Tourism Management 75 (2019) 595–608

Contents lists available at ScienceDirect

Tourism Management

journal homepage: www.elsevier.com/locate/tourman

Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing

Kun Zhang, Ye Chen*, Chunlin Li

College of Tourism and Service Management, Nankai University, No.38 Tongyan Road, Haihe Education Park, Jinnan District, Tianjin, 300350, China

ARTICLE INFO

Keywords:
Tourism destination
Tourists' behaviors
Tourists' perception
Visual content analysis
Computer deep learning model

ABSTRACT

Visual content analysis of tourist photos is an effective way to excavate tourist behavior and explore tourists' cognition in the tourism destination. With the development of computer deep learning and big data mining technology, identifying the content of massive numbers of tourist photos by Artificial Intelligence (AI) approaches breaks through the limitations of manual approaches of identifying photos' visual information, e.g. small sample size, complex identification process and results deviation. In this study, 35,356 Flickr tourists' photos in Beijing were identified into 103 scenes by computer deep learning technology. Comparison through statistical analysis for behaviors and perceptions of tourists from different continents and countries was conducted. Tourists' cognitive maps with different perceptual themes were visualized according to photos' geographical information by ArcGIS. The field of how to apply AI technology into tourism destination research was explored and extended by this trial study.

1. Introduction

The human society is stepping into the third information civilization. Information technology is connected to travel and tourism in so many ways and provides possibilities to access a variety of data, in massive quantities, in different formats, and potentially real-time. Information technology in tourism research has already shifted from primarily a marketing-driven tool to a knowledge creation tool (Xiang, 2018).

In such a knowledge-creating era, unconventional, big data sources are emerging, with the potential to improve our knowledge of tourism at unprecedented detail for vast world regions (Batista E Silva et al., 2018), and it is one of the most representative paradigms and represents an emerging investigative field for researchers and practitioners (Vecchio, Mele, Ndou, & Secundo, 2018). By data sources, the tourism-related big data fall into three primary categories: users generated content (UGC) data, including online textual data and online photo data; device data (by devices), including GPS data, mobile roaming data, Bluetooth data, etc.; and transaction data (by operations), including web search data, webpage visiting data, online booking data, etc. (Li, Xu, Tang, Wang, & Li, 2018). Especially, as web users are becoming increasingly enthusiastic about interacting, sharing, and working together through online collaborative media (Hussain & Cambria, 2018) and the mining possibility of big data in social networks are increasing in the era of social media and connectivity (Web 2.0), tourism research based on UGC data has increased rapidly in recent years.

Online textual data and online photo data are the main types of UGC data. For the former e.g., review data and blog data, mainly used for satisfaction studies and tourism recommendation and tourist sentiment (Deng & Li, 2018; Jang & Moutinho, 2019; Liu, Huang, Bao, & Chen, 2019). Compared to textual data, the photos uploaded by tourists involve a wealth of interesting information and are more informative in spatio-temporal attributes, providing a new perspective to study tourists and destinations. Mainly, there are three ways for the applications of online photo data in tourism research: the study of metadata imbedded by photos (such as location, time, etc.), the study of text attached by photos (such as title, description, tag, etc.), and the study of the photos' visual content itself. As limited to the technical threshold of information mining and content analysis, from metadata to text, and then to visual content, the carrying difficulty of the three subjects' study increases gradually. Photo data themselves involve a wealth of interesting information apart from metadata.However, most existing studies conducted analyses indirectly on the metadata or textural data embedded in photos. As a gap, mining the most abundant picture content is less performed, and powerful photo mining techniques acting directly

* Corresponding author.
E-mail addresses: zxzhangkun@163.com (K. Zhang), chenye@nankai.edu.cn (Y. Chen), lichunlinx@163.com (C. Li).

https://doi.org/10.1016/j.tourman.2019.07.002
Received 12 February 2019; Received in revised form 5 May 2019; Accepted 1 July 2019
Available online 11 July 2019
0261-5177/

on images are needed (Li et al., 2018). Facing the challenge of mining the picture content itself, computer vision, and image processing technology in the field of artifical intelligence (AI) have already made much progress. Through computer deep learning, effectively reading massive picture data and visual content recognition is possible. The popularization of Python and other coding techniques make interdisciplinary cooperation between computer and tourism research inevitable. It is time to make a better attempt to adopt the computer vision and image processing technology to identify the content of UGC photos.

With the aims of filling this gap and detecting more valuable information from visual metadata for tourism destination research, this study takes “how to effectively apply computer vision technology into tourists' behaviors and perception analysis in the field of tourism destination study” as the core question. By using photos' visual content as a new form of data, specific objectives are (1) to establish the tourists' photo dataset of 137265 photos of Beijing by data mining in social network Flickr and identify the visual contents of photo into 103 scenes and 11 categories by computer deep learning; (2) compare perception and behavior preference of tourists from different continents and countries and (3) explore the feasibility of using the computer deep learning model in the study of tourists' perception and behavior preference in tourism destinations.

2. Theoretical background

2.1. Photographs as the representation of tourism destination image (TDI)

In the Web 2.0 Era, the “Travel 2.0” phenomenon is catalyzed, which is characterized by a high level of social interaction and exchange of travel-related content between tourists on the Internet (Leung, Lee & Law, 2012). The building of TDI has transformed from the destination marketer's domination (Gunn, 1972) to a dynamic interaction process between tourist and promotion (Gilbert & Hancock, 2006; Hunter, 2013), and then to a new stage, which is primarily shaped by tourists' behavior, perception, experience and their dissemination through various social networks, and gives pressure to the destination management organizations (DMOs) of examining and modifying their projected destination image (PDI). In the research process, the current destination image study emphasizes the qualitative dimensions as the quantitative measurement (Chi, Wu, Morrison, Zhang, & Chen, 2015), and try to figure out a holistic destination image by detecting through several attributes or dimensions. For example, Echtner and Ritchie (1993) used the attribute-holistic framework to form a destination image (Echtner & Ritchie, 1993). Qu, Kim, and Im (2011) pointed out that the overall or holistic image of a destination is formed by the interaction with cognitive, affective, and unique images (Qu, Kim & Im, 2011).

In this sense, photos can be understood as a compression of the destination's image (Pan, Lee & Tsai, 2014a), as the visual images have a more significant impact on people's memories and attitudes than other forms of messages, such as text, sound, and so on (Kim, Kim & Wise, 2014). Photographs stand for tourists' perception, experience and even sentiment (Pan, Lee & Tsai, 2014b) towards tourism destinations. Thus, they serve as a pictorial representation of the cognition of TDI and provide a means to convey or construct a TDI (Hunter, 2016).

2.2. Photographs as a spatial-temporal mirror for tourists' behavior

Recently, GPS chips became a standard component inside many cameras and smartphones, and geographical information can be embedded into the photo's EXIF when taking a photo. Thus, the popular photo-sharing websites enable “geo-tagging” in their tagging interface. Latitude and longitude can be automatically extracted (Jiang, Yin, Wang, & Yu, 2013). These increasing geo-tagged photos provide opportunities in tourism research by providing high spatial and temporal data that make it possible to analyze the spatiotemporal patterns of a large number of tourists (Salas-Olmedo, Moya-Gómez, García-Palomares, & Gutiérrez, 2018) , like photo-takers' trails and movements, and trajectory patterns (Cai, Lee &Lee, 2018). In recent years, there have been numbers of spatiotemporal analyses of tourist behavior by using geographic information of UGC geotagged photos.

According to the overall complexity of the research question and data processing technology, the study of tourist behavior based on geotagged photos could be grouped into three categories. In the first type, geographic coordinate information is the independent and unique analysis data for obtaining spatial distributions characteristics, like urban area of interest (AOI), place of interest (PoI) (Cai, Hio, Bermingham, Lee, & Lee, 2014; García-Palomares, Gutiérrez & Mínguez, 2015; Han & Lee, 2015; Hu et al., 2015; Zhou, Xu & Kimmons, 2015) and geographic preferences (Su, Wan, Hu, & Cai, 2016), aggregated clustering frequently appeared in the methodology of this kind of study. In the second type, both spatial and temporal analysis are conducted using a dataset combination of geographic coordination and time sequence. Tourists behavior, tourist's trajectory, and pattern of tourist's trajectory are the primary research content (Salas-Olmedo, Moya-Gómez, García-Palomares, & Gutiérrez, 2017; Straumann, Çöltekin & Andrienko, 2014; Vu, Li, Law, & Ye, 2015). Based on this research, the further study aimed at travel route recommendation to the tourism spots are carried out with computer depth learning (Chua, Servillo, Marcheggiani, & Moere, 2016; Gavric, Culibrk, Lugonja, Mirkovic, & Crnojevic, 2011; Sun, Fan, Bakillah, & Zipf, 2015). For the third type, more attached metadata with geo-tagged photos are applied to research, semantic tourist's trajectory analysis appears as a more deep-going analysis for the characteristics of tourist behavior and smart itinerary recommendation (Cai, Lee, & Lee, 2018).

In all the three categories, the analytic dimension of behavioral differences between tourists and locals are covered. In addition to the geographic coordinate information, tag, capture, and other textual information describing the photo are usually combined into the spatial analysis as a complement. However, the interpretation of the visual content of the photo itself is rarely involved in any geo-tagged based study.

2.3. Visual content analysis of photos in tourism and related research fields

This paper reviews the literature about visual content analysis of the photo from three research aspects: the data source of photos, study theme and dimension, and methodology for visual content analysis.

Photo source. There are three main ways of photos material selection with chronological order: (1) the convention media resource, e.g., 145 photos published in The New York Times Travel Section “Why We Travel” from 2008 to 2012 were extracted for analysis (Pan, Lee, & Tsai, 2014b); (2) online engine website, online keywords search in the mainstream engine website is an approach for sampling, and it uses the specific research topics-online destination image (Choi, Lehto & Morrison, 2007; Hunter, 2013, 2016; Mak, 2017; Zhou, 2014); (3) UGC photos in the social network, keywords for tags searching still were used for samplings in the data collection, and social network like Flickr instead of online engine website (Stepchenkova & Zhan, 2013). The change in the selection of materials is closely related to the evolution and development of the media and social networks.

Study theme and dimension. How the pictorial materials reflect the destination image is the identified theme, and it unfolds in diverse aspect. First, photos are viewed as a representative semiotic for a destination. Semiotics theory is adopted for interpreting photos' representations of a tourism destination (Nöth, 1990), typical content-semiotic procedures with sampling, content analysis, reflexive semiotic, synecdoche, and synthesis are applied (Hunter, 2013, 2016; Leeuwen, 2003; Mcgregor, 2000). Second, photos that express the travel experience towards the destination. Tags, captions or other textural photo descriptions are the substitution of the photo itself for content analysis

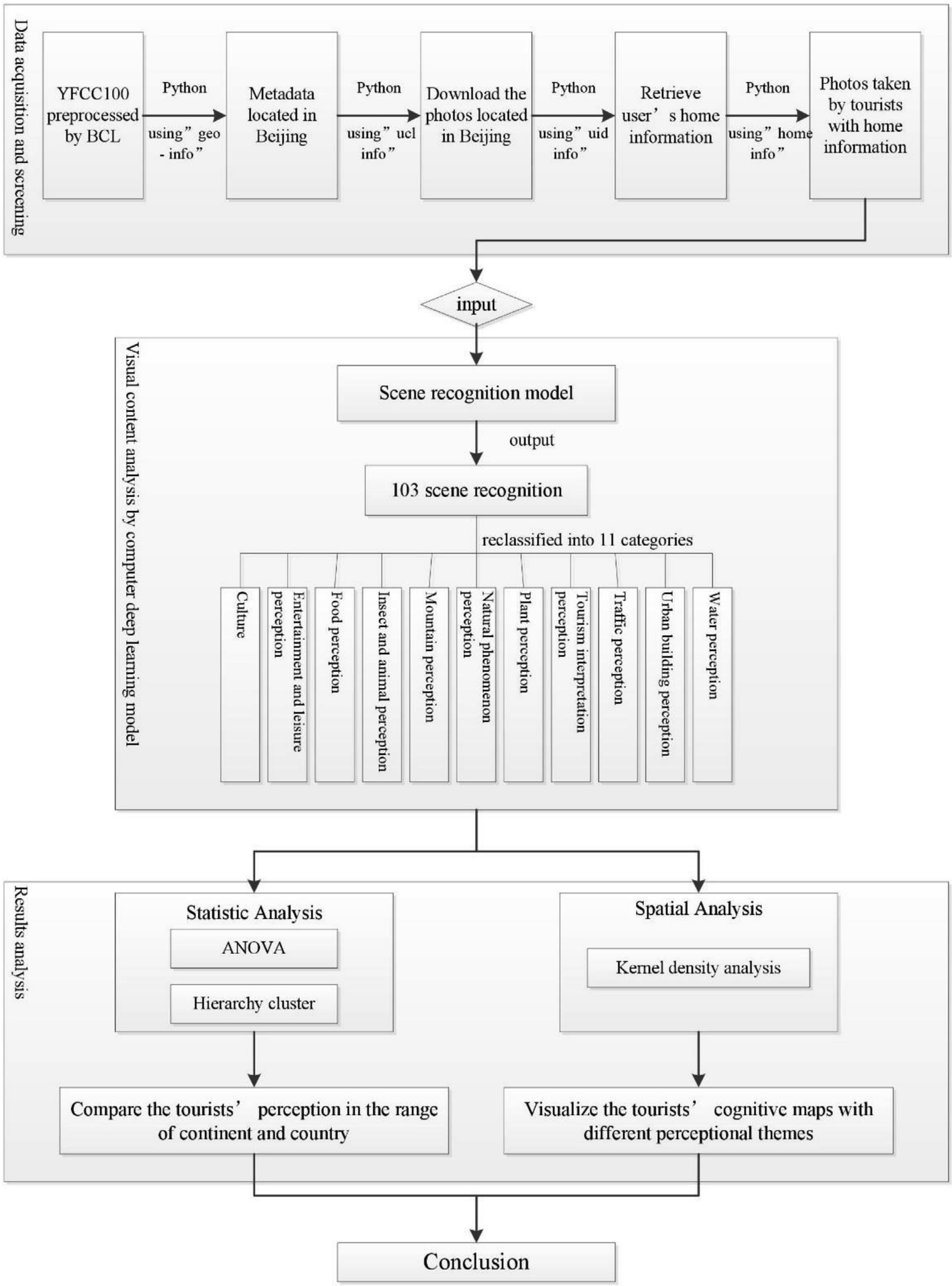


Fig. 1. Research process.

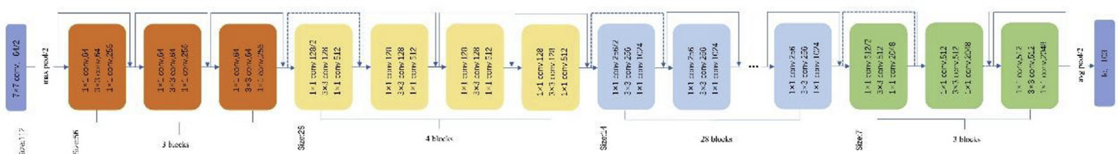


Fig. 2. Model structure of ResNet.

Table 1
11 Categories development from 103 recognized scenes.

perception type	scene
culture	dragon dance, fireworks, Fu character, library, lion dance, red envelope, stage, Xi character
entertainment	badminton court, bar, baseball court, billiard room, bowling alley, chess, football court, go, indoor basketball court, mahjong, ping-pong court, playground, swimming pool, tennis court
food	dining room, food, McDonald's, restaurant
insect and animal	bee, butterfly, camel, cat, deer, dog, dragonfly, elephant, giraffe, kangaroo, ladybug, leopard, lion, ornamental fish, panda, peacock, penguin, rabbit, rhinoceros, tiger, tortoise
mountain	mountain
natural phenomenon	night, overcast, snow, sunset, blue sky
plant	fallen leaves, flower, green plant
interpretation	map, text
traffic	aircraft, bicycle, cabin, car, helicopter, in car, motorcycle, ship, station, train
building	cathedral hall, corridor, European buildings, Islamic buildings, old buildings, overlook, skyscraper, wide-angle tower, worksite
water	beach, bridge, waterfall, waterside, water surface

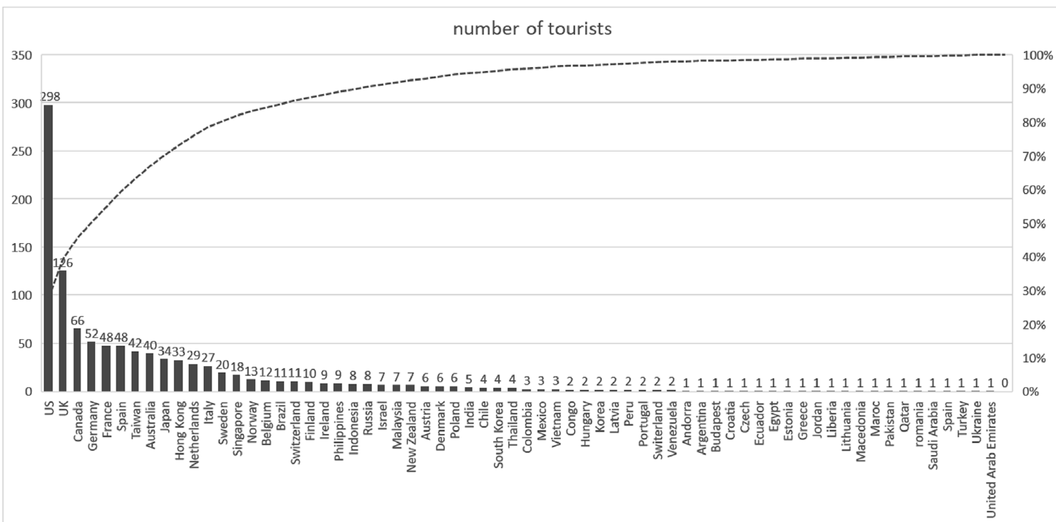


Fig. 3. Distribution of tourists from different countries.

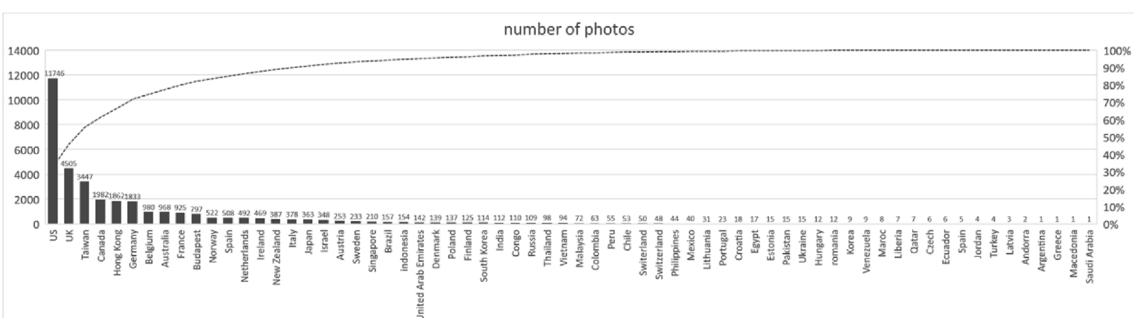


Fig. 4. Distribution of tourists' photos from different countries.

(Pan et al., 2014b), correlation analysis between affective qualities of places/landscape and the descriptive adjectives were conducted by "content analysis" and "semiotic analysis" (Albers & James, 1988). Third, photos as indicators for diagnosing the TDI divergences between DMO (destination management organizations) and tourists, tourists & residents, are different information resources (Choi, Lehto, & Morrison, 2007; Mak, 2017; Valek & Williams, 2018; Önder & Marchiori, 2017). Fourth, photos as logos serving for branding landmarks. The most suitable landmarks photos were found through the content analysis of tourist geo-tagged photos (Kennedy, Naaman, Ahern, Nair, & Rattenbury, 2007).

Methodology for visual content analysis. Visual content analysis is a systematic, observational method used for exploring how the studied phenomenon is represented (Bell, 2001). It allows the quantification of samples of observable content classified into distinct categories (Bell, 2001; Mak, 2017). Content analysis is attribute-based and is primarily concerned with quantitatively describing the appearance of specific themes and attributes in the collection of images, allowing the main focal items in the pictures to be identified and their frequencies, co-occurrence, clustering, and other related issues to be recoded (Stepchenkova & Zhan, 2013). Trained category and coding researcher are the crucial part of the procedure. In addition, Mise en scene refers to shot density, shot scale, shot angle, and shot composition, which has been employed as a complement for visual content analysis. NVivo, textblob, ThoughtView, and other related software are applied as tools for data statistics.

Table 2
Results of analysis of variance.

perception	quadratic sum	variance	mean square	F-value	significance
building	7.212	5	1.442	5.961	0
food	5.451	5	1.09	29.26	0
mountain p	1.951	5	0.39	11.164	0
natural phenomenon	3.329	5	0.666	24.785	0
entertainment	0.663	5	0.133	8.531	0
plant	4.146	5	0.829	11.155	0
cultural	11.321	5	2.264	11.139	0
insect and animal	0.346	5	0.069	4.014	0.001
traffic	0.403	5	0.081	3.065	0.009
interpretation	0.284	5	0.057	2.042	0.07
water	0.109	5	0.022	1.538	0.174

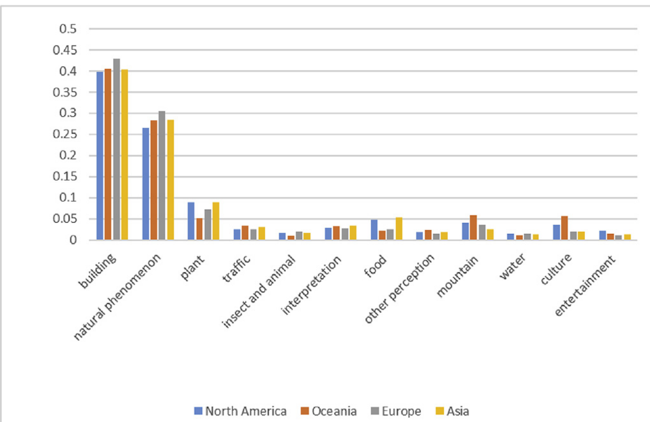


Fig. 5. Bar charts of tourists' perceptual percentage within 11 categories.

Compared to the tourism destination image, the concept of city image has existed in the research field of urban space for decades. Lynch maintained that city image formation is the result of a bilateral process led by the observer and the environment (Lynch, 1960). City image, which includes both of the tourists' and residents' perspectives, is utilized in the initial phase of a city's brand development (Kavaratzis, 2004; Oshimi & Harada, 2018). There are already several attempts in the field of city and urban planning with Visual content analysis of UGC photo. For example, Liu downloaded photos of 26 cities in the world (an average of 100,000 photos per city) with geographic information from Panoramio and Flickr using python scripting language, and categorized the photos into 102 scenes and 7 perceptions (the Green Perception, Water Perception, Transportation Perception, High-rises Perception, Architecture Perception, Socializing Perception, and Athletic Perception) by content analysis, which is built on a computer vision technique - scene understanding. The characteristics and spatial distribution of seven perceptions were statistically analyzed (Liu, Zhou, Zhao, & Ryan, 2016). Except the image issue, online data shared on social networks, particularly geo-tagged photos, are taken as an increasingly attractive source of information about cultural ecosystem services, a content analysis of 1.404 photos uploaded in Flickr and Panoramio has been conducted in five European cities for exploring how they represent cultural ecosystem services (Oteros-Rozas, Martín-López, Fagerholm, Bieling, & Plieninger, 2018).

Excepted using UGC photos as the resource for the exploration of city image, goolge street view (GSV) is another photo resource for content analysis in the urban study. Study issues include urban greening (Li, Ratti & Seiferling, 2018; Li & Ratti, 2018), Neighborhood walkability (Yin, Cheng, Wang, & Shao, 2015), Bikeability (Evans-Cowley & Akar, 2014), perceived naturalness (Hyam, 2017), public open spaces (Mygind et al., 2016) and so on. Nearly all the new emerging studies with content analysis using GSV has employed a deep learning method by computer vision technology (Rzotkiewicz, Pearson, Dougherty, Shortridge, & Wilson, 2018). All these give inspiration for the pictorial content analysis with big data in tourism research.

Therefore, big data has uniqueness and advantages compared to traditional data collection, and the content of the photo also holds great potential and effectiveness compared to textual material. It is the inexorable tendency to use deep-going mining for applying the visual pictorial contents into tourism research in the social network occupied era, as it is the most popular visual information in our present world and filled all the corners. Currently, the research concerning visual content-analysis fall into two approaches: one is the conventional way, which manually identifies visual contents of photo and taking attached textual information as complement, and the other is an emerging approach which interpreting the photo's content with the help of computer vision technology. For the conventional content analysis, the limited numbers of photos constrained the size of the sample, and the instability and irregularity of artificial coding for the visual content category show a common drawback. On the contrary, the emerging approach has an unparalleled advantage in the big datasets and relatively standardized classification processing by machine learning. However, the conventional approach holds a relatively mature conceptual and theoretical framework compared to the study using computer vision technology. Also, a well-designed and performed deep learning technology itself takes great time and energy in the phase of coding and machine training.

This study will refer to the theoretical framework of the conventional visual content study, and adopt an already well-performed visual content identification process in computer vision technology, and then explore the characteristics of tourist perception and the reflected destination image by combining the results of the pictorial content analysis and geo-tagged based spatial analysis.

3. Dataset and the study city

3.1. Dataset-Flickr YFCC 100M

In July 2015, Yahoo released a visual content dataset titled "Yahoo Flickr Creative Commons 100M" (YFCC 100M), the most extensive public multimedia collection that has ever been released. The dataset contains a total of 100 million media objects from the inception of Flickr in 2004 until early 2014, of which approximately 99.2 million photos and 0.8 million videos. Each media object in the dataset is represented by 23 pieces of metadata (Shamma et al., 2016). All the metadata are presented in line. In this study, five types of information are employed as: (1) photo/video identifier and user NSID (PID-picture identification/UID-user identification), used to identifying and counting the photo or user (2) date taken (day and time), used for tracking tourists' trajectory (3) title/description/user tags (comma-separated) used for analyzing tourists cognition and behavior (4) longitude and latitude, used for matching geographical coordinates (5) photo/video download URL, used for download of the original photos.

In all, YFCC 100M contains 48,366,323 geo-tagged photos (Deng & Li, 2018), and BCL[1] (Beijing city lab) selected all the 2,171,162 photos in China for researchers. It saves a lot of energy for handing the original dataset from Flickr and make it easier to cut out the data located in the study city.

3.2. The study city

Beijing is the capital of the People's Republic of China, it has been the political, economic, and cultural center of China for over 800 years from the Yuan Dynasty. It also serves as the most important transportation hub and port of entry. Now it has become one of the most

[1] https://www.beijingcitylab.com.

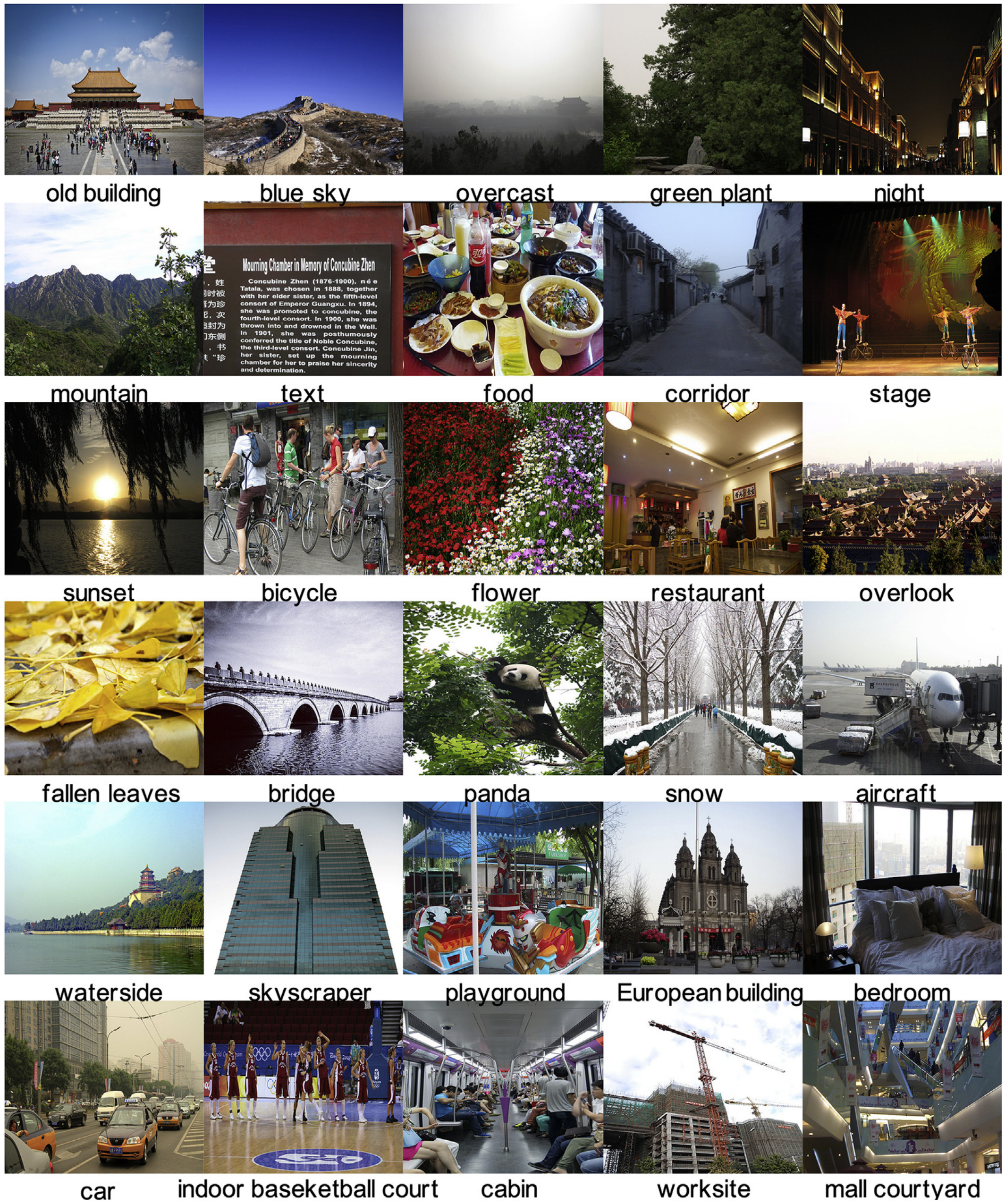


Fig. 6. Examples of the 30 most frequent scenes.

popular travel destinations in the world, world-famous scenic spots include Great Wall, Forbidden City, Summer Palace, Olympic Park, and so on. In 2017, 293.54 million Chinese tourists and 3.92 million international visitors had visited Beijing, according to the Beijing Statistics Bureau. Statistics by BCL labs show that Beijing has the largest number of photos; it indicates Beijing is the most popular destination for tourists.

For establishing the dataset in Beijing, both original metadata and the filtered BCL data are taken. Firstly, geo-tagged photos in Beijing were selected by Beijing administrative boundary in ArcGIS. Then, the completed metadata was supplemented by matching the PID (each photo has a distinctive PID) through writing script code in Python. 144967 metadata of photos were selected in Beijing, and each one has a download URL. By allocating the task to the computer with programming, the original uploaded photos by users were automatically downloaded. As users have deleted some of the photos, 137265 photos were found. By code designed in Python, information of the user's birthplace, city and country are retrieved by invoking API (Application Programming Interface) data. According to the three geographic locations filled by the users themselves, 1075 foreign users' countries were identified effectively, and 35356 photos uploaded by them were taken as the object of scene recognition. It should be noted that there is a restriction for Flickr in China; Flickr users who upload photos shot in Beijing are mainly tourists who come from overseas. For this reason, the perception of tourism destination image in this study is generated from oversea tourists.

4. Methodology

4.1. Research process

The overall research is divided into four steps, the first step is data collection and screening, the second step is the process of photo content analysis by computer deep learning model, the third step is the analysis

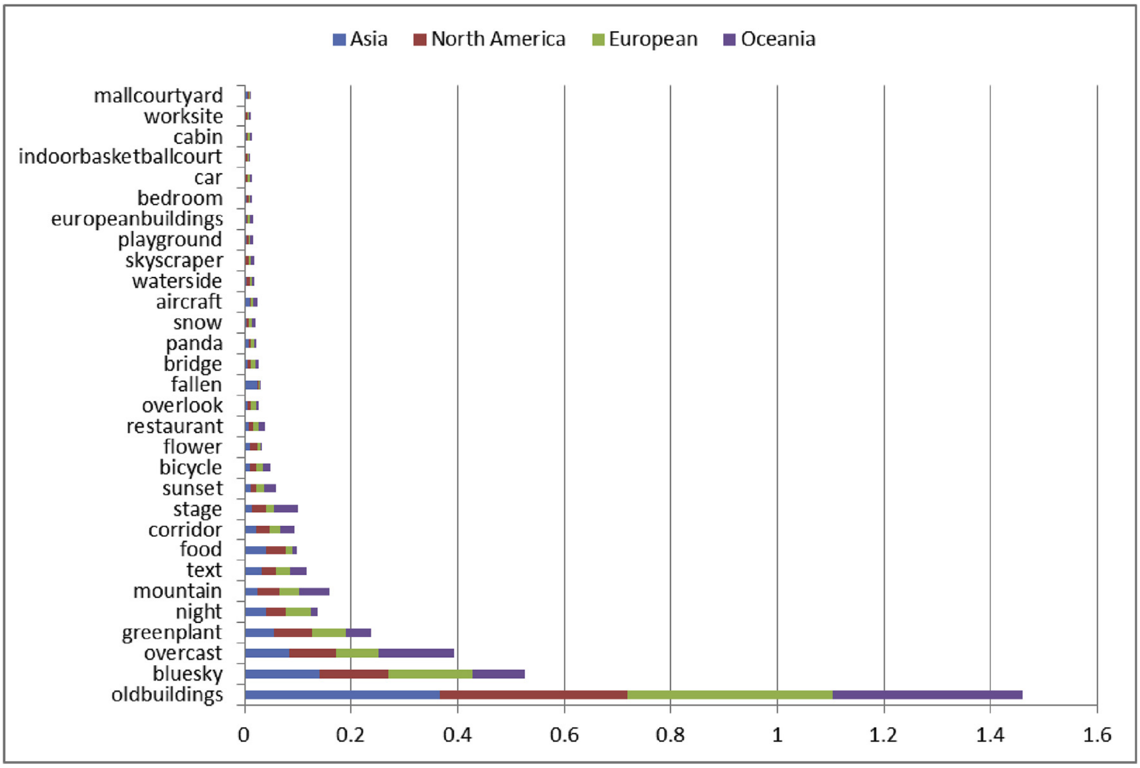


Fig. 7. Top 30 Scenes perceived by tourists from Asia, Europe, North American, and Oceania.

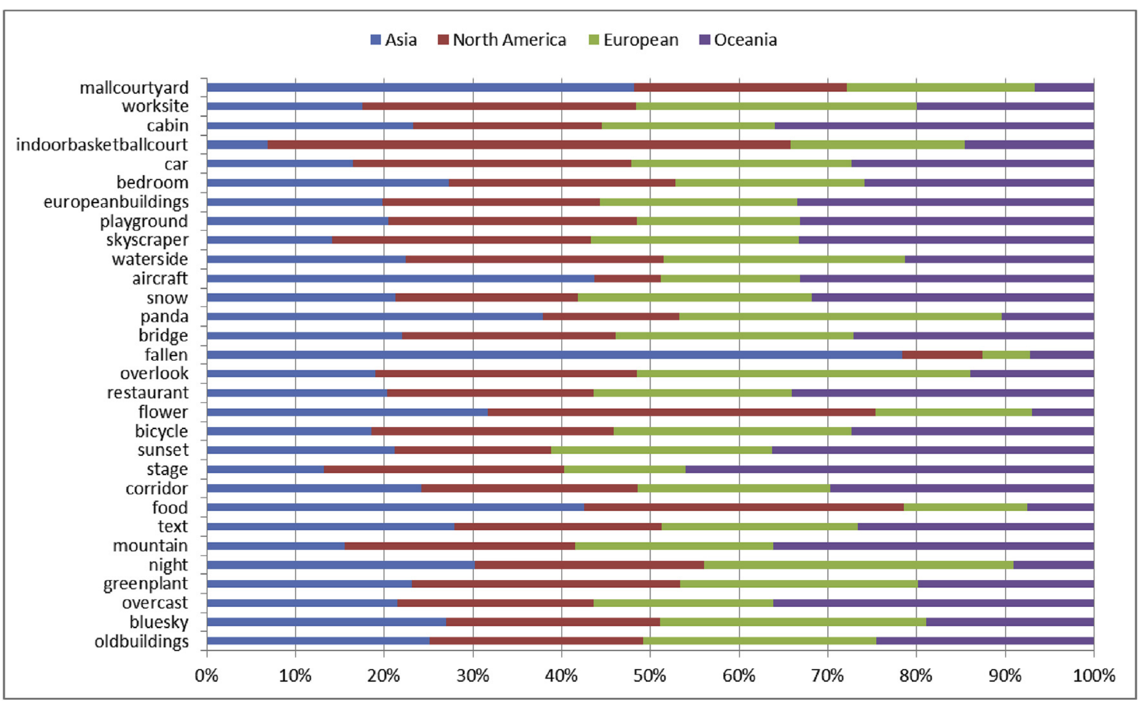


Fig. 8. Comparison of the top 30 scenes' perception in four continents.

of the results, which includes comparative perceptual analysis of tourists from different places in the world and spatial distribution with different preferences, and then the conclusion of the study is the last step (Fig. 1). Compared to the conventional photo-based tourism research methods, the substation of the artificial intelligence approach for natural intelligence approach in step 2 is the innovation, and the category developed based on the output of the machine learning model is a new way of induction for destination cognitive image analysis.

Data collection and screening has been introduced in section 3. In step 2, the deep learning model of scene recognition is employed for recognizing the contents of tourists' photos. The purpose of scene recognition is to identify the tourist photos referring to 103 scenes, and each photo would belong to one unique scene. In step 3, statistical analysis and spatial analysis are conducted in parallel as two independent and complementary approaches for analyzing the data results of scene recognition, tourists' home information (country and

Table 3
Specific proportions of the 30 scenes' perception in four continents.

Continent	Scene (> 30%)	Scene (< 20%)
Asia	mall courtyard (48%), aircraft (43%), panda (38%), fallen leaves (78%), flower (32%), food (31%)	mountain (16%), stage (13%), bicycle (18%), restaurant (19%), overlook (19%),skyscraper (14%), European buildings (19%),car (17%), worksite (18%), indoor basketball court (7%)
North American	Indoor basketball court (57%), overlook (30%), flower (42%), food (31%), skyscraper (30%), worksite (32%)	Aircraft (9%), panda (15%), fallen leaves (10%), sunset (16%)
Europe	Worksite (30%), panda (36%), overlook (39%), night (35%), blue sky (30%)	Cabin (19%), playground (19%), aircraft (16%), stage (14%), fallen leaves (5%), food (14%)
Oceania	Cabin (36%), European buildings (33%), playground (33%), skyscraper (33%), aircraft (33%), snow (32%), restaurant (34%), sunset (37%), stage (47%), mountain (37%),	Mall courtyard (7%), indoor basketball court (13%), panda (10%), fallen leaves (7%), flower (7%), overlook (14%), food (8%), night (9%), green plant (20%), blue sky (19%)

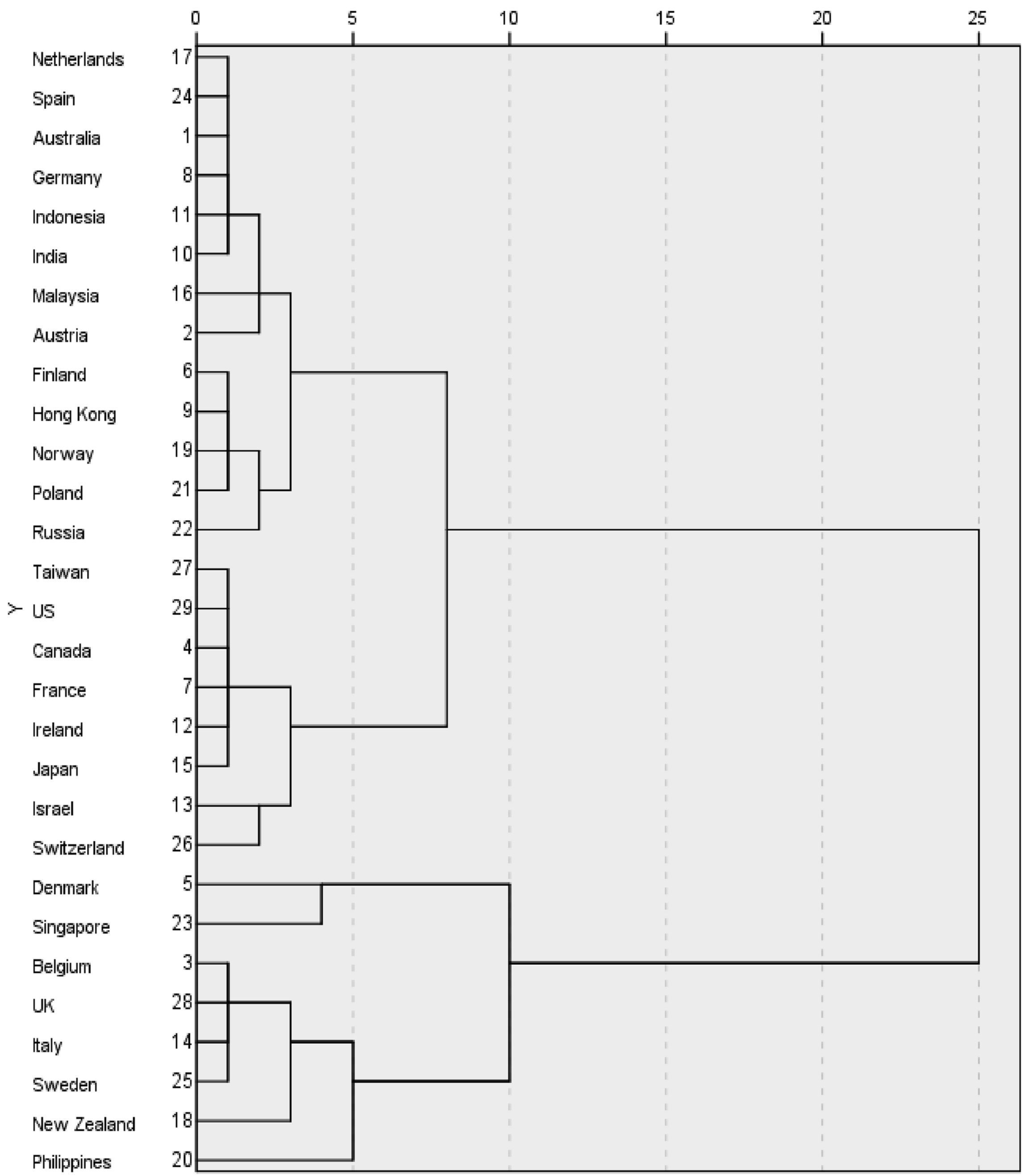


Fig. 9. Hierarchical cluster of countries referring to the photos' scenes.

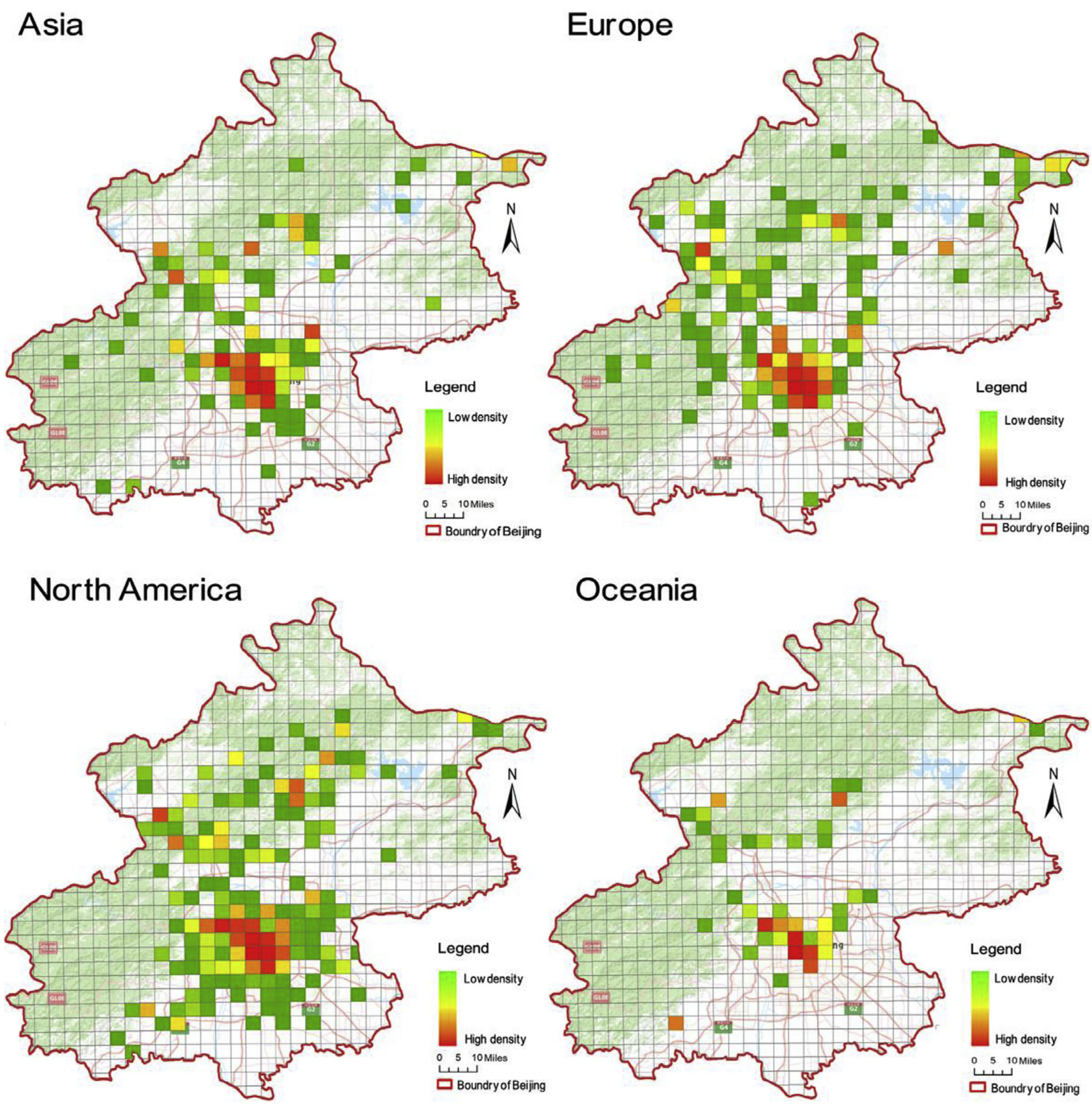


Fig. 10. Intercontinental differences of distribution of tourists' footprint.

continent), and photos' geographical information (longitude and latitude) are applied for comparative statistic analysis and spatial positioning. Statistical analysis is mainly to detect perceptual differences between tourists from different countries and continents. Hierarchical clustering analysis and ANOVA (analysis of variance) are the specific methods in statistical analysis. Spatial analysis is applied to discover the spatial distribution characteristics of tourists' behaviors preference with different scenes by the help of ArcGIS. The last part is responding to the research objective based on the analysis process and comes to a conclusion.

4.2. Deep learning models for recognizing photos' content

This study employs widely spread ResNet (He, Zhang, Ren, & Sun, 2015, pp. 770–778) for the image classification task-scene recognition. This deep learning structure obtained very successful results in these two computer vision competitions-ImageNet and MS-COCO. The core idea exploited in these models -residual connections, is found to improve gradient flow greatly, thus allowing training of much deeper models with tens or even hundreds of layers. Three realizations which are ResNet-50, ResNet-101 and ResNet-152 are proposed, then it chooses ResNet-101 in these experiments. A demonstrated model structure with only 101 layers is shown in Fig. 2. As an output, 103 scenes are recognized by taking tourists' photos as input. The designing of the machine learning model for scene recognition is the key technology of this study. As it is already a mature field in computer vision, there are many frontier papers about the details (structure and code) of the model (Cheng, Zhang, Fan, & Harris, 2018; Meiyin & Li, 2015).

This applied deep learning model has been tested with experimental data, experimental results show that it is efficient and robust for scene recognition with recall rate and high precision, the recall rate achieved 90% , and the false recognition rate is 0.1%. The recall rate indicates the ratio of the number of identified scenes to the total number of this scene. The false recognition rate means the proportion of the number of the false identified scenes to the total number of this scene.

4.3. Perceptual categories developed from 103 scenes

103 scenes are too fragmented to analyze the general characteristics of tourists' perception. By referring to the similarity of tourists' perception, 103 scenes are reclassified into 11 categories (Table 1), which covered most tourists' perception in tourism destinations.

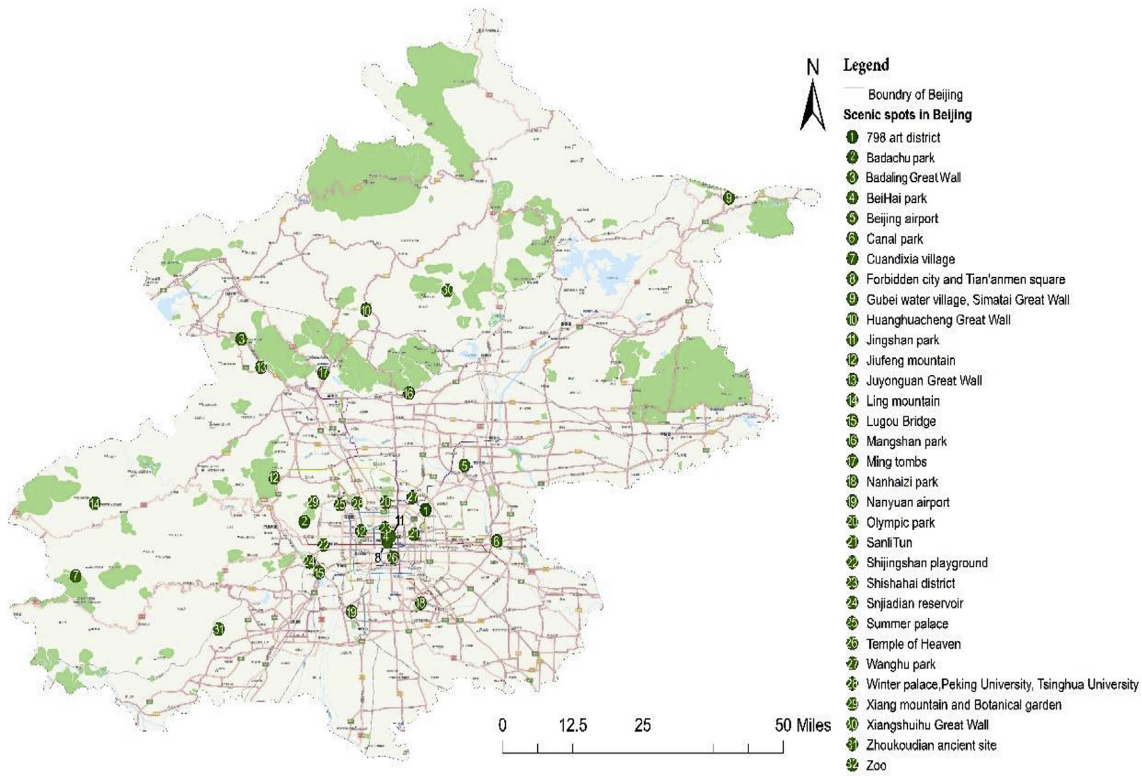


Fig. 11. Distribution of tourism resources in Beijing.

4.4. Tools for statistical and spatial analysis

In this study, the tools of analysis of variance and hierarchical cluster analysis were used for statistical analysis, and fishnet and kernel density analysis in ArcGIS were used for spatial analysis.

ANOVA is a collection of statistical models developed by Ronald Fisher (Fisher, 1921), which could be used to analyze the differences among group means in a sample. This model provides a statistical test of whether the population means of several groups are similar and to what extent, and therefore generalizes the *t*-test to more than two groups. ANOVA is useful for comparing three or more group means for statistical significance. In this study, ANOVA aims to detect whether a statistic significance exists in tourists from different places of the world. Hierarchical clustering is a method of cluster analysis which seeks to build a hierarchy of clusters. The principle of this method is to define the similarity or distance between network nodes by a given network topology structure, and then use single-link hierarchical clustering or full-link hierarchical clustering to form a tree-like graph hierarchical structure of network nodes. With this tool, the similarities and divergence of tourists' behaviors and perceptions from different continents and countries/regions could be explored.

By using fishnet in ArcGIS, Beijing was divided into 5 km * 5 km grids, with matching the longitude and latitude attributes of photos, the place where the photos were shot could be located, and the number of photographs in each grid is counted. Through this process, the distribution of the tourists' footprints could be displayed. The kernel density analysis in ArcGIS could calculate the density of point elements around each output raster pixel, which is the best option for visualizing the aggregation characteristics of tourists' behavior and perception in Beijing.

5. Result

5.1. Statistical analysis

5.1.1. Descriptive analysis

According to the statistics, 35356 photos uploaded by 1075 tourists from 64 countries/regions and 6 continents (Asia, Europe, North America, Oceania, South America and Africa) are collected (Fig. 3). The number of tourists and photos from North America (365/13728), Europe (443/12542), Asia (189/7205) and Oceania (47/1355) are relatively abundant, while South America (27/384) and Africa (5/143) have the least number of tourists and photos. Looking at the distribution in countries, as Flickr is most prevalent in the United States and Britain, these two countries hold the most significant number of tourists and photos comparing to other countries. Specifically, the United States ranked 1st place with 299 tourists and 11746 photos (Figs. 3 and 4), which means each American tourist uploaded average 39 photos. As the total number of tourists in South America and Africa is too small, the statistical analysis is restricted within four continents: Asia, Europe, North America and Oceania. With the same reason, only 30 countries whose photos have exceeded 100 were selected as the subjects for this research.

5.1.2. Intercontinental difference

According to Table 2, tourists from four continents had significant differences in the perception of the building, natural phenomenon, mountain, culture, plants, insects and animals, and traffic, which was supported by the considerable value (< 0.01). While, the difference of “water” and “interpretation” between continents is not apparent, as the significance value is higher than 0.05, or take another word, it shows commonalities in these two kinds of perception.

By comparing the proportion of tourists' photos in each perceptual category (Fig. 5), performances of similarities and differences could be detected. In general, tourists hold similar perceptions of building, natural phenomena, and plant. For differences, European tourists have a

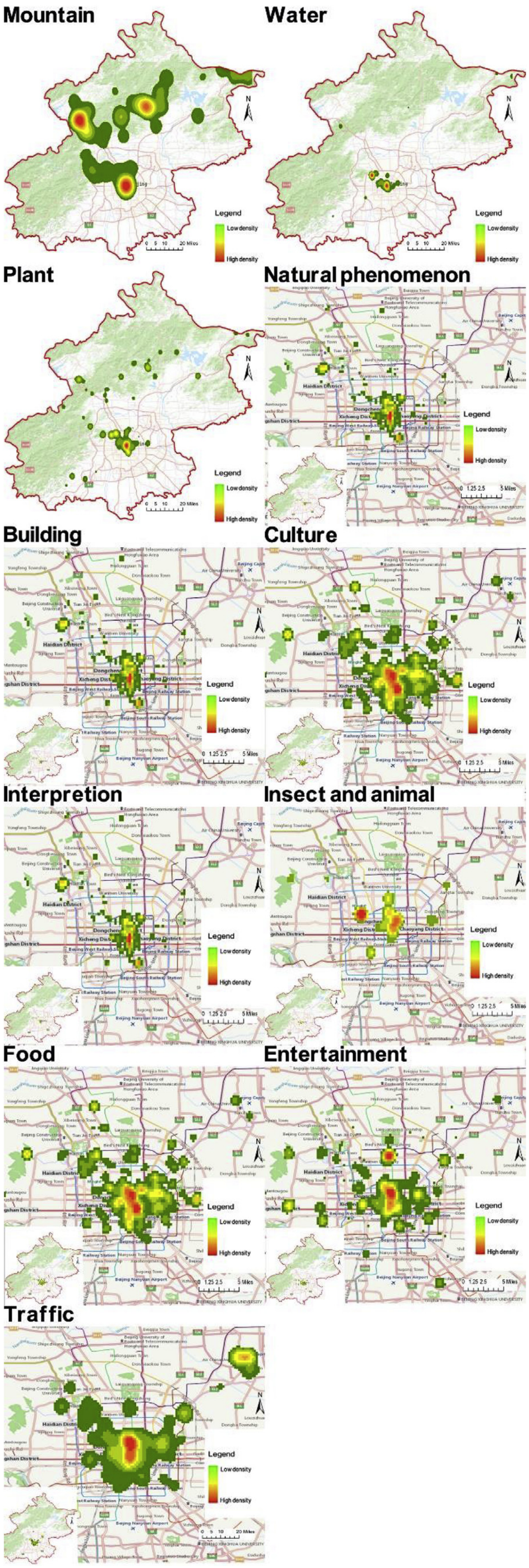


Fig. 12. Spatial distribution of different perception in the tourist destination.

stronger preference for building and natural phenomena, Oceania tourists have a stronger preference for mountain and culture, but less preference of plant, animal and insect. Both Asian and North American tourists have a stronger preference for plants and food.

In order to discover and compare the characteristics of tourists' perception from different continents, the 30 most frequent scenes' number (Fig. 6) and their proportion were counted and showed in Figs. 7 and 8 and Table 3. Referring to these statistical data, the most popular five scenes from tourists are traditional buildings, blue sky, overcast, green plants and mountain.

Intercontinental differences are discovered as follows. For Asian tourists, they are more interested in the mall courtyard (48%) and aircraft (43%), especially, pay more attention at fallen leaves (78%) than the tourists from the other three continents. Shopping malls and airports are the places that Asian tourists most willing to share. Panda (38%) is indeed important in the mind of Asian tourists, and they are more interested in food (31%) but pay less attention to the condition of restaurants (19%). They are totally not interested in sports and stage performances. They also pay little attention to overlook (19%), skyscraper (14%), European buildings (18%) and worksite (13%), which formed the urban visual space. For Oceania tourists, their enthusiasm for stage (47%) is much higher than the other three continents. At the same time, they are more interested in European buildings (33%), playground (33%), skyscraper (33%), aircraft (33%), snow (32%), sunset (37%) and mountain (37%) than the other three continents. But they particularly dislike Mall courtyard (7%), fallen leaves (7%), flower (7%), night (9%), and pandas (10%), which are representatives of Chinese tourism image. Besides, they are not very interested in food (8%), even they pay much attention to the restaurant environment (34%). For Europe tourists, they are more prefer to overlook (39%) and night (35%) relatively, and has a better recognition for panda (36%), but they are particularly insensitive for fallen leaves (6%), and are not so interested in food (14%) and stage (14%). North American tourists show more enthusiasm for indoor basketball court (59%) than the other three continents, and they are more interested in flowers (42%) than the other three continents. Aircraft (9%) and fallen leaves (10%) are not so attractive to them, and panda (15%), neither.

5.1.3. National differences

By applying the method of hierarchical cluster analysis, the preference of tourists from 30 countries was clustered according to the proportions of 12 categories of photos' scene cognition. With the group distance of 5, the tourists from different countries could be divided into four groups (Fig. 9).

Group I (13 countries/regions) includes 5 Asian countries or regions (Hong Kong, India, Indonesia, Malaysia, Russia), 6 European countries (Finland, Germany, Netherlands, Norway, Poland, Spain) and 2 Oceanian countries (Australia, Austria). The group characteristics of tourists are high preference for urban buildings, plants and natural phenomena, medium preference for tourism interpretation, and low preference for insects &animals and entertainment & leisure. Group II (8 countries/regions) includes 3 Asian countries or regions (Israel, Japan, Taiwan),2 North American counties (Canada and US) and 3 European countries (France, Ireland, Switzerland). The group characteristics of tourists are high preference for urban building and natural phenomenon, medium preference for traffic and food, and low preference for insects & animals and water. Group III (2 countries) includes Denmark and Singapore. The group characteristics of tourists are high preference for natural and plant scenes; medium preference for insects and animals and mountain, and low preference for tourism interpretation. Group IV (6 countries) includes 1 Asian country (Philippines), 1 Oceanian country (New Zealand) and 4 European countries (Belgium, Italy, Sweden, UK). The group characteristics of tourists are high preference for urban building and natural scenes, medium preference for mountain and plant, and low preference for shopping and entertainment.

5.2. Spatial distribution

5.2.1. Spatial distribution of tourists' footprints

Based on the result conducted in the software of ArcGIS (Fig. 10), the footprint's distributions of tourists from the four continents in Beijing are compared. North American tourists have the most extensive footprint distribution. The area's rank of tourist footprint distribution is like NA > EU > AS > Oceania. Tourists from four continents are all gathering around the traditional scenic spots. One dense area is the central urban area, composed by Forbidden City, Tiananmen Square and the Temple of Heaven. The other dense area is in Northwest Mountain, where are located the Great Wall scenic spots. Outside the central urban area, the distribution of tourists' footprint from four continents is quite different. The footprints of tourists from North America and Europe are more intensive in San-li-tun area than Asian and Oceanian tourists, while European tourists are less likely to enter the Winter Palace.

5.2.2. Spatial distribution of tourists with 11 scene categories

Referring to the spatial distribution of tourism resources in Beijing (Fig. 11) and tourists' footprints' distribution in tourism destinations (Fig. 12) conducted in the software of ArcGIS, several characteristics can be drawn. **Mountain and water** are clearly and precisely perceived by tourists. Mountain perception is mainly scattered in Badaling Great Wall and Mutianyu Great Walls. Water perception is displayed in the Beihai Park and the Summer Palace. **Insect & animal** perception shows a concentration in Beijing zoo. Tourists' perceptions of **building, culture, and interpretation** are mainly concentrated in traditional scenic spots and cultural heritage-such as Summer Palace, Forbidden City and so on, which possess a higher value of protection and a higher demand for interpretation. Tourists' perception of **food** is displayed in two types of scenic spots-traditional scenic spots and modern shopping areas. The former is around the Forbidden City, Beihai Park, Summer Palace and other scenic areas.The latter is around Wangfujing, Sanlitun commercial areas. “Sightseeing + food” and “shopping + food” become the main combination of tourists' behaviors. Tourists' perception of **entertainment** formed two centers, one is in the Olympic Sports Park, the other is in the traditional cultural communities. The Olympic Sports Park provides space for watching sports, and the traditional cultural communities provide space for tourists observing local recreational activities in the traditional neighborhoods, for example, chess, mahjong. Tourists' **traffic** perception gathers in the old central city and airport. The former mainly shows the mixed traffic landscape around the traditional scenic spots, and the other is illustrated by airplanes in the airport.

6. Conclusion and discussion

6.1. Evidence and theory for tourism destination development

In all, the traditional scenic spots are the main perceived subjective for inbound tourists in Beijing. Specifically, the traditional buildings in Forbidden City, Temple of Heaven Park, Summer Palace are the prominent perceived contents in the Beijing, and Great Wall which is located in the mountains are the secondary perceived content following the traditional building. Tourism destinations should extend the tourism attractions from traditional scenic spots to more modern scenic spots, and lead tourists to take more experiential activities. Compared to the traditional scenic spots, some specific attractions should be given more attention. The first is food, which is supposed to be the most popular destination representative in China, but did not receive strong recognition by all inbound tourists. In contrastto food, the perception of the stage is unexpectedly high from inbound tourists, it indicates that theatrical activities should be included as an attraction enhancement in Beijing, and more possibilities should be given to tourists who are interested in local cultural shows.

Also, influenced by cultural, social, and economic factors, tourists' preferences and recreational needs are complicated and diverse, but groups do exist among countries. All these specific results could give clues to more appropriate and detailed marketing promotion for DMOs. For Asian tourists, their recognition and preference of pandas, food, and other traditional Chinese images are more remarkable and shopping has become a behavioral feature of Asian tourists, while they are reluctant to pay time and money for participation in stage performances. North American show enthusiasm about basketball, and their footprints are the most widely distributed. European tourists pay more attention to the attributes of the sky and pandas, but they are not willing to try Chinese food and stage performances. Oceania has the smallest footprint, they show outstanding interests in the stage, but little interests in the blue sky, night, shopping malls, and flowers. From all these piece of evidence, differential marketing strategies for tourists from different continents could be experly made by DMOs for destination promotion.

The visual content analysis of big pictorial data provides a new way of understanding tourists’ preferences. This study also induced a new category with 103 attributes, which make a contributes to the theoretical framework of the tourism destination image. As the 103 scenes are selected from the most common scenes appearing in shoot photos, this study adds the details to the attribute-holistic framework of destination image proposed by Echtner and Ritchie (Echtner & Ritchie, 1993), and the developed category extends the topologic knowledge about “what” is perceived by tourists, which could be applied and examined in the destination cognitive image analysis.

6.2. Reflection of the newly developed methodology

The applicability of the newly developed methodology is one of the objectives of this study. Compared with previous studies, the innovation of this research method lies in the substitution of artificial intelligence analysis for natural intelligence analysis. The application of computer vision technology shows advantages in time-saving and data analyzing towards a huge number of photos. This research pioneered the application of computer vision technology for the study of destination cognitive image. The results show that scene recognition technology provides abundant evidence and information for revealing tourists' perception and behavior preferences in tourist destinations. As an initial attempt, this study opens a door for widely applying the interdisciplinary technologies- “computer vision” into tourism research. These visual AI technologies include but are not limited to the following aspects: visual semantic segmentation, visual emotional analysis, short video content recognition, etc.

Arising from this study, the merits of the new methods include the clarity of users' portrait, the capture of spatiotemporal information, and the richness and inductiveness of the recognized contents. By the data mining in social media, users' home location can be obtained, which makes it possible to compare tourists' behavior and preferences partly with demographic information. The spatiotemporal information attached by photos makes it possible to caption the tourists' behavior trajectory. Specifically, computer vision technology enables researchers to recognize the most common 103 scenes in human photography, which are flexibility and inductivity for topologic analysis.

Besides, the application of computer vision technology does have shortcomings, when the research topic is related to complex cultural and social issues, e.g., host-guest interaction, the symbolic meaning of pictures could not be well interpreted through this particular machine learning model. Or when the research needs to discover some new defined scene outside the list of 103 scenes, it is necessary to reassess the possibility of this AI-based method as it will consume a lot of energy and time to train a new machine learning model which could fit the research question.

7. Limitations and future work

Big data and computer vision technology are the highlights of this study. Limitations exist in both two aspects. For the former, tourists' photos shared on Flickr were taken as the big data resource, the popularity of Flickr varies from country to country in the world, which could influence the study results. For the latter, we selected and adopted a mature computer vision model that could optimum fit the study aim of the destination image, other than designing a new model, and this study's framework has somehow limited and adjusted to the outcomes of this adopted model. This limitation would be gradually broken through, as the supporting requirement for AI technology from tourism field becomes more specific and more precise after some pioneering attempts.

In the future, with the development of computer deep learning and big data mining technology, precisely identifying the content of massive tourist photos by artificial intelligence would continually deepen the tourism destination image study in the big data era. The whole spectrum of tourism marketing would be re-appraised and probably restructured as a result of the growing amount of sophisticated techniques emerging which offer more detailed information on what tourists do. It is also a changing time for DMOs, how to make use of AI as a viable and necessary part for the operations would need more practical exploration.

Acknowledgements

Thanks to National Natural Science Foundation of China (51608278) for supporting this research. We are also grateful to Professor Carson Lewis Jenkins, visiting professor, college of Tourism and Service management, Nankai University, for his valuable comments and suggestions on previous drafts of this article. We would like to thank Dr Zhibin Lin from Durham University Business School for his helpful comments.

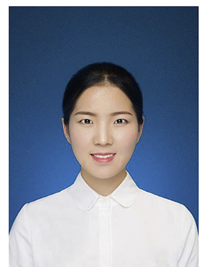
References

Albers, P. C., & James, W. R. (1988). Travel photography: A methodological approach. *Annals of Tourism Research, 15*(1), 134–158.
Batista E Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management, 68*, 101–115.
Bell, E. (2001). Content analysis of visual images. In C. J. T. Van Leeuwen (Ed.). *Handbook of visual analysis* (pp. 92–118). London: Sage.
Cai, G., Hio, C., Bermingham, L., Lee, K., & Lee, I. (2014). Sequential pattern mining of geo-tagged photos with an arbitrary regions-of-interest detection method. *Expert Systems with Applications, 41*(7), 3514–3526.
Cai, G., Lee, K., & Lee, I. (2018). Itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos. *Expert Systems with Applications, 94*, 32–40.
Cheng, F., Zhang, H., Fan, W., & Harris, B. (2018). Image recognition technology based on deep learning. *Wireless Personal Communications, 102*(2), 1917–1933.
Chi, T., Wu, B., Morrison, A. M., Zhang, J., & Chen, Y. C. (2015). Travel blogs on China as a destination image formation agent: A qualitative analysis using leximancer. *Tourism Management, 46*, 347–358.
Choi, S., Lehto, X. Y., & Morrison, A. M. (2007). Destination image representation on the web: Content analysis of Macau travel related websites. *Tourism Management, 28*(1), 118–129.
Chua, A., Servillo, L., Marcheggiani, E., & Moere, A. V. (2016). Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy. *Tourism Management, 57*, 295–310.
Deng, N., & Li, X. R. (2018). Feeling a destination through the “right” photos: A machine learning model for DMOs' photo selection. *Tourism Management, 65*, 267–278.
Echtner, C. M., & Ritchie, J. R. B. (1993). The measurement of destination image: An empirical assessment. *Journal of Travel Research, 31*(4), 3–13.
Evans-Cowley, J. S., & Akar, G. (2014). StreetSeen visual survey tool for determining factors that make a street attractive for bicycling. *Transportation Research Record Journal of the Transportation Research Board, 2468*(-1), 19–27.
Fisher, R. A. (1921). On the “probable error” of a coefficient of correlation deduced from a small sample. *Metron, 1*, 3–32.
García-Palomares, J. C., Gutiérrez, J., & Mínguez, C. (2015). Identification of tourist hot spots based on social networks: A comparative analysis of european metropolises using photo-sharing services and GIS. *Applied Geography, 63*, 408–417.
Gavric, K. D., Culibrk, D. R., Lugonja, P. I., Mirkovic, M. R., & Crnojevic, V. S. (2011). Detecting attractive locations and tourists' dynamics using geo-referenced images. *Paper presented at the international conference on telecommunication in modern satellite cable and broadcasting services*.
Gilbert, D., & Hancock, C. (2006). New York city and the transatlantic imagination. *Journal of Urban History, 33*(1), 77–107.
Gunn, C. A. (1972). *Vacationscape: Designing tourist regions.* Austin: University of Texas.
Han, J., & Lee, H. (2015). Adaptive landmark recommendations for travel planning: Personalizing and clustering landmarks using geo-tagged social media. *Pervasive and Mobile Computing, 18*, 4–17.
He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep residual learning for image recognition*.
Hu, Y., Gao, S., Janowicz, K., Yu, B., Li, W., & Prasad, S. (2015). Extracting and understanding urban areas of interest using geotagged photos. *Computers, Environment and Urban Systems, 54*, 240–254.
Hunter, W. C. (2013). China's chairman Mao: A visual analysis of hunan Province online destination image. *Tourism Management, 34*(Complete), 101–111.
Hunter, W. C. (2016). The social construction of tourism online destination image: A comparative semiotic analysis of the visual representation of seoul. *Tourism Management, 54*(2), 221–229.
Hussain, A., & Cambria, E. (2018). Semi-supervised learning for big social data analysis. *Neurocomputing, 275*, 1662–1673.
Hyam, R. (2017). Automated image sampling and classification can Be used to explore perceived naturalness of urban spaces. *PLoS One, 12*(1), e169357.
Jang, S., & Moutinho, L. (2019). Do price promotions drive consumer spending on luxury hotel services? The moderating roles of room price and user-generated content. *International Journal of Hospitality Management, 78*, 27–35.
Jiang, K., Yin, H., Wang, P., & Yu, N. (2013). Learning from contextual information of geo-tagged web photos to rank personalized tourism attractions. *Neurocomputing, 119*, 17–25.
Kavaratzis, M. (2004). From city marketing to city branding: Towards a theoretical framework for developing city brands. *Place Branding, 1*(1), 58–73.
Kennedy, L., Naaman, M., Ahern, S., Nair, R., & Rattenbury, T. (2007). How flickr helps us make sense of the world: Context and content in community-contributed media collections. *Paper presented at the ACM international conference on multimedia*.
Kim, S. B., Kim, D. Y., & Wise, K. (2014). *The effect of searching and surfing on recognition of destination images on Facebook pages.* Elsevier Science Publishers B. V.
Leeuwen, T. (2003). Semiotics and iconography. In T. van Leeuwen (Ed.). *Handbook of visual analysis* (pp. 92–118). London: Sage.
Leung, D. H., Lee, A., & Law, R. (2012). Examining hotel manager's acceptance of web 2.0 in website development: A case study of hotels in Hong Kong. *Social Media in Travel, Tourism and Hospitality,* 53–65.
Li, X., & Ratti, C. (2018). Mapping the spatial distribution of shade provision of street trees in Boston using Google Street View panoramas. *Urban Forestry and Urban Greening, 31*.
Li, X., Ratti, C., & Seiferling, I. (2018). Quantifying the shade provision of street trees in urban landscape: A case study in boston, USA, using google street view. *Landscape and Planning, 169*.
Liu, Y., Huang, K., Bao, J., & Chen, K. (2019). Listen to the voices from home: An analysis of Chinese tourists' sentiments regarding Australian destinations. *Tourism Management, 71*, 337–347.
Liu, L., Zhou, B., Zhao, J., & Ryan, B. D. (2016). C-IMAGE: City cognitive mapping through geo-tagged photos. *Geojournal, 81*(6), 817–861.
Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management, 68*, 301–323.
Lynch, K. (1960). *The image of the city.* MIT Press.
Mak, A. H. N. (2017). Online destination image: Comparing national tourism organisation's and tourists' perspectives. *Tourism Management, 60*, 280–297.
Mcgregor, A. (2000). Dynamic texts and tourist gaze. *Annals of Tourism Research, 27*(1), 27–50.
Meiyin, W., & Li, C. (2015). Image recognition based on deep learning. 2015-01-01 *Paper presented at the 2015 Chinese automation congress (CAC)*.
Mygind, L., Bentsen, P., Badland, H., Edwards, N., Hooper, P., & Villanueva, K. (2016). Public open space desktop auditing tool—establishing appropriateness for use in Australian regional and urban settings. *Urban Forestry and Urban Greening, 20*, 65–70.
Nöth, W. (1990). *Handbook of semiotics.* Indiana University Press.
Önder, I., & Marchiori, E. (2017). A comparison of pre-visit beliefs and projected visual images of destinations. *Tourism Management Perspectives, 21*, 42–53.
Oshimi, D., & Harada, M. (2019). Host residents' role in sporting events: The city image perspective. *Sport Management Review, 22*(2), 263–275.
Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C., & Plieninger, T. (2018). Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecological Indicators, 94*, 74–86.
Pan, S., Lee, J., & Tsai, H. (2014a). Travel photos: Motivations, image dimensions, and affective qualities of places. *Tourism Management, 40*(1), 59–69.
Pan, S., Lee, J., & Tsai, H. (2014b). Travel photos: Motivations, image dimensions, and affective qualities of places. *Tourism Management, 40*(1), 59–69.
Qu, H., Kim, L. H., & Im, H. H. (2011). A model of destination branding: Integrating the concepts of the branding and destination image. *Interdisciplinary Journal of Research in Business, 32*(3), 465–476.
Rzotkiewicz, A., Pearson, A. L., Dougherty, B. V., Shortridge, A., & Wilson, N. (2018). Systematic review of the use of Google Street View in health research: Major themes, strengths, weaknesses and possibilities for future research. *Health & Place, 52*, 240–246.
Salas-Olmedo, M. H., Moya-Gómez, B., García-Palomares, J. C., & Gutiérrez, J. (2017). Tourists' digital footprint in cities: Comparing Big Data sources. *Tourism Management, 66*, 13–25.
Salas-Olmedo, M. H., Moya-Gómez, B., García-Palomares, J. C., & Gutiérrez, J. (2018).

Tourists' digital footprint in cities: Comparing Big Data sources. *Tourism Management, 66*, 13–25.
Shamma, D. A., Shamma, D. A., Friedland, G., Elizalde, B., Ni, K., Poland, D., et al. (2016). YFCC100M: The new data in multimedia research. *Communications of the ACM, 59*(2), 64–73.
Stepchenkova, S., & Zhan, F. Z. (2013). Visual destination images of Peru: Comparative content analysis of DMO and user-generated photography. *Tourism Management, 36*(3), 590–601.
Straumann, R. K., Çöltekin, A., & Andrienko, G. (2014). Towards (Re)constructing narratives from georeferenced photographs through visual analytics. *The Cartographic Journal, 51*(2), 152–165.
Sun, Y., Fan, H., Bakillah, M., & Zipf, A. (2015). Road-based travel recommendation using geo-tagged images. *Computers, Environment and Urban Systems, 53*, 110–122.
Su, S., Wan, C., Hu, Y., & Cai, Z. (2016). Characterizing geographical preferences of international tourists and the local influential factors in China using geo-tagged photos on social media. *Applied Geography, 73*, 26–37.
Valek, N. S., & Williams, R. B. (2018). One place, two perspectives: Destination image for tourists and nationals in Abu Dhabi. *Tourism Management Perspectives, 27*, 152–161.
Vecchio, P. D., Mele, G., Ndou, V., & Secundo, G. (2018). Creating value from social big data: Implications for smart tourism destinations. *Information Processing & Management, 54*(5), 847–860.
Vu, H. Q., Li, G., Law, R., & Ye, B. H. (2015). Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. *Tourism Management, 46*, 222–232.
Xiang, Z. (2018). From digitization to the age of acceleration: On information technology and tourism. *Tourism Management Perspectives, 25*, 147–150.
Yin, L., Cheng, Q., Wang, Z., & Shao, Z. (2015). 'Big data' for pedestrian volume: Exploring the use of Google Street View images for pedestrian counts. *Applied Geography, 63*, 337–345.
Zhou, L. (2014). Online rural destination images: Tourism and rurality. *Journal of Destination Marketing & Management, 3*(4), 227–240.
Zhou, X., Xu, C., & Kimmons, B. (2015). Detecting tourism destinations using scalable geospatial analysis based on cloud computing platform. *Computers, Environment and Urban Systems, 54*, 144–153.

Dr. Kun Zhang, is a lecturer in College of Tourism and Service Management at Nankai University. Her research interests are tourism planning, with the focus on tourism destination image and AI technology, tourists' demands and sustainable tourism planning.

Dr. Ye Chen, is a Professor in College of Tourism and Service Management at Nankai University. His research interests include: destination marketing, service marketing and consumer behavior. His research has been published in academic journals such as Tourism Management, Current Issues in Tourism, Total Quality Management & Business Excellence, Tourism Tribune, among others.



Chunlin Li, is an undergraduate student in College of Tourism and Service Management at Nankai University.