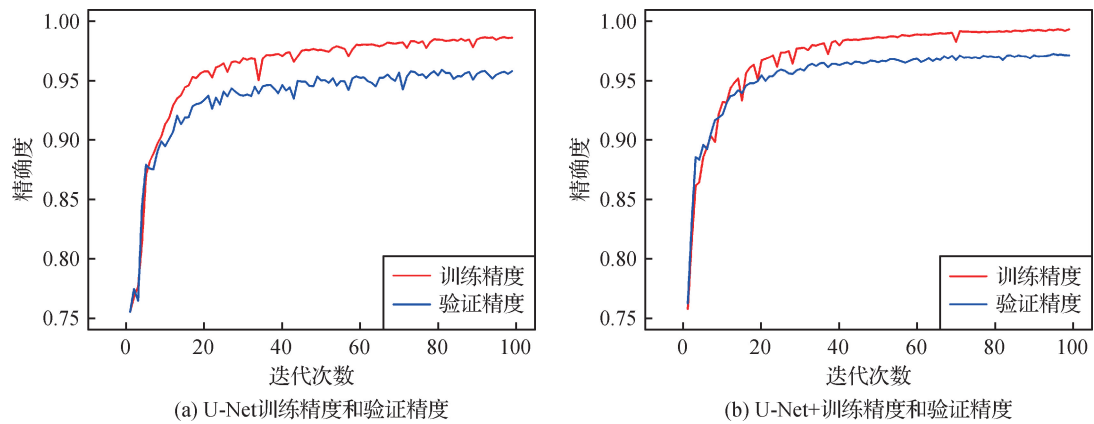


图 5 随迭代次数变化的精确度曲线

Fig. 5 Curve of accuracy with epochs

((a) training and validation accuracy of U-Net; (b) training and validation accuracy of U-Net +)

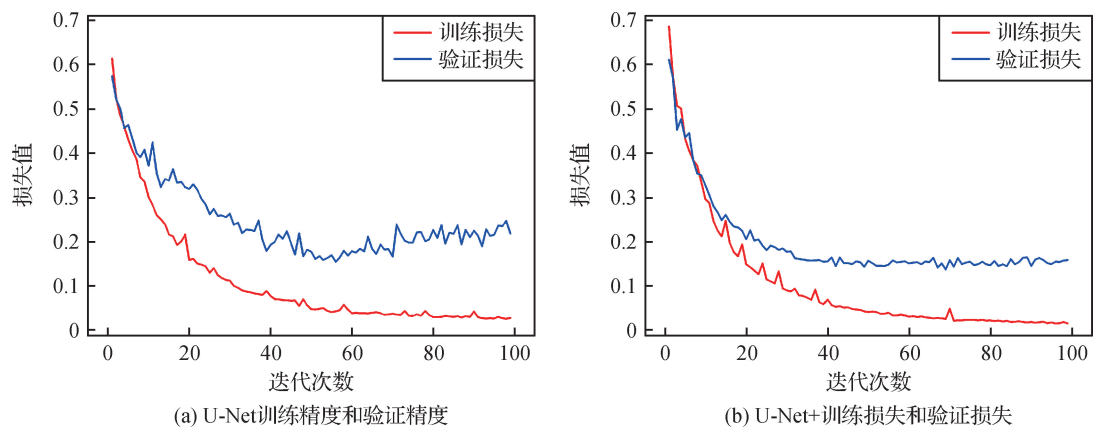


图 6 随迭代次数变化的损失值曲线

Fig. 6 Curve of loss with epochs ((a) training and validation loss of U-Net; (b) training and validation loss of U-Net +)

表 2 U-Net + 模型分类结果混淆矩阵与评价精度

Table 2 Confusion matrix and evaluation accuracies of classification results with U-Net + model

标签		预测类型及数量				评价指标				
类型	数量	落叶树	常绿树	草地	其他	用户精度/%	制图精度/%	F_1 /%	总体分类精度/%	kappa 系数
落叶树	345 565	311 496	4 347	2 011	27 711	83.23	90.14	86.55	92.73	0.89
常绿树	341 859	11 552	316 866	824	12 617	94.64	92.69	93.65		
草地	173 243	8 269	1 029	158 587	5 358	93.58	91.54	92.55		
其他	1 028 151	42 957	12 574	8 051	964 569	95.48	93.82	94.64		

注:加粗字体为各列预测正确的数量。

能较好区分。草地面积在研究区中整体占比较小,主要错分为其他用地和落叶树。总体来看,基于 U-Net + 卷积神经网络模型的遥感分类方法较好地完成了高分影像的城市绿地分类任务,3 种绿地类型边界划分清晰,分类精度较高,通过深层网络学习

除光谱特征以外的更多深层语义特征,能够有效降低分类过程中产生的“椒盐噪声”,在一定程度上减少地物像元混分现象,对分布零散且破碎的落叶树也能较完整地提取出来,证明该方法能够实现高分辨率遥感影像城市绿地分类的自动化分类提取。

3.2 地物分类结果对比

为验证本文方法的有效性,与 MLE(maximum likelihood estimation)(金杰等,2014)、NNs(neural networks)(Pacifi 等,2009)、SVM(support vector machine)(Liu 等,2017)等传统分类方法和 SegNet(Badrinarayanan 等,2017)、DeepLabv3+(Chen 等,2018)、U-Net(Ronneberger 等,2015)等语义分割模型进行对比实验,MLE、NNs 和 SVM 的分类实验均在 ENVI(environment for visualizing images)软件中完成。

图 7 为本文方法和 6 种比较方法对城市公园遥感影像分类的结果图。可以看出,SVM、MLE 和 NNs 等传统分类方法错分漏分现象严重,呈现出以下共性:1)对分布较为规则集中的常绿树的分类效果略好,对西南角居民区内分布较为零散的落叶树误分现象严重;2)草地在 3 种绿地类型中面积占比最

小,光谱特征与落叶树相似,分类精度最差,其中 NNs 未能识别出草地;3)MLE 由于高分辨率遥感影像类内差异大,存在同物异谱和同谱异物现象,落叶树与草地混淆现象较为严重,存在明显的“椒盐噪声”;4)SVM 在高分影像城市绿地分类中未发挥优势,对 3 种绿地的提取效果一般。相比之下,SegNet、DeepLabv3+ 和 U-Net 等深度学习模型在高分影像城市绿地分类中更具优势,3 种方法均能有效抑制地物混分现象,降低了分类过程中出现的“椒盐噪声”。相比 SegNet 和 DeepLabv3+ 模型分类结果,U-Net 模型对城市绿地的细节信息表达更精确,边界与真实绿地更吻合,且对两种以上混交绿地也能较好区分,例如排列整齐的树林间隙中的草地也能够实现有效提取。在 7 种分类方法中,本文提出的 U-Net+ 模型分类精度最好,在 U-Net 模型的基础上总体提高了 0.85%。

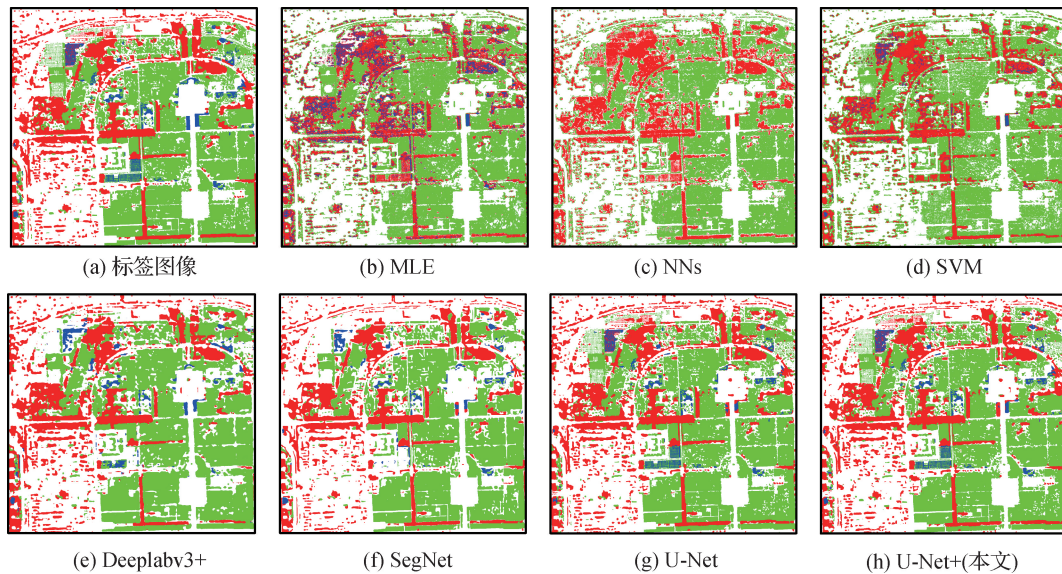


图 7 不同方法对城市公园遥感影像分类结果对比

Fig. 7 Comparison of classification results of urban parks remote sensing images among different methods

((a) label image; (b) MLE; (c) NNs; (d) SVM; (e) Deeplabv3+; (f) SegNet; (g) U-Net; (h) U-Net+(ours))

表 3 为本文方法和 6 种比较方法对城市公园遥感影像分类的精度对比。可以看出,按总体分类精度(OA)排序,分类效果依次为 U-Net+(93.80%)、U-Net(92.95%)、SegNet(87.31%)、DeepLabv3+(84.51%)、SVM(77.94%)、MLE(76.25%)和 NNs(75.85%)。显然,本文提出的 U-Net+ 模型分类精度最优。

图 8 为本文方法和 6 种比较方法对高尔夫球场遥感影像分类的结果图。高尔夫球场主要以种植人

工草地为主,因此草地在 3 种绿地类型中的辨识度最高,分类精度较优;常绿树在 3 种绿地类型中种植面积最小,错分漏分现象严重,从图 8(b)可以看出常绿树面积异常增多;中部地区(六郎庄烈士陵园)分布着许多细碎的落叶树斑块,传统分类方法和 DeepLabv3+ 均未能有效识别,SegNet 模型能够识别出破碎的绿地图斑,但是识别效果低于 U-Net 模型。在 7 种分类方法中,本文提出的 U-Net+ 模型在高尔夫球场遥感影像分类中表现最优异,不同类别边

表 3 不同方法对城市公园遥感影像分类精度对比

Table 3 Comparison of classification accuracies of urban parks remote sensing images among different methods

方法	$F_1/\%$					OA/%	kappa 系数
	落叶树	常绿树	草地	其他	平均		
MLE	57.22	83.21	36.74	76.04	63.3	76.25	0.61
NNs	61.89	83.53	0.00	72.81	54.56	75.85	0.60
SVM	62.77	82.81	23.54	78.71	61.96	77.94	0.63
DeepLabv3 +	75.65	89.40	59.33	81.73	76.53	84.51	0.74
SegNet	84.54	90.88	67.32	84.68	81.86	87.31	0.79
U-Net	85.72	95.89	73.18	92.53	86.83	92.95	0.88
U-Net + (本文)	88.40	96.21	81.25	93.08	89.73	93.80	0.90

注:加粗字体表示各列最优结果。

界规则且清晰,草地的 F_1 为 93.64%,识别效果最佳;落叶树的 F_1 为 83.40%,其中密集落叶林的识别效果优于稀疏落叶林;常绿树的样本量较少,特征

学习不充分,导致 F_1 仅有 44.02%。与 U-Net 模型相比,本文提出的 U-Net + 模型总体分类精度提高了 1.71%。

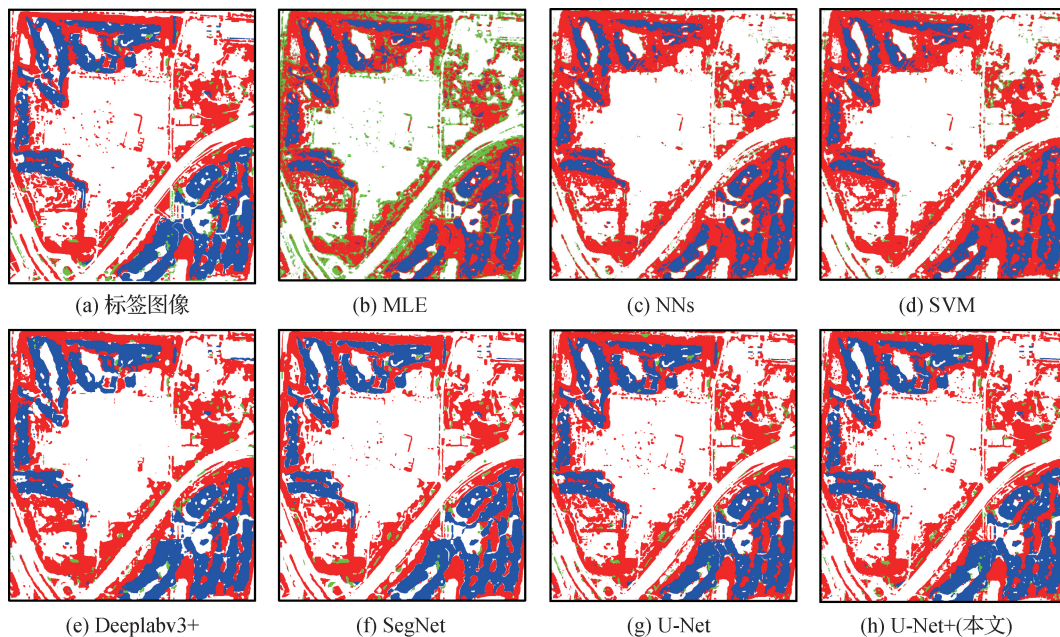


图 8 不同方法对高尔夫球场遥感影像分类结果对比

Fig. 8 Comparison of classification results of golf courses remote sensing images among different methods

((a) label image; (b) MLE; (c) NNs; (d) SVM; (e) Deeplabv3 +; (f) SegNet; (g) U-Net; (h) U-Net + (ours))

表 4 为本文方法和 6 种比较方法对高尔夫球场遥感影像分类的精度对比。可以看出,按总体分类精度 (OA) 排序,分类效果依次为 U-Net + (87.15%)、U-Net (85.44%)、SegNet (84.60%)、DeepLabv3 + (83.14%)、NNs (77.21%)、SVM (76.72%) 和 MLE (70.64%)。显然,本文提出的 U-Net + 模型分类精度依然最优。

图 9 为本文方法和 6 种比较方法对住宅区遥感影像分类的结果图。住宅区绿地主要包括住宅小区庭院绿地、花园绿地和城市道路等交通运输用地中的附属绿地,分布较破碎零散,主要以落叶树为主,常绿树为辅,草地类型极少。从图 9 可以看出,MLE 分类结果中常绿树出现大量错分现象,3 种传统分类结果中落叶树的 F_1 均低于 80%;4 种深度学习模

型均能实现高分影像住宅区零散绿地的提取,与 SegNet 和 DeepLabv3+ 相比,U-Net 预测结果与真实地物更加吻合;本文提出的 U-Net+ 模型的提取效

果最佳,比 U-Net 模型总体分类精度提高了 0.64%, 落叶树 F_1 为 90.57%, 比 U-Net 模型提高了 2.05%。

表 4 不同方法对高尔夫球场遥感影像分类精度对比

Table 4 Comparison of classification accuracies of golf courses remote sensing images among different methods

方法	$F_1/\%$					OA/%	kappa 系数
	落叶树	常绿树	草地	其他	平均		
MLE	65.64	17.22	79.09	82.18	61.03	70.64	0.59
NNs	69.11	9.46	78.50	86.74	60.95	77.21	0.66
SVM	68.64	12.02	76.80	86.89	61.09	76.72	0.65
DeepLabv3+	78.91	24.43	90.08	85.11	69.63	83.14	0.75
SegNet	81.48	31.65	91.22	86.21	72.64	84.60	0.77
U-Net	81.04	38.00	92.06	87.84	74.73	85.44	0.78
U-Net+(本文)	83.40	44.02	93.64	88.52	77.39	87.15	0.81

注:加粗字体表示各列最优结果。

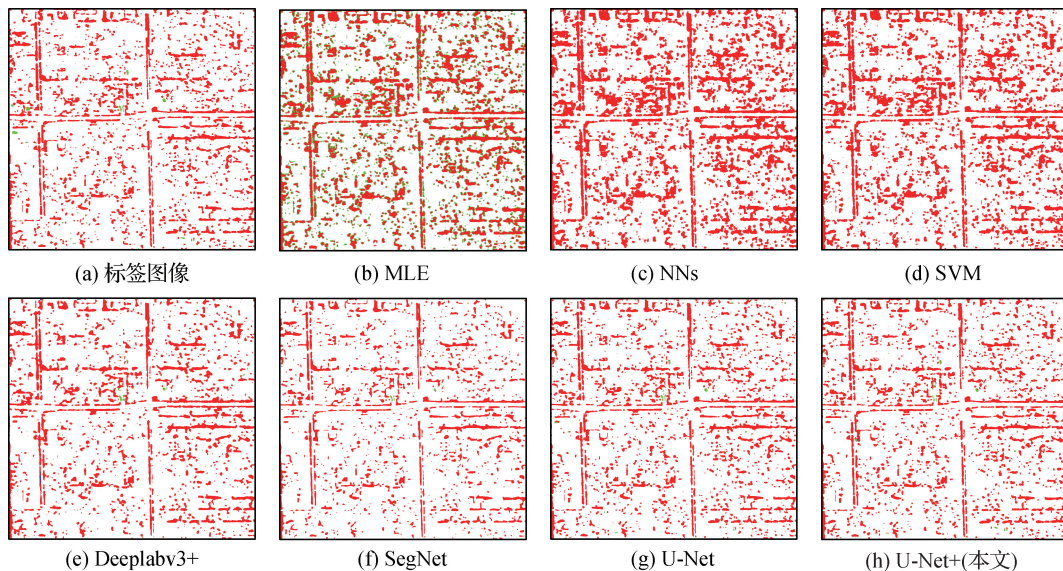


图 9 不同方法对住宅区遥感影像分类结果对比

Fig. 9 Comparison of classification results of residential areas remote sensing images among different methods

((a) label image; (b) MLE; (c) NNs; (d) SVM; (e) Deeplabv3+; (f) SegNet; (g) U-Net; (h) U-Net+(ours))

表 5 为本文方法和 6 种比较方法对住宅区遥感影像分类的精度对比。可以看出,按总体分类精度(OA)排序,分类效果依次为 U-Net+(96.92%)、U-Net(96.28%)、SegNet(94.79%)、DeepLabv3+(94.34%)、SVM(89.04%)、NNs(86.55%)和 MLE(84.36%)。显然,本文提出的 U-Net+ 模型分类精度最优。

通过上述对比分析,得出以下结论:面对高分辨

率遥感影像,基于深度学习技术的遥感分类方法相比传统机器学习算法能得到更优异的城市绿地遥感影像分类结果。MLE、NNs 和 SVM 等方法属于机器学习较浅层的学习算法,受计算单元限制,无法扩展到大型数据集解决较为复杂的分类问题;而 U-Net+、U-Net、SegNet 和 DeepLabv3+ 等深度学习网络模型能将城市绿地遥感影像进行较好的分类,在于其能够更充分地挖掘海量数据的深层信息,更好地进行

表 5 不同方法对住宅区遥感影像分类精度对比

Table 5 Comparison of classification accuracies of residential areas remote sensing images among different methods

方法	$F_1/\%$					OA/ $\%$	kappa 系数
	落叶树	常绿树	草地	其他	平均		
MLE	77.31	-	-	90.23	83.77	84.36	0.59
NNs	70.00	-	-	91.34	80.67	86.55	0.62
SVM	73.93	-	-	93.07	83.50	89.04	0.67
DeepLabv3 +	82.79	-	-	96.64	89.72	94.34	0.79
SegNet	82.00	-	-	96.96	89.48	94.79	0.79
U-Net	88.52	-	-	97.81	93.16	96.28	0.86
U-Net + (本文)	90.57	-	-	98.18	94.38	96.92	0.89

注:加粗字体表示各列最优结果。“-”表示测试集中无该类型。

特征工程构建,对高维复杂问题进行函数表达,其分类结果不仅有效抑制了像素级分类过程中易产生的椒盐噪声,而且在一定程度上减少了地物像元混分现象。通过对比实验证明,本文采用的 U-Net + 模型在 3 个验证区的 7 种方法中总体分类精度最优,相比 U-Net 模型分类精度有一定程度的提高,说明该方法能够有效提升高分率城市绿地遥感影像分类提取精度,为今后城市绿地遥感影像分类提供了一种新的智能解译方法。

同时,实验结果发现,深度学习方法在城市绿地遥感影像分类过程中具有一定的容错能力。在标签数据制作时,由于人工目视解译费时费力,难免存在误判现象,样本的选取不能保证百分之百正确。图 10 展示了深度学习样本的容错性。可以看出,第 1 行影像中的落叶树在目视解译时出现混淆像元, U-Net + 的预测结果排除了这部分错误信息,预测结果更加接近真实地物;第 2 行影像中的道路周边零散分布的常绿树在目视解译时未采集,但 U-Net + 模型的预测结果却得到了有效表达,说明深度学习模型对训练样本具有较强的容错能力,通过对大量标签数据训练学习,能够像人脑一样“去伪存真”,有效减少误差的影响,提高识别能力。

4 结论

本文将深度学习卷积神经网络应用于高分率城市绿地遥感影像分类。由于人工构建深度学习标签数据集困难,为解决小样本数据集不足可能带来

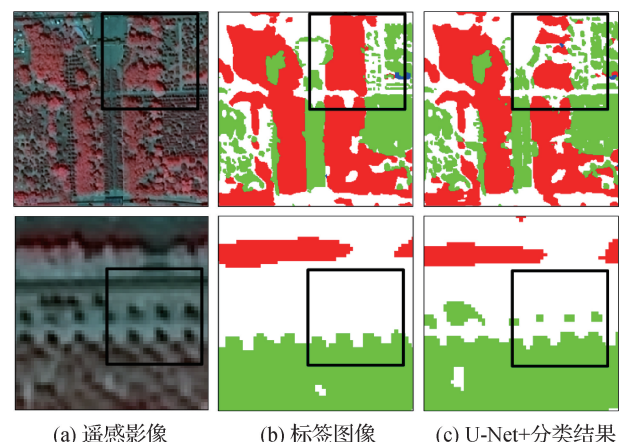


图 10 深度学习样本的容错性

Fig. 10 Fault tolerance of deep learning samples

((a) remote sensing images; (b) label images; (c) classification results of U-Net +)

的模型过拟合问题,本文对 U-Net 网络架构进行改进得到 U-Net + 模型。在此基础上,选取北京市城市公园、住宅区和高尔夫球场等 3 个典型城市绿地样本区进行 7 组不同分类模型的实验对比。实验结果证明,添加 BN 层、dropout 层和随机数据增强处理可以有效防止模型过拟合,提高了模型的泛化能力和鲁棒性。本文方法的总体分类精度为 92.73%、kappa 系数为 0.89、平均 F1 分数为 91.85%。7 种分类方法按照总体分类精度从大到小排序依次为 U-Net + (92.73%)、U-Net (91.67%)、SegNet (88.98%)、DeepLabv3 + (87.41%)、SVM (81.32%)、NNs (79.92%) 和 MLE (77.21%)。本文方法分类效果显著优于 MLE、NNs 和 SVM 等浅层

机器学习分类算法,较优于 DeepLabv3 + 和 SegNet 等深度学习模型,与 U-Net 模型相比,总体分类精度提升了 1.06%。本文方法有效克服了传统分类过程中的“椒盐现象”,实现了城市绿地遥感影像的高精度像素级分类。此外,深度学习模型对训练样本具有较强的容错能力,有效减少了误差的影响,提高了不同绿地遥感影像类型的识别能力。通过对比实验证明,U-Net + 卷积神经网络在精细化的城市绿地遥感影像分类提取中具有较强的可行性和准确性,为今后开展城市绿地遥感影像信息调查提供了一种高效的智能解释方法。

深度学习方法的学习能力和泛化能力强,能够从海量高分辨率遥感影像中挖掘待分地物光谱、纹理、形状等潜在的特征信息,自动构建特征工程,解决了高分辨率城市绿地遥感影像在分类中的异质性和类内差异等一系列问题。然而,分类效果的好坏还受到数据质量的影响,由于高分辨率遥感影像波段数有限,在一定程度上限制了特征学习的丰富度。在后续研究中,将尝试扩展多个通道来补充多源数据,例如增加植被特征来提高模型的分类精度;同时,本文模型属于基础性模型,后期工作将从模型的架构模式、损失函数等方面进一步提高模型性能。

参考文献 (References)

- Badrinarayanan V, Kendall A and Cipolla R. 2017. SegNet: a deep convolutional encoder-decoder architecture for image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12): 2481-2495 [DOI: 10.1109/TPAMI.2016.2644615]
- Beijing Municipal Institute of city Planning and Design. 2010. Planning of green space system in Beijing [EB/OL]. [2020-02-19]. <http://yllhj.beijing.gov.cn/zwgk/ghxx/gh/201911/P020191129502960993686.pdf> (北京市城市规划设计研究院. 2010. 北京市绿地系统规划[EB/OL]. [2020-02-19]. <http://yllhj.beijing.gov.cn/zwgk/ghxx/gh/201911/P020191129502960993686.pdf>)
- Chen L C, Zhu Y K, Papandreou G, Schroff F and Adam H. 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 833-851 [DOI: 10.1007/978-3-030-01234-2_49]
- Diakogiannis F I, Waldner F, Caccetta P and Wu C. 2020. ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162: 94-114 [DOI: 10.1016/j.isprsjprs.2020.01.013]
- Helber P, Bischke B, Dengel A and Borth D. 2018. Introducing EuroSAT: a novel dataset and deep learning benchmark for land use and land cover classification//2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: IEEE: 204-207 [DOI: 10.1109/IGARSS.2018.8519248]
- Hinton G E, Osindero S and Teh Y W. 2006. A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7): 1527-1554 [DOI: 10.1162/neco.2006.18.7.1527]
- Hirasawa T, Aoyama K, Tanimoto T, Ishihara S, Ozawa T, Ohnishi T, Fujishiro M, Matsuo K, Fujisaki J and Tada T. 2018. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer*, 21(4): 653-660 [DOI: 10.1007/s10120-018-0793-2]
- Ioffe S and Szegedy C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL]. [2020-02-19]. <https://arxiv.org/pdf/1502.03167.pdf>
- Jin J, Zhu H Y, Li Z X and Sun J W. 2014. The comparison of several kinds of supervised classification methods in ENVI remote sensing image processing. *Water Conservancy Science and Technology and Economy*, 20(1): 146-148, 160 (金杰, 朱海岩, 李子潇, 孙建伟. 2014. ENVI 遥感图像处理中几种监督分类方法的比较. *水利科技与经济*, 20(1): 146-148, 160) [DOI: 10.3969/j.issn.1006-7175.2014.01.058]
- Kussul N, Lavreniuk M, Skakun S and Shelestov A. 2017. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(5): 778-782 [DOI: 10.1109/LGRS.2017.2681128]
- Li R R, Liu W J, Yang L, Sun S H, Hu W, Zhang F and Li W. 2018. DeepUNet: a deep fully convolutional network for pixel-level sea-land segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(11): 3954-3962 [DOI: 10.1109/JSTARS.2018.2833382]
- Lin T Y, Goyal P, Girshick R, He K M and Dollár P. 2017. Focal loss for dense object detection//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 2999-3007 [DOI: 10.1109/iccv.2017.324]
- Liu D W, Han L and Han X Y. 2016. High spatial resolution remote sensing image classification based on deep learning. *Acta Optica Sinica*, 36(4): #0428001 (刘大伟, 韩玲, 韩晓勇. 2016. 基于深度学习的高分辨率遥感影像分类研究. *光学学报*, 36(4): #0428001) [DOI: 10.3788/AOS201636.0428001]
- Liu H, Luo J C, Huang B, Hu X D, Sun Y W, Yang Y P, Xu N and Zhou N. 2019. DE-Net: deep encoding network for building extraction from high-resolution remote sensing imagery. *Remote Sensing*, 11(20): #2380 [DOI: 10.3390/rs11202380]
- Liu P, Choo K K R, Wang L Z and Huang, F. 2017. SVM or deep learning? A comparative study on remote sensing image classification. *Soft Computing*, 21(23): 7053-7065 [DOI: 10.1007/s00500-016-2247-2]
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks

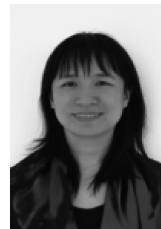
- for semantic segmentation//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE: 3431-3440 [DOI: 10.1109/CVPR.2015.7298965]
- Pacifici F, Chini M and Emery W J. 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing of Environment*, 113 (6): 1276-1292 [DOI: 10.1016/j.rse.2009.02.014]
- Qian Y G, Zhou W Q, Li W F and Han L J. 2015. Understanding the dynamic of greenspace in the urbanized area of Beijing based on high resolution satellite images. *Urban Forestry and Urban Greening*, 14(1): 39-47 [DOI: 10.1016/j.ufug.2014.11.006]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R. 2014. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15 (1): 1929-1958
- Wu L J. 2006. Studies on landscape pattern and biodiversity conservation of municipal open space in Beijing. Beijing: Beijing Forestry University (吴丽娟. 2006. 北京城市绿地景观格局与生物多样性保护研究. 北京: 北京林业大学)
- Xu H M, Qi H, Nan K and Chen M. 2019. High-resolution remote sensing image classification by combining deep learning with nDSM. *Bulletin of Surveying and Mapping*, (8): 63-67 (许慧敏, 齐华, 南轲, 陈敏. 2019. 结合 nDSM 的高分辨率遥感影像深度学习分类方法. *测绘通报*, (8): 63-67) [DOI: 10.13474/j.cnki.11-2246.2019.0253]
- Zhang Z X, Liu Q J and Wang Y H. 2018. Road extraction by deep residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5): 749-753 [DOI: 10.1109/LGRS.2018.2802944]
- Zheng M G, Cai Q G, Qin M Z and Yue T X. 2006. A new approach to accuracy assessment of classifications of remotely sensed data. *Journal of Remote Sensing*, 10(1): 39-48 (郑明国, 蔡强国, 秦明周, 岳天祥. 2006. 一种遥感影像分类精度检验的新方法. *遥感学报*, 10(1): 39-48) [DOI: 10.11834/jrs.20060107]

作者简介



徐知宇, 1996 年生, 女, 硕士研究生, 主要研究方向为基于深度学习的人居环境信息提取和分析评价。

E-mail: xuzhiyu@aircas.ac.cn



周艺, 通信作者, 女, 研究员, 主要研究方向为人居环境遥感。

E-mail: zhouyi@aircas.ac.cn

王世新, 男, 研究员, 主要研究方向为灾害与环境遥感应用。

E-mail: wangsx@aircas.ac.cn

王丽涛, 男, 副研究员, 主要研究方向为灾害与环境遥感应用。E-mail: wanglt@aircas.ac.cn

王振庆, 男, 博士研究生, 主要研究方向为深度学习语义分割与目标检测。E-mail: wangzhenqing19@mails.ucas.ac.cn