

MFBFS: 高分辨率多光谱遥感影像细粒度建筑物特征集

王振庆^{1,2}, 周艺¹, 王福涛^{1,2}, 王世新¹, 高郭瑞^{1,2}, 朱金峰¹,
王平³, 胡凯龙³

1. 中国科学院空天信息创新研究院, 北京 100094;

2. 中国科学院大学, 北京 100049;

3. 应急管理部国家减灾中心, 北京 100124

摘要: 遥感影像建筑物信息提取对城市信息管理与防灾减灾等领域具有至关重要的作用。本文建立了一个高分辨率多光谱遥感影像细粒度建筑物特征集 MFBFS。MFBFS 采用国产高分二号多光谱遥感影像作为数据源, 选择覆盖了总共 3668 km² 的中国各个灾害带的 21 个区县建筑物集中区域为研究区, 从光谱、纹理、边缘、指数 4 个角度, 生成了 17 种特征分量。MFBFS 中共包含超过 26 万个建筑物实例, 大规模的数据保证了较高的类内差异, 包括尺寸、形状、颜色、角度、背景等差异, 为后续泛化性模型的建立提供了数据支撑。此外, MFBFS 特地将建筑物按照结构类型分为钢及钢筋混凝土结构、砌体结构以及砖石和其他结构 3 种。不同结构类型的建筑物抵御灾害的能力以及可使用时间区别显著, 细粒度的设计使得遥感提取建筑物任务将发挥更大的作用, 尤其是灾害领域的灾前损失预测和灾后损失评估。为保证地面真实值的高度准确性, 进行了严格的质量流程控制和实地考察验证工作, 最终得到 191GB 高质量特征和标签数据。初步的深度学习实验表明了 MFBFS 的有效性。该特征集 (下载方式: <https://github.com/WangZhenqing-RS/MFBFS> [2023-12-11]) 可为建筑物结构细粒度提取研究提供良好的数据支持, 也可促进国产高分遥感数据应用发展。

关键词: 高分遥感, 多光谱影像, 细粒度类别, 建筑物提取, 特征集

中图分类号: TP701/P2

引用格式: 王振庆, 周艺, 王福涛, 王世新, 高郭瑞, 朱金峰, 王平, 胡凯龙. 2024. MFBFS: 高分辨率多光谱遥感影像细粒度建筑物特征集. 遥感学报, 28(11): 2780-2791

Wang Z Q, Zhou Y, Wang F T, Wang S X, Gao G R, Zhu J F, Wang P and Hu K L. 2024. MFBFS: A fine-grained building feature set for high-resolution multispectral remote sensing images. National Remote Sensing Bulletin, 28(11): 2780-2791 [DOI: 10.11834/jrs.20243526]

1 引言

高分辨率遥感影像技术的迅猛发展已经成为现代地理信息科学和遥感领域的一个重要突破, 高分辨率影像为我们提供了更加详尽和准确的地球表面信息。建筑物作为城市环境的重要组成部分, 其位置、类型、面积等信息在遥感影像中准确的提取对于城市规划和管理以及防灾减灾等领域至关重要 (王振庆等, 2021; Wang 等, 2023a)。

遥感影像建筑物提取常用的方式包括无监督学习和有监督学习。无监督学习不需要进行可训练样本的制作, 通过对影像进行特定运算突出建筑物的先验特征, 进而得到建筑物提取结果。该类方法以建筑物指数 (Nosrati 和 Saeedi, 2009; Pesaresi 和 Gerhardinger, 2010; Huang 和 Zhang, 2011) 为代表, 然而由于建筑物的多样性和相似地物的干扰, 应用效果较差。随着机器学习和深

收稿日期: 2023-12-11; 预印本: 2024-05-23

基金项目: 高分遥感在电子政务地理空间基础信息库建设与服务中的应用示范 (二期) (编号: 00-Y30B01-9001-22/23-CY-01); 国家重点研发计划 (编号: 2021YFB3901201)

第一作者简介: 王振庆, 研究方向为遥感信息智能提取。E-mail: wangzhenqing19@mails.ucas.ac.cn

通信作者简介: 王福涛, 研究方向为重大灾害应急监测与风险评估。E-mail: wangft@aircas.ac.cn

关联数据: <https://github.com/WangZhenqing-RS/MFBFS>

度学习的发展，有监督学习在建筑物提取任务中应用越来越多。数据驱动的有监督学习需要大量的可训练样本才能保证最终模型的泛化性。因此，国内外学者在建筑物样本集方面做了大量工作。Massachusetts 样本集 (Mnih, 2013)、国际摄影测量和遥感学会 ISPRS (International Society for Photogrammetry and Remote Sensing) 的 Vaihingen 和 Potsdam 样本集 (<https://www2.isprs.org>[2023-12-11])、Inria 样本集 (Maggiori 等, 2017)、WHU-Building 样本集 (Ji 等, 2019) 以及 UBC 样本集 (Huang 等, 2022) 等被广泛使用。除了上述光学遥感建筑物样本集之后，以合成孔径雷达 SAR (Synthetic Aperture Radar) 为数据源的 SARBuD 样本集 (吴樊 等, 2022) 以及由光学和 SAR 两种模态数据组成的 SpaceNet 6 (Shermeyer 等, 2020) 和 DFC2023 样本集 (<https://www.grss-ieee.org/technical-committees/image-analysis-and-data-fusion>[2023-12-11]) 也在促进建筑物提取中发挥了很大的作用。如表 1 所示，Massachusetts、ISPRS、Inria、WHU-Building、SARBuD、SpaceNet 6 仅将建筑物作为单个类别看待，UBC 将建筑物从屋顶类型（如平屋顶、斜屋顶等）和功能用途（如办公楼、学校等）

角度分为了 61 类，DFC2023 则从屋顶类型分为了 12 类，加速了建筑物提取在城市中的多方面应用。除 WHU-Building 和 DFC2023 覆盖面积较大且涉及城市较多外，其他样本集覆盖面积相对较小或影像区域较为集中，训练得到的模型的真实泛化性难以保证。SARBuD、SpaceNet 6 以及 DFC2023 由于使用了 SAR 数据，使建筑物提取突破了天气的限制，但是由于 SAR 影像的成像机制，使得在天气较好情况下的建筑物提取精度不及光学影像 (Xia 等, 2021)。在光学样本集中，除了 ISPRS 样本集含有少量近红外光谱特征外，其他均只包含可见光谱特征，较少的特征限制了模型学习能力的上限。此外，在防灾减灾及建筑物信息普查等领域中，建筑的结构类别信息十分重要。在地震灾害事件中，人员伤亡和财产损失主要是由建筑物损毁引起的。不同结构类型的建筑抵御灾害的能力不同，造成的破坏程度也不同 (Geiß 等, 2015)。所以在地震风险管理等领域上，不仅需要建筑物空间分布的轮廓信息，还需要建筑物结构等其他属性信息 (Qi 等, 2017; Liu 等, 2020)。因此，构建一个大规模、特征多样、结构细粒度类别的建筑物特征集具有重要实际意义。

表 1 常用建筑物样本集基本情况

Table 1 The basic information on common building sample sets

样本集	分辨率/m	面积/km ²	特征	类别数
Massachusetts	1.0	340	可见光谱	1
ISPRS	0.05/0.09	2/11	可见光谱(少量近红外)	1
Inria	0.3	405	可见光谱	1
WHU-Building	0.075/0.3—2.5	450/550	可见光谱	1
UBC	0.5–0.8	65.6	可见光谱	61(屋顶类型+功能用途)
SARBuD	10	约 35000	SAR	1
SpaceNet 6	1.0	120	可见光谱、SAR	1
DFC2023	0.5—1.0	约 780	可见光谱、SAR	12(屋顶类型)
Ours	0.8	3668	多光谱、纹理、边缘、指数	3(结构类型)

为推动高分辨率遥感影像建筑物提取的研究进展，本文创建了一个大型的用于建筑物细粒度提取的高质量多光谱遥感影像特征集 MFBFS (Multispectral Fine-grained Building Feature Set)。MFBFS 采用高分二号多光谱卫星为数据源，生成了光谱、纹理、边缘、指数 4 大类 17 种特征，并且将建筑物按照结构类型分为 3 种细粒度类别：钢及钢筋混凝土结构、砌体结构、砖石和其他结构。

它包含了中国 20 个常发生灾害的区县区域，含有超过 26 万个建筑物实例。

2 数据来源及预处理

2.1 数据来源

为了使特征集在灾害领域发挥更大价值，我们在选择采集图像区域时以中国自然灾害带为依据。Zhang 等 (1995) 据灾害系统理论、地域分异

理论、地学信息图谱理论，以县为基本单元，对自然灾害程度进行了系统评价。在此基础上，制定了包括一级区划（6个自然灾害带），二级区划（26个自然灾害区），三级区划（93个自然灾害小区）的中国自然灾害区划。我们根据该自然灾害区划，选取了21个区县中的建筑物集中区域作为特征集中的研究区，相对均匀地分布在陆上的5个灾害带，如表2所示。

表2 研究区在中国自然灾害带的分布情况
Table 2 The distribution of the research area in regionalization of natural disasters in China

灾害带	研究区
东南沿海灾害带	永嘉县、徐闻县、万宁市
东部灾害带	宁安市、开原市、涑源县、寿光市、新乡县、庐江县、衡东县、松北区
中部灾害带	大宁县、恩施市、腾冲市、水城县
西北灾害带	喀什市、伊州区、平罗县
青藏灾害带	迭部县、玉树市、米林县

特征集采用高分二号数据作为数据源。高分二号（GF-2）卫星是中国自主研制的首颗空间分辨率优于1 m的民用光学遥感卫星。其搭载两台高分辨率1 m全色和4 m多光谱传感器，星下点空间分辨率可达0.8 m。GF-2图像兼具空间分辨率高和视野开阔两大优点，允许在大范围内观察详细的地面信息。GF-2已被用于灾害应急评估、土地覆盖调查和其他重要应用。

2.2 数据预处理

为了获得高质量高分辨率的遥感影像，我们对高分二号影像采用了一系列的预处理方法。首先，我们在选择数据之前检查了原始数据的质量，去除了具有较多云、噪声和亮点的图像。其次，对多光谱图像进行辐射定标、大气校正以及正射校正，对全色图像进行辐射定标和正射校正。最后，我们使用Gram-Schmidt变换通过融合全色图像来提高多光谱图像的空间分辨率，得到空间分辨率为0.8 m的多光谱结果图像。

3 特征生成算法

获取地物在遥感影像中的特征是将其从遥感影像中高精度提取的先决条件。在使用机器学习进行提取时，需要使用人工设计的特征进行训练，通常包括光谱、纹理、边缘、指数特征（陶万成

等，2022；赵成等，2023；张悦琦和任鸿瑞，2023）。人工设计的特征属于浅层特征，是基于成熟的数学理论设计而来的，具有较强的可解释性。但是由于特征深度不足导致表征性不够强，限制了机器学习的精度上限。而深度学习的出现弥补了机器学习的缺陷，通过对光谱特征运算得到更深的语义特征，在遥感地物提取中广泛应用。然后，深层次的特征描述更加抽象，可解释性较差。一些学者将深度特征与人工设计的浅层特征相结合，使分类器获得了更好的性能（Xu等，2020；Zhang等，2022；Wang等，2023b）。

除遥感影像最基本的光谱特征除外，我们选取了研究人员常用的纹理、边缘、指数特征，这些特征被证实了在建筑物提取实验中的正向作用（游永发等，2019；李金香等，2019；Zhang等，2022；Wang等，2023b）。最终选择的特征分量如图1所示，其中光谱特征包括蓝波段、绿波段、红波段、近红外波段特征，纹理特征包括由灰度共生矩阵衍生的对比度、相异性、同质度、相关性、角二阶矩和局部二值模式以及方向梯度直方图，边缘特征包括一阶边缘特征和多阶边缘特征，指数特征包括建筑物指数、阴影指数、植被指数以及水体指数。每个特征分量的具体计算方法在3.1—3.4节详细给出，并将计算代码一并共享在特征集下载链接中。为方便特征集后续使用，统一将值域拉伸到[0, 1]。

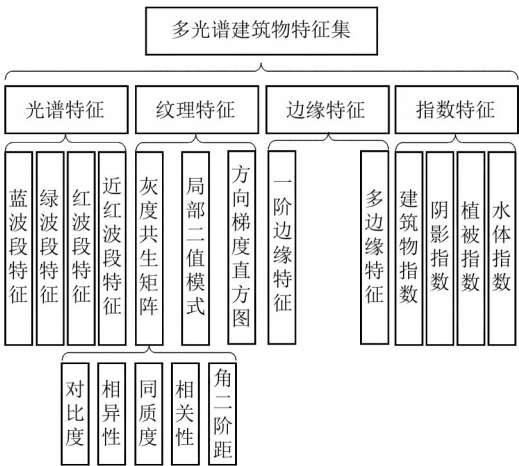


图1 建筑物特征集的组成
Fig. 1 The composition of the building feature set

3.1 光谱特征

地物在遥感影像中的光谱特征是指通过遥感

技术获取的影像所包含的不同波长范围内的光谱信息,通常使用地表反射率来表示。不同类型的地物对不同波段的光谱具有特定的反射特性,使得光谱特征成为地物最基本也是最重要的特征。

高分二号卫星覆盖了蓝、绿、红和近红外4个光谱范围。在经过辐射定标、大气校正后,得到每个波段的光谱特征,如图2所示。

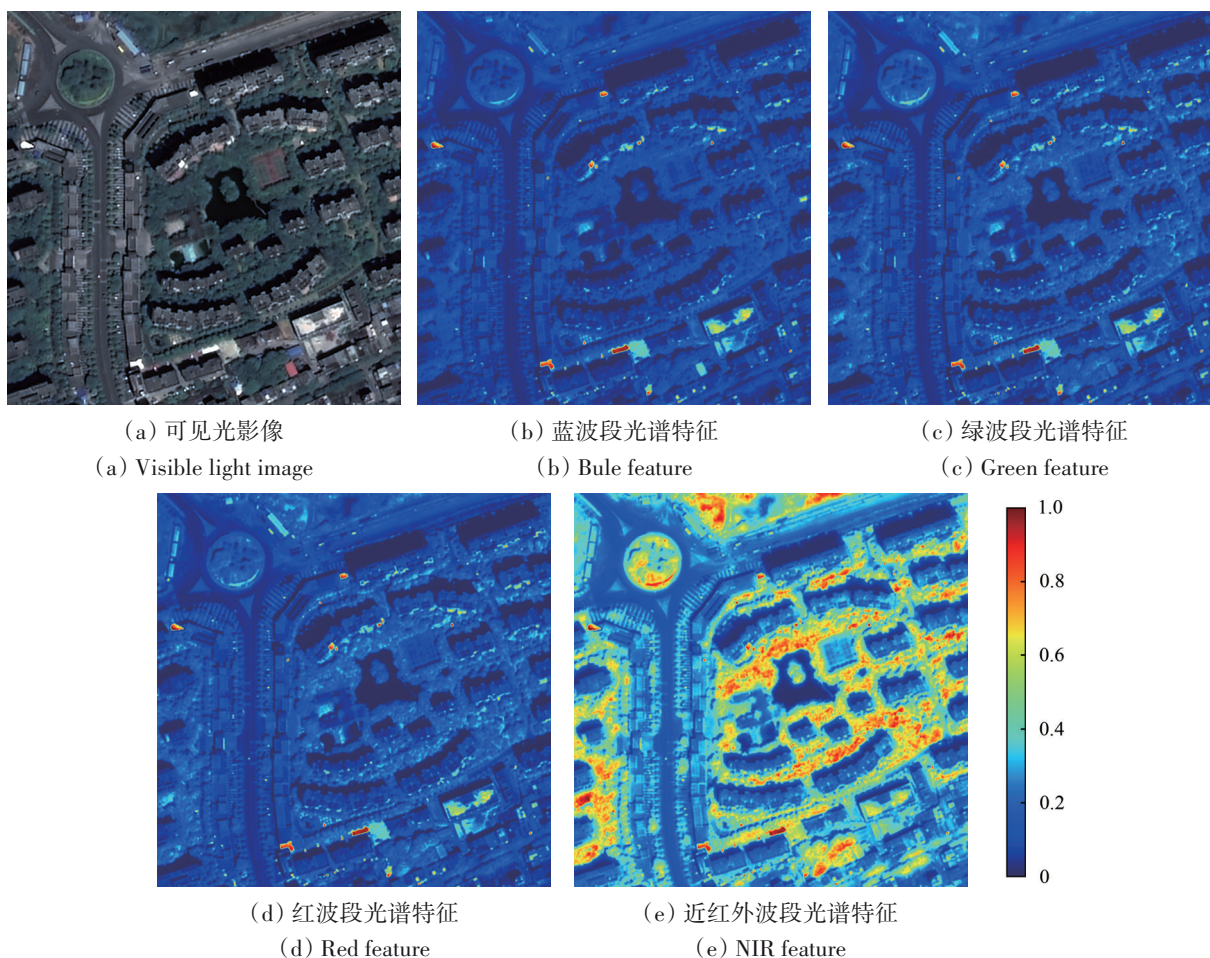


图2 地物光谱特征示例

Fig. 2 Example of spectral features of ground objects

3.2 纹理特征

地物在遥感影像中的纹理特征是指地物表面上的细节和纹理模式,可以用来描述地物的外观、结构和复杂性。较为规则的纹理特征是区别建筑物与非建筑物的重要依据。我们使用常用的灰度共生矩阵GLCM(Gray-Level Co-occurrence Matrix)、局部二值模式LBP(Local Binary Pattern)、方向梯度直方图HOG(Histogram of Oriented Gradients)来表示地物的纹理特征。

灰度共生矩是一种用于描述图像纹理特征的统计工具,该矩阵中的元素值表示灰度级之间联合条件概率密度 $P(i, j|d, \theta)$,即在给定空间距离 d 和方向 θ 时,灰度以 i 为起始点,出现灰度级 j 的概率。根据灰度共生矩阵可以进一步计算对

比度(Contrast)、相异性(Dissimilarity)、同质度(Homogeneity)、相关性(Correlation)以及角二阶矩ASM(Angular Second Moment)等直接纹理特征。

对比度用于描述图像中灰度级别相邻像素对的差异程度,具体公式如式(1)所示,结果如图3(b)所示。

$$\text{Contrast} = \sum_{i=0}^L \sum_{j=0}^L p(i, j) \times (i - j)^2 \quad (1)$$

相异性与对比度类似,但是考虑的是绝对值,具体公式如式(2)所示,结果如图3(c)所示。

$$\text{Dissimilarity} = \sum_{i=0}^L \sum_{j=0}^L p(i, j) \times |i - j| \quad (2)$$

同质度度量了图像中不同灰度级别的像素之间的相似性,具体公式如式(3)所示,结果如图3(d)所示。

$$\text{Homogeneity} = \sum_{i=0}^n \sum_{j=0}^n \frac{p(i,j)}{1 + |i - j|} \quad (3)$$

相关性用于衡量图像中不同灰度级别像素之间的线性关联度, 具体公式如式 (4) 所示, 结果如图 3 (e) 所示。

$$\text{Correlation} = \frac{\sum_{i=0}^n \sum_{j=0}^n (i - \mu_i) \times (j - \mu_j) \times p(i,j)^2}{\sigma_i \times \sigma_j} \quad (4)$$

式中, μ_i 和 σ_i 是指行方向的均值和标准差, μ_j 和 σ_j 是指列方向的均值和标准差。

角二阶矩表示图像中像素值共现的均匀程度, 也被称为反熵, 具体公式如式 (5) 所示, 结果如图 3 (f) 所示。

$$\text{ASM} = \sum_{i=0}^n \sum_{j=0}^n p(i,j)^2 \quad (5)$$

局部二值模式对图像中的每个像素的邻域进行编码来描述纹理特征。首先对每个邻域像素的灰度值与中心像素值进行比较, 如果邻域像素值大于或等于中心像素值, 则将该像素标记为 1; 否则, 标记为 0。根据这些比较结果, 将二进制标记

串转化为一个十进制数, 即得到 LBP 值。具体计算公式如式 (6) 所示, 结果如图 3 (g) 所示。

$$\text{LBP} = \sum_{i=0}^8 s(i) \times 2^i \quad (6)$$

方向梯度直方图对图像局部区域的梯度信息进行编码, 以捕捉的纹理特征。HOG 特征的计算基于图像中的梯度信息。使用 Sobel 算子计算图像的水平 and 垂直梯度, 从而得到每个像素点的梯度强度和梯度方向。之后将图像分割为一组互不重叠的细胞, 每个细胞覆盖一个相对较小的图像区域。对于每个细胞, 计算其内部梯度方向的直方图。直方图将梯度方向分为若干个方向区间, 通常为 9 个。对细胞内的像素进行投票, 将其梯度方向分布统计到对应的直方图区间中。为了提高 HOG 特征的鲁棒性, 将图像划分为若干块, 每个块包含多个细胞, 然后对每个块内的细胞直方图进行归一化。最后, 将所有块内归一化的细胞直方图串联得到 HOG 特征。可视化结果如图 3 (h) 所示。

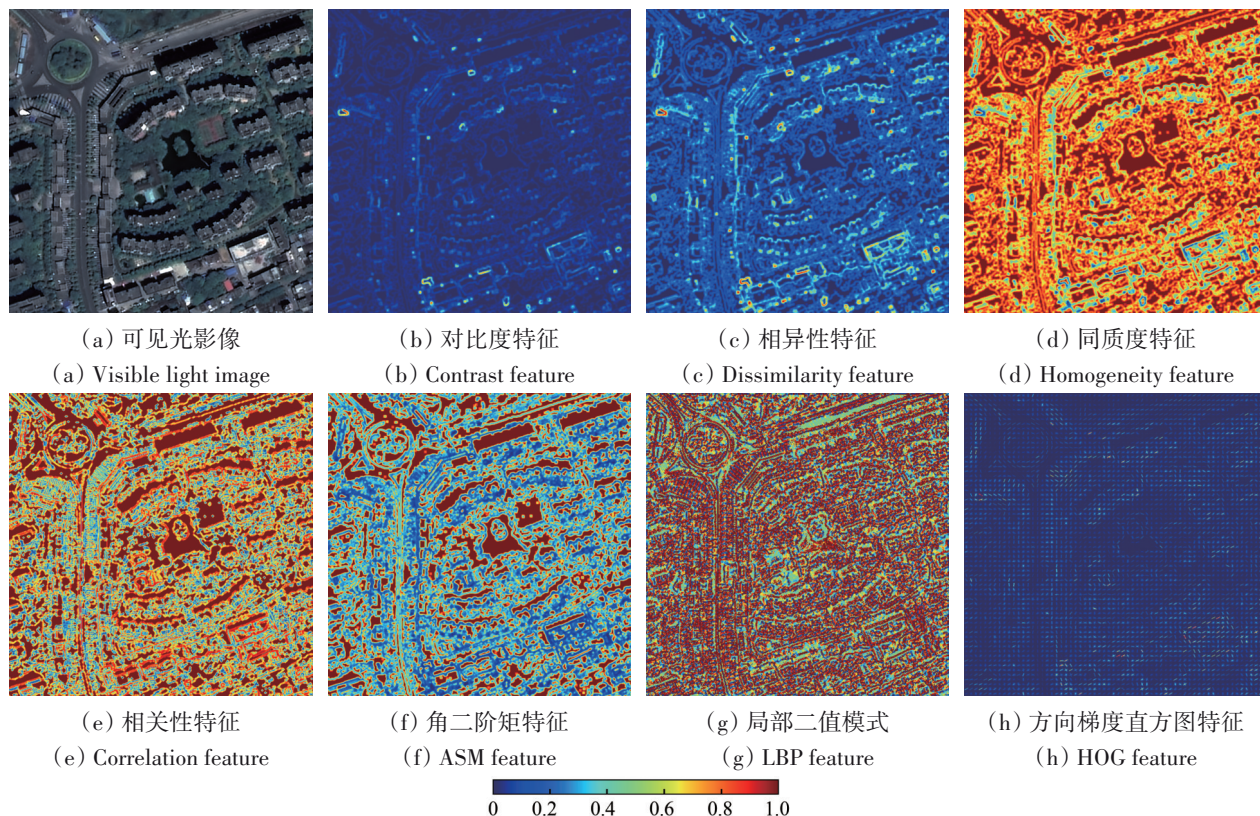


图 3 地物纹理特征示例

Fig. 3 Example of texture features of ground objects

3.3 边缘特征

边缘特征用于描述图像中不同区域之间的灰

度值变化明显的边缘或轮廓。由于期望的建筑物提取结果的边界一般与建筑物的边缘重合, 所以

边缘特征对于精确提取建筑物十分关键。虽然和纹理特征都用于描述图像的局部特征,但它们的关注点不同。边缘特征主要关注图像中不同区域的边界或边缘部分。

Sobel算子是一种常见的一阶导数边缘检测算子,该算子使用两个 3×3 矩阵与原始图像进行卷积,得到水平 E_x 和垂直 E_y 的梯度值,分别如式(7)和式(8)所示。最后结合上述两种结果得到边缘特征 E 。可视化结果如图4(b)所示。

$$E_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * I \quad (7)$$

$$E_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * I \quad (8)$$

$$E = \sqrt{E_x^2 + E_y^2} \quad (9)$$

相对于Sobel算子,Canny算子的计算复杂度较高,需要执行多个步骤。首先,使用高斯滤波以减少噪声的影响。之后使用Sobel算子来计算图像的水平 and 垂直梯度。这产生了每个像素点的梯度强度和梯度方向。最后在梯度图像中,通过在每个像素点处查找局部梯度强度的极大值,来细化边缘。这意味着只有在梯度方向上具有最大梯度强度的像素被保留,其余的像素被抑制。可视化结果如图4(c)所示。

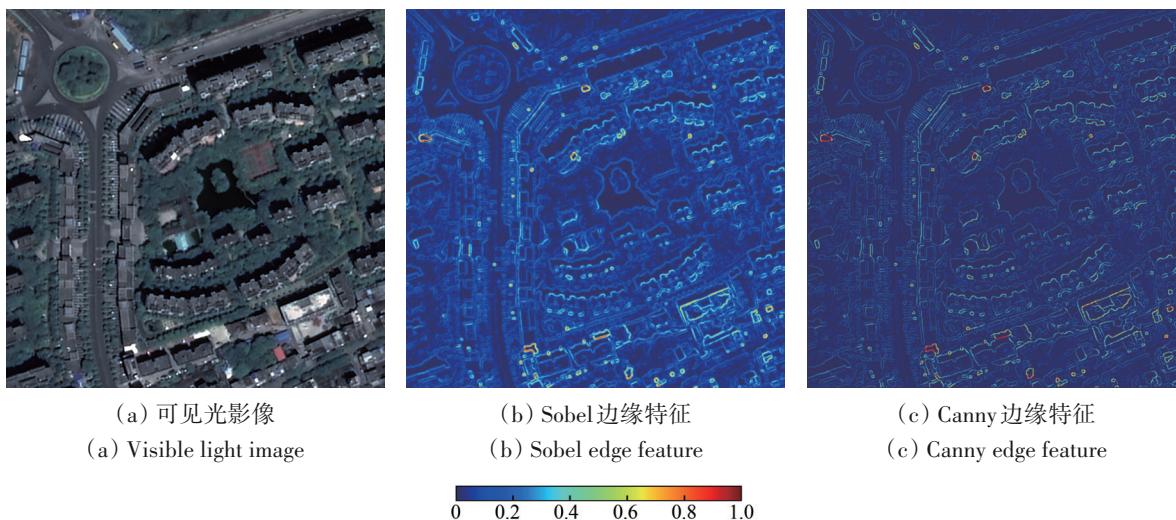


图4 地物边缘特征示例

Fig. 4 Example of edge features of ground objects

3.4 指数特征

遥感图像指数特征是通过通过对遥感图像中不同波段的数据进行运算和分析而得到的。它们提供了丰富的专题信息,有助于更好地理解和分析感兴趣地物。

形态学建筑物指数MBI (Morphological Building Index) 是根据形态学和光谱特征计算得到的与建筑物直接相关的特征,用于突出可能的建筑物的响应(Huang和Zhang, 2011)。首先取光谱带最大值最为后续计算数据。之后进行多角度形态学白帽重构,并基于此进行计算差分形态学剖面DMP (Differential Morphological Profiles),最后由DMP计算得到MBI。具体公式如式(10)所示,结果如

图5(a)所示。

$$MBI = \frac{\sum_{d,s} DMP(d,s)}{D \times ((S_{\max} - S_{\min})/\Delta s + 1)} \quad (10)$$

式中, d 和 s 分别代表线性结构元素的方向和尺度, D 为方向数, $S_{\max}$ 和 $S_{\min}$ 以及 Δs 分别代表结构元素尺度最大值和最小值以及间隔值。

阴影指数SI (Shadow Index) 可以帮助提取地表上的阴影信息,可以用于增强建筑物的特征。由于阴影的存在,建筑物的某些部分可能与周围环境有明显的对比度差异。通过分析阴影指数,可以突出建筑物的边缘和细节,使建筑物更容易识别。由于缺少了太阳光的直射,阴影区域具有高色调、低亮度的特征。计算像素的色调和亮度

之间的比率以突出阴影与非阴影区域的值的差异:

$$SI = \frac{h + 1}{L + 1} \quad (11)$$

式中, h 和 L 分别是像素在 CIELCh 色彩空间的色调和亮度, 在计算之前归一化至 $[0, 1]$ 。在阴影指数图像中, 阴影区域中的像素将具有比非阴影区域中的像素更高的值, 如图 5 (b) 所示。

归一化植被指数 NDVI (Normalized Difference Vegetation Index) 和归一化水体指数 NDWI

(Normalized Difference Water Index) 虽然不直接作用于建筑物提取任务, 但是起到背景去除的辅助作用, 结合其他特征一起实现准确的建筑物提取。具体公式如式 (12) (13) 所示, 结果如图 5 (c) (d) 所示。

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (12)$$

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad (13)$$

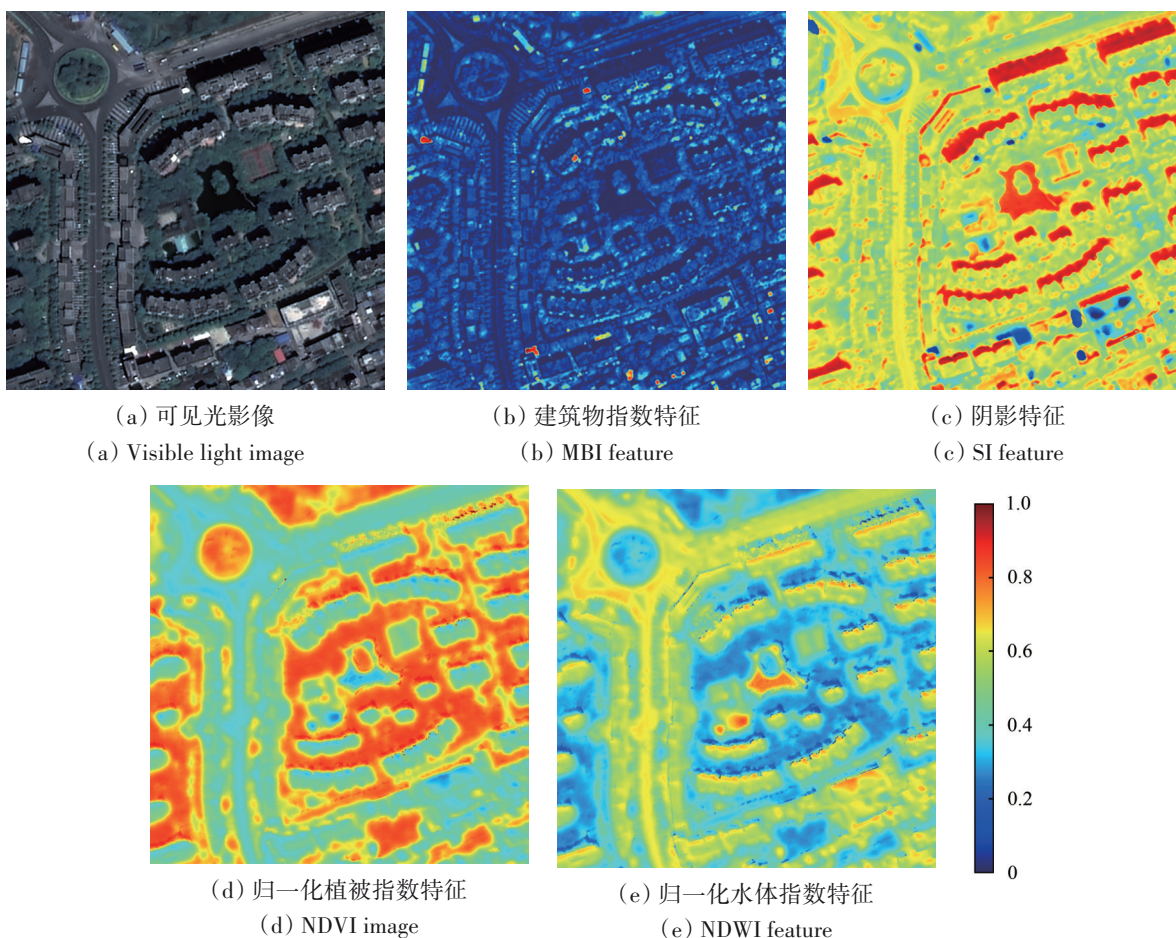


图5 地物指数特征示例

Fig. 5 Example of index features of ground objects

4 地面真实值制作与验证

4.1 地面真实值制作

为了使该特征集在防灾减灾领域发挥更大价值, 对建筑物进行了结构细粒度区分。中国国务院第一次全国自然灾害综合风险普查领导小组办公室印发的第一次全国自然灾害综合风险普查技术规范对建筑物结构类型区分做了规范。其中,

城镇房屋建筑调查技术导则 (FXPC/ZJ G-02, 2021) 将城镇建筑物分为了钢结构、钢筋混凝土结构、砌体结构、木结构和其他结构。农村房屋建筑调查技术导则 (FXPC/ZJ G-03, 2021) 将农村建筑物结构类型共分为钢结构、钢筋混凝土结构、砖石结构、土木结构、混杂结构、窑洞以及其他结构。考虑到遥感图像所获得的建筑特征有限, 将以上建筑物类别全部从遥感图像中识别出来难以实现。综合不同类别建筑物的特征异同,

将建筑物分为3种不同结构类型：钢和钢筋混凝土结构、砌体结构、砖石和其他结构。通过实地地面调查与高分二号影像目视解译相结合的方法，从建筑物的几何特征、光谱特征以及空间特征方面建立了建筑物细粒度地面真实值解译标准：

钢及钢筋混凝土结构：(1) 几何特征：单体面积大，层数高（20层左右），墙体边界不规则凹凸明显；(2) 光谱特征：拍摄视角倾斜导致墙体与墙身色调不一致，建筑阴影长；(3) 空间特征：单体呈点状分布，单体之间间距较大，多个单体组合呈簇状分布；单体周围一般有停车场、花园、植被等附属地物存在。

砌体结构：(1) 几何特征：单体面积较大，多为6层以下建筑，边界较为规则，边缘较为整齐；(2) 光谱特征：建筑阴影较长；(3) 空间特征：单体间距较窄，多个单体在空间上形成大面积规整建筑群。

砖石及其他结构：(1) 几何特征：单体较低

矮，面积较小。(2) 光谱特征：无阴影特征或阴影特征不明显。(3) 空间特征：分布情况较为极端，即零星分布或者大面积密集分布。部分会配有院落，院落多伴有杂物。

为了获得高度准确的地面真实值，我们进行了质量控制流程。除了严格按照以上标准进行解译之外，我们还制定了几项额外规定。它们包括：对于相邻且边界清晰的建筑物，每栋建筑物分别对应一个面状要素；对于边界模糊的相邻建筑，可以多个建筑对应一个面状矢量；对于拍摄到楼身的建筑物，需要统一标注楼顶；明显的凹凸边界、拐角都需要进行标注；对于建筑物之上的建筑物不需要单独标注。接下来，我们邀请数据检查员进一步检查地面真实值的质量，包括建筑物的边界和细粒度类别。最后，邀请遥感图像领域的专家检查遥感图像和对应地面真实值的质量，最终得到建筑物实例个数超过26万个，示例结果如图6所示。

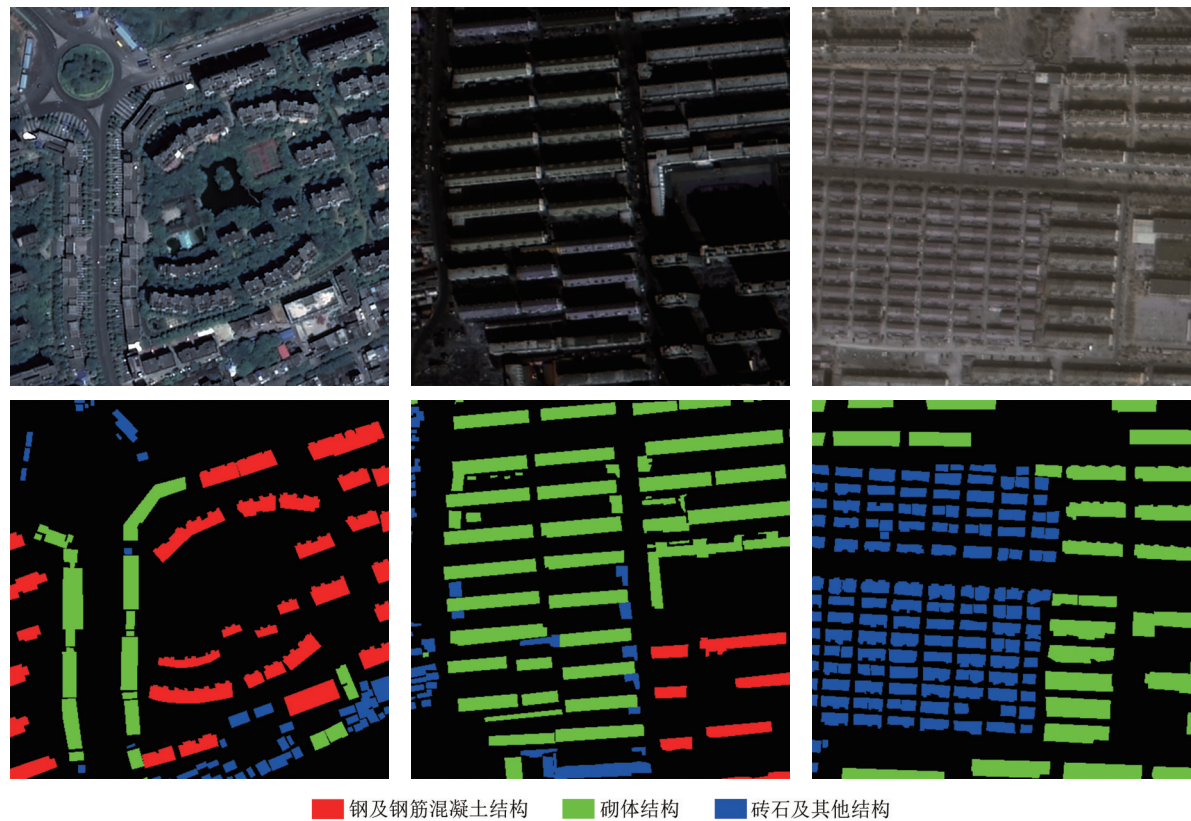


图6 地面真实值示例
Fig. 6 Example of ground truth values

4.2 实地考察验证

为了检验地面真实值的精度，我们进行了实

地调查。我们对万宁县、松北区、迭部县进行了实地调查。调查选用的设备主要是GPS接收仪、

手持相机以及卷尺, GPS用于记录建筑物的地理坐标, 手持相机用于记录建筑物外观以及所处环境, 卷尺用于测量建筑物的粗略长宽高等要素。在高分二号影像上选择建筑物实例并做好标记, 根据提前规划好的区域前往指定地区实地调研, 测量建筑物的高度、占地面积等信息, 咨询建筑物的实际用途等信息, 观测并记录建筑物的结构材料等信息, 同时拍摄建筑物正面、背面以及左右侧照片, 将以上信息记录在建筑物调查信息采集表对应位置。经过内业整理后, 实地调查建筑物数量96栋, 目视解译数据在细粒度类别上与实地调查不相符的仅有4栋, 准确率为95.83%。足以说明特征集中专家目视解译地面真实值数据是高度可信的。

5 特征集产品描述

高分辨率多光谱遥感影像细粒度建筑物特征集, 分为光谱、纹理、边缘、指数4类, 共有17个具体特征成分, 分别为蓝波段光谱、绿波段光谱、红波段光谱、近红外波段光谱、对比度、相异性、同质度、相关性、角二阶矩、局部二值模式、方向梯度直方图、Sobel边缘强度、Canny边缘强度、形态学建筑物指数、阴影指数、归一化植被指数、归一化水体指数。每个特征共有11005张512×512尺寸的样本切片, 所有成分分辨率均为0.8 m。

特征集内容主要包括(1)存储为tiff格式的11005张×17个特征集主要特征成分图;(2)存储为tiff格式的11005张样本解译地面真实值及存储

为txt格式的地面真实值介绍;(3)以6:2:2为划分比例的训练集、验证集以及测试集文件名汇总文本文档;数据总量约为191 GB。

6 数据实验与分析

针对本文的特征集, 利用深度学习方法进行了细粒度建筑物提取实验, 并对结果进行了简要分析。

卷积神经网络CNN(Convolutional Neural Network)具有局部感知和参数共享等多项优点, 在图像处理任务中表现优异。自注意力模型Transformer近年来由研究者将其从自然语言处理领域迁移到图像处理领域, 由于其强大的全局信息捕获能力, 表现出了优越的性能。我们分别以ConvNext(Liu等, 2022)和SwinTransformer(Liu等, 2021)为骨干网络, 以UperNet(Xiao等, 2018)作为解码器进行了两类深度学习模型的基线实验。均选用了最小参数量的版本作为实验实际骨干网络, 来提升实验效率。为初步探究纹理、边缘、指数特征是否具有正向作用, 增加了只输入光谱特征以及输入光谱、纹理、边缘、指数4种特征的对比实验。为综合考虑精确率和召回率, 使用F1分数定量评判不同结构类别建筑物的提取精度。实验结果如表3所示, 其中F1-1表示钢及钢筋混凝土结构建筑物的F1分数, F1-2代表砌体结构建筑物的F1分数, F1-3表示砖石及其他结构建筑物的F1分数, F1-B代表将3种子类建筑物看作一个大类建筑物时的F1分数。

表3 深度学习实验结果

Table 3 Experimental results of deep learning

模型	输入	F1-1/%	F1-2/%	F1-3/%	F1-B/%
UperNet-ConvNext	光谱特征	82.02	82.53	80.82	86.06
UperNet-ConvNext	光谱、纹理、边缘、指数特征	82.62	83.13	81.33	86.52
UperNet-SwinTransformer	光谱特征	81.81	82.56	80.56	86.14
UperNet-SwinTransformer	光谱、纹理、边缘、指数特征	81.60	82.93	81.23	86.54

实验结果表明, CNN和Transformer支持的深度模型在MFBFS特征集上取得的效果基本相当, F1-B分别达到了86.52%和86.54%。3个不同结构的建筑物子类的F1分数均在80%—84%, 多个类别涉及到类间竞争, 精度低于F1-B属于正常现象。在模型输入端增加了纹理、边缘、指数特征比仅输入光谱特征的F1分数要高, 但是最高仅高

出了0.46%, 如何更好的使用除光谱之外的特征, 进一步提升模型的上限是下一步考虑的重点。

MFBFS特征集包含了超过26万个实例各式各样的建筑物, 覆盖21个区县(包括农村和城镇)的重点区域, 具有较强的普适性。由于在结构角度分为了钢及钢筋混凝土、砌体、砖石及其他3种类别, 比常规建筑物提取难度更大。部分钢及钢

钢筋混凝土建筑物与砌体建筑物外观相似性高, 难以有效区分。而砖石及其他建筑物大多分布在农村, 较小的面积和较多的建筑物附属杂物的影响导致砖石及其他建筑物提取效果较差。不过, 该数据集的提出, 可以使更多的研究人员在多特征融合、建筑物细粒度以及弱小建筑目标提取方面进行更多的科学探索, 进一步推动建筑物提取的技术发展。

7 结 论

高分辨率多光谱遥感影像细粒度建筑物特征集 MFBFS 以国产高分二号多光谱遥感影像为数据源, 利用规范的预处理方案和高效的特征生成算法, 结合严格的地面真实值控制流程以及同步开展地面实地调研验证进行构建。该特征集空间分辨率为 0.8 m, 覆盖面积为 3668 km², 包含多光谱、纹理、边缘和指数 4 大类蓝波段、绿波段、红波段、近红外波段、对比度、相异性、同质度、相关性、角二阶矩、局部二值模式、方向梯度直方图、一阶边缘特征和多阶边缘特征、建筑物指数、阴影指数、植被指数和水体指数 17 种特征分量, 每个特征分量包含 11005 张 512×512 尺寸的特征图。在制作地面真实值过程中, 将建筑物按照结构属性分为钢及钢筋混凝土结构、砌体结构、砖石和其他结构, 并为每个类别制定了解译标准。在经过质量控制和实地验证之后, 得到超过 26 万个建筑物实例。经过初步的深度学习实验, 证明了该特征集的有效性, 下载链接为 <https://github.com/WangZhenqing-RS/MFBFS> [2023-12-11], 希望能够促进国产高分遥感建筑物提取的研究进展。团队后续将继续对 MFBFS 进行扩充, 欢迎感兴趣研究团队参与并提出宝贵意见。

参考文献 (References)

Geiß C, Pelizari P A, Marconcini M, Sengara W, Edwards M, Lakes T and Taubenböck H. 2015. Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 104: 175-188 [DOI: 10.1016/j.isprsjprs.2014.07.016]

Huang X and Zhang L P. 2011. A multidirectional and multiscale morphological index for automatic building extraction from multi-spectral geospatial imagery. *Photogrammetric Engineering and Remote Sensing*, 77(7): 721-732 [DOI: 10.14358/PERS.77.7.721]

Huang X L, Ren L B, Liu C L, Wang Y X, Yu H F, Schmitt M, Hänsch

R, Sun X, Huang H and Mayer H. 2022. Urban building classification (UBC)-a dataset for individual building detection and classification from satellite imagery//*Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. New Orleans: IEEE: 1412-1420 [DOI: 10.1109/CVPRW56347.2022.00147]

Ji S P, Wei S Q and Lu M. 2019. Fully convolutional networks for multisource 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 J X, Zhao S, Jin H, Li Y F and Guo Y. 2019. A method of combined texture features and morphology for building seismic damage information extraction based on GF remote sensing images. *Acta Seismologica Sinica*, 41(5): 658-670 (李金香, 赵朔, 金花, 李亚芳, 郭寅. 2019. 结合纹理和形态学特征的高分遥感影像建筑物震害信息提取. *地震学报*, 41(5): 658-670) [DOI: 10.11939/jass.20230067]

Liu Y H, So E, Li Z Q, Su G W, Gross L, Li X L, Qi W H, Yang F, Fu B, Yalikun A and Wu L J. 2020. Scenario-based seismic vulnerability and hazard analyses to help direct disaster risk reduction in rural Weinan, China. *International Journal of Disaster Risk Reduction*, 48: 101577 [DOI: 10.1016/j.ijdr.2020.101577]

Liu Z, Lin Y T, Cao Y, Hu H, Wei Y X, Zhang Z, Lin S and Guo B N. 2021. Swin transformer: Hierarchical vision transformer using shifted windows//*Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 9992-10012 [DOI: 10.1109/ICCV48922.2021.00986]

Liu Z, Mao H Z, Wu C Y, Feichtenhofer C, Darrell T and Xie S N. 2022. A convNet for the 2020s//*Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans: IEEE: 11966-11976 [DOI: 10.1109/CVPR52688.2022.01167]

Maggiori E, Tarabalka Y, Charpiat G and Alliez P. 2017. Can semantic labeling methods generalize to any city? The Inria aerial image labeling benchmark//*Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium*. Fort Worth: IEEE: 3226-3229 [DOI: 10.1109/IGARSS.2017.8127684]

Mnih V. 2013. *Machine Learning for Aerial Image Labeling*. Toronto: University of Toronto.

Nosrati M S and Saeedi P. 2009. A novel approach for polygonal rooftop detection in satellite/aerial imageries//*2009 16th IEEE International Conference on Image Processing*. Cairo: IEEE: 1709-1712 [DOI: 10.1109/ICIP.2009.5413641]

Pesaresi M and Gerhardinger A. 2011. Improved textural built-up presence index for automatic recognition of human settlements in arid regions with scattered vegetation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(1): 16-26 [DOI: 10.1109/JSTARS.2010.2049478]

Qi W H, Su G W, Sun L, Yang F and Wu Y. 2017. "Internet+" approach to mapping exposure and seismic vulnerability of buildings in a context of rapid socioeconomic growth: a case study in Tangshan, China. *Natural Hazards*, 86(1): 107-139 [DOI: 10.1007/

- s11069-016-2581-9]
- Shermeyer J, Hogan D, Brown J, Van Etten A, Weir N, Pacifici F, Hansch R, Bastidas A, Soenen S, Bacastow T and Lewis R. 2020. SpaceNet 6: Multi-sensor all weather mapping dataset//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle: IEEE [DOI: 10.1109/CVPRW50498.2020.00106]
- Tao W C, Zhang Y, Xie Z X, Wang X S, Dong Y, Zhang M Z, Su W, Li J Y, Xuan F. 2022. Intelligent recognition of corn residue cover area by time-Series Sentinel-2A images. *Spectroscopy and Spectral Analysis*, 42(6): 1948-1955 (陶万成, 张颖, 谢苾萱, 王新盛, 董懿, 张明政, 苏伟, 李佳雨, 轩阜. 2022. 基于时序Sentinel-2A影像的玉米秸秆覆盖区智能识别研究. *光谱学与光谱分析*, 42(6): 1948-1955) [DOI: 10.3964/j.issn.1000-0593(2022)06-1948-08]
- Wang Z Q, Zhou Y, Wang F T, Wang S X, Qin G, Zou W J, Wang Z C, Liu S M and Zhu J F. 2023a. Fine-grained building extraction with multispectral remote sensing imagery using the deep model. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1-13 [DOI: 10.1109/TGRS.2023.3327370]
- Wang Z Q, Zhou Y, Wang F T, Wang S X, Qin G, Zou W J and Zhu J F. 2023b. A multi-scale edge constraint network for the fine extraction of buildings from remote sensing images. *Remote Sensing*, 15(4): 927 [DOI: 10.3390/rs15040927]
- Wang Z Q, Zhou Y, Wang S X, Wang F T and Xu Z Y. 2021. House building extraction from high-resolution remote sensing images based on IEU-Net. *National Remote Sensing Bulletin*, 25(11): 2245-2254 (王振庆, 周艺, 王世新, 王福涛, 徐知宇. 2021. IEU-Net 高分辨率遥感影像房屋建筑物提取. *遥感学报*, 25(11): 2245-2254) [DOI: 10.11834/jrs.20210042]
- 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 (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) [DOI: 10.11834/jrs.20220296]
- Xia J S, Yokoya N, Adriano B, Zhang L C, Li G Q and Wang Z G. 2021. A benchmark high-resolution GaoFen-3 SAR dataset for building semantic segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 5950-5963 [DOI: 10.1109/JSTARS.2021.3085122]
- Xiao T T, Liu Y C, Zhou B L, Jiang Y N and Sun J. 2018. Unified perceptual parsing for scene understanding//Proceedings of the 15th European Conference on Computer Vision. Munich: Springer: 432-448 [DOI: 10.1007/978-3-030-01228-1_26]
- Xu Z Y, Zhou Y, Wang S X, Wang L T, Li F, Wang S C and Wang Z Q. 2020. A novel intelligent classification method for urban green space based on high-resolution remote sensing images. *Remote Sensing*, 12(22): 3845 [DOI: 10.3390/rs12223845]
- You Y F, Wang S Y, Wang B, Ma Y X, Shen M, Liu W H and Xiao L. 2019. Study on hierarchical building extraction from high resolution remote sensing imagery. *Journal of Remote Sensing (in Chinese)*, 23(1): 125-136 (游永发, 王思远, 王斌, 马元旭, 申明, 刘卫华, 肖琳. 2019. 高分辨率遥感影像建筑物分级提取. *遥感学报*, 23(1): 125-136) [DOI: 10.11834/jrs.20197500]
- Zhang L S, Shi P J, Wang J A and Zhu L. 1995. Regionalization of natural disasters in China. *Journal of Beijing Normal University (Natural Science)*, 31(3): 415-421.
- Zhang T W, Zhang X L, Ke X, Liu C, Xu X W, Zhan X, Wang C, Ahmad I, Zhou Y, Pan D C, Li J W, Su H, Shi J and Wei S J. 2022. HOG-ShipCLSNet: A novel deep learning network with hog feature fusion for SAR ship classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5210322 [DOI: 10.1109/TGRS.2021.3082759]
- Zhang Y Q and Ren H R. 2023. Remote sensing extraction of paddy rice in Northeast China from GF-6 images by combining feature optimization and random forest. *National Remote Sensing Bulletin*, 27(9): 2153-2164 (张悦琦, 任鸿瑞. 2023. 融合特征优选与随机森林算法的GF-6影像东北一季稻遥感提取. *遥感学报*, 27(9): 2153-2164) [DOI: 10.11834/jrs.20221338]
- Zhao C, Liang Y Y, Feng H, Wang Z, Yu Q, He J Q. 2023. Plastic-mulched farmland recognition in loess plateau based on Sentinel-2 Remote-sensing Images. *Journal of Agricultural Machinery*, 54(8): 180-192 (赵成, 梁盈盈, 冯浩, 王钊, 于强, 何建强. 2023. 基于Sentinel-2 遥感影像的黄土高原覆膜农田识别. *农业机械学报*, 54(8): 180-192) [DOI: 10.6041/j.issn.1000-1298.2023.08.017]

MFBFS: A fine-grained building feature set for high-resolution multispectral remote sensing images

WANG Zhenqing^{1,2}, ZHOU Yi¹, WANG Futao^{1,2}, WANG Shixin¹, GAO Guorui^{1,2},
ZHU Jinfeng¹, WANG Ping³, HU Kailong³

1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;

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

3. National Disaster Reduction Center of the Ministry of Emergency Management, Beijing 100124, China

Abstract: Building information extraction from remote sensing images plays an essential role in urban information management and

disaster prevention and mitigation. This study establishes a fine-grained building feature set, namely, MFBFS, for high-resolution multispectral remote sensing images. MFBFS uses the domestically produced Gaofen-2 multispectral remote sensing images as data source and selects 21 districts and counties with concentrated buildings in various disaster zones in China, covering 3668 km² as the study area. These regions include Yongjia County, Xuwen County, and Wanning City in the southeastern coastal disaster belt; Ning'an City, Kaiyuan City, Laiyuan County, Shouguang City, Xinxiang County, Lujiang County, Hengdong County, and Songbei District in the eastern disaster belt; Daning County, Enshi City, Tengchong City, and Shuicheng County in the central disaster belt; Kashgar City, Yizhou District, and Pingluo County in the northwest disaster belt; and Diebu County, Yushu City, and Milin County in the Qinghai – Tibetan disaster belt. To obtain high-quality and high-resolution remote sensing images, a series of preprocessing procedures was applied to the Gaofen-2 images. Initially, poor-quality images were removed, followed by radiometric and orthorectification corrections on multispectral and panchromatic images, respectively. Finally, the panchromatic images were fused to enhance the spatial resolution of the multispectral images, resulting in a spatial resolution of 0.8 m. Seventeen feature components were generated from four perspectives: spectral, texture, edge, and index. Spectral features include features from the blue, green, red, and near-infrared bands. Texture features consist of contrast, dissimilarity, homogeneity, correlation, angular second moment, local binary pattern, and histogram of oriented gradients. Edge features comprise first-order and multi-order edge characteristics. Index features include building, shadow, vegetation, and water indexes. MFBFS encompasses over 260000 building instances, ensuring high intra-class diversity in terms of size, shape, color, orientation, background, and structural type. These instances are classified into three structural types, namely, steel and reinforced concrete, masonry, and block stone structures, significantly reflecting the abilities of buildings to resist disasters and their usable lifespans. The fine-grained design will cause the task of extracting buildings through remote sensing to play a greater role, particularly in pre-disaster loss prediction and post-disaster loss assessment in the disaster field. Rigorous quality control processes and field inspections were conducted to ensure the high accuracy of ground truth values. This procedure involved adherence to interpretation standards and inviting data inspectors and remote sensing image experts to assess the quality of remote sensing images and corresponding ground truth values. Ultimately, 191 GB of high-quality feature and label data were obtained. Each of the 17 feature components comprises 11005 512×512-sized feature maps with a spatial resolution of 0.8 m, uniformly expanded to a value range of [0,1]. Initial deep learning experiments demonstrate the effectiveness of MFBFS. This feature set, available for download at <https://github.com/WangZhenqing-RS/MFBFS>, provides robust data support for fine-grained building structure extraction research and promotes the development of domestic high-resolution remote sensing data applications.

Key words: high-resolution remote sensing, multispectral imagery, fine-grained categories, building extraction, feature sets

Supported by National Key Research and Development Program of China (No. 2021YFB3901201)