

引用格式: Gao Ziwei, Sun Weiwei, Cheng Penggen, *et al.* Identify Urban Functional Zones Using Multi Feature Latent Semantic Fused Information of High-spatial Resolution Remote Sensing Image and POI Data[J]. Remote Sensing Technology and Application, 2021, 36(3): 618-626. [高子为, 孙伟伟, 程朋根, 等. 融合高分辨率遥感影像和POI数据的多特征潜在语义信息用于识别城市功能区[J]. 遥感技术与应用, 2021, 36(3): 618-626.]
doi:10.11873/j.issn.1004-0323.2021.3.0618

融合高分辨率遥感影像和POI数据的多特征潜在语义信息用于识别城市功能区

高子为¹, 孙伟伟², 程朋根¹, 杨刚², 孟祥超³

(1. 东华理工大学 测绘工程学院, 江西 南昌 330013;

2. 宁波大学 地理与空间信息技术系, 浙江 宁波 315211;

3. 宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要: 准确识别和划分城市功能区对合理规划城市发展和解决城市问题具有重要作用。遥感影像拥有丰富的光谱纹理特征但难以表征建筑物的社会经济属性, 而社交媒体数据等城市数据为城市研究与应用提供了丰富的数据资源, 补充了遥感影像所缺失的建筑物内在特征。融合高分辨率遥感影像和POI数据的多特征信息, 利用嵌入主题模型挖掘其潜在语义信息识别城市功能区。以宁波市2个典型的商业区为研究区设计3个实验, 验证该方法的效果和性能。研究结果表明: 该方法能够取得85.67%和85.78%的分类精度, 并准确识别出城市功能区。同时, 光谱、纹理、几何和POI特征组合的多特征信息能够明显提升城市功能区的识别精度, 并且嵌入主题模型能够更好地挖掘多特征的高层次潜在语义信息, 效果明显优于pLSA、LDA和STM 3种主流模型。

关键词: 高分辨率遥感影像; POI; 城市功能区识别; 多特征信息; 主题模型

中图分类号: TP75 **文献标志码:** A **文章编号:** 1004-0323(2021)03-0618-09

1 引言

随着城市化的飞速发展, 城市空间以“增量扩张”为主的发展模式难以维持, 精细化的空间“存量挖潜”成为城市发展的新趋势。城市功能区是存量空间的重要组成部分, 厘清城市功能区布局对加强存量利用和优化城市产业结构具有重要作用, 包括提升城市产业集聚效能、提高土地利用效率、辅助政府部门管理决策和推进城市化建设的可持续发展等。

遥感技术具有重访周期短、成本低、覆盖范围广等优点^[1], 尤其高分辨率遥感影像如IKONOS和

WorldView-II具有亚米级的空间分辨率, 获取的地物空间细节信息更加清晰丰富, 为准确识别城市土地利用和功能区提供了可能^[2]。因此, 有学者开始利用IKONOS高分辨率遥感影像来提取城市土地利用变化信息^[3], 也有学者利用WorldView-II高分辨率遥感影像来划分城市功能区^[4]。然而, 高分辨率遥感影像仅使用了地物低层次的灰度或纹理特征, 无法准确识别较为相似的不同类别地物。后来, 有学者考虑挖掘高分辨率影像的高层次语义特征信息来缩小场景分类的“语义差距”^[5], 通过构建词袋(Bag of words, BoW)模型来融合光谱、纹理和Scale-Invariant Feature Transform(SIFT)等多特征

收稿日期: 2020-09-15; 修订日期: 2021-04-13

基金项目: 国家重点研发计划项目(2017YFB0503704), 国家自然科学基金项目(41861052、41861062、41971296、41671342、41801256), 浙江省自然科学基金项目(LR1901D0001、LQ18D010001)。

作者简介: 高子为(1996—), 男, 浙江宁波人, 硕士研究生, 主要从事城市遥感环境监测方面的研究。E-mail: nbgaiziwei@sina.cn

作者简介: 孙伟伟(1985—), 男, 河南巩义人, 博士, 教授, 主要从事地理信息系统和遥感理论、方法及应用研究。E-mail: sunweiwei@nbu.edu.cn

信息,提高土地利用的分类精度^[6]。然而,遥感影像所提取的特征仅代表地物自然物理属性,仍无法完全对接经济社会属性明显的城市土地利用类型。

研究表明,社交媒体数据能够表征人类经济社会活动的内在特征,填补高分辨率遥感影像数据无法精细描述微观个体与用地空间内在联系的空白,有助于识别城市功能区^[7]。社交媒体数据在城市功能区识别中已得到了许多应用,如车辆轨迹辅助识别城市区域功能^[8]、手机和社交媒体的微博签到数据用于城市动态描述^[9]与城市功能区划分^[10]、带有地理标记的社交媒体和街景图像用于城市功能研究^[11]、兴趣点(Point of Interest, POI)数据用于城市土地利用类型分类^[12]。考虑到地物的自然物理特性与社会经济活动信息的互补优势,学者开始研究融合高分辨率遥感图像和社交媒体数据,以提升城市功能区识别的准确度。如Liu等^[7]提出了一种基于Latent Dirichlet Allocation(LDA)主题模型融合了高分辨率遥感数据和POI数据的场景分类框架对城市功能分区。然而,LDA的分析处理性能依赖于构建的多个特征的严格精简词汇表,对于大型词汇表难以保证模型的泛化预测能力^[13]。后来,Zhang等^[14]构建稀疏主题模型,通过度量区域中模式数量及对应的特征比例,整合高分辨率遥感影像和POI数据的多特征语义信息,尝试改善城市功能区的识别精度,但是稀疏主题模型在预测会考虑一

些权重较小的特征信息,从而干扰预测结果的正确性,导致识别精度的下降。为此,研究引入词嵌入主题模型(Embedded Topic Model, ETM),融合高分辨率影像和POI数据的多特征信息,挖掘融合数据的内在高层语义知识,提升城市功能区的识别准确性。相比LDA和稀疏主题模型,ETM词嵌入主题模型是一种整合传统主题模型与词嵌入方法的新型文档生成模型。该模型同时包含可解释的潜在语义结构和主题特征,对大型词汇表和复杂城市区域的处理能力较强,在主题质量和预测性能方面优于现有的LDA和稀疏主题模型^[15]。基于宁波市2个典型的新旧城市功能区,利用国产高分二号高分辨率影像和POI数据来设计一系列实验,综合验证本文方法的效果和性能。

2 研究方法和原理

图1为研究方法的技术流程图。首先,对遥感影像按照规则格网进行分割,提取每一块格网中影像特征和POI特征信息作为具有一定表达能力的视觉单词。在此基础上,利用词袋模型建立影像和POI特征信息的高层次语义表达,提升视觉特征的表达能力。其次,利用概率主题模型提取上述多特征信息的潜在语义特征。最后,选择迭代选择最优和最优价值的样本用于训练支持向量机模型,对城市功能区进行识别。

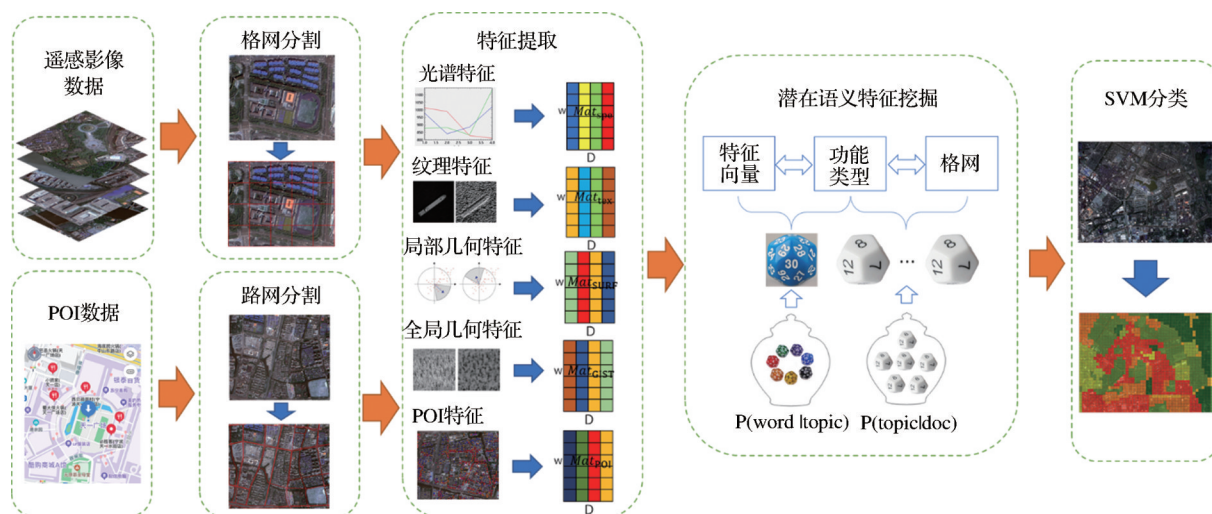


图1 技术流程图

Fig.1 Methodology in this study

2.1 高分辨率影像和POI数据的多特征提取

利用高分辨率影像和POI数据来提取多特征信息,以刻画每个影像格网单元的视觉单词。高分辨率遥感影像的特征提取包括光谱、纹理和空间几

何特征信息。POI数据的特征提取包括功能类型和经纬度信息。

(1)光谱特征。光谱特征反映地物表面的物理属性^[16]。本文采用波段均值与标准差来描述格网

单元的光谱特征 $spe = [mean, std]$, 公式如(1)~(2)所示。

$$mean = \frac{\sum_{i=1}^n v_i}{n} \quad (1)$$

$$std = \sqrt{\frac{\sum_{i=1}^n (v_i - mean)^2}{n}} \quad (2)$$

其中: n 为每个格网中的像元数量, v_i 为格网中第 i 个像元的灰度值。

(2) 纹理特征。纹理特征反映地物表面周期性变化的表面结构排列属性^[17]。研究采用局部二值模式(Local Binary Pattern, LBP)来提取各个格网单元中影像的局部纹理特征。LBP是一种描述图像纹理特征的算子, 具有灰度不变性和旋转不变性特点^[18]。

首先定义一个 3×3 的窗口, 以窗口中心像素的灰度值作为阈值, 比较该阈值与相邻的8个像素的灰度值。计算公式如(3)所示:

$$LBP(x_c, y_c) = \sum_{p=1}^8 s(I(p) - I(c)) \times 2^p \quad (3)$$

其中: $I(p)$ 表示除窗口像元之外的第 p 个像素点的灰度值; $I(c)$ 为中心像素的灰度值。若 $s(I(p) - I(c)) \geq 0$, 则为1; 否则为0。遍历窗口内所有像元, 得到一个8位二进制数作为中心像素点的LBP特征值 lbp 。

(3) 几何特征。研究采用 Speeded Up Robust Features (SURF) 算子和 Evaluation of GIST (GIST) 算子来分别提取影像的局部和全局几何特征。SURF是一种局部特征描述子, 能够克服仿射变换、光照变化和三维视点变化等传统几何特征描述的不足, 提取影像局部几何特征^[19]。SURF比SIFT算子具有更低的数据维度和更优的运算效率。

首先为保证尺度不变性, 使用不同的模板尺寸对格网进行盒式滤波, 构建多尺度斑点响应的金字塔图像, 并利用公式(4)的Hessian矩阵判别式来寻找不同尺度下的特征点位置。

$$\det(H) = \frac{\partial^2}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 \quad (4)$$

其中: x, y 分别为像素点的横纵坐标, $f(x, y)$ 为该点的像素值。当Hessian矩阵判别式取得局部最大值时, 该点即为特征点。

其次对每个特征点分配一个主方向, 以保证旋

转不变性。定义一个以特征点作为圆心, 0.2 弧度作为步长的 60° 的扇形滑动窗口, 对扇形窗口进行 Haar 小波响应运算, 统计窗口中每个点在水平和垂直方向的小波响应, 并累加求和 (m_w, θ_w) , 特征点的主方向即为小波特征总和最大的方向 θ , 公式如(5)~(7)所示。

$$m_w = \sum_w dx + \sum_w dy \quad (5)$$

$$\theta_w = \arctan \left(\frac{\sum_w dx}{\sum_w dy} \right) \quad (6)$$

$$\theta = \theta_w | \max \{m_w\} \quad (7)$$

最后, 以特征点为中心, 沿着主方向将图像划分为 4×4 的子块, 计算每个子块4个方向上的 Haar 值作为特征向量, 得到一个64维特征向量组成的 SURF 描述子, 并统计格网中的所有特征点描述子作为 SURF 特征 $surf$ 。

GIST 是一种空间包络特征描述符, 从自然度、开放度、粗糙度、膨胀度和险峻度5个方面描述影像的全局特征信息, 激活对场景类别的记忆表征, 具有较好的识别能力^[20]。首先用公式(8)~(10)的 Gabor 滤波器对格网在4个尺度、8个方向上进行卷积, 得到一个32维特征图。

$$g(x, y; \lambda, \omega, \psi, \sigma, \gamma) = \quad (8)$$

$$\exp \left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2} \right) \cos \left(2\pi \frac{x'}{\lambda} + \psi \right)$$

$$x' = x \cos \omega + y \sin \omega \quad (9)$$

$$y' = -x \sin \omega + y \cos \omega \quad (10)$$

其中: λ 为波长参数, ω 为平行条带方向, ψ 为相位偏移, γ 为空间纵横比, σ 为函数标准差。其次, 分别将32张特征图平均分割成16个区域。最后, 计算每个区域均值并进行拼接, 得到一个512维向量作为该格网的 GIST 特征 $gist$ 。

(4) POI 特征。POI 数据是包含名称、地址、功能和经纬度等空间属性信息的特征点集, 蕴含着丰富的人文经济特征以及自然特征, 能够揭示城市土地的利用功能^[21]。根据道路来分割研究区域, 计算每个路网区块中的第 i 种 POI 类型比例 P_i , 并将其赋予该区块中的每一个影像格网单元。每个影像格网单元的 POI 特征向量表示为 $poi = \{P_1, P_2, \dots, P_i\}$ 。

2.2 多特征融合的 BoW 字典构建

采用词袋模型来构建视觉语义字典, 获取多源数据中特征信息的词汇表达向量。BoW 模型起源

于文档分析,利用文档集中单词和主题的聚类分布来构建语义字典^[22]。应用于遥感领域,该模型将格网单元影像看作文档集,城市功能看作主题,单元格网中的各类特征看作单词。即任一格网单元可用城市功能的分布来表示,任意城市功能可用特征分布来表示。

BoW模型构建语义字典的流程如下。首先,对影像中第*i*个格网单元提取的多个特征 $\{spe_i, lbp_i, sur_i, gist_i, poi_i\}$ 进行K-means聚类,每个聚类中心构成一个单词,所有聚类中心组成一个单词表。其次,计算该格网单元的特征值与各聚类中心的距离,将特征值分配给单词表中距离最近的单词。最后,统计各个单词的词频,生成一个高维向量并拼接,得到该格网单元中特征词汇的表达向量 $\{word_{spe}, word_{lbp}, word_{surf}, word_{gist}, word_{poi}\}$ 。

2.3 潜在语义特征挖掘

采用ETM来挖掘多特征词汇表征中隐藏的高层语义结构信息挖掘高层次潜在语义信息。在传统的主题模型中,每个主题在词汇表上是一个完整的分布,而ETM采用深度学习模型构建,将主题和单词共同构建向量空间,融合Hierarchical Softmax实现的词嵌入方法,为每个主题构建其在语义空间的分布表征,并采用Logistic-normal分布简化推断算法来提高运算效率^[15]。

ETM推测主题的过程如下。首先,随机给每篇文档中的每个单词分配*k*个主题中的一个主题。其次,对于每篇文档*D*,重新分配单词的主题。接下来,假设除了当前文档之外的所有主题分配符合logistic-normal分布,计算两个概率:①当前文档*D*中的单词被分配到主题*t*的概率;②所有文档中被分配到主题*t*的单词中单词*w*的概率。然后,基于这两个概率的乘积和词向量与主题分布向量的关系分配给单词*w*一个新的主题。不断重复上述步骤,最终达到稳定状态,得到文档的主题分布。ETM主题模型将BoW字典中的特征合集进行挖掘,提取每个格网中多源特征的潜在语义特征,最终通过训练支持向量机(Support Vector Machine, SVM)分类器实现对各个格网单元的功能识别。

3 研究区域和数据源

3.1 研究区域

宁波市位于中国东南沿海,浙江宁绍平原东端,东有舟山群岛为天然屏障,北邻杭州湾,西接绍

兴市。宁波既是长江三角洲南翼经济中心和浙江省经济中心之一,经济社会发展水平高,第三产业发展,城市功能类型丰富,因而本文选取宁波作为研究城市。考虑到POI数据的有限性和分布情况,实验选取宁波市2个新旧商业中心作为代表性研究区域,如图2所示。研究区1包含镇海区万科广场及周边区域,面积约为15.3 km²。万科广场于2016年建成,是宁波市新兴发展起来的商业中心,是集购物、休闲和观光为一体的商业综合体。研究区2包含海曙区天一广场及周边区域,面积约13 km²,天一广场于2001年建成,是宁波最大的融休闲、商贸、旅游、餐饮、购物于一体的大型城市中心商业广场,拥有浓厚的商业文化特色。研究区1和2作为宁波市的新旧经济中心,高度集中了城市的商业、工业、居住和公共服务设施用地,区域内建筑分布密度高且外观相似,各类设施分布密集,POI数据丰富,适合作为本文的城市功能分区的实验区域。

3.2 数据源

实验采用GF-2卫星的遥感影像作为高分辨率影像数据源。研究区1和2的遥感影像的获取时间分别为2019年6月16日和4月18日,数据来源于地理空间数据云平台(<http://www.gscloud.cn/>)。行政区划数据来源于全球行政区域网(<http://www.gadm.org/>),路网数据来源于OpenStreetMap(<https://www.openstreetmap.org/>)。

POI数据来源于百度地图(<https://map.baidu.com/>),数据获取时间为2019年6月,包含名称、功能、地址和经纬度等4类属性。研究区1和2分别包含3 009和27 145个POI点。按照2011年中华人民共和国住房和城乡建设部颁布的新版《城市用地分类与规划建设用地标准》,同时考虑到城市功能的普遍性和研究区中所包含的地物类型,将格网功能划分为公共管理与公共服务用地(A)、商业服务设施用地(B)、绿地与广场用地(G)、工业用地(M)、居住用地(R)、水域(W)和道路(S)共7个类别。

有学者利用格网单元分割遥感影像划分城市功能区并取得了较高的精度^[10,23]。考虑到研究区大小以及POI分布密度,本文采用100 m×100 m大小的格网单元,并通过实地考察的方式标注格网单元的真实功能区类别。研究区1和2分别包含1 564和1 147个格网单元,各类功能区的真实样本信息如表1所示。

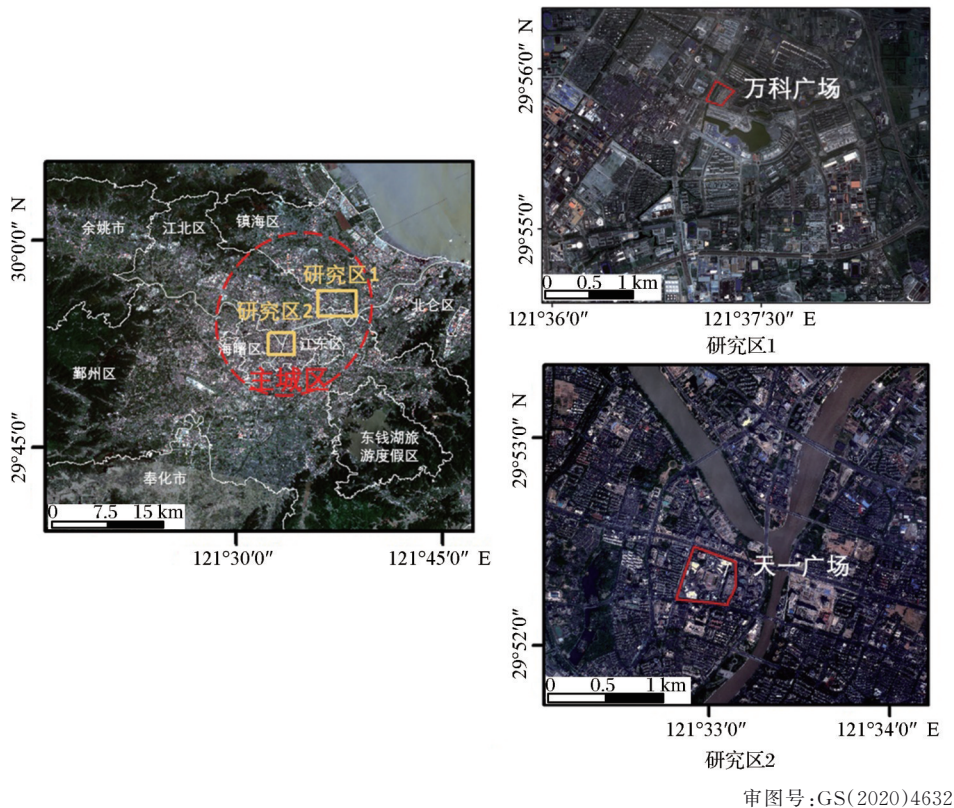


图2 研究区范围

Fig.2 Study area

表 1 各功能区的真实样本信息一览表
Table 1 The ground truth information of each functional area

类别	A	B	G	M	R	W	S	总计
研究区 1	230	421	255	169	439	8	42	1 564
研究区 2	132	525	80	83	167	135	25	1 147

3.3 预处理

为达到更好的识别效果,在提取影像多特征之前,对 GF-2 影像进行正射校正、辐射定标、图像融合、大气校正等预处理,得到 1 m 分辨率的高分辨率多光谱影像数据。在此基础上,根据研究区域范围对预处理后的影像进行裁剪,并采用 100 m×100 m 格网单元来分割研究区并提取多特征信息。

POI 数据虽然蕴含了能反映建筑物内部社会经济属性的语义信息,但并非所有 POI 数据都能辅助识别城市功能区,甚至可能造成干扰,如公厕、报刊亭、交通站点等,需要从原属数据中剔除。本文从以下几个步骤对 POI 数据进行预处理,首先对 POI 数据进行了筛选,排除无详细类别标识的数据和无详细坐标信息的数据。其次对 POI 类型按照所需类别进行重分类。最后将 POI 数据从百度坐标系纠偏为与遥感影像一致的 WGS84 坐标系。

4 实验与分析

设计了 3 类实验来综合验证提出的方法的性能和效果。实验 4.1 证明了该方法用于城市功能分区上的可行性。实验 4.2 剖析不同的特征组合对本文的 ETM 模型来挖掘高层次语义特征和识别城市功能区的性能影响;实验 4.3 比较了 3 种主流语义模型与本文 ETM 模型的识别效果。

4.1 融合多特征的城市功能识别实验

实验使用光谱特征、纹理特征、几何特征和 POI 特征,结合 BoW 模型和 ETM 模型来识别城市功能区。利用交叉验证原则,两个研究区域以均匀随机抽样的方式分别取 80% 的格网数据作为训练集,通过人工校核方式保证每种功能类型格网的数据都包含在训练集中。模型参数主题个数设置为 250,聚类中心设置为 600,SVM 分类器采用高斯核函数,通过 grid search 选取最优参数。城市功能分区识别结果的分类混淆矩阵如表 2 和表 3 所示。

两片研究区分别有 1 340 和 984 格网功能属性定义符合,识别结果精度分别达 85.67% 和 85.78%。公共管理与公共服务用地识别精度较低,主要因为该功能区中包含的服务型 POI 数量与类

型较少,且其外部物理特征存在多样性。商业服务业设施用地识别精度较高,主要原因是商业类型POI在空间中密集分布凸显了POI特征。居住用地识别效果相对较差,部分原因是居住用地中包含一些商业类型POI导致住宅类型POI所占比例较低,影响了识别效果。工业用地具有较高的识别精度,因为其带有鲜明的纹理和几何特征明显如蓝色屋顶和高大的烟囱。水域拥有较高识别精度是因为物理特征上的连续性以及光谱特征的独特性。道路识别精度相对较低,主要因为道路和建筑物顶部相似的物理性质,并且缺乏社交媒体数据辅助。

表2 研究区1的分类混淆矩阵

Table 2 Classification confusion matrix of study area 1

类别	A	B	G	M	R	W	S	总计
A	144	2	2	8	32	0	1	189
B	11	450	4	4	19	0	2	490
G	9	5	243	2	21	0	2	282
M	2	7	3	144	5	0	3	164
R	27	6	20	4	323	0	9	389
W	0	0	1	0	0	7	0	8
S	2	6	0	0	5	0	29	42
总计	195	476	273	162	405	7	46	1 564

4.2 不同特征的融合性能分析实验

采用实验来验证高分辨率遥感影像和POI数据中不同特征对高层次语义特征挖掘和功能识别的影响。实验中融合不同的特征组合,并利用ETM模型来挖掘高层次语义特征信息,并实现城市功能区识别的精度评价。采用的多特征包括POI

表3 研究区2的分类混淆矩阵

Table 3 Classification confusion matrix of study area 2

类别	A	B	G	M	R	W	S	总计
A	89	13	2	5	4	1	2	116
B	17	501	3	2	19	0	1	543
G	3	13	55	0	1	1	0	73
M	7	5	1	82	4	0	1	100
R	6	29	3	2	117	1	2	160
W	0	3	0	0	1	131	0	135
S	3	4	0	1	3	0	9	20
总计	125	568	64	92	149	134	15	1 147

特征“POI”,光谱特征“Spe”,光谱和纹理特征组合“Spe+Tex”,光谱、纹理和几何特征组合“Spe+Tex+SURF”和“Spe+Tex+SURF+GIST”,光谱、纹理、几何和POI特征组合“Spe+Tex+SURF+GIST+POI”。

为了验证每种特征对城市功能区识别是否有正向作用,使用不同的特征组合方式分别进行实验,并取10次的平均结果。实验中模型参数主题个数和聚类中心数分别设置为250和600。表4为使用ETM主题模型时不同的特征组合方法得到的功能区识别精度。与单独使用一类特征相比,融合多个特征进行分类能取得更好的结果。实验主要是对城市中建筑物功能类型进行分类,当仅使用光谱、纹理或几何特征时无法有效地挖掘建筑物内部功能属性,其分类精度结果较低。再加入POI特征后,两个研究区域的总体准确率分别提高了8.97%和8.88%。这表明融合多源数据的多特征信息可以大幅提高城市功能区识别的准确度。

表4 ETM模型中不同特征组合得到的分类精度对比(单位:%)

Table 4 Comparison of classification accuracy of different feature combinations in ETM Model (Unit:%)

研究区		Spe	POI	Spe+Tex	Spe+Tex+SURF	Spe+Tex+SURF+GIST	本文方法
1	OA	37.02	62.86	44.70	64.76	76.70	85.67
	Kappa	24.66	53.94	32.39	55.79	69.75	81.68
2	OA	39.26	67.13	49.78	66.96	76.90	85.78
	Kappa	26.72	58.04	38.31	57.38	70.24	80.03

4.3 ETM的语义特征挖掘性能分析实验

实验使用pLSA、LDA、STM 3种模型与ETM模型作对比,验证ETM模型的效果和性能。pLSA模型,利用似然函数来挖掘隐含变量,并利用EM算法(Expectation-Maximization Algorithm)进行求解,得到隐含变量的值,获得潜在语义特征^[24]。LDA模型将文档集中每篇文档的主题以概率分布的形式给出,并通过分析每篇文档中单词的主题分布推断

出文档的主题分布^[25]。STM模型是LDA的简化改进,将文档进行稀疏化表达,以克服LDA因词汇表过大导致运行效率过慢问题^[13]。

表5~6展示了4种主题模型用于研究区1和2的城市功能区的识别精度。实验中采用光谱、纹理、几何和POI特征组合,模型参数的主题数设置为250,K-means聚类中心为600。结果表明利用ETM主题模型提取潜在语义信息后对两个研究区

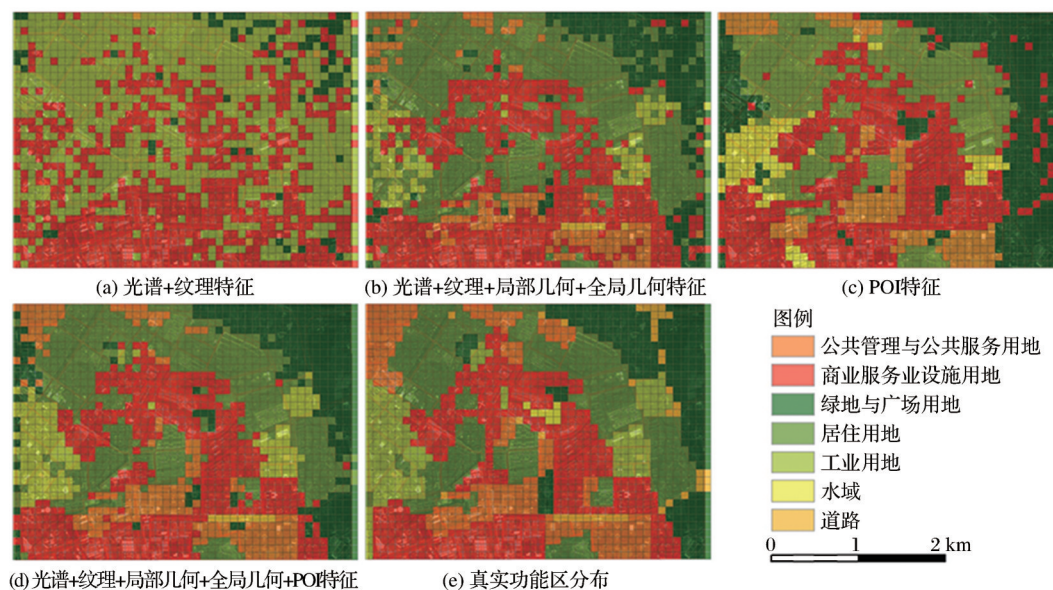


图 3 研究区 1 不同特征组合的识别结果图

Fig.3 Recognition results of different feature combinations in study area 1

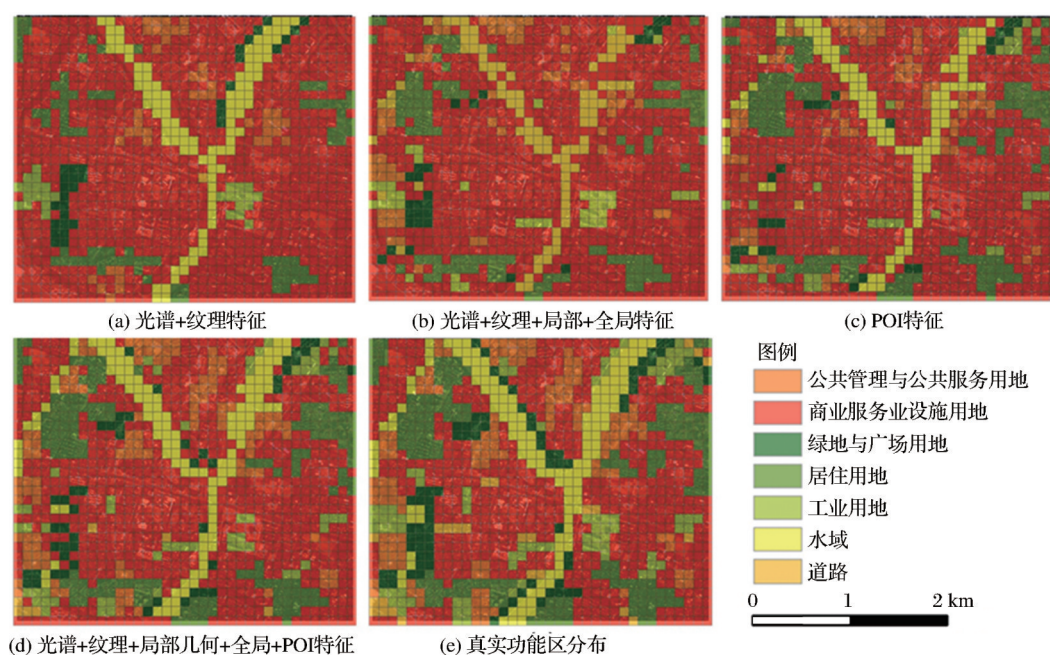


图 4 研究区 2 不同特征组合的识别结果图

Fig.4 Recognition results of different feature combinations in study area 2

的总体识别结果都优于其他 3 种主题模型。pLSA 模型效果不理想是因为文档数量过大时将导致其识别结果不佳。LDA 模型相较于 ETM 模型的识别能力较差,因为 Dirichlet 分布随机向量和各分量间仅仅存在弱相关性,而城市功能区的类型存在较强的相关性。STM 模型虽然有较快的运行速度,但仅仅在识别公共管理与公共服务用地时识别精度略高于其他几种主题模型。

而 ETM 模型在识别各类功能区时精度都较高,因为其良好的概率分布结构和词嵌入特性,在

处理大型词汇表时鲁棒性强。

5 结 语

基于 ETM 模型,提出了一种融合遥感影像和 POI 数据中多特征的潜在语义信息来识别城市功能区的方法,并采用宁波市 2 个典型的研究区来设计实验并进行方法验证。研究结果表明,该方法能够取得良好的城市功能区识别效果。同时,“光谱+纹理+局部几何+全局几何+POI”的多特征组合信息,明显优于传统的 POI 信息和其他多特征组

表 5 4 种语义模型用于识别研究区 1 中城市功能区的精度对比(单位: %)

语义模型	A	B	G	M	R	W	S	OA	Kappa
pLSA	60.00	85.99	80.00	81.07	67.43	62.50	45.24	78.32	71.46
LDA	61.74	89.55	82.35	82.84	71.07	75.00	57.14	81.2	74.68
STM	65.22	90.74	92.94	82.25	72.21	75.00	64.29	82.74	76.73
ETM	62.61	93.11	95.29	85.21	73.58	87.50	69.05	85.67	81.68

表 6 4 种语义模型用于识别研究区 2 中城市功能区的精度对比(单位: %)

语义模型	A	B	G	M	R	W	S	OA	Kappa
pLSA	59.85	89.14	61.25	89.16	63.47	90.37	28	80.47	72.19
LDA	62.88	90.67	62.5	86.75	67.07	95.56	28	80.99	73.53
STM	70.45	92.19	66.25	90.36	68.26	92.59	32	82.91	77.86
ETM	67.42	95.43	68.75	98.8	70.06	97.04	36	85.78	80.03

合,更利于城市功能区识别。此外,ETM 模型相较于 pLSA、LDA 和 STM 的语义挖掘性能更加优良,能够确保城市功能区的准确识别。然而此方法未考虑影像分割的最优尺度,各类型 POI 的最优权重,以及城市建筑的复杂或混合功能属性。下一步研究中可采用影像分割或者超像素作为分类单元,结合深度学习算法,探索多源社交媒体数据,并细致考虑混合功能区类型,进一步提升城市功能区的识别准确性。

参考文献(References):

[1] Sun Weiwei, Yang Gang, Chen Chao, *et al.* Development Status and Literature Analysis of China's Earth Observation Remote Sensing Satellites[J]. Journal of Remote Sensing, 2020, 24(5): 479-510. [孙伟伟, 杨刚, 陈超, 等. 中国地球观测遥感卫星发展现状及文献分析[J]. 遥感学报, 2020, 24(5): 479-510.]

[2] Yang Zhaobin, Zhang Shuwen, Bu Kun, *et al.* Application of High Spatial Resolution Remote Sensing Image in Urban LUCC[J]. Journal of University of Chinese Academy of Sciences, 2016, 33(3): 289-297. [杨朝斌, 张树文, 卜坤, 等. 高分辨率遥感影像在城市 LUCC 中的应用[J]. 中国科学院大学学报, 2016, 33(3): 289-297.]

[3] Sun Danfeng, Yang Jihong, Liu Shunxi. Application of High-Spatial IKONOS Remote Sensing Images in Land Use Classification and Change Monitoring[J]. Journal of Agricultural Engineering, 2002, 13(2): 160-164. [孙丹峰, 杨冀红, 刘顺喜. 高分辨率遥感卫星影像在土地利用分类及其变化监测的应用研究[J]. 农业工程学报, 2002, 13(2): 160-164.]

[4] Zhang X, Du S, Wang Q, *et al.* Multiscale Geoscene Segmentation for Extracting Urban Functional Zones from VHR Satellite Images[J]. Remote Sensing, 2018, 10(2): 281. doi: 10.3390/rs10020281.

[5] Tokarczyk P, Wegner J D, Walk S, *et al.* Features, Color Spaces, and Boosting: New Insights on Semantic Classification of Remote Sensing Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 53(1): 280-295. doi:

10.1109/TGRS.2014.2321423.

[6] Zhong Y, Zhu Q, Zhang L. Scene Classification based on the Multifeature Fusion Probabilistic Topic Model for High Spatial Resolution Remote Sensing Imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(11): 6207 - 6222. doi:10.1109/TGRS.2015.2435801.

[7] Liu X, He J, Yao Y, *et al.* Classifying Urban Land Use by Integrating Remote Sensing and Social Media Data[J]. International Journal of Geographical Information Science, 2017, 31(8): 1675-1696. doi:10.1080/13658816.2017.1324976.

[8] Gao Q, Fu J, Yu Y, *et al.* Identification of Urban Regions' Functions in Chengdu, China, based on Vehicle Trajectory Data[J]. PLOS One, 2019, 14(4): e0215656. doi: 10.1371/journal.pone.0215656.

[9] Tu W, Zhu T T, Xia J Z, *et al.* Portraying the Spatial Dynamics of Urban Vibrancy Using Multisource Urban Big Data[J]. Computer Environment and Urban Systems, 2020, 80: 101428. doi:10.1016/j.compenvurbsys.2019.101428.

[10] Tu W, Hu Z, Li L, *et al.* Portraying Urban Functional Zones by Coupling Remote Sensing Imagery and Human Sensing Data[J]. Remote Sensing, 2018, 10(1): 141. doi:10.3390/rs10010141.

[11] Zhu Y, Deng X, Newsam S. Fine-grained Land Use Classification at the City Scale Using Ground-level Images[J]. IEEE Transactions on Multimedia, 2019, 21(7): 1825-1838. doi:10.1109/TMM.2019.2891999.

[12] Yao Y, Li X, Liu X, *et al.* Sensing Spatial Distribution of Urban Land Use by Integrating Points-of-interest and Google Word2vec Model[J]. International Journal of Geographical Information Science, 2017, 31(4): 825 - 848. doi: 10.1080/13658816.2016.1244608.]

[13] Than K, Ho T B. Fully Sparse Topic Models[C]// Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, Berlin, Heidelberg, 2012: 490-505. doi:10.1007/978-3-642-33460-3_37.

[14] Zhang X, Du S, Zheng Z. Heuristic Sample Learning for Complex Urban Scenes: Application to Urban Functional-zone Mapping with VHR Images and POI Data[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 161: 1-12. doi:10.1016/j.isprsjprs.2020.01.005.

[15] Dieng A B, Ruiz F J R, Blei D M. Topic Modeling in Embed-

- ding Spaces [J]. ArXiv Preprint ArXiv: 1907.04907, 2019. doi:10.1162/tacl_a_00325.
- [16] Hu Yong, Liu Liangyun, Peter C, *et al.* Landsat Time-series Land Cover Mapping with Spectral Signature Extension Method[J]. Journal of Remote Sensing, 2015, 19(4): 639-656. [胡勇, 刘良云, Peter C, 等. 光谱特征扩展的时间序列 Landsat 数据地表覆盖分类[J]. 遥感学报, 2015, 19(4): 639-656.]
- [17] Zhang Fafa, Li Weizhong, Lu Liuye, *et al.* Technologies of Extracting Land Utilization Information based on SVM Method with Multi-window Texture [J]. Journal of Remote Sensing, 2012, 16(1): 67-78. [张伐伐, 李卫忠, 卢柳叶, 等. 康乐 SVM 多窗口纹理土地利用信息提取技术[J]. 遥感学报, 2012, 16(1): 67-78.]
- [18] Zhang B, Gao Y, Zhao S, *et al.* Local Derivative Pattern Versus Local Binary Pattern: Face Recognition with High-order Local Pattern Descriptor[J]. IEEE transactions on Image Processing, 2009, 19(2): 533-544. doi:10.1109/tip.2009.2035882.
- [19] Bay H, Ess A, Tuytelaars T, *et al.* Speeded-up Robust Features (SURF) [J]. Computer Vision and Image Understanding, 2008, 110(3): 346-359. doi:10.1016/j.cviu.2007.09.014.
- [20] Oliva A, Torralba A. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope [J]. International Journal of Computer Vision, 2001, 42(3): 145-175. doi:10.1023/A:1011139631724.
- [21] Jiang S, Alves A, Rodrigues F, *et al.* Mining Point-of-interest Data from Social Networks for Urban Land Use Classification and Disaggregation[J]. Computers, Environment and Urban Systems, 2015, 53: 36-46. doi:10.1016/j.compenvurb-sys.2014.12.001.
- [22] Csurka G, Dance C, Fan L, *et al.* Visual Categorization with Bags of Keypoints[C]// Workshop on Statistical Learning in Computer Vision, ECCV. 2004, 1(1-22): 1-2. doi:10.1080/01621459.1949.10483312.
- [23] Jia Y, Ge Y, Ling F, *et al.* Urban Land Use Mapping by Combining Remote Sensing Imagery and Mobile Phone Positioning Data[J]. Remote Sensing, 2018, 10(3): 446. doi:10.3390/rs10030446.
- [24] Bosch A, Zisserman A, Muñoz X. Scene Classification Via pLSA[C]// European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2006: 517-530. doi: 10.1007/11744085_40.
- [25] Blei D M, Andrew Y Ng, Jordan M I. Latent Dirichlet Allocation[J]. Journal of Machine Learning Research, 2003, 3(33): 993-1022. doi:10.5555/9449.19.944937.

Identify Urban Functional Zones Using Multi Feature Latent Semantic Fused Information of High-spatial Resolution Remote Sensing Image and POI Data

Gao Ziwei¹, Sun Weiwei², Cheng Penggen¹, Yang Gang², Meng Xiangchao³

(1.School of Surveying and Mapping Engineering, East China University of Technology, Nanchang, 330013, China;

2.Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo 315211, China;

3.Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China)

Abstract: Accurate identification and division of urban functional zones play an important role in rational planning of urban development and solving urban problems. Remote sensing images have rich spectral texture features, but it is difficult to characterize the social and economic attributes of buildings, while urban data such as social media data provide rich data resources for urban research and application, and supplement the internal characteristics of buildings missing from remote sensing images. In this study, multi-feature information of high-resolution remote sensing image and POI data is integrated and embedded topic model is used to mine its potential semantic information to identify urban functional areas. Three experiments were designed with two typical urban business districts in Ningbo as the study area to verify the effect and performance of the research method. The results show that this method can achieve 85.67% and 85.78% classification accuracy, and can accurately identify urban functional zones. At the same time, the multi-feature information of spectral, texture, geometry and POI feature combination can significantly improve the identification accuracy of urban functional zones, and the embedded topic model can mine the high-level potential semantic information of multi-features better than the three mainstream topic models of pLSA, LDA and STM.

Key words: High spatial resolution remote sensing image; POI; Urban functional zones identification; Multi-feature information fusion; Topic model