Research Article

Research on the Application of Deep Learning Technology Oriented to the Construction and Innovation of Smart City Image Cognition

Dan Shi and Lixin Song

School of Design, Hunan University, Changsha 410000, Hunan, China

Correspondence should be addressed to Dan Shi; shidan0402@hnu.edu.cn and Lixin Song; songlixin0221@hnu.edu.cn

Received 30 November 2021; Accepted 23 December 2021; Published 5 January 2022

Academic Editor: Gengxin Sun

Copyright © 2022 Dan Shi and Lixin Song. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

City image is the observer's subjective impression of the city image. It is an important content of urban geography and planning research and has important guiding significance for shaping a unique urban space. Cognitive research on traditional urban imagery is mainly by means of questionnaires and image sketches. It has problems such as high cost, low update frequency, and limited data coverage, which cannot meet the needs of quantitative research on smart cities and urban economic development in the information age. With the advent of the era of big data and the development of Internet technology, there are more and more quantitative research results on smart city image cognition with the help of big data and deep learning technology. It will be a feasible way to apply it to urban image research. This article combines the development and transformation of smart cities with the transformation of urban planning and leads to an innovation in the construction of urban image cognition based on urban image, active representation data as the data source, and deep learning as the core technology. The theoretical connotation and cognitive dimension of urban imagery are expanded to establish a cognitive model of urban imagery. The city image is cognitively analyzed from three dimensions: image structure, image type, and image evaluation. Specific cities are taken as examples to verify the applicability and scientificity of the cognitive methods and models, so as to enhance the practicality and applicability of urban imagery in urban planning. At the same time, this research is used to answer the development dilemma of big data, summarize the development trend of big data, and explore the new changes that artificial intelligence brings to urban planning. The experimental results show that the model we designed efficiently evaluates the image of the city and can also effectively recognize the image of the city in the main urban area of Chongqing.

1. Introduction

Image is an abstract feeling and perception, generally refers to the mental image of a certain area or place in people's mind, and is a comprehensive manifestation of people's impressions, opinions, cognition, evaluation, and feelings. A city is a special spatial structure, and many urban elements are readable fragments, which then form a complete form. Urban space is an open and complex giant system, and people of different social classes have different observation perspectives and subjective evaluations. From the garden city, the urban beautification movement to the emerging new urbanism in the United States, everyone is seeking a

balance between social justice, urban efficiency, ecological environment, physical environment, and public psychology. However, the research on the relationship between the urban environment and the public's psychological cognition is still in a blank state. Until the 1960s, the American urban planning theorist Lynch K proposed the concept of urban image: it refers to the strong psychological image of the observer evoked by the city or the environment. At the same time, his book titled "The Image Of The City" has also been recognized as a milestone in the study of urban imagery. The theoretical results in the book "Urban Image" have a profound and lasting influence on urban designers, psychologists, and geographers in various countries, and the

perceptual research methods he advocated have also been widely used [1–5].

Urban space research is usually divided into two research fields: urban physical space and urban social space. In the middle of the 20th century, Western scholars launched research on urban entities, mainly from the spatial distribution, spatial structure, and temporal and spatial evolution of various geographic phenomena (such as population, commerce, and transportation) in the city. With the advancement of humanistic concepts and the arrival of postmodern thoughts, researches related to human geography and urban geography have begun to tend to the discussion of urban social space. This type of research pays more attention to themes such as the characteristics and structure of urban social space, social events, social spatial differentiation, and community issues. At present, the above-mentioned researches have relatively mature theoretical and practical foundations, but they generally rely on survey statistics and economic census data for research. Data acquisition usually takes a lot of time, manpower, and material costs, and it is difficult to ensure the accuracy and comprehensiveness of the research. In addition, the frequency of data update is low, and it is difficult to adapt to the needs of urban business quantitative research and urban economic development in the information age. Especially with the construction and development of smart cities, ordinary analysis methods are no longer suitable for the cognitive construction of urban imagery [6–10].

With the continuous development of Internet technology and information technology, cutting-edge technologies such as cloud computing technology, 3S technology (remote sensing, geography information systems, global positioning systems), and deep learning artificial intelligence have become more and more complete, which has promoted the transformation of smart city research methods from single to diversified. At the same time, big data such as point of interest data, open map data, cell phone signaling data, public transportation card swiping records, and open data from commercial service websites and government departments have jointly promoted the arrival of the big data era and the formation of a new data environment. Deep learning is one of the most efficient data analysis techniques, and the basic logic of its algorithm is to analyze massive amounts of data, automatically obtain rules from it, and use these rules to make judgments and predictions on unknown data. It has huge advantages in research fields such as computer vision, language semantics, and speech analysis. The speech recognition and image recognition commonly used in daily life are also benefited from this. If big data provides new research ideas and data sources for urban research, then deep learning will provide the possibility for more in-depth data mining, a deeper understanding of the city, and the discovery of the underlying laws and characteristics behind the city. Using network photo data to study urban imagery can reflect people's overall perception of the city in a more comprehensive and intuitive way, providing a new perspective and thinking for urban imagery [11–15].

From this point of view, this work combines urban development transformation and urban planning changes

and leads to an innovation in the construction of urban image cognition based on urban image, active representation data as data source, and deep learning as the core technology. The theoretical connotation and cognitive dimension of urban imagery are expanded to establish a cognitive model of urban imagery. The city image is cognitively analyzed from three dimensions: image structure, image type, and image evaluation. A certain city is taken as an example to verify the applicability and scientificity of the cognitive method and model, so as to enhance the practicality and applicability of urban imagery in urban planning. At the same time, this research is used to answer the development dilemma of big data, summarize the development trend of big data, and explore the new changes that artificial intelligence brings to urban planning.

2. Related Work

This article mainly started from the perspective of urban planning and related majors, as well as from the researcher's main research purpose, and summarized the research results of foreign scholars into the five following aspects.

2.1. Research on Structural Urban Image. Structural urban imagery was an early research direction, and it was also the most concerned research content of urban planning and geography, focusing on the structural analysis and characteristic research of the urban space environment. Literature [16] conducted a survey on Boston, Jersey, and Los Angeles. The author found that the image of Boston had a clear structure and distinctive features. However, Jersey was a city lacking features because of its weak integrity and lack of center. The road network of Los Angeles rules lacked recognition. There were many independent architectural landmarks in the city, but there was no strong symbol. Literature [17] took image survey of Guyana city in Venezuela and pushed the urban image system into a more detailed analysis. Literature [18] explored the characteristics, structure, and functions of the three major tourist areas in Paris and gained a more comprehensive understanding of the tourist areas of large-scale multicentral urban areas.

2.2. Evaluative Urban Image Research. Evaluative imagery mainly used psychological investigation and research methods, focusing on how people perceive and evaluate the urban environment. It could be used to guide urban design and environmental transformation, urban spatial layout, reasonable allocation of resources, and so forth. In the research process, more attention was paid to the analysis of the psychological experience attributes of the respondents, so it became the focus of social psychology. Literature [19] focused on exploring residents' environmental evaluation research on residential areas. It drew the evaluation map of urban residents' environment, the preference of residents to choose residential areas, and the main influencing factors, which could be used to guide the environmental design of residential areas. Literature [20] took residents of two cities in the United States, Knoxville and Tennessee, as the survey

subjects, investigated the areas they liked and disliked, and obtained an evaluation map of the city. Literature [21] built a comprehensive tourism behavior model based on the image of the tourist destination and the traveler's perception evaluation to study the influence of all variables on the traveler's behavior.

2.3. Research on the Formation Mechanism of City Image.

The influencing factors of the image or the cause of the image are mainly analyzed, and the inner mechanism of the formation of the image of the city is discussed. Literature [22] found the opposite conclusion and found that residents' perception map of the urban environment would change over time. Literature [23] investigated the perceived image of children from different races in their neighborhoods and cities. The study found that age and gender had little effect on children's image, but ethnicity had a great influence. Literature [24] investigated the disagreement between tourists and residents of Simla. The image of tourists was based on the natural and cultural landscape, while the image of residents was based on their familiarity with the city. Literature [25] took Rotterdam as an example, and it was found that major events in the city were conducive to enhancing the image of the city.

2.4. Urban Image Planning and Construction Research.

How to create or plan a better urban image system was explored. Literature [26] believed that the study of urban imagery could provide relevant data basis for urban planning and design. Literature [27] took the Lambro River Valley as an example, and, starting from the regional scale, proposed a planning method for urban imagery. Literature [28] proposed 10 media strategies that could effectively change the negative image of the city. Literature [29] took Barcelona as an example to explain the method of creating city images.

2.5. Exploring the Research Methods of Urban Imagery.

With the integration of disciplines and the development of technology, new methods were continuously introduced to enrich the research on urban imagery. Literature [30] proposed a method to measure the relationship between the five elements of Kevin Lynch and used the analytic hierarchy process for measurement and analysis. Literature [31] used GIS software to conduct an image map survey of urban imagery of residents and found that residents usually had their own understanding of the boundaries and neighborhood elements of their neighborhood. Literature [32] put forward the applicability of traditional survey techniques in the study of urban image and proposed a survey method that combines visual collage and group interviews. Literature [33] proposed a measurement method of city image and found that residents and tourists had similarities and differences in the perception of the city.

3. Construction of Cognitive Model of City Image

The theoretical foundation and application logic of the current big data-urban planning research are integrated to construct a complete research paradigm, as shown in Figure 1. It includes three basic elements: basic theory, data sources, and technical methods. The three support and restrict each other. At the application level, the theoretical basis is urban imagery, the data source is active representation data, and the technical method is deep learning. The combination of city imagery and deep learning can expand the dimension and breadth of city cognition. The combination of urban intention and active representation data can supplement the data foundation of urban research. The combination of active characterization data and deep learning can provide technical possibilities for deep mining of data.

3.1. Theoretical Basis. The five elements of city image space are signs, nodes, paths, boundaries, and regions. City image space is used as a research paradigm for cognition of urban image and urban spatial structure. The urban image cognition method based on the five elements of space only expresses the material space structure of the urban image and does not involve the natural landscape, cultural life, and other nonmaterial elements of the city. The development of the connotation of urban imagery has long since surpassed the category of pure space. It inevitably contains material and nonmaterial elements. It is a mixed and diverse cognitive collection. Natural landscape, urban culture, and urban life have also become important components of urban imagery. In order to adapt to the new development of urban image connotation, through the expansion of traditional urban spatial image cognition methods, based on deep learning technology and the support of big data, this research proposes a smart city image cognition model under the new data environment. In terms of cognitive dimension, the model innovated traditional cognitive methods and expanded into an urban image cognitive model that includes three cognitive dimensions: image structure, image type, and image evaluation. Image structure cognition uses spatio-temporal data to continue the cognition method based on the five elements of sign, node, path, boundary, and regional space proposed by Lynch. Image type recognition uses a deep learning algorithm to perform image recognition processing on the photo image data in the active characterization data and builds a label-type correspondence table based on the characteristics of the algorithm to represent the type of city image. Image evaluation cognition uses deep learning algorithms to perform cognitive learning on language data to determine the emotions in a specific image environment and to evaluate the specific image. In the process of model cognition, the model includes three steps, namely, data acquisition, model calculation, and planning application. The specific process is illustrated in Figure 2.

FIGURE 1: The research paradigm construction of the cognitive model of city image.

FIGURE 2: Framework of city image cognitive model.

In data acquisition, Internet open API (Application Programming Interface) is the main data source. The data is acquired by setting the characteristic data collection rules, among which active characterization data is the main data of this research. The data types are mainly spatiotemporal information, image photos, language, and text. In model calculations, different technical methods are used according to different data types. The image data adopts image recognition deep learning algorithm to characterize the type of city image. The language data uses natural language deep learning algorithms to represent the evaluation of city intentions. Standardized processing methods are adopted for conventional time-space information and other data to characterize the time-space characteristics of urban imagery. Various cognition results are combined to form a three-dimensional or multidimensional cognition of city images. In the planning application, the urban image recognition model enhances the practicality of traditional urban image and expands the application scope of urban image theory. It can be applied to planning fields such as urban design, urban style planning, urban color planning, urban space quality evaluation, social surveys, and shaping of urban characteristics.

3.2. Data Sources. Sina Weibo data is a typical active characterization data. Generally speaking, a complete microblog contains at least three parts of data content, user attributes, time location, and published content. In addition, Weibo has a huge amount of data, fast update, and easy access. These characteristics have a high match with the complex system of the city. Although Weibo data has been widely used in the fields of urban functional area identification and city scale research in recent years, these studies

only use Weibo's passive characterization data such as time information and location information. Due to technical limitations, the content part of Weibo which users actively publish has not been effectively used. The emergence of deep learning just provides technical support for data mining of Weibo content.

In this study, Weibo data was obtained through the Sina Weibo development platform and used as the data basis for the research. It focuses on the content of photos and images, text content, spatiotemporal information, and user attributes of microblog data. After the processing of deep learning algorithms, the photo image data can objectively reflect the content of the photo. In this way, the scene can be distinguished, and the text content can accurately reflect the emotions expressed by the user in a specific environment. The spatiotemporal information includes longitude, latitude, and release time, which can accurately reflect the distribution of users in the city and the characteristics of time and space. User attributes can reflect the clustering of imagery evaluation sources and the distribution characteristics of different groups of people in urban space. Using these four data sources, the city image can be portrayed in depth from different dimensions, and a three-dimensional or multidimensional city image cognition model can be established.

In this study, the location service dynamic reading interface provided by the Sina Weibo open platform was selected. After obtaining the authorization to use, according to the latitude and longitude coordinate information of the spatial point, the microblog information within a certain period of time and within a certain distance around the point can be obtained. The research first set the collection coordinate point, collection radius, and time range through programming. The collection feature condition is set to only collect microblog data that contains three types of

information: photos, text, and location. Then the compiled program is run to call the interface, and the corresponding data from the Sina Weibo development platform is obtained.

In the specific data collection, the data collection and storage structure of this research are designed according to the research needs and the relevant data provided by the application interface. In addition to the feature information of the above three settings, Weibo data also contains more than 20 items of information such as user nickname, user ID, user gender, user residence, release time, equipment used, number of historical Weibo, and number of friends. This research finally selected eight data items: user ID, gender, address, release time, longitude, latitude, text, and photo ID. These, respectively, constitute four analysis modules of crowd attributes, spatiotemporal data, language data, and image data.

3.3. Technical Method. In order to process the content data of Weibo, the unstructured active characterization data is converted into structured data. This research uses two currently more mature deep learning algorithms: photo image recognition and natural language processing.

$$O_I = \text{IRP}(I), \quad (1)$$

$$O_L = \text{NLP}(L), \quad (2)$$

where I is image, L is language, IRP is image recognition processing, and NLP is natural language processing.

Among them, the photo image recognition uses the ImageNet computer vision challenge champion Microsoft Cognitive Service, and the natural language processing uses the Bosen Chinese semantic open platform, which is the leading Chinese language processing field.

Microsoft has launched an artificial intelligence service platform based on deep learning: Microsoft Cognitive Services. This service integrates a variety of deep learning cognitive services. With these cognitive services, computer programs can see, hear, and speak and can understand and interpret the needs we convey through natural communication. At the same time, these services provide open API interfaces, which can help developers through powerful Internet cloud computing capabilities. With the help of Microsoft Cognitive Services, developers can easily apply Microsoft's Cognitive Services to various fields even without a background in artificial intelligence. In this study, computer vision in the visual API of Microsoft Cognitive Services was selected as the deep learning analysis platform for Weibo photos. Computer vision can extract rich information from images for classification and processing of visual data, and tags, descriptions, and domain-specific modules are used to identify the content in the image and perform a confidence score, and the color mode of the image can also be analyzed. The data returned by the system includes description, tags, image format, image dimensions, clip art type, line drawing type, black & white, adult content, racy content, faces, categories, dominant background color, dominant foreground color, and accent color. The returned data used in this study includes label, image category,

dominant foreground color, dominant background color, and heavy color. Among them, the label and image category can represent the image type of the scene that the photo refers to, and the dominant foreground color, dominant background color, and heavy color can represent the image color of the scene that the photo refers to.

Although Microsoft Cognitive Services also provides natural language analysis services, it currently does not support Chinese-related language analysis. Therefore, this study selected the leading Chinese semantic open platform in China: Bosen Chinese Semantic API. The platform provides parts of speech analysis, entity recognition, dependency grammar, sentiment analysis, news summary, news classification, keyword suggestion, and semantic connection. Among them, the sentiment analysis API contains six different corpus analysis functions: general, automobile, kitchenware, tableware, news, and Weibo. Registered users can easily call the corresponding interface by uploading text information to obtain standardized and formatted data results after analysis. This research chooses Weibo language analysis API in sentiment analysis as the deep learning platform for Weibo text. Weibo language analysis can analyze and distinguish the language logic and lexical characteristics of a text through a deep learning engine with a massive natural corpus database and judge the emotional tendency of the text. The nonnegative index and the negative index are returned, and the sentiment tendency of the text can be judged by the value of the nonnegative index and the negative index. In this way, the user's emotion to the space is judged, and the evaluation of the space image is represented.

3.4. City Image Cognitive System

3.4.1. Image Type. The computer vision deep learning model in Microsoft Cognitive Services can accurately identify the content in the microblog photo image and output the corresponding content tag. These generated tags can be used as a data source to characterize the types of city images. The classification rules for the types of urban image elements are formulated according to the computer vision algorithm in Microsoft Cognitive Services. The algorithm divides image content into four categories: object, abstract, scene, and others. Object includes food, animal, people, building, and plant. The abstract includes net, shape, and texture. The scene includes indoor, dark, and sky. There are 86 different content tags in the algorithm.

According to the expansion of the connotation of urban imagery, this research divides urban imagery into two categories: material elements and nonmaterial elements. Material elements can be seen and expressed concretely, including public spaces, landmark buildings, and natural landscapes. Nonmaterial elements cannot be expressed in a concrete form through cognitive maps, such as history, culture, and life. The model sums it up as cultural life. Then, the 86 items of labels are mapped to the city image classification one-to-one, among which the public space contains labels such as outdoor, vehicle, and city. Logo construction

includes tags such as architecture, building, and tower. Natural landscape includes labels such as nature, plant, and landscape. Civic life includes labels such as sport, art, and meal. Because Weibo photos contain a lot of invalid information, such as facial selfie photos, nonphotographic images, and long microblogs with pictures and texts, it may have a greater impact on the cognitive results. Therefore, it is necessary to clean the Weibo photo data before the model calculation process and then use the cleaned data to recognize the image type. The whole process is described as follows. First, the photo data is uploaded to the Microsoft Cognitive Service interface, and the returned content tags are associated with the photos. Then the photo data that contains the image of the city is filtered out based on the tag data; and, according to the classification rules of urban image elements, these cleaned photo data are identified by image type, and the final recognition result of urban image type is obtained.

3.4.2. Image Evaluation. The Bosen Chinese Semantic Open Platform API uses a deep learning model constructed by vocabulary features, language logic, and corpus. It can accurately judge the emotion conveyed by the language and return the type and degree of emotion expressed in the language in the form of a number. The larger the value, the more positive the emotion, and the smaller the value, the more negative the emotion. Language and semantics not only express the essence of the things described but also express the logical relationships between things such as cause and effect, as well as upper and lower positions. They contain the current emotions of the speaker. Urban space and environment can affect the expression of human emotions and behaviors; that is, the emotions expressed by people in a specific space have a high correlation with the space environment. When the number is small, the correlation is not strong, and the contingency of the conclusion is greater than the correlation. But as the number scale increases, this correlation becomes more apparent. When a large amount of data is used as research support, a high correlation between emotion and space can be established. The comprehensive results of emotions expressed in a large number of microblog languages can be used as the basis for the evaluation of urban imagery. The microblog language data obtained in the research has a one-to-one correspondence with data such as crowd attributes, location information, and photos and images. The overall emotional expression of a specific group of people in a specific space environment can be judged according to the microblog language; and, based on this, the positive or negative evaluation of the image is produced by the space.

Based on this, we can divide the evaluation of city image into three categories according to the numerical value. A score of 0–0.4 is a negative image, indicating that the image of the space has a low evaluation, and it has a negative impact on the crowd's emotional guidance, social interaction, and behavioral activities. A score of 0.4–0.6 is a neutral image, indicating that the image evaluation of the space is average, and it has no positive or negative influence on the emotional

guidance, social interaction, and behavioral activities of the crowd. A score of 0.6–1 is a positive image, indicating that the evaluation of the spatial image is relatively high, and it has a positive impact on the emotional guidance, social interaction, and behavioral activities of the crowd.

3.4.3. Image Structure. Image structure is the structural feature of urban image in the dimension of urban space, which is to analyze the distribution, combination, and change law of image elements in urban space. This research continues the cognitive method of urban image spatial structure. According to the distribution characteristics of the spatial location of microblog data, the five elements of urban image's signs, nodes, paths, borders, and regional space are used as cognitive methods to determine the spatial structure of urban image. Using microblog data as the data source can accurately locate the spatial location of the image elements, and it is more reasonable to analyze the spatial structure of the urban image from the actual urban residents' activities.

In the five elements of urban imagery, signs refer to reference objects used by people to identify directions and regions. In the cognitive model of this study, the sign refers to the area in the city where the image elements are most concentrated. The spatial point guides the overall characteristics of the city image, and the spatial position is usually related to the image node. Nodes in the five elements of traditional urban imagery refer to places where traffic meets and where residents gather. In the urban image cognition model, it refers to the area where the image elements are relatively concentrated, and the spatial activities and the generation of image elements in the nodes are more active. In the five elements of urban imagery, path refers to the road that people pass through. It includes the streets and sidewalks for communication and traffic within the city. In the cognitive model, it refers to the linear space where the image elements are distributed in the urban space, mainly distributed in the spaces such as urban roads and rivers. In the five elements of imagery, the boundary refers to the dividing line or obstacle between the regions. It includes various natural or man-made boundaries, such as rivers, lakes, and railways. In the cognitive model, it is roughly the same as traditional cognition, which refers to the dividing line between the distribution areas of image elements, including various natural or man-made boundary lines, as well as other flexible boundaries. In the five elements of urban imagery, a region refers to an area with specific functions, special culture, and economic attributes. In the cognitive model, it refers to the different areas formed by the relative aggregation of image elements due to differences in urban functions, cultural distribution, and economic background.

3.5. Application of City Image Cognitive Model. At present, most urban designs are carried out based on the existing urban structure. Analysis and research on the current situation and background of the city are a necessary prerequisite for the development of urban design. In terms of methods, the cognitive map in the city image is a combination of urban spatial analysis technology and social survey

methods. It has good applications in the practice of urban design at home and abroad, especially for designers to quickly understand the urban spatial structure and environmental characteristics. As a result, a good urban image represents the orderly structure of the city, the outstanding urban characteristics, and the excellent spatial quality. This is exactly the goal that urban design strives to pursue. Through the cognition and shaping of urban imagery, it is helpful to the correct development and implementation of urban design. The urban image cognition model proposed in this research can represent the characteristics of the crowd, image structure, image type, image evaluation, and image color in the city. This has high application value for the preliminary analysis and research of urban design.

4. Experiment and Discussion

4.1. Research Scope. The main body of this empirical study is the urban image of the core area of Chongqing's main city. The urban space of Chongqing is divided into the main urban area and the central urban area. The planning scope of the main urban area includes Yuzhong District, Dadukou District, Jiangbei District, Nan'an District, Shapingba District, Jiulongpo District, Beibei District, Yubei District, and Banan District. The central city is located between Zhongliang Mountain and Tongluo Mountain. It is the main area of the main city construction and the location of the old city, with an area of 1062 square kilometers. The scope of this study is the core area in the central city of Chongqing. Due to the technical characteristics of the Sina Weibo open platform, the basic method of data collection is centered on the central point. Taking the distance from the center point as the acquisition judgment condition, the range formed by the data will be a circle. Therefore, the research scope of this study is a circular area covering the core area of the central city of Chongqing.

4.2. Data Source. The authors obtained the Weibo data from January 1st to January 31st, 2018, through the open platform of Sina Weibo, which contained the three characteristic pieces of information of photo image, text content, and location information. After data cleaning and screening, a total of 543915 valid microblog pieces of data related to the study of urban imagery were obtained, which were used as the data basis for this study. The user ID, gender, address, release time, longitude, latitude, text, photo ID, and a total of eight data items were selected to generate the final research data according to the data format constructed above.

4.3. Cognitive Analysis of Image Structure. According to the cognitive method of urban image spatial structure, ArcGIS 10.2 is used as the software operating platform. According to the spatial position in the microblog data, each microblog is placed in the urban space in the form of dots. Each point is an image element point that constitutes the image structure. Then the nuclear density tool is used, the longitude and latitude are taken as the characteristic index, the spatial distance is used as the characteristic basis, and the image

structure heat distribution map is generated according to the spatial distribution of the element points to match the urban space according to the heat map. This can intuitively identify the image structure of Chongqing's main urban area. From the overall point of view, the image structure presents a multicenter cluster-like distribution pattern, and there is a clear separation between clusters. This distribution pattern is consistent with the spatial structure of Chongqing's main urban area. From the perspective of the distribution of clusters, according to the five-element cognition method of region, boundary, path, node, and sign, the main urban area of Chongqing has formed a clear image structure system.

From the analysis results, a region with a large degree of image aggregation, a certain scale, and a relatively independent spatial perception range is selected. It can be found that five image regions are formed in the entire generated result, as shown in Table 1.

These areas are mainly concentrated in the central area of each group and important urban traffic stations. Among these areas, the Yuzhong Peninsula image area has the largest scale and the highest concentration of image elements. The second is the Jiangbei imagery area. This distribution feature is basically consistent with the urban spatial structure of Chongqing.

From the perspective of the boundaries of the image area, mountains and rivers have the most obvious separation effect on the image area, as illustrated in Table 2.

In particular, the Jialing River separates the Jiangbei image area from the Yuzhong Peninsula image area, and the Yangtze River separates the Yuzhong Peninsula image area from the Nanping image area. Secondly, the division of the image area by the administrative boundary is more obvious. But the reason is that, due to the uniqueness of Chongqing's topography and urban structure, the boundaries of various administrative divisions are mostly determined based on natural mountains and rivers. Therefore, the function of administrative boundaries still originates from natural mountains and rivers. The impact of roads on the image area is very small, which is only reflected in the northwest section of the inner ring of Chongqing. This is also due to the frequent commuting activities and the rich content of commuting activities, which brings the frequent production of image elements on the road. As a result, the obvious separation effect of roads on space does not exist in the traditional way.

From the perspective of the entire research scope, most of the image paths are roads in the city, and these paths constitute the communication channel space between the various image areas of the main city of Chongqing. In these paths, six main image paths are formed, as shown in Table 3.

These paths are as follows: (1) Jiefangbei-Lianglukou-Daping-Shiqiaopu; (2) Ranjiaba-Xinpaifang-Guanyin Bridge-Lianglukou; (3) Wulidian-Chongqing Yangtze River Bridge-Wanda Plaza-Ertang; (4) Jiangbeizui-Qiansimen Bridge-Dongshuimen Bridge-Shangxin Street; (5) Dashiba-Shimen Bridge-Three Gorges Square-Ciqikou Ancient Town; (6) Three Gorges Square-Southwest Hospital. From a functional point of view, these seven main image paths are all traffic paths and are also the main roads connecting

TABLE 1: Image area.

Image area	Name
1	Yuzhong Peninsula imagery area
2	Jiangbei imagery area
3	Longtou Temple imagery area
4	Nanping imagery area
5	Shapingba imagery area

TABLE 2: Image boundary.

Image boundary	Name
1	Jialing River
2	Yangtze

TABLE 3: Image path.

Image path	Name
1	Jiefangbei-Lianglukou-Daping-Shiqiaopu
2	Ranjiaba-Xinpaifang-Guanyin Bridge-Lianglukou
3	Wulidian-Chongqing Yangtze River Bridge-Wanda Plaza-Ertang
4	Jiangbeizui-Qiansimen Bridge-Dongshuimen Bridge-Shangxin Street
5	Dashiba-Shimen Bridge-Three Gorges Square-Ciqikou Ancient Town
6	Three Gorges Square-Southwest Hospital

various groups in the main urban area of Chongqing. Among them, four paths are cross-river paths.

By subdividing each image area, 30 image nodes are formed. It can be found that there are multiple image nodes in a single image area. The nodes in the Yuzhong Peninsula image area include the Jiefangbei Metropolitan Area, Lianglukou, and Eling Park, and the nodes in the Jiangbei image area include Guanyin Bridge, Ninth Street, and California Garden. The node of the temple image area is Chongqing North Railway Station, the node of the Nanping image area is Wanda Plaza and the Convention and Exhibition Center, the node of the Shapingba image area is the Three Gorges Square, and the node of the Daping-Yangjiaping image area is Times Tianjie and Yangjiaping. There are many independent image nodes distributed outside the 6 image areas.

The image sign is the node where the image elements gather the most frequently, and it plays a characteristic guiding role in the perception of the image space. Comparing 30 image nodes, a total of 8 image signs, the Great Jiefangbei Metropolitan Area, Guanyin Bridge, Chongqing North Railway Station, Three Gorges Square, Ciqikou Ancient Town, Times Tianjie, Yangjiaping, and Wanda Plaza, have been formed.

4.4. Cognitive Analysis of Image Type. According to the urban image type recognition method in the urban image recognition model established above, the image types are divided into four types: landmark buildings, natural landscapes, public spaces, and cultural life. Through the classification and statistics of the image types of the cognitive points, the composition level of the urban image types in the main urban area of Chongqing can be calculated, as

illustrated in Figure 3. LB is landmark buildings, NL is natural landscapes, PS is public spaces, and CL is cultural life.

The images expressed as landmarks accounted for 27.88% of the total scale, those expressed as natural landscape images accounted for 26.57% of the total scale, those expressed as public spaces accounted for 22.34% of the total scale, and those expressed as cultural life accounted for 23.21% of the total scale. From the results, it can be found that the dominant city image in the main urban area of Chongqing is the iconic building, followed by the natural landscape, public space, and cultural life. The image of natural landscape has not become the dominant direction of Chongqing's urban image in the overall cognition level. Among them, the compositions of landmark buildings and natural landscape are relatively close, and the compositions of public spaces and cultural life are relatively close. The proportions of the four image types are relatively balanced, and there is no particularly prominent image type, indicating that Chongqing is not clear enough in terms of the characteristics of urban imagery.

4.5. Cognitive Analysis of Image Evaluation. According to the urban image evaluation cognitive method in the cognitive model, the sentiment tendency analysis of the text in the Weibo data is carried out by calling the API of the Bosen Chinese Semantic Open Platform. The emotion expressed in the Weibo text is used as the basis for evaluating the image of the city. Then the sentiment analysis results are matched with the microblog data one-by-one to realize the one-to-one correspondence between the sentiment analysis results and the text information, location information, image information, and crowd attributes in the microblog data,

FIGURE 3: Image type composition.

FIGURE 4: Image evaluation type.

which is used as the data basis for image evaluation. The organized data are imported into the ArcGIS platform, and image feature points are generated based on the longitude and latitude in the data. The value of sentiment analysis is taken as the characteristic expression of the element point. The points with a value of 0–0.4 indicate a negative evaluation. The point with a value of 0.4–0.6 indicates a neutral evaluation. The points with a value of 0.6–1 indicate positive reviews.

The overall situation of the image evaluation is shown in Figure 4. In all the microblogs of the research data, the average of all sentiment values is 0.67 and the standard deviation is 0.35, indicating that the urban image in the main urban area of Chongqing is neutral and positive in terms of overall tendency. Judging from the positive and negative distribution trends of imagery evaluation, people have strong feelings for Chongqing. The polarization is more obvious, with a large number of poles and a small number of smiles in the middle. Among them, positive reviews accounted for 60.5%, with an average of 0.91,

neutral reviews accounted for 13.9%, with an average of 0.53, and negative reviews accounted for 25.6%, with an average of 0.18.

From the perspective of the classification and evaluation of image types, the natural landscape has the highest evaluation, followed by the public space image; then the iconic building image and the cultural life image have the lowest evaluation. Among them, the natural landscape in the main urban area of Chongqing has the highest evaluation among the four types of urban imagery and has the highest consistency among a large amount of evaluation data. This shows that the natural landscape pattern and features of Chongqing's mountains and rivers have been widely recognized. The consistency of public spaces evaluation is higher than those of landmark buildings and cultural life.

In the five image areas, the ranking is based on the image evaluation value. The results are the Yuzhong Peninsula imagery area, Longtuo Temple imagery area, Nanping imagery area, Jiangbei imagery area, and Shapingba imagery area, as shown in Figure 5.

FIGURE 5: The statistics of city image elevation.

Among them, the Yuzhong Peninsula image area has the highest image evaluation, with positive evaluations accounting for 65.5%, neutral evaluations accounting for 31.7%, and negative evaluations accounting for 2.8%. The Shapingba image area has the lowest image evaluation, with positive evaluations accounting for 42.3%, neutral evaluations accounting for 48.6%, and negative evaluations accounting for 9.1%.

5. Conclusion

This work starts from the three dimensions of the structure, type, and evaluation of smart city imagery and uses spatiotemporal data, image data, and language data as data sources. Then Microsoft Cognitive Services and Bosen Natural Semantics Open Platform are used as deep learning technologies to establish a construction model of urban image cognition to recognize and analyze the image structure, image types, and image evaluation of the urban image in the main urban area of Chongqing. The following can be concluded: First is the successful application of the smart city image recognition model based on deep learning. This paper constructs the basic paradigm of this research with three elements, takes city image as the theoretical basis of the research, uses open data as the data source, and uses deep learning as the technical method to construct a brand-new city image cognition model. The urban image cognition of Chongqing's main urban area is taken as an empirical test to verify the applicability and feasibility of the modified model. Second is the cognition of the image of the city in the main urban area of Chongqing. The image structure is relatively complete, but there are still problems such as the lack of natural landscape elements in the structure and the insufficient connection strength between the image areas. The most prominent city in Chongqing is the iconic architecture, and the image of cultural life occupies the least proportion of the overall image of the city. The overall image

evaluation is neutral to positive, and the image evaluation of natural landscape is the highest. The evaluation of the image core area is significantly higher than that of the outer core area, and the image evaluation of the south coastline of the Yangtze River and Jialing River is significantly higher than the north coastline. In future work, we will devote ourselves to designing more efficient and lightweight models to complete the assessment of urban imagery.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

- [1] R. Ladhari and N. Souiden, "The role of mega-sports event experience and host city experience in explaining enjoyment, city image, and behavioral intentions," *Journal of Travel & Tourism Marketing*, vol. 37, no. 4, pp. 460–478, 2020.
- [2] A. S. Hussein, "City branding and urban tourist revisit intention: the mediation role of city image and visitor satisfaction," *International Journal of Tourism Policy*, vol. 10, no. 3, pp. 262–279, 2020.
- [3] A. Chan, D. Suryadipura, and N. Kostini, "City image: city branding and city identity strategies," *Review of Integrative Business and Economics Research*, vol. 10, pp. 330–341, 2021.
- [4] A. Jia and C. Xu, "Smart city image landscape design based on wireless sensors," *Microprocessors and Microsystems*, vol. 83, Article ID 104022, 2021.
- [5] J. Huang, H. Obracht-Prondzynska, D. Kamrowska-Zaluska, Y. Sun, and L. Li, "The image of the City on social media: a comparative study using "Big Data" and "Small Data" methods in the Tri-City Region in Poland," *Landscape and Urban Planning*, vol. 206, Article ID 103977, 2021.
- [6] F. Scorza and G. Fortunato, "Cyclable cities: building feasible scenario through urban space morphology assessment," *Journal of Urban Planning and Development*, vol. 147, no. 4, Article ID 05021039, 2021.
- [7] J. Stehlin, M. Hodson, and A. McMeekin, "Platform mobilities and the production of urban space: toward a typology of platformization trajectories," *Environment & Planning A: Economy and Space*, vol. 52, no. 7, pp. 1250–1268, 2020.
- [8] S. Yavo-Ayalon, T. Alon-Mozes, and M. Aharon-Gutman, "The shape of theatre in the city: a theoretical and methodological approach for analyzing artistic activity in urban space," *Journal of Urban Affairs*, vol. 42, no. 8, pp. 1202–1221, 2020.
- [9] J. J. Gieseeking, "Mapping lesbian and queer lines of desire: constellations of queer urban space," *Environment and Planning D: Society and Space*, vol. 38, no. 5, pp. 941–960, 2020.
- [10] P. Przewoźna, P. Jankowski, and A. Stach, "Solid waste management in urban space: the volume-weight relationship," *AIMS Environmental Science*, vol. 7, no. 6, pp. 575–588, 2020.

- [11] S. Law, C. I. Seresinhe, Y. Shen, and M. Gutierrez-Roig, "Street-frontage-net: urban image classification using deep convolutional neural networks," *International Journal of Geographical Information Science*, vol. 34, no. 4, pp. 681–707, 2020.
- [12] Z. Pan, J. Xu, Y. Guo, Y. Hu, and G. Wang, "Deep learning segmentation and classification for urban village using a worldview satellite image based on U-Net," *Remote Sensing*, vol. 12, no. 10, Article ID 1574, 2020.
- [13] J. Pei, K. Zhong, J. Li, J. Xu, and X. Wang, "ECNN: evaluating a cluster-neural network model for city innovation capability," *Neural Computing & Applications*, pp. 1–13, 2021.
- [14] R. Cao, W. Tu, C. Yang et al., "Deep learning-based remote and social sensing data fusion for urban region function recognition," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 163, pp. 82–97, 2020.
- [15] A. K. M. M. Rahman, M. Zaber, Q. Cheng et al., "Applying state-of-the-art deep-learning methods to classify urban cities of the developing world," *Sensors*, vol. 21, no. 22, Article ID 7469, 2021.
- [16] K. Lynch, *The Image of the City*, MIT Press, Cambridge, MA, USA, 1960.
- [17] D. Appleyard, "Styles and methods of structuring a city," *Environment and Behavior*, vol. 2, no. 1, pp. 100–107, 1970.
- [18] D. G. Pearce, "Tourist district in Paris: structure and functions," *Tourism Management*, vol. 19, no. 1, pp. 49–65, 1998.
- [19] W. Clark and M. Cadwallader, "Residential preferences: an alternate view of infra urban," *Environment & Planning A*, no. 5, pp. 693–703, 1975.
- [20] J. L. Nasar, "The evaluative image of the city," *Journal of the American Planning Association*, vol. 56, no. 1, pp. 41–53, 1990.
- [21] C.-F. Chen and D. Tsai, "How destination image and evaluative factors affect behavioral intentions," *Tourism Management*, vol. 28, no. 4, pp. 1115–1122, 2007.
- [22] Y. Lee and C. G. Schmidt, "Urban spatial cognition: a case study of Guangzhou, China," *Urban Geography*, vol. 7, no. 5, pp. 397–412, 1986.
- [23] R. Maurer and J. C. Baxter, "Images of the neighborhood and city among Black-, Anglo-, and Mexican-American children," *Environment and Behavior*, vol. 4, no. 4, p. 351, 1972.
- [24] R. S. Jutla, "Visual image of the city: tourists' versus residents' perception of Simla, a hill station in northern India," *Tourism Geographies*, vol. 2, no. 4, pp. 404–420, 2000.
- [25] G. Richards and J. Wilson, "The impact of cultural events on city image: Rotterdam, cultural capital of Europe 2001," *Urban Studies*, vol. 41, no. 10, pp. 1931–1951, 2004.
- [26] Y. Lee and C. G. Schmidt, "Evolution of urban spatial cognition: patterns of change in Guangzhou," *Environment & Planning A: Economy and Space*, vol. 20, no. 3, pp. 339–351, 1988.
- [27] A. Toccolini, N. Fumagalli, and G. Senes, "Greenways planning in Italy: the Lambro River valley greenways system," *Landscape and Urban Planning*, vol. 76, no. 1–4, pp. 98–111, 2006.
- [28] E. Avraham, "Media strategies for improving an unfavorable city image," *Cities*, vol. 21, no. 6, pp. 471–479, 2004.
- [29] A. Smith, "Conceptualizing city image change: the "re-imaging" of Barcelona," *Tourism Geographies*, vol. 7, no. 4, pp. 398–423, 2005.
- [30] R. Banai, "A methodology for the image of the city," *Environment and Planning B: Planning and Design*, vol. 26, no. 1, pp. 133–144, 1999.
- [31] E. Talen and S. Shah, "Neighborhood evaluation using GIS: an exploratory study," *Environment and Behavior*, vol. 39, no. 5, pp. 583–615, 2007.
- [32] P. Laaksonen, M. Laaksonen, P. Borisov, and J. Halkoaho, "Measuring image of a city: a qualitative approach with case example," *Place Branding*, vol. 2, no. 3, pp. 210–219, 2006.
- [33] S. Gilboa, E. D. Jaffe, D. Vianelli, A. Pastore, and R. Herstein, "A summated rating scale for measuring city image," *Cities*, vol. 44, pp. 50–59, 2015.