

SARBuD 1.0: 面向深度学习的GF-3 精细模式SAR建筑数据集

吴樊¹, 张红¹, 王超¹, 李璐^{1,2}, 李娟娟^{1,2}, 陈卫荣³, 张波¹

1. 中国科学院空天信息创新研究院 中国科学院数字地球重点实验室, 北京 100094;

2. 中国科学院大学, 北京 100049;

3. 中国资源卫星应用中心, 北京 100094

摘要: 合成孔径雷达SAR (Synthetic Aperture Radar) 是开展城市建筑区信息获取与动态监测的重要数据源。本文建立了一个面向深度学习建筑区提取的中高分辨率SAR建筑区数据集SARBuD1.0 (SAR BUilding Dataset)。该数据集包含了覆盖中国不同区域的27景高分三号(GF-3)精细模式SAR图像, 并从中获取了建筑区共计60000个SAR样本数据, 结合光学图像与专家解译, 制作了与样本数据对应的标签图像。SARBuD1.0数据集包含了不同地形场景类型、不同分布类型、不同区域的建筑区。该数据集可支持研究者对建筑区进行图像特征分析、辅助图像理解, 并可对当前热点深度学习方法提供训练、测试数据支持。本文以山区建筑为例, 使用传统纹理特征与深度学习特征对建筑区进行了特征分析与比较, 相比于传统的人工设计的纹理特征, 卷积神经网络具有更深、更多的特征, 利用网络模型浅层的不同卷积核采样可得到各种纹理特征, 在网络的深层卷积结构中可获取代表着类别的深层语义特征, 使得分类器能更好地检测并提取图像中指定的目标。基于本数据集利用深度学习方法对不同地形区域的建筑区进行提取实验。实验结果表明基于本数据集训练的深度学习模型, 对建筑区提取可以取得良好的结果, 说明该数据集可以很好支持面向大数据的深度学习方法。其他学者可以基于SARBuD1.0数据集开展建筑区图像特征分析与语义分割提取等方面的研究。

关键词: 遥感, 合成孔径雷达, 建筑, 数据集, 深度学习, 高分三号, 语义分割

引用格式: 吴樊, 张红, 王超, 李璐, 李娟娟, 陈卫荣, 张波. 2022.SARBuD 1.0: 面向深度学习的GF-3 精细模式SAR建筑数据集. 遥感学报, 26(4): 620-631

Wu F, Zhang H, Wang C, Li L, Li J J, Chen W R and Zhang B. 2022. SARBuD1.0: A SAR Building Dataset Based on GF-3 FSII Imageries for Built-up Area Extraction with Deep Learning Method. National Remote Sensing Bulletin, 26(4): 620-631[DOI:10.11834/jrs.20220296]

1 引言

合成孔径雷达SAR (Synthetic Aperture Radar) 是对地观测的重要手段。近年中国SAR技术飞速发展, 以高分三号(GF-3)及HJ-1C为代表的SAR卫星成功发射并运行, 开启了中国SAR卫星民用、商用时代。有别于光学图像, SAR传感器的主动斜距成像机制使其图像在几何和辐射等方面都具有独特的特征(张红等, 2009)。比如, SAR图像中除了由于平台稳定性等原因产生的几

何畸变, SAR图像中还存在由于斜距成像导致的透视收缩、有一定高度的地物易产生叠掩现象等。此外地物随方位不同而易导致在SAR图像中几何形状与散射强度的变化, 也加大了SAR图像解译和信息提取的难度。

SAR图像是城市遥感解译的重要数据源, 获取城镇建筑信息对国土资源监测、城市规划、灾害应急评估等应用研究都有重要意义。在SAR图像中, 建筑物本身存在的二面角、三面角结构以及建筑物之间存在微波多路径反射等因素, 使得

收稿日期: 2020-07-20; 预印本: 2021-06-16

基金项目: 国家自然科学基金(编号: 41930110, 41971395)

第一作者简介: 吴樊, 研究方向为SAR图像处理与信息提取方法。E-mail: wufan@aircas.ac.cn

通信作者简介: 张红, 研究方向为SAR图像处理、极化、干涉处理理论与应用。E-mail: zhanghong200307@aircas.ac.cn

关联数据获取: <https://github.com/CAESAR-Radi/SARbVd>

建筑物/区通常表现为强散射体/区域,并在2维平面上具有一定的排列结构。依据此图像特点,传统建筑提取方法通常遵循“特征提取—分类/分割”的流程,多从散射强度(Sun等,2019;Chini等,2018)、纹理特征(Ban等,2015;Esch等,2017;Cao等,2018;Tavares等,2019)、阴影叠掩等成像几何特征(Li等,2020;Adelipour和Ghassemian,2018)、相干特征(Thiele等,2007)等出发进行提取。在进行分类/分割前通常需要经过特征分析、特征选择或特征降维,获取建筑区最优特征表达。这一过程多依赖于研究人员的经验总结,但研究人员通常不能观察到待分类对象的所有样本,无法对样本的高维和低维特征进行理解,所设计出来的特征可能无法表达复杂的映射。因而要实现快速、准确大范围范围的建筑提取仍面临着很多挑战,如在地理环境复杂的盆地、丘陵等区域,森林植被和山地岩石经常与城镇、村落混在一起,具有较高的后向散射值和类似的局部纹理,易产生虚警,且易受建筑区分布面积、建筑目标结构多样化等因素的影响。

近年基于卷积神经网络的深度学习算法在图像多层次特征学习、特征提取方面取得了良好的成果,已成为地物分类的研究热点(Shahzad等,2019;Zhang等,2019)。深度网络模型通过卷积、池化层等结构与不同的激活函数组合,可实现对图像的多尺度特征编码、感兴趣区域自动获取等能力,为解决复杂背景下目标检测、分类等研究提供了技术支持。深度学习方法以数据为驱动,样本训练数据集是深度学习模型参数优化的关键和基础。从公开发表的文献看,基于SAR图像的建筑类样本数据集较少,公开报道的有OpenSARUrban(Zhao等,2020)。OpenSARUrban数据集从基本覆盖中国21个城市的19景IW(Interferometric Wide)模式Sentinel-1的数据中获取了3万多个100像素×100像素的城市样本切片。该样本集主要集中于城市区域,包括5大类(居民区、商业区、工业区、交通枢纽及其他区域)十小类城市设施。该数据集主要是面向场景分类的切片集,未提供面向分割的像素级标签。数据集的发布为城市目标的特征分析、城市区场景分类及基于内容的图像检索等应用提供了良好的数据支持。

近年随着各国全球化、城市可持续化发展的推进,面向大范围范围的居民区获取需求不断增加。但在大范围范围内,建筑物所处的场景不同

而且多变,以数据为依赖的深度学习模型,在模型训练阶段需要对不同情况下的样本进行训练,以使模型能适应不同场景情况下的建筑信息获取。比如不仅仅需考虑城市建筑区,对建筑区的分布类型,建筑区周边地形场景等因素也应该考虑。针对这一目的本文基于GF-3精细模式SAR数据,构建了一个面向大范围多种地形场景的SAR建筑区公开数据集。该数据集采用27景SAR图像,并从中获取了建筑区样本数据,结合光学图像与专家解译,制作了与样本数据相应的像素级标签图像。该数据集包含了不同分布类型、不同区域的建筑;包含了平原区、山地、高原等地形场景。数据集旨在促进SAR建筑区特性分析及大范围、大范围建筑物语义分割提取方面的研究,并推广国产SAR数据的深入应用。

2 建筑SAR数据样本集

2.1 数据源概况

SARBuD1.0(SAR BUilding Dataset)数据集采用GF-3 SAR图像为数据源。GF-3是中国首颗C波段多极化高分辨率SAR卫星,填补了中国民用自主高分辨率多极化SAR遥感数据空白(张庆君,2017)。为了保证数据集的多样性和尽可能的充分性,在中国陆地区范围内的20个省、市、自治区选取不同地形(云南、贵州高原区;京津冀、上海、江苏平原区;重庆、湖南、江西、浙江等山地丘陵区)、不同建筑分布类型(北京、上海等密集城区;江浙水乡条状分布;河北等聚集村落块状分布;重庆、云南等山地零散分布)、不同区域(黑龙江、吉林等东北地区;京津冀等华北区;上海、江苏、浙江等华东区;广东、海南、台湾等华南区;云南、贵州等西南区;新疆维吾尔自治区)的27景GF-3 SAR图像。考虑到面向大范围、多种地形场景建筑区制图的需求,采用目前订购和存档数据最多的精细条带FSII(Fine Stripmap II)模式的数据作为数据源。FSII模式数据标称分辨率为10 m,成像幅宽为100 km(Sun等,2019),其分辨率和幅宽对于大范围建筑区制图可以保证较好的平衡。基于现有数据条件及目标,本数据集以精细条带模式中的单极化图像为主,目前未加入全极化SAR数据。全极化SAR可以通过极化分解等方法基于散射机制提取建筑物,是地物散射机制分析与分类的重要数据源,在后

期数据集的完善过程中将会加入。所选 SAR 图像成像时间主要为 2019 年, 以 2018 年及 2017 年的图像作为补充。表 1 给出了各景 SAR 图像的基本参数信息。

表 1 样本数据集原始 SAR 图像基本信息

Table 1 Information of the SAR images

图像所属区域	成像日期	像素间隔/m	入射角/(°)
安徽省 1	2019-02	2.24×4.77	31.47—38.36
安徽省 2	2019-03	2.24×4.77	31.41—38.32
湖南省	2019-09	2.24×4.76	31.17—38.15
新疆维吾尔自治区	2019-04	2.24×4.76	30.92—37.95
重庆市 1	2019-03	2.24×4.76	31.22—38.18
重庆市 2	2019-01	2.24×4.76	31.33—38.27
重庆市 3	2019-04	2.24×4.77	31.31—38.25
甘肃省	2019-01	2.24×4.76	31.38—38.30
贵州省	2019-03	2.24×4.77	31.29—38.24
黑龙江省	2019-02	2.24×4.76	31.14—38.10
江苏省 1	2019-01	2.24×4.76	31.17—38.14
江苏省 2	2019-05	1.12×5.85	24.18—32.00
吉林省 1	2019-07	2.24×4.79	31.03—38.04
吉林省 2	2019-04	2.24×4.77	31.22—38.15
上海市	2019-05	2.24×4.77	31.47—38.36
京津冀 1	2019-06	2.24×5.86	37.25—43.46
京津冀 2	2019-06	2.24×5.70	46.37—51.21
京津冀 3	2019-07	2.24×5.85	37.13—43.36
京津冀 4	2019-06	2.24×4.80	42.31—47.42
京津冀 5	2019-06	2.24×4.80	42.29—47.40
浙江省	2018-11	2.24×4.77	31.39—38.30
海南省	2018-12	2.24×4.77	31.28—38.24
台湾省	2017-06	2.25×4.82	42.56—47.54
内蒙古自治区	2019-03	2.24×4.76	31.40—38.31
广东省	2019-04	2.24×4.77	31.49—38.38
云南省	2019-07	2.24×4.77	37.04—43.32
西藏自治区	2019-02	2.24×4.76	31.37—38.21

2.2 数据集

2.2.1 数据预处理

数据预处理包括原始 SAR 图像的预处理和图像像素级标签的制作。

对原始 SAR 图像进行预处理, 包括幅度图生成、辐射定标、地理编码等基本处理。所获取的 GF-3 数据为 SLC (Single Look Complex) 数据, 为了使不同地区的建筑区之间具有可比性, 对图像进行辐射标定, 将图像 DN (Digital Number) 值转换为 σ_{db}^0 值 (CRESDA, 2016)。

$$\sigma_{\text{db}}^0 = 10 \log_{10} \left(P^I \times \left(\frac{V}{32767} \right)^2 \right) - K_{\text{db}} \quad (1)$$

式中, $P^I = I^2 + Q^2$, I 为 1A 级产品实部, Q 为 1A 级产品虚部, V 为该图像量化前的最大值, 可通过元数据文件 QualifyValue 字段获取, 元数据文件中字段 CalibrationConst 的值对应 K_{db} 。目前绝大多数卷积神经网络语义分割方法都是基于 8 位的图像, 遥感领域中神经网络的应用也通常将遥感图像转换为 8 位图像进行处理。采用线性 2% 拉伸方式对经过辐射标定后的图像处理成 8 位图像。几何纠正根据 GF-3 数据产品 RPC (Rational Polynomial Coefficient) 参数文件中的参数, 构建 RPC 模型对图像进行纠正。在实际操作中也可基于 ENVI/SARscape 进行处理, 将 GF-3 数据导入后, 利用 “Geocoding and Radiometric Calibration” 工具进行处理得到经过辐射标定和几何纠正后的图像。

结合与 SAR 图像成像时间接近的 Google Earth、Sentinel-2 等光学图像进行人工目视解译制作像素级标签二值图。样本制作过程中, 经过图像的辐射定标后, 发现少部分沿海地区 SAR 图像强度图整体灰度强度偏高, 建筑区的纹理、结构等特征易与非建筑地物混淆, 利用深度学习方法提取很难获得较高的检测精度。因此本文基于对抗学习的 CycleGAN 算法 (Zhu 等, 2017) 对沿海地区整体强度偏高的 SAR 图像进行增强处理。增强效果示例如图 1 所示。图 1 (a) 为 Google Earth 高分辨率光学图像, 图中包括了大面积的建筑区, 在图像的中下部有一块覆盖植被的山地。图 1 (b) 为增强前的 SAR 图像, 图 1 (c) 为增强后的 SAR 图像。从结果图中可以看出, 经过增强后提高了近海区域的 SAR 图像建筑区与其它地物 (如: 山地植被区) 的对比度。

2.2.2 样本切片制作

图 2 为样本集制作的具体流程。将预处理后的 SAR 图像及标注的二值标签图进行裁剪, 裁剪尺寸为 256 像素×256 像素。为了避免裁剪过程中部分建筑区被切割成小的子区域, 确保每个完整的建筑区都能有对应的切片, 切片与切片之间保证 30 个像素值宽度的重叠区。完成切片裁剪后, 对切片中包含建筑区范围很少的非完整切片进行剔除。根据相应的二值标注图中是否包含有建筑区, 将样本切片区分为正样本 (或称为有效标记样本) 或负样本, 然后放在不同的子集中, 构建用于深度学习网络模型训练的样本集。

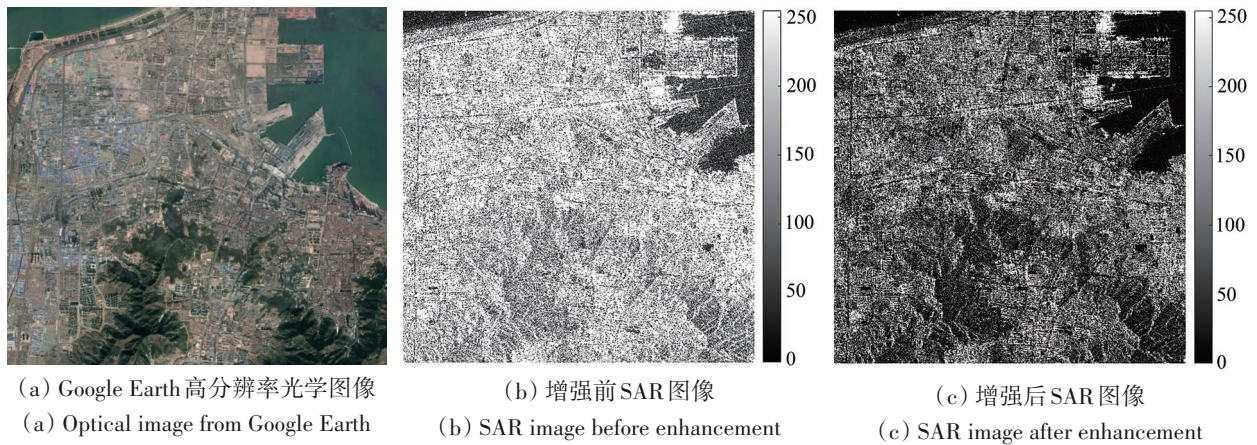


图1 SAR图像增强结果示例

Fig. 1 An example of SAR image enhancement

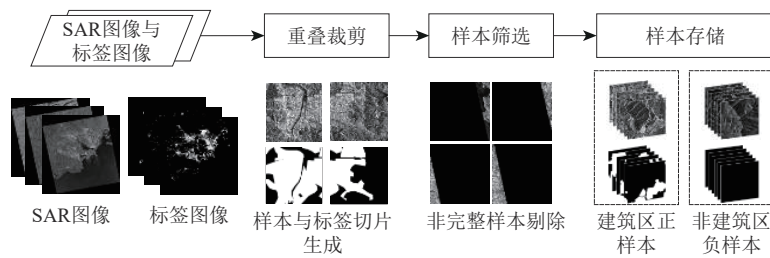


图2 样本制作流程

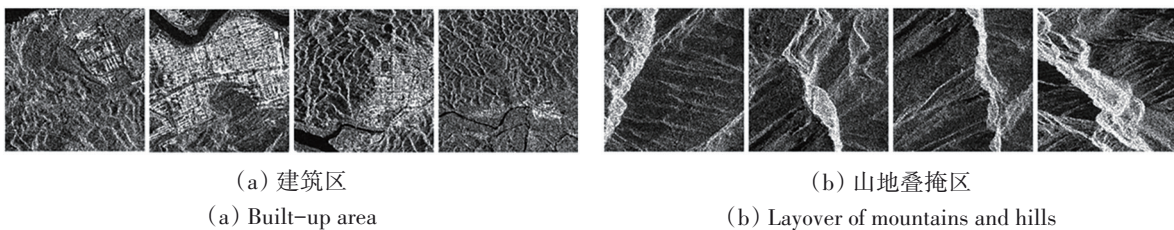
Fig.2 Workflow of compiling the SARBuD1.0 dataset

在样本组合方面，经实验发现，地势比较平坦的地区（如平原），仅使用正样本作为训练集就能达到很好的建筑区提取精度。这是因为在平原区，建筑区通常为强散射区，与周边环境存在相对差异性，仅以正样本作为训练集的训练模型可以很好区分建筑区和非建筑区。在山地丘陵区，建筑区多集中在省会城市周边，山区密集城区的建筑区切片包含的山区地形等信息较少。但在山地区，山地的叠掩区易与山区散落的建筑群混淆，容易产生虚警。因此，在山地丘陵区的建筑区提取中，训练样本集中需要增加一定比例的负样本切片，通过正负样本的模型训练可以使网络更好的将建筑区与山区区分开。实验研究发现山区样

本子集中加入 1.5 倍于正样本的负样本切片（即正、负样本比例为 1 : 1.5），可以有效弥补正样本中山区信息的不足。从目前的数据看，其他容易与建筑区混淆的负样本还包括与 SAR 图像中特征类似的结冰河流或湖泊（图 3（c））、部分农田植被区（图 3（d））以及戈壁沙丘区（图 3（e））。图 3 给出了建筑区及各类负样本的示例。

2.2.3 样本切片示例

图 4 展示了不同建筑区样本切片示例。从图 4 中可以看出，中国建筑分布类型比较多样化，建立多类型样本数据集有利于深度学习网络更好学习不同建筑分布类型的几何结构和纹理特征。



(a) 建筑区
(a) Built-up area

(b) 山地叠掩区
(b) Layover of mountains and hills

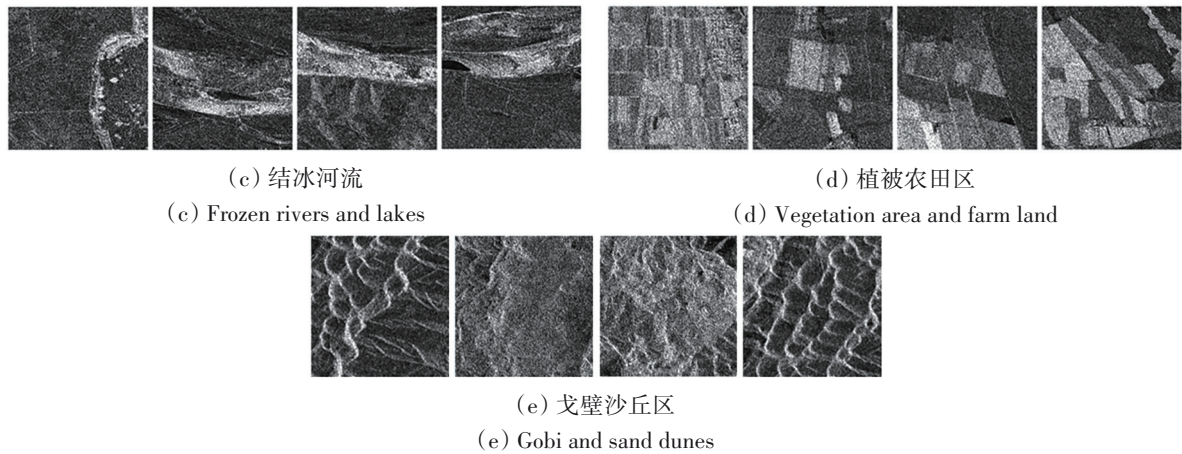


图3 建筑区与各类负样本示例

Fig. 3 Examples of built-up area and negative samples

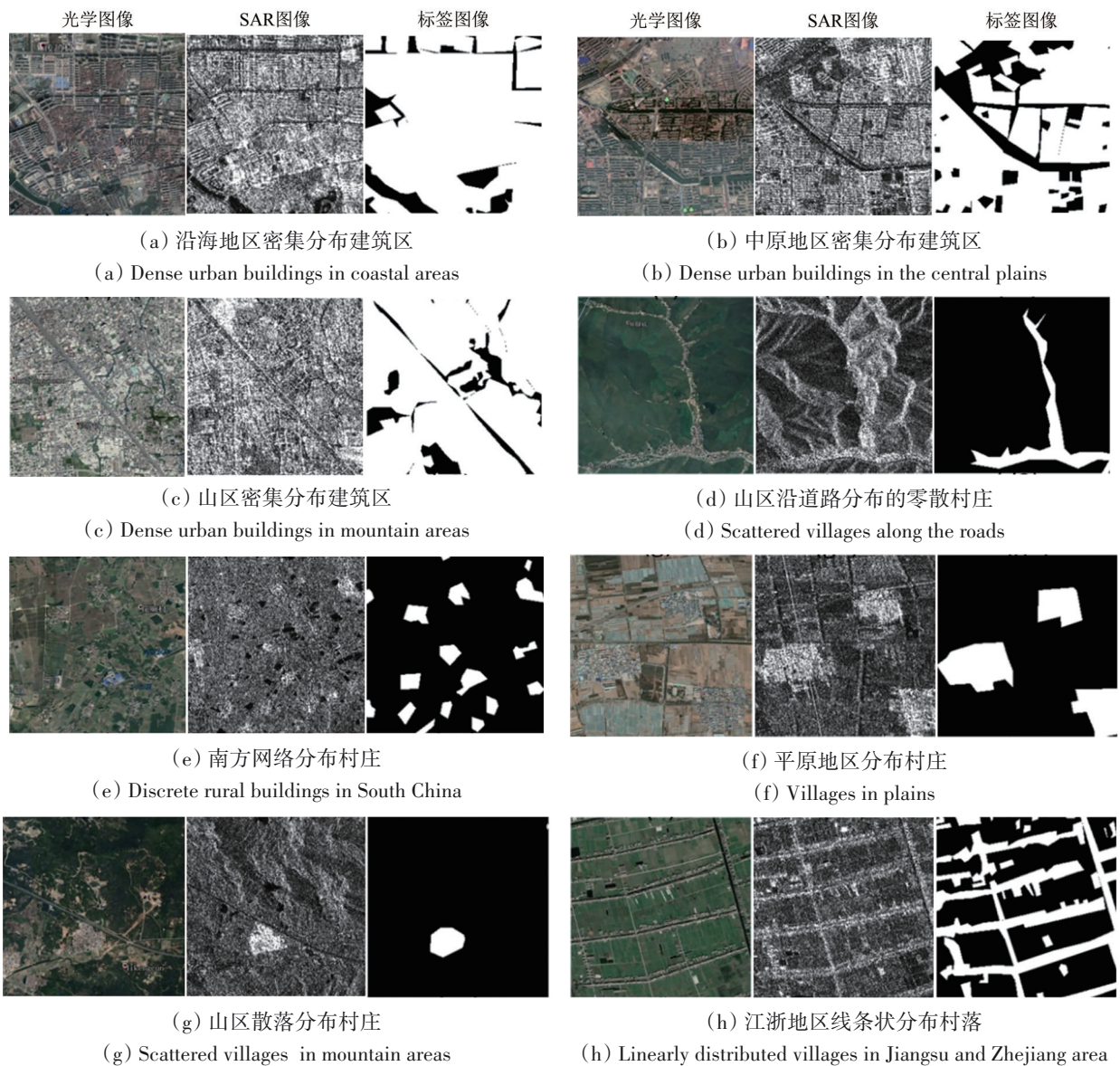


图4 不同建筑样本示例

Fig. 4 Examples of SARBuD1.0 dataset

2.2.4 数据集组织结构

SARBuD1.0 数据集目标是支持研究者对 SAR 图像进行建筑区与非建筑区的二元分割。因此本数据集参考语义分割数据集的方式 (Chen 等, 2019; Ji 等, 2019), 核心部分包含 SAR 图像样本及相应的图像二分类标签。二分类标签图由解译专家根据对应的高分辨率光学图像对 SAR 图像进行建筑区解译的结果。图 4 为数据集示例。

整个数据集的组织结构见图 5 所示。数据集目前包含 27 景 SAR 数据, 单景 SAR 数据以源数据命名。每景数据文件中包含对应的 SAR 图像强度图、SAR 图像样本切片, 相应的标签切片以及后向散射文件, 文件格式都为 TIFF 格式。还包括图像相应元数据文件, 格式为 XML, 元数据文件涵盖了 SAR 数据的成像时间、入射角、经纬度等信息。

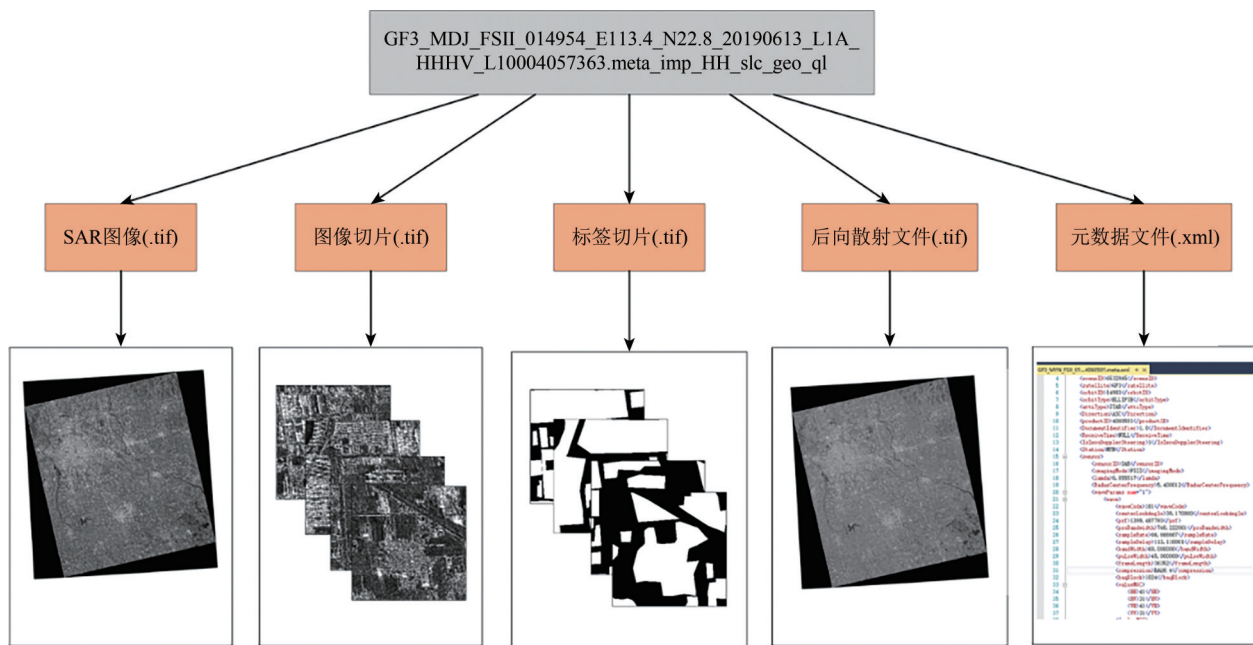


图 5 建筑 SAR 数据集组织结构
Fig. 5 Structure of SARBuD1.0 dataset

数据集覆盖了中国典型地区的建筑区分布类型和地形场景。获取了有效建筑区样本切片 60000 张, 非建筑样本切片共 80000 张。图 6 为各地区所采集的建筑区样本切片个数统计。

根据所获取的源数据情况, 数据集目前主要为 HH, HV 极化图像为主, 包含有不同入射角的样本。根据原图像的入射角范围, 对建筑区样本切片数量进行统计见图 7。根据不同极化方式对样本进行统计见图 8。图 9 给出了不同升降轨的切片数据量统计。以上的数量统计都为建筑样本的统计, 对于负样本, 主要包括山地叠掩区 (30000 张)、结冰河流或湖泊 (86 张)、部分农田植被区 (237 张)、戈壁沙丘区 (500 张)、其他负样本 (水域、道路等 49177 张)。

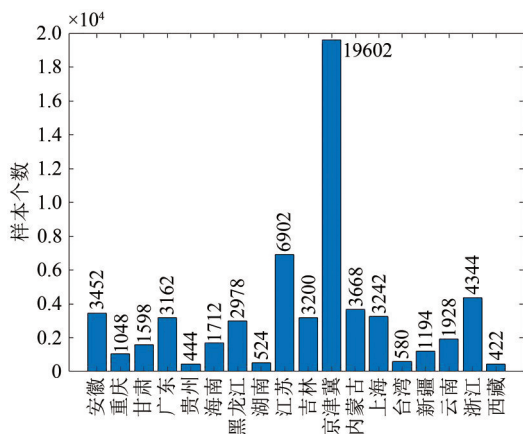


图 6 各地区采集样本个数统计
Fig. 6 The number of image patches of different regions

2.2.5 数据集分析

如前所述, 本数据集数据源为 GF-3 精细模式 II 数据。该数据为双极化数据, 从获取的数据看多为 HH+HV 极化。获取了北京地区升降轨的 GF-3 图像, 经辐射定标、地理编码后, 对建筑在 HH 和

HV 极化通道及升降轨图像中的散射强度进行分析。图 10 为选取的两块示例分析区，分别为图 10 (a) 示例区 1，为城市建筑密集区，图 10 (b) 示例区 2，为村落建筑区。

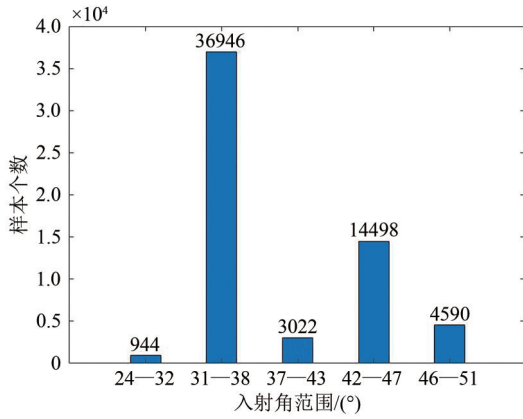


图 7 图像入射角范围样本数

Fig. 7 The number of image patches of different range of incidence angles of original SAR images

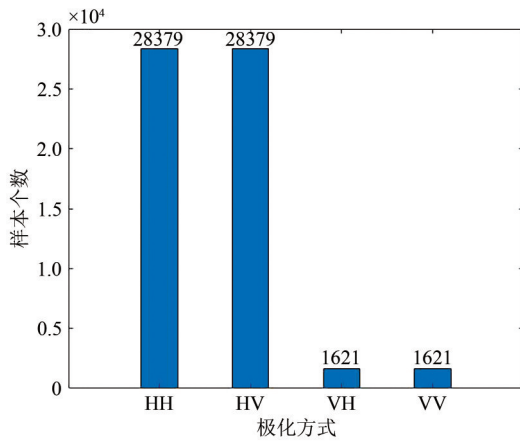


图 8 各极化方式样本数

Fig. 8 The number of image patches of different polarization

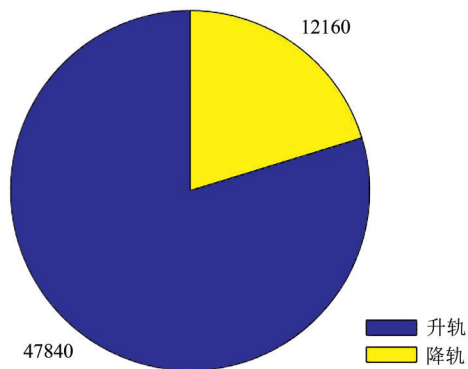


图 9 不同升降轨样本数

Fig. 9 The number of image patches of different orbit mode

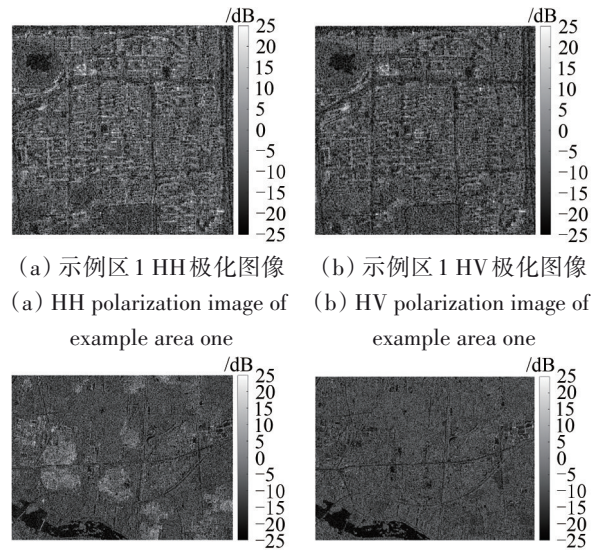


(a) 示例区 1 (b) 示例区 2
(a) Sample area one (b) Sample area two

图 10 示例区高分辨率光学图像(GoogleEarth)

Fig. 10 Optical images of sample areas

图 11 为图 10 两个示例区的不同极化 SAR 图像。从图 11 中可以看出，对于密集城区 HH 极化强度略高于 HV 极化，但两种极化方式图像中建筑区都具有相对高的后向散射强度 (图 11 (a) 和 (b))。对于城郊村落的低矮建筑，HH 极化图像中的建筑区相对于 HV 极化图像具有明显更高的散射强度 (图 11 (c) 和 (d))。导致这样的原因是由于 HV 极化的去极化机理特性，决定了其反射回波总体能量要弱于 HH 极化 (谷秀昌等, 2017)。此外，通常水平极化波 (H) 的入射波二次反射强度大于垂直极化波 (V) 的入射波二次反射强度 (谷秀昌等, 2017)。因而对于建筑通常具有明显二次反射特征的地物，相对于 VV 极化，HH 极化具有更强的散射回波。综上分析，HH 极化更有利于建筑区的提取。



(a) 示例区 1 HH 极化图像 (b) 示例区 1 HV 极化图像
(a) HH polarization image of example area one (b) HV polarization image of example area one

(c) 示例区 2 HH 极化图像 (d) 示例区 2 HV 极化图像
(c) HH polarization image of example area two (d) HV polarization image of example area two

图 11 不同极化 SAR 图像中的建筑

Fig. 11 Buildings in HH and HV SAR images

图 12 为不同升降轨的建筑区 SAR 图像, 图中箭头标示了升降轨方向。雷达的探测方向是地物或目标在雷达图像中呈现何形状或亮度的决定因素 (谷秀昌等, 2017)。不同升降轨将会改变雷达的探测方向。当建筑物墙面走向与雷达波入射方向垂直时, 后向散射强度最大, 当墙面走向与雷达波入射方向平行时, 后向散射强度最弱。通常为了便于采光, 在城市市区建筑物多呈东西向或南北向布置, 因此在升降轨条件下, 雷达波的入射方向与建筑物之间的相对角度关系相似。此外由于建筑密度大且楼层相对较高, 因而在城市建筑密集区不同升降轨条件下建筑区的散射回波强度变化不大 (图 12)。但是在城郊村落, 建筑物相对低矮, 且各建筑物的朝向变化较大, 因而城郊村落中的建筑区在升降轨图像中的散射强度可能出现明显变化 (图 12 (a) 与 (b); 图 12 (c) 与 (d))。因此对于大区域范围 SAR 建筑区提取, 选取升降轨均可。对于局部范围, 可以根据建筑物的走向选择升降轨。如果对时间、数据成本要求不高, 结合升降轨图像可以提高建筑区提取的准确性。

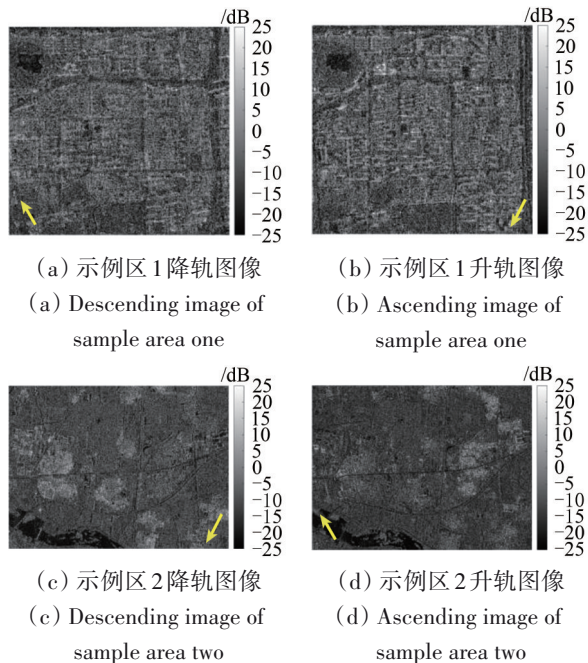


图 12 不同升降轨 SAR 图像中的建筑

Fig. 12 Buildings in SAR images of different orbit direction

3 数据实验与分析示例

本节以 SARBuD1.0 数据集为基础, 进行特征分析以及基于深度学习的建筑区提取的应用示例

分析。以山地区为例分析了建筑区的深度卷积特征; 利用本数据集对深度学习模型进行训练, 开展建筑区提取试验, 并对结果进行分析。

3.1 基于样本集的山地建筑区特征分析

SAR 图像纹理特征是建筑提取常用的特征, 本节以常用的纹理特征为例与深度学习卷积特征进行比较分析 (李璐, 2020)。所用纹理特征包括局部二值模式 LBP (Local Binary Pattern)、均匀二值模式 ULBP (Uniform Local Binary Pattern)、灰度共生矩阵 GLCM (Gray-level Co-occurrence Matrix) 方差与相关度特征。图 13 为两幅 SAR 示例图像。

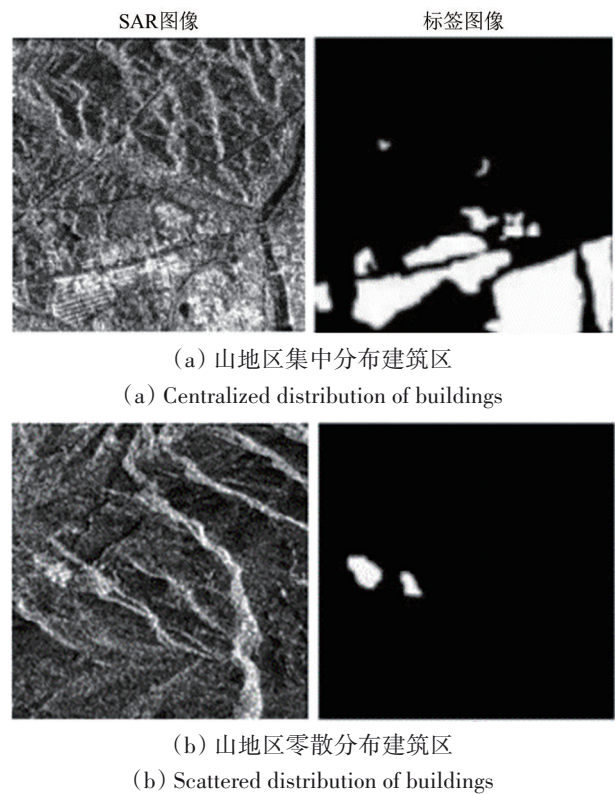


图 13 山地建筑区 SAR 图像。

Fig. 13 SAR image of buildings in mountain region

图 14 为图 13 两幅场景的不同纹理特征图。图 14 (a) 为图 13 (a) 的 LBP、ULBP、GLCM 方差和 GLCM 相关度特征。图 14 (b) 为图 13 (b) 的 LBP、ULBP、GLCM 方差和 GLCM 相关度特征。从图 14 (a) 和图 14 (b) 中可以看出, ULBP 特征图中建筑的线条纹理相对清晰、突出, 但图 14 (b) 右下角的山体叠掩区在 ULBP 特征图中也较为明显, 与建筑区的线条纹理相似。图 14 (a) 图 14 (b) LBP 特征图中建筑区能有所体现, 但相对于周边环

境并不突出。可以看出,虽然LBP算子适合SAR图像一些目标的检测、分类/分割任务,但不适合山地区域的建筑区提取。图14(a)图14(b)所示的GLCM方差特征在一定程度上保留了建筑区的亮度特征,但易与山区叠掩混淆。同样,图14(a)、图14(b)GLCM相关性特征图中,建筑区的纹理无法与山体纹理很好区分。

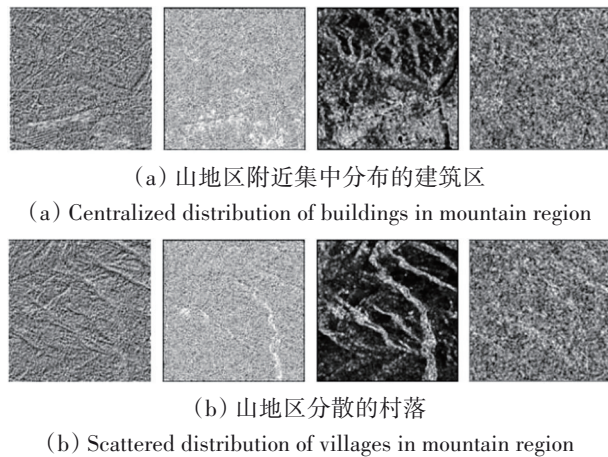


图14 两幅示例SAR图像的不同纹理特征图

Fig. 14 Different texture images derived from two SAR images

使用经典的卷积神经网络ResNet50(He等,2016)获取不同卷积层的多尺度纹理特征图开展分析(Li等,2020)。以图13(a)和图13(b)作为输入图像,列举出卷积网络中不同卷积层:‘Conv1’、‘add_3’、‘add_6’和‘add_13’的特征图,并展示每个卷积层的前8张纹理特征进行分析(图15)。其中,‘Conv1’层处于4个卷积层的最顶端,可获取较为详细的纹理;‘add_13’层处于4个卷积层的最底端,可获取不同地物类型的语义特征(本文分为建筑类和非建筑类)。从图15(a)的特征图中可以看出,随着网络卷积层的深度不断加大,建筑区的纹理细节逐渐模糊,从图15可以看出,‘add_6’层及更深的特征图中建筑区表现为粗略的亮斑形式;同时,在同一个卷积层中,由于不同的建筑体具有不同的范围、外形和强度,其特征也被分别保留在多个特征图中。因此,可以推断建筑区的特征在卷积神经网络中除了在同一特征图平面空间上有相关性外,特征图的通道间也是具有一定相关性的。图15(b)为山地村落建筑区的多个尺度特征图,从图中可以看出村落建筑区的特征较模糊,在4个卷积层中基本上都以亮

斑的形式被保留和传递。与此同时,从图15(b)的结果可以看出,部分山脊叠掩区呈线条状纹理也被保留在深层语义特征图中。如果对深层特征图进行局部采样观测,可以发现建筑区周围是相对平滑的较为均质的背景区,这使得建筑区在深层语义特征图像中能较好地被区分。然而,由于经过多层卷积后山体范围的一些细节纹理被平滑或丢失,部分山体周边可能存在与建筑区相似局部的特征,导致单一尺度下山体与村落等小区域的建筑区容易出现相似的纹理。

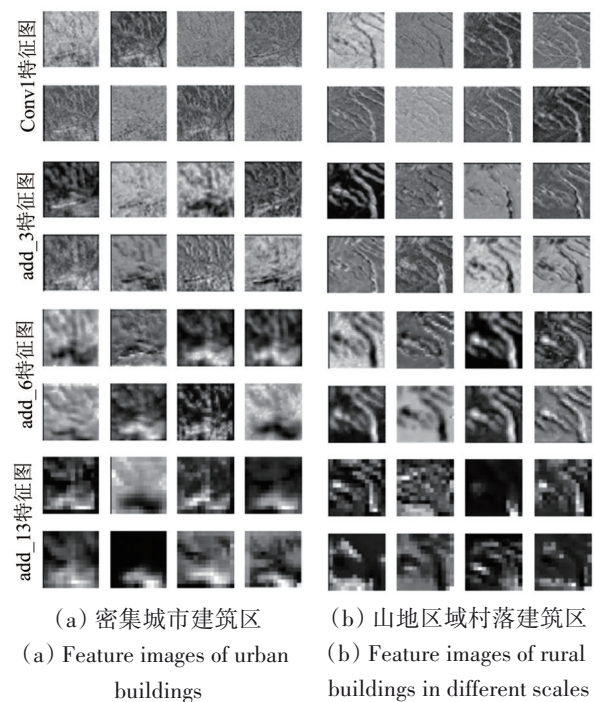


图15 ResNet50多尺度纹理特征图

Fig. 15 Multi-scale texture feature sample images by ResNet50

结合图14、图15的结果可以看出,相比于传统的LBP、ULBP、GLCM等人工设计的纹理特征,卷积神经网络可以产生更多、更深的特征。利用卷积神经网络模型不同卷积核采样可以得到体现建筑区各种纹理的浅层特征,在网络深层卷积结构中可以获得从类别上有较好区分度的深层语义特征,使得分类器模型能更好地检测并提取图像中的建筑区。

3.2 基于样本集的深度学习建筑区提取

基于本文的数据集,利用全卷积神经网络模型进行建筑区提取试验,考察基于本数据集的训练模型,在多类型地形场景的建筑区提取有效性,并对结果进行分析。

在图像中成功定位建筑区位置是提取的重要前提, 采用多尺度注意力U-Net网络结构MA-UNet (Multiscale attention Unet) 进行分析(李璐, 2020)。注意力机制主要启发于空间注意力机制模型(Woo等, 2018)。该注意力机制模型的核心思想是将原始图片中的空间信息变换到另一个空间中并保留了关键信息, 再利用这些信息寻找网络中最重要的部位进行处理。此外本文对深度卷积网络批归一化UNet方法(BN-UNet)(Wei等, 2019)和残差UNet(Res-UNet)方法(Zhang等, 2018)也进行了测试, 并对结果进行了分析。

选取了同时包含山区、平原等地形的河北省内整景SAR图像作为试验区, 数据为GF-3精细条带II模式图像, 幅宽约为100 km×100 km。该图像非样本采集图像。试验区同时包含密集城市建筑区、山区村镇和散落分布、规模不等的田间村落。这些村落与大量的种植耕地混在一起, 植被的生长态势会产生不同的后向散射强度并对周围村落识别产生一定的影响。根据高分辨率Google Earth光学影像随机选择SAR图像的建筑区和非建筑区各10000个样本点来验证结果。

图16和表2分别展示了这几种方法得到的建筑区提取结果及其精度评价。图16(a)为该区域的SAR影像, 可以看出主要城市区石家庄和邢台分布最为明显, 其余的城镇和村落在地理位置上呈聚集状的均匀分布。图16(b)——图16(d)分别展示了3种算法的建筑区提取结果, 从结果可以看出BN-UNet算法在密集城区具有较好的提取结果, 说明其算法在平原地区的密集城市提取方面具有一定的效果。然而, BN-UNet算法在石家庄市南部区域产生了大量的漏检, 在结果图中呈条带状的建筑区断层。且算法在山区、平原村落的提取方面也存在大量的漏检。Res-UNet中使用了残差连接结构来提高算法对多类型场景区域、小规模目标(这里指村落)的提取效果。对比图16(b)、图16(c)的结果可以看出, Res-UNet算法在不同区域均比BN-UNet提取到了更多的建筑区, 也弥补了石家庄市南部区域的大量漏检。然而, 由于该算法在西部的山区也提取到了大量的山体虚警, 甚至在结构上呈现出了山体的叠掩纹理。因此, 虽然Res-UNet将总体精度提升到了87.2%, 但是其用户精度却小于60%。相比之下, MA-UNet算

法在该区域的建筑区提取结果最优, 不仅在平原地区密集城区、村镇、村落等建筑区获取了较好结果, 在西部的山区也以较小的虚警率为代价获取了建筑区的提取。从统计结果可以看出, MA-UNet算法的总体提取精度达到了95.6%, 用户精度和生产精度也分别超过了90%。从以上分析可以看出, 3类网络模型的建筑区提取的总精度都在80%以上, 本文的数据集可以对基于深度学习方法的建筑区提取具有很好的支持。

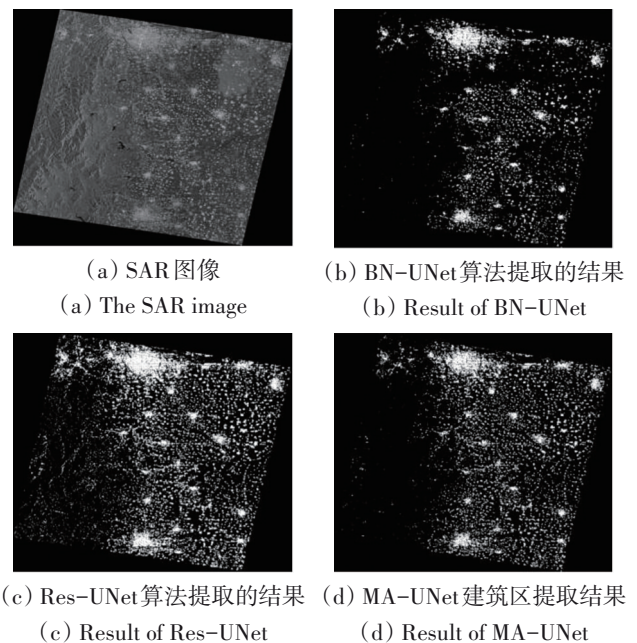


图16 3种方法建筑区提取结果

Fig. 16 Building extraction results of three methods

表2 3种深度学习方法的提取结果比较

Table 2 Results of three deep learning methods

方法	Kappa系数	总体精度/%	生产精度/%	用户精度/%
BN-UNet	0.77	82.6	74.5	79.6
Res-UNet	0.62	87.2	90.3	59.8
MA-UNet	0.84	95.6	93.6	92.4

4 结论

本文介绍了一个面向深度学习大区域建筑区提取的建筑样本集。该数据集采集于GF-3 FSII模式SAR图像。源数据获取于中国20个省、市、自治区, 包含了不同分布类型、不同区域、不同地形场景的建筑。通过结合光学图像与专家解译, 制作了与样本数据对应的标签数据。科研工作者

可以通过网站下载公开的数据, 下载地址为: <https://github.com/CAESAR-Radi/SARBuD> [2021-06-16]。可以在此数据集基础上开展建筑散射特性分析与建筑区语义分割提取方面的研究。

使用传统纹理特征与深度学习特征对山地村落以及密集建筑区进行了特征分析与比较, 从结果看, 相比于传统的LBP、ULBP、GLCM等人工设计的纹理特征, 卷积神经网络具有更深、更多的特征, 网络模型浅层特征可以更好表现地物的细节边缘信息, 网络的深层特征包含了可以显示类别的语义特征, 因此使用合适的深度模型能更好地检测并提取图像中指定的目标。基于本数据集利用深度学习方法对不同地形区域的建筑区进行提取试验。试验结果表明该数据集可以很好支持面向大数据的深度学习方法。

人工智能深度学习方法依赖于海量数据的训练学习与模型精化。虽然目前有仿真模拟、迁移学习等基于少量样本或样本不充分条件下的模型训练处理方法。但数据仍是提高模型处理能力和泛化能力的基础与关键。因此数据集的不断扩充与完善(如: 针对不同传感器、不同波段、不同分辨率、不同极化等参数的SAR图像, 在不同国家/地区的建筑类型等条件下获取样本数据)仍是以后的工作方向。

志 谢 本数据集所采用的高分三号SAR数据得到了中国资源卫星应用中心的大力支持, 在此表示衷心的感谢!

参考文献(References)

- Adelipour S and Ghasseman H. 2018. Building detection in very high resolution SAR images via sparse representation over learned dictionaries. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(12): 4808-4817 [DOI: 10.1109/JSTARS.2018.2876910]
- Ban Y F, Jacob A and Gamba P. 2015. Spaceborne SAR data for global urban mapping at 30 m resolution using a robust urban extractor. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103: 28-37 [DOI: 10.1016/j.isprsjprs.2014.08.004]
- Cao H, Zhang H, Wang C and Zhang B. 2018. Operational built-up areas extraction for cities in China using sentinel-1 SAR data. *Remote Sensing*, 10(6): 874 [DOI: 10.3390/rs10060874]
- Chen Q, Wang L, Wu Y F, Wu G M, Guo Z L and Waslander S L. 2019. TEMPORARY REMOVAL: aerial imagery for roof segmentation: a large-scale dataset towards automatic mapping of buildings. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147: 42-55 [DOI: 10.1016/j.isprsjprs.2018.11.011]
- China Centre for Resources Satellite Data and Application (CRESDA). 2016. GF-3 Basic Product Specification Document (中国资源卫星应用中心. 2016. 高分三号卫星产品用户手册)
- Chini M, Pelich R, Hostache R, Matgen P and Lopez-Martinez C. 2018. Towards a 20 m global building map from sentinel-1 SAR data. *Remote Sensing*, 10(11): 1833 [DOI: 10.3390/rs10111833]
- Esch T, Heldens W, Hirner A, Keil M, Marconcini M, Roth A, Zeidler J, Dech S and Strano E. 2017. Breaking new ground in mapping human settlements from space - the global urban footprint. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134: 30-42 [DOI: 10.1016/j.isprsjprs.2017.10.012]
- Gu X C, Fu K and Qiu X L. 2017. Fundamental Knowledge for SAR Image Interpretation. (谷秀昌, 付琨, 仇晓兰. 2017. SAR图像判读解译基础. 北京: 科学出版社)
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Ji S P, Wei S Q and Lu M. 2019. Fully convolutional networks for multi-source building extraction from an open aerial and satellite imagery data set. *IEEE Transactions on Geoscience and Remote Sensing*, 57(1): 574-586 [DOI: 10.1109/TGRS.2018.2858817]
- Li L. 2020. Built-up Area Intelligent Extraction and Change Detection with SAR Images Over Large Scale Space. Beijing: University of Chinese Academy of Sciences (李璐. 2020. SAR图像大空间尺度建筑区智能提取与变化检测研究. 北京: 中国科学院大学)
- Li X, Su J and Yang L. 2020. Building detection in SAR images based on bi-dimensional empirical mode decomposition algorithm. *IEEE Geoscience and Remote Sensing Letters*, 17(4): 641-645 [DOI: 10.1109/LGRS.2019.2928965]
- Shahzad M, Maurer M, Fraundorfer F, Wang Y Y and Zhu X X. 2019. Buildings detection in VHR SAR images using fully convolution neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 1100-1116 [DOI: 10.1109/TGRS.2018.2864716]
- Sun Z C, Xu R, Du W J, Wang L and Lu D S. 2019. High-resolution urban land mapping in China from sentinel 1A/2 imagery based on Google earth engine. *Remote Sensing*, 11(7): 752 [DOI: 10.3390/rs11070752]
- Tavares P A, Beltrão N E S, Guimarães U S and Teodoro A C. 2019. Integration of Sentinel-1 and Sentinel-2 for classification and LULC mapping in the urban area of Belém, eastern Brazilian Amazon. *Sensors*, 19(5): 1140 [DOI: 10.3390/s19051140]
- Thiele A, Cadario E, Schulz K, Thonnessen U and Soergel U. 2007. Building recognition from multi-aspect high-resolution InSAR data in urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 45(11): 3583-3593 [DOI: 10.1109/TGRS.2007.898440]
- Wei S S, Zhang H, Wang C, Wang Y Y and Xu L. 2019. Multi-temporal SAR data large-scale crop mapping based on U-Net model. *Remote Sensing*, 11(1): 68 [DOI: 10.3390/rs11010068]

- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//Proceedings of the 15th European Conference on Computer Vision. Munich: Springer: 3-19 [DOI: 10.1007/978-3-030-01234-2_1]
- Zhang H, Wang C, Zhang B, Wu F and Yan D M. 2009. Object Recognition with High Resolution SAR Image. Beijing: China Science Publishing and Media Ltd (张红, 王超, 张波, 吴樊, 闫冬梅. 2009. 高分辨率 SAR 图像目标识别. 北京: 科学出版社)
- Zhang Q J. 2017. System design and key technologies of the GF-3 satellite. *Acta Geodaetica et Cartographica Sinica*, 46(3): 269-277 (张庆君. 2017. 高分三号卫星总体设计与关键技术. *测绘学报*, 46(3): 269-277) [DOI: 10.11947/j.AGCS.2017.20170049]
- Zhang Z H, Guo W W, Yu W H and Yu W X. 2019. Multi-task fully convolutional networks for building segmentation on SAR image. *The Journal of Engineering*, 2019(20): 7074-7077 [DOI: 10.1049/joe.2019.0569]
- Zhang Z X, Liu Q J and Wang Y H. 2018. Road extraction by deep residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5): 749-753 [DOI: 10.1109/LGRS.2018.2802944]
- Zhao J P, Zhang Z H, Yao W, Datcu M, Xiong H L and Yu W X. 2020. OpenSARUrban: a Sentinel-1 SAR image dataset for urban interpretation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 187-203 [DOI: 10.1109/JSTARS.2019.2954850]
- Zhu J Y, Park T, Isola P and Efros A A. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE: 2242-2251 [DOI: 10.1109/ICCV.2017.244]

SARBuD1.0: A SAR building dataset based on GF-3 FSII imageries for built-up area extraction with deep learning method

WU Fan¹, ZHANG Hong¹, WANG Chao¹, LI Lu^{1,2}, LI Juanjuan^{1,2}, CHEN Weirong³, ZHANG Bo¹

1. Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. China Center for Resource Satellite Data and Applications, Beijing 100094, China

Abstract: Synthetic Aperture Radar (SAR) is one of the important data sources for built-up area information acquisition and dynamic monitoring. In this paper, SARBuD1.0, a SAR image patch dataset of built-up area with GF-3 Fine strip-map mode, is introduced.

The dataset is derived from 27 scenes of 10 m resolution GF-3 SAR images covering different regions in China. Approximately 60000 samples of built-up area are obtained from the SAR images. The dataset consists of built-up area SAR image patches and corresponding label images that are interpreted by experts with high-resolution optical images. The dataset contains built-up areas of different distribution types and different regions. The terrain scenes of the samples include plain area, mountain area, and plateau. The SARBuD1.0 dataset can support researchers to analyze the image features of built-up areas in different regions, assist in SAR image understanding, and provide training and test data for deep learning methods for built-up area segmentation in SAR images.

In this paper, using built-up areas in mountain areas as an example, the traditional texture features and deep learning features of the built-up area are analyzed. Experiments show that a convolution neural network can provide deeper features of built-up areas compared with traditional texture features. Features in different scales can be obtained by the shallow convolutional layers of the network, and semantic information can be obtained by the deep layers. Therefore, the deep convolutional neural network classifier can detect and extract the built-up area better. Based on the dataset, three deep learning methods are applied to extract the built-up area in different terrain areas. The experimental results show that the deep learning models can achieve good results for built-up area extraction based on the dataset.

The dataset can effectively support the deep learning method for big data processing. Based on the SARBuD1.0 dataset, scholars can carry out research on feature analysis and semantic segmentation of built-up areas with SAR images.

Key words: remote sensing, Synthetic Aperture Radar (SAR), building, dataset, deep learning, GF-3, semantic segmentation

Supported by National Natural Science Foundation of China (No. 41930110, 41971395)