引文格式：安文占,关涛,柴祥君,等. 利用高景卫星影像的超范围施工监测[J]. 地理空间信息, 2024, 22(8): 97-101. doi:10. 3969/j.issn. 1672-4623.2024.08.022

利用高景卫星影像的超范围施工监测

安文占[1,2]，关　涛[1,2]，柴祥君[1,2]，滕龙妹[1,2]，高青山[3]，赵梦琳[3]，邹　瑜[3]
（1. 浙江省国土空间规划研究院，浙江 杭州 311100；2. 浙江省微纳卫星研究重点实验室，浙江 杭州 311100；
3. 四维高景卫星遥感有限公司，浙江 杭州 310030）

摘　要：针对传统超范围施工建筑监测存在效率不高以及泛化能力较弱等问题，提出基于U-Net神经网络的高分辨率光学遥感影像超范围施工建筑提取方法。选用不同层数的Resnet特征提取主干网络替换Unet原始编码器作为ResUnet的主干网络，充分提取图像深层特征信息，通过轮廓化简及多边形化对语义分割结果进行轮廓规则化。以嘉兴海宁市为例，在区域内选取了大量城市和农村建筑物样本，构建了深度学习模型训练数据集。实验结果表明，采用改进后的U-Net模型在复杂地物背景下对超范围施工建设图斑定位识别较为精准，建筑物自动提取精度达到95%，召回率达到88%，减少了人工干预判别工作，显著提高国土空间用途管制评价和监管工作效率。

关键词：用途管制；超范围施工监测；深度学习；U-Net
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2024）08-0097-05

Over-range Construction Monitoring Based on SuperView Satellite Imagery

AN Wenzhan[1,2], GUAN Tao[1,2], CHAI Xiangjun[1,2], TENG Longmei[1,2], GAO Qingshan[3], ZHAO Menglin[3], ZOU Yu[3]
(1. Zhejiang Academy of Territorial Spatial Planning, Hangzhou 311100, China;
2. Key Laboratory of Micro-Nano Satellite Research Zhejiang Province, Hangzhou 311100, China;
3. Siwei Gaojing Satellite Remote Sensing Co., Ltd., Hangzhou 310030, China)

Abstract: In order to solve the problems of low efficiency and weak generalization ability in traditional monitoring of over-range construction buildings, we proposed an extraction method of over-range construction buildings from high-resolution optical remote sensing images based on U-Net neural network. We selected Resnet feature extraction backbone networks with different layers to replace the original U-Net encoder as the ResUnet backbone network, fully extracted the deep feature information of images, and regularized the semantic segmentation results through contour simplification and polygonization. Taking Haining District of Jiaxing City for example, we selected a large number of urban and rural building samples in the region to construct the training data set of deep learning model. The experimental results show that the improved U-Net model is more accurate in locating and identifying over-range construction map spots under the background of complex ground objects. The automatic extraction accuracy of buildings reaches about 95%, and the recall rate reaches about 88%, which can reduce the manual intervention and greatly improve the work efficiency.

Key words: usage regulation, over- range construction monitoring, deep learning, U-Net

超范围施工指项目实施过程中超出了用地审批范围进行施工建设，与批文不一致的情况。该监测评估指标作为国土空间用途管制监测评估的一项定量指标，在项目实施规范性审查、动态监测和管理等方面的作用重大[1]，尤其对有些超范围施工项目涉及侵占城市公共资源和公共安全，和对农村耕地及生态保护造成了一定的影响的情况。以往超范围施工建筑物的监测工作主要以人工判断为主，监测手段也多是实地探查或是群众监督举报，不仅效率较低，并且耗费大量人力物力。近年来，随着传感器与观测平台不断发展，对于农田、森林、城市等大面积区域监测，遥感技术能够提供更广的覆盖范围，更丰富的地物信息，更快速的动态化监测，使得超范围施工建筑物的自动识别提取成为可能[2]。超范围施工建筑物提取的前提是区域尺度整体建筑物提取，在建筑物提取结果的基础上进一步判断是否超范围。

传统建筑物识别方法多是依据建筑物的边线角点等几何形状特征以及光谱纹理等影像特征[3-4]，进而采用图像分割算法如区域增长[5]、边缘检测[6-7]、分水岭算法[8]、数学形态学[9-10]以及机器学习[11-13]等手段进行建筑物信息提取。但传统方法大多局限于人为特征工程以及建筑物特征明显、背景简单的情形，泛化能力一般。而

收稿日期：2023-03-17。
项目来源：国家重点研发计划资助项目（2022YFC3800805）。
第一作者简介：安文占（1987—），硕士，研究方向为遥感、土地资源管理和国土空间用途管制等，E-mail：261362060@qq.com。

深度学习能够利用多层神经网络结构，从海量数据中学习和存储更高层更抽象的表示特征，舍弃那些与判别无关的因素，用高效的算法替代手动特征工程，节省人力和时间，进而完成端到端的复杂语义识别问题[14]。因为模型对于不同场景具有较强的稳定性以及适用性，所以本文尝试使用深度神经网络实现大尺度建筑物识别提取，进而实现建筑物超范围施工识别。

1 研究方法

从遥感影像中提取建筑物可看作二分类的语义分割问题，即只有建筑物和背景两类，深度网络要做的就是对每个像素进行分类。目前深度学习中常用的语义分割网络有：全卷积神经网络（FCN）、SegNet、U-Net、Mask R-CNN、Deeplab系列等。综合研究区建筑物特点和训练时长，既考虑图像全局信息，又考虑图像细节信息，本文选用改进的ResUnet神经网络来训练模型完成建筑物语义分割。

Net模型是一种典型的语义分割模型，是对全连接神经网络（FCN）的改进，由编码器和解码器两部分构成[15-16]。U-Net模型的左半边是编码器，主要作用是对图像特征进行提取和压缩，右半边是解码器部分，是对提取的特征信息扩展映射[17]。U-Net模型共有23个卷积层，包括3×3卷积用于特征提取，解码器中的上采样2×2卷积用于恢复特征维度，以及最终1×1卷积用于分类结果映射[18]。编码器采用典型的卷积神经网络结构，即采用2个3×3卷积层进行特征提取和1个最大池化层进行特征压缩。解码器中先使用上采样操作，使用2×2的反卷积将特征图尺寸进行恢复，同时通道个数减半，再对恢复尺度的特征图进行2次3×3卷积运算，并重复3次。在最后的输出层，用全连接层将64维特征图映射成2维输出图。U-Net模型继承了FCN语义分割的思想，并融合了编码-解码结构和跳跃连接的特点，能够在小数据集上取得较为精确的分割结果。

针对建筑物具有密集离散、风格复杂等特点，高精度的影像使得影像纹理细节信息更加丰富，对网络的细节提取能力有较大的要求[1, 19]。为将U-Net网络用于超范围建设信息提取，本研究对原U-Net网络模型进行了针对性的优化，由于Unet仅使用9层卷积作为编码器，即无法充分提取图像更深层特征信息。为使编码器网络更为有效的提取图像深层特征，在ResUnet网络中，利用Resnet对深层语义特征提取能力较强这一特性，使用3通道作为网络输入，以适应RGB图像通道需求，通过选用不同层数的Resnet特征提取主干网络替换Unet原始编码器作为ResUnet的主干网络进行实验。本文使用Imagenet上预先训练好的ResUnet作为迁移学习的主干网络，保留计算机视觉中图像所要提取的共有特征，优化ResUnet特征提取方向。

在编码器阶段，模型使用Resnet作为ResUnet编码器主干网络，将Resnet4个采样模块与Unet编码器4个下采样模块相对应，并去除Resnet分类网络的全连接输出，以适应Unet语义分割网络需求，每批输入的样本通过批归一化操作化为正态分布，可以有效减轻模型训练过程中出现的梯度爆炸和消失现象，并有效加速了收敛[20]。解码器阶段保留Unet的解码器结构，使用反卷积操作作为上采样方式，并保持原始Unet相同的通道数，由于大型Resnet每层残差模块输出通道数量远大于原始Unet每次下采样输入横向连接的通道数量，参考FPN模块，使用1×1卷积，以较小的计算代价，压缩Resnet通道数量，保留较多的通道信息，将Resnet的输出通道数降到与原始Unet网络解码器相同的输入通道数量构成ResUnet，并与上采样的结果相连接。ResUnet结构如图1所示。

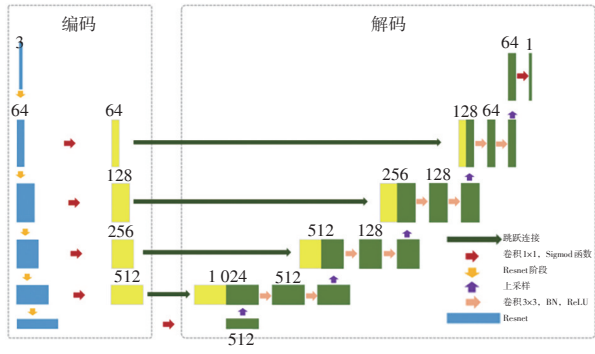


图1 ResUnet网络结构示意图

在ResUnet网络结构中，残差模块定义为本体映射和残差映射之和：

$$y_l = F(x_l) + x_l(1) \tag{1}$$

式中，y_l为残差块的输出；x_l为残差块的输入；F为神经元的求导结果。

$$\frac{\partial \text{Loss}}{\partial x_l} = \frac{\partial \text{Loss}}{\partial x_{l'}}\left(1 + \frac{\partial F(x_l)}{x_l}\right) \tag{2}$$

Loss指损失函数，常规神经元网络中的神经元求导结果：

$$\frac{\partial \text{Loss}}{\partial x_l} = \frac{\partial \text{Loss}}{\partial x_{l'}}\left(\prod_{i=l}^{l'-1} W_i\right) \tag{3}$$

式中，W_i为每层神经元权重。综合以上公式，ResUnet的神经元求导不会因为累乘W_i而出现弥散或爆

炸，并有效保证了收敛速度[21]。

由此可知，使用ResUnet，通道连接层结合下采样和上采样过程中的深浅层次信息，可以有效提取网格中的空间信息，有利于对建筑物信息进行监测和提取，有效提高识别准确度和收敛速度。在此基础上，引入空洞空间卷积池化金字塔（Atrous Spatial Pyramid Pooling），原始输入通过不同扩张率2、4、8的3×3卷积层，加上原有标准卷积和平均池化来提取不同尺度的特征。最后通过1×1卷积进行多尺度融合得到最终特征提取结果，通过不同扩张率卷积层可以增大感受野，获得不同尺度信息，有利于提高网络性能。在ASPP提取的特征图后增加通道和空间注意力机制（CBAM），CBAM在空间、通道两方面分别增加注意力，产生的权重和ASPP提取的特征相乘得到最终输入特征。能有效提高特征提取的关注点，减少不必要特征，提高最终训练效果。

此外，建筑物自动提取结果中会存在部分建筑物边界模糊、点数冗余、直角特征弱等问题，结果需进行优化操作。通过轮廓化简及多边形化对语义分割结果进行轮廓规则化。轮廓规则化算法主要包括2个步骤：①粗调整，消除分割和多边形化的明显误差；②精化，调整直线的方向和节点的位置。

1）粗调整过程。首先对512×512大小的预测结果图片进行中值滤波，去除孤立的噪声点；再进行连通域分析在图像中寻找出彼此互相独立的连通域并将其标记出来，进行轮廓定位，进行轮廓精简滤除冗余点。

2）精化过程。先获取识别结果图中所有多边形的轮廓线，计算轮廓线上每条边的长度和方位角，以最长边作为主方向，对其他边进行方向纠正，绕中心点旋转到与主方向平行或垂直，并将所有线连成多边形，得到轮廓规则化后的识别结果。

进行规则化处理后，建筑物边界清晰度提升、点数冗余降低、直角特征增强，建筑物的识别效果明显提升。

2　实验结果分析

2.1　数据集构建

实验区位于浙江省嘉兴市海宁市。根据海宁市项目红线结合建筑物分布特点，以高景一号卫星影像经正射融合、匀色镶嵌等数据预处理后得到海宁市完整RGB影像为地图选取样本数据。

由于本文识别目标为海宁市全部建筑物，而在影像中发现有建筑物遮挡等问题，建筑的底边轮廓无法作为准确的识别特征，因此以勾画建筑物屋顶进行标注。为保证样本的丰富度和深度学习模型的泛化能力，总计标注了10万多个样本，包括城市建筑与农村建筑样本，勾画比为5∶1。勾画样本和遥感影像的套合精度，合格率大于95%，勾画细分程度合格率高达98.8%。最后将标注的样本矢量转化为掩膜标签数据。

为适应网络输入，需要进行数据裁剪和格式转换[19]基于高分辨率。本研究将高分辨率遥感影像和掩膜标签数据裁剪为多个512×512尺寸，共计生成6 525张训练影像和标签。影像中包含背景，建筑物，标签数据一共有2个类别：0表示背景，1表示建筑物。并且由于原始影像和标签数据均为TIF格式，文件过大，影响程序运行速度，因此将TIF图像文件格式转换为PNG文件，在保证数据质量的同时降低空间使用。

2.2　精度评价指标

本文以像元为单位统计验证样本建筑物分类结果，得到混淆矩阵如表1所示，其中True Positive（TP）表示真实目标被正确分类；True Negative（TN）表示背景被正确分类；False Positive（FP）代表背景错分为目标；False Negative（FN）代表目标错分为背景[19-21]。基于上述指标值，使用图像语义分割中普遍的精度评估指标，分别为：准确率（Accuracy）、精确率（Precision）、召回率(Recall)和$F1$分数（F1-Score），计算公式如式（4）~（7）。

表1　混淆矩阵

		真实值	
		建筑物	非建筑物
预测值	建筑物	*TP*	*FP*
	非建筑物	*FN*	*TN*

准确率（Accuracy）表示正确分类的像素占总像素的比例，定义为：

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

式中，Precision为正确分类的目标像素与所有目标预测值的比例，定义为：

$$\text{Precision}=\frac{TP}{TP+FP} \tag{5}$$

式中，Recall为样本中有多少目标像素被正确的预测，定义为：

$$\text{Recall}=\frac{TP}{TP+FN} \tag{6}$$

式中，$F1$分数（F1-Score）为分类问题的一个综合指标，最大为1，最小为0。多分类问题的深度学习中，将F1-score作为最终评估指标。

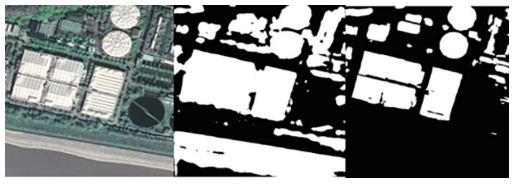
$$F1=\frac{2\times Precision\times Recall}{Precision+Recall} \quad (7)$$

2.3 实验结果分析

研究中，将数据集划分为80%训练集和20%测试集，基础学习率0.000 1，批次数量设置为8，训练和验证的Dice Loss损失函数，模型经过300 epoch迭代训练后，损失由开始的0.448下降至0.148左右，损失值基本趋于稳定，建筑物分类识别精度统计如表2，海宁市建筑物自动识别结果如图2所示。从精度指标统计可以看出，该模型的总体准确率为95%，召回率为88%，相比于Unet模型准确率提高了3%，召回率提高了1%。该模型建筑物的提取识别的精度较好。



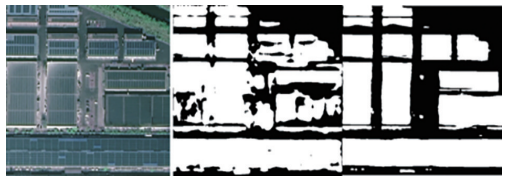
a 独立住宅



b 彩钢建筑



c 连片住宅



d 工业建筑

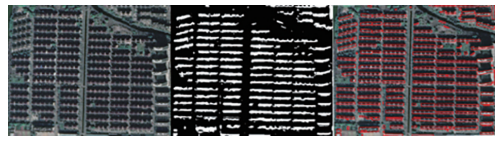
图2 海宁市建筑物提取结果

注：左边为原始影像，中间为Unet结果，右边为改进的ResUnet提取结果。

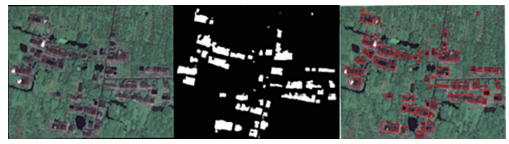
表2 精度指标统计

评估指标	Unet	改进的ResUnet
Accuracy	0.92	0.95
Precision	0.87	0.88
Recall	0.81	0.80
F1-Score	0.83	0.84

从分割效果看，对于建筑物而言，改进的ResUnet模型也明显优于Unet模型；整体分割结果中黏连、误提、小碎斑情况减少，轮廓更加清晰（图3）。



a 独立住宅



b 连片住宅



c 彩钢建筑



d 工业建筑

图3 海宁市建筑物识别结果轮廓规则化

注：左图为原始影像，中间的为改进的ResUnet分割结果，右边为轮廓规则化后结果。

针对不同建筑物类别，基于改进的ResUnet模型分割结果并结合轮廓规则化算法，首先将改进的ResUnet模型分割结果进行小图斑去除，再进行轮廓规则化，使得建筑物轮廓清晰度显著提高，能更好地服务于国土空间用途管制监测。

2.4 超范围施工建筑识别

利用训练模型得到海宁市全部建筑物的自动识别结果，与项目红线进行叠加分析得到超范围施工建设图斑，如图4所示。为了对深度学习模型提取出的超范围施工建筑进行验证，与原始影像进行叠加，在深度学习模型自动提取的建筑物图斑的基础上，结合已审批建设用地范围，进行人工判读提取超范围施工图斑。图斑之间以及图斑与已审批建设用地范围之间不存在重叠区域，疑似超范围施工图斑提取中最小上图面积为20 m^2。同一地块不同变化类型进行分类提取。经过人工检查，模型提取精度为92%。

3 结论与展望

3.1 结 论

针对国土空间用途管制监测需求，结合高景卫星亚米级高分辨率遥感影像，采用深度学习技术对超范围施工建筑进行动态监测，并选取了大量训练数据，提高了识别精度。对深度学习模型在超范围施工建筑图斑



图4　自动提取超范围建筑物图斑

注：红色为项目红线，黄色为自动提取结果。

提取中的适用性进行了探究和探索，取得如下结论：

1）高景一号0.5 m遥感影像具有丰富的地物信息，其像元精度也远高于普通的陆地观测卫星，是国土空间用途管制评价和监管应用的理想数据之一。

2）改进的ResUnet模型在自动化解译中更具优势，其精度更适用于国土空间用途管制评价和监管。

3）轮廓规则化算法通过粗调整，细化两步走，使得识别的建筑物轮廓更加清晰，更有利于超范围施工监测的精准判别，国土空间用途管制评价和监管应用。

3.2　展　望

本项目方法可基本满足复杂场景下的地物提取，但是在影像地物类型较多、地物形状多样且分布密集不规则的场景下，会存在个别地物单体提取边缘信息较差的问题。在今后的研究中，一是要继续扩大样本集，提升网络的泛化能力；二是尝试从不同的网络模型结构、优化器以及损失函数等方面考虑，利用消融实验设计出提升图像预测精度较高的神经网络模型；三是探索引入注意力机制来捕捉全局信息，增大感受野，避免卷积堆叠导致的计算效率减退问题。

参考文献

[1] 刘德祥. 基于深度学习的违章建筑智能识别[D]. 徐州:中国矿业大学,2021

[2] 苏健民,杨岚心,景维鹏. 基于U-Net的高分辨率遥感图像语义分割方法[J]. 计算机工程与应用,2019,55(7):207-213

[3] 王丹. 一种高分辨率遥感影像建筑物边缘提取方法[J]. 环境保护与循环经济,2009,29(10):26-28

[4] 林雨准,张保明,徐俊峰,等. 多特征多尺度相结合的高分辨率遥感影像建筑物提取[J]. 测绘通报,2017(12):53-57

[5] 傅刚. 高分辨率遥感影像分割方法综述[J]. 西部资源,2016,75(6):135-137

[6] 徐昌荣,葛山运. 基于面向对象的高分辨率遥感影像建筑物提取研究[J]. 城市勘测,2011(1):84-86

[7] 陈杰,邓敏,肖鹏峰,等. 基于分水岭变换与空间聚类的高分辨率遥感影像面向对象分类[J]. 遥感技术与应用,2010(5):597-603

[8] 林祥国,张继贤. 面向对象的形态学建筑物指数及其高分辨率遥感影像建筑物提取应用[J]. 测绘学报,2017,46(6):724-733

[9] 黄金库,冯险峰,徐秀莉,等. 基于知识规则构建和形态学修复的建筑物提取研究[J]. 地理与地理信息科学,2011,27(4):28-31

[10] 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 & Remote Sensing,2013,51(1):257-272

[11] 刘海飞,常庆瑞,李粉玲.高分辨率影像城区建筑物提取研究[J]. 西北农林科技大学学报(自然科学版),2013,41(10):221-227

[12] 孙宁. 面向高空间分辨率遥感影像的建筑物目标识别方法研究[D]. 杭州:浙江大学,2010

[13] Zhang Z,Liu Q,Wang Y. Road Extraction by Deep Residual U-Net[J]. IEEE Geoscience and Remote Sensing Letters,2017,15(5):749-753

[14] Zunair H,Ben H A. Sharp U-Net: Depthwise Convolutional Network for Biomedical Image Segmentation[J]. Computers in Biology and Medicine,2021,136(9):699-712

[15] 刘恒恒,张春森,葛英伟,等.多尺度特征融合深度学习建筑物的提取方法[J]. 地理空间信息,2022,20(2):97-100

[16] 张浩然,赵江洪,张晓光.利用U-net网络的高分遥感影像建筑提取方法[J]. 遥感信息,2020,35(3):143-150

[17] 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

[18] Zhang C,Rameau F,Lee S,et al. Revisiting Residual Networks with Nonlinear Shortcuts[J]. British Machine Vision Conference,2019,1(13):152

[19] 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

[20] 王曦.基于卷积神经网络的遥感图像语义分割方法研究[D]. 哈尔滨:东北林业大学,2021

[21] 景庄伟,管海燕,彭代峰,等.基于深度神经网络的图像语义分割研究综述[J].计算机工程,2020,46(10):1-17