

大语言模型驱动的人居环境特征智能识别

——以武夷山建成环境文献分析为例

The Intelligent Recognition of Human Settlements Characteristics Driven by Large Language Models: A Case Study of Built Environment Literature Analysis in Wuyishan

李凤娇 马悦 龙瀛

LI Fengjiao, MA Yue, LONG Ying

摘要：文章以武夷山建成环境相关文献为基础，提出一种基于大语言模型的人居环境特征智能识别应用方法。通过对武夷山 2015—2025 年中英文文献的系统筛选与分析，构建了“自然生态与景观”“空间结构与开发强度”“居住与社区”“功能与服务配套”“交通与可达性”“社会文化与感知体验”六个建成环境核心维度。基于 ChatGPT-4.1 模型，自动提取文献中的建成环境特征词并进行情感评分，进一步利用社交媒体数据提取结果与人工标注结果进行双重对比。结果显示：大语言模型在特征词提取与情感分析方面具备较高的稳定性和准确性，其中“自然生态与景观”“社会文化与感知体验”成为武夷山建成环境的核心优势维度；“功能与服务配套”“交通与可达性”则为亟须提升的重点；“居住与社区”属于中长期优化方向。这不仅拓展了大语言模型在城乡规划领域的应用范围，也为空间优化、地方产业发展和生态文明建设提供了理论支撑。

关键词：大语言模型；人居环境；武夷山；文献分析

【中图分类号】TU984

【文献标志码】A

【文章编号】1005-278X(2025)06-0029-09

Abstract: Based on a comprehensive review of literature related to the built environment of Wuyishan, this study proposes a large language model-based method for the intelligent identification of human settlement features. Through a systematic screening and analysis of Chinese and English publications on Wuyishan from 2015 to 2025, six core dimensions of the built environment were established: natural ecology and landscape, spatial structure and development intensity, housing and community, functional and service facilities, transportation and accessibility, and socio-cultural perception and experiential quality. Using the ChatGPT-4.1 model, built environment feature terms were automatically extracted from the literature and subjected to sentiment analysis. Furthermore, the results were examined through a dual comparison framework, integrating features derived from social media data and manually annotated references. The findings indicate that the large language model demonstrates high stability and accuracy in feature extraction and sentiment analysis. Natural ecology and landscape, as well as socio-cultural perception and experiential quality, emerge as the key strengths of Wuyishan's built environment, while functional and service facilities and transportation and accessibility are identified as priority areas for improvement. The housing and community dimension represents a medium- to long-term optimization direction. Overall, this research expands the application of large language models in urban and rural planning and provides theoretical support for spatial optimization, local industrial development, and ecological civilization construction.

Keywords: large language models; human settlements; Wuyishan; literature analysis

作者简介：

李凤娇 清华大学建筑学院博士后，lifengjiao@tsinghua.edu.cn

马悦 清华大学建筑学院博士研究生，mayue22@mails.tsinghua.edu.cn

龙瀛 清华大学建筑学院教授，通信作者，ylong@tsinghua.edu.cn

基金项目：

地方任务项目“武夷山生态环境与人体健康关联性课题研究”（04150400924）

收稿日期：2025-07-31

武夷山因其独特的丹霞地貌、丰富的生物多样性、珍贵的古闽越文化遗存等自然与人文价值,于1999年被联合国教科文组织列入《世界遗产名录》,成为世界自然与文化双遗产地^①。2021年,凭借其重要的生态功能和突出的人文景观,武夷山入选首批国家公园,成为展示“美丽中国”的重要窗口^②。至此,武夷山市成为同时拥有“双世遗”与国家公园双重殊荣的城市。武夷山市地处福建省西北部,隶属南平市管辖,截至2023年末常住人口约26万人^③。该市人居环境具有生态优势明显、人口密度较低、生活节奏舒缓等特征,为康养产业发展提供了良好的基础。深入挖掘武夷山市人居环境的优质特征,对于推动地方康养产业高质量发展,助力区域生态文明建设具有重要意义。

人居环境是指人类聚居生活的整体空间系统,包括五大子系统:自然系统、人类系统、社会系统、居住系统、支撑系统^[1],这些系统是人居环境建设的基本要素和人居环境发展的广阔空间^[2]。其中,吴良镛^[3]将建成环境纳入居住系统与支撑系统,作为自然、社会、人类系统的物质基础。由此可见,人居环境包含建成环境,建成环境是人居环境的物质载体。武夷山市以其优质的人居环境著称,但当前从建成环境视角对其优质特征的研究尚存空白。因此,本文立足城乡规划领域,重点从建成环境视角识别武夷山市人居环境的优质特征,为地方产业发展与空间优化提供科学依据。

当前,以ChatGPT、DeepSeek等为代表的大语言模型在计算机^[4-6]、教育^[7-9]、文学^[10-12]、医学^[13-15]、经济学^[16-18]、管理学^[19-20]等多个领域研究中得到了应用,在计算机、教育和文学领域的应用尤为广泛。然而,通用大语言模型在特定领域任务上通常存在针对性不足的问题,近年来研究者倾向于通过微调方式提升模型在特定任务与语境中的表现^[4-6,11,16]。微调模型能够更好地适应专业语料、输出结构化结果,并提升特征提取与情感识别的准确性与一致性。在多种语言与问答面前,学者们更倾向于直接将大语言模型作为分析工具^[7,12]。已有研究表明,不论是否经过微调,与其他模型相比,ChatGPT模型的整体表现通常较强^[4,12,21]。例如,周丽君等^[21]通过对比微调后的ChatGPT-4.5与DeepSeek-V3在康复指导分析中的表现,发现前者在准确性、完整性和可靠性等方面明显优于后者,且语言生成更为灵活。此外,通过Web of Science和中国知网对ChatGPT和DeepSeek关键词的检索,发现ChatGPT在国内外研究中更为主流。从分析内容看,经济学与管理学领域的研究同样证实,ChatGPT在文本分析和情感分析任务中具有较高的适用性和稳定性^[17,20]。

从文本与情感分析方法看,已有研究多使用隐含狄利克雷分布(Latent Dirichlet Allocation, LDA)主题聚类等经典文本挖掘方法深度分析酒店用户偏好^[22]、旅游信息^[23]及网络舆情^[24]等,近年来,自然语言处理技术及ROST CM6等传统分析工具也被广泛应用于情感测度。然而,上述方法仍主要依赖词典或统计建模等实现特征识别,对于隐含语义、跨语境解释、抽象概念推理等能力相对有限。相比之下,ChatGPT则具备更强的深度语义理解、概念聚合、类推推理及多场景表达解析能力,同时,以上大语言模型相关文献分析也进一步反映了ChatGPT在复杂文本分析中的应用潜力。

综上所述,鉴于ChatGPT在复杂文本分析中具有较大优势,已广泛应用于多个领域,但在城乡规划领域特别是人居环境领域的系统性应用仍相对有限,缺乏基于多源数据的人居环境特征智能识别与对比分析。此外,通过梳理现有武夷山相关文献发现,已有研究多涉及建成环境的单一要素或单一维度,缺乏系统性探讨。针对以上不足,本文以武夷山市为代表性案例,围绕“基于多源数据的人居环境特征智能识别与对比”这一核心命题,探索大语言模型在分析异构文本数据、感知人居环境特征方面的应用潜力。首先,由于已发表的学术文献具有系统性、理论性及同行评议下的可靠性,本文以中英文文献为数据基础,利用ChatGPT内嵌的多个大语言模型开展建成环境特征提取与情感识别的对比实验。基于人工识别的优质特征作为参照,最终选择与人工识别结果最为接近的ChatGPT-4.1模型对文献全文进行特征词提取与情感分析。其次,利用大模型对社交媒体数据进行特征提取与情感打分,并与文献提取维度进行对齐,进而对多源异构文本数据分析结果对比分析。最后,将人工识别与大模型分析文献结果进行一致性量化分析,以评估模型分析结果的有效性与准确性。本文旨在验证大语言模型在多源文本分析、人居环境特征智能识别中的有效性,不仅拓展了大语言模型在城乡规划领域的应用范围与场景,也为地方康养产业发展、区域生态文明建设与空间优化提供科学依据。

1 研究区概况与研究方法

1.1 研究区概况

武夷山市位于福建省西北部,以自然与文化景观而闻名,该地人居环境优越,建筑风格独特,基础设施较为完善,反映了其优质的建成环境特征。城区布局整体较为合理,传统民居、历史街区、现代公共设施与生态空间相互协调,展现了功能完善、宜居舒适的城市空间特征。然而,目前武夷山市建

成环境特征的系统性探讨较为缺乏。因此,本文从建成环境视角出发,深入识别并分析当地优质的人居环境特征,以期为城乡规划、康养产业发展与生态文明建设提供科学依据与实践指导。

1.2 数据来源与处理

中国知网和 Web of Science 分别是国内外权威的学术数据库,涵盖众多高质量期刊资源。本文以这两个数据库为文献来源,对武夷山市相关研究进行系统筛选与分析。筛选条件如下:

中国知网(CNKI): 主题为“武夷”或“南平”; 学科分类为“建筑科学与工程”或“环境科学与资源利用”; 发表时间为 2015 年 1 月至 2025 年 6 月; 文献类型为学术期刊。

Web of Science (WoS): 主题为“Wuyi”或“Nanping”; 研究领域涵盖地理学、遥感、建筑学、城市研究、环境科学与生态学; 文献发表时间范围同中文文献一致; 文献类型为 Article。

鉴于武夷山市为南平市下辖的县级市,部分研究以“南平”或“Nanping”为主题进行展开,因此本文将二者一并纳入检索范围。在分析过程中,提取南平市相关文献中涉及武夷山市建成环境特征的内容。为确保研究具有代表性和覆盖面,同时避免样本量不足,时间范围设定为 2015—2025 年,文献类型限定为正式发表的学术期刊。最终共获取中文文献 403 篇,英文文献 533 篇。而后,对 936 篇中英文文献的主题和摘要进行人工识别,共筛选出 56 篇与武夷山建成环境相关文献,其中英文 7 篇,中文 49 篇。

参考相关研究^[25-26]并结合武夷山市建成环境特点,本文归纳并构建了建成环境的六个核心维度,依逻辑顺序排列为:自然生态与景观、空间结构与开发强度、居住与社区、功能与服务配套、社会文化与感知体验^④。

此外,本文基于社交媒体数据对分析结果进行交叉验证,选取微博与小红书作为代表性平台。其中,微博数据来源于 2015—2020 年间公开发布的文本信息(<https://weibo.com>, 2025-07-30),共收集 3 亿余条内容,将其进行关键词与地理位置筛选,定位到武夷山市文本信息,用于分析长期、多样化的用户感知;小红书数据来源于 2024 年 8 月平台热度榜与时间排序榜(<https://www.xiaohongshu.com>, 2025-07-30),共抓取笔记 2 689 条(时间排序)与 2 587 条(热度排序),经数据清洗与去重处理后最终获得 3 861 条有效文本,用于补充近年用户感知。

1.3 研究方法

1.3.1 特征关键词提取与情感识别

由于 ChatGPT 在深度语义理解、概念聚合、

类推推理等文本分析方面具有优势,进一步突出其在复杂文本分析中的应用潜力。同时,为保证特征关键词提取与情感打分的科学性与稳定性,本文结合 ChatGPT 内嵌的多种主流模型,对武夷山建成环境相关文献进行了自动化分析,具体方法如下:

(1) 多模型性能对比与选择

本文采用随机抽样方式,从全部文献中选取 10% (共 6 篇) 作为测试集。首先,由人工对测试集进行特征词提取和情感评分,以此作为基准。随后,分别利用 ChatGPT-4o (适用于大多数任务)、ChatGPT-o3 (擅长高级推理) 和 ChatGPT-4.1 (擅长快速分析) 三种模型,对同一测试集开展特征词提取与情感评分对比。通过将各模型输出结果与基准进行多轮比对,重点评估模型在特征词提取准确性与情感识别能力方面的表现,并最终遴选出表现最优的 ChatGPT-4.1 模型作为后续正式分析的主要工具。

(2) 特征关键词提取与情感评分流程

采用统一提示词,调用大语言模型 API 接口,对 56 篇武夷山人居环境相关文献全文进行建成环境特征词自动提取与情感评分。具体流程为:提取文献中所有相关的建成环境特征词,并直接输出为 Excel 表格;根据语境对上述特征词进行情感打分,采用 5 分制(5 分为极为优质,1 分为极差),并输出 Excel 表格。

(3) 模型稳定性评估

通过 API 接口,随机抽取 10% (6 篇) 文献全文使用 ChatGPT-4.1 模型进行分析,采用相同的提示词独立提问三次,获取多组特征词及情感识别结果。邀请五位领域专家(其中城乡规划三人、计算机两人),对模型输出的稳定性进行 5 分制评分。

(4) 模型理解能力评估

针对同一问题,分别设计三种不同的提问方式,均由 ChatGPT-4.1 模型进行回答,并邀请同样的五位专家,采用 5 分制对模型的理解能力进行评分。

(5) 专家打分说明

为确保评价的客观性,所有专家评分均采用独立盲评方式,评分涵盖模型稳定性与理解能力两个维度。评分标准为 5 分制:5 分表示极强,4 分为较为优质,3 分为中等,2 分为较差,1 分为极差。

1.3.2 重要性-性能分析

重要性-性能分析(Importance-Performance Analysis, IPA)是一种用于评估产品或服务各属性重要性与用户满意度之间关系的常用工具^[27-28]。在 IPA 矩阵中,通常将用户对各属性的重要性(如期望)作为横轴(X轴),满意度作为纵轴(Y轴),以各自的平均值为界进行象限划分。本文运用

IPA 方法，将特征词频率作为 X 轴的重要性指标，情感评分作为 Y 轴的表现指标，通过绘制 IPA 矩阵，分析武夷山建成环境中各维度特征的关注意度与实际表现。

1.3.3 社交媒体数据分析

本文利用 ChatGPT-4o 对处理后的社交媒体数据（微博与小红书）进行武夷山人居环境优秀特征提取，以此与大语言模型对文献分析结果进行比对。一方面验证大语言模型对文献分析结果的准确性，另一方面，由于两类文本在语言风格与信息结构方面存在显著差异，交叉比较二者的特征提取结果可进一步检验大语言模型在跨文本语义识别中的鲁棒性。

1.3.4 人工验证

本文通过人工识别方法，对 56 篇文献进行分析。具体流程为：按照第 1.2 节所述的六个建成环境维度，从每篇文献中遴选不超过五个主要相关特征词，并结合语境对每个特征词进行情感打分。以人工标注结果作为真值集，采用准确率（Precision）、召回率（Recall）、F1 分数（F1 Score）及 Jaccard 系数对大语言模型的特征词识别一致性进行评估。对于情感打分的一致性，采用 ICC（Intraclass Correlation Coefficient，组内相关系数）方法将人工情感打分与大语言模型情感打分进行一致性分析，该方法是衡量多个评分者（或方法）对同一组对象评价结果一致性的统计指标。其中准确率的计算采用相似度或匹配得分，将其转化为二分类结果，即设定阈值大于等于 50 视为匹配，低于该阈值视为不匹配。

2 基于大语言模型的武夷山人居环境特征提取及情感识别

2.1 模型表现能力

ChatGPT-4.1 模型在特征词提取与情感打分任务中，展现出较强的模型稳定性和理解能力（见表 1）。其中，专家对模型稳定性评分均为 4 分，说明模型输出可靠、波动性小。从模型理解能力

看，平均分为 3.8 分，反映模型在大部分情境下能较为准确地理解问题，但在个别复杂任务或特殊情境中存在提升空间。总体来看，ChatGPT-4.1 模型对建成环境相关特征词的提取效果较好，在大部分情境下较为准确地理解问题，该模型表现良好，在武夷山建成环境相关文献特征词提取与情感打分中具备较高的实用价值。

2.2 特征词提取

以文献为单位，根据特征词统计各维度在 56 篇文献中出现的次数，以判断各维度的重要性。每篇文献中只要出现某一维度的任一特征词，即记为 1 次，否则为 0 次，得到表 2。由表 2 可知，“功能与服务配套”“自然生态与景观”是武夷山建成环境中最为重要的维度，分别在 52 篇和 49 篇文献中被提及；其次为“社会文化与感知体验”“空间结构与开发强度”“交通与可达性”，出现频率均大于 40 次，居住与社区出现频率最少，仅为 24 篇。

对大语言模型提取的特征词进行可视化展示，得到图 1，其中特征词越大，代表词汇出现频率越高。首先，“功能与服务配套”是武夷山建成环境中最为重要的特征。从图 1a 来看，典型特征词以“博物馆”“度假区”等旅游空间为主，在多数文献中被提及，说明旅游功能与服务在武夷山建成环境中具有重要地位。此外，茶馆、图书馆等功能与服务空间也被提及，反映出武夷山建成环境在功能多样性与文化休闲空间方面的重视。其次，“自然生态与景观”同样是武夷山建成环境的核心特征，相关特征词有“生态环境”“地形地貌”“园林绿化”“屋顶花园”等（见图 1b）。该维度不仅体现了武夷山在资源保护与利用方面的积极实践，也反映了其对自然生态保护的高度重视，推动了城市可持续发展。从“社会文化与感知体验”来看，该特征重要性位居第三，涉及的特征词包括“传统建筑”“建筑风格”“情景雕塑”“仿宋古街”等（见图 1c），这直接反

表 2 建成环境各维度重要性

Tab.2 Importance of built environment dimensions

资料来源：作者自绘

建成环境维度	出现频率 / 次
功能与服务配套	52
自然生态与景观	49
社会文化与感知体验	48
空间结构与开发强度	46
交通与可达性	40
居住与社区	24

表 1 ChatGPT-4.1 模型表现专家打分表

Tab.1 Expert scoring table for ChatGPT-4.1 model performance

资料来源：作者自绘

模型表现 / 专家打分	稳定性	理解能力
专家一	4	4
专家二	4	4
专家三	4	3
专家四	4	4
专家五	4	4
平均分	4	3.8

映了传统建筑与独特的建筑风格是武夷山建成环境中的关键要素,受到学界的广泛关注。第四,“空间结构与开发强度”维度关注度也相对较高,特征词涵盖“空间布局”“空间形态”“土地利用”“因地制宜”等(见图 1d),说明空间布局与土地利用是学者关注焦点,并强调规划中因地制宜的重要性。第五,“交通与可达性”维度中,常见特征词有“交通便利”“交通网络”“主干道”“步行街”等(见图 1e),这直接反映了武夷山交通系统的关注度较高。最后,“居住与社区”维度出现频率最低,典型特征词包括“民居”“住宅建筑”“住宿空间”“临河民居”等(见图 1f),表明该维度虽然在文献中关注相对较少,但该维度作为建成环境基础功能的重要性不容忽视。

综上,“功能与服务配套”与“自然生态与景观”维度的重要性最为突出,其他维度则分别从文化、空间结构、交通及社区等方面,为武夷

山建成环境提供了多维支撑。

2.3 情感识别

为进一步明确武夷山优质人居环境的关键特征维度,本文基于 ChatGPT-4.1 对特征词的情感分值进行平均计算,得到表 3。由表 3 可知,“自然生态与景观”与“社会文化与感知体验”两个维度整体表现最为突出,是武夷山建成环境的核心优势;其次为“空间结构与开发强度”“功能与服务配套”“居住与社区”三个维度;而“交通与可达性”维度情感分值相对较低,但仍为积极情感。

具体来看,“自然生态与景观”“社会文化与感知体验”均在4.5分以上(见表3),说明武夷山在生态保护与社会文化方面具有较高认可度,是该地建成环境的重要优势。例如,关瑞明、吴钦豪^[29]从聚落形态、民居形制与建造材料等方面系统梳理了武夷山传统建筑的地域特征,并



图 1 各维度特征词云

Fig.1 Word cloud of feature words for each dimension

资料来源：作者自绘

高度评价其建筑形式的文化表达，认为风火山墙的精美装饰与图案不仅展现了居者的审美情趣，也蕴含了深厚的文化内涵。“空间结构与开发强度”“功能与服务配套”的情感评分均在 4 分以上，具有积极情感（见表 3）。这反映了武夷山空间布局、建筑密度以及旅游功能与服务等方面表现良好。张坤^[30]以武夷山茶博园为例，指出作为当地现代观光旅游的重要休闲与娱乐场所，该项目合理划分了多个功能空间，并依据景区实际需求实现了科学规划与设计。此外，尽管“居住与社区”维度在相关文献中关注度相对较低，但其情感评分仍表现为积极，说明该建成环境维度具有一定发展

表 3 建成环境各维度情感打分

Tab.3 Sentiment scoring for each built environment dimension
资料来源：作者自绘

建成环境维度	情感打分
功能与服务配套	4.152 0
自然生态与景观	4.554 3
社会文化与感知体验	4.542 1
空间结构与开发强度	4.167 4
交通与可达性	3.898 8
居住与社区	4.111 1

基础且表现良好。“交通与可达性”维度的情感评分相对较低（3.898 8），但仍处于积极情感区间。总体来看，武夷山建成环境在自然生态与景观、社会文化与感知体验方面优势突出，在交通与社区等维度仍有进一步优化和完善的空间。

2.4 重要性-性能分析 (IPA)

从武夷山建成环境六个维度内部来看，由 IPA 分析结果（见图 2）可知，不同维度表现出明显差异。首先，“自然生态与景观”与“社会文化与感知体验”两个维度位于高重要性-高表现的优势象限，分别具有较高的出现频率（49 次和 48 次）与最高的情感分值（4.554 3 和 4.542 1）。这表明武夷山地区在自然生态环境与文化感知体验方面表现突出，是当地人居环境的重要优势，应持续予以保持和优化。其次，“功能与服务配套”维度出现频率最高（52 次），但情感分值相对适中（4.152 0），而“空间结构与开发强度”维度也呈现类似趋势（出现频率 46 次，情感分值 4.167 4），两者均处于或接近高重要性-低表现的改进象限。这表明虽然这两个维度被学界的关注度较高，但实际表现并未达到学者的高标准，可进一步被适当优化与提升。再次，“交通与可达性”维度表现出明显的不足，虽然其出现频率相对较高（40 次），但情感分值最低（3.898 8），与以上四个建成环境特征相比，落在低表现区，

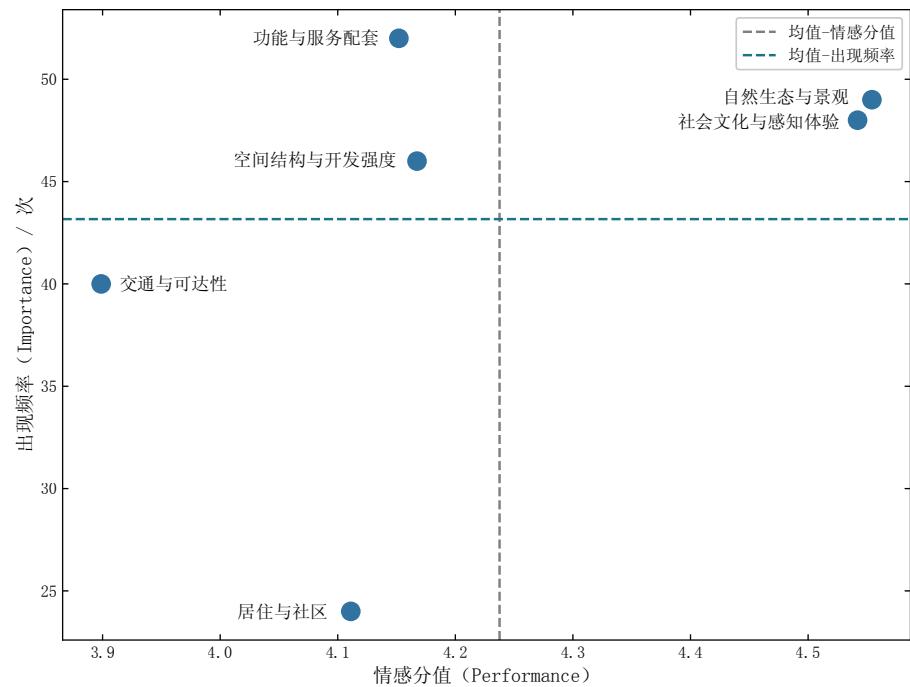


图 2 武夷山建成环境特征词维度重要性-性能 (IPA) 图

Fig.2 Importance-Performance (IPA) chart of built environment feature word dimensions in Wuyishan
资料来源：作者自绘

说明尽管部分文献指出武夷山内部交通较为便利，但仍有区域存在交通基础设施滞后的问题，整体可达性水平仍有待提升。最后，“居住与社区”维度无论在重要性（出现频率为 24 次）还是情感表现（评分为 4.111 1）方面，均处于相对次要位置，表明该维度的关注度较低，未来应予以适度关注与优化提升。综上所述，“自然生态与景观”“社会文化与感知体验”是武夷山建成环境突出的优质特征。

3 分析结果验证

3.1 社交媒体数据验证

3.1.1 维度一致性分析

基于社交媒体数据提取 10 类人居环境优质特征（包含绿化率高、交通便利、社区环境宜居等），将其按语义相似性对齐 2.2 节中的六个维度，并计算两类语料的维度频次与排序一致性。结果表明，社交媒体数据在“自然生态与景观”“交通与可达性”“功能与服务配套”“社会文化与感知体验”“居住与社区”5 个维度上与文献对应，维度一致性为 83.3%。进而采用 Spearman 相关性分析比较两类语料的维度重要性排序，得到相关系数约为 0.83，表明二者提取结果具有显著一致性。综上所述，大语言模型对文献分析结果具有较高准确性。

3.2.1 情感一致性分析

将基于社交媒体数据与文献两类语料得到情感分值进行相关分析，以检验两类文本在情感态度上的一致性。结果表明，Spearman 相关系数达到 0.90，表明公众在社交媒体中的情感表达与文献中的情感评价具有显著相似性。例如，“自然生态与景观”“社会文化与感知体验”在两类语料中均表现为情感最为积极的维度；“交通与可达性”在两类语料中情感值均最低。该结果进一步验证了大语言模型在测度文献情感方面的可靠性。

3.2 人工验证

3.2.1 特征词提取一致性

本文进一步以人工识别结果对第 2 部分分析结果进行验证，采用匹配得分进行相似性判定，将结果转化为二分类，设置相似度阈值大于等于 50 视为匹配，计算准确率、召回率及 F1 分数，得到表 4。结果表明，由于相似度阈值设定，大语言模型识别结果的准确率高达 100%，说明模型输出的特征词与人工标注结果较为一致，显示出大语言模型提取特征词较高的准确性与可靠性，具备较为稳定、可信的识别能力。然而，其召回率为 75.73%，表明仍有 24.27% 的人工标注特征词未被模型识别，存在一定的漏检问题。F1 分数达到 86.19%，表明模型整体一致性表现良好，在高精度的基础上兼顾了一定的覆盖能力。综上，大语言模型在当前任务中已展现出较强的特征词识别能力，特别是在精度方面表现突出。但在未来应用与研究中，仍需进一步提升召回率，以确保其在实际应用中的全面性和稳定性。

3.2.2 情感分析一致性

由表 5 可知，人工评分与大语言模型评分在固定评分者模型下（ICC3 和 ICC3k）具有高一致

表 4 人工识别与大语言模型特征词提取结果的一致性分析

Tab.4 Consistency analysis of feature word extraction results between manual recognition and large language model
资料来源：作者自绘

指标	数值 /%
准确率	100.00
召回率	75.73
F1 分数	86.19

表 5 人工识别与大语言模型情感打分的一致性分析

Tab.5 Consistency analysis of sentiment scoring between manual recognition and large language model
资料来源：作者自绘

类型	ICC 值	p 值	95% 置信区间	解释
ICC1	0.436	0.0004	[0.20, 0.63]	中等一致性，显著
ICC2	0.532	0.0000	[-0.09, 0.82]	较好一致性，但置信区间略宽
ICC3	0.806	0.0000	[0.69, 0.88]	高一致性，非常显著
ICC1k	0.607	0.0004	[0.33, 0.77]	较好一致性
ICC2k	0.694	0.0000	[-0.21, 0.90]	较好一致性，但置信区间略宽
ICC3k	0.893	0.0000	[0.82, 0.94]	极高一致性，非常显著

性。在单评分者及随机评分者模型下（如 ICC1, ICC2），一致性水平为中等到较高，且均具有统计学显著性（ $p < 0.001$ ）。整体而言，人工与机器评分具备良好的互换性和一致性。其中，ICC3 和 ICC3k 最高，表明在“人工与大语言模型评分直接对同一样本打分”应用场景下，机器分和人工分可以互换，自动评分系统表现良好。部分模型（ICC2 和 ICC2k）置信区间（Confidence Interval, CI）较宽，提示个别样本可能分歧较大。

4 结论与展望

随着大语言模型驱动研究不断深化，其在城乡规划领域的应用尚属起步阶段，尤其是在人居环境特征的系统性识别方面存在明显空白，缺乏基于多源数据的人居环境特征智能识别与分析。武夷山作为全国人居环境的典型优势区域，其建成环境视角下的优质特征尚未被充分揭示，因此，本文基于大语言模型开展武夷山人居环境特征的智能识别。通过对 ChatGPT 内嵌的三类主流模型进行对比实验，结合人工标注验证，最终采用 ChatGPT-4.1 模型对筛选后的 56 篇中外文献全文进行深度分析，系统提取了建成环境相关特征词并进行情感评分，从而识别出优质特征。进一步对模型识别结果进行双重比较，其中，社交媒体数据提取结果与其具有高度一致性。另外，还将模型分析结果与人工识别进行比对，双重对比表明大语言模型在人居环境特征识别与情感分析方面表现良好，具备在人居环境领域中的科研应用价值。自然生态景观和社会文化体验是武夷山建成环境的突出的优质维度，充分展示了武夷山人居环境以优质的生态本底和鲜明的文化特色为突出优势，功能与服务配套、交通可达性等应适当提升，在居住与社区维度，应提升研究与规划关注度，并加强该维度的优化与完善。未来应继续强化生态与文化保护，优化社区服务配套，推动人居环境高质量发展。

基于大语言模型的人居环境特征识别，有助于从文本感知层面深化对城市建成环境与人居环境特征的理解，为城市空间优化、地方产业发展和生态文明建设提供理论依据，为政府决策和政策制定提供科学支撑。建议相关部门持续推动武夷山生态环境和文化资源保护，补齐社区配套和功能服务短板，同时结合智能化交通与精细化开发，进一步提升居民生活质量与城市竞争力。本文聚焦于武夷山人居环境的现状特征提取，为后续区域人居环境演化及优质特征溯源研究提供了基础。本文仍存在一定局限性，首先，社交媒体等网络数据在使用人群、年龄结构与空间分布等方面存在覆盖偏差，所反映的人居环境感知特征

在一定程度上可能偏向特定群体，尚难以全面代表整体居民的真实感受。其次，受限于数据类型与获取方式，当前对人居环境特征的识别主要基于文本语义层面，对行为数据、空间实体特征及长期动态变化的刻画仍相对不足。未来研究可进一步拓展数据来源，引入调查问卷、传感数据及遥感信息等多源数据以应对社交媒体数据覆盖偏差，构建更加丰富的人居环境感知体系。同时，可通过对领域语料进行系统标注与预训练，结合微调后的大语言模型，开展更高精度、更具时序性的城市人居环境演变分析，从而更全面地支撑区域空间优化与人民美好生活需要的实现。

注释：

- ① http://www.ncha.gov.cn/art/2021/7/23/art_2539_170135.html，2025-03-01。
- ② <https://www.wyschina.com/cn/#/>，2025-03-01。
- ③ <https://baike.baidu.com/item/%E6%AD%A6%E5%A4%B7%E5%B1%B1%E5%B8%82/1660531>，2025-03-01。
- ④ 六个核心维度内涵界定。自然生态与景观：指建成环境中与自然生态系统、绿色空间、水体资源、植被景观等相关的要素，体现区域的生态质量、环境宜人性与可持续基础。空间结构与开发强度：关注空间格局、用地布局、建筑密度、体量、开发模式等反映物理空间利用与人类开发活动强度的特征，是建成环境的骨架和容量基础。居住与社区：反映居民居住环境、社区构成与邻里互动等与人居空间质量密切相关的内容，强调住宅空间、社区单元和居民日常生活的场所性。功能与服务配套：涵盖支撑居民和访客生活所需的各类功能性空间与公共服务设施，反映建成环境的服务能力和功能完善度。交通与可达性：指支撑人流、物流、信息流顺畅流动的道路系统、公共交通、步行系统等要素，体现城市空间的流动性与通达性。社会文化与感知体验：强调建成环境所承载的历史文化价值、社会认同、艺术表达以及居民或访客对空间的主观体验与情感归属，是空间人本价值和精神内涵的集中体现。

参考文献

- [1] 赵万民, 汪洋. 山地人居环境信息图谱的理论建构与学术意义 [J]. 城市规划, 2014, 38(4): 9-16.
- [2] 吴良镛, 毛其智. “数字城市”与人居环境建设 [J]. 城市规划, 2002, 26(1): 13-15.
- [3] 吴良镛. 人居环境科学导论 [M]. 北京: 中国建筑工业出版社, 2001: 38-48.
- [4] 孙蒙鸽, 王燕鹏, 付芸, 等. 面向短文本 - 多领域科技实体抽取的提示工程构建研究 [J/OL]. 数据分析与知识发现, 1-25[2025-06-01]. <https://link.cnki.net/urlid/10.1478.G2.20250508.1907.006>.
- [5] 张乐, 陈岩松, 张雷瀚. 大模型特征增强与多层次交叉融合的多模态情感分析方法 [J]. 数据分析与知识发现, 2025, 9(8): 47-58.

- [6] WU T, LI J, BAO J, et al. Large language model-driven multi-agent systems for improving production efficiency and reducing carbon emissions in manufacturing[J]. Computers & industrial engineering, 2025, 207: 111299.
- [7] 胡文玥, 谭维智. 让机器“可理解”: 大语言模型应用于语言学习的理解障碍及教育应对[J]. 现代教育技术, 2025, 35(6): 97-105.
- [8] 孙发勤, 王雯. 大语言模型驱动的 IPCRF 框架: 课堂问答数字化分析的创新与实践[J]. 远程教育杂志, 2025, 43(2): 103-112.
- [9] LIU W, HAO H, ZHOU A. An open-source large language model for Chinese education research[J]. Frontiers of digital education, 2025, 2(2): 23.
- [10] 李斌. 人工智能时代的语言研究: 走向基于大数据驱动的计算建模实证方法[J]. 外语学刊. 2025(5): 1-7.
- [11] 陆小飞, 金檀. 大语言模型微调技术在语言分析与测试中的应用与展望[J]. 现代外语, 2025, 48(3): 413-421.
- [12] 武光军, 丁泽. 大语言模型生成译文的语言特征探索: 基于平均依存距离及其与翻译质量间关系的考察[J]. 外语学刊, 2025(3): 48-55.
- [13] 陈静, 曹智勋. 对抗幻觉: 垂直领域中大语言模型的应用策略探讨: 以中医知识问答领域为例[J]. 数据分析与知识发现, 2025, 9(4): 1-13.
- [14] SU M J, LIANG X, TAO Z X, et al. Using a large language model for breast imaging reporting and data system classification and malignancy prediction to enhance breast ultrasound diagnosis: Retrospective study[J]. JMIR medical informatics, 2025, 13: e70924.
- [15] YU Z. Beyond hype: Separating large language model-driven health literacy from clinical decision support system diagnostic rigour[J]. Australian critical care, 2025, 38(5): 101265.
- [16] 金星晔, 左从江, 方明月, 等. 企业数字化转型的测度难题: 基于大语言模型的新方法与新发现[J]. 经济研究, 2024, 59(3): 34-53.
- [17] 马少辉, 王炜辰. 基于大语言模型的消费者社交媒体表征分析: 以淘宝退货服务为例[J]. 消费经济, 2024, 40(6): 85-98.
- [18] 陈澍, 韩清, 张伯超. 企业人工智能技术水平与劳动力需求结构变化: 基于大语言模型的新方法和新发现[J]. 技术经济, 2025, 44(1): 1-13.
- [19] 魏伟, 张坤, 徐哲淇. 从通用到垂直: 大模型赋能管理学研究的新路径[J]. 管理学报, 2025, 22(1): 1-11.
- [20] 高原, 李重阳, 曲博婷, 等. 基于大语言模型的 UGC 文本挖掘与城市旅游流网络结构研究[J]. 数据分析与知识发现, 2025, 9(4): 158-169.
- [21] 周丽君, 连昊, 刘思佳. ChatGPT-4.5 与 DeepSeek-V3 在膝关节关节炎康复指导中的表现比较[J]. 华西医学. 2025, 40(6): 870-875.
- [22] 赵志杰, 刘岩, 张艳荣, 等. 基于 Lasso-LDA 的酒店用户偏好模型[J]. 计算机应用与软件, 2021, 38(2): 19-26.
- [23] 林振宇, 解吉波, 杨腾飞, 等. 旅游多主题情感词典的构建方法[J]. 地理与地理信息科学, 2021, 37(4): 22-27, 98.
- [24] 邱泽国, 贺百艳. 基于 PCA-Spectral-LDA 的网络舆情聚类 and 情感演进分析: 一个微博文本挖掘研究[J]. 系统科学与数学, 2021, 41(10): 2906-2918.
- [25] LIU N, DENG Q J, HU P P, et al. Associations between urban exposome and recurrence risk among survivors of acute myocardial infarction in Beijing, China[J]. Environmental research, 2023, 238: 117267.
- [26] XIA J H, TONG Y, LONG Y. Advancements in the application of large language models in urban studies: A systematic review[J]. Cities, 2025, 165: 106142.
- [27] CHON K S, WEAVER P A, KIM C Y. Marketing your community: Image analysis in Norfolk[J]. Hotel & restaurant administration quarterly, 1991, 31(4): 31-37.
- [28] 王钦安, 彭建, 孙根年. 基于 IPA 法的传统型景区游客满意度评价: 以琅琊山景区为例[J]. 地域研究与开发, 2017, 36(4): 110-115.
- [29] 关瑞明, 吴钦豪. 武夷山地区建筑创作中的地域性表达: 以武夷山庄与九曲花街为例[J]. 华中建筑, 2015, 33(12): 153-156.
- [30] 张坤. 基于武夷茶博园的观光景观设计探讨[J]. 福建茶叶, 2018, 40(2): 131.