

基于卷积网络的遥感图像建筑物提取技术研究

付 发, 未建英, 张丽娜

(河北地质大学信息工程学院, 河北 石家庄 050000)

摘 要: Mask RCNN是当前最高水平的实例分割算法, 本文将该算法应用到高分辨率遥感图像建筑物提取中, 提出了一种高效、准确的高分辨率遥感图像建筑物提取方法。首先, 利用Tensorflow和Keras深度学习框架搭建Mask RCNN网络模型; 然后, 通过有监督学习方式在IAILD数据集上进行模型学习。利用训练出的模型对测试集进行建筑物提取实验, 通过与基于KNN和SVM等建筑物提取方法对比可以看出, 本文方法可以更加完整的、准确的提取出建筑物。采用mAP评价指标对实验结果进行定量评价, 本文算法的查全率和查准率均大于对比算法, 且多次实验中本文算法的mAP均在81%以上, 验证了基于卷积网络的高分辨率遥感图像建筑物提取的有效性和准确性。

关键词: 深度学习; 建筑物提取; Mask RCNN; 卷积网络

中图分类号: TP183 **文献标识码:** A

A Study of Building Extraction from Remote Sensing Imagery Based on Convolution Network

FU Fa, WEI Jianying, ZHANG Lina

(School of Information Engineering, Hebei GEO University, Shijiazhuang 050000, China)

Abstract: Mask RCNN, the state-of-the-art instance segmentation algorithm, is adopted in the building extraction from high-resolution remote sensing imagery. The paper proposes an efficient and accurate building extraction method from high-resolution remote sensing imagery. Firstly, Tensorflow and Keras deep learning framework are used to build the Mask RCNN network model; then, the model learning is carried out on the IAILD data set by supervised learning. By using the trained model, the experiment of building extraction is carried out on the test set. By comparing with the building extraction methods, such as KNN and SVM, it can be seen that this method can be more complete and accurate to extract the building. The mAP evaluation index is used to evaluate the experimental results quantitatively. The recall and precision of the algorithm are all larger than the contrast algorithms, and the mAP of the algorithm is more than 81% in this paper, which verifies the validity and accuracy of the high resolution remote sensing image building extraction based on convolution neural network.

Keywords: deep learning; building extraction; Mask RCNN; CNN

1 引言(Introduction)

提取遥感图像建筑物信息是遥感图像解译中的重要研究课题之一。一方面, 快速准确地提取遥感图像中的建筑物能够满足遥感图像制图、地理信息系统的数据获取、城市空间数据库的更新和建设“数字化城市”的需要; 另一方面, 建筑物提取方法也能为其他类型地物的提取提供一定的思路, 因此建筑物提取方法的研究具有重要意义, 特别是高空间分辨率遥感影像的投入使用, 对遥感信息处理与分析提出了更多的要求和挑战^[1]。

目前, 国内外对于遥感图像建筑物提取研究有很多成果, 主要分为基于边缘提取、基于纹理提取和基于分类提取三大类。Huertas等人^[2]通过检测建筑物边缘的线条, 然后根据空间关系, 进而提取出建筑物; 叶盛^[3]根据幅度谱信息将建筑物的边缘特征和纹理特征结合起来, 从而对建筑物提取; Atz M等人^[4]提出区域合并分割技术进行建筑物提取; Zheng等^[5]首次训练了一个将CNN和CRF结合的端到端的神经网络, 从而对建筑物提取; Noh等人^[6]设计了一个端到端的神经网络提高了对小物体、有遮挡物体的分割准确率。

由于遥感图像建筑物提取主要针对城区，但城区绿化区域较多，植被对建筑物提取严重干扰，而且建筑物类型复杂，除了单幢的、规则的矩形建筑物，还有不规则的、复杂的建筑物，这对建筑物提取也带来了困难。基于边缘和纹理特征提取建筑物的准确率不高，而且有较高的误检率，因此，本文采用Mask RCNN深度卷积网络进行遥感图像建筑物提取。实验表明，深度卷积网络模型对植被信息干扰、建筑物类型复杂等问题具有很好的鲁棒性，可以达到较高的查全率和查准率。

2 算法描述(Algorithm description)

2.1 网络结构

Mask RCNN^[7]是在Faster RCNN^[8]的基础上融合了FCN^[9]和FPN^[10]思想的一种多任务深度神经网络。Mask RCNN有两个阶段：第一阶段，由Mask RCNN主干网络(即ResNet101^[11]和FPN)提取出遥感图像的特征图，然后使用区域建议网络(Region Proposal Network, RPN)生成目标的建议框，并对建议框筛选得到感兴趣区域(Regions Of Interest, ROIs)；第二阶段，对每个ROIs预测类别、位置和对应的建筑物的二值掩膜。网络整体结构如图1所示。

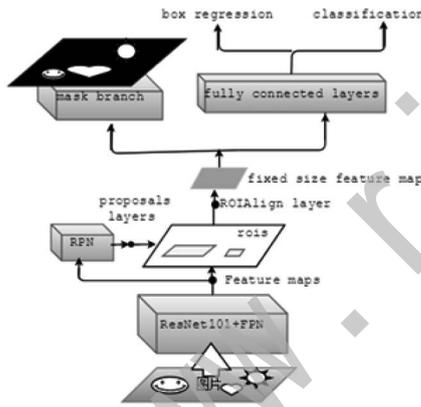


图1 Mask RCNN网络结构图

Fig.1 Mask RCNN structure chart

2.2 损失函数

损失函数表达了预测值与真实标签的差距程度，本文通过减小预测值与真实标签之间的损失函数来训练模型，Mask RCNN的损失函数定义为：

$$Loss = L_{cls} + L_{box} + L_{mask} \quad (1)$$

式(1)中 L_{cls} 表示分类误差； L_{box} 表示边界框回归误差； L_{mask} 表示掩码误差。

$$L_{cls} = \frac{1}{N_{cls}} \sum_i -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (2)$$

式(2)中 i 表示特征图中ROIs的下标； N_{cls} 表示类别数； p_i 表示第 i 个ROIs预测为正样本的概率；当ROIs为正样本时，

$p_i^*=1$ ，当ROIs为负样本时， $p_i^*=0$ 。

$$L_{reg} = \frac{1}{N_{reg}} \sum_i p_i^* R(t_i, t_i^*) \quad (3)$$

式(3)中 t_i 表示正样本ROIs到预测区域的四个平移缩放参数； t_i^* 表示正样本ROIs到真实标签的四个平移缩放参数； $R()$ 是 $smooth_{L_1}$ 函数，即 $smooth_{L_1} = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases}$

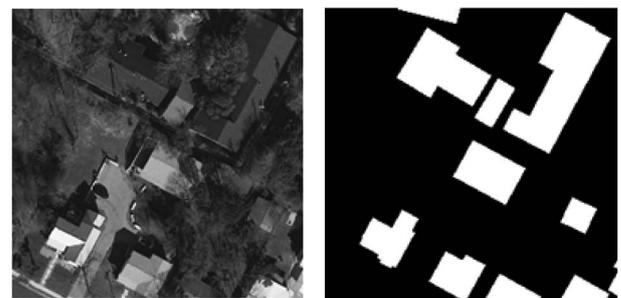
$$L_{mask} = [Km^2]_i \quad (4)$$

式(4)中 m^2 表示掩码分支对每个ROIs产生的 $m*m$ 大小的掩码； K 表示分类物体的种类数目； i 表示当前ROIs的类别。

3 数据集及预处理(Dataset and preprocessing)

在本文中，我们采用了法国国家信息与自动化研究所提供的Inria Aerial Image Labeling Dataset(简称IAILD数据集)。该数据集中每张图片是覆盖面积405公里，空间分辨率为0.3米的航空正射彩色图像，每张图像分为建筑物和非建筑物。这些图像覆盖了不同的城市居民点，从人口稠密地区(例如旧金山的金融区)到高山城镇(例如利恩茨在奥地利蒂罗尔，美国的一个城镇)。IAILD数据集包含180张5000*5000像素训练数据，以及相应的掩码图片。

我们将其分为训练集(150张)、验证集(5张)、测试集(25张)。由于图片尺寸太大，直接使用会造成内存溢出，所以我们使用黑色将图片的边缘填充60像素，得到了5120*5120的图片。然后，同时将图像和掩码分割为20行20列，共400张256*256大小的图片。最终，得到了训练集(60000张)、验证集(2000张)、测试集(10000张)。



(a)卫星图

(b)掩码图

图2 分割后的卫星图

Fig.2 Satellite map after segmentation

4 实验与算法评价(Experiment and algorithm evaluation)

4.1 算法实现

算法：使用Mask RCNN进行建筑物提取

Input：遥感图像(RGB)

Output：建筑物轮廓

- (1)将遥感图像进行分割，得到分割后的图像P。
- (2)将P输入ResNet101+FPN卷积网络中，得到图像的特征图集合。
- (3)将特征图集合输入到RPN中，得到候选框。
- (4)将候选框的特征交给预测器，得到每个候选框的类别、位置和掩码的预测值。
- (5)计算预测值与真实值的损失值loss。
- (6)优化损失函数loss，得到Mask RCNN网络模型。

我们使用内存为32GB，GPU(1080Ti)为11GB的主机对模型进行模型训练，每8张图片一个批次，训练25小时左右损失函数不再下降，停留在1.070左右。



图3 模型识别效果图

Fig.3 Model recognition effect diagram

4.2 算法评价

为了定量评价遥感图像提取算法的性能，本文采用的评价标准是mAP(mean Average Precision)。mAP是目标检测中衡量识别精度的指标，多个类别物体检测中每个类别都可以根据查准率(Precision)和查全率(Recall)绘制一条P-R曲线，AP就是该曲线下的面积，mAP是多个类别AP的平均值。分类结果混淆矩阵，见表1。

表1 分类结果混淆矩阵

Fig.1 Mixed matrix of categorized results

实际类别 预测类别	Positive	Negative
Ture	TruePositive(TP)	TrueNegative(TN)
False	FaslePositive(FP)	FalseNegative(FN)

TP(真正例)是将正类预测为正类的数量；TN(真负例)是将负类预测为负类的数量；FP(假正例)是将负类预测为正类的数量；FN(假负例)是将真累预测为负类的数量。

查全率、查准率和mAP的计算公式分别是：

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$mAP = \frac{\sum AveragePrecision_c}{Num(Classes)} \tag{7}$$

式(7)表示数据集中所有类别的平均精度的平均值

本文采用IOU(Intersection Over Union)来判断检测到的建筑物是否正确，即IOU>0.5时，说明检测到的是建筑物。

$$IOU = \frac{预测建筑物面积 \cap 标签建筑物面积}{预测建筑物面积 \cup 标签建筑物面积} \tag{7}$$

通过对测试集进行实验，得到mAP：0.8111。

为了验证本文算法在遥感图像建筑物提取方面优于其他算法。本文使用基于K-means、KNN、SVM的遥感图像建筑物提取方法对IAILD数据集进行了测试，评价结果见表2。

表2 算法性能评价

Tab.2 Algorithms performance evaluation

方法	K-means	KNN	SVM	本文方法
mAP	0.6838	0.6947	0.7075	0.8111

从上表可以看出本文算法的mAP高于其他方法10%以上，本文算法的查全率和查准率均大于对比算法，表明基于卷积网络的高分辨率遥感图像建筑物提取的有效性和准确性。

5 结论(Conclusion)

本研究基于MaskRCNN提出一种建筑物提取方法，通过使用深度学习框架Keras和Tensorflow搭建Mask RCNN网络模型，并通过IAILD数据集对模型进行有监督学习，从而提取出建筑物。并与K-means法、KNN法和SVM法的提取效果进行了对比。研究表明，本文方法提取的mAP较高，能有效提取出建筑物。提取建筑物时，只需遥感图像输入到网络模型中，即可提取建筑物的大致轮廓，不需另外的步骤，具有效率性和准确性。

本文方法仍存在一些不足，实验过程中发现在分割建筑物时，部分建筑物提取存在轮廓不规则；部分图像纹理特征类似建筑物的海洋和山地地区有4%左右的误检率。上述问题可作为今后研究的重点。

我们秉承开源，开放的技术分享思想，已经将代码放到了GitHub上：https://github.com/fusimeng/maskrcnn_building。该代码基于Tensorflow和Keras框架，简单易懂。

参考文献(References)

[1] 江华.基于eCognition面向对象技术的高分辨率遥感影像土地利用分类—以福州琅岐岛为例[J].海峡科学,2014(8): 12-17.

[2] A.Huertas and R.Nevatia.Detecting Building in Aerial

- Images.Computer Vision[J].Graphics and Image Processing,1998,41(2):131-152.
- [3] 叶盛.基于纹理特征的高分辨率遥感图像建筑物信息识别与提取方法研究[D].南京大学博士学位论文,2008:10-13.
- [4] Atz M and Schape A.Multiresolution segmentation:an optimization approach for high quality multiscale image segmentation[J].Journal of Photogrammetry and Remote Sensing,2000,58:12-23.
- [5] Zheng S,Jayasumana S,Romera-Paredes B,et al.Conditional Random Fields as Recurrent Neural Networks.Proceedings of 2015 IEEE International Conference on Computer Vision(ICCV),2015:1529-1537.
- [6] Noh H,Hong S,Han B.Learning Deconvolution Network for Semantic Segmentation.Proceedings of 2015 IEEE International Conference on Computer Vision(ICCV),2015:1520-1528.
- [7] Kaiming He,Georgia Gkioxari,Piotr Dollar,et al.Mask RCNN[C].In CVPR,2018,1:1703-1078.
- [8] Kaiming He,Shaoqing Ren,Ross Girshich,et al.Faster R-CNN:Towards Real-Time Object Detection with Region Proposal Networks[C].In CVPR,2016,1:1506-1611.
- [9] Jonathan Long,Evan Shelhamer,Trevor Darrell,et al.Fully Convolutional Networks for Semantic Segmentation[C].In CVPR,2015,5:1411-1445.
- [10] Tsung-Yi Lin,Piotr Dollar,Ross Girshick,et al.Feature Pyramid Networks for Object Detection[C].In CVPR,2017,4:1612-1618.
- [11] Kaiming He,Xiangyu Zhang,Shaoqing Ren,et al.Deep Residual Learning for Image Recognition[C].In CVPR,2015,12:1512-1516.

作者简介:

- 付发(1993-),男,硕士生.研究领域:人工智能与机器学习.
- 未建英(1993-),女,硕士生.研究领域:演化算法.
- 张丽娜(1993-),女,硕士生.研究领域:推荐系统.

(上接第10页)

- [7] 蒋慧,林东,谢希仁.UML状态机的形式语义[J].软件学报,2002,13(12):2244-2250.
- [8] 于晓玲,杨海波.UML状态图形式化方法的分析与比较[J].计算机与数字工程,2014,42(8):1488-1492.

(上接第13页)

网络模型,并通过参数测量与检验,发现相关系数值越大,网络越呈现小世界效应。研究表明,股票市场中存在着少数股票与大盘股指呈现密切的关联。通过度指数的建立与说明,在具体实验中得出金融体系中的度视角下的指数所反映出的股票份额占比最大,具体值为53.6%。通过在相似度指标体系的上述研究中,我们可以总结出,金融体系类股票在全部股票中的地位及其影响力作用是最多的,在整个股票复杂网络中扮演者举足轻重的影响作用,同时也说明的基于相似度的指数分析在对股票市场研究中具有一定的可操作性。

参考文献(References)

- [1] Cornelis A.Los,Bing Yu.Persistence characteristics of the Chinese stock markets[J].International Review of Financial Analysis,2008(17):64-82.
- [2] Yusoon Kima,Thomas Y.Choi,Tingting Yan.Structural investigation of supply networks:A social network analysis approach[J].Journal of Operations Management,2011(29):194-211.
- [3] NEWAN M E J.The structure and function of complex networks[J].SIAM Review,2003(45):167-256.
- [4] ALBERT R,BARABASI AL.Statistical mechanics of complex network[J].Review of Modern Physics,2002(74):47-97.
- [5] BARABASI A-L,ALBERT R.Emergence of scaling in random networks [J].Science,1999(286):509-512.
- [6] Kim H J,Lee Y,Kahng B,et al.Weighted scale-free network in financial correlations[J].J Phys Soc Jpn,2002,71(9):2133-2136.
- [7] L.Bakker,W.Hare,H.Khosravi,B.Ramadanovic.A social network model of investment behaviour in the stock market.Physica A,2010(389):1223-1229.
- [8] Chi K.Tse,Jing Liu,Francis C.M.Lau.A network perspective of the stock market.Journal of Empirical Finance,2010(17):659-667.
- [9] A Garas,P.Argyarakis,S Havlin.The structural role of weak and strong links in a financial market network[J].Eur Phys J B,2008(63):265-271.

作者简介:

- 惠宝锋(1977-),男,硕士,讲师.研究领域:信息管理与信息系统,大数据基础应用.
- 葛志远(1974-),男,博士,教授.研究领域:管理优化与决策.
- 王咏宁(1971-),男,硕士,副教授.研究领域:虚拟仪器技术及应用.