

引用格式:崔成,任红艳,赵璐,等.基于街景影像多特征融合的广州市越秀区街道空间品质评估[J].地球信息科学学报,2020,22(6):1330-1338.
[Cui C, Ren H Y, Zhao L, et al. Street space quality evaluation in Yuexiu district of Guangzhou city based on multi-feature fusion of street view imagery[J]. Journal of Geo-information Science, 2020,22(6):1330-1338.] DOI:10.12082/dqxxkx.2020.200072

基于街景影像多特征融合的广州市越秀区街道空间品质评估

崔成^{1,2},任红艳^{1*},赵璐^{1,2},庄大方¹

1. 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室,北京 100101;
2. 中国科学院大学资源与环境学院,北京 100190

Street Space Quality Evaluation in Yuexiu District of Guangzhou City based on Multi-feature Fusion of Street View Imagery

CUI Cheng^{1,2}, REN Hongyan^{1*}, ZHAO Lu^{1,2}, ZHUANG Dafang¹

1. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; 2. College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100190, China

Abstract: Street View Imagery (SVI) is one of the important data sources for the quantitative research of urban built environment. However, it is difficult to fully represent all the information with one type of feature in the SVI due to its complexity and diversity. In this paper, we proposed an effective multi-feature fusion method to evaluate the street space quality based on SVI. Taking Yuexiu district in Guangzhou city as the study area, the Baidu SVIs in the four orientations (front, behind, left, right) at the sample points were obtained. Speeded Up Robust Feature (SURF) and Histogram of Oriented Gradient (HOG) were derived from SVIs as handcrafted features. Semantic features were also derived from SVIs using ENet convolution neural network as features based on deep learning. Based on single feature and multi-feature fusion, Support Vector Machine (SVM) and Random Forest (RF) were used to train the street space quality evaluation model for the four orientations on the training set. The optimal model and the combination of features were selected according to the classification accuracy and Kappa coefficient on the test set. Results showed that: (1) The optimal classification accuracy of models based on SVM was 82.8% (front), 81.7% (behind), 76.6% (left), 76.6% (right), respectively. In the models based on single feature, the average accuracy of the models based on HOG feature (73.03%) or semantic feature (72.28%) was better than the SURF feature (56.00%). The optimal classification accuracy of the models based on RF algorithm was 82.8% (front), 85.0% (behind), 78.1% (left), 70.3% (right). (2) The accuracy of front and behind orientation model was slightly better than that of left and right orientation. The optimal models of each orientation all are multi-feature fusion models. The average classification accuracy and Kappa coefficient of

收稿日期:2020-02-11;修回日期:2020-03-20.

基金项目:国家自然科学基金项目(41571158);国家重点研发计划项目(2016YFC1302602)。[**Foundation items:** National Natural Science Foundation of China, No.41571158; National Key Research and Development Program of China, No.2016YFC1302602.]

作者简介:崔成(1995—),男,河北保定人,硕士生,研究方向为空间数据挖掘。E-mail: cuic@lreis.ac.cn

*通讯作者:任红艳(1979—),男,湖南岳阳人,博士,副研究员,研究方向为环境变化模拟与风险评估。E-mail: renhy@lreis.ac.cn

these optimal evaluation models was 80.6% and 0.62, respectively. These results showed that the proposed method could achieve a high recognition performance. (3) The selection and fusion of features were more determined to the model performance when the SVI were used to evaluate the street space quality, while the performance difference between the two algorithms was small. (4) There were obvious spatial differences in the street space quality of Yuexiu district. The street space quality in the southeast of Yuexiu district needed to be improved. A large scale and high precision street space quality evaluation method was proposed based on multi-feature fusion of SVI and achieved a high recognition performance in this study. And the street space quality score in Yuexiu district was obtained. These results could provide valuable clues for local authorities to conduct comprehensive renovations of urban built environment.

Key words: street view imagery; street space quality; multi-feature fusion; random forest; support vector machine; handcrafted feature; feature based on deep learning; Guangzhou city

***Corresponding author:** REN Hongyan, E-mail: renhy@reis.ac.cn

摘要:全面准确地描述街景影像的多层次特征在基于街景影像对街道空间品质进行评估的研究中具有重要意义。以广州市越秀区为例,获取前后左右各视角街景影像中手工设计的特征(SURF特征、HOG特征)和基于深度学习的特征(语义特征),基于单一特征和多特征融合采用支持向量机(Support Vector Machine, SVM)、随机森林(Random Forest, RF)训练各视角的评估模型。结果表明,以基于SVM建立的单特征模型为例,基于HOG特征(73.03%)、语义特征(72.28%)的模型平均精度优于SURF特征(56.00%),基于SVM前后左右各视角模型的最优分类精度为82.8%(前)、81.7%(后)、76.6%(左)、76.6%(右),而基于RF各视角模型的最优分类精度为82.8%(前)、85.0%(后)、78.1%(左)、70.3%(右)。前后视角的模型精度略优于左右视角。各视角最优模型均为多特征融合模型,最优模型平均分类精度和Kappa系数可达80.6%和0.62。利用街景影像评估街道空间品质时,各算法之间性能差异微弱,而特征选择及组合方式是提升精度的关键。越秀区街道空间品质存在明显空间分异,其西南部的街道空间品质亟待提升。本研究构建了基于街景影像多特征融合的大规模高精度街道空间品质测度方法,实现了对越秀区街道空间品质的评估,研究结果可为相关部门进行街道环境综合整治提供参考。

关键词:街景影像;街道空间品质;多特征融合;随机森林;支持向量机;手工设计特征;基于深度学习的特征;广州市

1 引言

自1978年至今,中国城镇化取得了令人瞩目的成就,人口城镇化率由1978年的17.92%提高到2018年的59.58%。然而城镇化背后的问题日益凸显,城市内部发展不平衡,部分大中城市出现了以“城中村”为代表的非正规居住空间。同常规城市建成区相比,非正规居住空间内部街道空间品质低下,如道路环境脏乱差、低矮老旧房屋密集等^[1]。随着我国新型城镇化建设的推进,高质量发展成为新时期的主题。街道是构成城市公共空间,提升城市生活品质的关键,因此对街道空间品质进行评估对于城市规划管理具有重要意义。基于高分遥感影像虽然可以提取“城中村”等非正规居住空间,却难以直接对某条街道的空间品质进行评估^[2]。而受限于时间和人力限制,问卷调查和实地踏勘等方式难以实现城市尺度上的精细化全覆盖。大范围、精细化地评估街道空间品质成为亟待解决的问题。

街景地图为用户提供了街道的360°全景影像,

街景影像(Street View Imagery, SVI)具有覆盖范围广、位置精度高、采集成本低等优势,且街景影像以人本视角呈现了街道景观真实形态,是城市景观的直接适宜表示,为城市建成环境的量化研究提供了重要数据源^[3-4],国内外学者尝试利用街景影像对城市环境进行大范围精细化的评估。

基于街景影像评价街道/社区/城市建成环境的研究主要包括客观物质要素评估,如行人数量^[5]、道路绿化水平等^[6],以及居民对街道环境的主观综合评价,如道路安全性^[7]、可步行性^[8]、街道空间品质等^[9-12]。基于街景影像评估街道空间品质常用方法包括3类。①以目视解译为主,依靠专家经验建立评价体系^[9],常用于小范围内道路空间品质评估,如赵雅芝^[10]通过分析广州市17个历史文化街区的266张街景影像,结合调查问卷发现其街道空间品质参差不一。②机器学习算法是目前应用较为广泛的方法,研究者提取街景影像中各类特征,在监督学习的框架下,建立低层视觉特征与高层语义场景描述之间的映射关系,进而可以在大规模尺度上进行

街道空间品质评估。例如,叶宇等^[11]以40 m的采样间隔获取上海市杨浦区和虹口区的百度街景影像,以SegNet网络作为特征提取器,获取影像中街道绿视率、天空可见度、道路机动化程度等6个特征,使用人工神经网络训练评估模型;刘伦等^[12]利用AlexNet等卷积神经网络,将最后一层隐含层的输出作为影像特征,结合支持向量机算法,对北京市建筑单体的外立面品质和建筑群体的房屋连续度进行评估。然而,上述研究中忽略了街景影像中低层次的纹理特征等信息。③以卷积神经网络(Convolutional Neural Network, CNN)为代表的深度学习算法也较为常用。甘欣悦等^[13]构建了基于残差网络的CNN对街景图片进行分类,在测试集上准确率约为79%,进而对北京老城的非正规性进行量化识别与评价。CNN直接将原始图像输入到网络中,避免了传统机器学习算法中显性的特征抽取过程。在海量数据的支持下,CNN在影像场景分类等任务上表现良好。然而,搭建参数众多,层级很深的CNN需要大规模数据集辅助训练,且CNN网络结构的设计依然是需要探索的问题^[14]。

传统的机器学习算法在图像场景分类任务中仍具有强大生命力,而图像特征的有效表示是获得良好分类效果的基础。传统的特征提取主要依赖于手工设计特征,手工设计特征是从图像特点、视觉心理学等角度出发,直接设计图像特征表达,反映图像内容的某些方面(包括纹理特征、形状特征等),如局部二值模式(Local Binary Pattern, LBP),尺度不变特征变换(Scale Invariant Feature Transform, SIFT),梯度方向直方图特征等。手工设计特征对图像的表达较为直接,但难以表达图像的高层语义信息。而随着深度学习的发展,基于深度学习的特征逐渐受到关注。如CNN使用图像灰度值作为神经网络输入层,随着网络层数的加深,可以逐步提取到深层次特征,将其卷积层或全连接层的输出作为图像特征用于构建分类器^[15]。彭清等^[16]在预训练CNN模型的基础上,提取所有隐含层的特征,结合支持向量机算法进行车辆识别,比常规CNN识别效果更好。2类特征各有所长,在图像分类识别任务中均有广泛应用。由于街景影像内容的复杂性与多样性,单一特征难以充分表达影像中所有信息,潜在地影响了模型精度,有必要融合多种特征用于提高特征表达能力,克服单一特征的不足^[17]。

基于此,本研究尝试融合手工设计特征和基于深度学习的特征,采用常用图像特征算子和ENet卷积神经网络等算法提取街景影像中不同层级、不同种类的特征,使用支持向量机、随机森林算法训练评估模型,以广州市越秀区作为案例区,对比单一特征模型和多特征融合模型在测试集上的分类精度,进而选择最优模型对越秀区街道空间品质进行评估,研究结果可为相关部门开展街道景观空间品质提升等工作提供数据支撑。

2 研究区概况、数据来源与研究方法

2.1 研究区概况

广州市是特大城市的典型代表,而越秀区是广州市11个城区中面积最小、人口密度最高的中心城区,越秀区面积为33.80 km²,2017年户籍人口约118万人。越秀区(113.24°E—113.32°E,23.11°N—23.17°N)位于广州市中部,东起广州大道,西至人民路,北部为白云山,南侧与海珠区隔江相望。越秀区属亚热带季风气候,四季常绿。越秀区老旧小区数量达148个,占全区222个社区的67%。

2.2 数据来源与预处理

2.2.1 数据获取

网络地图服务商如谷歌地图、腾讯地图、百度地图目前均提供街景地图服务。本研究所使用的街景影像来自于百度地图(<http://api.map.baidu.com/panorama/v2>),参数见表1。越秀区的道路信息来自

表1 百度街景影像参数信息

Tab. 1 Required parameters for Baidu map street view imagery

参数名	描述	取值
ak	开发者密钥,通过申请获取	32位字符
width	图片宽度,单位为像素,范围[10,1024]	1024
height	图片高度,单位为像素,范围[10,512]	512
location	街景采样点的位置	采样点经纬度坐标
coordtype	街景采样点的坐标类型	wgs84ll
heading	朝向角,表示与正北方向夹角,范围[0,360]	平行/垂直于道路方向
pitch	俯仰角,街景相机低头的角度,范围[0,90]	0
fov	视场角,范围[10,360]	90

注:百度街景影像数据涵盖范围不包括路宽过窄、禁止车行的道路和单位大院、门禁社区等内部道路。

于 OpenStreetMap (<https://www.openstreetmap.org>)。在越秀区路网上以 50 m 的间隔生成采样点,在保证了街道景观连续性的同时避免数据冗余。根据采样点所在道路的朝向,分别获取平行于道路(前向视角、后向视角)和垂直于道路方向(左向视角、右向视角)共 4 张街景影像,每幅影像的视场角设置为 90°,此时采集的街景影像即可全面囊括采样点周边环境。

采用 Python 语言编写网络爬虫批量下载越秀区的街景影像,剔除无街景影像的采样点后,最终获取广州市越秀区 14 930 个采样点上的 59 720 幅街景影像,街景影像的实际拍摄时间为 2017 年 5 月。

2.2.2 影像标注

街道空间品质是居民对道路环境的感知体验。它是与周边物理环境高度相关的主观感知指标,涉及到道路整洁程度、绿化水平、天空开阔度等多种要素^[11-12]。街道空间品质的概念较抽象,通过文献调研和实地踏勘,本研究确定了越秀区街道空间品质低下的具体表征:两侧房屋低矮密集,建筑布局混乱,建筑外立面老旧,管线布局杂乱;道路狭窄,人车混行,道路环境卫生脏乱差,绿化水平低,天空可见度差;部分街道上存在大量样式、颜色互不统一的商铺牌匾或广告牌,小商贩聚集。

以上述表征为依据对 4 个视角的街景影像进行标注(图 1),分为低质量和高质量 2 类。满足 1 条及



图 1 不同视角高/低空间品质的街景影像示例
Fig. 1 Sample of SVIs with high/low street space quality from 4 orientations

以上特征的影像标注为街道空间品质低,反之则标注为高空间品质街道。4 个视角各标记 200 张街景影像,其中街道空间品质低和高的街景影像样本各 100 张,作为机器学习算法的训练集和测试集。

2.3 研究方法

2.3.1 特征提取

街景影像的图像内容复杂,细节信息丰富。前后视角的街景影像常出现道路、汽车、天空,而左右视角道路景观则存在较大差异,如高速路两侧常存在隔音板,低等级公路两侧常为行道树或不同类型的建筑物(图 1)。基于街景影像对街道空间品质进行评估的关键是提取有效特征来准确表达图像内容。单一特征难以全面描述街景影像的复杂内容,本文获取了不同类型的图像特征,包括手工设计的特征(SURF 特征,HOG 特征)和基于深度学习的特征(语义特征),采用串行特征融合的策略进行组合,即图像的多个特征串接构成一个特征向量,然后选择合适的分类器进行处理。

(1) SURF 特征

快速健壮特征(Speed Up Robust Feature, SURF)是一种基于尺度空间的,具有良好几何、光照、遮挡和视角变化鲁棒性的图像局部不变特征,可以较好地描述图像结构和纹理信息。相比于 SIFT(Scale Invariant Feature Transform)算子,SURF 算子计算速度提高了 3 倍,而其性能与 SIFT 相近,具有良好实时性^[18]。

SURF 算子是以检测图像中的特征点为基础,通过生成特征点直方图并构建特征描述子来获取图像纹理特征。不同影像提取的特征点数量存在差异,因此采用 K 均值聚类算法对所有街景影像的特征点进行聚类,构建视觉词袋模型,最终为每幅街景影像生成一个 50 维的 SURF 特征。

(2) HOG 特征

梯度方向直方图(Histogram of Oriented Gradient, HOG)特征对光照、尺度、方向有良好的适应性,它通过计算和统计目标区域内的梯度方向分布来描述影像中局部目标的形状特征,因其对目标形状的良好刻画而获得了广泛应用^[19]。HOG 算法将图像分割成较小的胞元(cell),然后计算并统计胞元中各个像素点的梯度或边缘方向直方图构成该胞元的特征向量,把各个胞元组合成空间上连通的区块(block),并对梯度强度进行归一化,以保证其光照不变性。

提取 HOG 特征时 cell 尺寸设置为 64 像素×64 像素, block 尺寸设置为 2×2 胞元, 每个 block 的滑动增量为 32 像素×32 像素, 进而从每幅 1024 像素×512 像素的街景影像中获取一个 3780 维向量来描述 HOG 特征, 采用主成分分析技术将 HOG 特征降至 50 维。

(3) 语义特征

本文将街道景观中各类要素(道路、植被、天空等)视觉占比作为街景影像的语义特征。ENet 是目前针对影像语义分割领域实时性最好的深度卷积神经网络之一, 且具有较高的准确率^[20]。本文采用预训练的 ENet 网络作为特征提取器, 通过对街景影像进行像素级语义分割, 将各像素分类为天空、植被、道路、人行道、建筑物、墙、围栏、灯柱、交通信号灯、交通标志、山体、行人、骑行者、小汽车、卡车、公交车、火车、摩托车、自行车、背景共 20 类要素, 通过计算各类别要素的像素占比将其作为街景影像的语义特征。

2.3.2 模型构建

判定街景影像反映的街道景观属于高/低空间品质是二分类问题。支持向量机(Support Vector Machine, SVM)是解决二分类问题的常用方法, 通过非线性映射算法(核函数)将低维空间线性不可分样本转化至高维空间内线性可分的样本, 并通过间隔最大化的原则在高维空间内寻找最优分类超平面, 在解决非线性、小样本及高维模式识别问题上具有优势^[21]。本研究中 SVM 核函数采用径向基核函数, 算法主要参数正则化参数 C 与径向基核函数的带宽 gamma 采用网格搜索法进行筛选。

随机森林(Random Forest, RF)是一种基于分类树的算法, 利用 bootstrap 策略从原始样本中生成多个样本集, 再随机抽取特征子集, 对每个 bootstrap 样本集建立决策树, 组合多颗决策树的预测进而通过投票得出最终结果。RF 利用样本随机和特征随机的策略构建多颗互不依赖的决策树, 通过降低方差的策略提高了模型的泛化性能。算法主要参数包括决策树数量和决策树最大深度。本研究中构建 RF 的决策树数量为 100, 树最大深度同样采用网格搜索法确定。

基于街景影像中获取的 SURF 特征、HOG 特征和语义特征, 本研究采用串行方式对多个特征向量进行组合, 利用 SVM 和 RF 构建街道空间品质评估模型, 街道空间品质评估模型可以判定某一

幅街景影像所反映的街道景观是否为高空间品质。考虑到不同视角街景影像反映的城市景观存在差异(图 1), 因此针对不同视角的街景影像, 基于单一特征和多特征融合的方式分别构建 SVM 和 RF 分类器。在模型训练前对街景影像的 SURF 特征、HOG 特征和语义特征进行标准化处理, 随机选取 70% 人工标注的街景影像训练模型, 其余 30% 用来评估模型精度。利用模型在测试集上分类精度和 Kappa 系数衡量其性能, 其中分类精度是正确分类的影像占测试集整体影像的比例, Kappa 系数的计算基于混淆矩阵, Kappa 系数超过 0.6 表示预测结果与实际结果具有高度一致性。以上度量指标作为依据选择各视角最优模型及对应的特征组合方式, 进而对越秀区的街道空间品质进行评估, 整体技术路线如图 2 所示。

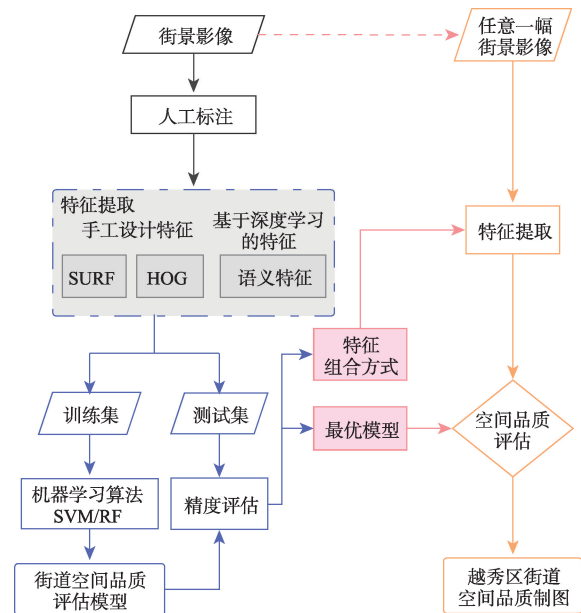


图2 街道空间品质评估的技术流程

Fig. 2 Technical flow chart of street space quality evaluation

3 实验结果与分析

3.1 模型性能对比

表 2 对比了采用不同特征组合方式、不同算法所构建模型的性能差异。结果表明, 基于 SVM 算法前后左右各视角街景的最优分类精度依次为 82.8% (前)、81.7% (后)、76.6% (左)、76.6% (右), 对应特征组合方式依次为 SURF 特征+HOG 特征+语义特征、HOG 特征+语义特征、SURF 特征+HOG 特征+语义特征、SURF 特征+HOG 特征+语义特征。单特征模

表2 模型分类精度与 Kappa 系数
Tab. 2 Model classification accuracy and Kappa coefficient

	各视角模型的分类精度/Kappa 系数							
	SVM算法				RF算法			
	前向街景	后向街景	左向街景	右向街景	前向街景	后向街景	左向街景	右向街景
SURF 特征	50.0%/0.00	60.0%/0.23	60.9%/0.22	53.1%/0.08	46.9%/-0.07	71.7%/0.43	60.9%/0.23	56.3%/0.13
HOG 特征	76.6%/0.53	73.3%/0.46	70.3%/0.40	71.9%/0.44	79.7%/0.59	70.0%/0.39	73.4%/0.47	70.3%/0.41
语义特征	75.0%/0.51	75.0%/0.51	70.3%/0.42	68.8%/0.38	81.3%/0.62	81.7%/0.63	71.9%/0.44	65.6%/0.32
SURF 特征+HOG 特征	68.8%/0.37	75.0%/0.50	73.4%/0.47	67.2%/0.35	76.6%/0.53	75.0%/0.50	68.8%/0.38	62.5%/0.25
SURF 特征+语义特征	81.3%/0.63	68.3%/0.37	70.3%/0.41	70.3%/0.41	82.8%/0.65	83.8%/0.67	73.4%/0.47	57.8%/0.15
HOG 特征+语义特征	82.8%/0.66	81.7%/0.63	76.6%/0.54	71.9%/0.44	82.8%/0.66	85.0%/0.70	78.1%/0.57	67.2%/0.35
SURF 特征+HOG 特征+语义特征	82.8%/0.66	78.3%/0.57	76.6%/0.54	76.6%/0.53	82.8%/0.65	81.7%/0.63	78.1%/0.57	62.5%/0.25

注: 各列加粗数值为当前列最优的分类精度和 Kappa 系数。

型的平均分类精度为 HOG (73.03%) > 语义特征 (72.28%) > SURF (56.00%)。整体而言, 基于 SURF 特征+HOG 特征+语义特征多特征融合的 SVM 分类器识别街景影像中街道空间品质的能力最强, 平均分类精度和 Kappa 系数分别为 78.6% 和 0.58。

而基于 RF 分类器不同视角模型的最优分类精度依次为 82.8% (前)、85.0% (后)、78.1% (左)、70.3% (右), 对应特征组合方式依次为 HOG 特征+语义特征、HOG 特征+语义特征、HOG 特征+语义特征、HOG 特征。单特征模型的平均分类精度为语义特征 (75.13%) > HOG (73.35%) > SURF (58.95%)。整体而言, 基于 HOG 特征+语义特征多特征融合的 RF 分类器判定街景影像中街道空间品质的能力最强, 分类精度和 Kappa 系数均值分别为 78.3% 和 0.57。

上述结果表明基于不同视角, 不同特征组合方式所构建模型存在性能差异。前后视角的模型精度略优于左右视角的模型精度, 单特征模型中 HOG 特征与语义特征的性能明显优于 SURF 特征的性能, 而多特征融合模型精度普遍优于基于单特征的模型。相比之下, 算法的差异则对模型性能影响较小, 基于 SVM 或 RF 构建的各视角最优模型性能相近。

前后左右各视角的最优空间品质评估模型依次是 RF (HOG 特征+语义特征)、RF (HOG 特征+语义特征)、RF (HOG 特征+语义特征)、SVM (SURF 特征+HOG 特征+语义特征), 均为多特征融合模型, 分类精度和 Kappa 系数均值为 80.6% 和 0.62。模型预测结果与实际标签具有高度一致性。该结果说明模型识别精度高, 基于街景影像的多特征融合策略能精确有效地评估街道的空间品质。

3.2 越秀区街道空间品质评估

各视角最优模型在测试集上表现良好, 因此用其评估越秀区 59 720 张街景影像中的街道空间品质。模型判断某幅街景影像属于高品质街道的概率值越高, 则表示街道空间品质越高, 通过计算各采样点 4 幅街景影像所反映街道空间品质的均值, 最终获取越秀区采样点尺度街道空间品质的空间分布 (图 3)。

广州市越秀区街道空间品质评分的均值为 0.62, 方差为 0.14, 变异系数为 23.05%。采样点尺度街道空间品质存在空间分异, 交通主干道的品质评分普遍高于次干道和支路。

越秀区西南部存在大量低空间品质的街道, 该区域密集分布着大量的老旧小区、城中村等 (图 3 区域 A), 建筑外立面老旧, 道路狭窄, 人车混行, 整体环境杂乱, 街道空间品质亟待改善提升。

高空间品质的街道分布范围广, 如图 3 中 C1 区域位于越秀区北部, 地处白云山脚下, 道路环境优美, 绿化程度高。越秀区东南部的东山湖公园, 二沙岛一带空间品质也较高 (图 3 中 C2 区域), 各类文化艺术场馆、景观雕塑、高档住宅小区和体育训练基地散布在青葱满目的绿植中。

中等空间品质的街道主要集中在交通主干路周边, 如图 3 中区域 B, 该类区域街道宽阔, 整洁度高, 两侧绿化程度良好, 但车辆或建筑物较多, 视野不开阔, 街道景观略显杂乱。

以上结果可以作为越秀区街道空间品质的现状基础评价资料。结果表明越秀区街道空间品质存在明显的空间分异, 越秀区西南部存在大量老旧小区等居住空间, 该区域的街道空间品质亟待提升。

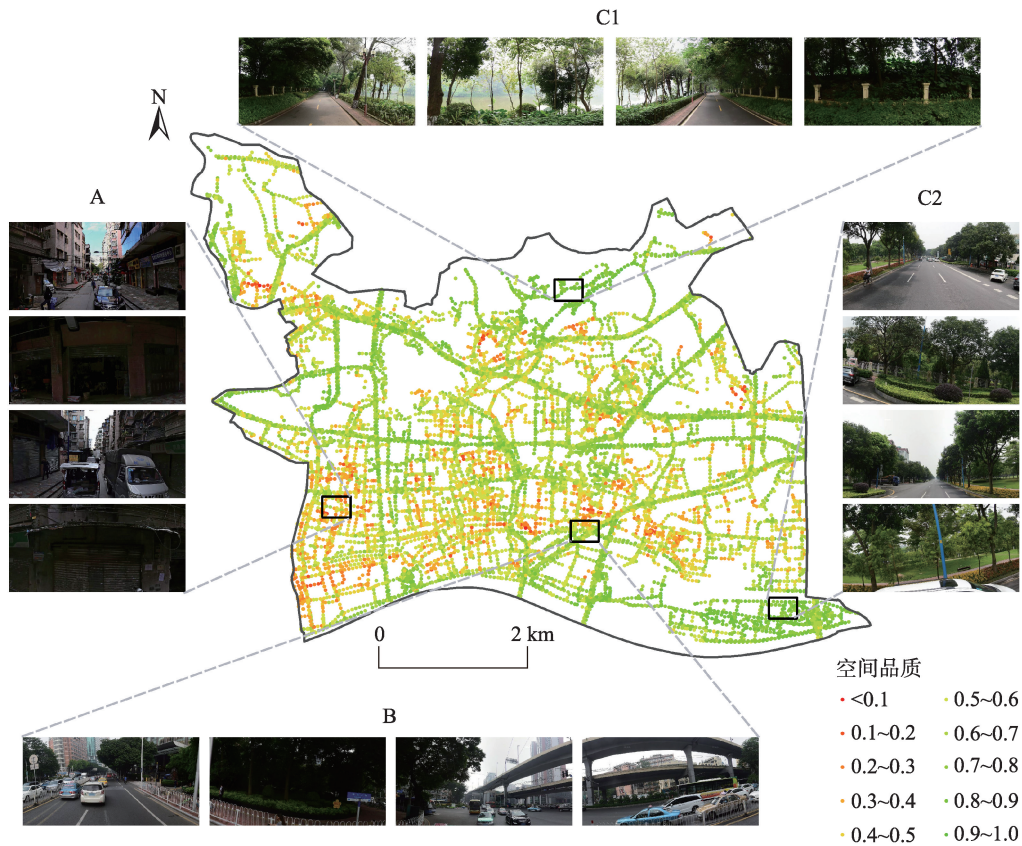


图3 越秀区采样点尺度道路空间品质
Fig. 3 Street space quality in Yuexiu District

4 讨论

多模型对比有助于获取更好的预测结果^[22]。本研究发现分别基于SVM和RF算法建立的各视角最优模型在测试集上表现相似。SVM和RF在二分类问题上的泛化能力不存在显著差异^[23]。而各视角最优模型均为融合了低层次细节信息(SURF特征,HOG特征)和高层次语义信息(语义特征)的模型,对街景影像更为全面的描述使得多特征融合模型性能突出。

而在3个特征之中,HOG特征通过滑动窗口的策略最大限度地保留了图像中纹理信息^[19]。实地调研发现空间品质较低的街道景观中植被、天空的占比相对较少,因此基于语义特征也能有效区分街道空间品质的高低。相比之下,在场景理解任务中,基于特征点检测方法获取的SURF特征无法得到图像中充足的信息^[24]。这可能影响了基于SURF特征所建立模型分类精度。

此外,本研究还发现模型精度存在方向上的差

异。这可能与前后视角的道路景观基本相同,而左右视角的道路景观则存在较大差异有关(图1)。道路景观的差异本质上仍为影像特征差异,现有特征集合无法完全表达左右视角街景影像的内容。各视角模型分类精度的差异说明有必要针对不同视角的影像分别建立模型,也进一步证实利用街景影像评估街道空间品质时,在训练样本量相当的情况下,不同算法之间的性能差异微弱,而特征选择和特征组合方式更能决定模型性能。

相比于基于单一特征建立的街道空间品质评价模型^[11-12],本研究证实了多特征融合策略能够有效提升模型分类效果,该策略可拓展到街道安全性评估、街道可步行性等主观综合评价中。在大数据时代,多数据的融合将帮助研究者对城市环境进行多角度的精细化研究,如耦合街景影像和遥感影像可以进行建筑区块级别的土地利用制图^[25]。本研究所涉及到的街道空间品质也可以纳入到类似研究中。

本研究虽取得了一定成果,但仍存在不足之

处:①目前将街景采样点间隔设置为50 m,而更为精细的采样策略(间隔为30 m,10 m等)能否显著提升模型性仍有待评估,不过目前已实现了一套完整的基于街景影像多特征融合的道路空间品质评估方法。②特征体系的构建是影响模型精度的关键,虽然本研究中考虑了不同层次的特征,但仍可能存在疏漏,如视觉熵、显著区域饱和度等均能影响居民对道路环境的评价^[26],需要在后续研究中进一步完善现有特征体系。③采用二分法(高/低)评价街道空间品质存在局限性,可进一步设计精细的评分策略,采用回归模型来对街道空间品质进行评分。④测度体系的可迁移性,通过A、B、C三类区域的对比证实了该测度体系在越秀区的有效性,该体系是否适用于广州市其余中心城区及其他城市核心区仍需要进一步验证。

5 结论

全面准确描述街景影像的多层次特征对于评估街道空间品质至关重要。本文基于广州市越秀区街景影像,通过对比基于不同特征组合方式、不同算法的模型精度,构建了基于街景影像多特征融合的大规模高精度的街道空间品质测度方法,并对越秀区的街道空间品质进行评估。本研究的结果可为城市规划和管理部门进行街道环境的综合整治、制定差异化的街道空间品质改善策略提供参考。本文的主要结论如下:

(1)基于不同特征组合方式构建的模型存在性能差异。单特征模型中HOG特征与语义特征的性能明显优于SURF特征的性能。各视角最优模型均为融合了低层次细节信息(SURF特征,HOG特征)和高层次语义信息(语义特征)的模型,各视角最优模型的分类精度和Kappa系数均值为80.6%和0.62。基于街景影像的多特征融合策略能有效地评估街道的空间品质。

(2)各视角最优SVM模型和RF模型性能相近。利用街景影像评估街道空间品质时,在训练样本量相当的情况下,不同算法间的性能差异微弱,而特征选择和特征组合方式是提升模型精度的关键。

(3)广州市越秀区街道空间品质评分均值为0.62,交通主干道的品质评分高于次干道和支路。越秀区街道空间品质存在明显空间分异,其西南部街道空间品质亟待提升。

参考文献(References):

- [1] Chung H. Building an image of villages-in-the-city: A clarification of China's distinct urban spaces[J]. *International Journal of Urban and Regional Research*, 2010,34(2):421-437.
- [2] Kuffer M, Pfeffer K, Sliuzas R. Slums from space- 15 years of slum mapping using remote sensing[J]. *Remote Sensing*, 2016,8(6):455.
- [3] 张丽英,裴韬,陈宜金,等.基于街景图像的城市环境评价研究综述[J].*地球信息科学学报*,2019,21(1):46-58. [Zhang L Y, Pei T, Chen Y J, et al. A review of urban environmental assessment based on street view images[J]. *Journal of Geo-information Science*, 2019,21(1):46-58.]
- [4] 谢润桦.基于城市街景影像的视觉定位研究[D].北京:北京建筑大学,2018. [Xie R H. Research on visual positioning based on urban street view[D]. Beijing: Beijing University of Civil Engineering And Architecture, 2018.]
- [5] Yin L, Cheng Q, Wang Z, et al. 'Big data' for pedestrian volume: exploring the use of google street view images for pedestrian counts[J]. *Applied Geography*, 2015,63: 337-345.
- [6] Li X, Zhang C, Li W, et al. Assessing street-level urban greenery using google street view and a modified green view index[J]. *Urban Forestry & Urban Greening*, 2015, 14(3):675-685.
- [7] Naik N, Philipoom J, Raskar R, et al. Streetscore-predicting the perceived safety of one million streetscapes[C]. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, USA Columbus, 2014:779-785.
- [8] 周垠,龙瀛.街道步行指数的大规模评价——方法改进及其成都应用[J].*上海城市规划*,2017(1):88-93. [Zhou Y, Long Y. Large-scale evaluation for street walkability: Methodological improvements and the empirical application in Chengdu[J]. *Shanghai Urban Planning Review*, 2017(1):88-93.]
- [9] 唐婧娴,龙瀛.特大城市中心区街道空间品质的测度——以北京二三环和上海内环为例[J].*规划师*,2017,33(2): 68-73. [Tang J X, Long Y. Metropolitan street space quality evaluation: second and third ring of Beijing, inner ring of Shanghai[J]. *Planners*, 2017,33(2):68-73.]
- [10] 赵雅芝.基于街景地图的历史文化街区街道空间品质评价[D].广州:广州大学,2019. [Zhao Y Z. Street space quality evaluation of historical and cultural streets based on street view map: Take Guangzhou as an example[D]. Guangzhou: Guangzhou University, 2019.]
- [11] 叶宇,张昭希,张啸虎,等.人本尺度的街道空间品质测度

- 结合街景数据和新分析技术的大规模、高精度评价框架[J]. 国际城市规划, 2019, 34(1): 18-27. [Ye Y, Zhang Z X, Zhang X H, et al. Human-scale quality on streets: A large-scale and efficient analytical approach based on street view images and new urban analytical tools[J]. Urban Planning International, 2019, 34(1): 18-27.]
- [12] Liu L, Silva E A, Wu C, et al. A machine learning-based method for the large-scale evaluation of the qualities of the urban environment[J]. Computers, Environment and Urban Systems, 2017, 65: 113-125.
- [13] 甘欣悦, 余天唯, 龙瀛. 街道建成环境中的城市非正规性基于北京老城街景图片的人工打分与机器学习相结合的识别探索[J]. 时代建筑, 2018(1): 62-68. [Gan X Y, She T W, Long Y. Understanding urban informality in street built environment: Combining manual evaluation with machine learning in processing the Beijing old city's street-view images[J]. Time Architecture, 2018(1): 62-68.]
- [14] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251. [Zhou F Y, Jin L P, Dong J. Review of convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.]
- [15] 葛芸. 基于CNN迁移特征融合与池化的高分辨率遥感图像检索研究[D]. 南昌: 南昌大学, 2019. [Ge Y. High-resolution remote sensing image retrieval based on fusion and pooling to transfer features from convolutional neural network[D]. Nanchang: Nanchang University, 2019.]
- [16] 彭清, 季桂树, 谢林江, 等. 卷积神经网络在车辆识别中的应用[J]. 计算机科学与探索, 2018, 12(2): 282-291. [Peng Q, Ji G S, Xie L J, et al. Application of convolution neural network in vehicle recognition[J]. Journal of Frontiers of Computer Science and Technology, 2018, 12(2): 282-291.]
- [17] 黄冬梅, 刘佳佳, 苏诚, 等. 多特征融合的复杂环境海洋涡旋识别[J]. 中国图象图形学报, 2019, 24(1): 31-38. [Huang D M, Liu J J, Su C, et al. Ocean eddies recognition based on multi-features fusion in complex environment[J]. Journal of Image and Graphics, 2019, 24(1): 31-38.]
- [18] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features[C]. Papers of the European Conference on Computer Vision, Austria Graz, 2006: 404-417.
- [19] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. Papers of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, USA San Diego, 2005: 886-893.
- [20] Paszke A, Chaurasia A, Kim S, et al. Enet: A deep neural network architecture for real-time semantic segmentation [EB/OL]. <https://arxiv.org/abs/1606.02147>.
- [21] 陆波, 尉询楷, 毕笃彦. 支持向量机在分类中的应用[J]. 中国图象图形学报, 2005, 10(8): 1029-1035. [Lu B, Wei X K, Bi D Y. Application of support vector machine in classification[J]. Journal of Image and Graphics, 2005, 10(8): 1029-1035.]
- [22] 周超, 方秀琴, 吴小君, 等. 基于三种机器学习算法的山洪灾害风险评价[J]. 地球信息科学学报, 2019, 21(11): 1679-1688. [Zhou C, Fang X Q, Wu X J, et al. Risk assessment of mountain torrents based on three machine learning algorithms[J]. Journal of Geo-information Science, 2019, 21(11): 1679-1688.]
- [23] 黄衍, 查伟雄. 随机森林与支持向量机分类性能比较[J]. 软件, 2012, 33(6): 111-114. [Huang Y, Zha W X. Comparison on classification performance between random forests and support vector machine[J]. Computer Engineering & Software, 2012, 33(6): 111-114.]
- [24] Li F F, Perona P. A bayesian hierarchical model for learning natural scene categories[C]. Papers of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, USA San Diego, 2005: 524-531.
- [25] Li X, Zhang C, Li W. Building block level urban land-use information retrieval based on google street view images [J]. GIScience & Remote Sensing, 2017, 54(6): 819-835.
- [26] Cheng L, Chu S, Zong W, et al. Use of tencent street view imagery for visual perception of streets[J]. ISPRS International Journal of Geo-information, 2017, 6(9): 265.