

# 基于深度学习的村镇砖(混)木房屋识别<sup>①</sup>



潘 健<sup>1</sup>, 董 翔<sup>1</sup>, 杨王永<sup>1</sup>, 娄世平<sup>1</sup>, 徐秀杰<sup>1</sup>, 王 宇<sup>2</sup>

<sup>1</sup>(山东省地震局, 济南 250014)

<sup>2</sup>(山东省国土测绘院, 济南 250100)

通讯作者: 潘 健, E-mail: panj1992@qq.com

**摘 要:** 破坏性地震发生后, 相较按照抗震设防标准建设的城市公共设施和居民住宅, 广大无抗震设防的村镇居民自建房屋, 更易发生倒塌甚至完全损毁. 以往地震灾情预评估、地震灾害风险调查、地震重点危险区调研, 依靠专家现场踏勘, 确定不同结构类型建筑物数量及所占比例. 本研究借助深度学习和倾斜摄影技术, 进行砖(混)木结构房屋识别, 郯庐断裂带山东境内砖(混)木房屋影像制作数据集, 训练得到 Faster R-CNN 模型, 该区域内砖(混)木房屋识别平均精度为 91.868%. 结果表明, 本文方法能够对砖(混)木房屋进行有效检测, 可应用于地震行业开展震前、震后各类现场工作, 提高政府部门应急管理能力和.

**关键词:** 深度学习; 倾斜摄影; 房屋识别; Faster R-CNN

引用格式: 潘健, 董翔, 杨王永, 娄世平, 徐秀杰, 王宇. 基于深度学习的村镇砖(混)木房屋识别. 计算机系统应用, 2020, 29(8): 261-265. <http://www.c-s-a.org.cn/1003-3254/7416.html>

## Recognizing Brick (Concrete) Wood Rural House Based on Deep Learning

PAN Jian<sup>1</sup>, DONG Xiang<sup>1</sup>, YANG Yu-Yong<sup>1</sup>, LOU Shi-Ping<sup>1</sup>, XU Xiu-Jie<sup>1</sup>, WANG Yu<sup>2</sup>

<sup>1</sup>(Shandong Earthquake Agency, Jinan 250014, China)

<sup>2</sup>(Shandong Institute of Land Surveying and Mapping, Jinan 250100, China)

**Abstract:** After the occurrence of destructive earthquake, compared with various urban public facilities and residential buildings built in accordance with the relatively seismic fortification standards, the vast number of villages and towns without seismic fortification houses are more likely to collapse or even completely damage. In the past, earthquake disaster risk investigation and disaster assessment relied on the field survey of experts to determine the number and proportion of buildings of different structural types. In this study, brick (concrete) wood structure houses are identified by deep learning and photography technology. The Faster R-CNN model is trained for the data set of brick (mixed) wood houses in Shandong Province of Tan Lu fault zone, with an average accuracy of 91.868%. The results show that this method can effectively detect brick (concrete) wood houses, and can be applied to earthquake disaster pre-assessment, earthquake disaster risk investigation, earthquake key-risk area investigation, and other related work.

**Key words:** deep learning; oblique photography; houses recognizing; Faster R-CNN

## 1 引言

近年来, 我国“高分”系列和各类商用遥感卫星投入使用, 无人机平台的快速普及, 各类高分辨率航空航

天遥感影像数据变得越来越容易获取. 高分辨率航空图像为图像识别提供了足够的结构和纹理信息, 同时也对现有的分割方法提出了新的挑战<sup>[1]</sup>. 近年来, 深度

① 基金项目: 山东省地震局一般科研项目 (YB1903); 山东省防震减灾社会服务能力提升工程 (SD135-3)

Foundation item: General Sci-Tech Research Program of Shandong Earthquake Agency (YB1903); Social Service Capacity Improvement Project for Earthquake Prevention and Disaster Reduction of Shandong Province (SD135-3)

收稿时间: 2019-10-19; 修改时间: 2019-11-20; 采用时间: 2019-11-22; csa 在线出版时间: 2020-07-29

学习、卷积神经网络技术作为机器学习的一个热门方向,被引入到遥感图像处理中,在遥感领域得到广泛关注.在高分辨率遥感图像分割研究中取得了显著效果<sup>[2,3]</sup>,同时也在图像预处理、基于像素的分类、场景理解、目标检测4个领域<sup>[4]</sup>,推动了遥感技术应用方法的创新.

目标检测属于遥感图像处理的一个子类,但具有特殊的难点与技术复杂度,深度学习由于其特征提取潜力可以在遥感图像目标检测中发挥重要作用.遥感图像目标由于背景复杂度高、噪声大、干扰强等的原因,其目标检测难度,较客观物体、人体行为等自然图像更大.Chen等<sup>[5]</sup>提出一种基于卷积神经网络的车辆检测算法,Zhang等<sup>[6]</sup>、Tang等<sup>[7]</sup>和YU等<sup>[8]</sup>分别实现了基于CNN、RCNN、FCN、U-NET等深度学习技术的油罐、舰船和飞机目标检测.

有学者Vakalopoulou等<sup>[9-11]</sup>在深度学习识别建筑物轮廓信息领域开展应用研究,拓展了深度学习的应用领域,但以往研究往往局限于城市地区,数据源多选择正视高分辨遥感影像.黑盒深度学习和无人机倾斜摄影技术方兴未艾,两种技术的交叉应用尚处于起步探索阶段,本文以无人机倾斜摄影图像为数据源,深度学习技术为主要技术手段,对村镇砖(混)木房屋进行识别研究.

## 2 图像提取方法

### 2.1 卷积神经网络图像提取

基于卷积神经网络CNN技术的目标检测识别、图像分类,丰富了传统监督分类、非监督分类、面向对象等遥感图像的解译方法,提高了遥感图像后期分类处理准确率<sup>[12]</sup>.Girshick等于2014年提出了基于候选框的目标检测分割算法R-CNN<sup>[13]</sup>,使用选择性搜索策略得到大量目标候选框,利用卷积网络获取所有区域特征,然后对所有区域逐一进行分类.这也造成了因候选区域重叠、冗余图像重复计算,而引发的算力浪费问题.He等<sup>[14]</sup>将空间金字塔池化层引入R-CNN,从特征图中提取特征取代了从原图获取特征,解决了R-CNN重复运算的缺点,提高了运算效率.Girshick等人在此基础上提出了Fast R-CNN,每张图像提取特征图只通过一次运算,再一次提高了算法的效能.候选区域生成速度缓慢,依然是Fast R-CNN算法的痛点.为解决此问题,Ren等<sup>[15]</sup>又提出了Faster R-CNN,增加了

区域建议网络,使得全局特征图中的目标可以在各个候选框共享,现了端到端的训练.在Faster R-CNN的基础上He等<sup>[16]</sup>增加了一个进行语义分割的分支得到了Mask R-CNN,并且将原来的ROI Pooling改为了ROI Align策略,使得Mask R-CNN可以保持Faster R-CNN快速的同时,可以完成包括目标分类、目标检测、语义分割、人体关键动作识别等多种任务,刷新了COCO数据集上的记录<sup>[16]</sup>.

### 2.2 Faster R-CNN 图像提取

本文基于Faster R-CNN模型,对村镇砖(混)木房屋进行识别提取的详细步骤如下:

(1) 将砖(混)木房屋训练集图片,导入训练网络进行训练,利用特征提取网络提取砖(混)木房屋的特征图像,得到的特征图像由所有候选区域RPN<sup>[17]</sup>网络和Fast R-CNN网络二者共享.较传统Fast R-CNN中Selective Search<sup>[18]</sup>方法实现了候选框提取,提高了图像检测精度,减少了重复训练次数,节约了计算机硬件资源.

(2) 候选区域RPN网络中的Softmax分类器,对特征图像进行二分类,确定所划分锚点(anchors)属于前景还是背景,同时利用锚点回归规则,得到候选框位置.

(3) Fast R-CNN结合特征图和候选框信息,通过多重卷积层和池化层处理,对图像中候选区域所在部分进行特征提取和学习,识别前景所属类别,对候选区域进行识别,判断其是否为砖(混)木房屋目标,随即产出检测框所在精确位置.

Fast R-CNN依据所制作图片集中训练图片的标注属性,类似机器学习中的训练样本的标签,经过迭代训练网络,网络中的参数趋近最优,增加目标识别精度.技术流程图,如图1所示.

### 2.3 RPN 网络

因RPN网络提取候选框的应用,使得Faster R-CNN克服了多任务模块串行模式,模型误差不断收敛,实现了从输入端到输出端的物体检测.在特征图传入全卷积网络RPN后,使用 $3 \times 3$ 的滑窗生成一个 $n$ 维长度的特征向量,然后将此特征向量分别传入回归层与分类层.在分类层中,使用Softmax分类器对anchors进行前景或背景的二分判断.在回归层中,通过调整锚点边框的中心坐标与长宽,拟合出候选框位置<sup>[19]</sup>.另外,在训练过程中,RPN网络需要使用损失函数分类层损失函数与回归层损失函数.

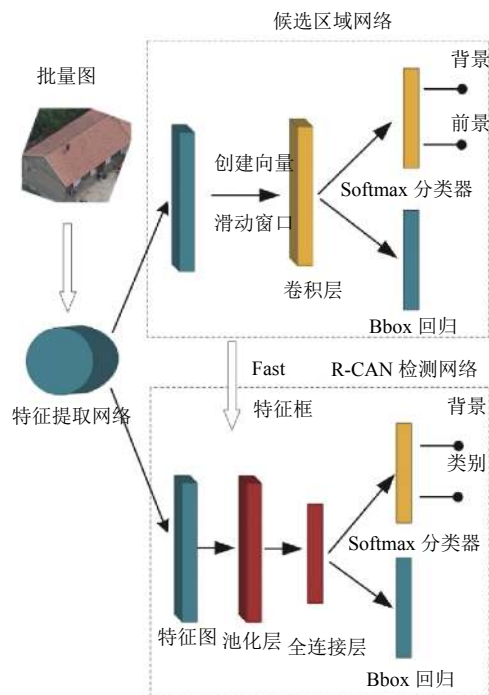


图1 Faster R-CNN 模型结构

## 2.4 特征共享模式

RPN 和 Fast R-CNN 两个网络相互独立, 单独网络进行训练很难得到最终的收敛结果, 本文采用交替训练的思路对 RPN 和 Fast R-CNN 两个网络进行训练.

首先采用 ImageNet 提供的预训练模型对 RPN 网络中的卷积层进行参数初始化, 获得砖(混)木房屋图像通用特征, 并生成候选区域框. 其次, 利用 RPN 网络输出的特征候选框对 Fast R-CNN 进行训练, 得到卷积层实时参数, 反作用于 RPN 网络, RPN 训练结束后, 仅更新 RPN 中特有网络层参数. 最后固定共享卷积层, 并对 Fast R-CNN 的全连接层进行细微调整. 通过以上操作, 将 RPN 与 Fast R-CNN 两个网络统一一致, 相同的卷积层在两个网络中得到共享.

## 3 实验数据与环境

### 3.1 原始数据获取与样本数据集

为保证样本集能充分涵盖不同建筑结构、建筑习俗的村镇房屋类型, 最大限度增加深度学习模型泛化性, 采用扩大研究区面积并随机选取的原则. 沿山东境内郯庐断裂带两侧 10~20 千米, 划定 54 行 5×5 千米正方形样本格网, 每行格网随机选取一个格网, 并在其内部随机选取一个村镇. 根据测区环境, 使用飞马 F200

(原始影像分辨率 5456×3632) 或大疆精灵 (原始影像分辨率 5472×3648) 无人机进行倾斜摄影航测. 实际航测区域, 如图 2 所示.

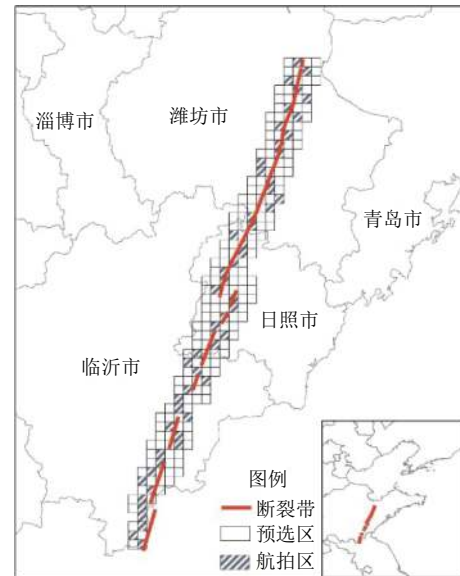


图2 倾斜摄影作业区

因航测现场存在平原、丘陵、山地多种地形, 及无人机型号、电量等客观条件限制, 原始航片地面采样间隔 GSD 不一, 但最大不超过 3 cm. 共获取 54 个村镇驻地的航空遥感影像, 原始影像集共计 732 GB.

由于无人机倾斜摄影可以从目标正视、左视、后视、后视、俯视 5 个方向进行拍摄, 故从每个视角各随机选取 1100 个样本, 组成样本集, 共计 5500 个样本, 如表 1 所示.

表1 样本数据集组成

角度	训练集	测试集
正视	1000	100
左视	1000	100
右视	1000	100
后视	1000	100
俯视	1000	100

### 3.2 实验平台

本次实验使用的处理器 CPU Intel i7-8700K, 显卡 GPU NVIDIA GTX1080Ti, 固态硬盘 512 GB, 内存 32 GB. 在开源 Caffe (Convolutional architecture for fast feature embedding) 深度学习框架<sup>[20]</sup>上, 采用 Python 作为编程语言, 实现本文砖(混)木房屋识别 Faster R-CNN 算法.



### 3.3 实验结果

本文从多个视角对砖(混)木房屋目标进行检测,房屋训练集图片的大小统一为 200×200,并且将检测出的房屋用红色框标出,为了验证该方法检测准确度,将检测框概率的阈值设置为 0.81,同时避免部分和待检测目标关联度较小的区域参与计算,图像中低于 0.81 阈值的目标不被框选.虽然 RPN 网络筛选出的候选区域数量较少,但若全部候选区域都进行分类判定,容易引起过拟合现象.本实验中,检测框上部的蓝色区域即表示房屋的概率大小,被标注出的房屋目标概率总体高于 0.88,如图 3 所示.



(a) 局部结果-1



(b) 局部结果-2

图 3 砖(混)木房屋识别结果

通过实验结果可以看出,在拍摄不同角度、光照条件、复杂背景等条件下,绝大多数砖(混)木房屋已被标记识别,但存在少量的漏检、误标问题发生.小部分土木、石木结构房屋被标记,被树木、高大建筑遮挡的房屋不能得到很好的识别.

### 3.4 实验结果精度评价

采用平均准确率  $Ap$  (Average precision) 作为砖(混)木房屋检测的评价指标,相关计算公式如下:

$$Pre = \frac{TP}{TP+FP} \times 100\%$$

$$Rec = \frac{TP}{TP+FN} \times 100\%$$

$$Ap = \int_0^1 PrecRec dRec$$

其中,  $Pre$ —精确率,  $Rec$ —召回率,  $TP$ —被正确划分为正样本的数量,  $FP$ —被错误划分为正样本的数量,  $FN$ —被错误划分为负样本的数量.  $Ap$  平均精度,平均精度代表模型识别效果,其值越大效果越好,反之越差.

精确率-召回率曲线,平均精度是对精确率-召回率曲线进行积分,曲线的横轴召回率表示分类器对正样本的覆盖能力,纵轴精确率表示分类器预测正样本的精准度,结果如表 2 所示.

表 2 精度评价结果

角度	$Ap$	$mAp$
正视	93.692	
左视	90.265	
右视	91.896	91.868
后视	92.564	
俯视	90.924	

综上所述可知,本文采用的 Faster R-CNN 在村镇砖石木房屋识别应用中取得了良好的效果,在复杂背景目标识别过程中体现了模型较强的鲁棒性.随着对测试集数据进行,人为降低目标大小、提高目标遮挡率等操作,模型识别精度随之降低,暴露了模型在弱小目标识别、强遮挡等方面的缺陷.同时,针对本文所用训练数据集、验证数据集采用 Faster R-CNN 进行对比实验,提升了 0.2% 左右准确度.

## 4 结论与展望

本文虽取得了一定的研究成果,但在此研究方法的基础上,可进一步丰富训练样本、改进算法、优化模型参数.提高倾斜摄影图像建筑物目标分类能力,实现自动识别多类建筑结构,达到产出区域性地震灾害风险报告的中长期目标.

推动深度学习技术在地震行业应用,是提升防震减灾能力的有力抓手,更是新时代防震减灾工作的内在要求.震前和震后开展的各项现场工作,离不开房屋结构数据的支持,本研究可起到减少人力成本,提供精准辅助决策数据的关键作用.

## 参考文献

- 1 Ma Y, Wu HP, Wang LZ, *et al.* Remote sensing big data computing: Challenges and opportunities. *Future Generation Computer Systems*, 2015, 51: 47–60. [doi: [10.1016/j.future.2014.10.029](https://doi.org/10.1016/j.future.2014.10.029)]
- 2 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, NV, USA. 2012. 1097–1105.
- 3 Wang HZ, Wang Y, Zhang Q, *et al.* Gated convolutional neural network for semantic segmentation in high-resolution images. *Remote Sensing*, 2017, 9(5): 446. [doi: [10.3390/rs9050446](https://doi.org/10.3390/rs9050446)]
- 4 Zhang LP, Zhang LF, Du B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 2016, 4(2): 22–40. [doi: [10.1109/MGRS.2016.2540798](https://doi.org/10.1109/MGRS.2016.2540798)]
- 5 Chen XY, Xiang SM, Liu CL, *et al.* Vehicle detection in satellite images by hybrid deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(10): 1797–1801. [doi: [10.1109/LGRS.2014.2309695](https://doi.org/10.1109/LGRS.2014.2309695)]
- 6 Zhang L, Shi ZW, Wu J. A hierarchical oil tank detector with deep surrounding features for high-resolution optical satellite imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(10): 4895–4909. [doi: [10.1109/JSTARS.2015.2467377](https://doi.org/10.1109/JSTARS.2015.2467377)]
- 7 Tang JX, Deng CW, Huang GB, *et al.* Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(3): 1174–1185. [doi: [10.1109/TGRS.2014.2335751](https://doi.org/10.1109/TGRS.2014.2335751)]
- 8 Yu YT, Guan HY, Zai DW, *et al.* Rotation-and-scale-invariant airplane detection in high-resolution satellite images based on deep-Hough-forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 112: 50–64. [doi: [10.1016/j.isprsjprs.2015.04.014](https://doi.org/10.1016/j.isprsjprs.2015.04.014)]
- 9 Vakalopoulou M, Karantzalos K, Komodakis N, *et al.* Building detection in very high resolution multispectral data with deep learning features. *Proceedings of 2015 IEEE International Geoscience and Remote Sensing Symposium*. Milan, Italy. 2015.1873–1876.
- 10 眭海刚, 刘超贤, 黄立洪, 等. 遥感技术在震后建筑物损毁检测中的应用. *武汉大学学报·信息科学版*, 2019, 44(7): 1008–1019.
- 11 范荣双, 陈洋, 徐启恒, 等. 基于深度学习的高分辨率遥感影像建筑物提取方法. *测绘学报*, 2019, 48(1): 34–41.
- 12 张兵. 遥感大数据时代与智能信息提取. *武汉大学学报·信息科学版*, 2018, 43(12): 1861–1871.
- 13 Oneata D, Revaud J, Verbeek J, *et al.* Spatio-temporal object detection proposals. *Proceedings of 13th European Conference on Computer Vision*. Zurich, Switzerland. 2014. 737–752.
- 14 He KM, Zhang XY, Ren SQ, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904–1916. [doi: [10.1109/TPAMI.2015.2389824](https://doi.org/10.1109/TPAMI.2015.2389824)]
- 15 Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. [doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)]
- 16 He KM, Gkioxari G, Dollár P, *et al.* Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 386–397. [doi: [10.1109/TPAMI.2018.2844175](https://doi.org/10.1109/TPAMI.2018.2844175)]
- 17 Felzenszwalb PF, Girshick RB, McAllester D, *et al.* Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(9): 1627–1645. [doi: [10.1109/TPAMI.2009.167](https://doi.org/10.1109/TPAMI.2009.167)]
- 18 Uijlings JRR, Van De Sande KEA, Gevers T, *et al.* Selective search for object recognition. *International Journal of Computer Vision*, 2013, 104(2): 154–171. [doi: [10.1007/s11263-013-0620-5](https://doi.org/10.1007/s11263-013-0620-5)]
- 19 孙哲, 张春龙, 葛鲁镇, 等. 基于 Faster R-CNN 的田间西兰花幼苗图像检测方法. *农业机械学报*, 2019, 50(7): 216–221. [doi: [10.6041/j.issn.1000-1298.2019.07.023](https://doi.org/10.6041/j.issn.1000-1298.2019.07.023)]
- 20 Jia YQ, Shelhamer E, Donahue J, *et al.* Caffe: Convolutional architecture for fast feature embedding. *Proceedings of the 22nd ACM International Conference on Multimedia*. New York, NY, USA. 2014. 675–678.