

基于 UGC 数据的南岳衡山旅游目的地形象感知研究

陆利军^{1, 2, 3} 廖小平¹¹

(1. 中南林业科技大学 旅游学院, 中国湖南 长沙 410004;

2. 湖南工学院经济与管理学院, 中国湖南 衡阳 421008;

3. “聚落文化遗产数字化技术与应用”湖南省重点实验室, 中国湖南 衡阳 421001)

【摘要】: 目的地旅游形象是吸引旅游者的关键因素, 而旅游者生成的 UGC 数据则在相当程度上较为精准地反映了旅游者对旅游目的地的偏好、兴趣和评价。以南岳衡山为研究对象, 以百度旅游等旅游门户网站爬取的游记和评论为研究素材, 采用文本分析与基于词向量的情感分析方法, 基于目的地形象构建的“认知—情感”三维模型, 分别从认知形象、情感形象和整体形象三个方面探讨了旅游者对其旅游目的地形象的感知。结果表明: ①根据词频统计分析、关键词聚类方法得出, 旅游者对南岳衡山的认知可归纳为人物、自然、交通、活动、建筑、文化、场所/地域等七个维度; ②根据情感和语义分析方法得出, 旅游者对南岳衡山的积极形象感知要素包括佛道共存、心愿之旅、五岳独秀、四季佳景、福寿文化、冰雪奇缘、精妙建筑、文明奥区和抗战名山等九个核心类属, 消极形象感知要素则包括服务、资源、环境、安全、体验和设施等 6 个核心类属、19 个二级子类; ③基于词向量的情感分析方法得出, 旅游者对南岳衡山的整体形象感知是积极的, 综合正面评论占比 91.58%。

【关键词】: UGC 数据 目的地形象 形象感知 旅游者 景区设施 南岳衡山

【中图分类号】: F592.68 **【文献标志码】:** A **【文章编号】:** 1000-8462 (2019) 12-0221-09

随着旅游市场日益成熟, 旅游目的地之间的竞争也愈发激烈。其中, 目的地形象更是“兵家必争之域”。如何通过目的地特色化、差异化发展来保持自己的竞争力以吸引更多的旅游者, 成为很多旅游目的地管理者、营销者和从业者所面临的一项严峻挑战。鉴于已有研究成果对目的地感知形象和体验质量等因素与游客行为意向、旅游地选择影响之间相关关系的充分验证^[1, 2], 本文基于目的地形象建构“认知—情感”模型, 采用旅游者生成的 UGC 数据, 更精确地把握旅游者对旅游目的地的感知形象, 并在其心目中树立积极正面的、富有吸引力的形象, 是旅游目的地将自己与对手区别开来, 进而赢得竞争的关键环节。

学界对目的地形象的关注已近半个世纪^[3]。1975 年, Hunt 率先提出了目的地感知形象这一基本概念, 并认为目的地感知形

作者简介: 陆利军(1983-), 男, 湖南衡阳人, 博士研究生, 副教授。主要研究方向为森林游憩与公园管理。

E-mail: rock7409@163.com。廖小平(1963-), 男, 湖南邵阳人, 博士, 教授, 博士生导师。主要研究方向为生态伦理。

E-mail: lxpjng6@163.com。

基金项目: 湖南省哲学社会科学基金项目(16YBA123); “聚落文化遗产数字化技术与应用”湖南省重点实验室开放基金项目(JL19K03)。

象是纯粹主观的概念，即外界作用于人脑所形成的意识^[4]。此后，国内外学者对目的地感知形象的内涵和外延^[5,6]、形成机理与建构过程^[7]、影响因素与测量评价^[8,9]等内容开展了深入研究。

就目的地感知形象基本结构而言，早期的研究只局限于感知形象的认知成分^[10]，之后，学者们逐渐意识到，除了个人对目的地的属性的认知评价外，还应包括人们对目的地的情感评价^[11,12]。随着研究的逐步深入，Baloglu 等首先提出目的地形象“认知—情感”三维模型，该模型认为，目的地感知形象是由认知形象、情感形象和整体形象等三维结构组成的，而目的地感知形象的认知则主要应当从功能、心理和整体等三个维度切入^[7]。其中，认知形象指人们实际掌握的或自认为了解的有关特定目的地的知识；情感形象是指人们基于这些知识对目的地产生的感觉，而此两种成分共同构建了游客内心对于目的地的“整体画像”。

在目的地感知形象的评价内容上，虽然学界对旅游形象是由认知形象、情感形象和整体形象等三维结构组成这一论断已经基本达成共识，但是，当具体到对目的地感知形象的测量与评价时，重“认知分析”而轻“情感形象”与“整体形象”的现象依旧不在少数。近年来，学界才逐渐意识到情感成分对游客满意度和重游意向有更大、更直接的影响^[13,14]。如有学者在“认知—情感”二阶因子结构模型对目的地感知形象进行系统阐述的基础上提出，“双管齐下”更有助于理解目的地感知形象^[15]。但是，学界对于旅游目的地整体形象感知研究仍然更加习惯于“只取其一”^[16]。而在目的地感知形象的评价方法上，国内外关于目的地感知形象的研究多以统计数据、专家咨询和问卷调查为主要途径获得分析数据，运用多向度量度法、对应分析、语义差别量表分析法和认知地图等量化或结构式方法，与项目网格法、照片导出法、内容分析法和自由揭露法等质化或非结构方法分析目的地形象的影响因素^[9]。随着人机交互协作的 Web2.0 时代到来和自带定位功能的移动通讯设备的普及，网络用户生成内容（User Generated Content, UGC）与日俱增^[17]。学界亦开始从旅游者角度出发，以旅游数字足迹、地理标记照片等自发地理信息（Volunteered Geographic Information, VGI）为数据来源，采用大数据收集、挖掘和内容分析方法，有针对性地采集旅游者信息，运用语义定位和机器学习等方法构建目的地形象测量框架，以此获得旅游者对目的地的认知形象、情感形象和整体形象，取得较为丰富的研究成果^[18,19,20]。

综上，目的地形象“认知—情感”三维模型已经得到了学界的基本认同，且以 UGC 数据为代表的新型数据的迅速增长也为目的地感知形象研究提供了不一样的研究视角和新的数据来源^[21]。近年来，学者们也借助以 UGC 数据为主的各种新型数据，从认知心理学、计算科学等学科背景出发，构建了诸多的目的地形象测量模型，分别探索了目的地感知形象认知因素、积极或消极的情感因素和旅游目的地整体感知形象。但对于如何利用旅游者在互联网中产生的海量数据捕捉旅游者偏好等情感信息，继而获得旅游者对旅游目的地的整体形象评价涉及不多^[22]。因此，本文以南岳衡山为案例，基于目的地形象“认知—情感”三维模型，以运用 Python 工具从百度旅游等旅游门户网站爬取的游记和评论为研究素材，采用文本分析与词频统计分析方法、语义分析方法和基于词向量的情感分析方法，分别从认知形象、情感形象和整体形象等三个方面对其目的地形象感知进行系统探索，为旅游目的地整体形象感知拓展研究方法，并为南岳衡山旅游形象提升提供模型参考和分析依据，具有一定的理论和实践意义。

1 数据来源与研究方法

1.1 数据来源

基于目的地形象“认知—情感”三维结构模型，根据认知形象、情感形象和整体形象的自身特点^[7]，将分析的对象分别聚焦于 UGC 数据中的网络游记与旅游评论（表 1）。其中 UGC 数据中包含的描述性名词可视为对旅游目的地的简单描述，视为旅游者对旅游目的地认知形象的感知；带有情感属性旅游评论则反映了旅游者对旅游目的地的积极或者消极的情感形象感知^[23,24]；借助极性情感分析模型，利用游记+评论构成的海量数据捕捉游客满意度等情感信息，继而获得游客对旅游目的地的整体评价获取用户整体评价。

表 1 目的地感知形象研究方法示意图

序号	语料来源	分析方法	形象感知维度
1	游记	高频名词	认知形象
2	评论	高频形容词	情感形象
3	游记+评论	极性情感分析模型	整体形象

1.1.1 网络游记

第一步，数据的抓取。使用 Python 工具从如下旅游门户网站爬取 2018 年 3 月 31 日以前的所有游记资料（表 2）。

表 2 南岳衡山旅游游记数据

序号	网站	篇数	序号	网站	篇数	序号	网站	篇数
1	百度旅游	277	4	马蜂窝	1315	6	去哪儿	186
2	驴妈妈	5	5	游多多	251	7	途牛	4
3	携程	435						

由于所获游记数量较多，故以下列标准提取数据：(1) 自游记发布以来点击量超过 500 次和 2015 年 1 月以来点击量超过 100 次且回复超过 5 条；(2) 游记对南岳衡山旅游有细致的和带有情感倾向的描述。经过筛选，最终得到 1203 篇网络日志。

第二步，数据的清洗。(1) 去除内容的 html 标签，再过滤掉同一作者发布的相同的游记和游记内容为纯图片的；(2) 去除获取到网页文字中的导航文字、地址文字、官网介绍文字；(3) 去除英文、空格和一些标点符号。经过数据清洗后获得有效游记为 1188 篇（共计 167.87 万字符）。

1.1.2 网络点评

使用 Python 工具从如下旅游门户网站爬取 2018 年 3 月 31 日以前的所有旅游点评资料，共得到 7768 条网友旅游点评，其中好评 6954 条，好评率 89.52%（表 3），共统计出对到南岳衡山旅游的 814 条负面反馈（因“社交积极倾向”，乐观评价偏多^[25]）。

1.2 分析方法

1.2.1 认知形象分析过程

表 3 南岳衡山旅游评论数据

序号	网站	评论数	好评数	好评率/%
1	百度旅游	606	523	86.3
2	大众点评	455	361	79.3
3	驴妈妈	1847	1797	97.3
4	马蜂窝	753	628	83.4
5	猫途鹰	196	153	78.1

6	去哪儿	364	328	90.1
7	途牛	14	12	85.7
8	携程	1562	1358	86.9
9	游多多	1971	1794	91.1

第一步：使用结巴中文分词工具对经过数据清洗后获得的 1188 篇与南岳衡山有关的有效游记进行分词处理；第二步：使用分词后的文本训练 word 2vec 模型；第三步：运用维基百科训练简体中文词向量；第四步：使用 sklearn 进行 k-means 聚类；第五步：统计各个维度的数量及分布情况。

1.2.2 情感形象分析过程

在采用自然语言分析工具对旅游评论文本进行处理和分析的基础上，运用扎根理论方法对评论进行提炼，逐一输入和归纳，经过三级编码过程，将评论凝练为若干个核心类属。

1.2.3 整体形象分析过程

基于 LSTM 搭建的文本情感分类的深度学习模型主要通过文本预处理、句子自动分词、载入情感词典、文本情感分类和模型结果检测等几个核心步骤来实现基于情感词典的整体形象感知。本文运用该模型分析南岳衡山游记，获得旅游者对南岳衡山的整体形象感知信息。

2 感知形象研究结果

2.1 认知形象

2.1.1 认知高频词

以过滤后的 1188 篇有效游记为研究基础，进行分析和统计，剔除了原数据中与内容无关的网址、昵称、域名等信息，共得到 11226 个高频词（频次>5），共计 361662 频次，平均词频为 32 次，表 4 为前 100 的高频名词信息汇总。

2.1.2 认知维度

将数据中词频大于 95 的 185 个高频名词（词频大于 95 的高频词共有 231 个，对上述 231 个高频词进行处理，将动词、形容词和连词剔除后共获得 185 个高频名词），运用开源程序 Gensim 训练 Word 2vec 模型，使用 K-means 算法进行词语聚类，整理出各分类维度所包含具体关键词，并分别统计样本文本在各维度间的数量及分布情况。共计处理总词频达 70456 项，总计得到 7 个认知内容维度：人物、自然、交通、活动、建筑、文化、场所/地域。

进一步通过计算机聚类分析程序，自动将所有高频名词按照 UGC 数据中的关键词进行维度分类，所得出的分析结果及各维度所占比例见表 5。

其中，自然类（34.93%）占比最高，词频为 24047 次，主要为南岳衡山的自然资源等。主要包括：冰雪、云海、日出等气候类旅游资源；花海、瀑布、流水等自然类旅游资源。

建筑类（20.34%）占比第二，词频为 14007 次，主要包括南岳大庙、忠烈祠、福严寺、祝融殿、牌坊、舍利塔、上封寺、铁佛寺、南台寺、祖师殿、胜利坊等历史建筑。

交通/设施类（15.39%）占比第三，词频为 10598 次，主要包括缆车、大巴、火车、摩的、环保车等景区内交通工具和火车站、高铁、车站、汽车、汽车站等景区内交通设施。

活动类（12.29%）占比第四，与交通/设施类差距较小，词频为 8457 次，主要为爬山、徒步、帐篷等反映南岳衡山旅游方式的活动。

文化类（6.42%）占比第五，主要包括菩萨、香烛、香火、烧香、道观、香炉、香火、佛教等南岳衡山最具盛名的文化景观，及以臭豆腐为代表南岳衡山饮食文化的旅游活动。

人物类（5.76%）占比第六，词频为 3962 次，主要包括美女、司机、老板、老板娘、师傅等旅游接待地人物形象、和驴友、朋友、学生、青年等旅游主要人群形象。

表 4 南岳衡山游记名词高频词统计表（前 100 位）

名词	词频	名词	词频	名词	词频	名词	词频	名词	词频
祝融峰	2442	雾凇	714	火车站	456	祖师殿	342	建筑	251
住宿	2233	大庙	697	太阳	444	淡季	333	大鼎	247
南岳	2209	公交	676	酒店	443	火车	324	西站	245
小吃	2179	上封寺	618	吃饭	430	学生	313	客栈	242
爬山	2159	香火	596	独秀	424	晚餐	311	亭	242
南天门	1980	寺庙	593	房间	418	抗日	300	沿途	235
风景	1543	索道	545	祝融寺	411	水	296	公路	231
南岳大庙	1464	大殿	536	步行	411	半山腰	295	灵芝	229
高铁	1407	雪	527	缆车	387	山峰	289	最高峰	228
长沙	1401	帐篷	526	汽车站	375	广场	286	水雾	226
半山亭	1177	书院	525	走路	370	穿岩诗林	284	拜佛	225
忠烈祠	1115	雾	517	景点	368	观日台	278	大门	225
徒步	951	路线	516	大巴	365	距离	273	旺季	224
广州	904	仙境	488	磨镜台	361	祝融殿	273	环境	220
老板	877	景色	487	水帘洞	360	佛教	270	缆车	220
景区	864	梵音	474	青年	358	体力	265	云海	218
五岳	858	寺庙	470	烧香	356	台阶	264	邨候	216
朋友	784	名胜区	464	牌坊	353	会仙桥	263	日落	216
日出	761	旅游者	462	宾馆	348	雨	261	岳麓	214
天气	726	麻姑	457	胜利坊	342	风景区	258	司机	213

表 5 南岳衡山旅游目的地认知形象维度一览表

序号	种类	词频	比重（%）	维度关键词
----	----	----	-------	-------

1	自然	24047	34.93	冰雪、诗林、景点、南岳、最高峰、景观、仙境、温度景色、雨、雪景、风景区、佛、延寿、日出、台阶、烟霞、云雾、阳光、邨候、南岳古镇、雾、空气、风、水、石阶、雾、环境、名胜区、风景、阶梯、云海、风光、梵音、景区、祝融峰、水帘洞、祝圣寺、云海、穿岩诗林、广场、磨镜台、雪、五岳、半山腰、海拔、麻姑、冬天、山峰、祝融寺、气温、天气、日落、自然、花、万寿、雾凇、瀑布、大门、灵芝、石头、流水、太阳、日出、紫竹林、独秀、天空、冰
2	建筑	14007	20.34	大鼎、亭、宫殿、岳麓、福严寺、寺庙、石鼓、观日台、南天门、南岳大庙、忠烈祠、半山亭、岳麓书院、观日台、会仙桥、祝融殿、牌坊、舍利塔、上封寺、铁佛寺、南台寺、观景台、建筑、庙宇、石碑、祖师殿、胜利坊、桤木潭、大殿、书院、华严寺、寺庙
3	交通/设施	10598	15.39	缆车、路线、大巴、火车、索道、摩的、西站、交通、火车站、高铁、车站、汽车、环保车、汽车站、线路、公路、公交、高速、地图、酒店、民宿、住宿、房间、客栈、宾馆、厕所
4	活动	8457	12.29	假期、晚餐、沿途、体力、祈福、照片、吃饭、淡季、旺季、爬山、农家乐、徒步、缆车、步行、拜佛、供奉、走路、帐篷
5	文化	4420	6.42	菩萨、香烛、香火、烧香、道观、香炉、香火、佛教、豆腐、小吃
6	人物	3962	5.76	美女、司机、老板、师傅、驴友、朋友、学生、青年、老板娘
7	场所/地域	3353	4.87	广东、深圳、长沙、距离、超市、广州、衡阳、武汉

场所/地域（4.87%）占比第七，词频为 3353 次，主要包含广东、深圳、长沙、广州、武汉、衡阳等南岳衡山代表性游客来源地区。

2.2 情感形象

2.2.1 南岳衡山目的地感知形象的积极感知因素分析

采用自然语言分析工具对 1188 篇网络日志进行处理和分析，并在此基础上运用扎根理论方法对网络日志进行提炼，可知，南岳衡山目的地感知形象的积极感知因素主要有 9 个，分别是：佛道共存、心愿之旅、五岳独秀、冰雪奇缘和抗战名山等（图 1）。其中南岳衡山的“佛道共存”在 1188 篇游记的 891 篇中提及，占到总样本的 75%，远远领先于其他积极感知因素；“心愿之旅”“五岳独秀”和“冰雪奇缘”可分为一组，对旅游者影响比较明显；其他“精妙建筑”“文明奥区”和“抗战名山”等三类积极感知因素，对旅游者的吸引力较为有限。

南岳衡山的“佛道共存”和“心愿之旅”无疑令旅游者印象最为深刻，为南岳衡山目的地感知形象的第一、第二大积极感知因素。其中，“佛道共存”在游记的 891 篇中提及，占到总样本的 75%；“心愿之旅”在 725 篇游记中提及，占到总样本的 61%。南岳之所以名扬天下，既在于她的自然风光，更在于她独具特色的三教一体的文化基因，如网友“瀚子”在游记写道：“儒、释、道三教共存一庙，这是我国乃至世界庙宇中绝无仅有”。

“五岳独秀”和“四季佳景”为南岳衡山目的地感知形象的第三、第四大积极感知因素。“五岳独秀”在游记的 582 篇中提及，占到总样本的 49%；“四季佳景”在 439 篇游记中提及，占到总样本的 37%。南岳衡山地临湘水之滨，享有“五岳独秀”美名。如网友“小號”在游记写道：“春赏花、夏观云、秋望日、冬赏雪，令人心旷神怡。”

“福寿文化”为南岳衡山目的地感知形象的第五大积极感知因素，在 392 篇游记中提及，占到总样本的 33%。南岳这片土地

孕育出了其无可复刻的“南岳寿文化”，享誉海内外，观“南岳寿文化”即可窥见“中华寿文化”的精华^[26]。

“冰雪奇缘”为南岳衡山目的地感知形象的第六大积极感知因素，在 369 篇游记中提及，占到总样本的 31%。高峰入云，寒气环舞，共同谱写了南岳衡山别样的清冽之美，其中冬季“雾凇冰雪”更是堪称一绝。如网友“阳光的味道”在游记写道：“衡山之旅，我有幸看到满山白雪、遍野雾凇，兴奋得神马火车折腾、上下山摔了 N 次、被冻成狗、手机相机冻死机啥的统统可以忽略……”

“精妙建筑”为南岳衡山目的地感知形象的第七大积极感知因素，在 321 篇游记中提及，占到总样本的 27%。南岳大庙素有“南国故宫”之称。如网友“来来小冉冉”在游记写道：“南岳大庙分为九进四重院落，四周围以红墙，角楼高耸，颇有故宫风貌。”

“文明奥区”为南岳衡山目的地感知形象的第八大积极感知因素，在 190 篇游记中提及，占到总样本的 16%。之所以被誉为“文明奥区”，是因为南岳衡山人文资源荟萃，远古文明、宗教文化、书院文化、寿文化、抗战文化等多种文化交融并存。

“抗战名山”为南岳衡山目的地感知形象的第九大积极感知因素，在 107 篇游记中提及，占到总样本的 9%。南岳产生了许多革命事件和历史人物。如网友“忙碌蚊子”在游记写道：“忠烈祠是大陆唯一一座纪念国民党抗日阵亡将士的大型烈士陵园。”

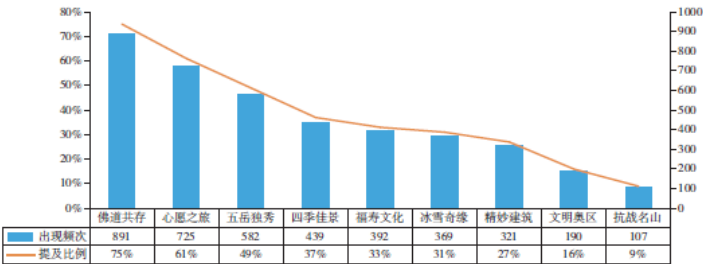


图 1 南岳衡山目的地感知形象的积极感知因素及其占比

2.2.2 南岳衡山目的地感知形象的消极感知因素分析

相比旅游者对南岳衡山目的地感知形象的积极感知，点评中所指出的南岳衡山的不足处则为南岳衡山旅游业发展提供了思路与方向。保持发扬自身优势自是不在话下，修复弥补短板劣势更是势在必行。在采用文本分析工具的基础上，使用扎根理论的方法对 814 条旅游者的负面反馈进行概念化、逐一登录并提炼，最终将负面反馈凝炼为六个核心类属、十九个二级子类。

在所列六个核心类属中，“旅游服务”类消极感知因素所占比列最高，为 24.8%。在“旅游服务”核心类属中，旅游者主要对南岳衡山工作人员的服务意识缺乏、人员素质不高和服务效率低下等有所不满。如网友“M20”在点评中写道：“很差！非常差的地方！每个工作人员服务态度超级差，就像我们欠了好几百万的表情。”可见，旅行的心境是极其不愿意被打破的，尤其是已然付费合理期待的服务。

在负面反馈中占比第二的核心类属是南岳衡山的旅游资源。旅游者认为，与中华五岳的地位相比，南岳衡山的资源表现有些名不副实。如网友“杰哥”在点评中写道：“衡山真不是一个来看风景的山，山不够陡峭，也不奇特，也不高，可以说平平无奇。”

在负面评论中占比第三的核心类属是南岳衡山的旅游环境。在有关南岳衡山环境问题的负面反馈中，旅游者的批评多指向南岳衡山的卫生环境。如网友“杰哥”在点评中写道：“衡山沿途卫生真差，满地的草鞋，各种垃圾，衡山的台阶路比较少，但就那几个台阶上的雪都没有工作人员清扫。”虽然在多数情况下卫生环境的破坏是由于部分旅游者乱扔垃圾造成的，但景区保护环境的宣传和管理不力也是原因之一，南岳衡山的卫生环境问题也越发严重。

在负面反馈中占比第四的核心类属是南岳衡山的旅游安全。在“旅游安全”核心类属中，“财产安全”问题更令旅游者担忧，其在负面反馈十九个二级子类中位列第五。无安全保障无异于变相“驱逐”游客。如网友“Saaaabrina”在点评中写道：“来往车辆频繁、车速快，而且道路崎岖，急弯比较多，人走在路上心惊胆战的。”

在负面评论中占比第五的核心类属是南岳衡山的旅游体验。旅游体验所蕴含的需求结构是构成旅游竞争力的重要因素^[27]，因此，思考如何丰富旅游者体验、探究了解进而多取向层次的满足旅游者需求是首要改善短板举措。当然，面对并不理智客观的评价，管理者依旧可以汲取事实层面的信息，能力范围内有所作为。诸如负面反馈中网友提到的因天气无缘日出，景区虽无呼风唤雨之能，但完善气象监测系统并预先告知旅游者还是可以做到的。杜绝非理性评价难如登天，降低比重还是大有可为。

在负面评论中占比第六的核心类属中包含了旅游者最为不满的二级子类问题，即南岳衡山的食宿问题。样本中 4%负面评论指向南岳衡山的食宿。如网友“M23”在点评中写道：“在旅馆里吃饭，一只鸡要 500，豆腐都 40 多……”另外，被旅游者广为诟病的景区厕所问题在南岳衡山同样存在，如网友“M23”在点评中写道：“厕所很少，并且大部分超级脏。希望管理方多拿点钱出来改善基础设施，要不然对不起门票价钱。”

2.3 整体形象

本文采用基于极性情感词库的情感分析技术对基于超过 1188 篇与南岳衡山相关游记数据判别南岳衡山旅游目的地整体形象。具体操作步骤如下：首先对游记进行自动分词，然后选择开源的“结巴中文分词”作为分词工具。另外为了情感的分析的准确性，从网上搜集了开源的停用词典，对分词后的数组进行处理，去除一些停用词（图 2）。

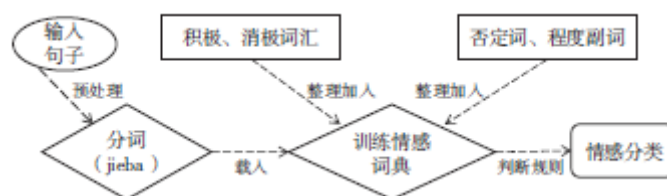


图 2 基于情感词典的极值分析流程图

注：引自 CSDN 专业 IT 技术社区。<https://blog.csdn.net/u012871493/article/details/72782744>, 2018-06-11.

基于情感词典的文本情感分类规则比较机械化^[27]。首先，将每个积极情感词语赋予权重 1，将每个消极情感词语赋予权重-1，并且假设情感值满足线性叠加原理；其次，将句子进行分词，如果句子分词后的词语向量包含相应的词语，就加上向前的权值，其中，否定词和程度副词会有特殊的判别规则，否定词会导致权值反号，而程度副词则让权值加倍。最后，根据总权值的正负性来判断句子的情感。

本文使用 KERAS 基于 LSTM 搭建一个深度学习模型，并运用该模型系统分析 1188 篇南岳衡山相关游记。其整体正面反馈 1088 条，整体负面反馈 100 条，综合正面评论占比 91.58%。即，总量中的 91.58%的游客对南岳衡山目的地感知形象的整体评价是积

极的、正面的。为了证明该结果的有效性，本文参照丁于思等人的做法^[28]，将测评结果与各旅游门户网站获得的南岳衡山旅游好评率做了简单对比，误差结果为±2.06%（表6）。

表6 南岳衡山目的地感知形象整体评价校验表

项目	方法	数据来源	结果/%
整体评价	深度学习模型	南岳衡山游记	91.58
网站点评	好评占比统计	南岳衡山点评	89.52

3 结论与讨论

3.1 结论

基于目的地形象构建的“认知—情感”模型，从认知形象、情感形象和整体形象等三个方面对目的地感知形象进行系统探索，这与目的地感知形象研究既有文献只分析认知形象，忽略情感形象和整体形象有所不同^[29]。论文证明，同时分析认知形象、情感形象和整体形象对更好地理解目的地感知形象大有裨益。

相对于统计数据、专家咨询和问卷调查等数据来源，以UGC数据为代表的新型数据为旅游者行为以及目的地形象研究提供了可信且相对易获取的数据^[20,30]，本文在运用游记和评论等研究素材的时候分别有所侧重。其中，用词频统计分析、聚类方法从游记中获得的描述性名词主要用于分析旅游者对旅游目的地认知形象的感知；用情感和语义分析方法从评论中获得的带有情感属性旅游评论则用于旅游者对旅游目的地的积极或者消极的情感形象感知；运用基于LSTM搭建的深度学习模型对由游记+评论组合构成的海量数据库进行系统分析获得的信息则用于捕捉游客满意度等情感信息，继而了解游客对旅游目的地的整体评价。

通过南岳衡山旅游形象感知研究发现，在认知形象上，旅游者对南岳衡山的基本认知仍然以自然风光和著名景点为主。在情感形象上，对南岳衡山的积极情感认知主要受到佛道共存、心愿之旅、五岳独秀、四季佳景、福寿文化、冰雪奇缘、精妙建筑、文明奥区和抗战名山等九大因素影响；对南岳衡山的消极情感认知主要受到包括旅游资源、旅游安全、旅游服务、旅游环境、旅游设施和旅游体验等六个核心类属的因素影响。在整体形象感知上，对南岳衡山的整体形象感知是积极的，综合正面评论占比较高。可见，旅游者基于对南岳衡山自然风光、人文景观、旅游设施设备等物理性、功能性要素的综合感官体验和对南岳衡山心理上的积极或消极情感倾向，共同形成了对目的地形象的整体评价，这一研究结论与Baloglu等提出的目的地形象“认知—情感”三维模型^[7]基本契合。

3.2 讨论

依托计算科学，基于以旅游者生成的UGC数据为主的新型数据，在更大的样本空间内对南岳衡山目的地感知形象进行了系统分析。然而，本文依然存在以下不足：（1）目的地情感评价分类采用人工判别UGC数据的方法，具有相当的主观性，后续研究需进一步完善语料库及相关模型的构建，提高语料分析的客观性和精准性；（2）研究数据仅限于国内各大旅游网站的文本数据，对图片数据和音频、视频数据并无涉及，故在后续研究中应引入诸如图片、语音和视频数据作为基础研究资料来源；（3）本文搭建了一个深度学习模型，但仅将该模型运算结果与各旅游门户网站的好评率进行简单对比验证，该模型在识别旅游者对目的地整体形象感知的有效性和可靠性有待进一步验证和改善。因此，关于目的地形象感知的后续研究，需不断改进对网络数据的分析应用方法。当然，其研究重点可拓展至认知形象、情感形象、整体形象三者关系等横向、共时性的研究，也应关注旅游者对目的地形象感知变化的纵向、历时性的研究。总之，旅游形象是吸引旅游者的关键因素，在当前日趋激烈的旅游市场竞争背景下，本文建议南岳衡山旅游目的地完善游览服务，优化旅游交通，保障旅游安全，整治环境卫生，加强旅游购物，同时挖掘地

方特色与文化, 提高从业人员服务质量, 推进综合治理管理, 以有效提升自身旅游形象。

参考文献:

- [1]Martín H S,Rodríguez del Bosque I A. Exploring the cognitiveaffective nature of destination image and the role of psychological factors in its formation[J]. Tourism Management,2008,29(2):263-277.
- [2]张红梅,梁昌勇,徐健,等. 特色旅游目的地形象对游客行为意愿的影响机制研究——以贺兰山东麓葡萄产业旅游为例[J]. 中国软科学,2016(8):50-61.
- [3]王素洁,刘海英. 国外目的地形象研究进展与展望——基于 TM、ATR、JTR 发表的相关文献[J]. 山东大学学报:哲学社会科学版,2017(1):137-147.
- [4]Hunt J D. Image as a factor in tourism development[J]. Journal of Travel Research,1975,13(3):1-7.
- [5]Pike S. Destination image analysis:a review of 142 papers from1973 to 2000[J]. Tourism Management,2002,23(5):541-549.
- [6]Gallarza M G,Saura I G,García H C. Destination image:towards a conceptual framework[J]. Annals of Tourism Research,2002,29(1):56-78.
- [7]Baloglu S,Mc Cleary K W. A Model of destination image formation[J]. Annals of Tourism Research,1999,26(4):868-897.
- [8]李宏. 旅游目的地形象测量的内容与工具研究[J]. 人文地理,2007(2):48-52.
- [9]刘国华,王红国. 旅游目的地形象测量:基于国外文献的研究[J]. 旅游学刊,2010,25(6):83-87.
- [10]黄震方,李想. 旅游目的地形象的认知与推广模式[J]. 旅游学刊,2002(3):65-70.
- [11]刘力. 旅游目的地形象感知与游客旅游意向——基于影视旅游视角的综合研究[J]. 旅游学刊,2013,28(9):61-72.
- [12]郭安禧,黄福才,杨晶,等. 目的地形象对感知吸引力及重游意向的影响——以厦门市为例[J]. 旅游科学,2015,29(6):50-67.
- [13]Lee S,Scott D,Kim H. Celebrity fan involvement and destination perceptions[J]. Annals of Tourism Research,2008,35(3):809-832.
- [14]王纯阳,屈海林. 旅游动机、目的地形象与旅游者期望[J]. 旅游学刊,2013,28(6):26-37.
- [15]Rio D,Nunes L Monitoring and evaluation tool for tourism destinations[J]. Tourism Management Perspectives,2012(4):64-66.

-
- [16]张宏梅, 陆林, 蔡利平, 等. 旅游目的地形象结构与游客行为意图——基于潜在消费者的本土化验证研究[J]. 旅游科学, 2011, 25(1):35-45.
- [17]赵振斌, 党娇. 基于网络文本内容分析的太白山背包旅游行为研究[J]. 人文地理, 2011, 26(1):134-139.
- [18]Stepchenkova S, Kim H, Kirilenko A. Cultural differences in pictorial destination images Russia through the camera lenses of American and Korean tourists[J]. Journal of Travel Research, 2014, 54(6):1 219-1 223.
- [19]Vu H Q, Gang L, Law R, et al. Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos[J]. Tourism Management, 2015, 46:222-232.
- [20]邓宁, 钟栎娜, 李宏. 基于 UGC 图片元数据的目的地形象感知——以北京为例[J]. 旅游学刊, 2018, 33(1):53-62.
- [21]方雨, 黄翔. 武汉市旅游目的地形象感知研究——基于网络文本分析方法[J]. 华中师范大学学报:自然科学版, 2017, 51(5):709-714.
- [22]刘逸, 保继刚, 朱毅玲. 基于大数据的旅游目的地情感评价方法探究[J]. 地理研究, 2017, 36(6):1 091-1 105.
- [23]Beerli A and Martin J D. Tourists' characteristics and the perceived image of tourist destinations:A quantitative analysis a case study of Linarite, Spain[J]. Tourism Management, 2004, 25(5):623-636.
- [24]熊伟, 高阳, 吴必虎. 中外国际高星级连锁酒店服务质量对比研究——基于网络评价的内容分析[J]. 经济地理, 2012, 32(2):160-165.
- [25]Qiu L, Lin H, Leung A K, et al. Putting their best foot forward:Emotional disclosure on facebook[J]. Cyberpsychology, Behavior, and Social Networking, 2012, 15(10):569-572.
- [26]蔡梅良, 钟志平. 南岳旅游目的地吸引力综合评价与对策研究[J]. 经济地理, 2010, 30(3):514-518.
- [27]胡均毅, 李金龙. 基于情感评分的分层文本表示情感分类方法[J/OL]. 计算机工程:1-7[2019-06-10]. <https://doi.org/10.19678/j.issn.1000-3428.0054521>.
- [28]丁于思, 肖轶楠. 基于网络点评的五星级酒店顾客满意度测评研究[J]. 经济地理, 2014, 34(5):182-186, 192.
- [29]Garrod B. Understanding the relationship between tourism destination imagery and tourist photography[J]. Journal of Travel Research, 2008, 47(3):346-358.
- [30]Stepchenkova S, and J E Mills. Destination Image:A MetaAnalysis of 2000-2007 Research[J]. Journal of Hospitality Marketing & Management, 2010, 19(6):575-609.
- [31]宋竹芳, 张高军, 李树民. 黑亦非黑:基于线上数据的黑色旅游体验及群体比较[J]. 旅游学刊, 2019, 34(5):90-104.