International Journal of Geo-Information

MDPI

Article

Mining the Tourism Destination Image and Analyzing Influence Mechanisms

Shan Huang, Xu Lu *, Jingqun Lu and Jinghua Zhang

School of Architecture and Urban Planning, Shenyang Jianzhu University, Shenyang 110168, China; huangshan@stu.sjzu.edu.cn (S.H.); ljq@stu.sjzu.edu.cn (J.L.); zjh49@stu.sjzu.edu.cn (J.Z.)
* Correspondence: luxu@sjzu.edu.cn

Abstract

Background: Research on spatial imagery as perceived by humans is an important frontier for deepening the theoretical understanding of Tourism Destination Image and promoting sustainable urban development. Significance: This study, from the perspective of tourists, explores the correlation mechanism between the cognitive image and affective image of urban space. This is of great significance for enhancing the overall spatial quality of cities, promoting the integration of the man–land relationship, and driving the sustainable development of tourism. Method: In this study, we took Harbin as the case site, collected 89,375 reviews and 23,561 review images of 488 scenic spots on the Mafengwo and Ctrip platforms, and constructed a multimodal dataset. We classified the image scenes with the help of the Places365-CNN model. We then extracted text emotional features by utilizing the SnowNLP deep learning algorithm. We constructed a map of the spatial influence mechanism acting on cognitive image and emotion through MGWR. Results: The experimental results showed that in the level of Pleasure, the five indicators NHS, HPA, RPA, PDS and WRV had significant spatial correlations with urban sentiment. In the level of Arousal, the three indicators PD, MaSD and WRV showed significant spatial characteristics. Conclusions: This study reveals the influence mechanism of urban spatial perception elements on tourists' emotions. It not only deepens the understanding of the Tourism Destination Image theory, but also provides a practical path based on the optimization of perception scenarios for the improvement of urban space, which has important implications for regional sustainable development.

Keywords: perceived image; emotional maps; tourism destination image; deep learning; Harbin

Check for updates

Academic Editors: Wolfgang Kainz, Wanzeng Liu, Wenhao Yu, Tian Lan, Wende Li and Peng Ti

Received: 8 December 2025
Revised: 2 February 2026
Accepted: 7 February 2026
Published: 12 February 2026

Copyright: © 2026 by the authors. Published by MDPI on behalf of the International Society for Photogrammetry and Remote Sensing. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Since the 20th century, China's urbanization rate has been steadily increasing, rising from 18% to over 67% from 1978 to 2024 [1]. However, with the continuous spread of the trend of globalization, the conflict of the man–land relationship has become increasingly prominent in urban spaces, making urban development homogenous and unsustainable. Therefore, accurately assessing the characteristics of the urban spatial environment as perceived by humans is a key focus of sustainable development [2–5]. Existing research relies on static and non-visual data and rarely incorporates human perception aspects into the evaluation of urban sustainable development [6].

Lynch (1960) put forward a core viewpoint: the value of a city lies not only in its objectively existing material form, but also in the way it is perceived and remembered

by its residents [7]. Subsequently, urban planners have been exploring and excavating the subject of imagery. An image of a city is generated in the brain and expressed in a Cognitive Map. A Cognitive Map is the result of the active processing and construction of visual cognitive information about urban space by the brain [8]. The formation of human perception often relies on the continuous repetition of specific scenes within a spatial range [9]. The construction of a perceived image takes high repeatability as an important prerequisite. This repetition is not only reflected in the frequency of the occurrence of related elements but is also significantly manifested in spatial agglomeration characteristics [10]. The centralized distribution of resources in geographical space makes it easier to shape a unified and distinct image perception.

Constructing a Cognitive Map of a city from the perspective of tourists is an effective way to analyze the formation mechanism and spatial laws of the city's perceived image. Tourists, as they do not live in the city for a long time, are more scattered and fragmented in their perceptions compared to local residents. Therefore, in the urban image, the saliency of scene elements represents the perceived images of tourists, which are the focus when exploring the images of tourist destinations [11].

Meanwhile, emotions are not only a main driving force of tourism consumption but also provide an important channel for exploring the construction of tourism-friendly cities. Although some scholars have interpreted the relationship between emotions and urban space from the perspective of residents [12–14], research into mapping and understanding urban emotional maps from a tourist perspective remains limited [15].

Analyzing the influence mechanism of the spatial distribution state of the perceived image on the emotional map is key to deepening the cognition of Tourism Destination Image (TDI) from the material dimension to the perceptual dimension [16]. In tourists' gazing behavior, visual feature density has been confirmed to be an important factor in predicting emotions [17]. In terms of emotional quantification, research has established a consensus framework, focusing primarily on the two core dimensions of Pleasure and Arousal. Pleasure characterizes the polarity of emotions, while Arousal represents the level of physiological activation.

This study adopts the unique perspective of tourists to conduct a spatial correlation analysis between the distribution of perceived image and the emotional map. The case study focuses on Harbin, a leading city for ice and snow tourism in China. As a widely recognized "Ice City", Harbin has developed a typical TDI based on its unique endowment of ice and snow resources and its cultural heritage. However, against the backdrop of the rapid expansion of the ice and snow tourism market, Harbin now faces the challenge of intensifying homogenized competition.

In response to the aforementioned issues, this study proposes a methodological framework. First, data (including coordinates, timestamps, comments and images) were crawled from the Mafengwo and Ctrip platforms to construct a multimodal dataset. Second, the Places365-CNN model was employed to classify Cognitive Image and construct a geographic information database of perceived images. Simultaneously, the SnowNLP algorithm was used to extract emotional keywords. Finally, a spatial correlation network linking cognitive and affective images was constructed. Using multi-scale geographically weighted regression (MGWR), the influence mechanisms of various cognitive elements on affective responses at the spatial level were determined. This framework aims to address the following key issues:

(1) What are the spatial layout characteristics of a city's perceived images?
(2) What is the influence mechanism of the spatial distribution of perceived images on the emotional map?

In conclusion, this study explores the mechanism through which perceived image influence emotional maps, thereby investigating the patterns of the man–land relationship. These findings are significant for optimizing tourism experiences, enhancing the quality of urban space, promoting sustainable urban development, and advancing interdisciplinary research.

2. Literature Review

2.1. Tourism Destination Image

Early TDI research defined the concept of tourism destination image as the overall impression people hold about a non-residential place [18]. It was then further refined into a collection of individuals' beliefs, thoughts and impressions about a destination [19]. As research progressed, scholars recognized that in addition to the evaluation of functional and psychological attributes, more attention should be paid to the composite image consisting of a mental picture and an overall view [20]. On this basis, subsequent studies introduced the Behavioral Conative Image, which deconstructed the TDI system into three interrelated components of Cognitive–Affective–Conative Image [21]. The theory has evolved to emphasize integration, a composite model integrating impressions, feelings and beliefs has been put forward, and the analytical paradigm of the "Cognitive-Affective-Overall Model" has been developed, which has moved the theory closer to being systematic and perfect [22].

In previous research on Cognitive Image, the exploration of its spatial features initially relied on traditional methods such as Cognitive Map. Subsequently, research has gradually shifted towards the integration of multi-source data and deep learning—for instance, analyzing scene characteristics through the physical spatial layout of POIs, thereby offering a more refined and novel perspective for understanding the composition of cognitive images [6]. The key role of visual elements in the construction and dissemination of TDI has also been further recognized in this process [23]. In recent years, the research perspective has been further expanded to group-level perception and its complex impact on urban emotions, deepening the understanding of the mechanisms of TDI shaping [24].

In terms of Affective Image, this study follows Russell's Affective Loop Theory and is centered around two core dimensions, Pleasure and Arousal [25]. Pleasure characterizes the positive and negative polarity of emotion (negative/positive), while Arousal represents the physiological activation level (low/high). Together, these two dimensions form a spatial model of emotion that can categorize complex emotions into basic types such as arousal, relaxation, irritation and boredom.

2.2. Deep Learning and Spatial Analysis

Early studies of TDI relied heavily on the official channels. These included brochures, postcards, guide maps and other traditional printed materials, as well as official websites and other media. With the widespread adoption of Internet technology and the explosive growth of User-Generated Content (UGC), the research data sources have gradually shown a diversified trend, and demand-side information, such as blogs, online reviews, and travel photos and videos shared on social media, have begun to be widely included in analysis frameworks [26].

Deep learning techniques have the powerful ability to enable the recognition and characterization of perceptual images, which can be classified into text and images based on the type of data source [23]. The main applications in text analysis are word cloud statistics, DeepSentiBank, SnowNLP, LDA topic modeling, and many other text semantic mining methods [27,28]. Image mining has realized multiple developments in Inception-v4

scene classification, Places365-CNN scene recognition, DenseNet-161 object recognition, target detection, and NIMA aesthetic evaluation [29,30].

The existing literature on TDI has made significant progress in the theoretical framework and computational methods, but there are still the following key limitations. First, most studies separate the Cognitive and Affective dimensions and lack an integrated explanation of the intrinsic mechanism of "how scenes trigger emotions". Second, the methodology often favors single-modal data and fails to take full advantage of the complementary nature of multimodal UGC data for collaborative analysis. Third, spatial analysis is mostly based on global or traditional spatial models, which makes it difficult to reveal the heterogeneity and multi-scale nature of the influence mechanism across geographical location [31].

To this end, this study proposes an integration framework. The specific contributions are first, the use of tourists' pictures and comments to achieve paired quantification of Cognitive and Affective dimensions through scene classification and sentiment analysis models, and second, the introduction of MGWR to reveal how the effects of different perceptual elements on the Affective dimensions vary spatially, so as to visualize the local context and spatial heterogeneity pattern. This study not only deepens and integrates existing theories and methods, but also provides a quantifiable and localizable empirical example of TDI theory at the micro-perception level.

3. Data and Method

3.1. Research Subjects and Data Sources

3.1.1. Research Subjects

The research subject is Harbin. It has unique climatic characteristics and rich tourism resources. Winter temperatures in Harbin average around −15 °C, with extremes below −30 °C [32]. The average number of days covered by snow is 105, which provides unique natural conditions for ice and snow tourism.

The research scope covers the area within the Third Ring Road of Harbin (Figure 1). This study established 168 continuous areal spatial units as the fundamental analytical framework, based on the street boundaries of Harbin. All point-based data were aggregated into their corresponding areal units via the spatial join function of a Geographic Information System.

3.1.2. Data Sources

Mafengwo and Ctrip are two major travel social networking websites in China, where users can share review information about their travel experiences with geographical locations. This study collected 488 scenic spot names and geographical coordinates, with 89,375 valid comments and 23,561 valid images. The collection period ends on 31 May 2025 (Table 1).

Table 1. Data sources.

Platform	Data Type	URL	Temporal Coverage
Mafengwo	Scenic spot coordinates, UGC reviews, UGC images	https://www.mafengwo.cn/mdd/map/10068.html (accessed on 10 June 2025)	January 2016–May 2025
Ctrip		https://www.ctrip.com/webapp/tripmap/travel?entranceId=Ctriponlinehomeside (accessed on 10 June 2025)	January 2016–May 2025

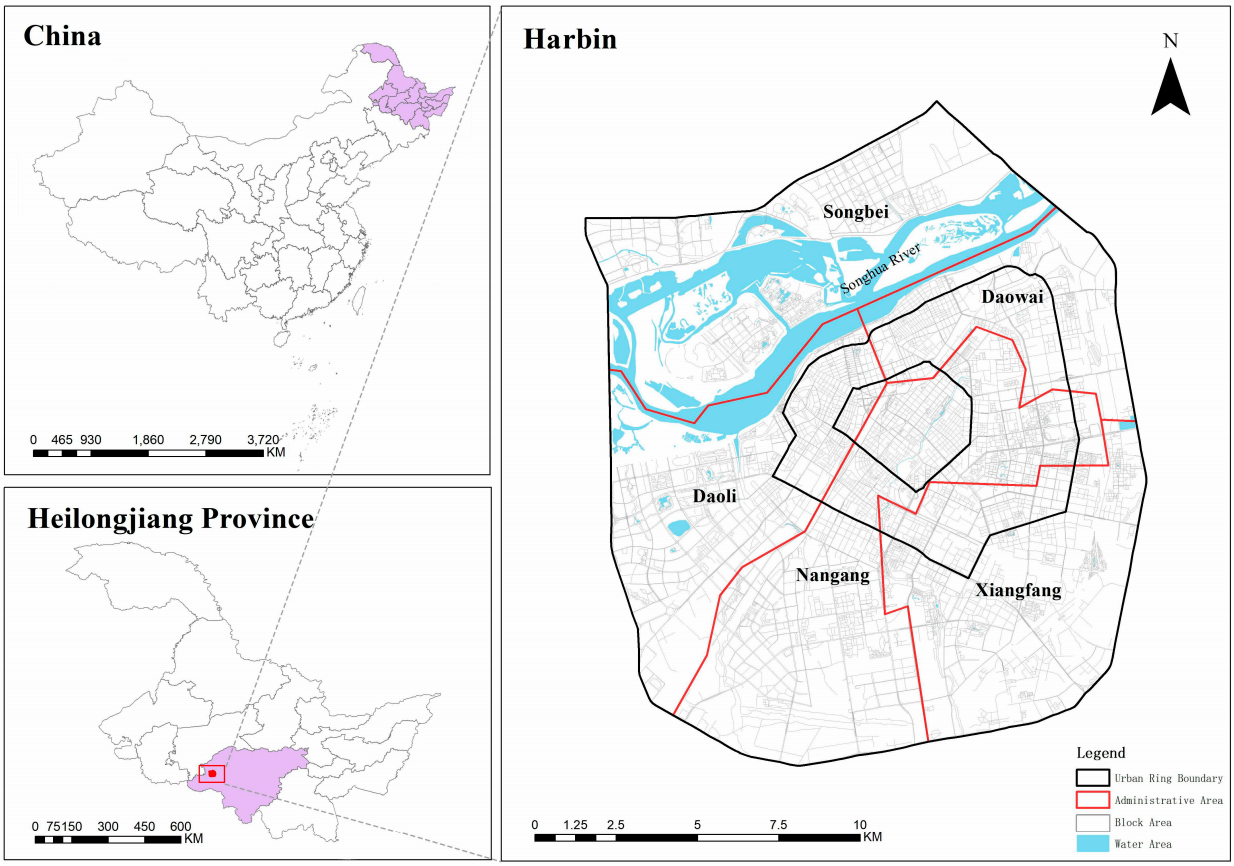


Figure 1. Schematic diagram of research scope.

3.2. Research Methods

3.2.1. Places365-CNN

Scenes in the scenic spots were extracted using the scene classification model Places365-CNN, and a convolutional neural network trained on the Places2 database was used to verify the images (Figure 2). Places365-CNN can recognize different types of scenes and category activation maps (Figure 3). The Places2 database contains over 10 million images and 365 unique scene categories, and its design follows the principles of human visual cognition [33]. This dataset contains 5000 to 30,000 training images for each category, which are consistent with occurrence frequency in the real world. The real world is complex, with various scenarios crisscrossing each other [34]. Therefore, this study determined that scene categories with a probability higher than 0.1 in the classification results of each photo are valid. Finally, occurrence frequencies of 365 scenes for each image were obtained.

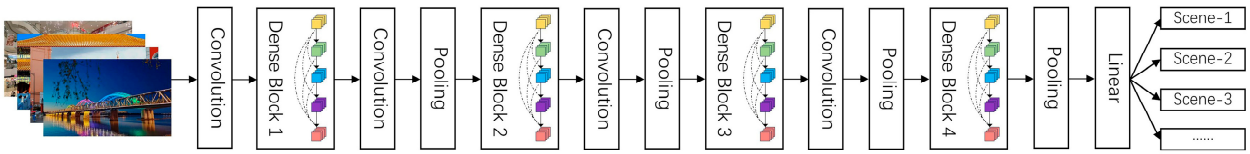


Figure 2. Convolution diagram.

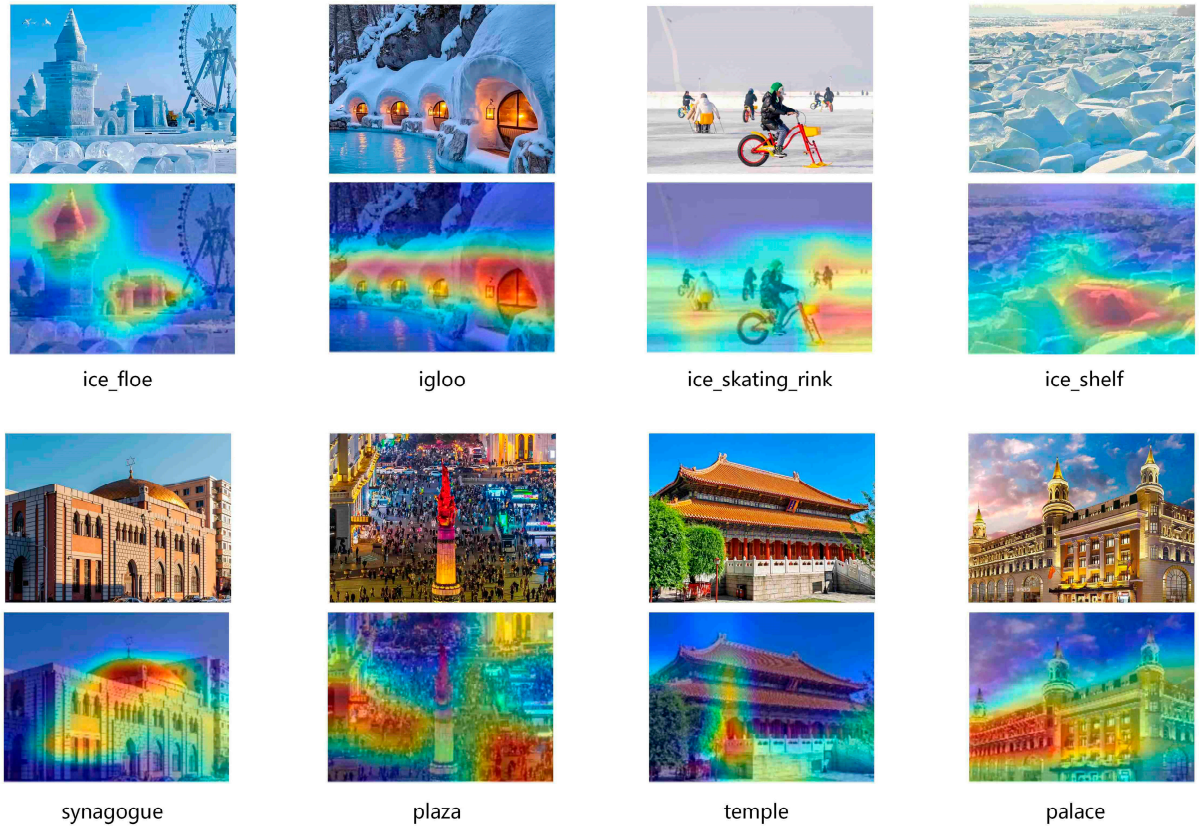


Figure 3. Scene category activation map.

3.2.2. SnowNLP

An algorithm for training deep learning models in combination with the corpus method is a mainstream research approach for text comment sentiment analysis currently [35]. The core of the corpus method lies in conducting a statistical analysis of a large amount of text data from real language usage scenarios, and thereby extracting the rules and patterns of language and providing a basis for the construction and optimization of language models.

Accuracy of emotion recognition is much lower than that of object recognition [36]. This is caused by a "semantic gap" between low-level features and high-level semantics [37]. Deep learning helps solve this problem. Photos on social networks are usually accompanied by text information, which contains the emotions of tourists visiting these places. This is ideal material for sentiment analysis [38]. Deep learning can automatically learn the features and patterns of language from large amounts of text data by constructing multi-layer neural network structures, thereby achieving efficient recognition and the classification of emotions [30]. Commonly used models for sentiment analysis include RNN, CNN, LSTM, BERT, Inception-v4, DeepSentiBank, and SnowNLP [9,15,39–41].

This study followed Russell's Affective Loop Theory [25]. Based on the sentiment classification results matched with a corpus, the SnowNLP deep learning model was trained for sentiment analysis, and then the levels of Pleasure and Arousal for all comment data within the study scope were obtained.

3.2.3. MGWR

This study adopted MGWR to analyze a spatial differentiated influence path of scene elements on tourists' emotions, and identified key driving factors of TDI and their intensity of effect.

Ordinary Least Squares (OLS) regression assumes globally homogeneous relationships, unlike MGWR. MGWR also differs from Spatial Lag Models (SLM) and Spatial Error Models (SEM), which mainly address spatial dependence. The core advantage of MGWR is its ability to deal with spatial non-stationarity and multiscale effects at the same time. This means that it allows the intensity, direction, and even effective range of the effect of each independent variable on the perceived image to vary with geospatial location. The scale of effect of each variable is determined by model data-driven, independent optimization.

Before establishing the model, it is necessary to test for multicollinearity among the independent variables. To reduce the influence of inter-variable correlations on the results, we excluded variables with VIF > 7.5. The formula is as follows:

$$VIF_j = \frac{1}{\left(1 - R_j^2\right)} \quad (1)$$

where VIF_j represents Variance Inflation Factor and R_j^2 the coefficient of determination for the j-th independent variable obtained from the regression model.

MGWR is a spatial statistical technique used to explore and model the spatial variation relationship between dependent and independent variables at different positions. MGWR is mainly applied in research such as in land use and block vitality evaluation [42]. In research related to emotion perception, scholars have utilized MGWR to analyze a spatial influence mechanism of the built environment on emotions [43]. The effectiveness of MGWR in research on spatial emotions has been proven. The calculation formula for MGWR is as follows:

$$Y_i = \beta_0(u_i, v_i, s_i) + \sum_{j=1}^{k} \beta_j(u_i, v_i, s_i) x_{ij} + \varepsilon_i \quad (2)$$

where Y_i represents the dependent variable for observation i; $\beta_0(u_i, v_i)$ and $\beta_j(u_i, v_i)$ are location-specific coefficients; x_{ij} is the value of the j-th independent variable for observation i; ε_i is the error term; (u_i, v_i) are the spatial coordinates for observation i; and s_i is the scale parameter for observation i.

3.3. Research Framework

3.3.1. Technical Route

This study took photos and comments in UGC as its main data sources and used deep learning and spatial analysis methods to construct a TDI analysis framework from the perspective of tourists. This framework is divided into three main parts: data collection and processing, quantitative analysis, and spatial visualization analysis (Figure 4).

Firstly, we utilized UGC on the tourism platform to obtain effective data through data crawling, data cleaning, and data quantification.

Secondly, we quantified the scene and emotional data perceived by tourists. Places365-CNN was utilized to obtain scene classification data perceived by tourists from images shared by tourists, while SnowNLP was utilized to obtain the levels of Pleasure and Arousal in emotions from visitor reviews.

Finally, based on ArcGIS 10.8, MGWR was utilized to determine a spatial influence mechanism of scenes on urban emotions. This provides a theoretical basis for managers to promote the sustainable development of tourism-friendly cities.

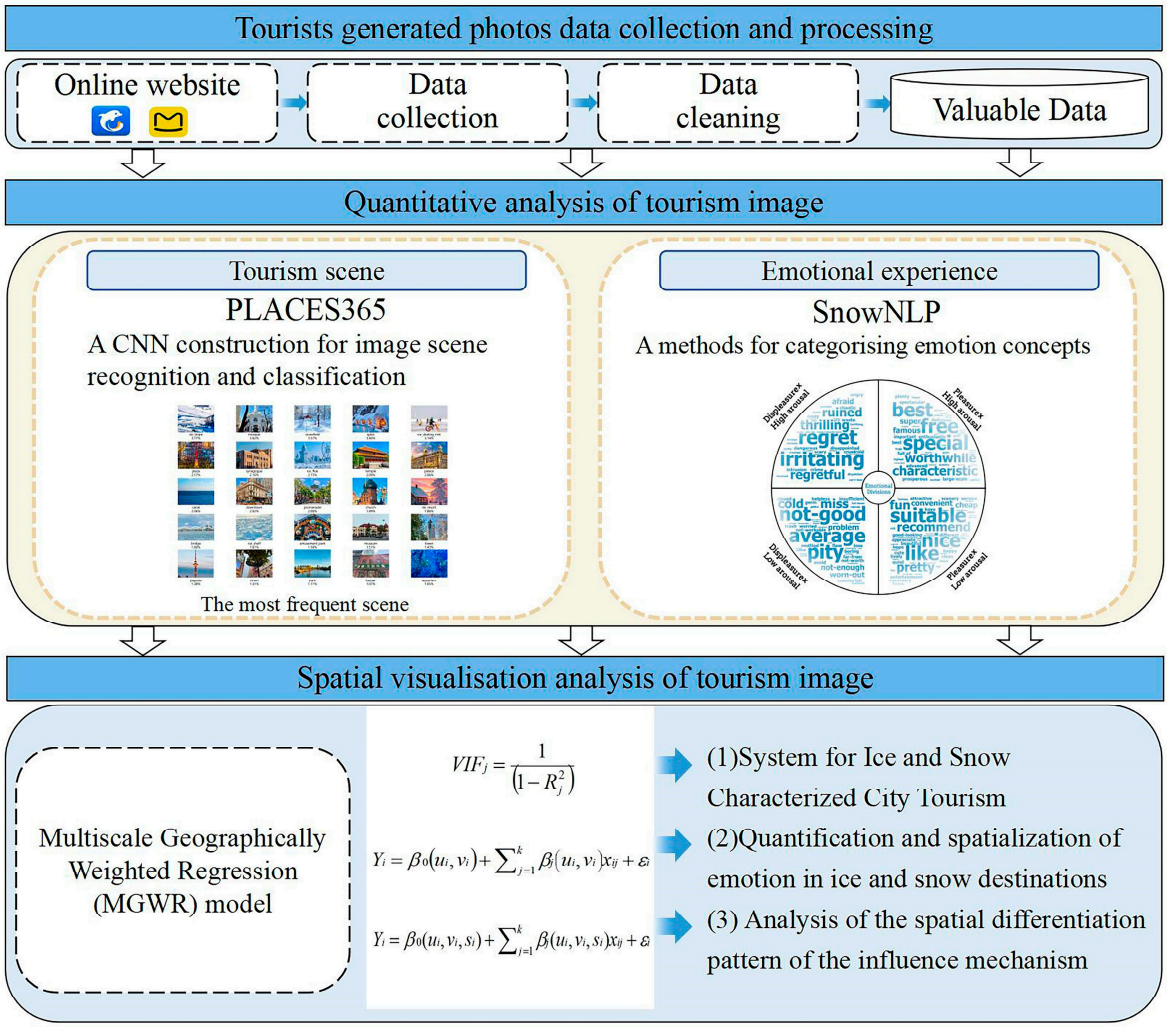


Figure 4. Technical roadmap.

3.3.2. Quantitative Framework

Guided by Clark's Scenes Theory and tailored to the regional characteristics of Harbin, this study developed a 33-indicator evaluation framework for analyzing the perceived image of tourism destinations [44]. The framework is organized into five dimensions—Physical Space, Visual Composition, Collective Perception, Sensory Experience, and Characteristic Information—which are detailed below (Table 2):

(1) The Physical Space dimension is represented by POI point information, reflecting the physical and spatial characteristics of the scene [6].
(2) The Visual Composition dimension is represented by scene labels recognized through the Places365 model, capturing the visual features of the environment [23].
(3) The Collective Perception dimension is represented by topic-related descriptions extracted from UGC texts, indicating shared perceptions among visitors [24].
(4) The Sensory Experience dimension is represented by characteristic regional features, conveying subjective visitor feelings and experiential richness [6].
(5) Finally, the Characteristic Information dimension is introduced to represent incidental information commonly noted in tourism reviews.

Table 2. Scene Indicator System.

Goal	Criteria	Indicator		Calculation Method	Formula	Note
Spatial cognitive image features	Physical Space	PD	POI Density	Number of Points of Interest (POIs) per unit area	$D = \frac{N}{A}$	D: POI Density; N: Total number of POIs; A: Area of spatial unit
		PDI	POI Diversity	Measuring the mixing degree of POI types based on the Simpson's Diversity Index	$SDI = 1 - \frac{\sum n_i(n_i-1)}{N(N-1)}$	SDI: Simpson's Diversity Index; n_i: Number of POIs in category i; N: Total number of POIs
		PC	POI Count	Total occurrence count of all POIs	N	N: Total number of POIs in the spatial unit
	Visual Composition	NHS	Number of Historical Scenes	Number of scenes identified by the model as historical buildings	$N_{historical}$	N: Total number of POIs in the spatial unit
		NAS	Number of Human Activity Scenes	Number of scenes identified by the model as related to human activities	N_{human}	
		NNS	Number of Natural Scenes	Number of scenes identified by the model as natural landscapes	$N_{natural}$	
		NRS	Number of Resort Scenes	Number of scenes identified by the model as related to tourism and resorts	N_{resort}	
		NIS	Number of Infrastructure Scenes	Number of scenes identified by the model as urban infrastructure	$N_{infrastructure}$	
		MaSR	Macro-category Scene Richness	Number of different macro scene categories present within the unit	R_{macro}	R: Scene Richness
		MiSR	Micro-category Scene Richness	Number of different fine-grained categories (from 365 classes) present within the unit	R_{micro}	
		MaSD	Macro-category Scene Diversity	Mixing degree of different scene categories based on Simpson's Diversity Index	$SDI_{macro} = 1 - \frac{\sum n_i(n_i-1)}{N(N-1)}$	SDI: Simpson's Diversity Index; n_i: Number of scenes in major category i; N: Total number of scenes
		MiSD	Micro-category Scene Diversity	Mixing degree of 365 fine-grained scene categories based on Simpson's Diversity Index	$SDI_{micro} = 1 - \frac{\sum n_i(n_i-1)}{N(N-1)}$	
		SD	Scene Density	Total number of scenes per unit area	$D_{scene} = \frac{N_{scene}}{A}$	D: Scene Density; N: Total number of scenes; A: Area of spatial unit
		DSC	Dominant Scene Concentration	Proportion of the most frequently occurring scene category	$C = \frac{n_{max}}{N}$	C: Concentration; n_{max}: Number of scenes in the most frequent category; N: Total number of scenes
		TSC	TOP25 Scene Concentration	Cumulative proportion of the top 25 most frequently occurring scene categories	$C_{top25} = \frac{\sum_{i=1}^{25} n_i}{N}$	

Table 2. *Cont.*

Goal	Criteria	Indicator		Calculation Method	Formula	Note
Spatial cognitive image features	Collective Perception	HPA	Historical Perception Abundance	Proportion of perception descriptions in UGC text identified as related to historical themes	$A = \frac{F_{historical}}{F}$	A: Perception Abundance; F: Perception frequency
		APA	Human Activity Perception Abundance	Proportion of perception descriptions in UGC text identified as related to daily human life themes	$A_{human} = \frac{F_{human}}{F}$	
		NPA	Natural Perception Abundance	Proportion of perception descriptions in UGC text identified as related to natural landscape themes	$A_{\text{natural}} = \frac{F_{\text{natural}}}{F}$	
		RPA	Resort Perception Abundance	Proportion of perception descriptions in UGC text identified as related to tourism and resort themes	$A_{\text{resortl}} = \frac{F_{\text{resort}}}{F}$	
		IPA	Infrastructure Perception Abundance	Proportion of perception descriptions in UGC text identified as related to infrastructure themes	$A_{\text{infrastructurel}} = \frac{F_{\text{infrastructuret}}}{F}$	
		MaPD	Macro-category Perception Diversity	Measuring the mixing degree of different macro perception themes based on Simpson's Diversity Index	$SDI_{percept-macro} = 1 - \frac{\sum n_i(n_i-1)}{N(N-1)}$	SDI: Simpson's Diversity Index; n_i: Frequency of perceptions in major category i; N: Total perception frequency
		MiPD	Micro-category Perception Diversity	Measuring the mixing degree of 365 fine-grained perception themes based on Simpson's Diversity Index	$SDI_{percept-micro} = 1 - \frac{\sum n_i(n_i-1)}{N(N-1)}$	
		PDS	Perception Density	Total frequency of UGC perception descriptions per unit area	$D_{percept} = \frac{N_{percept}}{A}$	D: Percept Density; N: Total perception frequency; A: Area of spatial unit
		DPC	Dominant Perception Concentration	Proportion of the most frequently occurring perception category	$C_{percept-dom} = \frac{n_{percept-\max}}{N_{percept}}$	C: Concentration; n_{max}: Number of percept in the most frequent category; N: Total number of scenes
		TPC	TOP25 Perception Concentration	Cumulative proportion of the top 25 most frequently occurring fine-grained perception categories	$C_{percept-top25} = \frac{\sum_{i=1}^{25} n_{percept-i}}{N_{percept}}$	
	Sensory Experience	SSS	Snow Scene Salience	Proportion of visual scenes identified as related to ice and snow	$P_{snow-scene} = \frac{N_{snow-scene}}{N_{scene}}$	P: Proportion; N: Total number of scenes
		SPS	Snow Perception Salience	Proportion of perception descriptions in UGC text identified as related to ice and snow themes	$P_{snow-percept} = \frac{F_{snow-percept}}{F}$	P:Proportion; F: Perception frequency
		ASR	Average Scene Richness	Ratio of the number of dominant macro scene categories to the number of POIs, reflecting the core visual theme intensity carried by each POI	$R_{avg-scene} = \frac{n_{dom-scene}}{N_{POI}}$	R: Scene Richness; n: Number of scene category; N: Total number of scenes
		FSR	Featured Scene Richness	Ratio of the number of top 25 high-frequency scenes to the number of POIs, reflecting the featured visual experience intensity carried by each POI	$R_{feature-scene} = \frac{n_{top25-scene}}{N_{POI}}$	
		APR	Average Perception Richness	Ratio of the frequency of dominant macro perception categories to the number of POIs, reflecting the core perception intensity triggered by each POI	$R_{avg-percept} = \frac{n_{dom-percept}}{N_{POI}}$	
		FPR	Featured Perception Richness	Ratio of the frequency of top 25 high-frequency perceptions to the number of POIs, reflecting the featured perception intensity triggered by each POI	$R_{feature-percept} = \frac{n_{top25-percept}}{N_{POI}}$	
	Characteristic Information	WRV	Winter Review Volume	Total number of UGC reviews published during the winter period (20 October–20 April)	N_{winter}	N: Total number of reviews
		NRV	Non-winter Review Volume	Total number of UGC reviews published during the non-winter period	$N_{non-winter}$	

4. Results and Analysis

4.1. Evaluation of TDI

4.1.1. Cognitive Image Evaluation

Photo scene extraction is a prerequisite and foundation for scientifically analyzing the perceived image of a tourist destination [9]. Data in this study involved 357 scenarios out of the 365. The top 20 tourism scenarios are presented in the figure. These photos are main indicators of how tourists perceived the Harbin TDI (Figure 5).

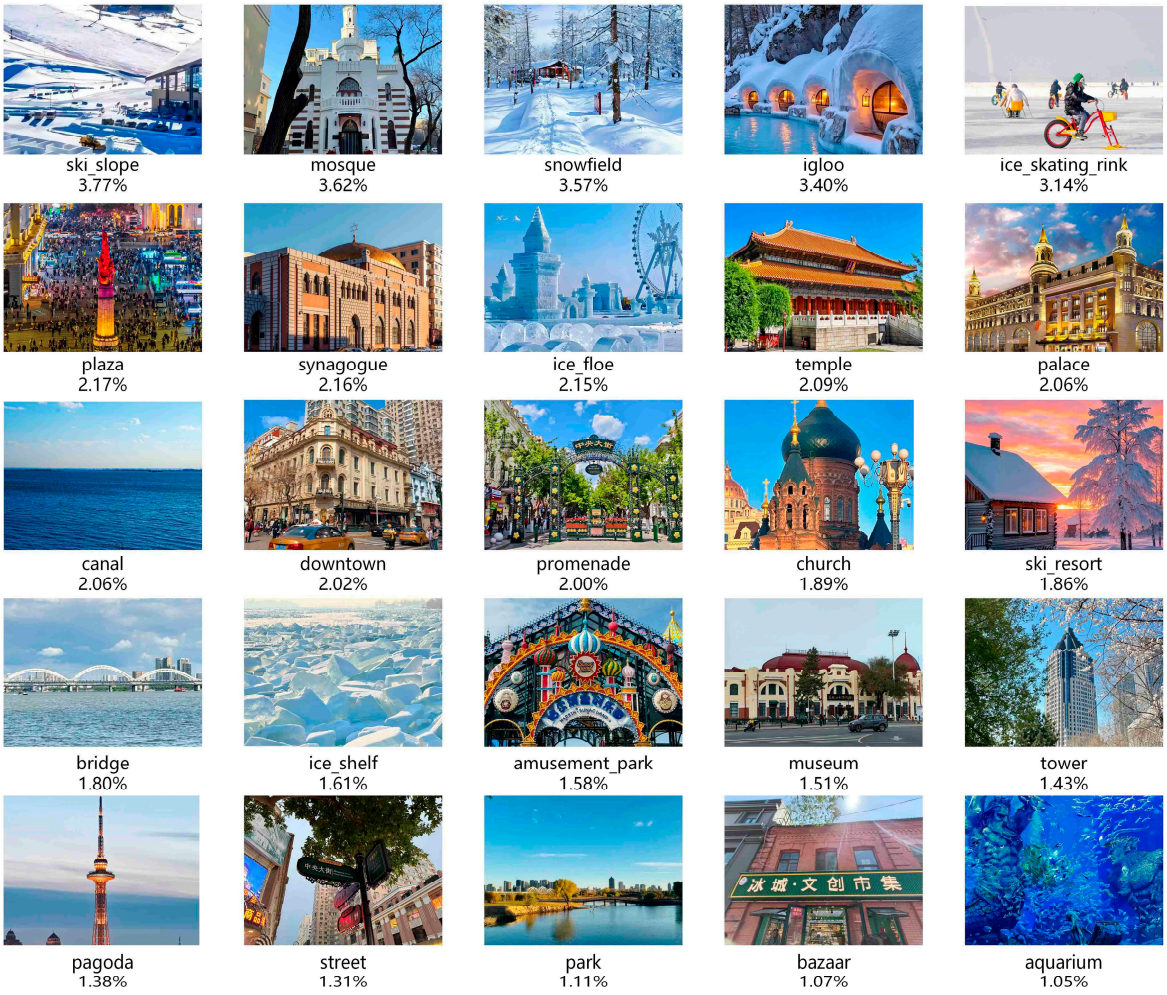


Figure 5. Schematic diagram of TOP25 scenes.

This study quantitatively analyzed scene types present in images shared by tourists. Statistical results show that: (1) among the top 20 most frequently occurring scene categories, those characterized by winter ice-and-snow themes accounted for 35%; (2) within the full set of 365 scene categories defined by the Places365 model, only 13 categories contained distinct ice-and-snow visual elements, representing merely 3.62% of the total category count. However, the combined occurrence frequency of these 13 ice-and-snow related categories constituted 12.02% of all images in the dataset.

These findings indicate that while the spectrum of visual scenes perceived by tourists in Harbin is highly diverse overall, ice-and-snow themed scenes are concentrated in specific categories and exhibit extremely high visibility. Although their absolute number of categories is limited, they form the core of the high-frequency scene set and account for a substantial proportion of the images actually captured and shared by visitors. Therefore, ice-and-snow scenes constitute a primary, highly salient scene type within Harbin's

destination visual image, playing a prominent role in shaping tourists' spatial perception and experience.

4.1.2. Affective Image Evaluation

Sentiment evaluation was calculated using the SnowNLP algorithm combined with Russell's Affective Loop model. Among the 89,375 comments, 53,046 (59.35%) were pleasure-high arousal, 13,148 (14.71%) were pleasure–low arousal, and 7149 (8.00%) were displeasure–low arousal. Finally, 16,032 (17.94%) were displeasure–high arousal. To deeply analyze the emotional experience of tourists regarding the tourist destination, we measured the occurrence frequency of different types of adjectives, and constructed a word cloud map of four emotional characteristics of tourists (Figure 6).



Figure 6. Sentiment Word Cloud.

4.2. Spatial Representation

4.2.1. Cognitive Image Spatial Representation

Through the quantitative processing of information, quantitative data were spatialized. Among the 168 research units, there were a total of 37,873 scene perceptions with weights greater than 0.1. Among them, history and architecture accounted for 39.23%, humanistic life for 22.06%, nature for 12.60%, vacation for 12.23%, and infrastructure for 13.88%. The spatial distribution of various indicators presents five significantly different structural patterns, revealing the spatial characteristics of the distribution of perception scenarios (Figure 7).

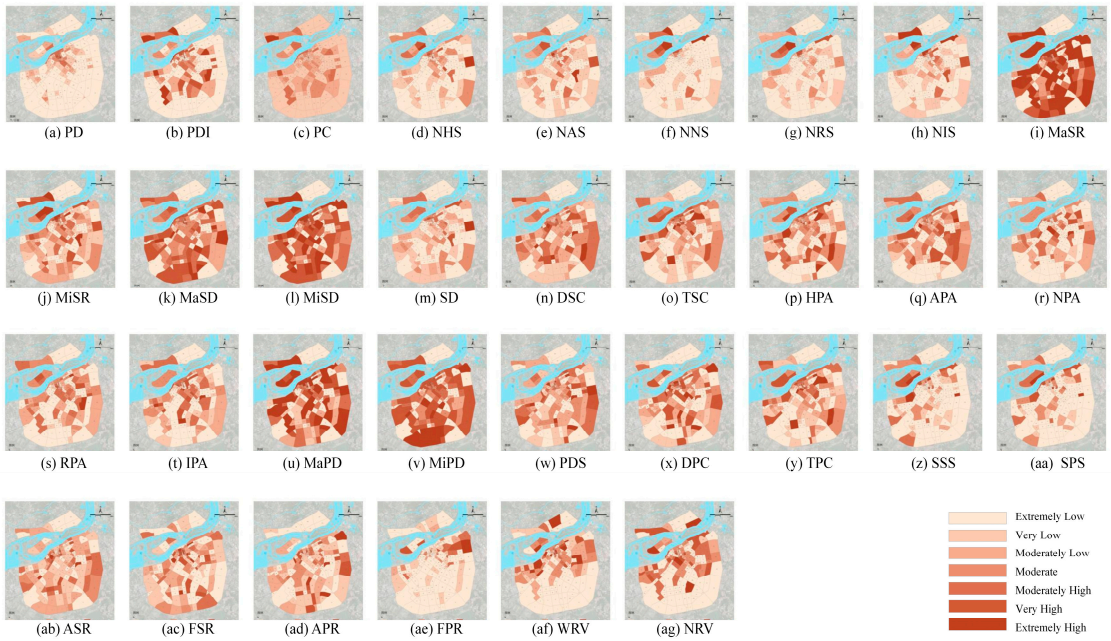


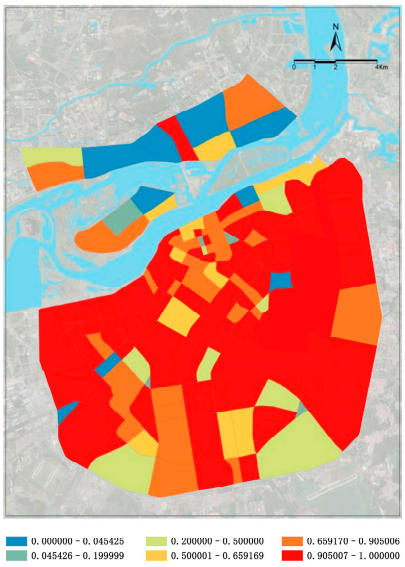
Figure 7. Schematic diagram of scene feature space.

According to the Quantitative Framework for Cognitive Image, the Spatial Representation of Cognitive Image reveals five distinct patterns. PD, NAS, TSC and NRV present a core-edge structure, and element density decreases exponentially from the urban center to the periphery. NHS, NNS, NRS, PDS, SSS, SPS, APR, FPR and WRV present a local central type. This type of structure exhibits a significant differentiation between the activity center and the urban geographical center. PDI, NIS, SD, DSC and ASR present a multi-center balanced layout. PC, HPA, NPA, RPA, IPA, TPC and FSR are presented in the form of strip corridors. These elements extend along linear spaces such as major traffic arteries or waterfront areas. MaSR, MiSR, MaSD, MiSD, APA, MaPD, MiPD and DPC exhibit distributed characteristics. Diverse scenes and perceptual elements are not evenly distributed but are scattered throughout the city according to their specific needs, forming fragmented and collaged spatial forms similar to cells.

4.2.2. Affective Image Spatial Representation

Through spatial clustering analysis of multi-source sentiment computing data, this study revealed that there were significant spatial heterogeneity characteristics in sentiment perception in Harbin, which were specifically reflected in two aspects: the spatial differentiation of the level of Pleasure and the circle structure of the level of Arousal.

At the level of Pleasure, positive and negative emotions show obvious spatial separation and multi-center agglomeration patterns (Figure 8). Positive emotions are manifested as a "dual-core" spatial structure, one on the south bank of the Songhua River, and another around the university town. Firstly, a continuous distribution corridor is formed along the riverside landscape belt on the south bank of the Songhua River. The proportion of positive emotions in this area was 92.31%, demonstrating the significant role that waterfront spaces played in carrying positive emotions. Secondly, a radiation circle is formed with the university town as the core. The proportion of positive emotions in this area was as high as 90%, of which the proportion of high-intensity positive emotion units was 80%, demonstrating the unique emotional vitality of the education and research area. In contrast, negative emotions are concentrated on the north bank of the Songhua River. In the analysis units on the north bank, the proportion of negative emotion units reached 50%, reflecting the emotional differences between the two banks of the Songhua River.



Emotion	Number	Proportion	Type	Number	Proportion
positive	143	85.12%	Ecstasy	100	59.52%
			Joy	25	14.88%
			Affection	18	10.71%
negative	25	14.88%	Disgust	6	3.57%
			Sadness	8	4.76%
			Anger	11	6.55%
total	168	100.00%		168	100.00%

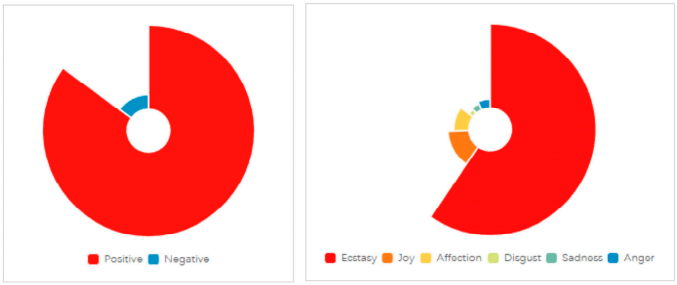
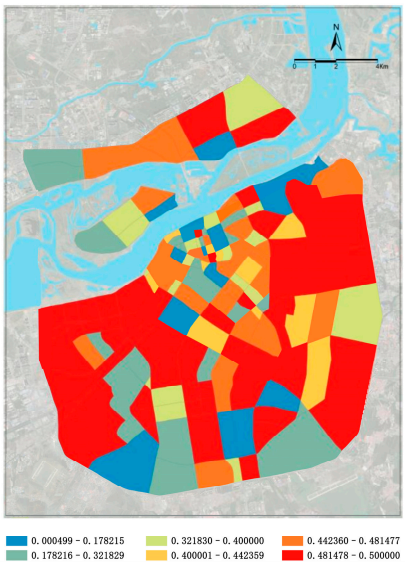


Figure 8. Spatial distribution characteristics map of pleasure.

At the level of Arousal, the research area presents a clear three-circle structure of "core–transition–periphery" (Figure 9). The core area of the city (within the First Ring Road) shows a distinct Low Arousal feature. The proportion of Low Arousal Unit was 35%, and the proportion of Median Arousal Unit was 15%. The urban transition zone from the First Ring Road to the Second Ring Road shows a significant High Arousal characteristic. The proportion of Low Arousal Unit dropped to 30.6%. This shows the emotional stimulation potential of this area as an active zone for urban development. The peripheral area from the Second Ring Road to the Third Ring Road forms a sub-High Arousal Belt. The Low Arousal Unit accounted for approximately 50.9%, revealing the emotional vitality characteristics of emerging areas during urban expansion.



Emotion	Number	Proportion	Type	Number	Proportion
high arousal	147	87.50%	I	99	58.93%
			II	12	7.14%
			III	36	21.43%
low arousal	21	12.50%	I	4	2.38%
			II	8	4.76%
			III	9	5.36%
total	168	100.00%		168	100.00%

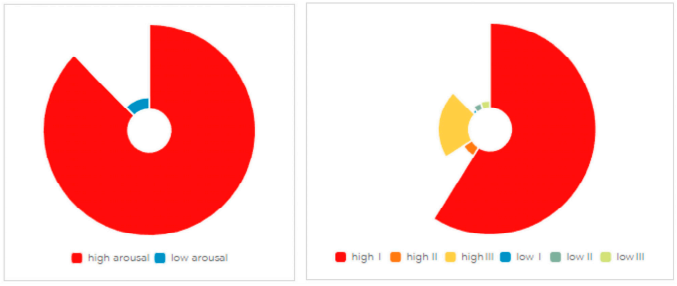


Figure 9. Spatial distribution characteristics map of arousal.

4.3. Heterogeneity of Image and Affective

4.3.1. MGWR Model Results and Descriptive Statistics

Since both perceptual images and emotional elements exhibited random distribution characteristics in space, the research assumed that some configuration features of scene

elements might be related to emotions. This inference needed to be further verified through spatial analysis methods such as MGWR.

To avoid deviations in estimation results due to the mutual influence of various factors, it was necessary to test each factor and select the optimal one as a variable to construct the MGWR model. In this study, the Pearson method was used to verify factor significance, and linear regression analysis was employed to screen for factor collinearity. Of these, four factors including SD failed the significance test. Nine factors, including PDI, were removed from the Pleasure influence factors due to excessive collinearity. Similarly, nine factors were excluded from the Arousal influencing factors. Finally, 20 explanatory variables were selected at the level of Pleasure and Arousal respectively, and sentiment value was taken as the dependent variable for collinearity analysis. The results showed that the variance inflation factor (VIF) values of the 20 explanatory variables were all less than 7.5. This indicated that there was no multicollinearity problem among the variables.

Through the MGWR method, it was determined that if the *p*-value was less than 0.1, it proved a significant spatial-level association between a given scene element and the level of Pleasure and Arousal (Table 3).

Table 3. MGWR coefficients.

Variable	Pleasure	Arousal
PD		0.010
NHS	0.090	
MaSD		0.065
HPA	0.045	
RPA	0.045	
PDS	0.008	
WRV	0.049	0.002

4.3.2. Spatial Analysis of MGWR

After spatial calculation, it was found that there was a significant spatial correlation between the five indicators of NHS, HPA, RPA, PDS and WRV and Pleasure. Three indicators of PD, MaSD, and WRV showed significant spatial characteristics related to Arousal.

(1) The Level of Pleasure

Five indicators showed a significant correlation with Pleasure:

Firstly, the association between NHS and the level of Pleasure shows a significant east–west spatial divergence (Figure 10a). This pattern is closely related to the spatially uneven implementation of historical preservation policies in Harbin. The east–west gradient in tourists' pleasure levels is systematically explained by how renewal paths (redevelopment, isolation, organic repair) shape the material integrity of historical spaces and thus the depth of cultural immersion.

Secondly, across Harbin's overall space, HPA is positively correlated with Pleasure (Figure 10b). This positive correlation is stronger in the southwest direction. Some blocks in the southwestern part of Nangang District are rich in well-preserved and highly attractive historical elements, and their mature tourism development jointly enhances the mechanism that positively converts historical attention into pleasurable experiences.

Thirdly, RPA shows a statistically weak correlation with Pleasure, indicating that the visual saliency of resort scenes has an emotional impact on tourists but is not a decisive spatial factor (Figure 10c). In Harbin's low-temperature environment, the visual appeal of resort scenes is diminished, making it difficult for them to become the dominant force driving emotional experiences.

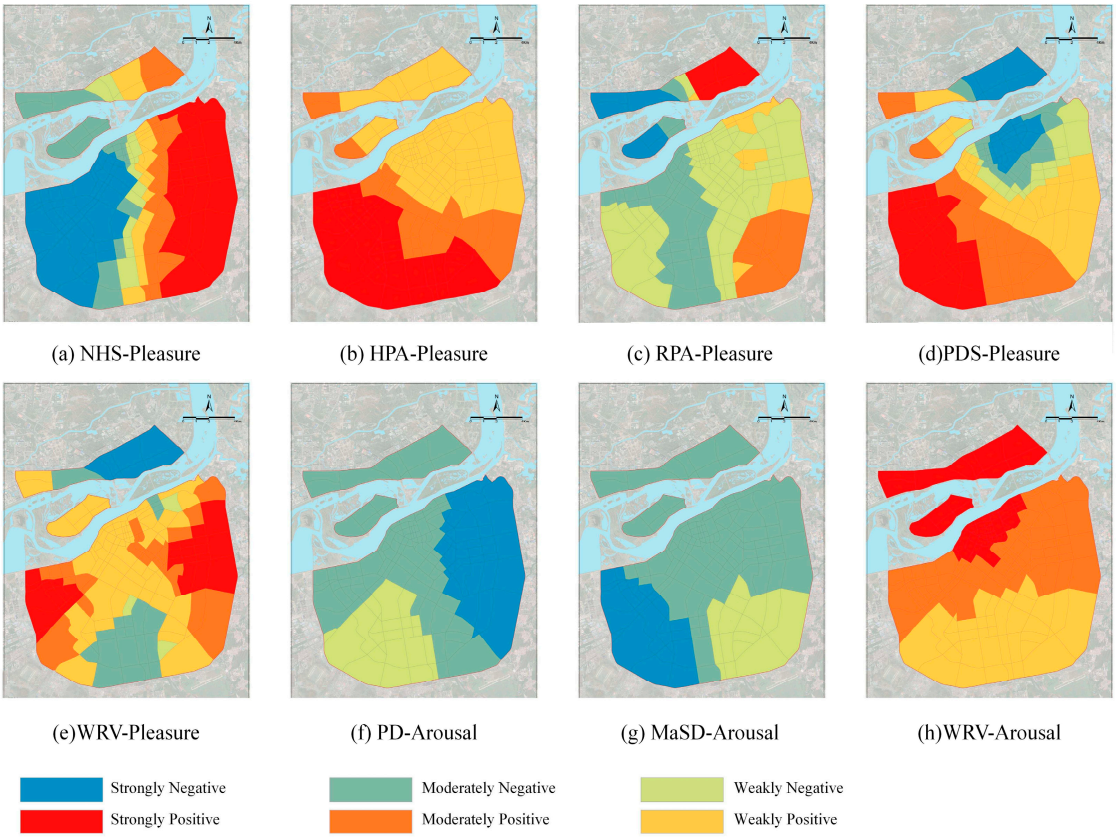


Figure 10. Coefficient map of spatial regression analysis.

Fourthly, the relationship between PDS and Pleasure shows spatial heterogeneity: a negative correlation in the northeast region, but a positive correlation in the southwest region (Figure 10d). The scarcity of visual attractions on the northeast side fails to support positive pleasurable experiences. In contrast, the urban space on the southwest side shows a positive correlation due to a relatively higher density of visual attractions per unit area, which can trigger emotional pleasure.

Fifthly, WRV and Pleasure are relatively weakly correlated in the central region, showing a positive correlation on the east and west sides, and a negative correlation on the north and south sides (Figure 10e). The east–west living axis integrates daily consumption with ice–snow experiences, fostering tourists' tendency to post positive reviews. Although the northeastern tourist area possesses excellent ice–snow resources, insufficient development means there are few reviews.

(2) The Level of Arousal

Three indicators showed a significant correlation with Arousal.

Firstly, PD is globally negatively correlated with Arousal, revealing the inhibitory effect of high-density facility layouts on perceived arousal (Figure 10f). Within Harbin's limited and cold outdoor activity spaces in winter, excessively dense POIs may induce aesthetic fatigue and choice pressure among tourists, thereby weakening the focused intensity of environmental stimulation on emotions.

Secondly, MaSD is globally negatively correlated with Arousal, indicating that over-homogenized scene configurations may inhibit emotional arousal (Figure 10g). The underlying mechanism lies in the fact that excessively homogenized scenes lead to attentional dispersion and a lack of focus on a dominant scene, making it difficult for tourists to form deep cognitive connections and emotional resonance. Moreover, homogeneous competi-

tion among scene types further weakens the hierarchy of emotional stimulation, which is particularly evident in areas dominated by a single ice-and-snow theme.

Thirdly, the spatial association between WRV and Arousal exhibits a distinct "stronger in the north, weaker in the south" pattern (Figure 10h), which profoundly reflects the polarization effect of winter tourism activities along the Songhua River waterfront in Harbin.

5. Discussion and Conclusions

5.1. Discussion

This study was the first to systematically spatialize the perceived image of tourists in Harbin, and found that it presented a clear geographic structure of "multiple composites and emotional differentiation". First, the high-frequency and high-emotion coupling of ice and snow scenes confirmed that they are not only visual symbols, but also core carriers of positive emotions and meanings, which strengthens the viewpoints of the TDI theory about iconic elements shaping overall perceptions [18,22]. Second, spatial representations of the five types of scenes were not homogeneously distributed, but formed an agglomeration and transition zone, with the Songhua River and the core scenic area as the axes, and this structure echoes the basic form of urban tourism flow and functional zoning [25]. Finally, the Affective Map presented some features that intuitively revealed that the emotional response is not randomly distributed, but is closely related to the urban vein [24]. This pattern suggests that the perceived image of the city is a geographic map with an inherent spatial logic and emotional gradient, rather than a simple collection of elements.

A complex pattern of spatial differentiation was found between pleasantness and arousal. Traditional studies have used global models (OLS) or models that can only handle spatial dependence (SLM/SEM) to explore affective influences, and their implicit assumption of "global homogeneity" makes it difficult to explain such complex spatial patterns [43]. This study introduced the MGWR model, whose core strength is its ability to deal with "spatial non-stationarity" [31], which fits perfectly with the theoretical presuppositions of this study, which emphasize "locality" under the framework of human–place relations.

The findings confirmed that significant spatial heterogeneity and scale differences do exist in the effects of different scene elements on affect. For example, the positive effect of NHS on pleasantness was stronger in the traditional neighborhoods in the east, whereas the boosting effect of WRV on arousal was most significant in the tourism polar core area on the north bank of the Songhua River. This finding went beyond studies that only describe the spatial patterns of emotion [23,24] and revealed differentiated formation mechanisms behind the patterns. This suggests that affective responses are not globally driven by some fixed factors, but are the result of a combination of scene-specific elements and localized geographic factors such as urban function, historical heritage, and the intensity of tourism activities. This provides a refined empirical basis for understanding the dynamics and locality of the TDI.

5.2. Conclusions

Through a multi-dimensional analysis of the tourism scenes in Harbin, this study systematically revealed the internal mechanism and spatial effect of the perception of TDI. The research results show the following:

(1) Ice-and-snow scenes constitute a primary, highly salient scene type within Harbin's destination visual image, playing a prominent role in shaping tourists' spatial perception and experience.
(2) The spatial representation of scene types was classified into five types (core-edge structure, local central, multi-center, strip corridors and distributed characteristics). This study found that the emotional geography of Harbin showed two key character-

istics: the spatial differentiation of the level of Pleasure and the circle structure of the level of Arousal.

(3) Through MGWR, it was found that, in terms of Pleasure, there was a significant spatial correlation between the five indicators of NHS, HPA, RPA, PDS and WRV and urban sentiment. In terms of Arousal, there were significant spatial characteristics for the indicators PD, MaSD, and WRV.

This research not only deepens the theoretical understanding of the urban tourism perception mechanism, but also provides an empirical framework for the transformation of cold-region cities into ice-and-snow tourism destinations, which has significant practical value for optimizing the man–land relationship in cold-region cities.

5.3. Theoretical Contributions, Practical Implications and Limitations

This study makes two contributions. First, it broke through a traditional deconstruction paradigm of physical space elements in urban research by incorporating a relatively abstract dimension of "perceived image", and verified the theoretical status of the "perceived image" as the core link of TDI. Secondly, a coupling mechanism between spatial features of urban scenes and tourists' emotional response was innovatively established, and the internal cognitive–emotional linkage mechanism and its spatial heterogeneity were finely portrayed through the cross-validation of spatial-landing and emotional-calculation models.

This study provides valuable geospatial insights for destination management in tourism cities. Managers can implement differentiated scenario optimization and marketing strategies based on the types of emotional dominance and influence mechanisms in different regions, thus translating the concept of TDI into a concrete spatial planning and experience design.

Although this study provided an innovative explanation for the impact of scenes in destinations on emotions, there are still limitations that need to be overcome. While this study employs MGWR to reveal spatially heterogeneous correlations between perceptual elements and tourist emotions, it does not establish causal relationships. Correlation alone cannot verify whether specific scene characteristics directly elicit emotional responses—a limitation for translating findings into planning practice, where causal evidence is often essential for designing effective interventions. Future research should adopt causal inference approaches to examine the directional pathways and potential feedback loops within the perception–affect system. This will help support the optimization of tourism experiences, the enhancement of urban spatial quality, the promotion of sustainable urban development, and the advancement of interdisciplinary research.

Author Contributions: Conceptualization, Xu Lu; methodology, Shan Huang; software, Shan Huang; validation, Jinghua Zhang; formal analysis, Shan Huang; investigation, Shan Huang; resources, Xu Lu; data curation, Jingqun Lu; writing—original draft preparation, Shan Huang; writing—review and editing, Shan Huang; visualization, Jingqun Lu; supervision, Xu Lu; project administration, Xu Lu; funding acquisition, Xu Lu. All authors have read and agreed to the published version of the manuscript.

Funding: Basic Research Projects for Higher Education Institutions by the Education Department of Liaoning Province (Grant No. LJ222510153004). Doctoral Research Startup Project under the 2025 Natural Science Foundation Program of Liaoning Province (Grant No. 2025-BS-0881).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: Authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Abbreviations

PDI	POI Diversity
PC	POI Count
NHS	Number of Historical Scenes
NAS	Number of Human Activity Scenes
NNS	Number of Natural Scenes
NRS	Number of Resort Scenes
NIS	Number of Infrastructure Scenes
MaSR	Macro-category Scene Richness
MiSR	Micro-category Scene Richness
MaSD	Macro-category Scene Diversity
MiSD	Micro-category Scene Diversity
SD	Scene Density
DSC	Dominant Scene Concentration
TSC	TOP25 Scene Concentration
HPA	Historical Perception Abundance
APA	Human Activity Perception Abundance
NPA	Natural Perception Abundance
RPA	Resort Perception Abundance
V	Infrastructure Perception Abundance
MaPD	Macro-category Perception Diversity
MiPD	Micro-category Perception Diversity
PDS	Perception Density
DPC	Dominant Perception Concentration
TPC	TOP25 Perception Concentration
SSS	Snow Scene Salience
SPS	Snow Perception Salience
ASR	Average Scene Richness
FSR	Featured Scene Richness
APR	Average Perception Richness
FPR	Featured Perception Richness
WRV	Winter Review Volume
NRV	Non-winter Review Volume
MGWR	Multi-scale Geographically Weighted Regression
UGC	User-Generated Content
OLS	Ordinary Least Squares
SLM	Spatial Lag Models
SEM	Spatial Error Models
VIF	Variance Inflation Factor
TDI	Tourism Destination Image

References

1. Zhang, D.; Peng, J.; Yang, Z. Identifying urban agglomeration's range: Integrating clustering and accessibility analysis. *Cities* **2026**, *171*, 106708. [CrossRef]
2. Xia, C.; Yeh, A.G.-O.; Zhang, A. Analyzing spatial relationships between urban land use intensity and urban vitality at street block level: A case study of five Chinese megacities. *Landsc. Urban Plan.* **2020**, *193*, 103669. [CrossRef]
3. Li, M.; Liu, J.; Lin, Y.; Xiao, L.; Zhou, J. Revitalizing historic districts: Identifying built environment predictors for street vibrancy based on urban sensor data. *Cities* **2021**, *117*, 103305. [CrossRef]
4. Li, Y.; Yabuki, N.; Fukuda, T. Exploring the association between street built environment and street vitality using deep learning methods. *Sustain. Cities Soc.* **2022**, *79*, 103656. [CrossRef]
5. Wu, C.; Ye, Y.; Gao, F.; Ye, X. Using street view images to examine the association between human perceptions of locale and urban vitality in Shenzhen, China. *Sustain. Cities Soc.* **2023**, *88*, 104291. [CrossRef]

6. He, Z.; Zhang, X. Towards more reliable measures for "perceived urban diversity" using point of interest (POI) and geo-tagged photos. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 91. [CrossRef]
7. Lynch, K. *The Image of the City*; MIT Press: Cambridge, MA, USA, 1960.
8. Zhou, X.; Cui, L.; Chen, S.; Tan, Z. Analysis on the visual salience of image elements of historical blocks based on eye-tracking: A case study of Yongqingfang, Guangzhou. *Mod. Urban Res.* **2024**, *10*, 101–107.
9. Wang, X.; Mou, N.; Zhu, S.; Yang, T.; Zhang, X.; Zhang, Y. How to perceive tourism destination image? A visual content analysis based on inbound tourists' photos. *J. Destin. Mark. Manag.* **2024**, *33*, 100923. [CrossRef]
10. Wu, C.; Wilkes, R.; Silver, D.; Clark, T.N. Current debates in urban theory from a scale perspective: Introducing a scenes approach. *Urban Stud.* **2019**, *56*, 1487–1497. [CrossRef]
11. Wang, C.; Chen, Y.; Zheng, S.; Yuan, Y.; Wang, S. Research on generating an indoor landmark salience model for self-location and spatial orientation from eye-tracking data. *Int. J. Geo-Inf.* **2020**, *9*, 97. [CrossRef]
12. Iwanczak, B.; Lewicka, M. Affective map of Warsaw: Testing Alexander's pattern language theory in an urban landscape. *Landsc. Urban Plan.* **2020**, *204*, 103910. [CrossRef]
13. Meenar, M.R.; Mandarano, L.A. Using photovoice and emotional maps to understand transitional urban neighborhoods. *Cities* **2021**, *118*, 103353. [CrossRef]
14. Ye, H.; Tussyadiah, I.P. Destination visual image and expectation of experiences. *J. Travel Tour. Mark.* **2011**, *28*, 129–144. [CrossRef]
15. Rani, S.; Kumar, P. Deep learning based sentiment analysis using convolution neural network. *Arab. J. Sci. Eng.* **2019**, *44*, 3305–3314. [CrossRef]
16. Agustí, D.P.; Rutllant, J.; Fortea, J.L. Differences in the perception of urban space via mental maps and heart rate variation (HRV). *Appl. Geogr.* **2019**, *112*, 102084. [CrossRef]
17. Mehrabian, A.; Russell, J.A. The basic emotional impact of environments. *Percept. Mot. Ski.* **1974**, *38*, 283–301. [CrossRef]
18. Hunt, J.D. Image as a factor in tourism development. *J. Travel Res.* **1975**, *13*, 1–7. [CrossRef]
19. Crompton, J.L. An assessment of the image of Mexico as a vacation destination and the influence of geographical location upon that image. *J. Travel Res.* **1979**, *18*, 18–23. [CrossRef]
20. Echtner, C.M.; Ritchie, J.R.B. The measurement of destination image: An empirical assessment. *J. Travel Res.* **1993**, *31*, 3–13. [CrossRef]
21. Gartner, W.C. Image formation process. *J. Travel Tour. Mark.* **1994**, *2*, 191–216. [CrossRef]
22. Baloglu, S.; McCleary, K.W. A model of destination image formation. *Ann. Tour. Res.* **1999**, *26*, 868–897. [CrossRef]
23. Xiao, X.; Fang, C.; Lin, H. Characterizing tourism destination image using photos' visual content. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 730. [CrossRef]
24. Tan, J.; Li, Y. Influence of the perceptions of amenities on consumer emotions in urban consumption spaces. *PLoS ONE* **2024**, *19*, e0304203. [CrossRef]
25. Russell, J.A. A circumplex model of affect. *J. Pers. Soc. Psychol.* **1980**, *39*, 1161–1178. [CrossRef]
26. Xu, L.; Zhou, B.; Yu, H.; Zhang, P. Research on the influence of 2022 Winter Olympic Games on the tourism destination image of Zhangjiakou: Based on UGC text analysis. *Geogr. Res.* **2023**, *42*, 422–439.
27. Li, C.; Cao, M.; Wen, X.; Zhu, H.; Liu, S.; Zhang, X.; Zhu, M. MDIVis: Visual analytics of multiple destination images on tourism user generated content. *Vis. Inform.* **2022**, *6*, 1–10. [CrossRef]
28. Tang, F.; Yang, J.; Wang, Y.; Ge, Q. Analysis of the image of global glacier tourism destinations from the perspective of tourists. *Land* **2022**, *11*, 1853. [CrossRef]
29. Lin, M.S.; Liang, Y.; Xue, X.J.; Pan, B.; Schroeder, A. Destination image through social media analytics and survey method. *Int. J. Contemp. Hosp. Manag.* **2021**, *33*, 2219–2238. [CrossRef]
30. Zhang, Q.; Xu, H. Understanding aesthetic experiences in nature-based tourism: The important role of tourists' literary associations. *J. Destin. Mark. Manag.* **2019**, *16*, 100429. [CrossRef]
31. Fotheringham, A.S.; Yang, W.; Kang, W. Multiscale geographically weighted regression (MGWR). *Ann. Am. Assoc. Geogr.* **2017**, *107*, 1247–1265. [CrossRef]
32. Jiang, B.; Zhang, C.; Cui, Y.; Zhu, J.L.; Liu, Z. Enhancement of Harbin ice and snow tourism destination competitiveness: A large-scale data study based on sentiment analysis and Latent Dirichlet Allocation. *PLoS ONE* **2025**, *20*, e0319435. [CrossRef]
33. Zhi, M.; Jia, F.; Tian, J.; Lin, D. Review of the research progress in extracting the landmark based on salience model. *Sci. Surv. Mapp.* **2017**, *42*, 48–54.
34. Xiao, X.; Fang, C.; Lin, H.; Chen, J. A framework for quantitative analysis and differentiated marketing of tourism destination image based on visual content of photos. *Tour. Manag.* **2022**, *93*, 104585. [CrossRef]
35. Jiang, Y.; Yu, R.; Cao, S.; Shan, X. A study on urban emotional map and spatial improvement based on social networking data: Take Chengdu City as an example. *Planners* **2023**, *39*, 56–62.

36. Jou, B.; Chen, T.; Pappas, N.; Redi, M.; Topkara, M.; Chang, S.F. Visual affect around the world: A large-scale multilingual visual sentiment ontology. In *MM'15: Proceedings of the 23rd ACM International Conference on Multimedia*; ACM: Brisbane, Australia, 2015; pp. 159–168. [CrossRef]
37. Zhao, R.; Grosky, W.I. Narrowing the semantic gap-improved text-based web document retrieval using visual features. *IEEE Trans. Multimed.* **2002**, *4*, 189–200. [CrossRef]
38. Deng, N.; Li, X.R. Feeling a destination through the "right" photos: A machine learning model for DMOs' photo selection. *Tour. Manag.* **2018**, *65*, 267–278. [CrossRef]
39. Arshed, M.A.; Mumtaz, S.; Liaqat, M.S.; Ul Haq, I.; Hussain, M. LSTM based sentiment analysis model to monitor COVID-19 emotion. *VFAST Trans. Softw. Eng.* **2022**, *10*, 70–78. [CrossRef]
40. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*; Association for Computational Linguistics: Minneapolis, MN, USA, 2019; pp. 4171–4186. [CrossRef]
41. He, Z.; Deng, N.; Li, X.; Gu, H. How to "read" a destination from images? Machine learning and network methods for DMOs' image projection and photo evaluation. *J. Travel Res.* **2022**, *61*, 597–619. [CrossRef]
42. Lu, X.; Huang, S.; Xie, W.; Sun, Y. The impact of built environment on urban vitality—A multi-scale geographically weighted regression analysis in the case of Shenyang, China. *Buildings* **2025**, *15*, 2989. [CrossRef]
43. Liu, W.; Li, D.; Meng, Y.; Guo, C. The relationship between emotional perception and high-density built environment based on social media data: Evidence from spatial analyses in Wuhan. *Land* **2024**, *13*, 294. [CrossRef]
44. Clark, T.N. Urban amenities: Lakes, opera, and juice bars: Do they drive development? In *The City as an Entertainment Machine*; Emerald Group Publishing Limited: Bingley, UK, 2003; pp. 103–140. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.