



文献 DOI:

10.11922/noda.2021.0013.zh

数据 DOI:

10.11922/sciencedb.00620

文献分类: 地球科学

收稿日期: 2021-02-11

开放同评: 2021-02-26

录用日期: 2021-03-25

发表日期: 2021-03-30

中国典型城市建筑物实例数据集

吴开顺¹, 郑道远², 陈妍伶², 曾林芸², 张嘉辉¹, 柴生华²,

徐文杰², 杨永亮², 李圣文^{1,2}, 刘袁缘^{1,2}, 方芳^{1,2*}

1. 中国地质大学(武汉)国家地理信息工程技术创新研究中心, 武汉 430074

2. 中国地质大学(武汉)地理与信息工程学院, 武汉 430074

摘要: 建筑物轮廓信息是国家基础地理信息的重要组成部分。高精度建筑物轮廓自动提取往往依赖于大规模的建筑物标注样本。为了丰富中国地区建筑物提取数据集, 本文研究以高分辨率遥感影像为数据源, 采用人工标注与交互式标注相结合的方式构建形成中国典型城市建筑物实例数据集。本数据集包含 7260 个影像区域样本, 共 63 886 栋建筑物, 分布在北京、上海、深圳及武汉 4 个城市。数据集由 MSCOCO 2017 格式的标注文件及相应的建筑物掩膜二值图构成, 可为研究高分辨率遥感影像的建筑物检测和提取提供基础数据。

关键词: 建筑物提取; 实例分割; 高分辨率遥感影像

数据库(集)基本信息简介

数据库(集)名称	中国典型城市建筑物实例数据集
数据作者	吴开顺, 郑道远, 陈妍伶, 曾林芸, 张嘉辉, 柴生华, 徐文杰, 杨永亮, 李圣文, 刘袁缘, 方芳
数据通信作者	方芳 (fangfang@cug.edu.cn)
数据时间范围	2017–2019年
地理区域	中国
空间分辨率	0.29 m
数据量	约5000 MB
数据格式	*.tif, *.json, *.png
数据服务系统网址	https://doi.org/10.11922/sciencedb.00620
基金项目	国家对地观测科学数据中心开放基金项目 (NODAOP2020015)
数据库(集)组成	数据集包括7260个区域样本, 由3部分信息构成: (1) *.tif存储了遥感影像瓦片; (2) *.json描述了建筑物标注, 包含训练集和测试集两个文件, 可用于实例分割任务; (3) *.png存储了建筑物区域的像素级语义标签, 可用于语义分割任务。

引言

* 论文通信作者

方芳: fangfang@cug.edu.cn

建筑物轮廓信息是最重要的基础地理信息之一, 在日常生活、经济建设和国防建设中发挥着重要的作用。遥感影像建筑物检测和提取在城市规划、人口估计、地形图制作和更新等应用中都具有极为重要的意义^[1]。相比人工遥感解译与矢量

化，自动化的遥感影像建筑物轮廓提取方法，不仅节省人力物力，而且效率高、信息提取周期短。

近年来，随着深度学习技术的不断发展，出现了基于深度学习的高精度建筑物轮廓信息自动提取方法，识别精度获得显著提升。大量的数据样本是训练深度学习模型、提升建筑物自动提取性能的关键。在计算机视觉领域，ImageNet^[2]、MS COCO^[3]等开放数据集极大地促进了深度学习方法的发展。然而，可供开放使用的建筑物提取数据集相对缺乏^[1]。目前常用的建筑物提取数据集主要有AIRS 数据集^[4]、WHU 建筑物数据集^[1]和 Inria 遥感影像数据集^[5]。此外，马萨诸塞数据集^[6]由于影像质量和分辨率较低，较少应用于建筑物提取的相关研究；ISPRS^[7]及竞赛数据集^[8]覆盖区域过小，很难反映出建筑物的多样性。需要指出的是，遥感影像中的建筑物影像在时空分布、形态、背景环境上存在较大差异，模型精度和泛化能力与训练时采用的数据集密切相关。但是目前尚未见发表中国地区的建筑物提取公开数据集，这在一定程度上制约着适应中国地区建筑物自动提取方法及应用的研究。

笔者基于高分辨率遥感影像构建了中国地区面向城市建筑物检测和提取的建筑物实例分割数据集。本数据集覆盖了国内 4 个具有代表性的城市，反映国内城市建筑物自身和背景环境的特点，为中国基础数据的构建提供数据支撑，以期推进相关学术的深入研究。

1 数据采集和处理方法

本数据集选取北京、上海、深圳、武汉等 4 个具有代表性的城市中心城区作为数据采集目标区域，具体位置如表 1 所示。原始数据源自由谷歌提供的 19 级卫星影像，地面分辨率为 0.29 m。为了提高数据集的通用性，数据区域的选取包括正射影像和非正射影像区域、建筑物稀疏分布和密集分布区域，同时考虑建筑物轮廓形状多样性等因素。数据集样本覆盖区域共计约 120 平方公里。

表 1 数据采集目标区域信息

城市名	数据所在经纬度范围	样本区域面积 (km ²)	区域建筑物数 (栋)
北京	116.1605°–116.4111°E, 39.8701°–40.0974°N	25.48	15205
上海	121.3288°–121.6262°E, 30.8819°–31.5582°N	20.85	13487
深圳	113.7552°–114.0738°E, 22.5432°–22.7815°N	42.66	27018
武汉	114.1418°–114.3718°E, 30.4768°–30.6800°N	35.33	8176

数据处理及标注的流程如图 1 所示，处理流程主要包含数据预处理阶段、人工标注阶段和人机交互标注阶段。

首先，在数据预处理阶段，针对 4 个城市的原始卫星影像，人工筛选并截取无重叠的 82 块 5000 × 5000 像素大小的影像区域。同时，参考现有标准实例分割数据集格式，将每幅影像统一切割为 100 幅 500 × 500 像素大小的瓦片。本数据集从以上瓦片数据中随机选取 7260 个区域作为样本进行标注，其中北京 2237 个，深圳 2344 个，上海 1231 个，武汉 1448 个。

在人工标注阶段，首先采用数据标注软件 (labelme^[9]) 对 20% 的遥感影像数据建筑物轮廓进行标注，形成初始样本集；随后，基于初始样本集训练神经网络模型，使模型具有初步处理的能力。

在交互式标注阶段，借鉴反向传播修正机制 (Backpropagating Refinement Scheme, BRS)^[10]对其余 80% 的数据进行交互式数据标注。最终获得全部数据的建筑物实例标签，完成数据集的制作。

2 数据样本描述

本数据集以“遥感影像+数据标注文件”的形式组织存储，分别存储原始遥感影像和所对应的建筑物轮廓标注信息。数据集样本包括训练集和测试集两个文件夹，每个文件夹包含一个遥感影像数

据文件夹和一个 json 格式标注文件。数据集的文件组织形式如表 2 所示。

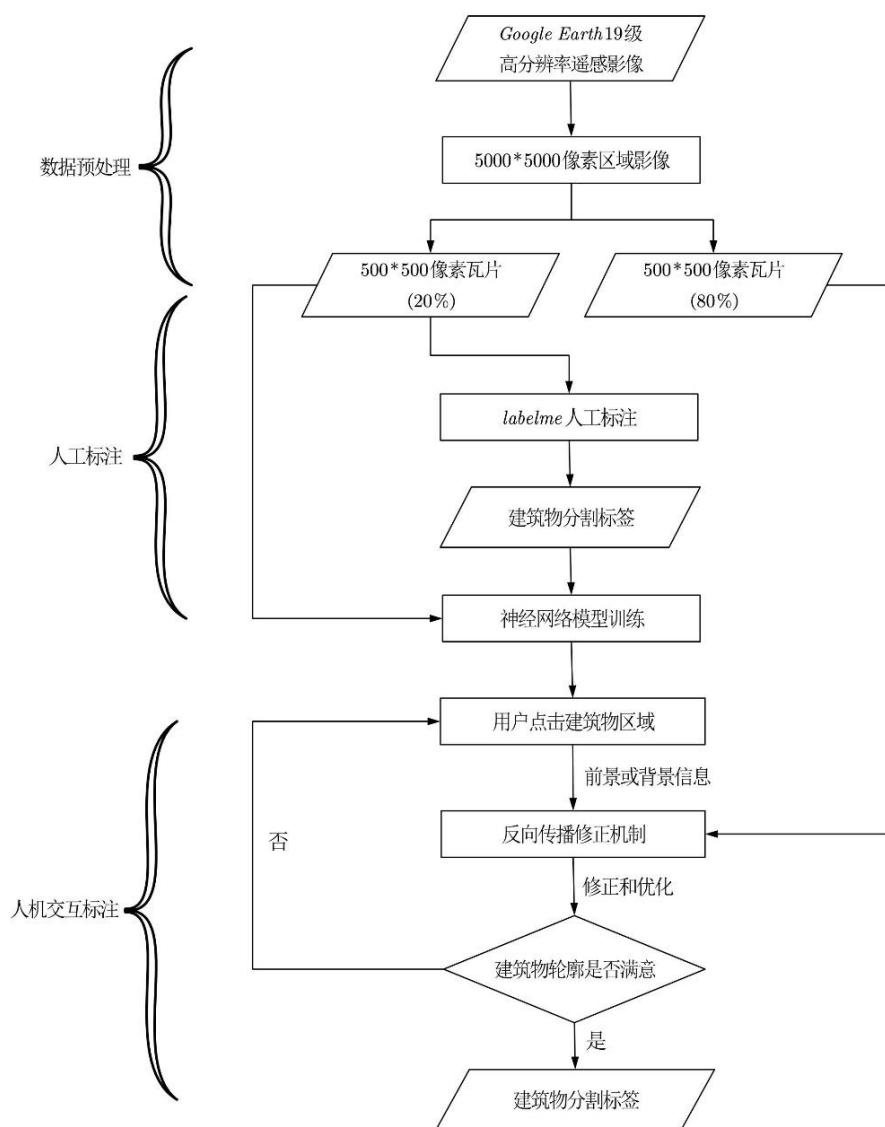


图 1 数据处理流程图

表 2 数据集文件组织形式

文件夹	文件内容	文件说明
train	Images/*.tif	训练集影像目录
	PNG/*.png	训练集像素标签
	Annotations.json	训练集标注文件
test	Images/*.tif	测试集影像目录
	PNG/*.png	测试集像素标签
	Annotations.json	测试集标注文件

其中, tif 格式文件为 500×500 像素的包含位置信息的原始遥感影像; json 格式文件则描述了该遥感影像所对应的建筑物轮廓标注信息, 主要包括 4 个方面的信息: info、images、categories 和 annotations。其中, info 记录的是数据集的制作年份、版本、描述等基本信息; images 记录了数据集

中影像的大小等信息；categories 将类别从名称映射到类别编号；annotations 则记录了标注编号、标注对应的影像编号、建筑物轮廓点坐标等标注信息。数据标注文件的组织形式如图 2 所示。

```
{
  "info": { #数据集相关信息
    {
      "description": "Chinese city building instance segmentation dataset", #数据集描述
      "version": "1.0.0", #数据集相关信息
      "year": 2020, #数据集年份
      "contributor": "Fang Fang, Kaishun Wu, Daoyuan Zheng etc.", #数据集贡献者
      "date_created": "2020-10-18 10:39:30.199994". #数据集创建时间
    }
  },
  "images": [ #数据集包含图片信息
    {
      "id": 0, #图片编号
      "file_name": "1-3-21.tif", #图片名称
      "width": 500, #图片宽度
      "height": 500, #图片高度
      "date_captured": "2020-10-13 18:36:15.880734", #图片获取时间
    }
  ],
  "categories": [ #类别信息
    {
      "id": 1, #类别编号
      "name": "building", #细类别名称
      "supercategory": "building" #粗类别名称
    }
  ],
  "annotations": [ #标注信息
    {
      "id": 0, #标注编号
      "image_id": 0, #标注对应图片编号
      "category_id": 1, #标注对应类别编号
      "iscrowd": 0, #标注是否多连通域
      "bbox": [ #标注对应目标框范围
        475.0, 102.0, 26.0, 29.5
      ],
      "area": 572, #轮廓面积
      "segmentation": [ #建筑物轮廓点坐标
        [
          498.0, 131.5, 490.0, 129.5, 484.0, 121.5, 474.5, 117.0, 476.0, 102.5, 499.5, 102.0, 500.5, 127.0, 498.0,
        ]
      ]
    }
  ]
}
```

图 2 数据标注文件的组织形式

图 3 展示了在 4 个城市中选取的具有代表性的遥感影像及相应的建筑物轮廓。

3 数据质量控制和评估

为保证数据集质量，我们在影像整理和检查、人工标注和交互式标注阶段均采用了完整的质量控制过程（如图 4 所示），通过多重检查保证标注数据的可靠性、完整性和统一性。

在影像整理和检查阶段，人工剔除了失真、畸变、模糊等影像，以减少原始影像引入的噪声。在人工标注阶段采用人工交叉检验方法检查标注结果并修正发现的问题，检查内容包括标注轮廓不

完整、建筑物标注遗漏以及非建筑物误标注为建筑物等。交互式标注阶段则重点检查标签文件的准确性、一致性，确保标注质量。为避免影像标签缺失、标签与影像匹配错误等问题，我们采用循环遍历算法进行检验，并对错误数据逐一确认和修改。



(a)北京 (b)上海 (c)深圳 (d)武汉

图3 数据集样例展示

此外，在制作最终的标注文件之前，为了确保数据组织按照既定的规则进行，所有的遥感影像及其所对应的标注信息均通过程序自动从原始数据中读取得到，并按照规则批量化自动命名，最后采取人工方式进行复核。通过以上步骤，本数据集的质量能够得到良好的保证。

4 数据价值

本数据集是首个公开的取材于国内城市的高分辨率遥感影像城市建筑物轮廓的实例分割数据集，从一定程度上反映了国内的建筑物特征和遥感影像的关系，丰富了遥感领域建筑物数据集，有望更好地支撑国内遥感影像建筑物提取的研究和生产。本数据集的组织 and 标注方式与常用实例分割数据集一致，可以很方便地服务于影像分割任务。

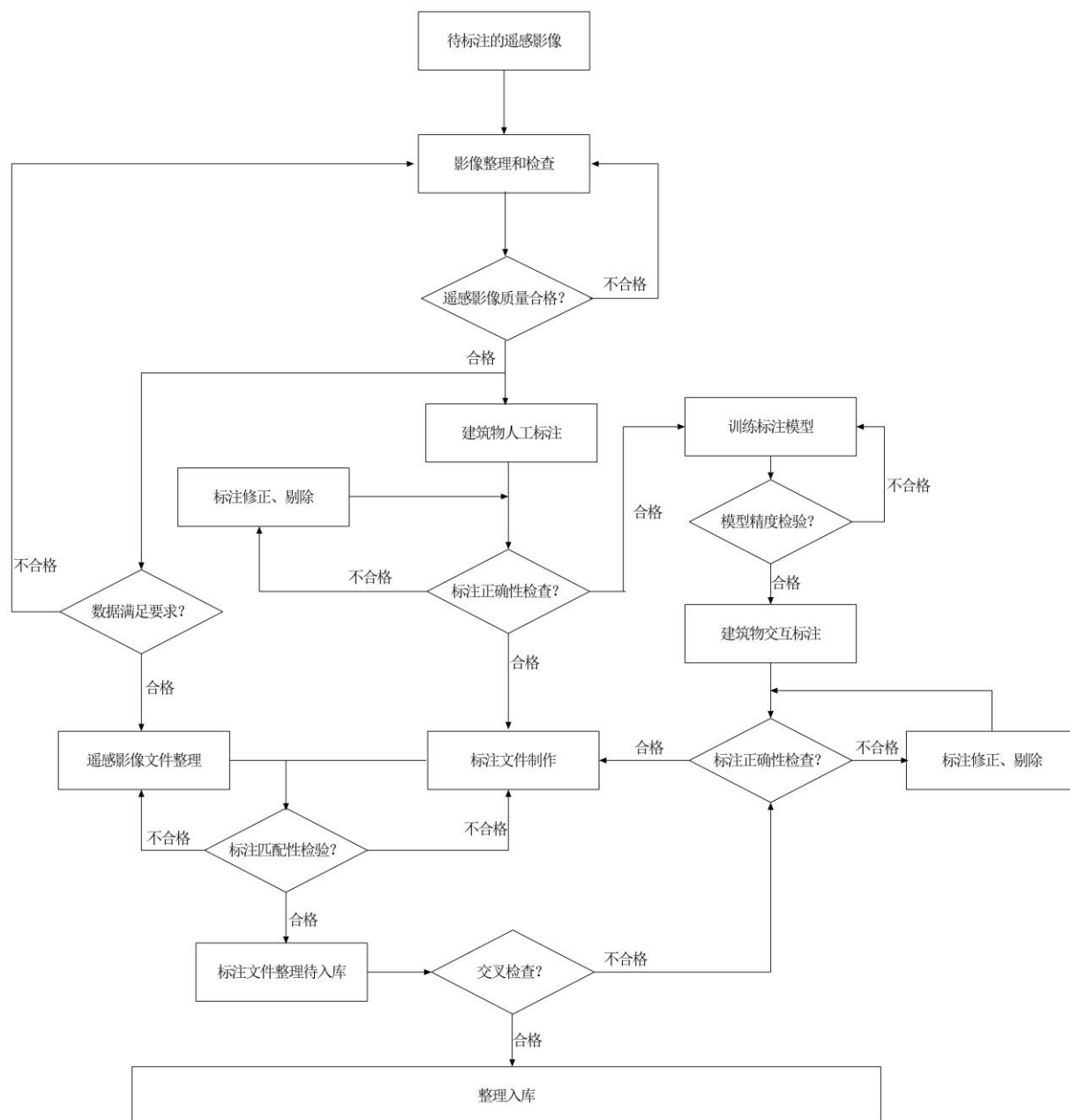


图4 数据集质量控制过程

5 数据使用方法和建议

本数据集可用于实例分割任务和基于像素级别的语义分割任务。数据集的使用方法与常用于上述两种分割任务的标准数据集使用方法一致，总体步骤如图5所示。

对于实例分割任务，本数据集采用的标注格式与实例分割数据集 MS COCO 相同，使用方式相同。实例分割任务可通过解析 json 文件，从 annotations 字段中批量读取建筑物轮廓标注；从 images 字段中获取影像存储路径，进而获取实例标注所对应的建筑物影像。实例分割模型可将获取的标注和影像两类信息作为输入，进行模型的训练和预测。

语义分割的类别标签是像素级的，其中建筑物区域对应的像素值为 1，非建筑物区域对应的像素值为 0。每个样本的标签存储在一个 png 文件中，其原始影像文件存储在与标签相同文件名的 tif 文

件中。语义分割任务依次读取每个样本的标注和影像文件，输入到语义分割模型中进行训练和预测。



图5 数据使用流程

后续过程需根据实际任务进行设计，主要包括模型训练、预测和精度验证等。两种任务使用中的模型也在不断发展，现有经典模型如 Mask R-CNN^[11]，DeepLabv3^[12]等，此文不再赘述。

数据作者分工职责

吴开顺（1995—），男，四川省彭州市人，在读硕士研究生，研究方向为图像智能处理及遥感应用。主要承担工作：数据集整体结构设计，交互式标注算法研究与实现。

郑道远（1999—），男，湖北省荆州市人，在读硕士研究生，研究方向为目标检测、图像分割及遥感应用。主要承担工作：数据集人工及交互式标注，论文初稿撰写。

陈妍伶（1997—），女，四川省射洪市人，在读硕士研究生，研究方向为大数据城市意向分析。主要承担工作：数据集人工及交互式标注。

曾林芸（1996—），女，四川省成都市人，在读硕士研究生，研究方向为图卷积神经网络及应用。主要承担工作：数据集人工及交互式标注。

张嘉辉（1998—），女，山东省潍坊市人，在读硕士研究生，研究方向为多模态图像数据融合及应用。主要承担工作：数据集人工及交互式标注。

柴生华（2000—），男，河北省承德市人，在读本科生。主要承担工作：数据集人工标注。

徐文杰（2000—），男，湖北省武汉市人，在读本科生。主要承担工作：数据集人工标注。

杨永亮（1999—），男，甘肃省武威市人，在读本科生。主要承担工作：数据集人工标注。

李圣文（1978—），男，山东省济宁市人，博士，副教授，研究方向为时空大数据挖掘与机器学习。主要承担工作：论文方向指导与质量把关。

刘袁缘（1984—），女，江西省景德镇人，博士，副教授，研究方向为计算机视觉。主要承担工作：算法指导。

方芳（1976—）女，湖北省黄冈市人，博士，副教授，研究方向为智能信息处理。主要承担工作：项目规划与论文质量把关。

参考文献

- [1] JI S P, WEI S Q, LU M. Fully Convolutional Networks for Multisource Building Extraction from an Open Aerial and Satellite Imagery Data Set[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 57(1): 574-586.
- [2] DENG J, DONG W, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009.
- [3] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common Objects in Context[C]. Proceedings of the European Conference on Computer Vision, Springer, Cham, 2014.
- [4] CHEN Q, WANG L, WU Y, et al. Aerial imagery for roof segmentation: A large-scale dataset towards automatic mapping of buildings[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 147: 42-55.
- [5] MAGGIORI E, TARABALKA Y, CHARPIAT G, et al. Can semantic labeling methods generalize to any city? The INRIA aerial image labeling benchmark[C]. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 2017.
- [6] MNIH V. Machine learning for aerial image labeling[M]. Canada: University of Toronto, 2013.
- [7] GERKE M, ROTTENSTEINER F, WEGNER J D, et al. ISPRS Semantic Labeling Contest[EB/OL].(2014) [2021-02-03]. <http://www2.isprs.org/semantic-labeling.html>.
- [8] LE S B, YOKOYAN, HANSCH R, et al. 2018 IEEE GRSS Data Fusion Contest: Multimodal Land Use Classification [Technical Committees][J]. IEEE Geoscience and Remote Sensing Magazine, 2018, 6(1): 52-54.
- [9] RUSSELL B C, TORRALBA A, MURPHY K P, et al. A Database and Web-Based Tool for Image Annotation[J]. International Journal of Computer Vision, 2008, 77(1-3):157-173.
- [10] KONSTANTIN S, ILIA P, OLGA B, et al. F-BRS: Rethinking Backpropagating Refinement for Interactive Segmentation[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020.
- [11] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017.
- [12] Liang-Chieh Chen, Yukun Zhu, George Papandreou, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation[C]. Proceedings of the European Conference on Computer Vision(ECCV), Munich, Germany, 2018.

论文引用格式

方芳, 吴开顺, 郑道远, 等. 中国典型城市建筑物实例数据集[J/OL]. 中国科学数据, 2021, 6(1). (2021-03-19). DOI: 10.11922/noda.2021.0013.zh.

数据引用格式

方芳, 吴开顺, 郑道远, 等. 中国典型城市建筑物实例数据集[DB/OL]. Science Data Bank, 2021. (2021-03-19). DOI: 10.11922/sciencedb.00620.

A dataset of building instances of typical cities in China

WU Kaishun¹, ZHENG Daoyuan², CHEN Yanling², ZENG Linyun²,
ZHANG Jiahui¹, CHAI Shenghua², WU Wenjie², YANG Yongliang²,
LI Shengwen^{1,2}, LIU YuanYuan^{1,2}, FANG Fang^{1,2*}

1. National Engineering Research Center of Geographic Information System, China University of Geosciences, Wuhan 430074, P.R.China

2. School of Geographic and Information Engineering, China University of Geosciences, Wuhan 430074, P.R. China

*Email: fangfang@cug.edu.cn

Abstract: Building contour information is an important part of the national basic geographic information. The performance of building automatic extraction is usually driven by a large number of training samples. To enrich the building extraction datasets of cities in China, we compiled a building instance dataset sourced from high-resolution remote sensing images through the combination of manual annotation and interactive annotation. This dataset consists of the samples in 7,260 regions, with 63,886 building instances in four China's cities: Beijing, Shanghai, Shenzhen and Wuhan. The annotations of the dataset consist of MS COCO 2017 format files and the corresponding building mask binary maps. This dataset provides fundamental data for the research on building detection of and extraction from high-resolution remote sensing images.

Keywords: building extraction; instance segmentation; high-resolution remote sensing images

Dataset Profile

Title	A dataset of building instances of typical cities in China
Data corresponding author	FANG Fang (fangfang@cug.edu.cn)
Data authors	WU Kaishun, ZHENG Daoyuan, CHEN Yanling, ZENG Linyun, ZHAN Jiahuig, CHAI Shenghua, XU Wenjie, YANG Yongliang, LI Shengwen, LIU YuanYuan, FANG Fang
Time range	2017–2019
Geographical scope	China
Spatial resolution	0.29 m
Data volume	about 5,000 MB
Data format	*.tif, *.json, *.png
Data service system	< https://doi.org/10.11922/sciencedb.00620 >
Source of funding	Open Research Fund of National Earth Observation Data Center (NODAOP2020015)
Dataset composition	The dataset consists of samples in 7,260 regions, with three types of files: (1) *.tif files, storing the information about high-resolution remote sensing images; (2) *.json files, used for instance segmentation tasks, and describing building annotation data, including the training set and the test set; (3) *.png files, the pixel-level semantic label of building area, used for semantic classification tasks.