

基于NLP定制模型的游客感知研究 ——以重庆市鹅岭公园为例

A Study on Tourists' Perception of Eling Park in Chongqing City Based on NLP Customized Model

张汇雯 邓宏 冯琬清
ZHANG Huiwen DENG Hong FENG Wanqing

(重庆大学建筑城规学院, 重庆 400044)
(School of Architecture and Urban Planning, Chongqing University, Chongqing, China, 400044)

文章编号: 1000-0283(2023)12-0122-09
DOI: 10.12193/j.laing.2023.12.0122.015
中图分类号: TU986
文献标志码: A
收稿日期: 2023-06-30
修回日期: 2023-10-31

摘要

游客感知对于了解人民需求、提升城市建设质量有着重要意义。以公园网络文本为数据训练NLP定制模型,更适用于风景园林领域的需求,使公园治理与设计更加智能与高效。基于深度学习平台,训练三个多标签文本分类、情感倾向分析、评论观点提取三个模型处理公园网络文本,从“时间—评价对象—评价对象下的感知要素”多层次分析公园游客情感倾向特征,挖掘重点感知要素。研究表明:(1)鹅岭公园游客感知整体积极性较高,6类评价对象中园外景观受关注度最高,自然景观与游客积极情绪成正比,设施配套消极情绪最高,停车位是亟需解决的问题。(2)在60个高频感知要素中,7个要素与游客积极情绪概率成显著正比,其中5个正相关,2个负相关。(3)采取“文本分类—高频词提取—情感分析”的分析顺序,可以挖掘词频低但有重要影响的感知要素。(4)NLP定制模型提供的属性级情感分析可以减少情感分析误差,使研究更准确。研究鹅岭公园游客情感与公园重点感知要素,为鹅岭公园的建设提升提出优化建议,为自然语言处理在风景园林中的应用提供了参考。

关键词

游客感知;网络文本分析;深度学习;情感倾向分析;评论观点抽取

Abstract

The perception of tourists holds significant importance in understanding people's needs and enhancing the quality of urban development. Training an NLP customized model using park network text as data is more suitable for the demands of the landscape architecture field, enabling intelligent and efficient governance and design of parks. Using a deep learning platform, three models were trained for multi-label text classification, sentiment analysis, and comment viewpoint extraction to process park network text. A multi-level analysis of "time-evaluated object-perceived elements under the evaluated object" was conducted to identify key perception features and explore the significant elements. The research revealed the following findings: (1) Overall, visitors' perception of Eling Park was predominantly positive, with the highest attention given to external landscapes among the six evaluated objects. Natural landscapes had the most positive impact on visitors' emotions, while negative emotions were associated with facilities and the need for parking space. (2) Among the 60 high-frequency perception elements, 7 elements are significantly proportional to the probability of tourists' positive emotions, of which 5 are positively and 2 are negatively correlated. (3) The analysis sequence of "text classification - high-frequency word extraction - sentiment analysis" allowed the identification of low-frequency perceived elements that had essential impacts on corresponding landscapes. (4) The attribute-level sentiment analysis provided by the NLP customized model reduced sentiment analysis errors, leading to more accurate research outcomes. This study examined the emotional perception of visitors and the key perceived elements of Eling Park, providing optimization suggestions for its development and contributing to the application of natural language processing in landscape architecture.

Keywords

tourist perception; text analysis; deep learning; sentiment analysis; opinion extraction from reviews

张汇雯

1998年生/女/四川盐亭人/在读硕士研究生/研究方向为风景园林规划与设计

邓宏

1963年生/男/重庆人/副教授/研究方向为风景园林规划与设计、风景园林历史与理论、城市色彩规划与研究

冯琬清

1999年生/女/福建闽清人/在读硕士研究生/研究方向为风景园林规划与设计

基金项目:

国家自然科学基金项目“巴渝传统场镇景观价值识别及活态保护研究”(编号: 51978094)

在社会高质量发展的目标下, 城市的建设愈发关注人民需求, “游客感知”近年来成为风景园林领域的研究热点, 其从使用者的角度提供了公园建设的现状优缺点信息, 帮助设计师识别潜在需要改进的空间。在现有研究中, 历史村镇与街区^[1-2]、国家公园^[3-5]、森林公园^[6-8]、乡村旅游地^[9-10]为游客感知的主要研究对象, 城市公园与以上研究对象有一定的相通性, 但其景观要素、服务设施、服务人群等又各有差异, 城市公园作为重要的城市空间, 是人们休闲娱乐的重要场所, 和人民的幸福感息息相关^[11]。在城市更新的背景下, 重庆市于2021年完成了公园更新提质工作, 其中鹅岭公园作为重庆地标公园, 游览人群众多且功能多样, 拥有较大数量的网络评价文本, 具有可研究性, 并且该公园经过多次环境更新与功能提升, 对其进行游客感知研究, 分析网络评价所体现的人民需求, 可以为建设更好的公园环境提供设计依据。

在研究数据上, 近几年网络文本以其获取性强、数量丰富的特点成为研究者常用的数据来源, 主要集中在游客感知评价^[12-13]、旅游地意象^[14-15]、情感分析^[16-19]等研究。在研究工具上呈现由传统工具转向人工智能的趋势^[20], 传统多采用ROSTCM6等软件对文本进行处理, 较为依靠软件内置的情感分析词典, 而词典无法及时更新。自然语言处理(Natural Language Processing, NLP)是应用深度学习方法来处理和理解人类语言, 更适用于处理当下词汇日益更新的网络文本。通过建立模型模仿人脑的神经网络, 对大量的文本进行特征学习, 并且可以对语言进行理解和分析。已有研究实现了利用深度学习下的NLP技术对城市群和城市公园^[21]进行情感分析和使用后评价^[22], 同时现有的深度学习平台通过

提供预训练好的通用模型降低了深度学习的使用门槛, 已有多名学者利用该平台分析建筑及公园的使用后评价, 并挖掘游客情感特征和景观意象, 取得了一定成果^[23-26]。然而通用模型所提供的分析也存在一定不足, 文本情感倾向分析分为文档级、句子级、属性级^[27-29], 属性级可以将情感落实到具体的单个感知要素上, 而通用模型的情感倾向分析仅在文档级和句子级, 这两个级别假设整个文档只有一个主题, 但公园网络文本存在着对象多样化、观点复杂的特点, 一个句子中可能涉及到多个评价对象多种观点, 通用模型在特定领域中效果较差。其次通用模型无法针对公园内某一个具体的感知要素进行属性级情感分析, 提取的观点也较为笼统, 难以排除其他要素的影响。而定制的评论观点抽取模型, 可以判断评价对象、感知要素以及对应的属性级情感, 能将游客情感对应到具体的景观要素, 有效提高情感分析的细粒度、针对性以及处理效率。

综上所述, 本研究主要从研究对象和研究工具上对游客感知进行深入分析。以重庆城市公园的代表之一鹅岭公园为研究对象, 以该公园的网络文本为研究数据, 使用深度学习平台Easydl对网络文本进行分析, 根据游客感知特征, 总结鹅岭公园建设现状及需要优化的地方, 并分析NLP定制模型在公园网络文本分析中的优点, 挖掘鹅岭公园重要的游客感知要素, 提出深度学习平台分析公园网络文本的技术路线, 为智能算法在风景园林领域中的应用提供参考。

1 研究区域

鹅岭公园位于重庆市渝中区, 为重庆地标式公园, 历史悠久, 选择该公园作为研究对象主要有以下几点原因: (1) 鹅岭公园

知名度高, 人流量大, 在各大社交媒体评论数据样本量大, 利于模型训练; (2) 鹅岭公园景观类型丰富, 其网络文本往往含有对公园要素较为细致的描述, 便于研究公园感知要素与游客情感的关系; (3) 鹅岭公园作为重庆的地标性公园之一, 研究鹅岭公园可以对该地区的景观发展和公园管理提供重要建议。

2 研究方法

研究步骤主要分为“数据采集—数据清洗与定制模型训练—结果分析”三步(图1)。本文以“鹅岭公园”为关键词, 爬取5个主流网络平台的游客评论, 利用深度学习平台EasyDL提供的文本清洗功能和三类NLP模型完成对游客多层次情感倾向分析和公园感知要素的提取, 并通过SPSS相关性分析两者之间的关系。模型训练需要多次迭代才能达到最佳效果, 使用者可以结合模型评估报告和校验结果来决定数据扩充以及调优。本研究所用到的三个定制模型迭代5到15次不等。

2.1 数据采集与清洗

利用八爪鱼平台和python等工具对5个互联网平台评论进行爬取, 选取2017年1月1日到2022年12月31日的评论数据, 共获得10 991条评论数据。

为保证评论数据的真实性和可靠性, 主要采用智能清洗结合手动清洗, 有以下几种清洗方式: (1) Easydl提供基础清洗工具, 包括文本去重、删除emoji和网页链接以及繁体转简体, 以上功能主要基于文本哈希值计算、正则表达式和开源的文本转化库来实现。(2) 删除广告、租房、非游客评论等与鹅岭公园无关的内容。(3) 删除官方账号发布的数据, 如“重庆旅游”“古城重庆”等。

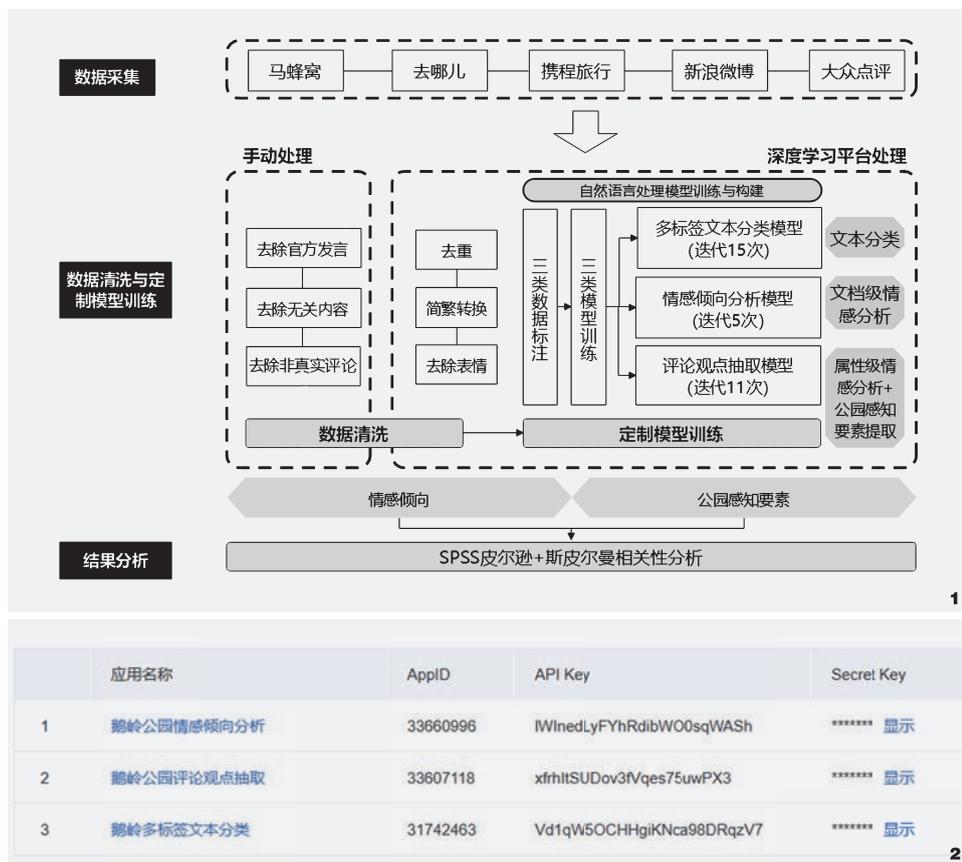


图1 技术路线图
Fig. 1 Technological roadmap

图2 定制模型应用列表
Fig. 2 Customized model application list

(4) 删除非个人真实使用后评论的数据，例如旅游攻略、计划等。(5) 网络评论数据用词较为随意，对评论中的同义词进行统一规范，例如“两江亭”“揽胜楼”等实际上都指“瞰胜楼”，用替换工具统一处理。清洗后数据为5 159条。

2.2 游客多层次情感倾向分析与公园感知要素提取

本文基于EasyDL平台训练了三类定制模型，对网络文本进行“文本分类—文档级情感分析—感知要素提取—属性级情感分析”的处理。三个定制模型经实测满足应用需求

并均已发布。在pycharm中调用模型API即可对文本数据进行批量化处理(图2)。其中多标签文本分类为后两个模型分析打下了基础，文档级情感分析在情感倾向分析模型中实现，属性级情感分析和感知要素提取在评论观点抽取模型中实现(表1)。

(1) 多标签文本分类模型。多标签文本分类模型对一段评价可预测出多个标签，本研究参考过往研究^[30]结合鹅岭公园景观要素，将评论划分为自然景观、建筑景观、人文历史景观、园外景观、公园活动、设施配套6类(表2)，并作为标签对数据进行标注训练定制模型，标注单位为整段文字。

(2) 文档级情感—情感倾向分析模型。本研究通过该模型分析鹅岭公园整体游客感知积极性，对网络文本进行标注和结果呈现，标注单位为整段文字，标注后训练定制模型，模型返回的结果分为三类，其中积极评论4 952条，消极评论174条，中性评论33条，说明积极情绪为鹅岭公园的情感主调。

(3) 属性级情感—评论观点抽取模型。经训练可提取评价片段、评论对象、评论维度、评论观点、属性级情感倾向，本研究主要使用该模型的抽取评价维度及情感倾向分析功能。根据官方文档，评价维度为基于评价对象的属性细分，本文将该模型提取出的评价维度定为感知要素，例如自然景观评价对象下的感知要素有树木、花卉等。并以上述6种分类为评论对象，标注单位为单句，标注数据后训练定制模型(图3)，并对6类文本进行感知要素提取和情感倾向分析。

3 结果分析

3.1 游客感知情感分析

3.1.1 不同年份情感特征

情感倾向分析可以反映游客对公园的满意程度，从积极情绪占比来看(图4)，鹅岭公园积极情绪占比一直保持在较高的水平。除2018年最低为92.27%，其余年份均高于年平均均值95.58%。2018年的消极情绪主要为“夜景一般，除了夜景外没什么好看的、公园普通、鹅岭拥挤杂乱、公园较小”，彼时鹅岭公园改造尚未开始，游客需求尚未被满足，总的来说，鹅岭公园经2018年、2021年的多次更新和修整，对游客积极情感的影响较小。

3.1.2 不同评价对象情感特征

研究根据不同月份对6类评价对象的评

表1 文本分类—文档级情感分析—属性级情感分析—感知要素提取示意图
Tab. 1 Illustration of the three deep learning models analyzing park network text

文本内容 Text	多标签文本分类 Multi-label text classification		评论观点抽取 Comment extraction	
	文档级情感倾向 Document-level sentiment analysis	积极情感概率 Positive emotion probability	感知要素 Perceptual element	
鹅岭公园交通很方便,一进门就是个大坡,左手边有一片水域,上面有廊桥;汽车可以进入,就是不知道是不是有什么要求;沿人行步道走一段就会有一个平台,可以坐着歇会儿;有个苏军烈士纪念碑,还有个瞰胜楼,登顶可以俯瞰重庆城市全景,作为城市公园还是不错的	建筑景观	0.99986	瞰胜楼,廊桥	
	人文历史景观	0.99932	纪念碑	
	园外景观	0.99965	重庆,全景,公园	
	设施配套	0.99988	交通	
鹅岭公园风景挺美的,可以登高望远,亦可放慢脚步观看园艺植物,曾经的私家园林,如今的开放公园,设计独特的绳桥,造型各异精心修剪的植物,别具一格的瞰胜楼,平时人不少,遇到菊花展等活动的时候更是人潮涌动,附近的鹅岭二厂也是不错的网红打卡地,唯一不好的就是不好停车,停在路边一不小心就多了一张罚单	自然景观	0.99991	风景,造型	
	人文历史景观	0.99991	设计,绳桥	
	设施配套	0.00020	停车	
天气凉快啦,去鹅岭公园爬个山,下了地铁跟着人群走个几分钟就到了,里面现在有几个地方在检修,不是全部开放,但也足够逛半天了,去了瞰胜楼,是看大重庆的好位置,等到晚上在顶层看夜景也是一绝,推荐外地朋友打卡。不知道是不是检修的原因,卫生间感觉设置不够,在里面没找着,出门右转外面路边有一个公共卫生间,还有一个不好的是公园里面晚上居然没有路灯,黑漆漆一片	建筑景观	0.99847	瞰胜楼,位置	
	园外景观	0.99975	夜景	
	设施配套	0.01608	卫生间,设置	
从李子坝站步行去了鹅岭公园二厂这边,其实可以打个车上来一路小上坡(就是爬山……),免得太累或者走错路……下来可以坐个公交(在二厂门口就有站牌),鹅岭公园就是个普普通通的公园……很多本地老年人在公园里打打牌下下棋唱唱歌散散心,景色也看不到啥,拍的话也都被树挡住了,不容易找到一个不那么挡视线的犄角旮旯……拍几张照片就撤了	公园活动	0.98551	散心,唱歌,下棋,打牌	

表2 公园6类评价对象说明
Tab. 2 Explanation of the six evaluation objects in the park

评价对象 Opinion target	说明 Description of the object of evaluation	文本数量/条 Text quantity
自然景观	公园内的自然要素	1 584
建筑景观	公园内的历史建筑及其周围的景观	2 115
人文历史景观	公园内具有人文历史的构筑物、园中园、雕塑等	779
园外景观	鹅岭公园地势较高,可以看到园外的城市景色	2 666
公园活动	公园举办的主题活动以及游人活动	1 596
设施配套	公园所提供的公共服务及设施	1 616

论数量进行统计如图5所示,游客游览呈现出较为明显的季节特征,集中在春秋两个季节,尤其是每年的5月与11月,与相关研究中城市公园最受关注的多为自然景观的结论相比,鹅岭公园的园外景观最受到游客关注,是该公园的一大优势景观。

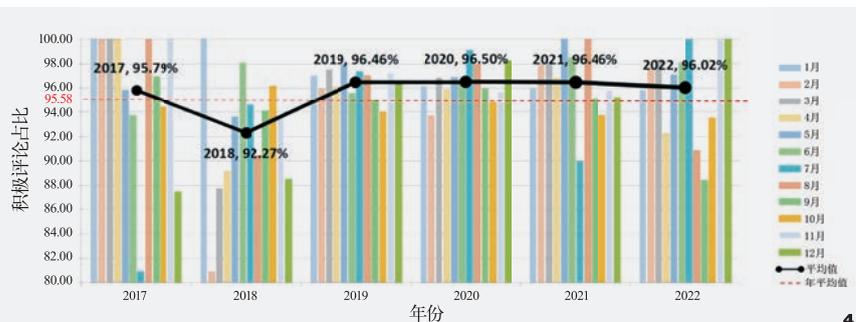
评价条数越多代表游客关注度越高,根据多标签文本分类的结果,6类评价对象关注度排名为园外景观>建筑景观>设施配套>公园活动>自然景观>人文历史景观。积极评论占比排名为人文历史景观>自然景观>建筑景观>公园活动>园外景观>设施配套(表3)。

人文历史景观占比最高,设施配套占比最低。人文历史景观虽然受到关注最少,但积极评价占比也是最高的,说明鹅岭公园作为城市历史公园,历史景点受到人们喜爱,但在人文历史价值的发扬方面略有缺失,鹅岭公园内的外国使馆旧址、鹅项山庄等都是重要的历史资源,在公园活动方面可以增加历史课堂、演出历史剧目等项目,在设计上织补历史建筑周围的绿化肌理,构成较为完整的历史公共空间^[31]。计算6类评价对象的积极情感均值,并与其对应的关注度做皮尔逊相关性分析(表3),其中成正相关的是自然景观和公园活动,成负相关的为设施配套,设施配套关注度排名第三,积极情绪占比最低,是鹅岭公园消极情绪的主要来源,设施配套不

渝中区最美的公园非鹅岭公园莫属，进门就长期有各种当季鲜花绽放，家里的长辈基本都来过这里打卡拍照，游玩路线也不容易迷路，从左或者右都可以完整逛一圈。盆景园里的盆景让人好生喜欢呀，都是出自名家之手做出的。自然景观-欣赏-喜欢-正向 走到公园临江的位置有回廊可以休息，顺便能看到整个江北的城市风景，这么好的视野 园外景观-视野-好-正向 还免费，怎么可以错过呢！公园里的瞰胜楼也值得登上去，可以从独一无二的视角欣赏美丽的重庆。建筑景观-视角-瞰胜楼-美丽-正向 公园里古树众多，绿化养护管理非常棒。

评价对象-评价维度-评价观点词-情感倾向

3



4



5

图3 评论观点抽取模型—数据标注示意图
Fig. 3 Data annotation illustration for comment opinion extraction model

图4 鹅岭公园积极情绪占比年变化图
Fig. 4 Yearly variation of positive sentiment proportion in Eling Park

图5 不同评价对象月评价数量变化图
Fig. 5 Chart of monthly changes in the number of evaluations of different opinion targets

从总体感知要素词频来看，“瞰胜楼”和“夜景”两词词频远超其他词语，说明这两处景观最能给游客留下深刻印象。并且建筑景观和园外景观的感知要素重合度较高，都有“视野”“江景”“夜景”，瞰胜楼地理位置高，可同时提供建筑景观与园外景观，是鹅岭公园的一大特点。而公园内的公使馆、石室等其他建筑物的吸引力相对较弱；自然景观中空气和植物最受关注，评价内容多为“空气清新”“植物茂盛”“绿化很好”，“菊花”“银杏树”“枫树”是最能引起游客注意的植物，说明整体上自然景观较好，在细节上有特色的植物最容易得到关注；人文历史景观中“广岛园”的词频最高，说明富有异国风情的日式园林最能吸引游客评价，其次是“绳桥”，其独特造型尤其受到游客关注；设施配套的感知要素主要集中在交通工具和停车设施上，游客较为关注到达公园的方式是否便捷，轻轨和公交等交通工具是否能到达，以及停车位是否充裕等情况；公园活动中“拍照”是游客最常开展的行为，“菊展”作为主题活动受到游客关注。

从时序来看，以月份为单位对感知要素词频进行统计(图6)，其中自然景观感知要素呈现明显的季节特征，尤其是“花草”一词，其作为公园自然景观的基底最容易受到人们感知，并在每年的1~3月及9~10月出

表3 评价对象关注度与积极情感均值相关性

Tab. 3 Correlation between attention level of evaluation objects and mean positive sentiment

评价对象 Opinion target	关注度/% Attention	皮尔逊相关性 Pearson correlation	积极评价占比/% Proportion of positive comments
自然景观	15.30	0.929**	98.68
建筑景观	20.42	0.384	98.30
人文历史景观	7.52	0.150	99.49
园外景观	25.74	-0.589	98.04
设施配套	15.61	-0.948**	91.71
公园活动	15.41	0.894*	98.18

注：**在0.01级别(双尾)，相关性显著，*在0.05级别(双尾)，相关性显著。

齐全，停车不方便等容易引起消极情绪。

3.2 游客感知要素分析

3.2.1 6类感知要素特征

公园感知要素是引起游客发表评价的公园要素，对6类文本进行评论观点抽取分析，

得到公园感知要素和积极或消极情感概率，为了便于理解，后文都采用积极情感概率的数值进行计算与相关性分析。对公园感知要素进行词频统计，选取词频前10的公园感知要素统计数量，6类文本中词频前10的感知要素及总体词频如表4所示。

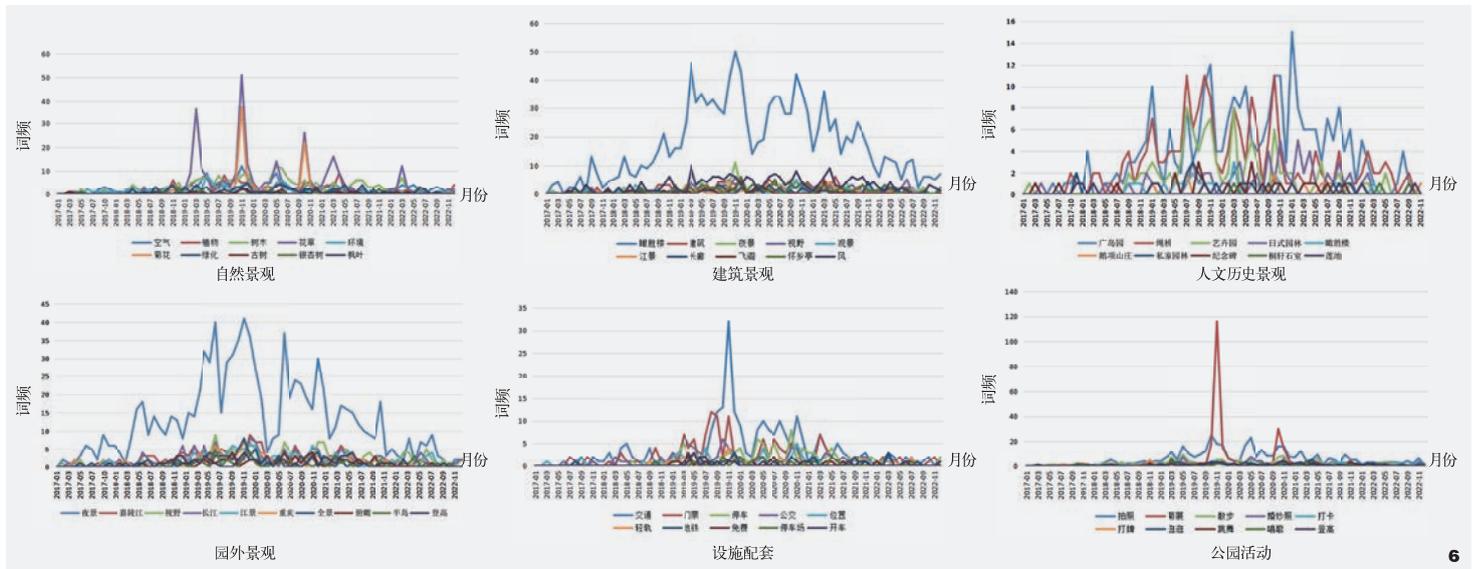


图6 6类感知要素月份统计折线图
Fig. 6 Line chart of monthly statistics of six categories of perception elements

表4 数据示例及6类评价对象前10的感知要素
Tab. 4 Data examples and the top ten perceptual elements of six categories of evaluation objects

文本内容 Text	识别评价对象 Identify and evaluate objects	积极情感概率 Positive emotion probability	公园感知要素及词频 Park perception elements and word frequency	评价对象 Opinion target	感知要素词汇及总体词频 Perceptual element vocabulary and overall word frequency
鹅岭公园交通方便, 位置好找, 绿树成荫, 鸟语花香, 近期举办的菊展, 更是人山人海, 菊展的品种很多, 有百八十种, 几万盆菊花造型艺术很美, 瞰胜楼可以观望两江美景, 还不收费用都是免费观赏, 是个值得游览观光拍照打卡的好地方	自然景观	0.100	拍照 (1)、菊展 (2)、散步 (0)、婚纱照 (0)、打卡 (1)、打牌 (0)、逛逛 (0)、跳舞 (0)、唱歌 (0)、登高 (0)	自然景观	空气(166)、植物(162)、树木(130)、花草(107)、环境 (95)、菊花 (87)、绿化 (86)、古树 (21)、银杏树 (16)、枫叶 (14)
				建筑景观	瞰胜楼(1274)、建筑(103)、夜景(101)、视野(81)、观景 (64)、江景 (44)、长廊 (42)、飞阁 (31)、怀乡亭 (31)、风 (26)
	人文历史景观		广岛园 (283)、绳桥 (133)、艺卉园 (91)、石绳桥 (62)、日式园林 (47)、鹅项山庄 (41)、瞰胜楼 (26)、纪念碑 (25)、私家园林 (21)、桐轩石室 (21)		
	园外景观		夜景 (937)、嘉陵江 (157)、视野 (156)、长江 (149)、江景 (124)、重庆 (81)、全景 (75)、俯瞰 (63)、渝中半岛 (53)、登高 (50)		
	设施配套		交通 (256)、门票 (155)、停车 (71)、公交 (59)、位置 (55)、轻轨 (26)、地铁 (25)、免费 (23)、停车场 (22)、开车 (17)		
	公园活动		拍照 (407)、菊展 (250)、散步 (134)、婚纱照 (86)、打卡 (60)、打牌 (59)、逛逛 (54)、跳舞 (44)、唱歌 (36)、登高 (35)		

现峰值, 除此之外, 每年秋季的“菊花”最常被游客感知, 与公园活动下的“菊展”呈现了相近的峰值时间, 也说明“菊展”是游

客最关心的公园活动, 作为鹅岭公园多年来的传统活动, 菊展曾停办过一段时间, 于2019年重启后词频便达到峰值。建筑景观中

“瞰胜楼”词频一直较高, 从2017年开始即是游客最易感知的要素, 是公园里重要的建筑景观。人文历史景观中, 除“莲池”外各

表5 正负感知要素相关性结果
Tab. 5 Table of positive and negative perceptual factors and ρ -values

评价对象 Opinion target	正相关感知要素 Positive correlation perception elements	负相关感知要素 Negative correlation perception elements
自然景观	—	—
建筑景观	江景** (0.073)	—
人文历史景观	私家园林** ((0.130)	—
园外景观	视野** (0.073)	—
设施配套	门票* (0.086)	交通* (-0.089)
公园活动	散步* (0.066)	菊展** (-0.098)

注：**在0.01级别(双尾),相关性显著,*在0.05级别(双尾),相关性显著;括号内为斯皮尔曼相关系数。

历史景点词频较为稳定且接近,“广岛园”在其他历史景点频次低的时间段仍能保持较高词频,并于2019年后词频再度升高,“私家园林”是游客对于鹅岭的重要印象。园外景观中“夜景”超90%的时间都是最易感知的要素,“嘉陵江”和“长江”即是公园所能眺望之江景的名称,也常常被提及。

3.2.2 重点感知要素分析

重点感知要素是对游客积极情绪影响较大的公园要素,对每条评论的高频词词频与积极情感概率进行斯皮尔曼相关分析,分析和游客情绪呈显著性相关的重点感知要素如表5所示。

正相关感知要素说明这些感知要素与积极情绪概率成正比,是公园建设较好的部分,可总结为“江景壮阔”“像私家园林一样”“视野开阔”“门票免费”“适合散步”。其中“江景”和“视野”两个感知要素从不同角度证明了瞰胜楼所提供的观景体验能有效地提升游客积极情绪。“私家园林”代表游客对于鹅岭公园人文历史景观的总体印象。设施配套上,“门票”免费与游客积极情感成明显正比,而散步则是游客开展最多的活动。

负相关感知要素说明该感知要素与积极

情绪概率成反比,是公园之后应重点提升的部分,结合文本具体内容可总结为“交通不是很方便”“菊展没有那么丰富”,其中前者的评价多出现在2019年前,近两年对于交通的评价已经多为“交通便利”,说明随着城市的建设,交通已经逐步优化。后者还需要鹅岭公园在菊展形式及菊花种类上进行丰富,满足游客需求。

3.3 NLP定制模型在公园游客感知研究中的优点

3.3.1 公园游客情感分析提升至属性级

情感倾向分析模型提供的是文档级情感,而评论观点抽取定制模型可以提取属性级的情感,情感判断更准确。现有研究中多采用的是文档级情感分析,可以高效地分析整段文字的情感,但文字中常涉及到多个评价对象,时常出现整体情感积极但个别消极情绪被忽略的判断。而后者提供的属性级情感倾向分析功能,可以分析句子中多个评价对象的情感,使游客感知情感分析精细化。如针对鹅岭公园的同一段网络文本,评论观点抽取可以提取出多类评价对象,并针对每一类评价对象给出情感倾向分析,第一段文字情感倾向分析模型判定为消极情感,积极

概率仅有0.39537,但评论观点抽取模型分析后判定“建筑景观”和“园外景观”情感为积极,且概率均大于0.99,由此可见属性级情感结果更为精准(表6)。

3.3.2 公园感知要素提取更深入

评论观点抽取模型使每句话的情感与公园感知要素一一对应,便于分析两者关联性。既有研究中常采用Excel搜索的方法计算感知要素词频,忽略了句子情感与感知要素的联系,而经过训练的定制模型可以根据输入的公园评论对象,智能提炼出每句话的游客感知要素、评论观点词以及属性级情感。如表7中的第二段文字,通用模型提取的感知要素仅有“地点”,而NLP定制模型可以具体到“长江、嘉陵江、夜景”,可以更精确分析情感与感知要素的相关性。

挖掘公园感知要素更全面。在既有研究^[32,33]中,对于感知要素常采用“高频词提取—词汇分类—情感分析”的分析顺序,同样的方法若应用在此次研究中,可能会遗漏一些词频较低的重要感知要素,如人文历史景观下的“私家园林”,总词频仅为21次,若采用传统词频排序则可能因为排名靠后被忽略。本研究采取“文本分类—高频词提取—情感分析”的顺序,则能让词频较低但重要的感知要素突显,使分析结果更加准确。

4 结语与展望

研究通过NLP定制模型对鹅岭公园多类景观游客情感倾向及其感知要素进行分析,得出鹅岭公园游览呈现明显季节性特征,春秋游客多,且游客一直保持较高水平的积极情绪,设施配套是主要的消极情绪来源。公园特色优势景观为园外景观,人文历史景观较为弱势。

表6 情感倾向分析模型与评论观点抽取定制模型分析游客情感结果对比

Tab. 6 Comparison of tourist sentiment analysis results between the sentiment analysis model and the comment opinion extraction custom model

文本内容 Text	使用模型 Model	评价对象 Opinion target	积极情感概率 Positive emotion probability	情感倾向 Emotional tendency
天气凉快啦，去鹅岭公园爬个山，下了地铁跟着人群走个几分钟就到了，里面现在有几个地方在检修，不是全部开放，但也足够逛半天了，去了瞰胜楼，是看大重庆的好位置，等到晚上在顶层看夜景也是一绝，推荐外地朋友打卡。不知道是不是检修的原因，卫生间感觉设置不够，在里面没找着，出门右转外面路边一个公共卫生间，还有一个不好是公园里面晚上居然没有路灯，黑漆漆一片	定制情感倾向分析	—	0.395	消极
	定制评论观点抽取	自然景观	—	—
		建筑景观	0.998	积极
		人文历史景观	—	—
		园外景观	0.100	积极
		设施配套	0.001	消极
		公园活动	—	—
鹅岭公园风景挺美的，可以登高望远，亦可放慢脚步观看园艺植物，曾经的私家园林，如今的开放公园，设计独特的绳桥，造型各异精心修剪的植物，别具一格的瞰胜楼，平时人不少，遇到菊花展等活动的时候更是人潮涌动，附近的鹅岭二厂也是不错的网红打卡地，唯一不好的就是不好停车，停在路边一不小心就多了一张罚单	定制情感倾向分析	—	0.998	积极
	定制评论观点抽取	自然景观	0.100	积极
		建筑景观	—	—
		人文历史景观	0.100	积极
		园外景观	—	—
		设施配套	0.040	消极
		公园活动	—	—

表7 评论观点抽取通用模型与定制模型分析公园网络文本效果对比

Tab. 7 Comparison of park network text analysis results between the general comment opinion extraction model and the custom comment opinion extraction model

文本内容 Text	评论观点抽取模型 Review view extraction model	评价对象 Opinion target	感知要素 Perception elements	评价观点词 Evaluative viewpoint word	情感倾向 Emotional tendency
一个重庆本地大爷大妈占领的公园，里面全是打麻将和扑克的，还有跳舞的唱歌的，非常巴适，不过在这美丽的公园旁边还有个豪宅，好像叫鹅岭公寓啥的，公园环境很好，而且可以找到合适的角度看到江边，空气也是非常得好，适合本土老人生活	定制模型	公园活动	扑克、跳舞、唱歌	适合	正向
	通用模型	—	环境、感觉	好	—
到重庆必须看夜景，鹅岭公园是俯瞰长江和嘉陵江的绝佳地点。鹅岭公园市中心，坐地铁到李子坝站往前走就可以到达。鹅岭公园地处于长江、嘉陵江南北挟持而过的陡峻狭长山岭上，因其形似鹅之颈项，故而得名。我们去的时候起雾了，就没有登上瞰胜楼。站在鹅岭公园平台上等待夜幕降临时，不经意目睹了传说中的李子坝站地铁穿过居民楼的情景。而老天爷也垂青我们，拍摄夜景时雾散了，让我看到了灯光照耀下的嘉陵江和长江，璀璨夺目，形成一道灯的海洋	定制模型	园外景观	长江、嘉陵江、夜景	绝佳	正向
	通用模型	—	地点	不错	—

然而，随着城市建设加快，对于鹅岭这样资源丰富的城市历史公园，固然应该继续提升现有的特色景观，但也应将建设力量向历史景点倾斜，鹅岭公园有较多的历史建筑，从研究结果可发现游客对于人文历史景

观的积极情绪占比最高，但评价数量最少，应加强对于人文历史文化的宣传，对人文历史景观进行重点提升与打造，并开展相关教育宣传活动，实现活态化保护。

从研究适用性来看，本文采取的深度学习

习三类定制模型，通过专业数据训练，使模型更适用于风景园林专业领域的文本分析，提升了情感分析的等级，并于每类评价对象下挖掘10个高频感知要素，以及7个重点感知要素。从多方面以量化的方式挖掘了公园特色景

观和亟需提升的景观，专业适用性较强。

从研究局限性来看，主要体现在以下几点：(1) 研究数据方面。网络评论作为一种有门槛的评价方式，本身存在局限性，会使用网络发布景点评价的用户多为年轻人，游客群体有限，挖掘出的信息也有限，难以反映其他年龄阶段人员对公园的意见，且数量有限，在相关性分析时可能会因为数据不足导致相关系数较低。(2) 研究方法方面。本文虽训练了三个深度学习模型，但评论观点抽取模型的评价观点词功能使用较少，在未来的研究中还可应用该功能，进一步探索观点词和公园感知要素的关联，挖掘公园网络文本中存在的信息，拓展智能算法在风景园林中的应用深度。(3) 感知要素方面。影响游客感知的要素有多种，除了本文所提到的，还包括使用者的性别、年龄以及公园周围环境的影响等，在之后的研究中应将这些要素也纳入考量。

本研究通过深度学习平台训练的NLP模型，实现了更精细化的情感分析和感知要素提取。当前，可利用NLP定制模型更精准地分析游客情感，使设计师能将网络意见落实到相应景观上并作出调整措施。未来还可以将其集成为针对风景园林的文本分析软件，相较传统软件ROSTCM6，其对风景园林领域的文本更具针对性，并且可以通过大量数据的标注，提升对于风景园林领域文本的理解能力和处理能力，为风景园林设计更快更好地提供使用者的意见。

注：文中图表均由作者绘制。

参考文献

- [1] 曹雨平, 刘松. 基于网络文本分析的大运河遗产游客感知研究——以常州市青果巷历史文化街区为例[J]. 江苏海洋大学学报(人文社会科学版), 2022, 20(01): 134-140.
- [2] 徐虹, 张行发. 基于主成分分析和熵值法的历史文化街区游客感知评价研究——以天津市五大道为例[J]. 天津商业大学学报, 2021, 41(03): 3-11.
- [3] 李经龙, 朱敏. 武夷山国家公园投射形象与游客感知形象比较研究[J]. 资源开发与市场, 2023, 39(01): 114-119.
- [4] 任庆柳, 杨兆洋, 韩芳, 等. 国家公园“资源价值”与“游客感知”对比研究——以大熊猫国家公园四川片区为例[J]. 中国科学院大学学报, 2023, 40(03): 333-342.
- [5] 张佳宝, 乌思. 基于游客感知价值的国家公园游憩功能研究——以黄石国家公园和武夷山国家公园为例[J]. 世界地理研究, 2023, 32(02): 146-157.
- [6] 刘巧辉, 王小平, 刘晶岚. 森林公园游客感知价值、满意度与环境责任行为的关系研究[J]. 生态经济, 2022, 38(02): 137-141.
- [7] 余尤骋, 程南洋. 森林公园生态环境服务的游客感知评价研究——以江苏省7家森林公园为例[J]. 林业经济, 2020, 42(08): 39-49.
- [8] 王妍方, 何平鸽. 基于森林环境禀赋与游客感知的康养旅游研究[J]. 西南林业大学学报(社会科学), 2021, 5(04): 47-50.
- [9] 郭苏明, 冯美玲. 基于游客感知的乡村景观评价因子分析——以苏州市树山村为例[J]. 西北林学院学报, 2022, 37(01): 253-258.
- [10] 王铭杰, 孟凯, 张世泽, 等. 乡村性和乡村旅游吸引力: 基于游客感知视角的要素结构辨识和认知机制解析[J]. 热带地理, 2021, 41(06): 1325-1337.
- [11] 谢征. 习近平关于城市治理重要论述的科学内涵与实践路径[J]. 观察与思考, 2021(02): 36-43.
- [12] 李念, 陈其兵, 姜涛. 基于网络数据的国家湿地公园游客感知分析[J]. 地域研究与开发, 2020, 39(01): 89-93.
- [13] 王中德, 张少丽, 杨玲. 基于网络文本分析的重庆市网红打卡点空间感知研究[J]. 西部人居环境学刊, 2022, 37(03): 126-131.
- [14] 张瑞, 张建国. 基于网络文本与IPA模型分析的上海辰山植物园旅游形象感知研究[J]. 中国园林, 2019, 35(08): 83-87.
- [15] 谭红日, 刘沛林, 李伯华. 基于网络文本分析的大连市旅游目的地形象感知[J]. 经济地理, 2021, 41(03): 231-239.
- [16] PLUNZ R A, ZHOU Y, CARRASCO VINTIMILLA M I, et al. Twitter Sentiment in New York City Parks as Measure of Well-being[J]. Landscape and Urban Planning, 2019, 189: 235-246.
- [17] HEIKINHEIMO V, TENKANEN H, BERGROTH C, et al. Understanding the Use of Urban Green Spaces from User-generated Geographic Information[J]. Landscape and Urban Planning, 2020, 201: 103845.
- [18] KOBLET O, PURVES R S. From Online Texts to Landscape Character Assessment: Collecting and Analysing First-person Landscape Perception Computationally[J]. Landscape and Urban Planning, 2020, 197: 103757.
- [19] 戴代新, 陈语娴. 基于情感计算的城市历史公园更新改造文化服务感知评价——以鲁迅公园更新改造为例[J]. 同济大学学报(社会科学版), 2022, 33(03): 81-90.
- [20] HOU Z, CUI F, MENG Y, et al. Opinion Mining from Online Travel Reviews: A Comparative Analysis of Chinese Major OTAs Using Semantic Association Analysis[J]. Tourism Management, 2019, 74: 276-289.
- [21] 栗恩波, 王静文. 基于自然语言处理的北京城市公园使用后评价[J]. 中国城市林业, 2023, 21(03): 101-108.
- [22] 李凤娇, 张书颖, 刘家明, 等. 京津冀和长三角城市群的旅游形象感知对比研究[J]. 经济地理, 2023, 43(04): 194-205.
- [23] 罗俊杰, 雷泽鑫, 胡一可, 等. 基于深度学习的城市公园游客情感倾向分析——以天津水上公园为例[J]. 中国园林, 2021, 37(09): 65-70.
- [24] 王炯. 自然语言处理技术在建筑使用后评价中的应用[J]. 南方建筑, 2019(01): 82-87.
- [25] 李舒扬, 李晓宇, 刘芳. 基于社交媒体数据的城市公共空间审美意象研究——以深圳地区为例[J]. 新建筑, 2022(06): 31-36.
- [26] 叶林, 江伦, 韩贵锋. 基于文本情感分析的城市公园使用感知评价研究——以重庆36个公园为例[J]. 西部人居环境学刊, 2022, 37(04): 147-154.
- [27] MEDHAT W, HASSAN A, KORASHY H. Sentiment Analysis Algorithms and Applications: A Survey[J]. Ain Shams Engineering Journal, 2014, 5(4): 1093-1113.
- [28] TSYTSARAU M, PALPANAS T. Survey on Mining Subjective Data on the Web[J]. Data Mining and Knowledge Discovery, 2012, 24(3): 478-514.
- [29] TANG H, TAN S, CHENG X. A Survey on Sentiment Detection of Reviews[J]. Expert Systems with Applications, 2009, 36(7): 10760-10773.
- [30] 肖竞, 张晴晴, 张芮珠, 等. “城园关联”视角下城市历史公园遗产价值的多尺度解析与双系统识别——以重庆市鹤岭公园为例[J]. 中国园林, 2022, 38(05): 56-61.
- [31] 周向频, 陈路平. 基于遗产价值评估的历史公园更新策略[J]. 风景园林, 2021, 28(06): 50-55.
- [32] 李娅琪, 袁鸿菲. 新冠肺炎疫情前后游人感知对比研究——以武汉东湖风景区为例[J]. 西北林学院学报, 2022, 37(03): 268-274.
- [33] 王琳, 白艳. 基于网络点评的城市公园使用后评价研究——以合肥大蜀山森林公园为例[J]. 中国园林, 2020, 36(06): 60-65.