

引用格式: Wang Jing, Gao Shuai, Guo Liang, *et al.* Impervious Surface Extraction from High-resolution Images based on Multi-scale Feature Fusion in U-Net Network[J]. Remote Sensing Technology and Application, 2022, 37(4): 811-819. [王晶, 高帅, 郭亮, 等. 基于多尺度特征融合的U-Net网络高分影像不透水面提取研究[J]. 遥感技术与应用, 2022, 37(4): 811-819.]
DOI: 10.11873/j.issn.1004-0323.2022.4.0811

基于多尺度特征融合的U-Net网络高分影像不透水面提取研究

王晶¹, 高帅², 郭亮^{3,4}, 汪云⁵

(1. 中国地质大学 地球科学与资源学院, 北京 100083;

2. 中国科学院空天信息创新研究院, 北京 100094;

3. 广州市城市规划勘测设计研究院, 广东 广州 510060;

4. 广东省城市感知与监测预警企业重点实验室, 广东 广州 510060;

5. 北京林业大学 园林学院, 北京 100083)

摘要:不透水面作为监测城市生态环境的重要指标,其信息提取具有重要意义。由于城市地表的复杂性及细化的城市管理需要,急需提取高精度的城市不透水面。但是基于传统方法提取高精度的城市不透水面面临巨大困难。而深度学习方法因其自动化提取影像特征的特点逐渐成为遥感影像地物提取的新兴方法。基于此,采用多尺度特征融合的U-Net深度学习方法以提升语义分割精度,开展高分辨率遥感影像不透水面的精确提取研究。模型引入残差模块代替普通卷积以加深网络,提取更多影像特征;加入金字塔池化模块增强网络对复杂场景的解析能力;利用跳跃连接方式融合不同尺度特征,有利于恢复空间信息。以广州市航摄正射影像为数据源,通过卷积神经网络将遥感影像分割为背景、其他、植被、道路和房屋5种地物类型,将其与人工目视解译的地面真值进行验证,最终提取研究区域不透水面。实验证明:多尺度特征融合的U-Net模型总体精度和Kappa系数分别为87.596%和0.82。在定性与定量两个方面均优于传统的监督分类法、面向对象分类法和经典U-Net模型法。结果表明:该模型利用多维度影像特征信息,有效提升了复杂场景图像的分割精度,分割效果好,适用于高分辨率遥感影像不透水面提取,该研究成果可为城市环境监测提供数据支撑。

关键词:不透水面;U-Net模型;残差模块;金字塔池化模块;高分辨率遥感

中图分类号: TP79 **文献标志码:** A **文章编号:** 1004-0323(2022)04-0811-09

1 引言

随着城市化进程加快,土地利用类型逐渐从以植被、土壤为主转变为以城市不透水面用地为主^[1]。不透水面是指城市中由各种不透水性材料组成的

人工硬质表面,主要包括建筑屋顶、交通用地和其他硬化地表^[2]。城市不透水面是评估城市生态环境质量的重要指标,也是一系列生态环境问题的主要影响因素^[3]。高精度的不透水面数据常被用于研究城镇化进程、城市气候变化、热岛效应等问题^[4]。因

收稿日期:2021-09-08;修订日期:2022-07-14

基金项目:国家自然科学基金项目(42171377),国家重点研发计划(2017YFA0603004),高分项目(30-Y20A15-9003-17/18),广东省重点领域研发计划(2020B0101130009),广东省城市感知与监测预警企业重点实验室基金项目(2020B121202019)。

作者简介:王晶(1997-),女,天津人,硕士研究生,主要从事机器学习与遥感信息处理研究。E-mail:wj970205@126.com

通讯作者:高帅(1983-),男,山东高密人,副研究员,主要从事辐射传输遥感建模、机器学习与空间数据挖掘等研究。

E-mail: gaoshuai@radi.ac.cn

此,准确快速地获得高精度城市不透水面数据对城市监测与建设具有重要意义^[5]。

遥感作为一种远距离的探测技术,具有监测范围广、时间分辨率高、成本低等优势^[6]。因城市地表空间范围广,地物类型复杂等特点,国内外学者相继使用遥感方法实现对城市不透水面的提取。提取方法主要有指数法、回归模型法、光谱混合分解法和分类法^[7]。指数法是根据遥感影像中不透水面的光谱信息进行不同波段组合的方法。曹勇等^[8]基于裸地区域指数构建了抑制裸地的不透水面指数,实验表明该指数方法可有效消除裸地在不透水面区域提取中的影响。杨栩等^[9]利用无人机影像在可见光波段范围内构建了绿-蓝不透水面指数,解决了无人机影像提取不透水面的部分错分问题。指数法虽操作简便,但阈值大多依靠经验知识,难以准确划定。光谱混合分解法主要针对中低分辨率遥感影像,是由 Ridd 等^[10]于 1995 年首次提出,其基于多光谱遥感数据提出“植被—不透水层—土壤”分布模型,将城市土地类型表达为植被、不透水层、土壤的线性组合,用以进行土地分类。Wu 等^[11]在此基础上,采用多光谱数据提出“植被—高反照率—低反照率—土壤”四端元模型,将高、低反照率丰度组合为不透水面。实验证明混合光谱分解法可有效解决中低分辨率遥感影像提取不透水面的混合像元问题,但由于城市背景的复杂性,端元难以确定。李方刚等^[12]使用多源数据从多角度构建不透水面信息的多指数特征,利用多元回归方法建立不透水面覆盖率估算模型,并在典型城市进行验证,实验表明回归模型法对于不透水面的提取在一定程度上优于传统方法,但对于大区域不透水面覆盖率的估算精度仍有待提高。

尽管传统提取不透水面方法简单便捷,但对于地物细节丰富的高分辨率遥感影像仍具有挑战性。高分影像具有高空间、高时间分辨率、多波段等显著优势^[13],但也因此加大了同类地物间混合像元等因素的影响^[14]。并且,由于城市地物空间分布紧密,影像“同物异谱、异物同谱”现象尤为明显^[15]。因此直接在高分分辨率遥感影像上使用传统地物提取方法会出现错分严重、精度偏低等问题。

深度学习作为一种模拟人脑自主学习特征的机器学习方法^[16],已逐渐应用于高分分辨率遥感影像的处理中。刘文雅等^[17]基于高分二号数据,利用 DeepLabV3+深度学习网络实现对城市绿地的自

动化提取。罗琪等^[18]基于高分 1 号与 Sentinel-2 融合数据,使用深度学习算法对河北省邯郸市邱县的地膜进行识别。张泽瑞等^[19]利用深度学习语义分割模型,结合高分分辨率遥感影像与雷达数据对广西贵港市新增建设用地区域进行自动检测。研究表明,深度学习算法可有效解决高分分辨率影像复杂信息识别及提取问题,对城市高精度生态环境监测具有重要意义。

随着深度学习的发展,语义分割领域出现多种卷积神经网络以实现图像的像素级分割^[20]。Long 等^[21]提出全卷积神经网络用于图像分类,将图像分类级别扩展到像素级,开启了语义分割的大门。Ronneberger 等^[22]在全卷积神经网络的基础上,提出呈现 U 型对称的 U-Net 网络,该模型通过将低维与高维影像特征进行融合拼接,达到提升分割精度的目的。Zhao 等^[23]提出 PSPNet(金字塔场景解析网络),该模型通过融合不同区域的上下文信息,以提高预测全局信息的能力,并在多个数据集上取得了良好的效果。武花等^[24]通过引入膨胀卷积模块改进 PSPNet 实现对高分分辨率遥感影像的建筑物提取。卢麒等^[25]利用密集连接机制与空洞空间金字塔模块实现从 GF-2 遥感影像中提取建筑物信息。宋廷强等^[26]通过引入通道注意力机制与空间金字塔池化模块对 U-Net 结构进行改进,实现对道路的提取。研究表明深度学习可有效提取遥感影像中的道路与建筑物等信息。尽管以上方法都能实现遥感影像的图像分割任务,但在提取精度和提取细节等方面仍有提升空间。

实验将多尺度特征融合的 U-Net 卷积神经网络应用于广州市越秀区高分分辨率遥感影像分类中,提出一种面向高分分辨率遥感影像的城市不透水面自动分类技术。依据广州市越秀区航摄正射影像,制作研究区高分分辨率遥感影像深度学习数据集。模型在 U-Net 结构的基础上引入残差结构及金字塔池化模块以融合多维度特征信息,提高影像分割的准确性。使用改进的 U-Net 深度学习模型对研究区进行地物分类,提取不透水面区域,并将提取结果与最大似然法分类结果、面向对象法分类结果和经典 U-Net 分类结果进行对比分析。

2 研究方法

实验中不透水面提取过程分为 4 个阶段,技术路线如图 1 所示。第一阶段为高分分辨率数据预处理

理。将研究区域航摄正射影像图进行图像校正、融合、裁剪等预处理操作,得到 $1\,024\times 1\,024$ 图幅高分影像数据集。按6:2:2的比例将数据集划分为训练集、验证集、测试集。训练集与验证集数据通过人工目视解译方式获得地面真值标签,整理得到研究区深度学习数据集。第二阶段为模型训练,使用的模型为改进的U-Net深度学习网络。该结构在U-Net模型的基础上采用ResNet50进行特征提取,通过加深网络层数,以提取更多影像特征。同时加入金字塔池化模块,结合上下文特征信息,提高影像整体分割精度。对训练集数据及标签进行数据增强操作后,输入网络中进行训练。利用验证集数据调整网络参数,保存最优训练模型。第三阶段为模型测试,将测试集数据输入训练好的卷积神经网络进行逐像素预测,得到影像分类结果图并进行精度分析。第四阶段为结果对比,将研究区内本模型提取结果与最大似然法、面向对象法和经典U-Net模型法分类结果进行对比分析,以验证本文方法的有效性。

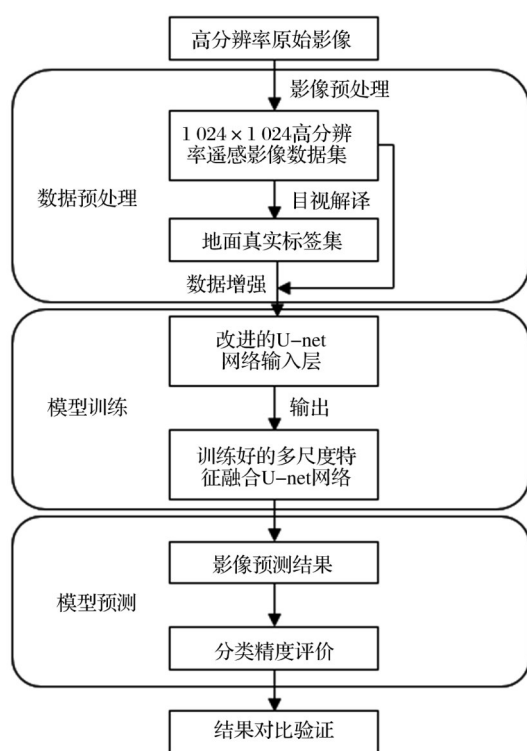


图1 技术路线图

Fig.1 Technical flowchart

2.1 数据预处理

首先对研究区域航摄正射影像数据进行图像校正、图像融合等预处理,由于深度学习网络对影像输入大小有限制,因此将融合后的影像裁剪为固

定大小的影像块,整理得到研究区 $1\,024\times 1\,024$ 影像数据集。为准确获取本研究区不透水面,使用Labelme软件构建标签数据集,将影像中的地物分类为背景、其他、植被、道路、房屋5类。其中水体类光谱特征明显,可准确提取水体的方法众多,因此使用已有数据集先验知识提取水体,分类为背景类,整理得到研究区深度学习数据集。由于卷积神经网络模型参数众多,在数据量不足的情况下极易出现模型过拟合等问题,使得模型泛化能力降低。因此,实验针对卷积神经网络的准确性及泛化性要求,将训练集数据及对应标签进行数据增强处理。主要为图像的平移、翻转、色彩变换以及增添噪声等操作,增加了数据的多样性,如图2所示。

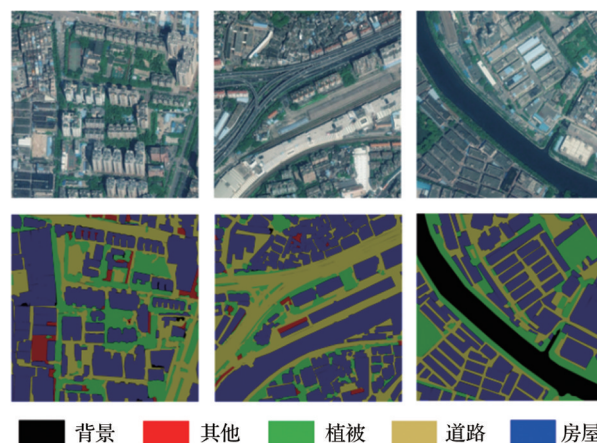


图2 数据集影像及其标签

Fig.2 Data set image and label

2.2 深度学习网络结构

深度学习网络结构采用基于U-Net网络的“编码器—多特征融合—解码器”结构。经典U-Net模型主要应用于医学影像分割,对于地物类型复杂的遥感影像需要有更深层的卷积神经网络来提取多样的影像特征,以实现较好的分割效果。Resnet残差网络相比于U-Net网络具有更深层次,且其独有的残差块学习方式可以在较少增加计算内存的基础上,学习更多影像特征。残差网络已被用于多种遥感影像信息提取的场景中,均获得较好效果。

网络编码器使用Resnet50结构代替经典U-Net网络中的普通卷积,相较于U-Net的特征提取结构,Resnet50的网络结构更深,可有效提取不透水面的高维度特征,更有利于复杂场景遥感影像的地物分类。残差结构如图3所示,分为两部分,一部分为输入特征的通路,另一部分对输入特征进行卷积运算,最后将两部分的特征相加。假设网络输入

特征为 x , 学习到的特征与输入之间的残差为 $F(x)$, 则学习到的特征为 $J(x) = x + F(x)$ 。当残差 $F(x)$ 为 0 时, 该结构为恒等映射, 但实际网络中 $F(x)$ 不为 0。此结构使得在输入特征的基础上又学习到新的特征, 可有效提升网络性能。卷积神经网络共包括 4 个下采样层, 逐层将图像维度扩大为原来的 2 倍, 图像尺寸缩小为原来的 $1/2$, 此操作使网络训练参数减少, 有利于网络收敛。卷积操作后使用 ReLu 函数进行非线性激活。

由于实验所提取的不透水面信息无规则形状, 且相邻地物多样, 环境复杂。随着卷积神经网络深

度的增加, 下采样提取的特征图易丢失大量信息, 影像部分低维度特征无法有效提取, 从而导致预测结果的边缘轮廓不清晰, 细节处表达错误等问题。为实现精确提取不透水面的目的, 需要利用深度学习网络获取多维度特征。因此在模型下采样操作后, 增加金字塔池化模块, 用以实现多维度特征融合提取, 提高分割精度。金字塔池化模块将输入的特征图通过 4 种不同尺度的池化得到不同尺度的特征图, 感受野大的池化操作获取影像全局信息, 感受野小的池化操作获取影像局部信息。经过 1×1 的卷积层和上采样后, 将特征图还原至输入尺寸并进行融合。由此可获得不同尺度下的影像特征信息, 实现多尺度特征提取。

在网络右半部分的上采样过程中, 特征图与下采样过程同一级的低维特征图进行连接, 以达到增强低维特征, 防止细节信息丢失的目的。为防止卷积神经网络在训练过程中出现梯度消失及梯度爆炸问题, 网络在每次卷积操作之后都加入 Batch Normalization 层, 并将前一层结果输入到 ReLu 激活函数中。最后将特征图输入到 1×1 的卷积层中, 利用 Softmax 函数进行分类处理, 得到与原图像尺寸、类别数相同的输出图像。这样的网络结构在一定程度上增强了低维特征信息, 提高了目标分割精度。改进网络如图 4 所示。

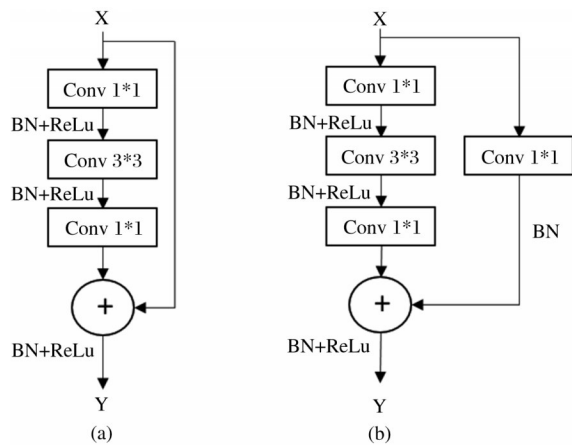


图 3 残差模块

Fig.3 Residual module

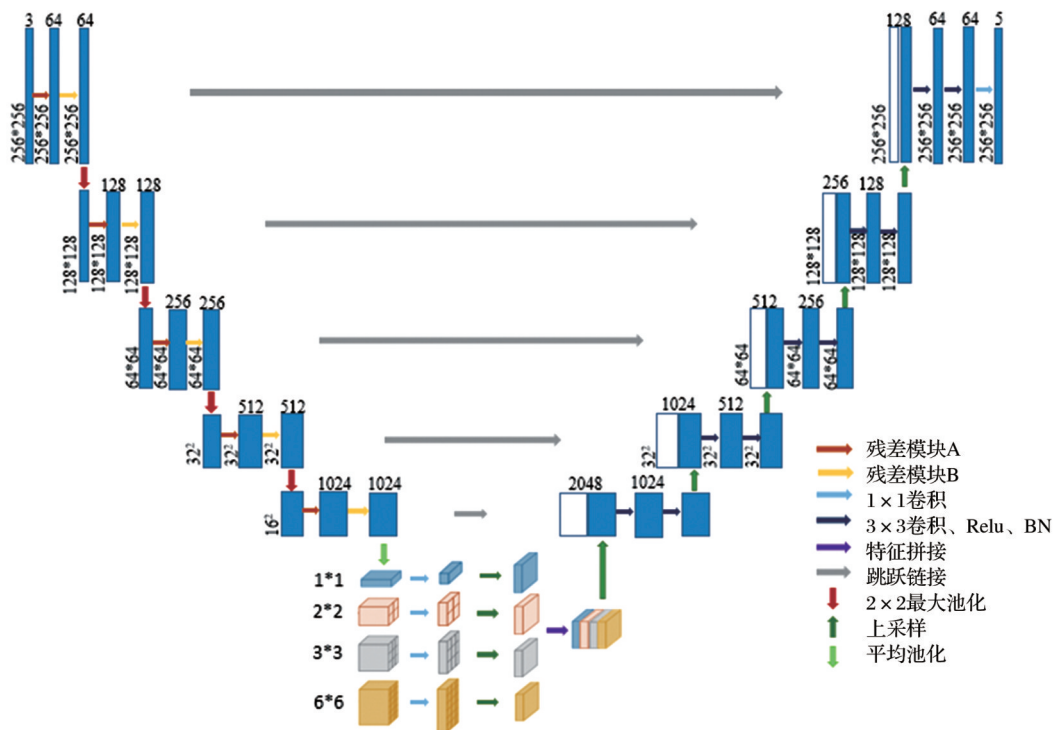


图 4 多尺度特征融合 U-Net 网络

Fig.4 Multi-scale feature fusion U-Net network

2.3 训练模型

网络模型采用深度学习框架Pytorch实现,按照

主流深度学习配置搭建实验环境。基础平台设置如表1所示。

表 1 实验环境配置

Table 1 Experimental environment configuration

项目	硬件系统	CPU	GPU	Pytorch	Python	CUDA
内容	Ubuntu 18.04	Intel E5-2603	NVIDIA TITAN X (Pascal)	1.6	3.6	10.1

综合考虑到模型计算量,硬件条件和结果精度要求等方面,经过多次试验,最终将模型批量大小设置为8,迭代轮数设置为300,使用Adam作为算法优化器,ReLU作为非线性激活函数,损失函数为交叉熵损失函数,初始学习率为0.000 1。

模型训练主要分为4步。首先,将数据增强操作后的训练样本及对应标签图像输入网络进行训练,影像类别先验知识通过标签图像输入进网络。其次,数据在卷积神经网络中进行前向传播,通过收缩路径提取图像特征,经过多特征融合后,上采样至原始图像尺寸,在1×1卷积层对影像进行逐像素分类,得到预测结果。使用交叉熵损失函数计算网络预测值与标签值之间的损失。最后采用Adam优化算法计算网络梯度,将参数反向传播并进行更新,以此降低损失值。重复上述步骤直至损失函数收敛,保存当前最优深度学习模型。

2.4 精度评价指标

为定量分析模型分割精度,采用交并比(IoU)、Kappa系数、总体分类精度、准确率、召回率及F1分数作为评价指标。计算方法如表2所示。

3 实验数据及结果分析

3.1 研究区概况

广州市为广东省省会城市,坐落于中国南部,广东省中南部。全境共12个区,包括越秀区、荔湾区、天河区、黄埔区、海珠区、番禺区、南沙区、萝岗区、花都区、白云区、增城市和从化市,总面积达7 435 km²,常住人口为1 530.59万人。广州市地域辽阔,地物类型丰富,涵盖了水域、山地、林地、农田及城市用地等多种类型。其中越秀区作为广州市政治、科技、教育及文化中心,是人口最密集的中心城市。

随着城镇化进程的加快,广州市经济迅猛发展,逐步将以植被、土壤为主的土地利用方式改变为以城市用地为主,此类现象在中心城市尤为显著。因土地利用结构的变化,不透水面占比持续增加,由此带来的城市环境问题亟待解决。

表 2 分类结果评价指标

Table 2 Evaluation index of classification results

评价指标	计算方法
IoU	$IoU = \frac{ A \cup B }{ A \cap B }$
	其中A代表预测结果的像素集合,B代表真实值的像素集合。
Kappa系数	$Kappa = \frac{p0-pe}{1-pe},$
	$P_e = \frac{a1*b1+a2*b2+\cdots+an*bn}{N*N}$
	其中P ₀ 表示每一类正确分类的像元数占总像元的比例,a代表每个类别的实际像元数,b代表每个类别的预测像元数,N代表总像元数。
总体分类精度	正确分类的像元数占总像元数的比例
准确率	被正确分类的像元数占预测为正确的总像元数的比例
召回率	被正确分类的像元数占实际正确的总像元数的比例
F1分数	$F1=2*\frac{precision*recall}{precision+recall}$

3.2 实验数据

以2008年广州市1:2 000彩色数字正射影像数据作为主要研究数据,数据来源于广州市航空摄影资料,搭载的航空遥感传感器能够同时获得立体影像和彩色多光谱影像,具有多波段成像、高定位精度、快速机动能力等特点。影像空间分辨率0.2 m,覆盖广州市全境7 435 km²,包括越秀区全域。对其进行遥感影像校正、融合与裁剪等预处理,得到越秀区1 024×1 024图幅大小的高分辨率遥感影像数据集。同时使用研究区域的矢量边界数据作为辅助数据。

3.3 实验结果及分析

3.3.1 模型训练结果

网络模型分割结果如表3所示。模型总体分类精度为87.596%,Kappa系数为0.82,Kappa系数结果位于[0.81,1],表示预测结果与真实值几乎完全一致。在各类别的预测结果中,植被IoU指数最高,为0.85,其次为房屋,IoU为0.81,道路IoU指数为0.66。在准确率评估中,背景类、植被类和房屋类均在0.9以上,说明网络对于此类地物的分割效果良好。由于不透水面类包含道路类与房屋类,故总体

精度较高。整体来看,该网络可有效分割不透水面类。

表 3 模型精度评价						
Table 3 Model accuracy evaluation						
	精度	召回率	F1分数	IoU	Kappa	总体精度
背景	0.94	0.65	0.77	0.62		
其他	0.18	0.16	0.17	0.09		
植被	0.91	0.94	0.92	0.85	0.82	87.596%
道路	0.74	0.86	0.80	0.66		
房屋	0.92	0.88	0.90	0.81		

基于此,实验完成了对广州市越秀区不透水面的提取,提取结果如图 5 所示。通过对广州市越秀区不透水面提取结果统计发现,越秀区全域不透水面占比 63.75%,植被占比 32.08%,背景占比 3.71%,其他占比 0.46%。

3.3.2 结果对比分析

为验证本文深度学习模型提取不透水面的有效性,需要与其他提取不透水面方法进行对比。在使用同样训练样本及测试数据的前提下,同时使用最大似然法、面向对象法和经典 U-net 网络进行实



图 5 越秀区不透水面提取结果

Fig.5 Extraction results of impervious water surface in Yuexiu District

验,结果如表 4 所示。其中,最大似然法实验使用 ENVI 软件完成。面向对象法实验使用易康软件完成,通过设置分割尺度,解决多类别归属问题,实现地物自动分类,分类器采用支持向量机 SVM 分类器。

由表 4 的实验对比结果可知,最大似然法受椒盐噪声影响严重,虽可基本识别大面积不透水面类,但提取过程夹杂其他类别(区域 1、区域 3),无法

表 4 研究区提取结果比较					
Table 4 Comparison of extraction results in the study area					
	原图	最大似然法	面向对象法	经典 U-net 法	本文方法
区域 1					
区域 2					
区域 3					
区域 4					
区域 5					

准确表达。对于形状复杂,分布不均的建筑物(区域4、区域5),无法有效识别,易与道路类混淆,出现错分现象,预测结果混乱。面向对象方法基本克服影像椒盐噪声现象,对于植被类的分类结果较好,对于不透水面类出现严重的错分现象,无法有效识别建筑物与道路及背景的区别,将建筑物间的阴影部分错分为背景类(区域3、区域4)。且在提取过程中,不透水面的边界出现锯齿状现象(区域5),无法清晰表达。传统分类方法结果可明显看出影像的错分,漏分问题,尤其在房屋与道路、道路与其他以及背景与房屋之间,无法准确提取边界信息。U-Net深度学习方法有效解决了传统提取方法所出现的噪声,锯齿状边界等现象,可基本识别建筑物及道路并进行清晰的表达。但由于U-Net网络对于复杂场景遥感影像的学习能力有限,分类结果出现内部空洞(区域3)、无法区分阴影(区域2),边界分割不精细(区域5),部分类别错分(区域1)等问题。本文网络方法总体分类结果良好,能够较好地提取不透水面区域,对地物细节表述准确。消除了噪声和阴影的影响,对内部空洞现象也有一定改善,错分漏分问题被有效解决。提取的不透水面类边界清晰、位置准确,相较于其他3种方法有更好的发挥。

在研究区内,4种实验方法均能在一定程度上识别和提取不透水面区域,但精度不一。对于建筑物类,本文模型方法分类精度最高,能完整准确提取房屋信息。经典U-Net方法次之,部分房屋出现内部空洞现象,且房屋的阴影无法有效识别。面向对象法能准确识别房屋,但提取边界呈现锯齿状,提取结果凌乱。最差的为最大似然法,无法准确识别房屋,对于形状复杂的建筑物,提取结果与其他类别混淆。对于道路类,本文方法精度最高,可有效提取道路信息。U-Net方法次之,出现错分及内部空洞现象。最大似然法可识别道路,但受椒盐噪声影像严重且道路边界不清晰。面向对象法提取道路精度最差,无法准确描绘道路形状,漏分严重。

由于改进后的U-Net网络通过增加网络层深和多尺度特征信息融合模块,模型可充分结合上下文信息,有效融合多维度特征,在挖掘影像深层规律的同时,保留图像基本原始信息,实现了对复杂场景高分辨率遥感影像的精确提取。

从定量角度分析,本文网络分类结果最优,总体分类精度为87.596%,Kappa系数为0.82。其次

是经典U-Net网络,总体分类精度为86.9976%,Kappa系数为0.81。面向对象法次之,总体分类精度为75.1047%,Kappa系数为0.65。最大似然法分类结果最差,总体精度为73.8467%,Kappa系数为0.61。本文方法较其他3种方法精度分别提高了13.7493%,12.4913%和0.5984%。验证了本文方法对复杂场景的高分辨率遥感影像提取具有优势,分类结果准确可靠。

4 结 论

实验基于经典U-net模型提出一种融合多尺度特征的改进模型实现对广州市越秀区高分辨率遥感影像不透水面的提取。该模型在U-Net的基础上,引入残差网络进行特征提取,并加入金字塔池化模块实现多尺度特征融合。为进一步增强模型鲁棒性与泛化性,增加BN层与数据增强操作。实验结果证明:

基于多尺度特征融合的U-Net深度学习算法对广州市越秀区高分辨率遥感影像分类精度较高。总体精度为87.596%,Kappa系数为0.82。改进的U-Net模型通过残差网络和金字塔池化模块进行多维度特征融合,有效增强了模型的学习能力,解决了类别错分漏分、提取区域内部空洞等问题,对细小的不透水面区域也能准确识别提取,在高分辨率遥感影像精细化的城市不透水面分类提取中有较好的分类效果。本文模型与最大似然法、面向对象法和经典U-Net模型方法进行比较,分类精度最高。最大似然法与面向对象法存在明显的错分漏分现象,最大似然法椒盐噪声严重,面向对象法边界锯齿化。尤其体现在不透水面类。经典U-Net模型方法易出现内部空洞,边界划分不清等问题。证明改进的U-Net模型方法适用于高分辨率遥感影像的不透水面提取。

深度学习算法具有强大的数据表达能力,在遥感影像分类领域有巨大潜力。本文基于深度学习算法得到较好的实验结果,但仍有改进空间。首先可在大型数据集上进行实验,该模型的推广性有待进一步讨论。其次,样本数据有限,后期可探索多种训练样本数据增强方法。最后,深入研究广州市不透水面与城市影响因子之间的定量关系。将不透水面提取结果与城市影响因子进行综合分析,为城市建设提供理论支撑。

参考文献 (References):

- [1] Arnold C L, Gibbons C J. Impervious surface coverage: The emergence of a key environmental indicator[J]. Journal of the American Planning Association, 1996, 62(2): 243-258.
- [2] Weng Q. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends[J]. Remote Sensing of Environment, 2012, 117: 34-49. DOI: 10.1016/j.rse.2011.02.030.
- [3] Peng Wenfu, Zhang Donghui, He Zhengwei, *et al.* Study on response of urban surface temperature to impervious surface area in the city Chengdu[J]. Remote Sensing Information, 2010, 25(2): 98-102. [彭文甫, 张东辉, 何政伟, 等. 成都市地表温度对不透水面的响应研究[J]. 遥感信息, 2010, 25(2): 98-102.]
- [4] Arnold C L, Gibbons C J. Impervious surface coverage: The emergence of a key environmental indicator[J]. Journal of the American Planning Association, 1996, 62(2): 243-258.
- [5] Zhang H S, Lin H, Wang Y P. A new scheme for urban impervious surface classification from SAR images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018(139): 103-118.
- [6] Deng C B, Wu C S. A spatially adaptive spectral mixture analysis for mapping subpixel urban impervious surface distribution[J]. Remote Sensing of Environment, 133: 62-70. DOI: 10.1016/j.rse.2013.02.005]
- [7] Xu Hanqiu, Wang Meiya. Remote sensing-based retrieval of ground impervious surfaces[J]. Journal of Remote Sensing, 2016, 20(5): 1270-1289 [徐涵秋, 王美雅. 地表不透水面信息遥感的主要方法分析[J]. 遥感学报, 2016, 20(5): 1270-1289.]
- [8] Cao Yong, Tao Yuxiang, Deng Lu, *et al.* An impervious surface index construction for restraining bare land[J]. Remote Sensing for Land & Resources, 2020, 32(3): 71-79. [曹勇, 陶于祥, 邓陆, 等. 一种抑制裸地的不透水面指数构建[J]. 国土资源遥感, 2020, 32(3): 71-79.]
- [9] Yang Xu, Zhu Daming, Yang Runshu, *et al.* A visible-band remote sensing index for extracting impervious surfaces[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(8): 127-134. [杨栩, 朱大明, 杨润书, 等. 一种提取不透水面的可见光波段遥感指数[J]. 农业工程学报, 2020, 36(8): 127-134.]
- [10] Ridd M K. Exploring a V-I-S (Vegetation-Impervious Surface Soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities[J]. International Journal of Remote Sensing, 1995, 16(12): 2165-2185.
- [11] Wu C S, Murrar A T. Estimating impervious surface distribution by spectral mixture analysis[J]. Remote Sensing of Environment, 84(4): 493-505. DOI: 10.1016/S0034-4257(02)00136-0.
- [12] Li Fanggang, Li Erzhu, Samat Alim *et al.* Estimation of large-scale impervious surface percentage by fusion of multi-source time series remote sensing data[J]. Journal of Remote Sensing, 2020, 24(10): 1243-1254. [李方刚, 李二珠, 阿里木·赛买提, 等. 融合多源时序遥感数据大尺度不透水面覆盖率估算[J]. 遥感学报, 2020, 24(10): 1243-1254.]
- [13] He Guizhen, Cheng Xiaojun. Study on monitoring and evaluating of flood disaster based on RS and GIS[J]. Remote Sensing Information, 2012, 27(1): 84-87. [何桂珍, 程效军. 基于 RS 和 GIS 的洪涝灾害监测与评估研究[J]. 遥感信息, 2012, 27(1): 84-87.]
- [14] Zuo Jiaqi, Wang Zegen, Bian Jinhu, *et al.* A review of research on remote sensing for ground impervious surface percentage retrieval[J]. Remote Sensing for Land & Resources, 2019, 31(3): 20-28. [左家旗, 王泽根, 边金虎, 等. 地表不透水面比例遥感反演研究综述[J]. 国土资源遥感, 2019, 31(3): 20-28.]
- [15] Oon A, Sharfi H Z M, Lechner A M, *et al.* Discriminating between large-scale oil palm plantations and smallholdings on tropical peatlands using vegetation indices and supervised classification of Landsat-8[J]. International Journal of Remote Sensing, 2019, 40(19): 7312-7328.
- [16] Weng Q, Hu X. Medium spatial resolution satellite imagery for estimating and mapping urban impervious surfaces using LSMA and ANN[J]. IEEE Transactions on Geoscience and Remote Sensing, 2008, 46(8): 2397-2406.
- [17] Liu Wenya, Yue Anzhi, Ji Jue, *et al.* Urban green space extraction from GF-2 remote sensing image based on DeepLabv3+ semantic segmentation model[J]. Remote Sensing for Land & Resources, 2020, 32(2): 120-129. [刘文雅, 岳安志, 季珏, 等. 基于 DeepLabv3+ 语义分割模型的 GF-2 影像城市绿地提取[J]. 国土资源遥感, 2020, 32(2): 120-129.]
- [18] Luo Qi, Liu Xiaolong, Shi Zhengtao, *et al.* Study on plastic mulch identification based on the fusion of GF-1 and Sentinel-2 images[J]. Geography and Geo-Information Science, 2021, 37(1): 39-46. [罗琪, 刘晓龙, 史正涛, 等. 基于 GF-1 与 Sentinel-2 融合数据的地膜识别方法研究[J]. 地理与地理信息科学, 2021, 37(1): 39-46.]
- [19] Zhang Zerui, Liu Xiaoping, Zhang Honghui, *et al.* Automatic detection of newly increased construction land based on deep learning and multi-source remote sensing data[J]. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2022, 61(2): 28-37. [张泽瑞, 刘小平, 张鸿辉, 等. 基于深度学习与多源遥感数据的新增建设用地自动检测[J]. 中山大学学报(自然科学版), 2022, 61(2): 28-37.]
- [20] Ji Mengyu, Ji Xiaoming, Yu Zhilou. A review of semantic segmentation based on deep learning[J]. Information Technology and Informatization, 2017(10): 137-140. [计梦予, 裘肖明, 于治楼. 基于深度学习的语义分割方法综述[J]. 信息技术与信息化, 2017(10): 137-140.]
- [21] 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.
- [22] Ronnebergero, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C] // International Conference on Medical Image Computing and Computer-assisted Intervention, 2015: 234-241.

- [23] Zhao H, Shi J, Qi X, *et al.* Pyramid Scene Parsing Network [R]. IEEE Computer Society, 2016.
- [24] Wu Hua, Zhang Xinchang, Sun Ying, *et al.* Building extraction in complex scenes based on the fusion of multi-feature improved PSPNet model [J]. Bulletin of Surveying and Mapping, 2021(6):21-27.[武花,张新长,孙颖,等.融合多特征改进型PSPNet模型应用于复杂场景下的建筑物提取[J].测绘通报,2021(6):21-27.]
- [25] Lu Qi, Qin Jun, Yao Xuedong, *et al.* Buildings extraction of GF-2 remote sensing image based on multi-layer perception network[J]. Remote Sensing for Land & Resources, 2021,33(2):75-84.[卢麒,秦军,姚雪东,等.基于多层次感知网络的GF-2遥感影像建筑物提取[J].国土资源遥感,2021,33(2):75-84.]
- [26] Song Tingqiang, Liu Tongxin, Zong Da, *et al.* Research on road extraction method from remote sensing images based on improved U-Net network[J]. Computer Engineering and Applications, 2021,57(14):209-216.[宋廷强,刘童心,宗达,等.改进U-Net网络的遥感影像道路提取方法研究[J].计算机工程与应用,2021,57(14):209-216.]

Impervious Surface Extraction from High-resolution Images based on Multi-scale Feature Fusion in U-Net Network

Wang Jing¹, Gao Shuai², Guo Liang^{3,4}, Wang Yun⁵

(1.School of Earth Sciences and Resources, China University of Geosciences, Beijing 100083, China;

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

3.Guangzhou Urban Planning & Design Survey Research Institute, Guangzhou 510060, China;

4.Guangdong Enterprise Key Laboratory for Urban Sensing, Monitoring and Early Warning, Guangzhou 510060, China;

5.The College of Forestry of Beijing Forestry University, Beijing 100083, China)

Abstract: As an important index for monitoring urban ecological environment, the extraction of impervious surface is of great significance. Due to the complexity of urban surface and the need of detailed urban management, it is urgent to extract high-precision urban impervious surface. However, it is difficult to extract high precision impervious surface based on traditional methods. Deep learning method has gradually become a new method of remote sensing image feature extraction because of its characteristics of automatic image feature extraction. Based on this, this paper uses the U-Net deep learning method based on multi-scale feature fusion to improve the semantic segmentation accuracy, and carries out the research on the accurate extraction of impervious surface from high resolution remote sensing images. The residual module is introduced instead of convolution to deepen the network and extract more image features, the pyramid pooling module is added to enhance the network's ability to resolve complex scenarios. It is beneficial to recover spatial information by combining different scale features with jump connection. In this paper, aerial orthophoto images of Guangzhou were taken as the data source. Through convolutional neural network, the remote sensing image is segmented into five types of features: background, others, vegetation, road and building. Verify it with the ground truth value of manual visual interpretation, finally, the impervious surface of the study area was extracted. Experiments show that the overall accuracy and Kappa coefficient of the U-NET model are 87.596% and 0.82, respectively. It is superior to traditional supervised taxonomy, object-oriented taxonomy and classical U-Net model in both qualitative and quantitative aspects. The results show that the model can effectively improve the segmentation accuracy of complex scene images by using the multi-dimensional image feature information, and the segmentation effect is good, which is suitable for the extraction of impervious water from high resolution remote sensing images. The research results in this paper can provide data support for urban environmental monitoring.

Key words: Impervious surface; U-Net; Residual module; Pyramid pooling module; High resolution remote sensing