TU Delft

Delft University of Technology

Document Version
Final published version

Licence
Dutch Copyright Act (Article 25fa)

Citation (APA)
Zhang, C., Cheng, Y., Zhu, L., Mao, H., Du, Y., & Yu, X. (2026). Understanding Landscape Perception in Mountainous Cities: Mapping its Patterns and Influencing Factors. *Applied Spatial Analysis and Policy*, *19*(2), Article 119. https://doi.org/10.1007/s12061-026-09893-1

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Applied Spatial Analysis and Policy (2026) 19:119
https://doi.org/10.1007/s12061-026-09893-1

RESEARCH

Understanding Landscape Perception in Mountainous Cities: Mapping its Patterns and Influencing Factors

Changyu Zhang[1] · Yu Cheng[2,3] · Lingna Zhu[1] · Huasong Mao[1] · Yizhao Du[4,5] · Xinrui Yu[1]

Received: 15 December 2025 / Accepted: 12 May 2026
© The Author(s), under exclusive licence to Springer Nature B.V. 2026

Abstract

In mountainous cities, mountain landscapes are a crucial component of urban nature, offering rich perceptual layers and angles due to their complex topography, and serving as a vital embodiment of urban culture and place identity. Studying the visual perception patterns of these landscapes from a public perspective provides new dimensions for landscape value assessment. This study, taking Chongqing, China as a case study, aims to construct a multi-dimensional mountain perception analysis framework that integrates multi-source data and advanced technologies. We extracted spatial information and element proportions from geo-tagged images and used a Geographically Weighted Regression (GWR) model to quantitatively analyze the correlation between various factors and mountain proportionate density. The results show that public visual perception of mountain landscapes exhibits two spatial characteristics: a high frequency in mid-to-short distance riverside areas and a high perceptual intensity in high-elevation locations. Building Proportion (BP), Green View Index (GVI), and Sky View Factor (SVF) are highly correlated with mountain perception and demonstrate significant spatial heterogeneity. Furthermore, the Shannon–Wiener Index (SWI) shows a weak negative correlation with mountain perception. This study reveals the spatial patterns and influencing mechanisms of mountain landscape perception to better formulate urban planning and landscape renewal strategies, thereby fostering a harmonious relationship between humans and nature.

Keywords Visual perception · Mountain landscape · GEO-tagged photos · Spatial distribution analysis · Crowd-sourced data

Introduction

Urban natural landscapes are fundamental components of urban ecosystems that make significant contributions to urban sustainability (Ahern, 2013; Wu, 2010), residents' well-being (Gong et al. 2024; Ives et al. 2017; Matsuoka and Kaplan 2008;

Extended author information available on the last page of the article

Published online: 18 May 2026

Zhang et al. 2022a, b), and quality of life. However, the pervasive process of urbanization has led to the encroachment of urban natural spaces, which threatens the integrity of urban ecosystems and their ecological and aesthetic value (Molla, 2015; Parnell et al., 2013). This trend highlights a growing demand among urban residents and visitors for nature-based experiences (Xu et al., 2020). As the European Landscape Convention defines, a landscape is 'an area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors (Europe, 2000). Existing research indicates that approximately 80% of human sensory experiences in landscape perception come from vision (Jeon & Jo, 2020). Thus, it is essential to conduct in-depth research on the visual characteristics of public perception of urban natural landscapes. Early landscape perception research was limited by technology and often employed an objectivist paradigm, quantifying visual aesthetic quality by classifying and assigning values to a landscape's physical features (Lothian, 1999). Although some studies attempted to explore perception from a public perspective, they were restricted to small-scale surveys due to insufficient sample sizes (Arriaza et al., 2004), making it difficult to capture large-scale, spontaneous public perception behaviors. The advent of the Web 2.0 era provided an opportunity to address this challenge. Social media, as an interactive digital medium, offers an unprecedented data source for urban planning and perception research, characterized by its diverse, voluminous, and spontaneous nature (Evans-Cowley and Hollander 2010; Li et al. 2025a, b; Yang and Liu 2022). Crowdsourced images have been proven effective in reflecting public perceptions and preferences for urban natural landscapes (Bubalo et al., 2019; Jang & Kim, 2019). Furthermore, recent breakthroughs in the field of computer vision, particularly with Large Visual Models (LVMs), have enabled the automated semantic recognition and segmentation of image instances, making it possible to quantify public perception from image content (Minaee et al., 2021; Xie et al., 2025). Currently, research on the visual perception of urban landscapes mainly focuses on urban street greening (Wei et al., 2022), urban green spaces (Zhang et al., 2023), urban waterscapes (Luo et al., 2022, 2023), and urban parks (Gungor et al., 2018).

Mountain landscapes, as a crucial component of urban natural systems, have received increasing attention in studies of landscape aesthetics and environmental perception. Existing research indicates that scenic quality in mountain environments is strongly influenced by spatial configuration and topographic characteristics, with landscape diversity and morphological complexity identified as key drivers of visual preference, while high-elevation and open landscapes are generally perceived as more attractive and land-use changes such as urban expansion tend to degrade scenic value (Schirpke et al., 2013). Complementary studies in hilly contexts further show that people prefer natural elements, visually diverse compositions, and open spatial structures, while overly fragmented or enclosed environments weaken aesthetic perception (Long et al., 2023). Overall, mountain landscape preference emerges from the interaction between subjective perception and objective spatial attributes.

Their unique topographical features provide richer perceptual layers and viewing angles (Petrova et al., 2015), which gives rise to a more complex three-dimensional spatial perception pattern, where viewing distance and elevation jointly influence the distribution of perception. Concurrently, their cultural attributes make them a vital embodiment of urban culture and local identity (Bernbaum, 2022; Zhao et al., 2024). While a few cities, such as Hong Kong, have successfully managed their urban-mountain visual interfaces (Planning Department, 1992), a lack of conservation awareness in many other regions during development has left over 40% of mountain landscapes completely unprotected (Jacobs et al., 2021), leading to issues of low visibility and poor aesthetic quality. This makes research into the public's perception of mountain landscapes and its spatial distribution patterns particularly urgent. However, two main gaps persist in the current literature: First, existing research about mountain landscape perception is mostly concentrated in rural or natural reserves (Y. Zhang & He, 2021), with little exploration of mountain landscapes within high-density urban areas; second, although some studies have analyzed the influence of perceptual content (Guo et al., 2021; X. Zhang et al. 2022a, b), they generally overlook the differences in visual perception behavior in the spatial dimension.

Therefore, this study aims to fill these gaps by integrating multi-source data and advanced technologies to build a multi-dimensional visual perception analysis framework. Specifically, this research addresses two core questions: (1) What is the spatial distribution pattern of public's perception of mountain landscapes in specific cities? (2) What are the key factors influencing public mountain perception, and does their mechanism of action exhibit spatial heterogeneity? Besides, the findings are expected to provide quantitative and actionable methods for optimizing urban visual corridors and mountain landscape viewpoints.

Methodology

Figure 1 shows an integrated analytical framework to quantitatively analyze the visual perception characteristics and influencing mechanisms of urban mountain landscapes from a multi-source, multi-dimensional perspective. To achieve this, we first collected crowdsourced social media image data to capture spontaneous public perception behavior. This data, combined with image geotags and elevation information, forms a comprehensive perception database. Second, we apply computer vision techniques for semantic segmentation and quantification of image content, which allows us to extract multiple influencing factors related to perceptual content. Finally, we use spatial statistical analysis and a Geographically Weighted Regression (GWR) model to investigate the spatial relationships and heterogeneity between these factors and mountain perception, aiming to reveal underlying perceptual patterns and mechanisms. This chapter details the construction of this analytical framework, including the collection and filtering of crowdsourced data, spatial distribution analysis, and GWR model correlation analysis.

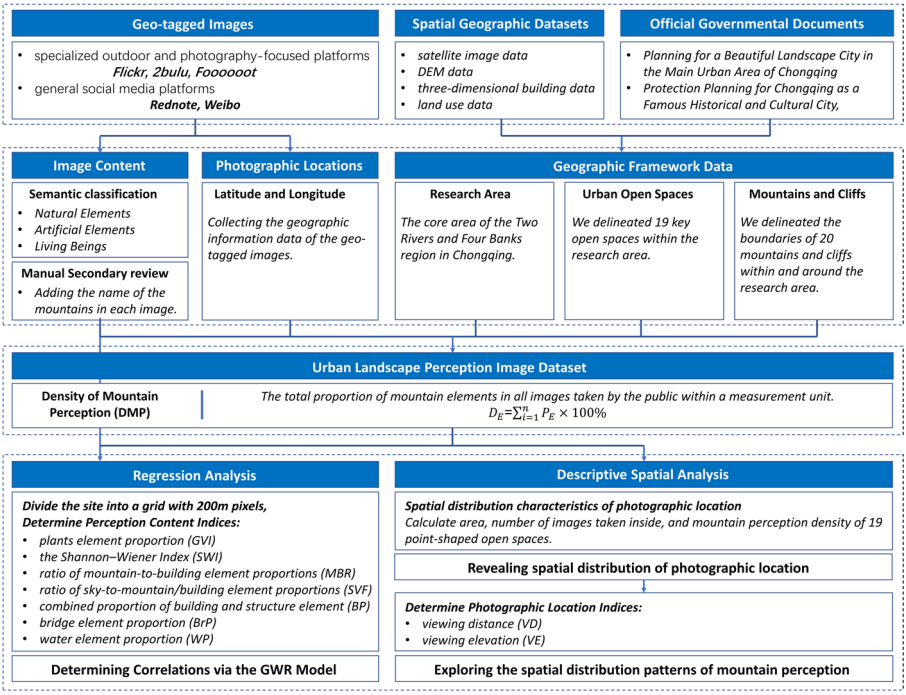


Fig. 1 Framework of the current study

Research Area and Study Objects

This research focuses on the core area of the Two Rivers and Four Banks in Chongqing, located in Southwest China. As a prototypical mountainous metropolis, Chongqing's central urban district encompasses over 40 individual mountains and cliffs, exhibits a relief of approximately 60 m, and features 38.4 km of scenic riverbank. The urban terrain is characterized by high mountain density, rich geomorphic diversity, and an intimate urban-mountain integration. Moreover, the mountain landscape of this city has undergone tremendous changes over the past several decades of urban development. This dynamic evolution has demonstrated the diversity of visual characteristics of urban mountain scenery, making it an ideal sample for exploring the laws of visual perception of urban mountains.

Figure 2 illustrates the research area and highlights the 20 mountains and cliffs boundaries visible from within and around the study zone. The delineation criteria for these mountain and cliff boundaries are based on three principles: (1) minimum area threshold, only land parcels exceeding 0.5 hectares may be classified as mountain landscapes, thereby excluding insignificant small, isolated landform units (Ge et al., 2022). (2) vegetation cover requirement, the mountain landscapes should maintain a green cover of no less than 30% to ensure their ecological and cultural significance in an urban environment (Koeser et al., 2023), following International Mountain Definition. Meeting international standards for mountain definitions in terms of slope, relative height difference, and other related criteria (Price et al., 2019).

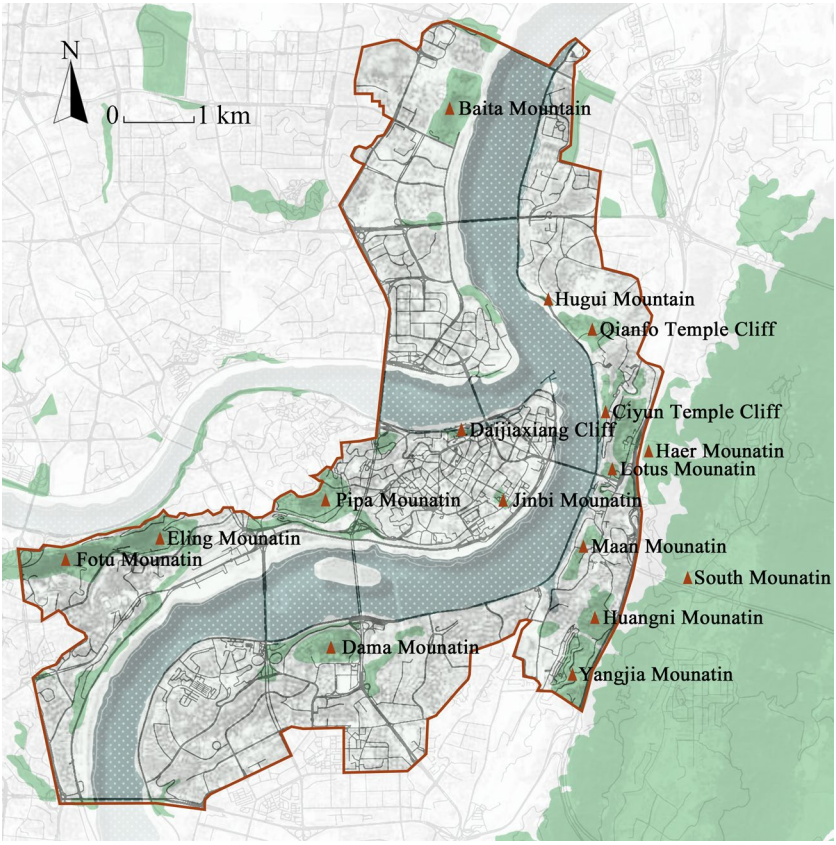


Fig. 2 Core area of the Two Rivers and Four Banks in Chongqing

Data Source and Processing

Data Acquisition

The data primarily originates from three sources: (1) geotagged images collected from online social media platforms; (2) satellite imagery and various spatial geographic datasets of the urban environment; and (3) public planning documentations released by official governmental agencies.

Among them, the geo-tagged images reflecting public perceptions of mountain landscapes. The spatial geographic data mainly include high-resolution satellite image data, DEM data (obtained from the ASTER GDEM v3, jointly released by the National Aeronautics and Space Administration [NASA] and the Ministry of Economy, Trade and Industry [METI] on August 5, 2019; available at: https://earthexplorer.usgs.gov/), three-dimensional building data, and land use data. The official government planning documents mainly consist of *Planning for a Beautiful Landscape City in the Main Urban Area of Chongqing*, *Protection Planning for Chongqing as a Famous Historical and Cultural City*, *Chongqing Territorial Spatial Master*

Plan (2021–2035)(https://ghzrzyj.cq.gov.cn/ztlm_186/ghgb/ztgh/202406), and *Comprehensive Planning for Governance and Enhancement of the "Two Rivers and Four Banks" in Chongqing's Main Urban Area*(https://zfcxjw.cq.gov.cn/zwxx_166/gsgg/202012).

Media Data Collecting and Filtering

Geo-tagged images were collected via crowdsourced data. Previous studies have demonstrated that compared to traditional data collection methods, crowdsourced data significantly enhance sample representativeness and reduce potential selection bias(Bubalo et al., 2019; Dunkel, 2015). The data sources are categorized into two groups: One is specialized outdoor and photography-focused platforms(Tieskens et al., 2018), including Flickr (https://www.flickr.com/), 2bulu (https://www.2bulu.com/), and Fooooooot (https://www.fooooooot.com/), which cater to enthusiasts and professionals. These platforms enable users to upload images with precise geographic metadata, allowing direct access to spatial coordinates. The other one is general social media platforms, including RedNote (https://www.xiaohongshu.com/) and Sina Weibo (https://weibo.com/) (Qi et al., 2023; Wang et al., 2025), which boast broad user bases. On these platforms, users can add location information through textual descriptions or manual map-based tagging, increasing coverage across diverse user demographics (Niu et al., 2025).

During the image data collection phase, this study utilized web-scraping tools to aggregate all image posts within the research area published between January 2019 and January 2025. The retrieval process employed keywords such as 'Two Rivers and Four Banks', riverside road names like 'Nanbin Road', 'Beibin Road', and administrative district names including 'Yuzhong District', 'Nan'an District'. Additionally, on platforms like RedNote and Weibo, broader keywords such as 'Chongqing Riverside', 'Yangtze River', and 'Jialing River' were added to enhance data coverage.

The scraped data elements encompassed three main dimensions: user ID, image content, and photographic location, all of which were imported into a unified database. To ensure research relevance, we conducted a detailed filtering process to retain only images portraying urban landscapes. Images were excluded if the content was unclear (e.g., overexposed, underexposed, or strongly stylized), thematically irrelevant (e.g., focusing on people, animals, food, or signage), of insufficient scale, or taken indoors.

To further reduce sampling bias, for highly similar images released by the same user within a short period, only the clearest and most representative one was reserved. After standardized processing of image format and resolution, a total of 1,560 high-quality urban landscape images were initially obtained.

Furthermore, spatial geographic information was extracted from qualified images and summarized into CSV files to support subsequent spatial statistical analysis. Considering that platforms such as Flickr and 2bulu allow users to upload images independently, not all pictures contain complete GPS positioning data. At this stage, a total of 230 images lacking valid geographic coordinates were further excluded. Ultimately, a formal research dataset consisting of 1,330 effective geo-tagged images was established for subsequent analysis.

Image Content Processing

This study employs image semantic segmentation to identify the content of geo-tagged images of urban mountain landscapes. Semantic segmentation utilizes deep learning on annotated datasets to automate batch processing of image instances. We utilize MMSegmentation, an open-source semantic segmentation toolbox built on PyTorch and maintained as part of the OpenMMLab project, which supports a wide array of segmentation architectures and tools (Contributors, 2023).

For quantitative analysis of image compositional components, we adopt the SegFormer model—a novel framework introduced in 2021 that combines a hierarchical Transformer encoder with a lightweight multilayer perceptron (MLP) decoder (Xie et al., 2021). SegFormer has demonstrated state-of-the-art performance across diverse tasks, including autonomous driving, medical imaging, and remote sensing.

Model training is conducted using the ADE20K dataset, released by MIT in 2016, which comprises 25,000 annotated images spanning 150 semantic categories (Zhou et al., 2017, 2019). This dataset offers extensive diversity and scene complexity, thereby enhancing the model's capability to recognize and segment urban landscape features. The examples of the outputs are shown in Fig. 3.

In Table 1, to tailor the segmentation outputs to urban landscape perception, we reclassify the semantic categories into three primary groups: (1) Natural elements—including mountain, vegetation, water, land, and sky; (2) Artificial elements—such as buildings, infrastructure, roads, transport features, and bridges; and (3) Living beings, currently limited to humans.

Trained on the ADE20K dataset, the SegFormer model often misclassifies vegetation-covered mountain areas as pure vegetation, leading to biased mountain proportion statistics. To eliminate such systematic errors, we performed manual post-correction via image processing software. For scenes with clear boundaries between foreground trees and background mountains, pixel area ratios were used to distinguish vegetation from mountain bodies. For highly blended vegetation–mountain landscapes that cannot be accurately segmented automatically, we manually delineated object boundar-

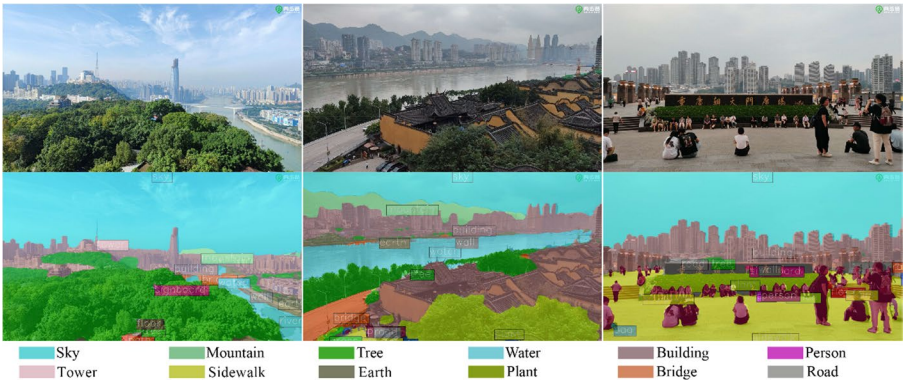


Fig. 3 Examples of the outputs of the image semantic segmentation

Table 1 Semantic classification system for urban landscape perception

Node number	Type	Detailed Elements
Natural Elements	Mountain	* mountain, hill
	Plants	* tree, palm
	Water	* water, sea, river, fountain
	Land	* earth, land, rock, grass, flower
	Sky	* sky
Artificial Elements	Building	* building, skyscraper
	Structure	* wall, fence, railing, streetlight, traffic light, sculpture, flag, signboard
	Road	* road, sidewalk, path, runway, dirt track
	Transport	* bus, car, track, airplane, ship, boat, pier
	Bridge	* bridge
Living Beings	Person	* person

ies and corrected misclassified pixels to refine mountain extraction. This workflow effectively improves the quantitative accuracy of mountain elements.

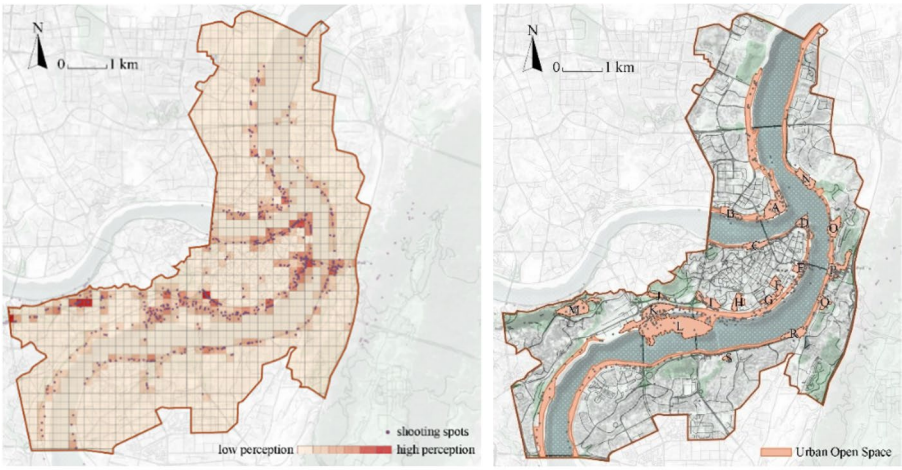
All segmentation results were further manually reviewed to reduce classification confusion over vegetated mountain surfaces. The final curated dataset records image IDs, element proportions, mountain names and spatial information for subsequent quantitative analysis.

Spatial Data Processing

To investigate the spatial distribution patterns of public perception of urban landscapes, we first imported a comprehensive dataset into ArcGIS. This dataset included image numbers, geolocated coordinates of photographic locations, and the proportion of various visual elements within each photograph. We then standardized all location data to the WGS84 coordinate system and overlaid it onto satellite imagery of Chongqing's Two Rivers and Four Banks area.

We employed two complementary approaches to assess the spatial patterns and intensity of public perception of mountain landscapes. The first approach utilized a spatial grid analysis to identify hot and cold spots of public perception. We divided the study area into 200 m × 200 m grid cells (Luo et al., 2021), with the perception density of each cell defined by the number of photographs taken by the public within its boundaries (Fig. 4a). This grid size was chosen to balance the need for clear differentiation between cells while ensuring a consistent perceptual context within each one, thus mitigating potential positional errors associated with photographic locations.

The second approach evaluated perception density at the level of specific open spaces to better integrate our findings with local planning realities. Based on field surveys and planning documents such as the Chongqing Historic and Cultural City Protection Plan, we delineated 19 key open spaces within the study area (Fig. 4b). The selection criteria for these spaces included: (1) high visual accessibility to urban landscapes, (2) an area no less than 10,000 square meters, and (3) a consistent perceptual context. This open-space-based analysis allows for a more refined understanding of public perception, providing more context-specific guidance for urban planning.



(a). Perception density of each cell. **(b).** 19 designated Urban open spaces.

Fig. 4 **(a).** Perception density of each cell. **(b).** 19 designated Urban open spaces

By integrating the spatial locations of photographs with the calculated proportions of various visual elements within each image, we can analyze the spatial distribution of public perception of mountain landscapes and derive the perception density sequences across different open spaces.

Quantification of Perceptual Density

This study explores the influencing factors of public mountain landscape perception from two perspectives: photographic location and perceptual content. We first quantified the dependent variable of this study, the Total Density of Mountain Element Proportion (DMP). This index is defined as the total proportion of mountain elements in all images taken by the public within a single analysis unit (e.g., a 200 m grid cell or a specific open space). This index is capable of reflecting the overall perceived intensity of the mountain landscape in a specific area.

It should be emphasized that DMP is an objective quantitative indicator that merely reflects the visual coverage and spatial occupation of mountain elements. A higher mountain proportion does not simply equate to better landscape aesthetic quality or stronger public visual preference. Even when mountains serve as auxiliary background elements rather than dominant landscape features, they can still exert a crucial influence on overall visual experience and landscape cognition. Therefore, this study strictly distinguishes objective mountain perception density from subjective visual preference to avoid conceptual confusion.

The calculation formula is as Eq. (1):

$$DMP_i = \sum_{i=1}^{n} P_{m_i} \times 100\% \tag{1}$$

Where DMP_i is the mountain proportionate density within the i-th analytical unit; n is the total number of images taken within that analysis unit; and P_{M_i} is the proportion of mountain elements in the i-th image.

In addition, if the basic units of analysis differ in area (such as open spaces), this study further explores the Average Density of Mountain Element proportion (AMP) metric on the basis of DMP, to measure the perception density sequence across different open spaces. The calculation formula is as Eq. (2):

$$AMP = \frac{1}{k}\sum_{i=1}^{k} DMP_i \tag{2}$$

Where AMP is the average mountain proportionate density; n is the total number of point-like open spaces included in the analysis; and DMP_i is the perception density of the i-th point-like open space.

Table 2 presents the detailed descriptions of the predictor variables of public perception of mountain landscapes. Drawing on relevant literature, this study identifies potential independent variables that may influence mountain landscape perception, classifying them into viewpoint distribution factors and viewpoint photographic feature factors. The subsequent analysis explores the correlation between these factors and the dependent variable, DMP.

Research Methodology

Descriptive Analysis of Spatial Indicators

The spatial distribution of mountain viewpoint perception is primarily determined by two factors: viewing distance and viewing elevation. Viewing distance (VD) is defined as the Euclidean distance from a photographic location to the boundary of the photographed mountain area, while viewing elevation (VE) corresponds to the ground elevation data of the photographic location.

Table 2 Predictive variable influencing public perception of mountainous landscapes

factors	variables	description	reference
Viewpoint Distribution Factors	VD	distance from the viewpoint to the mountain	(Schirpke et al., 2013; Tenerelli et al., 2017)
	VE	elevation of viewpoints	(Misthos et al., 2019)
Viewpoint Photographic Feature Factors	SWI	the Shannon–Wiener Index	(Y. Li et al. 2025a, b)
	GVI	plant element proportion	(Zhou et al., 2023)
	MCR	ratio of mountain-to-building element proportions	(Zhou et al., 2023)
	SVF	ratio of sky-to-mountain/building element proportion	(Song et al., 2025)
	BP	combined proportion of building and structure elements	(Zhou et al., 2023)
	BrP	bridge element proportion	(Zhou et al., 2023)
	WP	water element proportion	(Zhou et al., 2023)

To analyze the distribution characteristics of public mountain viewing distance, we categorized the distances into a series of equal-interval bins. Previous studies have shown that a viewing distance-to-vertical-height ratio (d/h) of approximately 3:1 is often ideal for a clear, full view of a scene and its background, while a ratio of 4:1 or more makes the object more of a background element (Korkmaz & Kim, 2022; Ladd, 1990). Considering Chongqing's unique geographical features, where the average height difference between mountains and the Yangtze River is approximately 400 m, we chose 400 m as the equal interval to explore how public perception changes at this specific geographic threshold. We calculated the total number of photographs (n) and the aggregated DMP within each distance bin and plotted their distributions. In addition, due to significant variations in perceptual distances for different mountains, we conducted a parallel comparative analysis of the viewing distance distribution for a selection of mountains.

Similarly, we analyzed the ground elevation data of all photographic locations. The data showed that most public photographs were taken from elevations ranging between 160 m and 400 m. In order to provide a detailed analysis of viewing behavior at different altitudes while ensuring a sufficient sample size in each bin, we divided the viewing elevations into bins with an equal interval of 20 m. We statistically analyzed the number of public perceptions and the aggregated DMP within each elevation bin and plotted their distributions to reveal the patterns of perception with respect to elevation.

Analysis of Influencing Factors at the Perceptual Content Level

This study identified potential influencing factors of DMP by combining relevant literature and the visual characteristics of urban landscapes (Table 2). These indicators, which represent the landscape's green cover, urban form, visual openness, and artificial structures, include Green View Index (GVI), Mountain-to-City Ratio (MCR), Sky View Factor (SVF), Building Proportion (BP), Bridge Proportion (BrP), Water Proportion (WP), and the Shannon–Wiener Index (SWI).

Before conducting the regression analysis, we used the Variance Inflation Factor (VIF) and tolerance to test for multicollinearity among all independent variables, ensuring they would not significantly affect the stability of the regression results. All variables had a VIF value of less than 10 and a tolerance value greater than 0.1, indicating no serious multicollinearity and confirming their suitability for inclusion in the regression model.

To establish a baseline for the global relationship between the variables, we first performed an Ordinary Least Squares (OLS) regression analysis. The OLS model can be expressed as Eq. (3):

$$y_i = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ki} + \varepsilon_i \tag{3}$$

Where y_i denotes the value of the dependent variable, the mountain landscape perception density (DMP) for the i-th analytical unit;

x_{ki} represents the value of the k-th independent variable (viewpoint distribution factor or photographic content factor) for the i-th analytical unit;

β_0 is the intercept term of the regression model;

β_k is the regression coefficient of the k-th independent variable;

p is the total number of independent variables included in the analysis;

ϵ_i denotes the random error term.

The OLS results, summarized in Table 3, provided an initial understanding of the average relationships between the perceptual content variables and DMP. However, we then performed a spatial autocorrelation test (Moran's I) on the OLS model's residuals. The test result (Moran's I index = 0.0358, z-score = 2.244039, p-value = 0.024830) indicated significant positive spatial autocorrelation in the residuals. This finding confirms the presence of spatial non-stationarity in the data, which the global OLS model is unable to account for.

Given this spatial heterogeneity, we adopted the Geographically Weighted Regression (GWR) model, which is capable of handling spatial non-stationarity, to identify and explore the localized associations between DMP and the independent variables (Brunsdon et al., 1996; Chang Chien et al., 2020). GWR works by computing a series of local linear regressions to reveal how the influence of the same factors varies in different regions. Its formula can be expressed as Eq. (4):

$$y_i = \beta_0 (u_i,\ v_i) + \sum_{k=1}^{p} \beta_k (u_i, v_i)\, x_{ki} + \varepsilon_i \tag{4}$$

where y_i represents the dependent variable value at the i-th observation point, x_{ki} represents the k-th independent variable value of the i-th sample, $\beta_k ?u_i,\ v_i ?$represents the local regression coefficient at location $?u_i,\ v_i ?$, and ϵ_i represents the error term.

The GWR model was constructed in ArcGIS with an adaptive Gaussian kernel function, and the optimal bandwidth was determined by the AICc minimization criterion. The final GWR model yielded an optimal bandwidth of 1172.09, an adjusted R^2 of 0.908, and an AICc of 9351.07. Compared with the global OLS model (adjusted R^2 = 0.603), the GWR exhibited greatly improved explanatory power, explaining 90.8%

Table 3 Summary of OLS results

Variable	Std Error	t-Statistic	Probability	Robust_SE	Robust_t	Robust_Pr	VIF	Tolerance
Intercept	7.53587	-2.33245	0.020431	5.197632	-3.38173	0.000844	–	–
GVI	0.167222	1.89605	0.059066	0.198302	1.598888	0.111076	1.103833	0.90593
SVF	0.704703	0.936655	0.349797	0.60446	1.091989	0.275847	1.183961	0.84462
SWI	4.150519	1.374775	0.170399	3.932326	1.451058	0.147986	1.263035	0.79174
BP	0.026737	13.7466	0	0.11188	3.285144	0.001172	2.053816	0.48690
MCR	0.913175	1.734032	0.084109	1.258279	1.258445	0.209368	1.020682	0.97974
BrP	0.078319	3.472899	0.000615	0.287794	0.945099	0.345478	1.543025	0.64808
WP	0.045438	-4.59727	0.000008	0.108409	0.345478	0.055086	1.638384	0.61036

of the total variance of DMP. Considering the significant residual spatial autocorrelation and spatial non-stationarity of the OLS model, the GWR effectively reduced spatial dependence via local regression and achieved better model fitness and applicability. The reliable GWR diagnostic indicators supported the rationality of local regression results. The local regression coefficients of each variable were further calculated to quantitatively reveal the spatial heterogeneity of driving effects on DMP.

Result

Semantic Segmentation Results

In Fig. 5, we performed content recognition and semantic segmentation on the 1,330 urban landscape images that were collected and filtered. Figure 5A presents the average proportion of each visual element. The results show that during the perception of urban landscapes, the sky element has the highest average proportion at 36.36%. This is followed by buildings at 16.38%, vegetation at 13.69%, water at 10.76%, and mountains at 6.63%.

We defined the presence of an element in an image as having a proportion greater than 0.1%. Based on this rule, approximately 72.63% of the images, totaling 975, contained the mountain element.

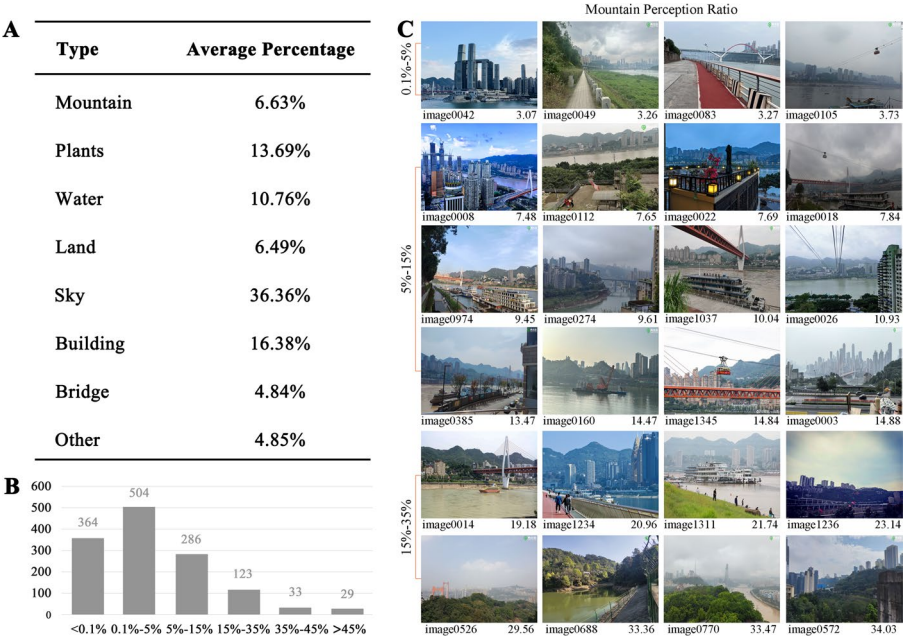


Fig. 5 Semantic segmentation results and visual element analysis. **(A)** average percentage of visual element in all images; **(B)** distribution of mountain element proportion across different ranges; **(C)** typical image examples within each proportion range

Figure 5B further details the distribution of mountain element proportion across different ranges, revealing the quantity of images that fall within each category. Concurrently, the examples shown in Fig. 5C illustrate typical scenes for each of these ranges. This reflects how the public's perception scale changes with the proportion of mountain elements. Specifically, when the mountain element proportion exceeds 30%, the perception scale tends to be smaller, with the mountain serving as a background and only a partial view of it being present. In contrast, when the mountain element proportion is between 0.1% and 5%, the mountain often acts as a complementary element to the perceived scene.

Spatial Distribution Characteristics

Overall Spatial Distribution

Through the visualization of public photography locations (Fig. 6a), we found that the hotspots of public urban landscape perception are primarily concentrated in the riversides, historical districts, and at elevated mountain points within the study area.

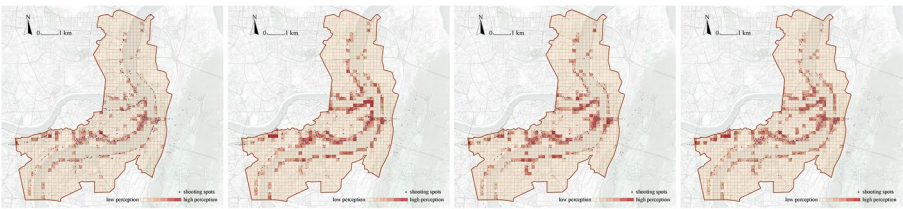
To further quantify this perceptual pattern, we calculated the DMP based on 200 m × 200 m grid cells. The resulting visualization (Fig. 6b) reveals the spatial hotspots of mountain perception in the city.

Furthermore, a comparative visualization of the perception densities for different landscape elements (mountains, water, bridges, and buildings) (Fig. 6a-d) revealed their unique spatial distribution characteristics:

- The public's perception of water elements is relatively high and is concentrated along the banks of the Yangtze and Jialing Rivers.
- The perception of mountains is generally weaker than that of other elements.
- The hotspots for bridge perception are clustered around major Yangtze River bridges that are representative city landmarks.

Perception Density Sequence

Based on the calculation of perception density for the 19 key point-like open spaces, we sorted the results by their DMP and AMP (Table 4). The results indicate that the DMP and AMP are both relatively high in areas such as Node Space E, I, J, K, L, and



(a)Mountain Perception Density (b)Water Perception Density (c)Bridge Perception Density (d)Building Perception Density

Fig. 6 Spatial Distribution Map of Visual Perception Density

Table 4 Perception Density Sequence for the 19 Key Open Spaces

Node space	Area (10,000 m²)	NMVV	DMP	AMP
I	11.8	50	43.56	10.28
J	4.1	26	41.66	6.57
K	2.6	25	36.92	3.84
E	7.9	47	33.73	5.67
M	22.5	82	26.97	7.4
D	10.5	43	17.40	4.25
C	8.4	25	12.56	4.22
P	13.7	29	5.19	2.45
B	7.3	12	4.70	2.86
A	22.1	43	4.26	2.19
O	10.7	6	2.90	5.17
S	1.6	2	1.98	1.58
L	72	32	1.86	4.19
G	5.1	3	1.64	2.78
Q	6.8	3	1.43	3.25
R	9	6	0.75	1.12
H	8.3	6	0.61	0.85
N	11.5	3	0.12	0.45
F	7.6	1	unconsidered	unconsidered

NMVV Number of Mountain View Visits, *DMP* Total Density of Mountain Element Proportion, *AMP* Average Density of Mountain Element Proportion

M. These areas correspond to public spaces such as Huguang Guild Hall, Coral Dam Park, Eling Park, and Fotuguan Park, which are considered important public activity hotspots.

Analysis of Perceptual Viewpoint Characteristics

Viewing Distance (VD) Analysis

Figure 7 presents the curve plots of the AMP and the number of photos taken (n) within each viewing distance bin. From the figure, it can be seen that the public's mountain viewing frequency (n) is mainly concentrated in the range of 800–2800 m, with the highest peak appearing at a distance of 1200–1600 m. However, the AMP generally shows a decreasing trend as the viewing distance (VD) increases. This indicates that although the public most frequently takes photos at medium distances, their highest perceptual intensity of mountains occurs at a closer range.

The study also found that there are significant differences in the perceptual viewing distance for different mountains. Therefore, we further analyzed the distribution of public viewing distances and the frequency of photography for different mountains. Figure 8 presents the boxplot of viewing distance and frequency for different mountains. The analysis reveals that South Mountain is the mountain with the highest public perception frequency, and it also has a wider distribution of viewing distances. This may be attributed to its highest altitude and largest area. Other mountains with

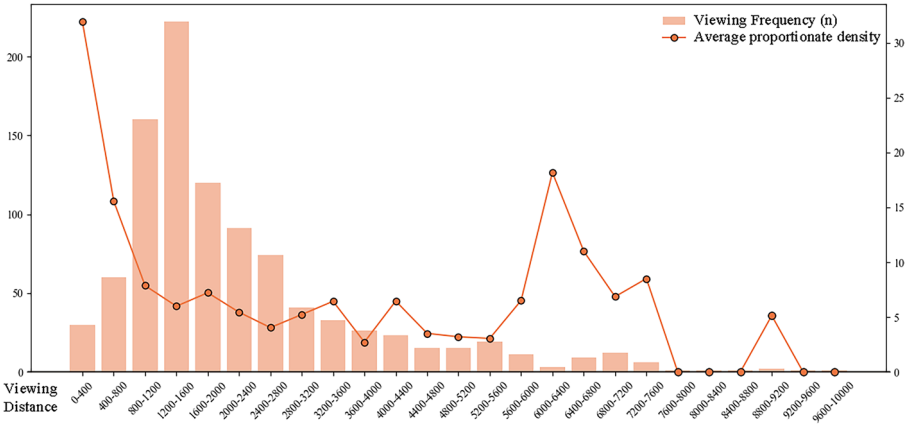


Fig. 7 Visual Distance from Mountain

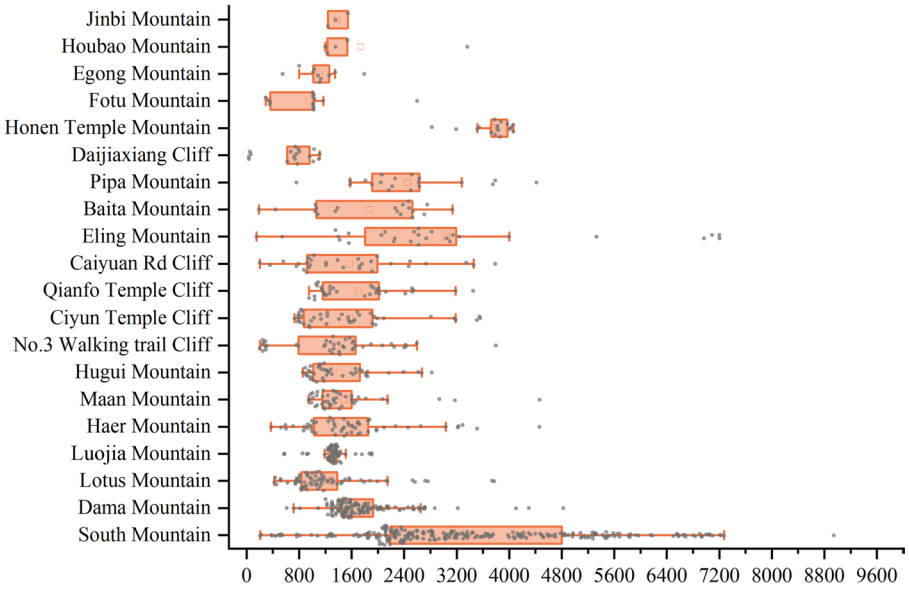


Fig. 8 Boxplot of VD per Mountain

relatively high perception frequency, such as Dama Mountain and Lotus Mountain, have viewing distances mostly concentrated in the 800–2000 m range, which is consistent with the overall mountain perception frequency distribution data. The viewing distance distribution for Luojia Mountain, however, is more concentrated, possibly due to the existence of specific photo spots or viewing platforms that make the perception angle more focused.

Viewing Elevation (VE) Analysis

Figure 9 presents the curve plots of the DMP and perception frequency (n) within each elevation bin. From the figure, it can be observed that the perception frequency is highest in the 175–195 m elevation range, followed by the 155–175 m range. This suggests that the riverside areas are where the public most frequently perceives the urban landscape. However, the AMP shows a trend of gradually increasing with the increase in viewing elevation, showing a significant increase after the viewing elevation exceeds 375 m. This indicates that while low-elevation areas are the most frequented, higher-elevation areas can also increase the frequency of mountain photos taken by the public.

Summary: The Influence of Viewpoint Characteristics on Mountain Perception

This section's analysis shows that public perception of urban mountains is not only influenced by the content elements but is also closely related to the physical characteristics of the viewpoint. Although the public most frequently takes photos in low-elevation, close-range areas (e.g., riverside areas), the viewing behavior from high-elevation viewpoints (e.g., Eling, Fotuguan Parks) and at long distances (e.g., South Mountain) brings a higher perceptual intensity. These findings provide a crucial foundation for understanding the local heterogeneous effects of variables like SVF and GVI in the GWR model.

Influencing Factors at the Perceptual Content Level

GWR Regression Analysis

Figure 10, which includes a series of maps, displays the spatial heterogeneity of the associations between each explanatory variable and the dependent variable, DMP.

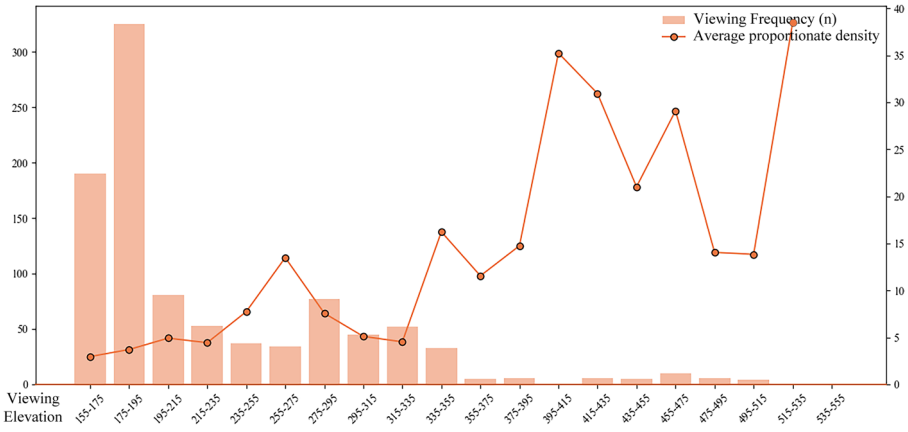


Fig. 9 Elevation of Viewpoints

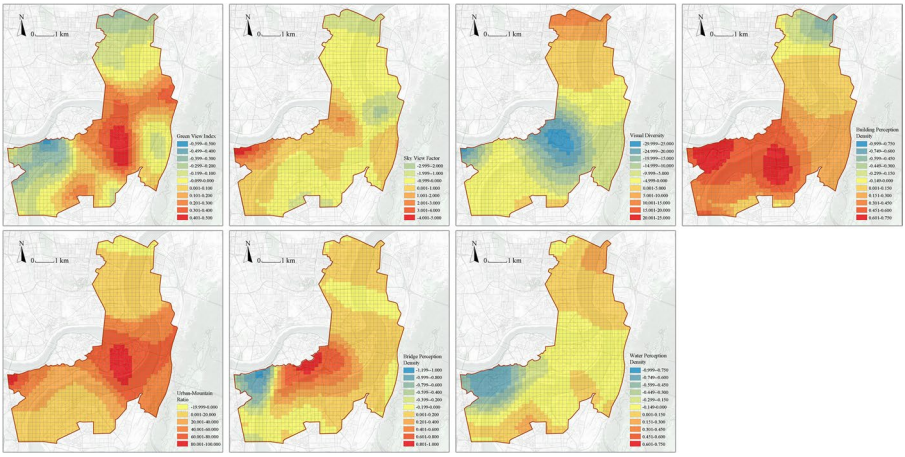


Fig. 10 Spatial Distribution of GWR Regression Coefficients

The following is a detailed interpretation of the spatial distribution of each element's influence:

(1) GVI and mountain perception. The relationship between GVI and mountain perception shows significant spatial heterogeneity. GVI exhibits a positive correlation in areas such as the Yangtze River riverside and riverside parks, but a negative correlation in high-altitude mountain viewpoints like Eling and Fotuguan parks.
(2) WP, SVF, BrP, and Mountain Perception. While WP, SVF, and BrP were not globally significant in the overall model, they express strong local correlations in specific areas. At high-altitude viewing platforms like Eling and Fotuguan parks, WP shows a significant negative correlation, while SVF shows a strong positive correlation.
(3) BP, MCR, and Mountain Perception. Both BP and MCR show a significant positive correlation in both the global and local models.
(4) SWI and Mountain Perception. The association between SWI and DMP is not significant in either the global or local models. However, a negative correlation is observed in certain areas, such as along the Yangtze River riverside and in riverside parks.

The maps from top to bottom and left to right represent the spatial distribution of regression coefficients for GVI, SVF, SWI, BP, MCR, BrP, and WP, respectively.

Discussion

Innovation of the Integrated Visual Perception Analysis Framework

This study integrates multi-source data and leverages cutting-edge digital technologies to construct a more comprehensive and multi-dimensional visual perception

analysis framework. On one hand, the method based on social media image data can capture the spontaneous perception behaviors and visual foci of the public on a large scale. This, to a certain extent, objectively reflects the public's aesthetic preferences for mountain landscapes and overcomes the issues of insufficient sample size and narrow demographic coverage in previous studies. On the other hand, the collection and analysis of spatial geographic information from images tightly links visual perception content with urban space. This allows the research to move beyond a holistic analysis of different influencing factors and more accurately identify the spatial heterogeneity exhibited by these factors, providing more localized and refined strategies for landscape planning and management.

Comprehensive Interpretation and Implications of Key Findings

In this study, we revealed the characteristics of public mountain perception from two dimensions—perception content and spatial distribution—and identified their influencing factors. Our analysis uncovered a seemingly contradictory yet profoundly insightful phenomenon: the public's average perception density of mountains increases as the viewing distance gets closer, but at the same time, the AMP also significantly increases with the elevation of the viewing point. This phenomenon suggests that the perception patterns of the public differ in focus across various spatial viewpoints. In low-elevation, mid-range areas (500–2000 m) (e.g., the riverside) (Lee et al., 2019), which are the most frequent places for public daily activities, mountains often serve as a backdrop for daily life. In this case, the public's viewing distance directly influences the proportion of the mountain landscape in the images. In contrast, high-elevation, long-range areas (e.g., Eling Park) offer a more open panoramic view. The public often visits these places with a specific purpose to perceive the overall urban-mountain form, which consequently leads to a higher intensity of perception.

This difference in viewpoint spatial characteristics provides a crucial clue for understanding the spatial heterogeneity of influencing factors. We found that the BP has the most significant positive association with mountain perception among the content indicators, which aligns with previous findings that artificial elements are often more attractive than natural ones (Zhou et al., 2023) and have a significant impact on landscape visual quality (Guo et al., 2021). However, a difference from previous studies is that the increase in the proportion of building elements in image scenes may enhance the public's perception of mountain elements. This might be related to Chongqing's unique urban form. The GWR model results show that this positive association is particularly significant in areas like the Yangtze River riverside and Eling Park, which also proves that urban constructions and mountain landscapes are not mutually exclusive but can form a harmonious and symbiotic relationship. In practical planning, such findings indicate that moderate and controlled building development along riverfront areas and key mountain-view corridors can help enrich urban skyline layers and jointly shape a distinctive city-mountain landscape interface, rather than simply restricting urban construction.

Furthermore, the GWR model revealed the dynamic spatial changes in the influence of elements such as GVI and SVF. Our analysis shows that as the viewing elevation increases, SVF exhibits a stronger positive correlation, while GVI has a negative correlation with public mountain perception. This result is consistent with the prospect-refuge theory (Chen et al., 2023; Wartmann et al., 2021; Zhang et al., 2021), indicating that in high-elevation viewpoints, the public places more value on openness and broad views, and foreground vegetation may be perceived as a visual obstruction. Conversely, in low-elevation areas, the green visual index has a positive effect on public perception. These findings collectively point to one conclusion: the public's perception patterns and preferences are influenced by their different spatial locations. Correspondingly, differentiated landscape renewal strategies can be formulated for mountainous cities: for high-altitude panoramic parks and overlook platforms, excessive dense green vegetation should be avoided to guarantee sky visibility and mountain viewing openness; for low-lying riverside public spaces, reasonable green space allocation and gentle plant collocation are encouraged to optimize daily mountain landscape experience.

The conclusion regarding SWI is consistent with previous research, that is, the public's attention to mountains may decrease due to the presence of complex and diverse landscape scenes (Y. Zhang & He, 2021). It is suggested that in the renewal of complex multi-type waterfront and composite open spaces, planners should properly control the richness of miscellaneous landscape elements, maintain concise visual hierarchy, and effectively protect the legibility and visual dominance of mountain backgrounds in daily public view.

Limitations and Future Work

Although this study has revealed some influencing factors and their correlations in the public's visual perception of mountain landscapes, some limitations should be considered. Firstly, despite the large amount of public perception image data collected, it is still difficult to cover all populations due to the influence of social media platform algorithms and their inherent biases.

Moreover, the adopted social media platforms differ greatly in user attributes: Rednote and Weibo are mainly dominated by young tourist groups, while Flickr users are mostly professional landscape photographers. Such differences in user identity, aesthetic experience and viewing background may lead to heterogeneous perception characteristics of mountain landscapes. This study does not further distinguish and compare the perceptual differences among different crowds, which needs to be explored in targeted research in the future.

Secondly, this study focuses on the influence of various elements on the public's mountain visual perception density, but the correlation between these elements and public visual preferences still needs to be verified. In addition, in the subsequent GWR model construction process, the correlations between independent variables such as GVI, SVF, SWI, and MCR and DMP were not significant in the global OLS model. Since OLS assumes that the influence of variables is globally linear and constant, it may underestimate the local effects or curvilinear trends of these factors.

Finally, as a special type of urban landscape element, the complexity of mountains is reflected in their frequent vertical interweaving with the city. The role of mountains in constructing the city's cultural image has not been fully discussed. In future research, artificial structures can be further sub-classified to understand how different types of structures affect the public's perception of mountain landscapes.

Conclusion

This study establishes a comprehensive, multi-dimensional analytical framework for the visual perception of mountain landscapes. By integrating crowdsourced, geotagged social media imagery with semantic segmentation technology and spatial statistical analysis, we successfully quantified the public's visual perception patterns of mountain landscapes in Chongqing.

Our research addresses two core questions. First, from a spatial perspective, we reveal that public perception behavior for mountains is concentrated in mid-to-short distance riverside areas in terms of frequency, yet a higher perceptual intensity is associated with high-elevation locations. Second, we confirm a strong correlation between factors like building proportion, green visual index, and sky view factor with mountain landscape perception, and demonstrate the significant spatial heterogeneity exhibited by these influencing factors.

On a theoretical level, this study proposes a novel approach that integrates public perception imagery with spatial analysis, emphasizing the importance of localized and contextual research. In practical terms, our conclusions provide a theoretical basis for local governments to quantify and optimize the visual perception experience of urban mountain landscapes. At the viewpoint level, we suggest adding more viewing platforms on mountaintops and ensuring open sightlines from urban vistas. For shaping the urban-mountain scenery, the visual relationship between buildings and mountains should be considered to create a harmonious interface, and the diversity of landscape elements should be appropriately limited in important visual display areas to highlight the city's natural and cultural features.

By understanding the spatial distribution patterns and influencing factors of mountain perception, we can optimize the city's spatial structure, highlight its natural characteristics, and gain a deeper understanding of the complex, dynamic relationship between people and urban landscapes.

Author Contributions Changyu Zhang: Conceptualization, Methodology, Software, Writing – original draft, Writing – review and editingYu Cheng: Funding acquisition, Project Administration, SupervisionLingna Zhu: Funding acquisition, Investigation, Methodology, VisualizationHuasong Mao*: Conceptualization, Funding acquisition, Supervision, Writing – review and editingYizhao Du: Data curation, ValidationXinrui Yu: Writing – review and editing.

Data Availability The data used in this study consist of publicly accessible user-generated images collected from social media platforms, including 2bulu, fooooooot, Rednote, Sina Weibo, and Flickr. Due to platform terms of service, copyright protection, and privacy considerations, the original images cannot be publicly shared. The derived datasets, including visual indicators and spatial analysis results generated by the authors, are available from the corresponding author upon reasonable request.

Declarations

Competing Interests The authors declare no competing interests.

References

Ahern, J. (2013). Urban landscape sustainability and resilience: The promise and challenges of integrating ecology with urban planning and design. *Landscape Ecology*, *28*(6), 1203–1212. https://doi.org/10.1007/s10980-012-9799-z

Arriaza, M., Cañas-Ortega, J. F., Cañas-Madueño, J. A., & Ruiz-Aviles, P. (2004). Assessing the visual quality of rural landscapes. *Landscape and Urban Planning*, *69*(1), 115–125. https://doi.org/10.1016/j.landurbplan.2003.10.029

Bernbaum, E. (2022). The Spiritual and Cultural Importance of Mountains. In F. O. Sarmiento (Ed.), *Montology Palimpsest: A Primer of Mountain Geographies* (pp. 213–224). Springer International Publishing. https://doi.org/10.1007/978-3-031-13298-8_12

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, *28*(4), 281–298. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x

Bubalo, M., van Zanten, B. T., & Verburg, P. H. (2019). Crowdsourcing geo-information on landscape perceptions and preferences: A review. *Landscape and Urban Planning*, *184*, 101–111. https://doi.org/10.1016/j.landurbplan.2019.01.001

Chang Chien, Y. M., Carver, S., & Comber, A. (2020). Using geographically weighted models to explore how crowdsourced landscape perceptions relate to landscape physical characteristics. *Landscape and Urban Planning*, *203*, 103904. https://doi.org/10.1016/j.landurbplan.2020.103904

Chen, Z., Sheng, Y., Luo, D., Huang, Y., Huang, J., Zhu, Z., Yao, X., Fu, W., Dong, J., & Lan, Y. (2023). Landscape Characteristics in Mountain Parks across Different Urban Gradients and Their Relationship with Public Response. *Forests*, *14*(12), 2406. https://doi.org/10.3390/f14122406

Contributors, M. (2023). *MMSegmentation: OpenMMLab Semantic Segmentation Toolbox and Benchmark. 2020*.

Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning Special Issue: Critical Approaches to Landscape Visualization*, *142*, 173–186. https://doi.org/10.1016/j.landurbplan.2015.02.022

Europe, C. O. (2000). European landscape convention. *Report and Convention*.

Evans-Cowley, J., & Hollander, J. (2010). The New Generation of Public Participation: Internet-based Participation Tools. *Planning Practice & Research*, *25*(3), 397–408. https://doi.org/10.1080/02697459.2010.503432

Ge, Y., Chen, H., Zhang, M., & Li, X. (2022). Area threshold interval of urban forest patches required to maintain the synergy between biodiversity conservation and recreational services: Case study in Beijing, China. *Forests*, *13*(11), 1848.

Gong, C., Yang, R., & Li, S. (2024). The role of urban green space in promoting health and well-being is related to nature connectedness and biodiversity: Evidence from a two-factor mixed-design experiment. *Landscape and Urban Planning*, *245*, 105020. https://doi.org/10.1016/j.landurbplan.2024.105020

Gungor, S., & Polat, A. T. (2018). & others. Relationship between visual quality and landscape characteristics in urban parks. *Journal of Environmental Protection and Ecology*, *19*(2), 939–948.

Guo, S., Sun, W., Chen, W., Zhang, J., & Liu, P. (2021). Impact of Artificial Elements on Mountain Landscape Perception: An Eye-Tracking Study. *Land*, *10*(10), 1102. https://doi.org/10.3390/land10101102

Ives, C. D., Oke, C., Hehir, A., Gordon, A., Wang, Y., & Bekessy, S. A. (2017). Capturing residents' values for urban green space: Mapping, analysis and guidance for practice. *Landscape and Urban Planning*, *161*, 32–43. https://doi.org/10.1016/j.landurbplan.2016.12.010

Jacobs, P., Beever, E. A., Carbutt, C., Foggin, M., Juffe-Bignoli, D., Martin, M. T., Orchard, S., & Sayre, R. (2021). Identification of global priorities for new mountain protected and conserved areas: Final Report. *USGS Report*, pp. 51. https://pubs.usgs.gov/publication/70217649

Jang,K. M., & Kim, Y. (2019). Crowd-sourced cognitive mapping: A new way of displaying people's cognitive perception of urban space. *PloS One*, *14*(6), e0218590. https://doi.org/10.1371/journal.pone.0218590

Jeon, J. Y., & Jo, H. I. (2020). Effects of audio-visual interactions on soundscape and landscape perception and their influence on satisfaction with the urban environment. *Building and Environment*, *169*, 106544. https://doi.org/10.1016/j.buildenv.2019.106544

Koeser, A., Hauer, R., Andreu, M., Northrop, R., Clarke, M., Diaz, J., Hilbert, D., Konijnendijk, C., Landry, S., Thompson, G., & Zarger, R. (2023). *Using the 3-30-300 Rule to Assess Urban Forest Access and Preferences in Florida (United States)*. https://doi.org/10.20944/preprints202305.1760.v1

Korkmaz, S., & Kim, I. (2022). The optimal D: H ratio assessment for sense of enclosure in virtual landscapes. *J Digit Landsc Archit*, *7*, 425.

Ladd, B. (1990). *Urban planning and civic order in Germany, 1860–1914* (Vol. 105). Harvard University Press.

Lee, K. Y., Seo, J. I., Kim, K. N., Lee, Y., Kweon, H., & Kim, J. (2019). Application of Viewshed and Spatial Aesthetic Analyses to Forest Practices for Mountain Scenery Improvement in the Republic of Korea. *Sustainability*, *11*(9), 2687. https://doi.org/10.3390/su11092687

Li, X., Yang, T., Meng, B., Chen, S., & Zhang, S. (2025a). Analysis of Tourist Behavior Patterns and Perceptions in Beijing Based on User-Generated Content Data. *Applied Spatial Analysis and Policy*, *18*(3), 108.

Li, Y., Cao, L., & Gao, W. (2025b). Perceptual dimensions of heritage canal landscapes and their impact on conservation value cognition. *Ecological Indicators*, *176*, 113626.

Long, K., Wang, N., & Lin, Z. (2023). Assessing scenic beauty of hilly and mountain villages: An approach based on landscape indicators. *Ecological Indicators*, *154*, 110538.

Lothian, A. (1999). Landscape and the philosophy of aesthetics: Is landscape quality inherent in the landscape or in the eye of the beholder? *Landscape and Urban Planning*, *44*(4), 177–198. https://doi.org/10.1016/S0169-2046(99)00019-5

Luo, S., Liu, Y., Du, M., Gao, S., Wang, P., & Liu, X. (2021). The influence of spatial grid division on the layout analysis of urban functional areas. *ISPRS International Journal of Geo-Information*, *10*(3), 189.

Luo, J., Zhao, T., Cao, L., & Biljecki, F. (2022). Water View Imagery: Perception and evaluation of urban waterscapes worldwide. *Ecological Indicators*, *145*, 109615. https://doi.org/10.1016/j.ecolind.2022.109615

Luo,S., Xie, J., & Furuya, K. (2023). Effects of perceived physical and aesthetic quality of urban blue spaces on user preferences–A case study of three urban blue spaces in Japan. *Heliyon*, *9*(4), e15033. https://doi.org/10.1016/j.heliyon.2023.e15033

Matsuoka, R. H., & Kaplan, R. (2008). People needs in the urban landscape: Analysis of Landscape And Urban Planning contributions. *Landscape and Urban Planning*, *84*(1), 7–19. https://doi.org/10.1016/j.landurbplan.2007.09.009

Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(7), 3523–3542.

Misthos, L. M., Nakos, B., Krassanakis, V., & Menegaki, M. (2019). The effect of topography and elevation on viewsheds in mountain landscapes using geovisualization. *International Journal of Cartography*, *5*(1), 44–66.

Molla, M. B. (2015). Urbanization process in developing countries: A review on urban ecosystem degradation and public health effect. *Research Journal of Agriculture and Environmental Management*, *4*(4), 291–298.

Niu, H., Seraphim, A. P., Morgado, P., Miranda, B., & Silva, E. A. (2025). Mapping urban emotion from geotagged social media data: Age, gender and spatial heterogeneity. *Applied Geography*, *185*, 103768.

Parnell, S., Schewenius, M., Sendstad, M., Seto, K. C., & Wilkinson, C. (2013). *Urbanization, biodiversity and ecosystem services: Challenges and opportunities*. Springer.

Petrova, E. G., Mironov, Y. V., Aoki, Y., Matsushima, H., Ebine, S., Furuya, K., Petrova, A., Takayama, N., & Ueda, H. (2015). Comparing the visual perception and aesthetic evaluation of natural landscapes in Russia and Japan: Cultural and environmental factors. *Progress in Earth and Planetary Science*, *2*(1), 6. https://doi.org/10.1186/s40645-015-0033-x

Planning Department, H. K. G. (1992). *Hong Kong planning standards and guidelines*.

Price, M. F., Arnesen, T., Gløersen, E., & Metzger, M. J. (2019). Mapping mountain areas: Learning from Global, European and Norwegian perspectives. *Journal of Mountain Science*, *16*(1), 1–15.

Qi, Z., Duan, J., Su, H., Fan, Z., & Lan, W. (2023). Using crowdsourcing images to assess visual quality of urban landscapes: A case study of Xiamen Island. *Ecological Indicators*, *154*, 110793. https://doi.org/10.1016/j.ecolind.2023.110793

Schirpke, U., Tasser, E., & Tappeiner, U. (2013). Predicting scenic beauty of mountain regions. *Landscape and Urban Planning*, *111*, 1–12. https://doi.org/10.1016/j.landurbplan.2012.11.010

Song, H., Chen, J., & Li, P. (2025). Decoding the cultural heritage tourism landscape and visitor crowding behavior from the multidimensional embodied perspective: Insights from Chinese classical gardens. *Tourism Management*, *110*, 105180.

Tenerelli, P., Püffel, C., & Luque, S. (2017). Spatial assessment of aesthetic services in a complex mountain region: Combining visual landscape properties with crowdsourced geographic information. *Landscape Ecology*, *32*(5), 1097–1115.

Tieskens, K. F., Van Zanten, B. T., Schulp, C. J. E., & Verburg, P. H. (2018). Aesthetic appreciation of the cultural landscape through social media: An analysis of revealed preference in the Dutch river landscape. *Landscape and Urban Planning*, *177*, 128–137. https://doi.org/10.1016/j.landurbplan.2018.05.002

Wang, C., Xu, W., & Zhang, J. (2025). Geographies of Online Social Interaction: A Novel Big Data Analytics Approach to Social Media Platform–Sina Weibo. *Applied Spatial Analysis and Policy*, *18*(4), 1–28.

Wartmann, F. M., Frick, J., Kienast, F., & Hunziker, M. (2021). Factors influencing visual landscape quality perceived by the public. Results from a national survey. *Landscape and Urban Planning*, *208*, 104024. https://doi.org/10.1016/j.landurbplan.2020.104024

Wei, J., Yue, W., Li, M., & Gao, J. (2022). Mapping human perception of urban landscape from street-view images: A deep-learning approach. *International Journal of Applied Earth Observation and Geoinformation*, *112*, 102886. https://doi.org/10.1016/j.jag.2022.102886

Wu, J. (2010). Urban sustainability: An inevitable goal of landscape research. *Landscape Ecology*, *25*(1), 1–4. https://doi.org/10.1007/s10980-009-9444-7

Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. *Advances in Neural Information Processing Systems*, *34*, 12077–12090. https://proceedings.neurips.cc/paper/2021/hash/64f1f27bf1b4ec22924fd0acb550c235-Abstract.html

Xie, Y., Zhu, M., Wang, F., Zhao, J., Sun, W., Qin, Y., & Zhang, H. (2025). Unveiling changes in urban visual perception inequality in high-density cities: Insights from housing types and travel distances using time-series street view imagery. *Applied Geography*, *185*, 103776.

Xu, M., Luo, T., & Wang, Z. (2020). Urbanization diverges residents' landscape preferences but towards a more natural landscape: Case to complement landsenses ecology from the lens of landscape perception. *International Journal of Sustainable Development & World Ecology*, *27*(3), 250–260. https://doi.org/10.1080/13504509.2020.1727989

Yang, C., & Liu, T. (2022). Social media data in urban design and landscape research: A Comprehensive literature review. *Land*, *11*(10), 1796.

Zhang, Y., & He, J. (2021). Understanding the visual image of Kailash Sacred Landscape through geotagged landscape photos mapping. *Environmental Challenges*, *5*, 100360. https://doi.org/10.1016/j.envc.2021.100360

Zhang, G., Yang, J., Wu, G., & Hu, X. (2021). Exploring the interactive influence on landscape preference from multiple visual attributes: Openness, richness, order, and depth. *Urban Forestry & Urban Greening*, *65*, 127363. https://doi.org/10.1016/j.ufug.2021.127363

Zhang, N., Zheng, X., & Wang, X. (2022a). Assessment of Aesthetic Quality of Urban Landscapes by Integrating Objective and Subjective Factors: A Case Study for Riparian Landscapes. *Frontiers in Ecology and Evolution*, *9*, 735905. https://doi.org/10.3389/fevo.2021.735905

Zhang, X., Xu, D., & Zhang, N. (2022b). Research on Landscape Perception and Visual Attributes Based on Social Media Data—A Case Study on Wuhan University. *Applied Sciences*, *12*(16), 8346. https://doi.org/10.3390/app12168346

Zhang, X., Lin, E. S., Tan, P. Y., Qi, J., & Waykool, R. (2023). Assessment of visual landscape quality of urban green spaces using image-based metrics derived from perceived sensory dimensions. *Environmental Impact Assessment Review*, *102*, 107200. https://doi.org/10.1016/j.eiar.2023.107200

Zhao, Y., Zhao, Y., Liu, X., & Zhou, Y. (2024). Enhancing conservation and management of mountain area through a biocultural diversity evaluation approach: A case study of Taishan Mountain Area, China. *Journal of Cleaner Production*, *449*, 141716.

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). *Scene Parsing Through ADE20K Dataset*. 633–641. https://openaccess.thecvf.com/content_cvpr_2017/html/Zhou_Scene_Parsing_Through_CVPR_2017_paper.html

Zhou, B., Zhao, H., Puig, X., Xiao, T., Fidler, S., Barriuso, A., & Torralba, A. (2019). Semantic understanding of scenes through the ade20k dataset. *International Journal of Computer Vision*, *127*(3), 302–321.

Zhou, X., Cen, Q., & Qiu, H. (2023). Effects of urban waterfront park landscape elements on visual behavior and public preference: Evidence from eye-tracking experiments. *Urban Forestry & Urban Greening*, *82*, 127889. https://doi.org/10.1016/j.ufug.2023.127889

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Changyu Zhang[1] · Yu Cheng[2,3] · Lingna Zhu[1] · Huasong Mao[1] · Yizhao Du[4,5] · Xinrui Yu[1]

✉ Huasong Mao
Maohuasong@cqu.edu.cn

Changyu Zhang
changyuzhang@cqu.edu.cn

Yu Cheng
e1514009@u.nus.edu

Lingna Zhu
zhuln777@163.com

Yizhao Du
Y.du-4@tudelft.nl

Xinrui Yu
accylau@gmail.com

[1] School of Architecture and Urban Planning, Chongqing University, Chongqing, China

[2] Department of Architecture, National University of Singapore, Singapore, Singapore

[3] School of Architecture and Urban Planning, Chongqing University, Chongqing, China

[4] Department of Urbanism, Delft University of Technology, Delft, Netherlands

[5] Department of Land Economy, University of Cambridge, Cambridge, United Kingdom