



Article

# Less-Known Tourist Attraction Discovery Based on Geo-Tagged Photographs

Jhih-Yu Lin <sup>1,\*</sup> , Shu-Mei Wen <sup>2</sup>, Masaharu Hirota <sup>3</sup> , Tetsuya Araki <sup>4</sup> and Hiroshi Ishikawa <sup>1</sup>

<sup>1</sup> Graduate School of Systems Design, Tokyo Metropolitan University, Hino-shi, Tokyo 191-0065, Japan; ishikawa-hiroshi@tmu.ac.jp

<sup>2</sup> Department of Statistics and Information Science in Applied Statistics, Fu Jen Catholic University, Taipei 24205, Taiwan; 126531@mail.fju.edu.tw

<sup>3</sup> Department of Information Science Faculty of Informatics, Okayama University of Science, Okayama 700-0005, Japan; hirota@mis.ous.ac.jp

<sup>4</sup> Graduate School of Science and Technology, Gunma University, Gunma 371-8510, Japan; tetsuya.araki@gunma-u.ac.jp

\* Correspondence: lin-jhihyu@ed.tmu.ac.jp

Received: 5 August 2020; Accepted: 16 October 2020; Published: 19 October 2020



**Abstract:** Most existing studies of tourist attraction recommendations have specifically emphasized analyses of popular sites. However, recommending such spots encourages crowds to flock there in large numbers, making tourists feel uncomfortable. Furthermore, some studies have discovered that quite a few tourists dislike crowded destinations and prefer to avoid them. A ready solution is discovery and publicity of less-known tourist attractions. Especially, this study specifically examines discovery of less-known Japanese tourist destinations that are attractive and merit increased visits. Using this approach, crowds can not only be dispersed from popular tourist attractions, but more diverse spots can be provided for travelers to choose from. By analyzing geo-tagged photographs on Flickr, we propose a formula that incorporates different aspects such as image quality assessment (IQA), comment sentiment, and tourist attraction popularity for ranking tourist attractions. We investigate Taiwanese and Japanese people to assess their familiar Japanese cities and remove them from ranking results of tourist attractions. The remaining spots are less-known tourist attractions. As reported from results of verification experiments, most less-known tourist attractions are known by only a few people. They appeal to participants. Additionally, we examined some factors that might affect respondents when they decide whether a spot is attractive to them or not. This study can benefit tourism industries worldwide in the process of discovering potential tourist attractions.

**Keywords:** point of interest; potential places; ranking formula; image quality assessment; geolocation; Japan; Flickr

## 1. Introduction

Along with development of information and communication technologies, almost all mobile devices have come to be equipped with global positioning system (GPS) sensors. The sensors and devices can support users by confirming their current position, but they can also annotate their photographs with “geo-tagging” on social networking services (SNSs). Numerous studies have analyzed “geo-tagged” photographs to elucidate user behaviors and preferences. Through that process, one can discover popular tourist attractions and can recommend some tour plans for users according to their preferences [1–4]. In addition to geolocation, diverse information such as comments and photographs are available from SNS users. That information includes important and useful data for research. For instance, Hausmann et al. [5] pointed out that social media contents might provide a swift

and cost-efficient substitute for traditional surveys. Ghermandi et al. [6] monitored and analyzed the activity of tourists to sites of environmental and historical importance by photographs with geo-tagging from Flickr. The result can help understand the spatial patterns of visitation and differences in how cultural benefits are accrued to various sectors of the population. Liu et al. [7] adopted an approach to discover areas of interest (AOIs) by analyzing geo-tagged photographs and check-in information to suggest popular scenic locations and popular spots for travelers. Another study with similar aims to those of the present study used SNS users' information and geo-tagged photographs to suggest obscure sightseeing locations [8]. Nevertheless, most earlier studies have specifically undertaken analyses of popular tourist attractions, points of interest (POIs) [9–15] or AOIs [16,17] while neglecting other unnoticed places.

Recently, tourism has become a development strategy for many countries because international tourism can generate enormous revenues; it can have positive long-run effects on higher economic growth. Several reports have explained that international tourism can bring benefits by boosting foreign exchange revenues, spurring investment in new infrastructure, stimulating other economic industries indirectly, and generating employment [18–23]. Although the numbers of tourists continue to increase, generating enormous revenues for tourism-associated industries, tourism benefits have been accompanied by negative effects such as overtourism and larger carbon footprints. Kakamu et al. [24] reported that crime rates and police forces increase when the numbers of foreign visitors increase. Rising crime rates reduce tourists' willingness to visit, thereby reducing tourism income [25]. Unfortunately, since the COVID-19 outbreak, the whole travel and tourism industry has been put on hold. Results of some studies show that COVID-19 has led to severe consequences for international tourism [26–29]. Achieving tourism recovery has become an exceedingly important task nowadays.

An important tourist phenomenon has been observed: most tourists receive sightseeing information through travel websites or SNSs. Nevertheless, almost all such sources present well-known tourist attractions. Consequently, although the attractions become crowded and congested, visitors will continue to be guided there. However, some studies have revealed that quite a few tourists dislike crowded destinations and prefer to avoid them [30–34]. Luque-Gil et al. [32] point out that this crowding situation can reduce people's satisfaction, attitude, and loyalty. Jacobsen [33] reported a source of negative traveler reactions from those crowded destinations. In addition, Yin et al. [34] described that physical crowding and human crowding have significantly negative impacts on destination attractiveness. Furthermore, they also indicated human crowding should be regarded as an important factor that negatively affects destination attractiveness. These conditions make it difficult to promote the further development of tourism industries. Therefore, we provide a novel approach to discover less-known but attractive tourist attractions, which might ameliorate the difficulties described above and which might support tourism to regions other than popular regions. Moreover, our earlier study [35] investigated Taiwanese and Japanese participants' preferences of tourist attractions. The results demonstrated that over half of Taiwanese and Japanese respondents are interested in less-known tourist attractions. In fact, less-known tourist attractions are a worthy issue to probe and discuss. They offer the potential of benefiting tourism industries worldwide.

This study specifically examines less-known "scenic" tourist attractions, which represent a clearly defined research scope because natural landscapes can make travelers realize beauty intuitively. To accomplish our aim, we analyze scenic geo-tagged photographs taken in Japan obtained from Flickr. Additionally, we clustered Japanese prefectures and cities via X-means based on their number of photographs. These clusters were used to survey unfamiliar clusters to Japanese and Taiwanese people. Our earlier research [36] revealed that participants have different preferences for scenic spots and that they truly care about photograph quality. These factors affect tourists as they decide whether a spot is attractive to them or not. Consequently, to provide more reliable results for tourists, we add image quality assessment (IQA) in this study and image classification to our research structure. To discover attractive less-known tourist attractions, the photograph quality was evaluated using IQA approaches. Using image classification techniques, scenic photographs are classified using nine labels such as

forests, oceans, and mountains. In this way, tourists can choose their tourist spot preferences easily. Finally, these geo-tagged photographs are ranked using our formula. The verification experiments specifically investigate Japanese people, Taiwanese people, and their differences.

The remainder of the paper is organized as follows: Section 2 introduces related work. Section 3 presents the methodology of discovering less-known tourist attractions. In Section 4, we illustrate less-known tourist attraction estimation and demonstrate the current results. Section 5 explains results of our verification experiments. Section 6 discusses our experiment results and the improvable aspects of this research. Section 7 interprets conclusions and future work.

## 2. Related Work

### 2.1. Points of Interest (POIs)

A point of interest (POI) is used with a technique positing a particular spot that someone might find useful or interesting. Such spots can be landmarks, sightseeing spots, or a commercial institution of any type such as a restaurant, a hospital, or a supermarket. Based on data types and discovery procedures, the approaches developed for POI are divisible into two types. The first type is top-down: discovery of POI from an existing POI repository or database, such as check-in data or yellow pages that are used frequently or which fit for a specific theme or target [9–11]. The second type is bottom-up: raw data (e.g., geo-tagged photos, digital footprints with implicit geographic information or metadata that involve latitude and longitude) to construct a new database or dataset that includes the POI [12–15]. Skovsgaard et al. [13] demonstrated a clustering technique that incorporates consideration of both spatial and textual attributes of microblog posts to obtain clusters that represent POI. Based on Flickr geo-tagged photographs, Kuo et al. [15] used pattern discovery, the spatial overlap (SO) algorithm, and the naming and merging method for attractive footprint clustering. From the peak value and range of clusters, the POI and region of interest (ROI) can be extracted, indicating the most popular location and range for appreciating attractions.

Many studies have combined a POI and a recommender system to provide various travel plans for tourists [37–40]. To recommend POIs for a given user at a specified time in a day, Yuan et al. [37] developed a collaborative recommendation model that is able to incorporate temporal information. Massimo et al. [38] presented a new recommender system technique for tourists' behavior learning and next-POI recommendations. The technique clusters users with similar POI visit trajectories and then learns a general user behavior model via inverse reinforcement learning (IRL).

Discovering new tourist attractions is an important task for tourism industry. Nevertheless, in accordance with our observation, those studies are only related to popular tourist attractions, but they neglect other places. Different from existing POI studies, this research provides a POI method for discovering less-known tourist attractions and specifically analyzes those unnoticed places which might include some attractive spots for tourists.

### 2.2. Image Quality Assessment (IQA)

Photograph is an important factor that affects tourists to make the decision about travel destinations, and also influenced their behaviors and reflected their satisfaction with tourism places [41,42]. Molina et al. [43] found out that the good quality photographs influenced tourist destination choice. On the other hand, in the previous research [36], we observed a phenomenon that most participants really cared about photographs' quality when evaluating the tourist attractions. Consequently, to discover appealing photographs, the image quality assessment is applied in this research.

Image quality assessment, an image processing technique, can use subjective and objective methods. Subjective methods rely on the intuitive appreciation of human observers for image attributes. Such methods are classifiable into two types: absolute evaluation and relative evaluation. Objective methods are based on computational models that can predict perceptual image quality.

They include three evaluative approaches: full-reference (FR), reduced-reference (RR), and no-reference (NR) approaches.

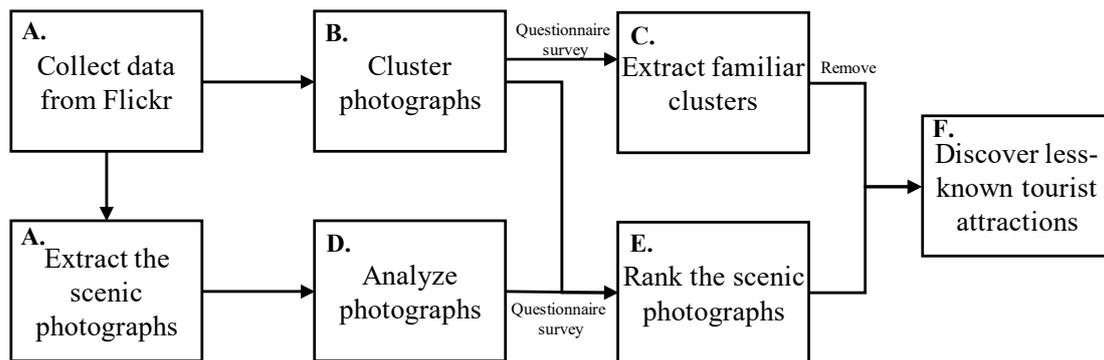
Numerous early studies have been conducted to automate NR-IQA to assess photograph quality using machine learning techniques. Most of these studies applied binary labels (“good” or “bad”) to assess image quality [44–49]. Although Dong et al. [50,51] developed a method to extend the image quality representation (“good”, “medium”, and “bad”), the results nevertheless leave great difficulty in ranking the images. Talebi et al. [52] proposed an approach called neural image assessment (NIMA), which differs from methods of other studies in that they predict the distribution of human opinion scores and assess techniques used for photography. Those IQA studies applied a large-scale database for aesthetic visual analysis (AVA) dataset [53] as their training data to machine learning model. The AVA dataset contains about 255,000 images, rated based on aesthetic qualities by different viewers (include amateur, professional, novice photographers, etc.). By using AVA dataset, those studies can train the model to classify the photographs into different levels or predict the scores of image quality. Especially, NIMA will be used to rank scenic photographs in this study and will be used to compare them with other ranking results to choose the best method for our study.

### 2.3. Image Classification

Image classification technique is used to discern the contents of images and classify these images into distinct categories or to assign a probability that the image is of a particular category. Traditional image classification is feature description and detection, which might be effective for some sample images, but the high dimensionality of the feature space is difficult to process in a factual situation. Recent studies applied machine learning techniques to create automatic image classification and to alleviate the shortcomings of traditional methods. This technique is widely applied in diverse fields such as medical field [54–57] and image quality assessment [44–49]. Raj et al. [56] improved classifier to recognize image of lung cancer, brain image, and Alzheimer’s disease for Internet of Medical Things (IoMT). In addition, Shankar Et al. [56] ameliorated the model to distinguish image of diabetic retinopathy. To find out guide-suitable pictures for improving the touristic experience, Kleinlein et al. [58] presented an approach to classify photographs into three labels base on aesthetic perception. Different from object detection, image classification only can annotate one label for the photograph. To our knowledge, in the tourism field, most research administered object detection to recognize the content of photographs for analyzing tourist photographs. However, considering nature scenes do not have fixed features (e.g., shape), it is hard to correctly annotate multiple labels for scenic photographs (which we used in this research) by object detection. Therefore, we used technique of image classification to simply classify scenic photographs into different types that include mountains, oceans, and nightscapes. By doing so, tourists can readily choose their preferences for natural landscapes.

## 3. Methodology

This section describes our proposed method for extracting less-known tourist attractions. Figure 1 presents an overview of our method. Our method is divided into five parts. Every step is explained in the subtask. For the first, we introduce the dataset for the research in Step A. In Step B, the dataset will be classified into distinct clusters by the number of photographs in each prefecture and city. Step C, these clusters will be used to investigate participants’ unfamiliar clusters. As Step D, we analyze and extract the positive comments of photographs. Subsequently the quality of photographs is evaluated using 5 IQA methods. Steps E and F are introduced in Section 4.



**Figure 1.** Overview of workflow. Red blocks represent updated parts used for this study.

### 3.1. Definition of Less-Known Tourist Attractions

To differentiate well-known and less-known tourist attractions, we adopt two definitions of less-known tourist attractions as the following:

- Definition 1: Only some people know about this tourist attraction.
- Definition 2: The tourist attraction is attractive for tourists and deserves to be visited.

On top of that, if tourists could view well-known landscape from a certain place, but only a few people know about this place, that is also regarded as a less-known tourist attraction. Moreover, we then assume that the less-known tourist attractions might be included in unfamiliar cities of tourists.

### 3.2. Data Collection and Extract the Scenic Photographs (Step A)

Using Flickr API, 769,749 photographs taken in 2017 at geolocations throughout Japan were collected. To obtain full addresses, we apply geocoding to photograph latitude and longitude using Google Geocoding API. Nevertheless, 309 photographs have no details of addresses because these photographs were taken on the ocean. As a result, our dataset includes 769,440 photographs. Additionally, we collected the information of these photographs such as comments, numbers of views, and numbers of favorites they earned. To swiftly filter the scenic photographs, the photographs with the tags that related to scenic descriptions in English, Japanese, and Chinese (e.g., “scenery”) are extracted from the dataset. Further, we manually sifted the inappropriate photographs which are not related to the nature scene and extracted 1159 scenic photographs as second dataset for this research. The content of scenic photographs includes over 80 percent natural scene without human.

Subsequently, these photographs are classified into different prefectures and cities according to the full address of photographs. Later, we calculated the numbers of photographs of 47 prefectures and 1158 cities (those cities include special wards). Table 1 presents the top 10 prefectures and cities in terms of the number of photographs. Figure 2 shows the distribution of photographs for Special ward of Tokyo. In Figure 2, we can realize that most photographs are shoot in Shibuya and Shinjuku where are popular spots.

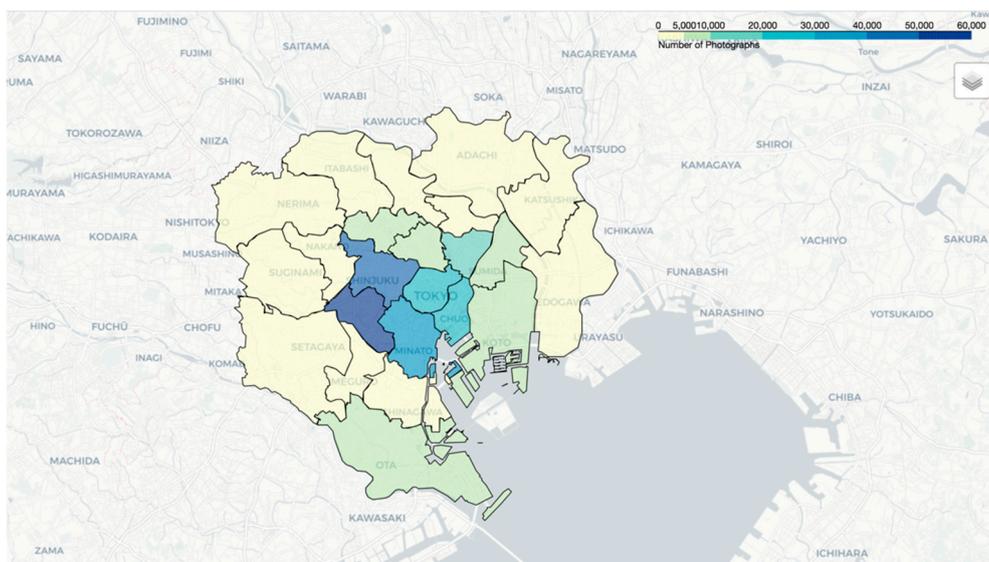
### 3.3. Clustering Prefectures and Cities (Step B)

For this study, the less-known tourist attractions are assumed to exist in cities that are unfamiliar to tourists. Moreover, people from different countries have distinct familiarity with Japanese cities. To ascertain and compare residents and foreign visitors’ unfamiliarity with Japanese cities, we intend to conduct a questionnaire to investigate Japanese and Taiwanese. However, surveying the degrees of familiarity for each city (1158 cities) from respondents was difficult. For that reason, to reduce the respondent burden, X-means was used to cluster prefectures and cities to administer the questionnaire survey easily in this step.

The X-means algorithm is a clustering technique presented by Pelleg and Moore [59] to improve the shortcomings of K-means. Moreover, X-means algorithm can determine the optimum number of clusters automatically from a user setting of only the minimum and maximum of clusters. Here, we refer to the results of elbow method to set the minimum of clusters. Additionally, this approach greatly reduces the probability of being trapped into a local optimum. Considering the outliers existing in the data, we used this method to distribute the prefectures and cities into different clusters based on their respective characteristics. For the features of X-means, we adopted the number of photographs in each prefecture as the most appropriate feature for analyzing the less-known level of prefectures in current work. Furthermore, the four features are applied for cities' cluster: the number of photographs in each city, the rate of number of photographs in each city, the rate of number of photographs in each prefecture and the average of the number of photographs in each prefecture.

**Table 1.** Top 10 Japanese prefectures and cities for the number of photographs.

Rank	Prefecture	Number of Photographs	City	Number of Photographs
1	Tokyo	274,530	Tokyo-Shibuya	56,911
2	Kyoto	55,747	Tokyo-Shinjuku	43,788
3	Chiba	41,279	Kyoto-Kyoto	43,300
4	Kanagawa	35,345	Tokyo-Minato	32,311
5	Aichi	33,356	Aichi-Nagoya	29,746
6	Osaka	26,567	Tokyo-Chiyoda	21,717
7	Hiroshima	22,560	Osaka-Osaka	20,396
8	Hokkaido	22,222	Tokyo-Chuo	20,164
9	Saitama	20,381	Tokyo-Taito	16,942
10	Gunma	17,369	Hiroshima-Hiroshima	14,096

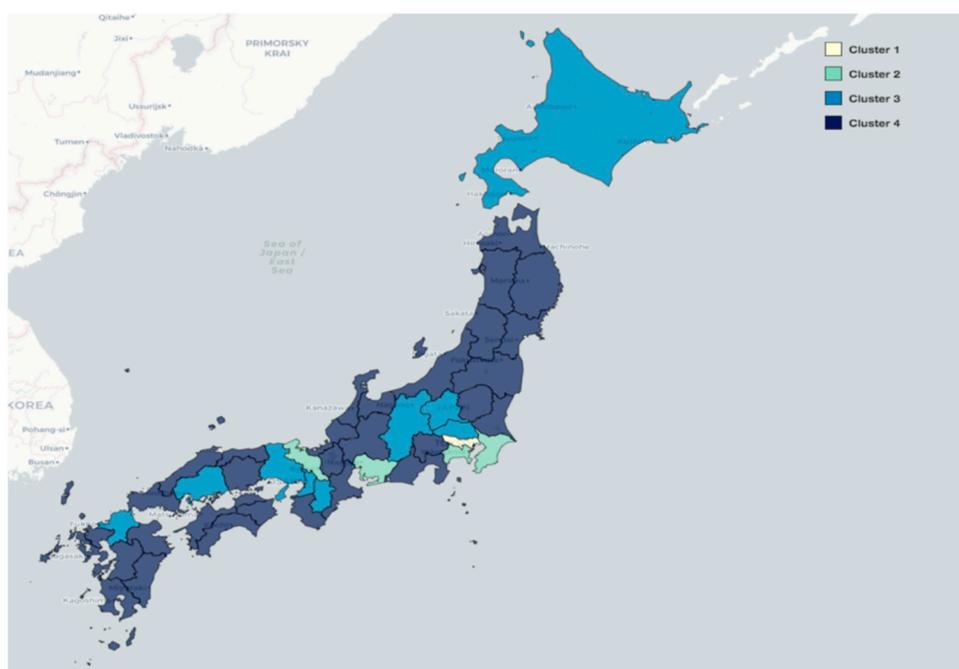


**Figure 2.** Distribution of photographs for Special ward of Tokyo.

The 47 prefectures are clustered into four clusters, as shown in Table 2 and Figure 3. The 1158 cities are distributed into 14 clusters. In Table 2, the third column represents the score of each cluster which we defined in this step. These scores will be used in our developed formula. In addition, the clustering result roughly matches the distribution of the population in Japanese prefectures. Which means that the clustering results of cities might have sufficient validity. Furthermore, the city cluster score is defined according to questionnaire survey responses, as explained in Section 3.4.

**Table 2.** Clustering result of prefectures.

Cluster	Prefectures	Score
Cluster 1	Tokyo	4
Cluster 2	Kyoto, Chiba, Kanagawa, Aichi	3
Cluster 3	Osaka, Hiroshima, Hokkaido, Saitama, Gunma, Nara, Nagano, Okinawa, Hyogo, Fukuoka	2
Cluster 4	Mie, Tochigi, Shizuoka, Yamanashi, Oita, Okayama, Ibaraki, Aomori, Miyagi, Gifu, Ishikawa, Wakayama, Kagawa, Niigata, Shiga, Ehime, Kumamoto, Akita, Toyama, Fukushima, Nagasaki, Yamagata, Kagoshima, Tottori, Saga, Fukui, Tokushima, Kochi, Yamaguchi, Iwate, Shimane, Miyazaki	1

**Figure 3.** Distribution of prefecture clusters.

### 3.4. Evaluating Familiarity of City Clusters (Step C)

As described in this section, we administered an online questionnaire survey to elicit information from foreign visitor (115 Taiwanese) and local residents (123 Japanese): their degrees of familiarity with Japanese cities. The reason why we invite Taiwanese as our participants is according to the news (<https://www.nippon.com/en/japan-data/h00375/overseas-visitors-to-japan-in-2018-top-31-million.html>), Taiwan was reported as the third place in the “Top 20 Countries/Regions by Number of Visitors”. However, this ranking did not consider the population in each country. Considering the population, Taiwan will be the first place in the average of each person visiting Japan. Thus, Taiwanese are the most suitable participants for this research.

For this research, all of participants meet five requirements as follow,

1. Have travel experience in Japan.
2. Preferring natural tourism.
3. They do not mind visiting unknown places.
4. Taiwanese who has the economic ability for overseas travel.
5. They only can participate in the questionnaire survey for one time.

According to the number of cities in each cluster, 30 city names were selected randomly for the questionnaire. Additionally, the five options are provided for participants to select, with higher scores indicating greater familiarity with this city. The question is as follows,

Do you know this “random name of city”? (point 1–5)

- (1) I totally have no idea.
- (2) I have heard of this city, but I don’t know the relevant tourist attractions.
- (3) I have heard of this city and know the relevant tourist attractions.
- (4) I have been to this city, but I don’t know the relevant tourist attractions.
- (5) I have been to this city and know the relevant tourist attractions.

Afterward, the familiar clusters of participants are extracted. Then these clusters are removed from the final ranking result. Tables 3 and 4 show the average scores of respective clusters, which imply that local residents and foreign visitors have different degrees of familiarity with Japanese cities. Furthermore, the average scores of city clusters will be applied to our developed formula.

**Table 3.** Taiwanese unfamiliar city clusters.

Cluster	Sample Mean	<i>t</i> -Test Value	<i>p</i> -Value	Unfamiliar
Cluster 1	1.63	−1.40	0.08	
Cluster 2	1.89	2.05	0.98	
Cluster 3	1.40	−7.63	0.00	√
Cluster 4	3.74	18.07	1.00	
Cluster 5	2.90	8.55	1.00	
Cluster 6	1.70	0.12	0.55	
Cluster 7	1.56	−2.50	0.01	√
Cluster 8	2.55	6.49	1.00	
Cluster 9	1.74	0.62	0.73	
Cluster 10	1.69	−0.06	0.47	
Cluster 11	1.59	−1.11	0.13	
Cluster 12	1.52	−2.82	0.00	√
Cluster 13	1.37	−7.22	0.00	√
Cluster 14	1.31	−11.33	0.00	√
Population mean	1.69	−	−	−

**Table 4.** Japanese unfamiliar city clusters.

Cluster	Sample Mean	<i>t</i> -Test Value	<i>p</i> -Value	Unfamiliar
Cluster 1	2.02	0.22	0.59	
Cluster 2	2.96	7.14	1.00	
Cluster 3	1.61	−7.24	0.00	√
Cluster 4	4.47	29.79	1.00	
Cluster 5	3.51	10.23	1.00	
Cluster 6	1.80	−3.53	0.00	√
Cluster 7	1.47	−11.01	0.00	√
Cluster 8	3.24	9.55	1.00	
Cluster 9	2.32	3.46	1.00	
Cluster 10	2.22	2.28	0.99	
Cluster 11	1.59	−4.42	0.00	√
Cluster 12	2.09	0.88	0.81	
Cluster 13	1.64	−6.50	0.00	√
Cluster 14	1.40	−15.37	0.00	√
Population mean	2.01	−	−	−

Subsequently, to extract the unfamiliar clusters of participants, we use statistical methods and assumed that half of the participants are unfamiliar with the cluster when the sample means of

the cluster are less than the population means. Considering that we used the survey sampling approach to conduct the questionnaire survey, it might include sampling error. To decrease the inaccuracy from the sampling error, we categorized the cluster as a less-known one using  $t$ -tests and  $p$ -values. After calculating the  $t$ -test values, we used the  $p$ -value to ascertain whether the sample mean was greater than the population mean or not. If the  $p$ -value of cluster was less than 0.05, then we inferred this cluster as an unfamiliar cluster. Conversely, the cluster will be categorized into familiar clusters when the  $p$ -value is greater than 0.05.

Tables 3 and 4 present results of application of  $t$ -tests, with  $p$ -values obtained for the respective clusters. Table 3 presents that five clusters were regarded as unfamiliar by Taiwanese people. Moreover, Table 4 shows that six clusters were regarded as unfamiliar by Japanese people.

### 3.5. Analysis of Photograph (Step D)

The attractive spots include various factors that affect decision-making about travel destinations of tourists. To find the attractive tourist attractions and enhance the formula we propose, we specifically analyze comment sentiment of the photographs and the photograph quality in Step D.

#### 3.5.1. Analysis of Comment Sentiment

To discover attractive tourist attractions from unfamiliar areas, positive comments about photographs are assumed to be a factor affecting whether this sightseeing spot is attractive for tourists to visit or not. Therefore, we extract the comments about scenic photographs collected in Section 3.2 through Web crawler. Additionally, comments written by the photograph owner are removed because almost all of these comments are merely responses to the viewer comments. For this study, we specifically examined English, Chinese, and Japanese comments using Google natural language API, which yielded a score of sentiment representing the probability of positive meaning. In this way, one can detect whether the sentiment of comments is positive or not. Table 5 presents the results of applying Google natural language API and the number of positive comments in each language.

**Table 5.** Result of positive comment.

	English	Chinese	Japanese	Total
<b>Viewer Positive Comment</b>	828	413	157	1398
<b>Owner Positive Comment</b>	112	216	16	344
<b>Total</b>	940	629	173	1742

#### 3.5.2. Evaluation of Photograph Quality

In previous investigation [36], we observed that most participants really care about the quality of photographs that affects tourists as they decide whether a spot is attractive to them or not. Thereby, photograph quality, as judged by attributes such as aesthetics and composition, is important for evaluating the attractiveness of sightseeing spots. To assess photograph quality, we discussed five approaches using heuristic and image processing methods, as described below. Subsequently, we administered the questionnaire survey to ascertain the best parameter for improving the performance of our proposed formula (Figure 4).

1. Method 1 (number of favorites): Users of Flickr can collect their favorite photographs. Then Flickr counts how many users like this photograph. We infer that a photograph with a higher number of favorites indicates high quality.
2. Method 2 (number of views): Flickr counts views for each photograph. We consider that a photograph with a higher number of views represents a strong interest of other users. The higher number of views implies that this photograph might have high quality.

3. Method 3 (followers of photographers): Presumably, a user with many followers tends to post high-quality photographs. For method 3, we collect 6361 photographers’ information (such as the number of followers, number of photographs, and the year they joined Flickr) from our dataset. Considering that the year of joining Flickr influences the number of followers of photographers, our presumption might be unfair to Flickr novices. Therefore, we calculate the average annual followers of users up to the end of 2018 (Table 6). Then we ranked the scenic photographs through this information (In Flickr, if the number of followers is greater than 1000, then the value will become “1K”. We cannot ascertain details of the numbers of followers in such cases. For that reason, in Table 5, the “K” of the followers is changed to “1000”). In addition, the photographers’ works, including only one photograph were chosen as representative based on the number of favorites
4. Method 4 and Method 5 (aesthetics and technique of photographs): Using the fourth and fifth methods, we adopt a method proposed by Talebi et al. [46]. They presented a novel approach called neural image assessment (NIMA), which can predict both technical and aesthetic qualities of photographs. Using the model of NIMA, the aesthetic and technique of 1159 scenic photographs can be evaluated as shown in Figure 5.

Table 6. Photographer information components.

Rank	Owner Name	Followers	Number of Photos	Joined	Average Followers
1	141****1@N05	62,100	6000	2016	31,050
2	54****5@N06	44,100	44	2010	5512
3	25****4@N05	54,200	516	2008	5420
4	59****3@N07	37,400	72,300	2011	5343
5	60****3@N06	31,700	137	2011	4529

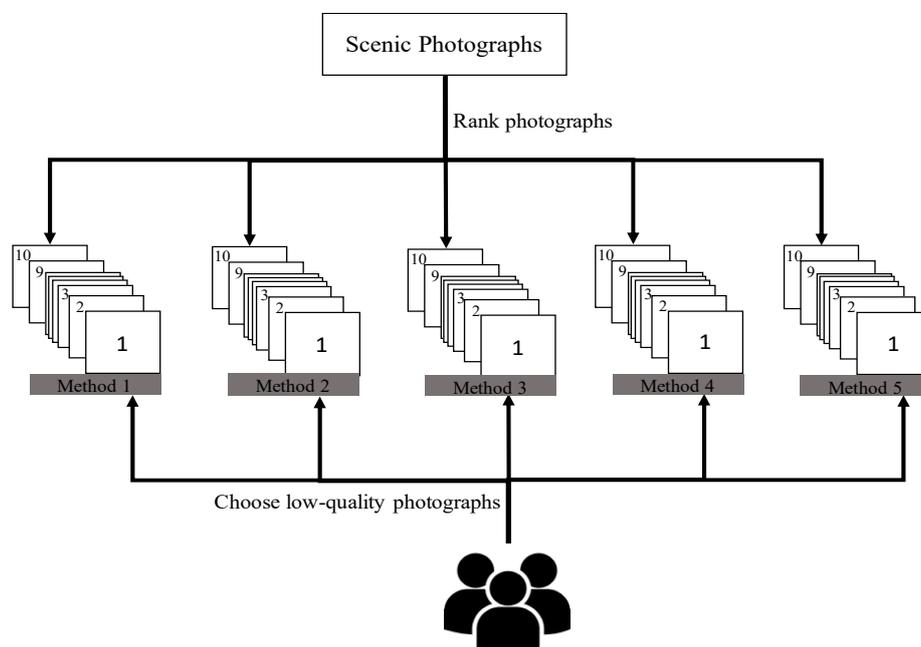


Figure 4. Schema of questionnaire survey.

The top 10 photographs obtained using each method are extracted for the questionnaire survey. For this questionnaire survey, 50 participants (25 Japanese people and 25 Taiwanese people) who meet five requirements (as explain in Section 3.4) were asked to select, intuitively, those photographs having normal or below-normal quality. If the method yields many low-quality photographs, then it shows agreement with human perception. Table 7 presents the questionnaire survey results.

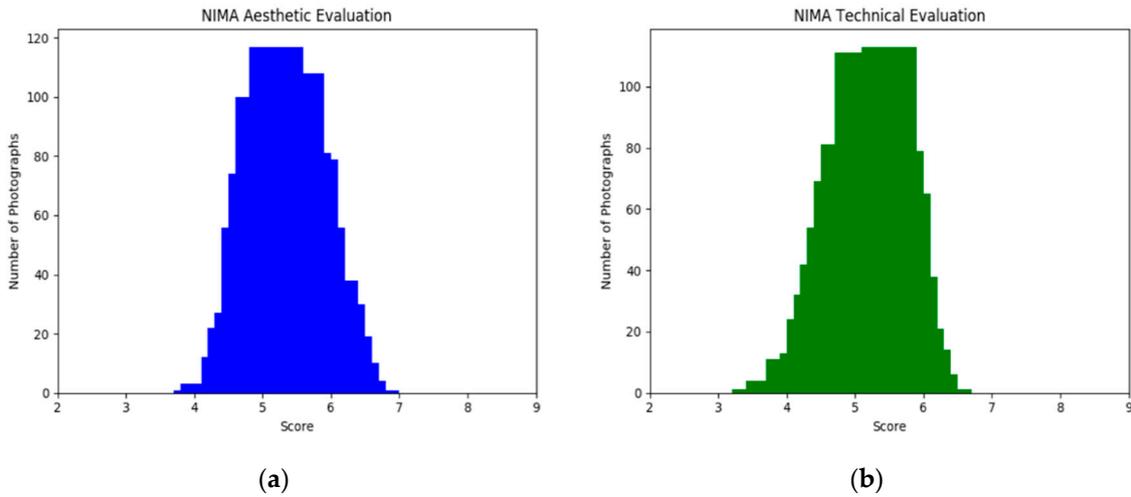


Figure 5. Distribution of NIMA result: (a) aesthetics and (b) techniques.

Table 7. Voting results for low-quality photographs obtained using several methods.

	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10	Total
Method 1	3	16	9	14	14	3	5	6	7	8	85
Method 2	2	3	1	1	7	4	1	11	0	9	39
Method 3	14	4	11	3	15	5	5	11	4	4	76
Method 4	4	9	6	2	3	3	9	2	23	2	63
Method 5	12	21	6	22	27	28	16	16	20	4	172

In Table 7, the first row presents the rankings of the respective photographs. The first column shows the name of each method. The last column represents how many votes the method received for the top 10 photographs. When regarding this table, one can realize how many people vote the photograph as the low-quality photograph. Method 2 (the number of views) received the lowest number of votes in this questionnaire, which means that this method is the most applicable approach to our research. Subsequently, the number of photograph views is used in our developed formula (Section 4.1). All the tourist attractions can be evaluated and ranked. Finally, the participants’ familiar clusters are removed from the ranking results. The remaining spots are the less-known tourist attractions, which is our goal.

#### 4. Rank the Scenic Photographs

Step E and Step F of our workflow will be presented in this section. Combining the result of Section 3, we propose a formula to assess the score of photographs. Finally, the participants’ familiar clusters will be removed from the ranking result. The remaining spots are less-known tourist attractions, which are our goal.

##### 4.1. Evaluation of Formula (Step E)

Considering the definitions of less-known tourist attractions and data construction, we propose a formula to rank the photographs. Using this formula, we can calculate the score of photographs for ranking.

$$\sum_{p=1}^3 F_{pi}W_p + R_i, 0 < W_p < 1 \text{ and } \sum_{p=1}^3 W_p = 1 \tag{1}$$

In Equation (1), the following variables are used:  $i$  represents each photograph;  $F_{1i}$  and  $F_{2i}$  respectively express the cognitive scores of Japanese prefectures and cities, as defined in

Sections 3.3 and 3.4, with weights  $W_1$  and  $W_2$  are their weights;  $F_{3i}$  stands for the number of view in each photograph (describe in Section 3.5.2);  $W_3$  is the  $F_{3i}$  weight;  $R_i$  represents the number of positive comment of the photographs; and  $W_3$  and  $R_i$  are processed by feature scaling. Particularly,  $R_i$  stands for an additional point in that we obtain the weight of  $R_i$  as almost equal to 0 by entropy weight method (EWM). The reason is that most photographs have no associated comments. However, before visiting tourist attractions, most tourists refer to related comments and information. They then decide whether to go there, or not. Therefore, we presume  $R_i$  as a necessary parameter because the positive comments might affect the perspectives of the other viewers. In this formula, the quality of photographs and positive comments were assumed as factors attracting someone to visit.

For the weight of Equation (1), we must set optimal weights for each parameter, but we do not know the importance of the respective parameters. Therefore, we applied EWM to calculate the optimal weights. Because it depends solely on the discreteness of data, EWM is an objective set weight method. Actually, EWM is used widely in the fields of engineering, socioeconomic studies, etc., [60–62]. In information theory, entropy is a kind of uncertainty measure. When information is greater, uncertainty and entropy are smaller. Based on entropy information properties, one can estimate the randomness of an event and can estimate the degree of randomness through calculation of the entropy value. Furthermore, entropy values are used to gauge a sort of degree of discreteness for an index. When the degree of discreteness is larger, the index affecting the integrated assessment is expected to be greater.

To complete the setting of the formula weights, we require the steps presented below.

1. Calculate the ratio ( $P_{ij}$ ) of the  $i$ -th index under the  $j$ -th index. Therein,  $x_{ij}$  denotes the  $j$ -th index of the  $i$ -th sample.

$$P_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}}, (i = 1, \dots, n; j = 1, \dots, m) \quad (2)$$

2. Calculate the entropy value ( $e_j$ ) of the  $j$ -th index as shown below.

$$e_j = -k \sum_{i=1}^n P_{ij} \ln(P_{ij}), (j = 1, \dots, m; k = \frac{1}{\ln(n)} > 0) \quad (3)$$

3. Calculate the discrepancy of information entropy ( $d_j$ ).

$$d_j = 1 - e_j, (j = 1, \dots, m) \quad (4)$$

4. Calculate the weight ( $w_i$ ) of each index.

$$w_i = \frac{d_j}{\sum_{j=1}^m d_j}, (j = 1, \dots, m) \quad (5)$$

The prefecture cluster score ( $F_{1i}$ ), the city cluster score ( $F_{2i}$ ), and the number of views ( $F_{3i}$ ) of 1159 scenic photographs are used to calculate the weight of the formula by EWM. The weight results are presented in Table 8. The Taiwanese  $W_1$  is equal to 0.2554,  $W_2$  is equal to 0.2559, and  $W_3$  is equal to 0.4887. Additionally, the Japanese  $W_1$  is equal to 0.2725,  $W_2$  is equal to 0.2061, and  $W_3$  is equal to 0.5214. In Equation (1), the Taiwanese and Japanese weights differ in that their city clusters are assigned distinct scores based on questionnaire survey results, which affect all weights of parameters.

**Table 8.** Taiwanese and Japanese weights.

	$W_1$	$W_2$	$W_3$
Taiwanese weight	0.2554	0.2559	0.4887
Japanese weight	0.2725	0.2061	0.5214

#### 4.2. Current Result

Using this formula, all scenic photographs can be ranked; then Taiwanese and Japanese familiar city clusters (defined in Section 3.4) are removed from the ranking result. The remnant spots are less-known tourist attractions, as shown in Tables 9 and 10. Tables 9 and 10 present some Taiwanese and Japanese ranking results. The second and third columns are photograph cluster scores, as defined in Sections 3.3 and 3.4. The fourth column is the photographs' number of views collected from Flickr; the results of 5 IQA surveys are shown in Section 3.5.2. The fifth column shows positive comments about photographs, which are defined in Section 3.5.1. The last column presents the scores of places, as calculated using our formula. Particularly, before calculating the scores of places, the number of views and numbers of positive comments are processed by feature scaling. High scores are associated with places that might be attractive to travelers. Comparison of these results indicates great differences between Taiwanese and Japanese results; the differences of levels of results are distinct.

**Table 9.** Top 5 Taiwanese ranking results.

Address	Prefecture Score	Score of City Cluster (Taiwan)	Views	Positive Comments	Score
Kendou 388 sen, Inuma, Kawanehon-cho Haibara-gun, Shizuoka, 428-0402, Japan	1	1.31	3762	46	1.49
Narukodamu, Narukoonseniwanobu2-8, Osaki Shi, Miyagi Ken, 989-6100, Japan	1	1.37	3721	41	1.42
Kendou32sen, Yotsuya, Shinshiro Shi, Aichi Ken, 441-1942, Japan	3	1.56	4370	2	1.28
Kokudou134sen, KosigoeItyoume, Koshigoe, Kamakura Shi, Kanagawa Ken, 248-0033, Japan	2	1.52	2531	1	1.21
Motosu-michi, Minamikoma Gun Minobu Cho, Yamanashi Ken, 409-2401, Japan	1	1.4	3018	27	1.16

**Table 10.** Top 5 Japanese ranking results.

Address	Prefecture Score	Score of City Cluster (Japan)	Views	Positive Comments	Score
Bi-chisaidokondominiamuIII, 2-16-1, Chatan, Nakagami Gun Chatan Cho, Okinawa Ken, 904-0116, Japan	2	1.8	8251	32	1.69
Kendou 388 sen, Inuma, Kawanehon-cho Haibara-gun, Shizuoka, 428-0402, Japan	1	1.4	3762	46	1.49
510, Tangocho Takano, Kyotango Shi, Kyoto Fu, 627-0221, Japan	3	1.59	10,394	4	1.47
Narukodamu, Narukoonseniwanobu2-8, Osaki Shi, Miyagi Ken, 989-6100, Japan	1	1.64	3721	41	1.42
Minamihanda53, Shirasu, Soraku Gun Wazuka Cho, Kyoto Fu, 619-1222, Japan	3	1.59	3430	9	1.4

#### 5. Verification Experiment

This section describes the verification experiment design and the experimentally obtained results. The results verify that the proposed method is reliable. First, our earlier experiment [36] revealed that participant preferences affect tourists as they decide whether a spot is attractive to them, or not. Thereby, we use image classification to categorize scenic photographs. Afterward, based on image classification results, we design three questionnaires from which participants can choose. In the last subsection, we present the questionnaire results and discuss them.

### 5.1. Image Classification

Considering the diverse preferences of various tourists, scenic photographs can be categorized into different labels using image classification. Subsequently, tourists can choose their favorite type of tourist attraction rapidly. With a view to building the image classifier model, we adopt the technique of transfer learning to retrain the Inception-v3 [63] model, which can save much time in training the model. Some parameters that Inception has already learned can be reused. We can build a highly accurate classifier using fewer training data. The Inception-v3 model is a convolutional neural network trained on more than a million images from the ImageNet. It has learned rich feature representations for widely diverse images. Moreover, Inception-v3 can identify images with 1000 object categories such as animals, vegetation, and landscapes.

For this step, based on the contents of 1159 scenic photographs, we defined nine labels to assign to these photographs. Especially, those labels include nightscape and snow in regard to the fact that the model of image classification is difficult to distinguish the dim photographs; likewise, some places covered with snow are also hard to identify the content of photographs. Considering that some scenic photographs were taken during the night/evening (when their subjects include the starry sky and evening seaside), nightscape and snow should be added to the label list.

Furthermore, 14,662 images were collected from Flickr and Google as our training dataset. Of the data, 10% were used to test the model. The remaining data were used to train the model. After 9000 training steps, the training accuracy of our model achieved 0.85, with validation accuracy of 0.83. Using this image classifier model, the 1159 scenic photographs are classifiable into distinct labels, as presented in Table 11.

**Table 11.** Image classification results.

Label	Number of Photographs	Label	Number of Photographs
Lake	80	Forest	226
River	136	Farmland	156
Ocean	163	Snow	82
Mountain	218	Nightscape	34
Flower, Grass	64	Total	1159

### 5.2. Questionnaire Design

To verify the validity of our approach that discovers the less-known tourist attractions, we design the questionnaires for Taiwanese and Japanese participants in this section. Although full addresses of less-known tourist attractions are known, the cognitive levels of less-known tourist attractions are complex and difficult to delimit. Furthermore, in the previous study [35], most participants reported that except the address of their home and company, it is difficult for remembering other places' addresses. Therefore, our verification experiment must specifically address the recognition of Japanese cities (in which less-known tourist attractions exist) and provide photographs of less-known tourist attractions from these cities to respondents. The 10 questions are extracted for Japanese cities from Taiwanese and Japanese ranking results of less-known tourist attractions respectively. Since their ranking results are different, the distinct contents of questionnaires are provided for them. However, each label has no more than 10 Japanese cities because no city will contain all labels of scenes. For better calculation, less-known tourist attractions are classified into three categories based on feedback from earlier results [36]. Each category includes labels of similar properties: Category 1 includes mountains, forest, flowers, grass, and farmland; Category 2 includes oceans, rivers, and lakes; the third category is a composite, comprising category 1, category 2, snow and nightscapes. In this way, the questionnaire can be administered and analyzed easily.

Results demonstrate that 19 Taiwanese people (all of them are office workers with average age of 27) and 22 Japanese people (most engineering students with average age of 25) were recruited for the questionnaire survey. All of the participants meet five requirements as explain in Section 3.4. None

participated in an earlier questionnaire survey. They were instructed to choose their preference of category which engenders dissimilar questionnaire contents.

Two questions were asked for each Japanese city. Scenic photographs were provided for the respondents' reference.

1. Do you know this city? (Yes/No)
2. According to these scenic photographs, do you want to visit this place of city? (1–5)
  - (1) Strongly do not want to visit.
  - (2) Somewhat do not want to visit.
  - (3) Neutral.
  - (4) Somewhat want to visit
  - (5) Strongly want to visit

For the first question, if participants probably knew the city, then the answer was "Yes". For the second question, respondents were instructed to assign a score of 1–5 to the attraction's photographs.

Table 12 presents the preferred categories of respondents. In this experiment, most participants selected category 3 as their preference. Category 1 and category 2 were chosen by five people each.

**Table 12.** Preferences of participants.

	Category 1	Category 2	Category 3	Total Participants
Taiwan	5	5	9	19
Japan	5	5	12	22

### 5.3. Verification Experiment Results

Results obtained for categories 1–3 are explained in this subsection. In Figures 6–8, each point represents a Japanese city in the questionnaire. The  $x$ -axis shows what percentages of respondents know the Japanese city. The  $y$ -axis shows the attraction level of less-known tourist attractions. In these scatter plots, some points are overlapping because these places were assigned the same estimation. For example, if the two places were known by 20% of respondents and if the attraction levels of these places were equal, then their points would overlap in the scatter plot.

Figure 6 indicated results obtained for category 1. One can infer that these places are known by only a few people. The Taiwanese result presents the average scores of three places as greater than four points; scores of one of these places approach the full mark, meaning that these places are attractive for Taiwanese respondents. By contrast, for the place with the lowest score, the scenic photographs show scenery similar to that in their own country. As a result, the respondents assigned few points to this place. For the result obtained for Japanese people, the average scores of four places are more than four points. However, a few places are assigned a low score because of the fact that the scenery is common in Japan. Especially, one place is known by 80% of people because that city is close to Tokyo, where the respondents live. The respondents are familiar with this city.

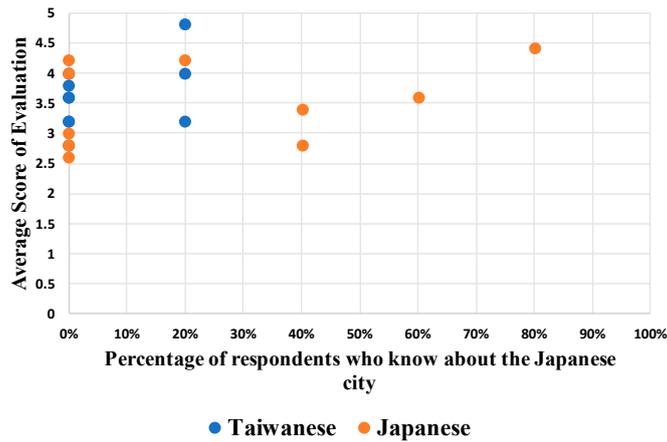


Figure 6. Evaluation result: Category 1.

Figure 7 depicts results obtained for category 2. The Taiwanese result shows that only the average scores of three places are less than four. Moreover, one place score approaches full marks. Nevertheless, no one knows about this place. In other words, most category 2 places are known by only a few respondents, but the place appeals to them. For Japanese results, the average scores of four places are greater than four points. Particularly, one place is known by 80% of respondents. The reason is the same as that in the case of category 1. Furthermore, we investigated the answers of the respective Japanese respondents deeply and detected that the disparity between their decisions decreased the average. Additionally, we observed an interesting phenomenon: one place was assigned greatly different scores by Taiwanese and Japanese respondents. For the Japanese evaluation, this place is estimated as having the lowest score, but Taiwanese respondents assigned this place over four points. This situation expresses that the evaluation of less-known tourist attractions is subjective for respondents.

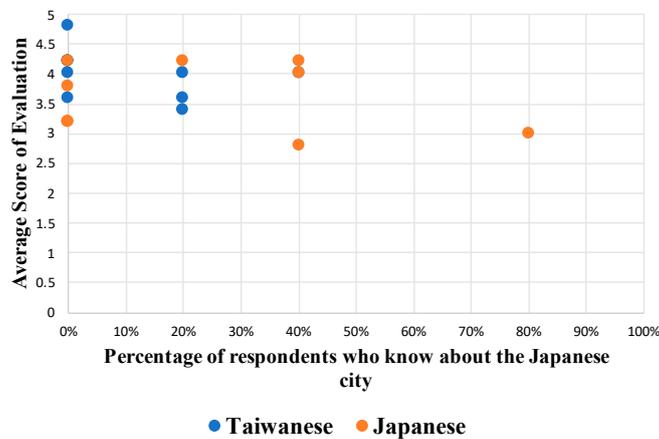


Figure 7. Evaluation result: Category 2.

Regarding the result obtained for category 3 (Figure 8), although the average score of only one place is over four points, the average scores of other places are over 3.5 points, indicating that Taiwanese respondents are not excluded from visiting these spots. Additionally, we detected that contents of photographs with the highest scores included snowscapes and Japanese castles, which are scarce in Taiwan. Taiwanese respondents reported that some places seem difficult to reach, which might influence their decision. The required cost of Taiwanese includes a monetary cost and time cost, which are higher than those of Japanese people. Consequently, Taiwanese prefer to choose tourist attractions that include local characteristics or exceptional landscapes. For Japanese results, we obtained the surprising result that the average scores of eight places are over four points: Japanese respondents are very satisfied with the less-known tourist attractions of category 3. The place with the lowest score

is a view of snow-covered mountains. Japanese respondents think this place looks very chilly and report that there is nothing nearby.

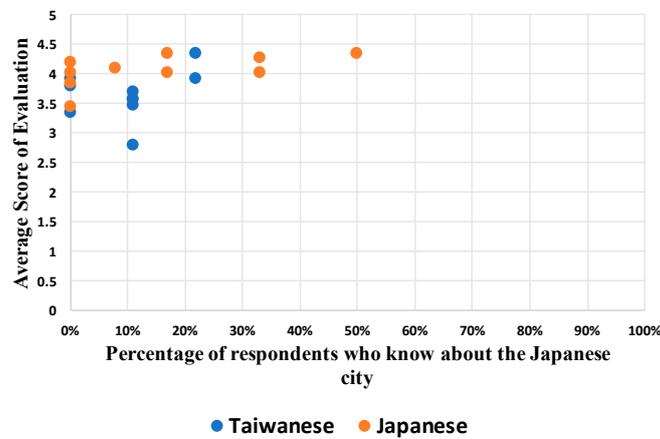


Figure 8. Evaluation result: Category 3.

In summary, this verification experiment indicates that most of these places are known by a few people, but the evaluation of less-known tourist attractions is the objective for respondents. Although two cities are known by most respondents, they do not know the details of the locations of scenic photographs. Furthermore, this experiment demonstrates that local residents and foreign visitors differ greatly in their evaluation of less-known tourist attractions. Table 13 shows that we organize the answers of Taiwanese and Japanese respondents for what percentages of people want to visit the less-known tourist attractions (who assign more than four points for the place). Table 13 presents the responses to 10 questions of each category. Then we can realize that almost all of these places are sufficiently attractive for someone to visit. Moreover, more than half of respondents are interested in the lesser-known tourist attractions that we provided. We discovered less-known tourist attractions that are attractive to some people.

Table 13. Percentage of respondents who want to visit the less-known tourist attractions.

	%	City 1	City 2	City 3	City 4	City 5	City 6	City 7	City 8	City 9	City 10
Taiwan	Cat. 1	20	40	60	40	60	100	60	80	20	60
	Cat. 2	80	80	100	60	80	40	60	80	60	40
	Cat. 3	22	33	100	67	67	67	67	33	67	33
Japan	Cat. 1	20	80	80	40	80	60	40	0	80	20
	Cat. 2	40	100	80	100	20	60	80	40	40	40
	Cat. 3	75	92	83	83	75	67	83	92	75	42

## 6. Discussion

Since the less-known tourist attractions are assumed might be included in unfamiliar cities of tourists, we conducted the questionnaire survey to understand tourists’ cognitive level of Japanese cities. However, people from different country might have various perspectives with Japanese cities. Thereby, to compare local residents and foreign visitors’ difference, Taiwanese and Japanese are invited to participate in the questionnaire survey. It is interesting to note that in this survey, interviews of some Taiwanese participants to ascertain what factors lead them to prefer to travel in Japan indicated four main reasons which are attractive to Taiwanese. The first reason is that air fare is cheaper and the flight time is short. The second reason is that the Japanese environment is neat and tidy. Furthermore, public security is high. The third reason is that Japanese food is delicious and exquisite. The fourth reason is that Japanese language characters and culture are similar to those of Taiwan, which can help Taiwanese people travel easily in Japan.

Considering that in a previous investigation [36], most participants really cared about the quality of photographs which affects tourists as they decide whether a spot is attractive to them or not. Hence, to find out attractive less-known tourist attractions and strengthen the formula which we proposed in this research, the 5 IQA methods were applied to assess the quality of photographs. To choose the best IQA method for the formula, we conducted the questionnaire survey and invited 50 participants to participate. An important finding was that after the IQA questionnaire survey, interviews of some participants were conducted to ascertain what factors led them to choose the photograph as the low-quality one. The main reasons were the photograph brightness and color saturation. Most participants prefer brilliant photographs and dislike obscure photographs. This survey only provided scenic photographs for participants to choose from, which might have led them to prefer brilliant and colorful photographs. The second reason is that a few participants were concerned about the photograph composition. Those participants know basic photography principles, which caused them to choose low-quality photographs often. In the next IQA survey, we investigate expert photographers and laymen (who have no knowledge of photography) along with their differences in choosing low-quality photographs.

To verify the result of less-known tourist attraction, the verification experiments are conducted which revealed interesting points: while we provide the same seascape photographs for Taiwanese respondents and Japanese respondents, for Taiwanese results, this place received high evaluations and attracted respondents to visit there, but this place received the lowest score among Japanese results. The reason is that these photographs show “torii”, which are traditional gates of Japanese shrines. “Torii” are truly rare in Taiwan, but they are very common in Japan. Consequently, in this case, we can observe that foreign visitors are interested in special landmarks that their country does not have. That is to say, the scenic photographs including some special landmarks are expected to increase the attractiveness of these spots.

Investigation of potential tourist attractions is important for the tourism industry and for academics. Potential tourist attractions can not only promote economic development for a country, they can also enhance cultural communications. In academic assessment, very little was found in the literature on the issues of using social big data to identify those potential tourist attractions currently. Therefore, this study can encourage more researchers to assign importance to potential tourist attractions. The present study revealed some attractive less-known tourist attractions, which have insufficient information to estimate whether this place is safe or not. We expect that these potential places can be assessed further through field surveys by experts in the future.

Study limitations include the following: (i) this research collected scenic photographs as the dataset to discover less-known tourist attractions. The complex scenes did not exist in those scenic photographs which only included one to three subjects such as mountains surround the lake and forest with the river. Nevertheless, the nature scenes do not have fixed features (e.g., shape) that is hard to use object detection to annotate multiple labels for photographs and provide more information to tourists. Thereby, we use image classification simply to classify photographs into one label currently. (ii) From early investigation, most of Japanese participants minded to leave their background information. This situation makes it difficult to conduct a questionnaire survey and collect more samples. Hence, we decided to eliminate the questions of their background in this survey to prevent raising privacy concerns. In addition, the participants are strictly selected i.e., who meet five requirements as explain in Section 3.4. In this way, we can rely on their viewpoint and ensure the reliability of the result. (iii) In light of the limited sample used for the present study, a more comprehensive survey is expected to investigate more participants from different countries.

## 7. Conclusions

This study applied a novel method to identify less-known tourist attractions for people of different nationalities. The construction of the approach was undertaken based on two ideas. The first is ascertainment of local residents’ and foreign visitors’ unfamiliarity with Japanese cities.

Second, we propose a formula to evaluate the degree of tourist attraction, which includes different aspects such as image quality assessment (IQA), comment sentiment, and tourist attraction popularity for ranking tourist attractions. Cities that are familiar to participants are eliminated from the ranking results; the remnant spots are our target. Finally, through verification experiments, we confirmed that our result represents success.

Because COVID-19 has brought enormous damage worldwide and it has particularly influenced the tourism industry, most countries have lost great amounts of revenue. After the pandemic, tourism recovery efforts will be of paramount importance. Apart from original popular tourist attractions, tourism to other potential places can be developed, providing tourists with various tourist attractions. The use of the discovering less-known tourist attractions approach in future applied studies could contribute to developing the tourism industries worldwide as well as has the potential benefit to aid tourism recovery. Additionally, in accordance with our observations, most existing tourism recommender systems only recommend popular tourist attractions for tourists. However, certain tourists might feel tired of visiting those popular tourist attractions and interested in new places. Combing the information about less-known tourist attractions with existing tourism recommender system, tourists can be served with helpful and more comprehensive results. Besides, how to popularize and conduct propaganda for less-known tourist attractions is an important issue for the tourism industry in the future.

As future work, after collecting and analyzing more photographs taken in certain years, we expect to distinguish between local residents and foreign visitors in terms of their characteristic preferences. Considering more factors related to less-known tourist attractions (e.g., geography and population), we expect to improve the formula and cluster analysis used in this study. Less-known tourist attractions can be classified by season, weather, days, and nights according to the photograph times and contents. Furthermore, less-known tourist attractions can be assessed according to whether a place is readily accessible, or not, which can support tourists in their judgment about visiting a place. Currently, less-known tourist attractions have insufficient photographs to which tourists can refer. Therefore, we are working on simulating photographs at different times and seasons using a generative adversarial network (GAN). Moreover, other information related to SNSs (e.g., Instagram, Twitter, and Facebook) will be added to our dataset and subjected to cross-validation with our results of a questionnaire survey to verify the correctness of our result. Finally, we want to apply our results to travel recommendation services and provide various travel plans for tourists.

**Author Contributions:** Conceptualization, J.-Y.L., S.-M.W., M.H., T.A., H.I.; methodology, J.-Y.L., S.-M.W., M.H., T.A., H.I.; software, J.-Y.L., S.-M.W.; formal analysis, S.-M.W.; data curation, J.-Y.L.; writing-original draft, J.-Y.L.; writing-review and editing, S.-M.W., M.H.; supervision, M.H., T.A., H.I. All authors have read and agreed to publish version of the manuscript.

**Funding:** This work was supported by JSPS KAKENHI Grant Number 20K12081 and Tokyo Metropolitan University Grant-in-Aid for Research on Priority Areas.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Lim, K.H.; Chan, J.; Leckie, C.; Karunasekera, S. Personalized tour recommendation based on user interests and points of interest visit durations. In Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI), Buenos Aires, Argentina, 25–31 July 2015; pp. 1778–1784.
2. Memon, I.; Chen, L.; Majid, A.; Lv, M.; Hussain, I.; Chen, G. Travel recommendation using geo-tagged photos in social media for tourist. *Wirel. Pers. Commun.* **2014**, *80*, 1347–1362. [[CrossRef](#)]
3. Jiang, S.; Qian, X.; Mei, T.; Fu, Y. Personalized travel sequence recommendation on multi-source big social media. *IEEE Trans. Big Data* **2016**, *2*, 43–56. [[CrossRef](#)]
4. Peng, X.; Huang, Z. A Novel Popular Tourist Attraction Discovering Approach Based on Geo-Tagged Social Media Big Data. *ISPRS Int. J. Geo-Inform.* **2017**, *6*, 216. [[CrossRef](#)]

5. Hausmann, A.; Toivonen, T.K.; Slotow, R.; Tenkanen, H.T.O.; Moilanen, A.J.; Heikinheimo, V.V.; Di Minin, E. Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. *Conserv. Lett.* **2017**, *11*, e12343. [[CrossRef](#)]
6. Ghermandi, A.; Camacho-Valdez, V.; Trejo-Espinosa, H. Social media-based analysis of cultural ecosystem services and heritage tourism in a coastal region of Mexico. *Tour. Manag.* **2020**, *77*, 104002. [[CrossRef](#)]
7. Liu, J.; Huang, Z.; Chen, L.; Shen, H.T.; Yan, Z. Discovering areas of interest with geo-tagged images and check-ins. In Proceedings of the 20th ACM international conference on Multimedia-MM '12, Nara, Japan, 26–30 October 2020; Association for Computing Machinery (ACM): New York, NY, USA, 2012; pp. 589–598.
8. Zhuang, C.; Ma, Q.; Liang, X.; Yoshikawa, M. Discovering Obscure Sightseeing Spots by Analysis of Geo-tagged Social Images. In Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015—ASONAM '15, Paris, France, 25–28 August 2015; Association for Computing Machinery (ACM): New York, NY, USA, 2015; pp. 590–595.
9. Chuang, H.-M.; Chang, C.-H.; Kao, T.-Y.; Cheng, C.-T.; Huang, Y.-Y.; Cheong, K.-P. Enabling maps/location searches on mobile devices: Constructing a POI database via focused crawling and information extraction. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 1405–1425. [[CrossRef](#)]
10. Jonietz, D.; Zipf, A. Defining fitness-for-use for crowdsourced points of interest (POI). *ISPRS Int. J. Geo-Inform.* **2016**, *5*, 149. [[CrossRef](#)]
11. Rousell, A.; Hahmann, S.; Bakillah, M.; Mobasheri, A. Extraction of landmarks from OpenStreetMap for use in navigational instructions. In Proceedings of the 18th AGILE International Conference on Geographic Information Science, Lisbon, Portugal, 9–12 June 2015.
12. Cheng, Z.; Caverlee, J.; Lee, K.; Sui, D.Z. Exploring Millions of Footprints in Location Sharing Services. In Proceedings of the 5th International AAAI Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, 17–21 July 2011; pp. 81–88.
13. Skovsgaard, A.; Idlauskas, D.; Jensen, C.S. A clustering approach to the discovery of points of interest from geo-tagged microblog posts. In Proceedings of the 2014 IEEE 15th International Conference on Mobile Data Management, Brisbane, Australia, 15–18 July 2014; Institute of Electrical and Electronics Engineers (IEEE): New York, NY, USA, 2014; Volume 1, pp. 178–188.
14. Vu, D.D.; To, H.; Shin, W.; Shahabi, C. GeoSocialBound: An Efficient Framework for Estimating Social POI Boundaries Using Spatio-Textual Information. In Proceedings of the Third International ACM SIGMOD Workshop on Managing and Mining Enriched Geo-Spatial Data, San Francisco, CA, USA, 26 June–1 July 2016; ACM: New York, NY, USA, 2016.
15. Kuo, C.-L.; Chan, T.-C.; Fan, I.-C.; Zipf, A. Efficient method for POI/ROI discovery using flickr geotagged photos. *ISPRS Int. J. Geo-Inform.* **2018**, *7*, 121. [[CrossRef](#)]
16. Spyrou, E.; Korakakis, M.; Charalampidis, V.; Psallas, A.; Mylonas, P. A Geo-Clustering approach for the detection of areas-of-interest and their underlying semantics. *Algorithms* **2017**, *10*, 35. [[CrossRef](#)]
17. Hu, Y.; Gao, S.; Janowicz, K.; Yu, B.; Li, W.; Prasad, S. Extracting and understanding urban areas of interest using geotagged photos. *Comput. Environ. Urban. Syst.* **2015**, *54*, 240–254. [[CrossRef](#)]
18. Schubert, S.F.; Brida, J.G.; Risso, W.A. The impacts of international tourism demand on economic growth of small economies dependent on tourism. *Tour. Manag.* **2011**, *32*, 377–385. [[CrossRef](#)]
19. Andriotis, K. Scale of hospitality firms and local economic development-Evidence from Crete. *Tour. Manag.* **2002**, *23*, 333–341. [[CrossRef](#)]
20. Croes, R.R. A paradigm shift to a new strategy for small island economies: Embracing demand side economics for value enhancement and long term economic stability. *Tour. Manag.* **2006**, *27*, 453–465. [[CrossRef](#)]
21. Fagence, M. Tourism as a feasible option for sustainable development in small island developing states (SIDS): Nauru as a case study. *Pac. Tour. Rev.* **1999**, *3*, 133–142.
22. Lin, B.; Liu, H. A study of economies of scale and economies of scope in Taiwan international tourist hotels. *Asia Pac. J. Tour. Res.* **2000**, *5*, 21–28. [[CrossRef](#)]
23. Crouch, G.I.; Ritchie, J. Tourism, Competitiveness, and Societal Prosperity. *J. Bus. Res.* **1999**, *44*, 137–152. [[CrossRef](#)]
24. Kakamu, K.; Polasek, W.; Wago, H. Spatial interaction of crime incidents in Japan. *Math. Comput. Simul.* **2008**, *78*, 276–282. [[CrossRef](#)]
25. Altindag, D.T. Crime and International Tourism. *J. Labor Res.* **2014**, *35*, 1–14. [[CrossRef](#)]

26. Niewiadomski, P. COVID-19: From temporary de-globalisation to a re-discovery of tourism? *Tour. Geogr.* **2020**, *22*, 651–656. [[CrossRef](#)]
27. Gössling, S.; Scott, D.; Hall, C.M. Pandemics, tourism and global change: A rapid assessment of COVID-19. *J. Sustain. Tour.* **2020**, 1–20. [[CrossRef](#)]
28. Higgins-Desbiolles, F. Socialising tourism for social and ecological justice after COVID-19. *Tour. Geogr.* **2020**, *22*, 610–623. [[CrossRef](#)]
29. Polyzos, S.; Samitas, A.; Spyridou, A.E. Tourism demand and the COVID-19 pandemic: An LSTM approach. *Tour. Recreat. Res.* **2020**, 1–13. [[CrossRef](#)]
30. Jurado, E.N.; Damian, I.M.; Fernández-Morales, A. Carrying capacity model applied in coastal destinations. *Ann. Tour. Res.* **2013**, *43*, 1–19. [[CrossRef](#)]
31. Marusic, Z.; Horak, S.; Tomljenović, R. The socioeconomic impacts of cruise tourism: A case study of Croatian destinations. *Tour. Mar. Environ.* **2008**, *5*, 131–144. [[CrossRef](#)]
32. Luque-Gil, A.M.; Gómez-Moreno, M.L.; Peláez-Fernández, M.A. Starting to enjoy nature in Mediterranean mountains: Crowding perception and satisfaction. *Tour. Manag. Perspect.* **2018**, *25*, 93–103. [[CrossRef](#)]
33. Jacobsen, J.K.S. Anti-tourist attitudes. *Ann. Tour. Res.* **2000**, *27*, 284–300. [[CrossRef](#)]
34. Yin, J.; Cheng, Y.; Bi, Y.; Ni, Y. Tourists perceived crowding and destination attractiveness: The moderating effects of perceived risk and experience quality. *J. Destin. Mark. Manag.* **2020**, *18*, 100489. [[CrossRef](#)]
35. Lin, J.; Wen, S.; Hirota, M.; Araki, T.; Ishikawa, H. Analysis of rarely known tourist attractions by geo-tagged photographs. In Proceedings of the 11th International Conference on Advances in Multimedia MMEDIA 2019, Valencia, Spain, 24–28 March 2019; pp. 13–18.
36. Lin, J.; Wen, S.; Hirota, M.; Araki, T.; Ishikawa, H. Less-known tourist attraction analysis using clustering geo-tagged photographs via X-means. *Int. J. Adv. Syst. Meas.* **2019**, *12*, 215–224.
37. Yuan, Q.; Cong, G.; Ma, Z.; Sun, A.; Thalmann, N.M. Time-aware point-of-interest recommendation. In Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval-SIGIR '13, Dublin, Ireland, 28 July–1 August 2013; Association for Computing Machinery (ACM): New York, NY, USA, 2013; pp. 363–372.
38. Massimo, D.; Ricci, F. Clustering Users' POIs Visit Trajectories for Next-POI Recommendation. In *Information Communication Technology Tourism 2019*; Springer Science and Business Media LLC: New York, NY, USA, 2018; pp. 3–14.
39. Bin, C.; Gu, T.; Sun, Y.; Chang, L.; Sun, W.; Sun, L. Personalized POIs travel route recommendation system based on tourism big data. In Proceedings of the PRICAI 2018: Trends in Artificial Intelligence, Nanjing, China, 28–31 August 2018; Springer Science and Business Media LLC: New York, NY, USA, 2018; pp. 290–299.
40. Ye, M.; Yin, P.; Lee, W.-C.; Lee, D.-L. Exploiting geographical influence for collaborative point-of-interest recommendation. In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information-SIGIR '11, Beijing, China, 24–28 July 2011; Association for Computing Machinery (ACM): New York, NY, USA, 2011; pp. 325–334.
41. Garrod, B. Exploring place perception a photo-based analysis. *Ann. Tour. Res.* **2008**, *35*, 381–401. [[CrossRef](#)]
42. Garrod, B. Understanding the relationship between tourism destination imagery and tourist photography. *J. Travel Res.* **2008**, *47*, 346–358. [[CrossRef](#)]
43. Molina, A.; Esteban, A. Tourism brochures: Usefulness and image. *Ann. Tour. Res.* **2006**, *33*, 1036–1056. [[CrossRef](#)]
44. Deng, Y.; Loy, C.C.; Tang, X. Image aesthetic assessment: An experimental survey. *IEEE Signal. Process. Mag.* **2017**, *34*, 80–106. [[CrossRef](#)]
45. Wang, Z.; Liu, D.; Chang, S.; Dolcos, F.; Beck, D.; Huang, T. Image aesthetics assessment using Deep Chatterjee's machine. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017.
46. Tang, X.; Luo, W.; Wang, X. Content-Based Photo Quality Assessment. *IEEE Trans. Multimed.* **2013**, *15*, 1930–1943. [[CrossRef](#)]
47. Lu, X.; Lin, Z.; Jin, H.; Yang, J.; Wang, J.Z. RAPID: Rating pictorial aesthetics using deep learning. In Proceedings of the 22nd ACM International Conference on Multimedia, Orlando, FL, USA, 3–7 November 2014; ACM: New York, NY, USA, 2014; pp. 457–466.
48. Tian, X.; Dong, Z.; Yang, K.; Mei, T. Query-dependent aesthetic model with deep learning for photo quality assessment. *IEEE Trans. Multimed.* **2015**, *17*, 2035–2048. [[CrossRef](#)]

49. Lu, X.; Lin, Z.; Jin, H.; Yang, J.; Wang, J.Z. Rating image aesthetics using deep learning. *IEEE Trans. Multimed.* **2015**, *17*, 2021–2034. [[CrossRef](#)]
50. Dong, Z.; Tian, X. Multi-level photo quality assessment with multi-view features. *Neurocomputing* **2015**, *168*, 308–319. [[CrossRef](#)]
51. Dong, Z.; Shen, X.; Li, H.; Tian, X. Photo quality assessment with DCNN that understands image well. *Lect. Notes Comput. Sci.* **2015**, *8936*, 524–535. [[CrossRef](#)]
52. Talebi, H.; Milanfar, P. NIMA: Neural Image Assessment. *IEEE Trans. Image Process.* **2018**, *27*, 3998–4011. [[CrossRef](#)]
53. Murray, N.; Marchesotti, L.; Perronnin, F. AVA: A large-scale database for aesthetic visual analysis. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 16–21 June 2012; Institute of Electrical and Electronics Engineers (IEEE): New York, NY, USA, 2012; pp. 2408–2415.
54. Spanhol, F.A.; Oliveira, L.S.; Petitjean, C.; Heutte, L. Breast cancer histopathological image classification using Convolutional Neural Networks. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; Institute of Electrical and Electronics Engineers (IEEE): New York, NY, USA, 2016; pp. 2560–2567.
55. Zhang, J.; Xie, Y.; Wu, Q.; Xia, Y. Medical image classification using synergic deep learning. *Med. Image Anal.* **2019**, *54*, 10–19. [[CrossRef](#)]
56. Raj, R.J.S.; Shobana, S.J.; Pustokhina, I.V.; Pustokhin, D.A.; Gupta, D.; Shankar, K. Optimal feature selection-based medical image classification using deep learning model in internet of medical things. *IEEE Access* **2020**, *8*, 58006–58017. [[CrossRef](#)]
57. Shankar, K.; Zhang, Y.; Liu, Y.; Wu, L.; Chen, C.-H. Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. *IEEE Access* **2020**, *8*, 118164–118173. [[CrossRef](#)]
58. Kleinlein, R.; García-Faura, Á.; Jiménez, C.L.; Montero, J.M.; Díaz-De-María, F.; Fernandez-Martinez, F. Predicting image aesthetics for intelligent tourism information systems. *Electronics* **2019**, *8*, 671. [[CrossRef](#)]
59. Pelleg, D.; Moore, A.W. X-means: Extending k-means with efficient estimation of the number of clusters. In Proceedings of the 7th International Conference on Machine Learning, Tokyo, Japan, 14–15 June 2020; pp. 727–734.
60. He, Y.; Guo, H.; Jin, M.; Ren, P. A linguistic entropy weight method and its application in linguistic multi-attribute group decision making. *Nonlinear Dyn.* **2016**, *84*, 399–404. [[CrossRef](#)]
61. Ji, Y.; Huang, G.; Sun, W. Risk assessment of hydropower stations through an integrated fuzzy entropy-weight multiple criteria decision making method: A case study of the Xiangxi River. *Expert Syst. Appl.* **2015**, *42*, 5380–5389. [[CrossRef](#)]
62. Delgado, A.; Romero, I. Environmental conflict analysis using an integrated grey clustering and entropy-weight method: A case study of a mining project in Peru. *Environ. Model. Softw.* **2016**, *77*, 108–121. [[CrossRef](#)]
63. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; Institute of Electrical and Electronics Engineers (IEEE): New York, NY, USA, 2016; pp. 2818–2826.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).