引文格式：陶从辉,高青山,赵梦琳. 改进U-Net网络的高分辨率遥感影像建筑物提取方法[J]. 地理空间信息, 2025, 23(1): 9-12. doi:10. 3969/j. issn. 1672-4623.2025.01.002

改进U-Net网络的高分辨率遥感影像建筑物提取方法

陶从辉[1]，高青山[1]，赵梦琳[1]
（1. 四维高景卫星遥感有限公司，浙江 杭州 310012）

摘　要：针对高分辨率遥感影像建筑物提取精度不高、易出现误提和漏提等问题，提出了一种改进U-Net网络的建筑物提取方法。以ResNet50为U-Net模型的编码器部分，同时引入CBAM混合注意力机制和FPN特征金字塔结构对网络进行优化，从而提高网络提取建筑物信息的准确度和稳健性。基于高景一号遥感影像，制作512×512大小的样本进行训练，并与U-Net、基于ResNet50骨干网络的U-Net网络和DeepLabv3+进行对比验证。结果表明，该算法具有更强的分割效果和性能，适用于不同类型的高分辨率建筑物提取任务。

关键词：高景一号；建筑物提取；U-Net；ResNet50；FPN；CBAM
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2025）01-0009-04

Building Extraction Method in High-resolution Remote Sensing Images Based on Improved U-Net Network

TAO Conghui[1], GAO Qingshan[1], ZHAO Menglin[1]
(1. Siwei Gaojing Satellite Remote Sensing Co., Ltd., Hangzhou 310012, China)

Abstract: In view of the low accuracy and misrepresentation and omission of building extraction in high-resolution remote sensing images, we proposed a building extraction method based on improved U-Net network. We used ResNet50 as the encoder part of U-Net model, and introduced CBAM hybrid attention mechanism and FPN feature pyramid structure to optimize the network to improve the accuracy and robustness of network in building information extraction. Based on SuperView-1 remote sensing images, we produced 512×512 samples for training, and compared with U-Net, the U-Net network based on ResNet50 backbone network and DeepLabv3+. The results show that this method proposed in this article has a stronger segmentation effect and performance, and can apply to different types of high-resolution building extraction tasks.

Key words: SuperView-1, building extraction, U-Net, ResNet50, FPN, CBAM

遥感影像包含地物丰富的纹理、形状、光谱信息，已成为许多应用任务的重要基础数据[1-3]，其中建筑物的高效精准提取尤为关键。然而，由于高分辨率遥感影像的复杂性以及建筑物类型的多样性，传统建筑物提取方法[4-5]往往效果有限，精度无法满足实际需求。近年来，深度学习技术在遥感影像分析中得到了广泛应用。U-Net网络作为一种强大的图像语义分割基础模型，在建筑物提取任务中取得了显著效果，如杨栋杰等[6]在U-Net网络中设计了多重高效通道注意力模块和多尺度特征融合注意力模块，突出了建筑物的浅层有效特征，提高了网络对不同尺度建筑物的分割性能；金澍等[7]在U-Net网络中引入卷积块注意力模块和空洞空间金字塔模块，增强了模型识别建筑物重要特征的能力，改善了复杂场景下小型建筑物的提取结果；秦梦宇等[8]在U-Net编码末端添加双重注意力模块，并改进了损失函数，进一步提高了模型对城市建筑物的提取能力；陈嘉浩等[9]利用全连接条件随机场对U-Net进行改进，在一定程度上解决了建筑物提取的边缘模糊问题；卢彻等[10]利用残差块代替了U-Net编码部分的卷积层，同时引入Dropout和BN层加速模型的训练效率，提高了建筑物的提取精度。尽管深度学习技术在建筑物提取方面已取得了一定的进展，但对于高分辨率遥感影像，建筑物提取仍存在提取精度不高、误提、漏提等问题，因此需要设计更加高级和复杂的深度学习网络。鉴于此，本文提出了一种改进U-Net网络的高分辨率遥感影像建筑物提取方法。

1　研究方法

1）U-Net网络[11]常用于图像分割任务。整个网络呈U形网络架构，由对称的编码器和解码器组成。编

项目来源：浙江省生态环境科研和成果推广项目（2021HT0061）。
第一作者简介：陶从辉（1992—），硕士，助理工程师，从事基于深度学习的遥感影像信息自动化提取工作，E-mail：937312331@qq.com。

码器部分通过卷积和池化层逐步降低图像尺寸，并提取高级的语义特征；解码器部分则通过上采样和卷积操作逐步恢复图像尺寸，并还原图像信息。U-Net网络采用跳跃连接对编码器中的特征图与解码器中对应层级的特征图进行拼接，使得网络能够同时利用低层和高层的特征信息，从而提高分割结果的准确性。

2）ResNet50网络。ResNet[12]是一种深度残差网络，可解决训练深层网络时遇到的梯度消失问题，通过引入残差连接和残差模块的方式实现。ResNet50具有50层网络结构，主要由Stage0、Stage1、Stage2、Stage3 、Stage4、Stage5六个阶段组成，其中Stage1、Stage2、Stage3、Stage4阶段分别利用3个、4个、6个、3个残差块进行反复堆叠提取深层次不同等级的语义信息，每个残差块由1×1、3×3和1×1三个卷积层组成，每个卷积层后接一个批归一化层和ReLU激活函数，并引入跳跃连接直接将输入与输出相加传递信息。

3）CBAM[13]混合注意力机制将通道注意力和空间注意力结合在一起，形成一个注意力模块。首先将输入特征图经过通道注意力模块获得通道注意力加权的特征图，再将其经过空间注意力模块获得最终的注意力加权特征图，最后输入后续的网络层进行特征提取和分类。

4）FPN特征金字塔结构。图像中目标通常以不同的尺度和大小出现，因此需要一个能够捕捉多尺度信息的网络结构，而FPN[14]特征金字塔网络能有效解决不同尺度目标的识别问题。FPN构建一个特征金字塔，通过融合不同层次的特征来处理不同尺度的目标。金字塔结构包含多个层次，每个层次代表输入图像的不同分辨率，最底层特征具有高分辨率但语义信息较少，最顶层特征具有低分辨率但语义信息较丰富。

5）改进的U-Net网络。鉴于高分辨率影像建筑物信息提取存在的问题，为更准确地提取高景一号影像中的建筑物信息，本文设计了一种改进的U-Net网络。将ResNet50网络Stage0的第一层7×7卷积的输出特征图Efeat1、Stage1的输出特征图Efeat2、Stage2的输出特征图Efeat3、Stage3的输出特征图Efeat4、Stage4的输出特征图Efeat5替换为U-Net网络编码器对应的特征层，以此增强网络对建筑物目标的特征提取能力。在编码器部分，为强化浅层特征的丰富空间信息和深层特征的抽象语义信息，在特征图Efeat1和特征图Efeat5后分别加上CBAM混合注意力机制。在通道注意力阶段，对输入特征进行全局平均池化和全局最大池化，通过共享的MLP结构学习这两个池化结果的通道注意力权重，将二者的权重相加并通过Sigmoid激活函数得到最终的通道注意力；在空间注意力阶段，计算拼接输入特征图在通道维度上的平均值和最大值，通过7×7卷积操作学习、拼接后得到特征图的空间注意力权重，通过Sigmoid激活函数归一化得到最终的空间注意力；将通道注意力和空间注意力相乘。在解码器部分，上采样操作可能导致空间分辨率下降，边缘和细节信息丢失，为有效解决该问题，采用FPN特征金字塔结构，自上而下对解码器每一层的输出特征图Dfeat1、Dfeat2、Dfeat3、Dfeat4进行信息融合，得到特征图Dfeat5，再将Dfeat5进行8倍上采样处理，并与Dfeat4融合得到Dfeat6，最后将Dfeat6进行2倍上采样处理并输出。

2 实验结果与分析

2.1 实验数据

高景一号卫星是我国自主研发的高分辨率遥感卫星，全色分辨率为0.5 m，多光谱分辨率为2 m，轨道高度为530 km，幅宽为12 km，过境时间为10:30。高景一号影像建筑物自绘数据来源于浙江省东北部区域的真彩色影像，空间分辨率为0.5 m，已进行辐射校正、正射融合、匀光匀色等预处理。将绘制好的数据集裁剪为512×512大小，同时采用随机旋转、水平翻转、垂直翻转和对角镜像等操作来扩充数据集，得到19 515张建筑物样本影像，按照8∶1∶1的比例划分训练集、验证集和测试集。对应数据的增强结果见图1。

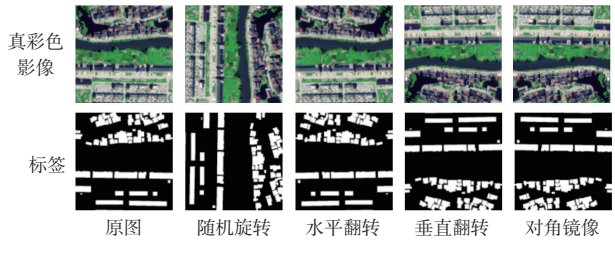


图1 数据增强

2.2 实验环境与参数

实验采用Windons10 64位操作系统，内存64 G，显卡为NVIDIA GeForce RTX 3060；采用Python3.7编程语言和深度学习框架Pytorch1.7进行网络环境搭建。实验训练采用Adam优化器，设置初始学习率为0.000 7，学习率衰减系数为0.96，同时采用Poly策略来动态调整学习率，Batch Size为4，Epoch为300；采用BCEloss与Diceloss之和作为网络的损失函数。

为更好地评估模型的性能和准确度，引入精确率（Precision）、召回率（Recall）、F1-score和交并比（IoU）进行定量分析。

2.3　结果分析

2.3.1　对比分析

为验证本文方法对高景一号影像建筑物提取的优势，选取U-Net、基于ResNet50骨干网络的U-Net（以下简称R50_U-Net）以及DeepLabv3+[15]进行对比分析。实验结果见表1，可以看出，相较于U-Net、R50_U-Net和DeepLabv3+，本文方法的Precision分别提高9.22%、3.12%、2.9%，Recall分别提高10.28%、9.31%、3.61%，F1-score分别提高9.75%、6.32%、3.26%，IoU分别提高10.79%、5.54%、2.68%，从定量分析的角度来看，本文算法在多个评价指标上都表现出明显优势。各方法的分割效果见图2，从可视化的角度来看，对于不同类型的建筑物，本文方法在分割效果上均能体现较强的稳健性和准确性，更接近标签图；数据1的结果表明本文方法在提取建筑物信息时漏提现象有所减少，保证了建筑物的完整性；数据2的结果表明本文方法能较好地保留建筑物的边缘信息，且具有清晰的分割轮廓线；数据3的结果表明本文方法在细节上较大程度地避免了建筑物粘连的情况，体现了较强的分割能力；数据4的结果表明本文方法不易受背景信息干扰，即使在较复杂的地物信息下仍能准确识别出建筑物，有效减少了建筑物误提的情况。

表1　建筑物提取指标对比/%

模型	Precision	Recall	F1-score	IoU
U-Net	84.33	83.45	83.89	77.25
R50_U-Net	90.43	84.42	87.32	82.50
DeepLabv3+	90.65	90.12	90.38	85.36
本文算法	93.55	93.73	93.64	88.04

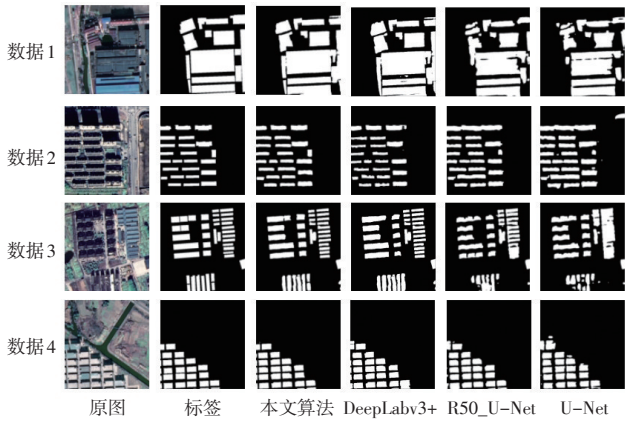


图2　高景一号影像建筑物提取结果对比

2.3.2　消融实验

为证明网络中CBAM混合注意力模块和FPN特征金字塔结构的有效性，本文设计了4组不同网络结构的消融实验，所涉实验在高景一号影像建筑物自绘数据上进行训练和测试，每组实验均以ResNet50为骨干网络。第一组实验为基础网络R50_U-Net，第二组实验增加了CBAM混合注意力机制（以下简称C_U-Net），第三组实验增加了FPN特征金字塔结构（以下简称F_U-Net），第四组实验为本文网络结构。实验结果见表2，可以看出，CBAM混合注意力模块和FPN特征金字塔结构在建筑物提取中发挥了积极的作用，C_U-Net使得IoU提升了1.24%，而F_U-Net使得IoU提升了3.95%，二者结合IoU可达到88.72%的精度，同时可使建筑物的定位更准确，边界更清晰，且明显增强了复杂场景的处理能力，充分彰显了两个模块的协同效应（图3）。

表2　消融实验提取指标对比/%

模型	Precision	Recall	F1-score	IoU
R50_U-Net	90.56	84.68	87.52	83.02
C_U-Net	91.83	89.27	90.53	84.26
F_U-Net	92.35	92.16	92.25	86.97
本文算法	93.82	94.02	93.92	88.72

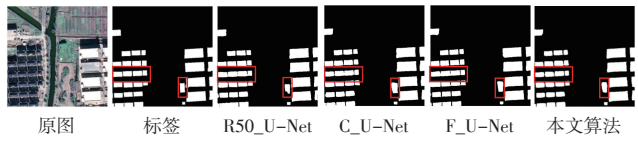


图3　消融实验提取结果对比

2.3.3　开源数据集验证

为体现本文方法不仅仅局限于高景一号建筑物提取，选取由武汉大学季顺平教授所在团队制作的高质量开源建筑物数据集[16]进行验证。随机选取武汉大学建筑物数据集中的7 670张影像，并采取相同的处理方式进行模型训练，实验结果见表3和图4，可以看出，本文方法在建筑物提取任务上均表现出较高的水平，IoU达到了89.92%，对建筑物的识别能力更精准，说明该算法不仅在高景一号建筑物提取任务上表现出色，而且面对其他开源建筑物数据时依然能保持较高的准确度，进一步印证了其在不同数据集上的优越性能，展现出较好的鲁棒性。

表3　武汉大学建筑物提取指标对比/%

模型	Precision	Recall	F1-score	IoU
U-Net	90.28	89.77	90.02	81.68
R50_U-Net	92.36	91.94	92.15	84.74
DeepLabv3+	93.15	92.06	92.60	87.07
本文算法	94.07	93.28	93.67	89.92

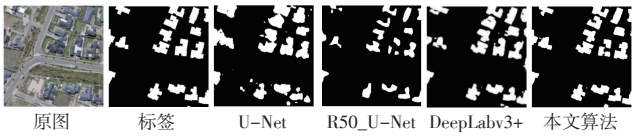


图4 武汉大学建筑物提取结果对比

3 结 语

为提升高分辨率遥感影像中建筑物的识别能力，本文基于高景一号卫星影像数据，提出了一种改进U-Net网络的建筑物提取方法。实验结果表明，本文方法的Precision、Recall、F1-score、IoU均优于U-Net、R50_U-Net、DeepLabv3+模型，表现出更高的准确度和更出色的提取能力，且在不同类型的建筑物提取上都能体现出较强的稳健性，有效减少了建筑物错提、误提和粘连现象，提取的建筑物轮廓更加清晰、规则；但本文方法对于连片密集的建筑物提取效果不是很理想，有待进一步优化。

参考文献

[1] 魏麟. 基于高分辨率遥感影像的震灾建筑物损毁检测[J]. 地理空间信息,2022,20(3):68-71

[2] 陆妍如,毛辉辉,贺琰,等. 基于Deeplab v3+的高分辨率遥感影像地物分类研究[J]. 地理空间信息,2022,20(6):1-6

[3] 许开宏,施招,马磊超,等. 基于机载激光雷达与高景一号数据的草原地上生物量反演研究[J]. 草业学报,2023,32(5):40-49

[4] Huang X,Zhang L. An SVM Ensemble Approach Combining Spectral,Structural,and Semantic Features for the Classification of High-resolution Remotely Sensed Imagery[J]. IEEE Transactions on Geoscience and Remote Sensing,2013,51(1):257-272

[5] 高贤君,郑学东,沈大江,等. 城郊高分影像中利用阴影的建筑物自动提取[J]. 武汉大学学报(信息科学版),2017,42(10):1 350-1 357

[6] 杨栋杰,高贤君,冉树浩,等. 基于多重多尺度融合注意力网络的建筑物提取[J]. 浙江大学学报(工学版),2022,56(10):1 924-1 934

[7] 金澍,关沫,边玉婵,等. 基于改进U-Net的遥感影像建筑物提取方法[J]. 激光与光电子学进展,2023,60(4):59-65

[8] 秦梦宇,刘勇,张寅丹,等. 基于改进U-Net模型的高分辨率遥感影像中城市建筑物的提取[J]. 兰州大学学报(自然科学版),2022,58(2):254-261

[9] 陈嘉浩,邢汉发,陈相龙. 融合级联CRFs和U-Net深度学习模型的遥感影像建筑物自动提取[J]. 华南师范大学学报(自然科学版),2022,54(1):70-78

[10] 卢彻,徐胜华,朱军. 改进U-Net的高分影像建筑物提取方法[J]. 测绘科学,2021,46(12):140-146

[11] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]. International Conference on Medical Image Computing and Computer-assisted Intervention,2015:234-241

[12] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, 2016:770-778

[13] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional Block Attention Module[C]//Proceedings of the European Conference on Computer VIsion(ECCV),2018:3-19

[14] Lin T Y, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection[C]. IEEE Conference on Computer Vision and Pattern Recognition,2017:936-944

[15] Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,40(4):834-848

[16] 季顺平,魏世清. 遥感影像建筑物提取的卷积神经元网络与开源数据集方法[J]. 测绘学报,2019,48(4):448-459

（上接第8页）

[EB/OL]. (2023-10-12)[2024-11-30]. http://ghj. enshi. gov. cn/xwzx/xsdt/202310/t20231012_1488896. shtml

[52] 丹江口市租用卫星全天候守水护水 保一库碧水永续北送[EB/OL]. (2024-07-04)[2024-11-30]. https://www. hubei. gov. cn/hbfb/xsqxw/202407/t20240704_5259604. shtml

[53] 郧阳区:建设监测管控平台 确保一库碧水永续北送[N]. 郧阳新闻,2024-06-29(2)

[54] 作者不详. 罗建军:数字城市地理空间框架建设及应用[J]. 中国建设信息,2011(23):13-15

[55] 作者不详. 国土测绘司负责人这样解读《智慧城市时空大数据平台建设技术大纲(2019版)》[J]. 国土资源,2019(4):38-39

[56] 自然资源部. 加快测绘地理信息事业转型升级 更好支撑高质量发展[J]. 习近平经济思想研究,2024(6):22-25

[57] 孙中原,邵成立,徐洪峰,等. 自然资源信息化三大信息平台及其应用[J]. 自然资源信息化,2024(2):29-36

[58] 自然资源部. 自然资源部地理信息管理司负责人解读《新一代地理信息公共服务平台(天地图)建设总体实施方案》[J]. 青海国土经略,2024(2):37-39

[59] 常钦. 40个智慧城市时空大数据平台建成[J]. 人民日报,2023-05-18(4)

[60] 罗平. 地理信息公共服务平台及其在智慧城市中的应用[J]. 智能城市,2019,5(6):80-81