Article

Where and Why Travelers Visit? Classifying Coastal Tourism Activities Using Geotagged Image Content from Social Media Data

Gang Sun Kim [1] , Choong-Ki Kim [1,*] and Woo-Kyun Lee [2]

1   Divisions for Natural Environment, Korea Environment Institute (KEI), Sejong 30147, Republic of Korea; gskim@kei.re.kr
2   Department of Environmental Science & Ecological Engineering, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea; leewk@korea.ac.kr
*   Correspondence: ckkim@kei.re.kr; Tel.: +82-44-415-7007

Abstract: Accurate information regarding the size, activity, and distribution of coastal tourists is essential for the effective management and planning of coastal tourism. In this study, geotagged photos uploaded to social network services were classified to identify coastal tourism activities. These activities were linked with spatial-scale data on tourist numbers estimated from social media data. To classify the activities, which included recreation, appreciation, education, and other activities, an image-supervised classification model was trained using 12,229 images, and the test accuracy was found to be 0.7244. On the Flickr platform, 43% of the image data located in the coastal land of South Korea are other activities, 39% are appreciation activities, and 18% are recreation and education activities. Other activities are mainly located in urban areas with a high population density and are spatially concentrated, while appreciation activities are mainly located in the natural environment and tend to be spatially spread out. Data on tourist activity categorization through content classification, combined with traditional tourist volume estimates, can help us understand previously overlooked information and context about a space.

Keywords: coastal tourism management; geotagged social media; spatial data analysis; image classification; tourist behavior insights; data integration

1. Introduction

Coastal tourism refers to a broad range of travel, leisure, and recreational activities that take place in coastal areas and involves close interactions between humans and the coastal environment [1–3]. Coastal tourism depends on a well-managed coastal environment, and it provides a rationale for conserving and managing such environments [4,5]. The coast is an ideal place for recreation and leisure, and the benefits that people derive from coastal tourism, such as stress relief and relaxation, can be considered ecosystem services [6]. The 14th goal of the UN's Sustainable Development Goals (SDGs), "Life under Water", includes the effective management of coastal and marine tourism and the equitable distribution of its benefits to communities [7]. However, there are other demands on coastal space, including demand for ports, aggregate extraction, and so on [8,9]. Spatial planning is the process of identifying the current use and demand for space and reconciling conflicting activities and demands to achieve goals [10,11]. To carry out spatial planning for coastal tourism, it is important to know how tourists are distributed in a space and the activities that they enjoy in a given space [12].

With the proliferation of smartphones since the late 2000s, people have started sharing their travel experiences on social media platforms. Geotagged social media data, which include the location information of users, have emerged since the early 2010s, and there has been a surge in research utilizing these data in the field of tourism [13–15]. Social media data

are categorized as user-generated content (UGC) in the big data category, which is defined as content such as photos and text produced by users [16]. By utilizing UGC with geotags, it is possible to gain insights into the geographical distribution of tourists. When combined with official tourist statistics, it becomes possible to estimate the number of tourists in areas lacking visitor counts [17]. Over the past decade, research has proven that geotagged social media data can be a good proxy for visitor numbers; however, such data also have several limitations [14,15], such as data preprocessing problems that prevented the extraction of complete tourism-related records and the inability to check detailed tourism activities, which were left for future research [18].

Recent advancements in AI and deep-learning algorithms have significantly contributed to the analysis and classification of qualitative content, such as text and images [19–21]. These techniques offer new possibilities for analyzing the content generated by tourists, thus providing insights into their characteristics and tourism activities. Specifically, leveraging AI and deep learning to analyze UGC can further enhance existing research on geotagged social media data [22]. Moreover, it opens up opportunities for qualitative research in addition to quantitative research focused on estimating visitor numbers because it allows for an understanding of the specific tourist activities that take place [23]. Finally, more detailed information and spatial insights can be obtained by integrating geographic information with advanced content analysis [22].

Spatial insights reveal certain spatial contexts and trends rather than a simple distribution. The data and methods mentioned in the previous paragraphs will allow for a better understanding of the quantitative and qualitative context of "Where and why travelers visit". Understanding this information is crucial when formulating spatial plans [24,25]. Previous studies have primarily explored the feasibility of image classification using unsupervised techniques applied to social media data [26–28]. However, limited research has been conducted on applying classification criteria specifically to coastal tourism or utilizing such criteria for spatial planning. Unified standards and methodologies that can be applied equally to wide areas are required to establish spatial planning policies, and a supervised classification method with a top–bottom process based on tourism activity classification criteria is considered suitable.

In this study, a UGC image classification model for coastal tourism activity classification was developed using deep-learning technology and applied to coastal areas on a national scale. This model will allow for the identification of patterns in how coastal tourism activities are spatially distributed and enable discussions on how qualitative information can be combined with existing quantitative information to inform spatial planning. To this end, we first created tourism activity classification criteria and training data to apply the supervised classification method for coastal tourism activity classification and then examined the classification accuracy and validity of the created model. Second, the classification model was applied to the study area (South Korean coastal area) to create a map of the distribution of coastal tourism activities and to understand the patterns revealed by the spatial environment. Finally, the combination of quantitative and qualitative information to understand spatial information and context is explored through examples.

2. Materials and Methods

2.1. Study Area

The study area included the coastal land and island areas of South Korea, and social media and other data were collected based on this target area (Figure 1). Coastal land is defined as the area buffered by 500 m landward from the coastline and 1 km landward from certain areas, such as ports, according to the South Korean Coast Management Act. The study area was divided into a 30 s square grid, and social media data and other data were tagged with grid information for analysis.

Five major coastal cities and one island in South Korea were selected as hotspot target areas, where a large amount of UGC data can be obtained. All are located by the sea, represent popular tourist destinations, and are easily accessible by rail or air transport

(Figure 1, Table 1). Incheon and Mokpo are on the west side of Korea, Yeosu is on the south side, and Busan and Gangneung are on the east side. The western and southern parts of Korea have very complex coastlines, with many islands and tidal flats. The eastern side of Korea has a simpler coastline and no tidal flats; instead, the sea is deeper than that in the west. Jeju Island is a volcanic island far to the south of the peninsula and has a unique tropical climate that differs from that of the rest of the peninsula.
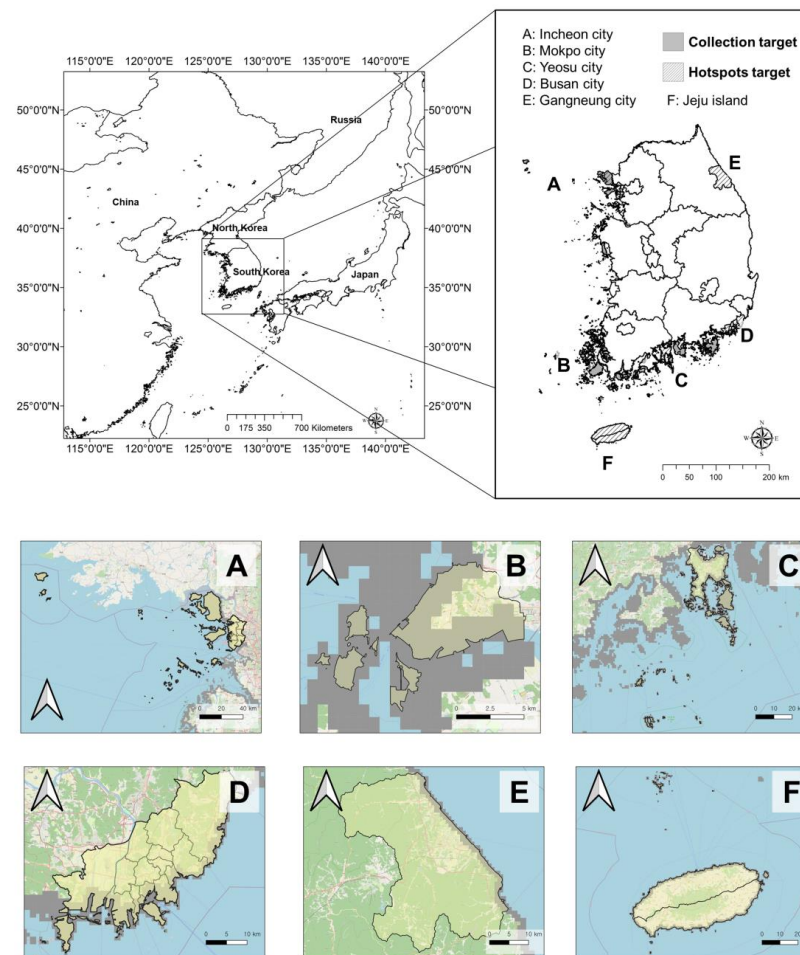
Figure 1. Data collection and hotspot targets in this study (study area).

Table 1. Characteristics of six hotspot target cities/regions in the study area (2022).

City/Region	Area (km^2)	Population per Area (People/km^2)	Transportation Infrastructure	Characteristics
Incheon City	1066	2764.6	Maritime transport Air transport	• Korea's largest international airport is located here and this city is considered as a gateway city adjacent to Seoul; • Incheon is a very large city, and the distance between the islands is long.
Mokpo City	51.66	4231.3	High-speed rail Maritime transport	• Although Mokpo has a small area, it is a densely populated city; • There is a port to move to the major famous islands in the West Sea.
Yeosu City	512.3	540.28	High-speed rail Maritime transport Air transport	• Tourist attractions are formed around the EXPO Ocean Park, where the Expo was held in 2012.
Busan City	770.1	4350.4	High-speed rail Maritime transport Air transport	• Busan, the second largest city in Korea, and Busan Port, the largest port in Korea, exist in this city.

Table 1. *Cont.*

City/Region	Area (km²)	Population per Area (People/km²)	Transportation Infrastructure	Characteristics
Gangneung City	1041	204.58	High-speed rail	• A city bordered by the sea to the east and rugged mountains to the west.
Jeju Island	1850	365.76	Maritime transport Air transport	• There are many natural resources which can only exist on volcanic islands; • Three types of protected areas designated by UNESCO exist here.

2.2. Research Process

This study was conducted as detailed in Figure 2 and divided into five major steps: (1) data collection, (2) preprocessing of social media data into spatial grids, (3) estimation of visitation rate by a spatial grid and derivation of the spatial hotspot grid, (4) setting image classification criteria and model training, and (5) application of the image classification model to the study area. Step (1) is the preparatory stage for this study and includes collecting social media data and other data based on the study area. Steps (2) and (3) are performed to estimate the quantitative size and distribution of coastal tourists and spatial grids, as well as to gather and analyze the social media data and other additional data to identify the locations visited by tourists. Steps (4) and (5) are conducted to classify the image data recorded on the social network site (SNS) Flickr using a deep-learning model to monitor the distribution of coastal tourism activities.
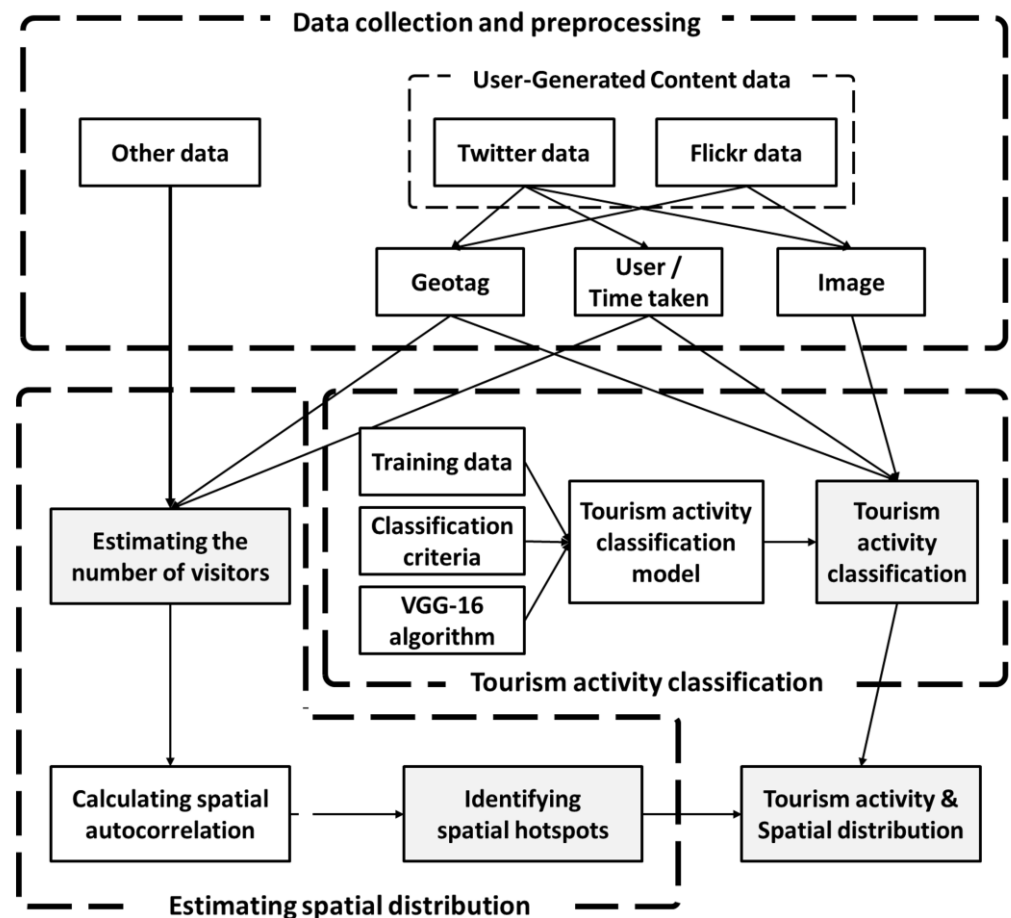

Figure 2. Research process.

2.3. Data Collection

The data collected were social media, geographic information, and other spatial data (Table 2). Social media data and other spatial data containing geographic information were

used to estimate the size and distribution of coastal tourism, and the image contents of social media data were used to analyze the ratio of coastal tourism activities in the study area ("Data collection and preprocessing" in Figure 2).

Table 2. Data used in this study.

Category	Data	Geospatial Type	Time Range	Data Components	Source
Social media data	Flickr	Geographic coordinates (point)	2013–2020	• User ID (de-identified); • Time taken; • Geographic coordinates; • Picture content.	Flickr SNS platform
	X (formerly known as Twitter)	30 s grid polygon units	2013–2018	• User ID (de-identified); • Time taken; • Grid-feature ID.	X SNS platform
Other spatial data	Population density	Raster format (100 m resolution)	2020	• Population count (people).	[29,30]
	Beach distribution map	Polygon format	-	• Location and shape of beaches.	[30–32]
	Administrative region map	Polygon format	2019	• Region name.	[33]

The social media data platforms used in this study were Flickr and X (formerly Twitter), and we chose them for several reasons. First, both platforms have been heavily utilized in tourism research, and their reliability and validity have been proven in various studies [15,34]. Second, data are relatively accessible compared with other platforms and can be obtained through APIs or crawling techniques. Third, X is a widely used platform in Korea; therefore, these data have a high spatial density and are highly representative of the total number of visitors [35]. Fourth, because Flickr is a platform that mainly deals with photography, it is easy to collect photo data, making it suitable for this study. Finally, an equation was developed to estimate the marine tourism visitation rate in South Korea using data from both platforms [36].

Flickr data were collected using the Flickr API and the R open library "Photosearcher", which simplifies the utilization of the Flickr API in R programming [37]. This library facilitated the collection of Flickr images captured within the study area, along with the associated metadata for each image. The collected data encompassed the period from 1 January 2013 to 31 December 2020, and included the following data components: user ID (de-identified), time taken, geographic coordinates, and picture content.

Geotagged X data within the data collection target area were collected from 1 January 2013 to 31 December 2018. Unlike on Flickr, data were collected at the level of the 30 s grid used in this study. It was assumed that tweets searched within a 600 m radius from the center of each grid were recorded within a specific 30 s grid. The process began by extracting the center-point coordinates for each grid used in this study. Subsequently, the X website's Tweet Search filter was utilized to search for tweets within a 600 m radius of a specific point. The searched information was converted into text format using Python, the Selenium library, and Chromedriver [38]. These processes were repeated iteratively using Python programming. The collected data included the following components: user ID (de-identified), time taken, and grid-feature ID.

Other spatial data were collected to estimate the number of visitors to coastal tourism. The data used were a population density map, a beach distribution map, and an administrative district map. The "Constrained Individual Countries" provided by Worldpop in 2020 was used for this study as the population density data, and the resolution of the data was 100 m [29,39]. Beach polygon distribution data were created using the Open Street Map and Esri Satellite map based on a list of beaches provided by the Korea Coastal

Portal [30–32]. The administrative district map shows the location and shape of South Korean administrative districts [33].

2.4. Data Processing and Estimating the Visitor Distribution Map

To estimate the number of visitors to coastal tourism on a grid unit, Equation (1) was used, where Y is the number of annual coastal tourism visitors, X1 is the annual average of the sum of Flickr Photo Users per day (PUD), X2 is the annual average of the sum of Twitter Users per day (TUD), X3 is the population, D1 is the Dummy variable for the Chungcheongnam-do region, D2 is the Dummy variable for Gangneung city, D3 is the Dummy variable for the Gyeonggi-do region, E1 is the Dummy variable for the beach, and when X1, X2, and X3 are all 0, Y is regarded as 0.

$$Ln(Y) = 10.30618 + [1.17358 \times \ln(X1)] + [0.81358 \times \ln(X2)]$$
$$+ [-0.13465 \times \ln2(X2)] + [0.18876 \times \ln(X3)] + (0.45986 \times D1) + (-1.66881 \times D2) \quad (1)$$
$$+ (-1.35672 \times D3) + (-0.67596 \times E1)$$

Equation (1) is an empirical regression derived from actual tourist visitor statistics that can estimate the number of coastal tourism visitors using Flickr and X (formerly Twitter) [36,40]. During data processing, the input variables were processed into square grids, and Equation (1) was applied to each grid. Only those grids that spatially overlapped with the study area (coastal land and island areas in South Korea) were used.

The PUD variable is the number of users who upload photos to Flickr in a day and represents an indicator that allows us to determine the actual number of users, even if the same user uploads many Flickr photos. After creating point distribution data based on geographic coordinates, overlapping Flickr data were extracted for each grid, and duplicates were removed through date and user to calculate the PUD for each grid. The TUD variable refers to the number of users who tweeted X in a day, and the TUD for each grid was calculated using date and user data. The population per grid was calculated using population density data. Population density raster data were extracted into grids, and the values of each raster pixel were summed. For pixels that only partially overlapped inside the grid, the ratio of the area of the pixel to the area of the pixel overlapping the grid was multiplied by the pixel value for each pixel, and the values were summed up. The variable region of each grid was determined using the administrative district map, and for grids that overlapped more than one administrative district, the region value was assigned based on the administrative district with the larger overlapping area. For beach areas, each grid was assigned a beach value of 1 or 0 depending on whether it overlapped with beach polygon data ("Estimating spatial distribution" in Figure 2).

2.5. Identifying Spatial Hotspots by Using the Visitor Distribution Map

Spatial hotspots generally refer to areas with higher values compared to their surroundings [41–43]. In this study, hotspots were identified based on estimated visitor count values. Because the effect of distance is not meaningful for islands that are not connected by land to the mainland or the main island, the hotspot analysis was based only on the mainland, main island, and areas connected by land. Getis–Ord Gi is a method for calculating statistics that determines local spatial autocorrelation and represents an indicator for finding spatial hotspots in the ArcGIS program developed by ESRI [44,45]. Getis–Ord Gi can be derived through Equations (2)–(4), where x_j is the attribute value for feature j, $\omega_{i,j}$ is the spatial weight between feature i and j, and n is equal to the total number of features [44].

$$\frac{\sum_{j=1}^{n} \omega_{i,j} x_j - \overline{X} \sum_{j=1}^{n} \omega_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} \omega_{i,j}^2 - \left(\sum_{j=1}^{n} \omega_{i,j} \right)^2 \right]}{n-1}}} \quad (2)$$

$$\overline{X} = \frac{\sum_{j=1}^{n} x_j}{n} \tag{3}$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\overline{X}\right)^2} \tag{4}$$

The hotspot analysis tool in the ESRI ArcGIS program (10.1 ver) calculates the Gi value for each feature and measures the intensity of clusters with high or low values. We consider Gi Z-score values greater than 1.96 as a hotspot, which means that the significance level of hotspots is below 0.05. To calculate the Getis–Ord Gi statistic, it is necessary to specify the "Conceptualization of Spatial Relationships" setting. Because each hotspot target may have its own unique local spatial autocorrelation, the fixed-distance method was deemed appropriate.

Global Moran's I is a statistic that identifies the overall spatial autocorrelation of a particular space [18,46,47]. A high Moran's I z-score indicates that the high and low values are clustered spatially, rather than randomly distributed spatially [48]. Moran's I can be derived through Equations (5) and (6), where z_i is the deviation of an attribute for feature i from its mean, $\omega_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features, and S_0 is the aggregate of all the spatial weight [48].

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2} \tag{5}$$

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{i,j} \tag{6}$$

Peak distance refers to the distance at which the z-score of Global Moran's I reaches its highest point as the distance increases. This can be considered an appropriate threshold distance, indicating the distance at which spatial autocorrelation becomes significant [44]. For each site, the Global Moran's I value and z-score were obtained by increasing the distance by 700 m, starting at 803 m, which was the minimum distance between the grids. The first peak distance, which is the distance at which the peak tendency first appeared, was set as the threshold distance for each site to highlight the factors causing hotspots ("Estimating spatial distribution" in Figure 2).

2.6. Creating Image Classification Criteria and a Model to Classify Tourism Activity

A top–down supervised classification was conducted to classify images of coastal tourism activity. For this purpose, classification criteria were set based on existing studies on coastal tourism. In this study, we chose a distinction between "marine-dependent" and "marine-related" tourism based on coastal space [49,50]. Marine-dependent tourism involves the direct utilization of the ocean and coastal environment, and it includes recreational activities, such as surfing, swimming, and snorkeling. Marine-related tourism refers to activities that indirectly utilize the ocean and coast, such as enjoying scenery and cultural experiences. In this study, a classification method that considers the level of ocean dependence was employed to establish criteria for categorizing coastal tourism activities. The established categories are "recreation activity", "appreciation activity", and "education activity". Additionally, an "other activities" category was created to account for activities recorded in coastal spaces that are not directly related to the coastal environment, such as indoor activities and foods (Table 3).

To conduct supervised classification, it is essential to gather a sufficient amount of training data. A total of 10,000 images were randomly selected from the Flickr dataset collected to make training data. Three assistants, including two master's students specializing in big data and one undergraduate student focused on statistics, visually categorized the images based on the established criteria. In cases of disagreement among the assistants, the category value was assigned based on a majority vote. If all three assistants assigned

different category values, the authors made the final judgment. Initially, a significant number of images were classified as appreciation and other activities, whereas relatively low counts were recorded for recreational and educational activities. To balance the amount of data across categories, additional images corresponding to recreation or educational activities were collected from Flickr and the Internet. A total of 12,226 training images were collected and classified. After removing 17 erroneous files and black-and-white images, the final training dataset consisted of 12,209 images.

Table 3. Coastal tourism activity criteria used in this study.

Coastal Tourism Activity Category	Explanation
Recreation **(Recreation/Sports)**	• Activities related to sandy beaches, such as sea bathing; • Beach camping and beach sports; • Participation in the beach; • Swimming, surfing, yachting, etc.; • Cruise activities such as passenger ships and submersibles; • Fishing, scuba diving, etc.
Appreciation **(Appreciation of Scenery/Aesthetics)**	• Appreciating the scenery; • Views from the observatory or coastal path; • Visiting a place with a good view.
Education **(Education/Culture)**	• Observing animals and plants; • Environmental experience education accompanied by a guide; • Visiting a museum or exhibition hall; • Tidal flat ecosystem observation.
Other	• Images that do not fall into the above categories (food, building interior, etc.)

The transfer-learning method, which leverages a ready-made algorithm architecture, was performed using the VGG16 (Visual Geometry Group 16) algorithm developed by the VGG team at the University of Oxford [51,52]. The VGG16 model is a convolutional neural network algorithm that first appeared in the Imagenet Large Scale Visual Recognition Challenge 2014 (ILSVRC, 2014). It is recognized for its ability to classify images and has been used in various studies.

The algorithm was trained and implemented using Python and the Keras library [53]. Cross-validation is essential in machine learning to mitigate overfitting and obtain a generalized model. To perform cross-validation, the data were divided into training, validation, and test datasets. The model training process was iterated multiple times, where the training data were used to train the model, and the validation data were used to assess the model's performance and guide adjustments to achieve a better fit in subsequent training iterations. Finally, the test data were used to select the best-performing model among several models created. A total of 999 images, which accounted for 1/10 of the total data, were allocated as test data, whereas the remaining data were split into training and validation data in a 7:3 ratio. Ten models were built during training iterations, and the model demonstrating the best performance was selected for analysis

Accuracy was used as the main indicator to evaluate the model's performance, whereas sensitivity and precision served as additional metrics for assessing the classification accuracy of each category. Accuracy represents the percentage of correctly classified data out of the total number of classified data points, with a higher accuracy indicating a better performance. However, accuracy has limitations when the distribution of data across classes is uneven because it can potentially overestimate the model's classification ability. Sensitivity and precision were calculated for each individual class to provide additional insights into the classification performance. Sensitivity measures the proportion of correctly

classified data within the target class, whereas precision measures the proportion of data classified as the target class that is truly part of the target class. (Table 4).

Table 4. Number of data points used to train and validate the image classification model.

Total Data	Training Data	Validation Data	Test Data	Removed Data
12,226	8789	2198	1222	17
Category	**Training/Validation Data**		**Test Data**	**Total Data**
Recreation	1571		170	1741
Appreciation	5044		583	5627
Education	1084		131	1215
Other	3288		338	3626
Total	10,987		1222	12,209

The selected model was applied to the collected geotagged Flickr images. Following the classification process, duplicates were removed based on the date, user, and tourism activities in each grid. The ratios of tourism activities were then calculated at the grid level ("Tourism activity classification" in Figure 2).

3. Results

### 3.1. Distribution Map and Spatial Hotspots of Coastal Tourism Visitors

Equation (1) was applied to estimate the annual number of marine tourist visits by grid, and the results are shown in Figure 3. It can be seen that grids located in big cities like Busan or famous tourist destinations like Seongsan Sunrise Peak (a UNESCO World Heritage Site since 2007) have higher values.
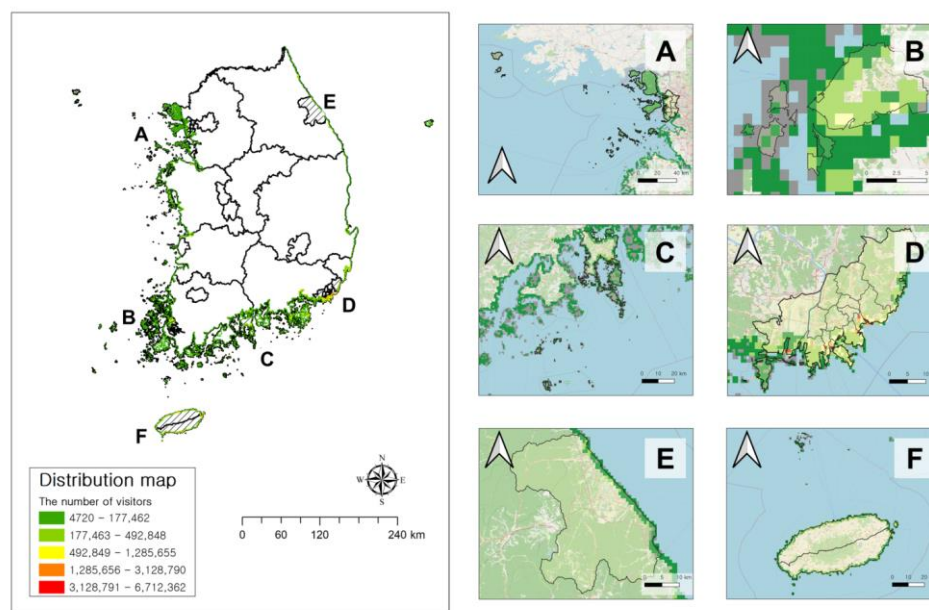
Figure 3. Coastal tourism visitor distribution map (A is Incheon City, B is Mokpo City, C is Yeosu City, D is Busan City, E is Gangeung City, and F is Jeju Island).

The hotspots are shown in Figure 4, and their details and characteristics are listed in Table 5. Incheon City (Target A) was divided into northern Incheon (Target A-1) and southern Incheon (Target A-2) based on the characteristics of being separated by land and then analyzed for spatial autocorrelation in each area. Overall, hotspots were most often located in urban centers around transport hubs, and also famous natural tourist destinations such as Songsan Sunrise Peak. Yeosu and Busan showed a pattern of having a

single giant hotspot; thus, they seem to represent a location type where giant hotspots can develop owing to the high population density and relatively simple and short coastlines. For less densely populated targets, the first peak distance was shorter; therefore, there were several relatively small hotspots.
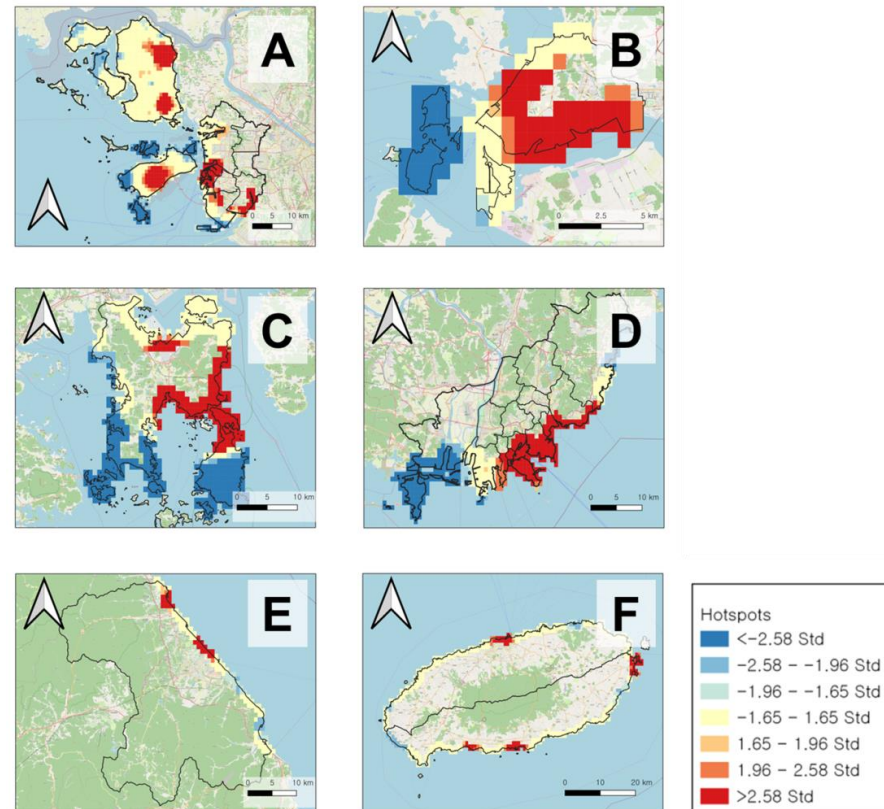
Figure 4. Coastal tourism hotspot map for each hotspot target ((**A**) is Incheon City, (**B**) is Mokpo City, (**C**) is Yeosu City, (**D**) is Busan City, (**E**) is Gangeung City, and (**F**) is Jeju Island).

Table 5. Spatial autocorrelation statistics and number of hotspot grids by hotspot target city/region.

City/Region	First Peak Distance (m)	Number of Grids	Number of Hotspot Grids	Spatial Patterns and Features
Incheon City	2203 (North) 3603 (South)	1268	183	• Different peak distance patterns between the islands north of the city and the land and islands to the south; • In the north of the island, hotspots were located around the town office. In the south, hotspots were located around Incheon International Airport, Incheon Port, the downtown area, and tourist attractions Wolmido Island and Sorye Wetland Ecological Park (Figure S1).
Mokpo City	4303	101	35	• One hotspot covering Mokpo Station, Mokpo Port, and the surrounding urban center (Figure S2).
Yeosu City	8503	580	136	• Hotspots include Yeosu Port, Yeosu Station, the marine cable car, the Expo Exhibition Center, and the surrounding city center; • Hotspots exist around Yeosu Airport and industrial complexes in the north of the country (Figure S3).
Busan City	7803	445	143	• One giant hotspot centered around Busan Port, the surrounding city center, and tourist destinations Haeundae and Bexco (Figure S4).

Table 5. *Cont.*

City/Region	First Peak Distance (m)	Number of Grids	Number of Hotspot Grids	Spatial Patterns and Features
Gangneung City	1503	128	24	• Hotspots exist around Gyeongpo Beach and Jisujin Beach in the northern part of the city; • Hotspot area is relatively small compared to other targets (Figure S5).
Jeju Island	2903	518	59	• Hotspots exist in the city center around Jeju Airport and Jeju Port, and in the city center of Seogwipo City. • Hotspots exist around Jungmun Tourist Complex, where hotels are located, and Seongsan Ilchulbong Peak, a UNESCO natural heritage site (Figure S6).

3.2. Accuracy of the Classification Model of Coastal Tourism Activities

The evaluation results for the classification model selected as the optimal model are listed in Table 6. The test accuracy of the classification model was 0.7422, which indicated that it has an appropriate classification ability for use in the field. The difference between this and the validation accuracy value of 0.7707 was small, which indicated that the model is not overfit and has acquired generality.

**Table 6.** Results of the cross-verification of coastal tourism activity classification.

Test Indicators		Indicator Value	Explanation
Accuracy	Training data	0.9185	$\frac{\text{Number of correct classifications}}{\text{Number of all classifications}}$
	Validation data	0.7707	
	Test data	0.7422	
Sensitivity	Recreation class	0.4824	$\frac{\text{Number of data corresponding to and classified to class}}{\text{Number of data corresponding to class}}$
	Appreciation class	0.8422	
	Education class	0.4962	
	Other class	0.7959	
Precision	Recreation class	0.5857	$\frac{\text{Number of data corresponding to and classified to class}}{\text{Number of data classified to class}}$
	Appreciation class	0.8143	
	Education class	0.5652	
	Other class	0.7390	

In addition, it is necessary to check the sensitivity and precision, which express the classification ability of each item, because the amount of data for each activity item used for training is different. The sensitivity and precision of the classification model are listed in Table 6. The sensitivity and precision of the appreciation item were 0.8422 and 0.8143, respectively, while those of the other activities item were 0.7959 and 0.7390, respectively. This means that the classification ability for appreciation and other activities is good, which is likely due to the large amount of training data. On the other hand, the sensitivity and precision of the recreation item and education item are relatively low compared to those of the previous two items. This is likely because of the relatively small amount of data used for training. It is expected that the classification ability will be improved if more sample data can be used for training in the future.

3.3. Distribution of Coastal Tourism Activities in the Study Area

A main coastal tourism activity map was created to illustrate the primary tourism activities for each spatial grid. For each grid, the tourism activity with the highest number was considered the main activity (Figures 5 and S7). The main activity was dynamically distributed according to the geography and characteristics of the space, and the number

of grids corresponding to each activity was 146 for recreation, 219 for education, 1419 for appreciation, and 629 for other activities.
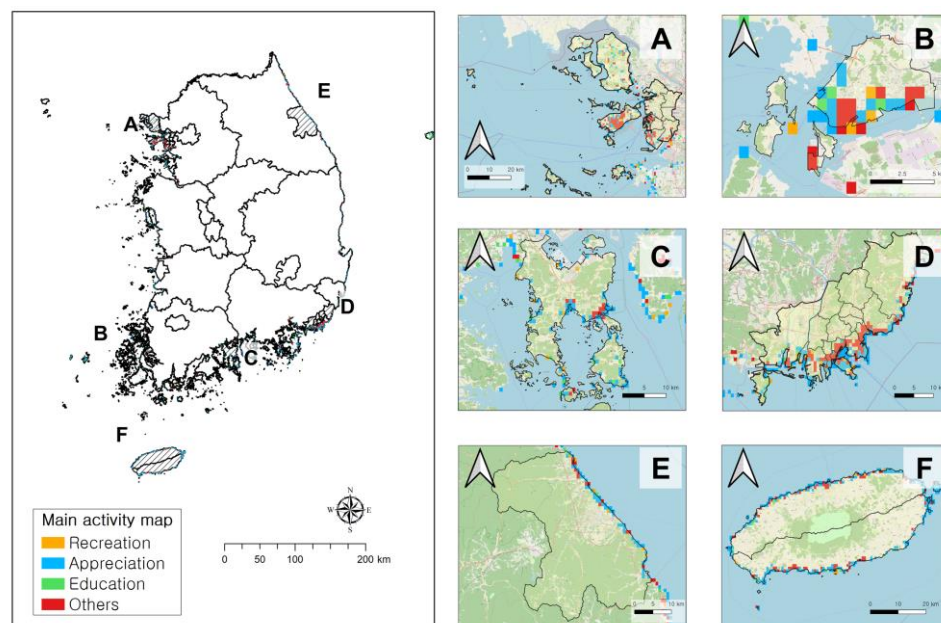

Figure 5. Main coastal tourism activity map representing tourism activities with the highest number of counts in each grid (number of grids corresponding to each activity is 146 for recreation, 219 for education, 1419 for appreciation, and 629 for other activities; A is Incheon City, B is Mokpo City, C is Yeosu City, D is Busan City, E is Gangeung City, and F is Jeju Island).
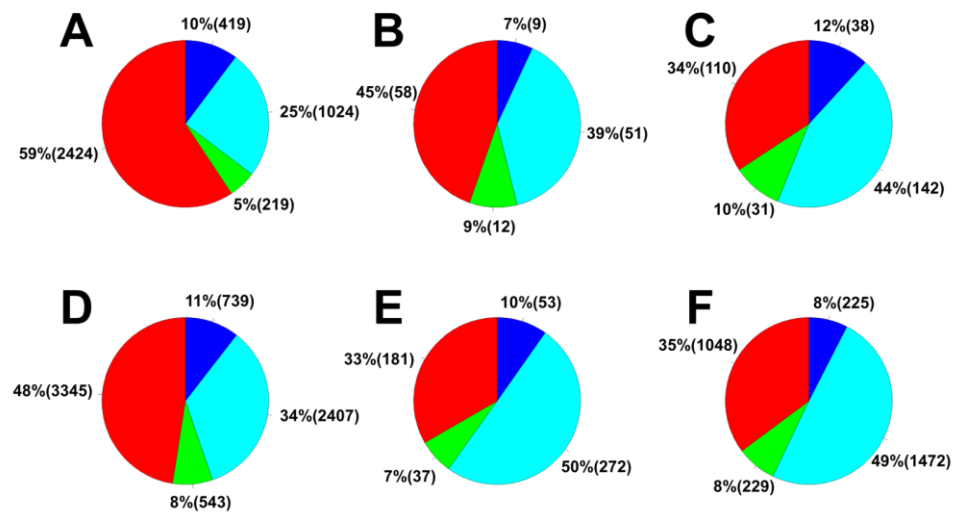
The distribution of coastal tourism activity categories across the study area is shown in Figure 6. Other and appreciation activities accounted for high percentages at 43% and 39%, respectively, whereas education and recreation had low values of 10% and 8%, respectively. In areas that included beaches, a significant increase in appreciation activities to 48% and a significant decrease in other activities to 35% were observed. In contrast, in grids with a denser-than-average population density (309 people/grid), other activities increased to 48%, and appreciation activities decreased to 34%. The data showed that in densely populated urban areas, other activities increased, while in natural coastal environments, such as beaches, appreciation activities that rely on the natural environment increased. The six hotspot target areas had a high percentage of other activities, which was similar to the distribution in densely populated areas.

The distribution of coastal tourism activities by hotspot target city/region can be found in Figure 7 and Table 7. Targets A, B, and D (Incheon, Mokpo, and Busan) have a large percentage of other activities (45–59%) and a small percentage of appreciation activities (25–39%), while targets C, E, and F (Yeosu City, Gangneung City, and Jeju Island) have a small percentage of other activities (33–35%) and a large percentage of appreciation activities. Recreation and education have fewer data points and relatively small proportions; therefore, significant patterns were not observed except for the lower educational activities of target A and lower recreational activities of target B.

In the grids corresponding to the hotspots, it was found that the overall proportion of other activities increased and the overall proportion of appreciation activities decreased in all hotspot target cities/regions. However, in the case of target D (Busan), which is already a huge metropolis and has a very large hotspot, the ratio of other activities to appreciation activities did not change dramatically. When comparing the ratio of counts recorded in hotspots to total counts for each activity, all six targets had the lowest ratio (0.45–0.83) for appreciation activity (Table 6). This indicates that appreciation activities are more likely to occur in spaces that are not hotspots.
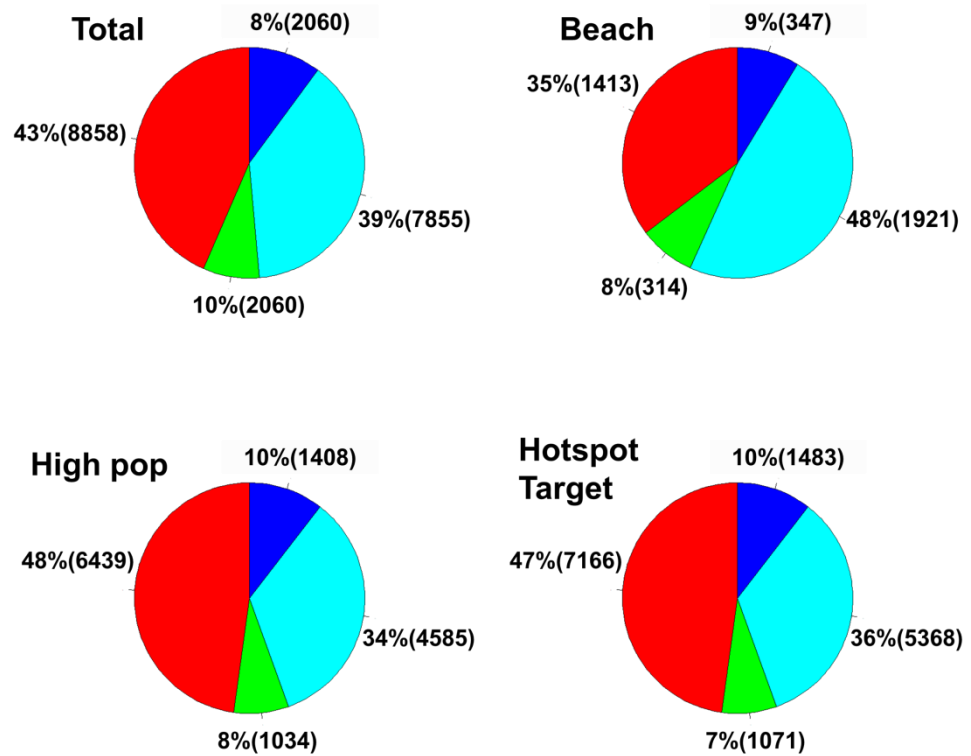
Figure 6. Pie graphs expressing the proportion of coastal tourism activities (**upper left** is for data collection target grids, **upper right** is for grids corresponding to beaches, **lower left** is for grids with a population density greater than the average (309 people/grid), and **lower right** is for hotspot target city/region lattice targets; blue is for recreation, sky blue is for appreciation, green is for education, and red is for other activities).

Figure 7. Pie graphs expressing the proportion of coastal tourism activities ((**A**) is Incheon City, (**B**) is Mokpo City, (**C**) is Yeosu City, (**D**) is Busan City, (**E**) is Gangeung City, and (**F**) is Jeju Island; blue is for recreation, sky blue is for appreciation, green is for education, and red is for other activities).

Table 7. Spatial autocorrelation statistics and number of hotspot grids by hotspot target city/region.

City/Region	Total/Hotspots	Recreation	Appreciation	Education	Other
Incheon City	Total	10% (419)	25% (1024)	5% (219)	59% (2424)
	Hotspots	11% (340)	20% (580)	5% (142)	64% (1898)
	Ratio recorded in hotspots	0.81	0.57	0.65	0.78
Mokpo City	Total	7% (9)	39% (51)	9% (12)	45% (58)
	Hotspots	6% (6)	32% (30)	9% (8)	53% (50)
	Ratio recorded in hotspots	0.67	0.59	0.67	0.86
Yeosu City	Total	12% (38)	44% (142)	10% (31)	34% (110)
	Hotspots	13% (30)	40% (95)	10% (24)	38% (91)
	Ratio recorded in hotspots	0.79	0.67	0.77	0.83
Busan City	Total	11% (739)	34% (2407)	8% (543)	48% (3345)
	Hotspots	11% (672)	33% (2021)	8% (486)	48% (2954)
	Ratio recorded in hotspots	0.91	0.83	0.90	0.88
Gangneung City	Total	10% (53)	50% (272)	7% (37)	33% (181)
	Hotspots	8% (21)	49% (125)	7% (18)	35% (89)
	Ratio recorded in hotspots	0.40	0.46	0.49	0.49
Jeju Island	Total	8% (225)	49% (1472)	8% (229)	35% (1048)
	Hotspots	8% (110)	48% (669)	7% (102)	37% (509)
	Ratio recorded in hotspots	0.49	0.45	0.45	0.49

4. Discussion

4.1. Distribution of the Number of Coastal Tourism Visitors and Characteristics of Spatial Hotspots

Figure 3 shows the distribution of coastal tourism visitors in South Korea, and it shows that the number of tourists in major cities is generally high based on the collected data. In the hotspot target city/region data, there was a high concentration of visitors in urban centers and popular tourist attractions. This is the same pattern that has been noted in other previous studies, with the data suggesting that visitors tend to visit popular tourist destinations and stay in urban centers along the coast [54–56].

Hotspot analysis allows the identification of areas containing clusters of high values among values distributed across a broader range. In addition, the degree of clustering and degree of concentration vary depending on the perspective of the target site class that researchers and policymakers are interested in. Therefore, some of the spaces that are hotspots from a regional-level perspective may not be hotspots from a local-level perspective [17]. In this study, the first peak distance was used to reveal local hotspot patterns and factors. However, if others wanted to find the strongest spatial autocorrelation pattern from a specific perspective, they could use the maximum peak distance to identify hotspots. For example, target E (Gangneung City) has a maximum peak distance of 16,903 m, in which case a huge hotspot will be formed that covers the north of the target (Figures 4 and A1). In conclusion, spatial hotspots tend to form around major urban areas or tourist attractions, and the size of the hotspots varies depending on the target site's area, degree of spatial autocorrelation, and specific conditions of the target site.

4.2. Tourism Activity Classification Methodology Based on the Deep-Learning Model and Supervised Classification

Previous studies have used unsupervised classification to classify images and identify tourism activities [26–28]. The unsupervised classification method is a bottom-to-top process that fits a classification system based on data and has the advantage of proceeding with classification without building training data. However, it cannot easily reflect the intention of the researcher, and the classification system changes each time the research target is changed. A unified methodology that can be applied equally to all areas is required to establish policies such as spatial planning, and a supervised classification method with a top–bottom process based on tourism activity classification criteria is considered suitable.

Image classification using deep-learning technology has become increasingly applicable across various fields. The advancements in deep learning have made it possible to achieve highly accurate results quickly and cost-effectively over large areas. The test accuracy of the model in this study was 0.7422 (Table 6), indicating a significant ability to classify images. This suggests that deep learning-based image classification technology offers a means to overcome the limitations of traditional qualitative analysis.

Most of the training data used in this study for the classification model were derived from a random selection of 10,000 images from the Flickr dataset collected during the data collection phase. Because of the nature of photo-based social networking platforms, there was an uneven distribution of images across the different tourism activity categories. There was an abundance of images related to appreciation and other activities and a relatively low proportion of images corresponding to recreational and educational activities. As shown in Table 6, the classification performance for recreational and educational activities is weaker than that for appreciation and other activities. To address this imbalance, additional data related to recreational and educational activities must be acquired. Considering the relatively higher classification performance achieved for appreciation activities and other activities, which had larger amounts of training data, obtaining a sufficient quantity of training data for recreational and educational activities is crucial for enhancing their classification capabilities.

4.3. Distribution Patterns of Tourism Activities according to Spatial Environments and Hotspots

In this study, a grid-based level analysis was conducted that covered the entire South Korean coast and islands. The data collection target and research area revealed that other (43%) and appreciation (39%) activities accounted for the largest proportion of tourism activities. In the natural environment, such as beach areas, the proportion of appreciation activities increased, and the proportion of other activities decreased. Conversely, in hotspots with high population densities, such as urban centers, there was an increase in other activities and a decrease in appreciation activities. This means that people's tourism activity patterns change depending on the surrounding environment or coastal space.

This pattern was also observed when each hotspot's target city and region were compared. Targets A, B, and D (Incheon, Mokpo, and Busan), which are all cities with relatively large population densities (2764.64–4350.41 people/km^2), had a large proportion of other activities (45–59%) and a small proportion of appreciation activities (25–39%; Figure 7, Table 1). In contrast, targets C, E, and F (Yeosu City, Gangneung City, and Jeju Island), which have relatively low population densities (204.58–540.28 people/km^2), had a small proportion of other activities (33–35%) and a large proportion of appreciation activities. This confirms that coastal areas with high population densities and urban centers tend to have a higher proportion of other activities and a lower proportion of appreciation activities.

In the main tourism activity map, there were differences in the spatial distributions of each tourism activity. The number of grids corresponding to each activity was 146 for recreation, 219 for education, 1419 for appreciation, and 629 for other activities. As shown in Figure 6, the number of appreciation and other activities was approximately the same, but the grid numbers are twice as different. This means that other activities were more densely distributed over smaller areas, whereas appreciation activities were less densely distributed over larger areas. This could also mean that there are spaces with a low number of coastal tourism visitors that are very prominent in appreciation. In this respect, quantitative analyses that estimate only the number of tourists have their limitations. If researchers only focus on areas with high visitor numbers, they will inevitably focus on urban areas, and other natural and scenic resources may be left out. Therefore, it is important to not only estimate the number of visitors but also conduct a qualitative analysis of the content data.

Owing to the nature of the data extracted from Flickr photos, appreciation and other activities accounted for a large proportion of the data, while recreation and education accounted for approximately 20% of the total. Unfortunately, it is difficult to draw meaningful

conclusions regarding recreational and educational activities based on current data and patterns. However, considering the nature of the data, the proportions of recreation and education were likely to be underestimated, and it is believed that certain weights or factors can compensate for this. If we can weigh the data based on field data, we may be able to obtain more meaningful numbers or patterns.

4.4. Combining Quantitative and Qualitative Data to Understand Spatial Context

Spatial planning is the process of identifying the current situation of a space and reconciling conflicting activities and demands. To make a spatial plan for coastal tourism, the current status must be understood. With geotagged social media data and deep-learning models, it is now possible to obtain quantitative and qualitative data information at a fraction of the cost and time.

Taking target C (Yeosu City) as an example, it has a spatial concentration of city halls, train stations, national parks, and industrial complexes, which together form a large hotspot (Figures 4 and A2). The area around the industrial complex in the north of Yeosu is a hotspot, but there are few appreciation activities and many other activities. The area around the Hyangilam hermitage, a tourist destination, is not a hotspot, but counts of other activities and appreciation activities are observed (Figure 5).

By analyzing the visitor and activity distribution layers together, complex information can be derived. Taking appreciation activities as an example, after creating a distribution with the number of appreciation activities and the number of visitors as the axes, the grid type for appreciation activities can be specified by dividing the areas by the axes (Figure 8). In Figure 8, Grid Type I represents grids with a high number of visitors and a large proportion of appreciation activities, Grid Type II represents grids with a low number of visitors but a significant presence of appreciation activities, Grid Type III represents grids with a high number of visitors but a limited number of appreciation activities, and Grid Type IV represents grids with both a low number of visitors and few appreciation activities. In the case of Yeosu City, Grid Type I is mainly concentrated in the vicinity of the national park around Yeosu EXPO Station, while Grid Type II, which occupies a smaller number of grids, is located within the national park area or near the Hangilam heritage site located to the south.
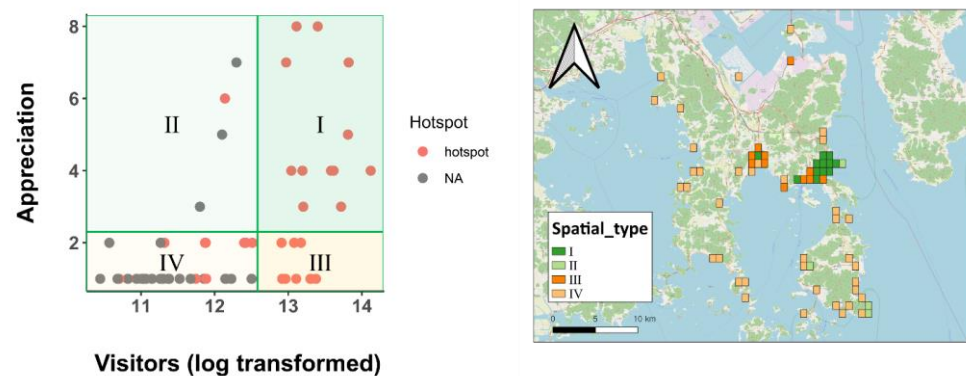
Figure 8. (**Left**) example of the classification types according to the number of coastal tourism visitors and appreciation activity patterns; (**right**) map displaying the spatial distribution of classification types.

Appreciation activities in Grid Types I and II can represent key locations in the management of coastal tourism in terms of coastal conservation and ecosystem services. In particular, Grid Type II is one of the spaces that would have been overlooked based on visitor count data alone and is evidence of why qualitative analyses should be conducted alongside it. Grid Type 1 may be categorized as a space that has good landscape resources but is heavily visited and requires attention for conservation, or conversely it may be a place that should be actively developed to meet current tourism demand. Grid Type 2

could also be considered a potential tourism destination that should be developed. It is up to the decision makers to decide, but the important thing is that they are informative and have context about the space. It is believed that qualitative analysis of data content is essential in order to gain in-depth information about a space in order to make decisions and plans.

5. Conclusions

This study demonstrated that coastal tourism activities can be classified through the content classification of geotagged social media data, and these classifications can then be applied to the study area, which includes the coastal land and island areas of South Korea. Qualitative information on coastal tourism obtained in this manner can be combined with quantitative information on the number of coastal tourists to reveal various spatial information and meanings. We found that the supervised classification model for tourism activities developed via transfer learning of deep-learning models has a high accuracy. The classification results can be combined with geotagged information to provide the various types of information. Through the data analysis of this study, the following findings were demonstrated: (1) approximately 43% of the SNS (Flickr platform) image data in the study area are other activities, 39% are appreciation activities, and 18% are recreation and education activities; (2) spatial hotspots are mainly formed in urban centers or around famous tourist attractions; and (3) natural environments such as beaches have more appreciation activities and fewer other activities, and the opposite trend is observed in urban areas with a high population density. Cities with high population densities have a higher proportion of other activities and a smaller proportion of appreciation activities than other cities. The combination of quantitative and qualitative information can be used to understand the context of the space.

This research contributes to the literature by introducing a novel methodological approach that leverages deep-learning and transfer-learning techniques for the classification of coastal tourism activities through geotagged social media data. Unlike previous studies that primarily relied on quantitative metrics, this study integrates both qualitative and quantitative data, offering a more comprehensive understanding of spatial tourism dynamics. By bridging qualitative spatial patterns with large-scale social media data, we are pushing the boundaries of how cultural ecosystem services, particularly in coastal regions, are quantified and interpreted, thus addressing a critical gap in both tourism and environmental studies.

The limitations of this study were as follows. First, although we used two SNS platforms and other data to estimate the number of tourists to control for data bias, we could only use data from Flickr for the tourism activity classification process because of data acquisition problems. Future classification studies using different platforms may yield less biased results. Second, the data and analyses were limited to coasts and islands. Therefore, the impact of the land adjacent to the coast was not considered in this study. Future studies that include all contiguous spaces will provide a more accurate picture of the factors driving these patterns.

When social media data initially emerged as a method for quantifying cultural ecosystem services in 2013, there was much debate. However, more than a decade later, social media data have been recognized as a useful tool for quantifying tourism and cultural ecosystem services that were once considered difficult to quantify. Although this study reaffirms previous quantitative estimation methods, it also presents new approaches and implications through qualitative estimation and discussion. As each piece of information is connected through space, complex tourism patterns and types can be identified. It is said that new insights can be obtained based on connections with old knowledge; thus, combining social media data with spatial analysis allowed us to discover new facts and expand their applicability.

Author Contributions: Conceptualization, Gang Sun Kim, Choong-Ki Kim, and Woo-Kyun Lee; Methodology, Gang Sun Kim and Choong-Ki Kim; Validation, Gang Sun Kim; Formal analysis, Gang Sun Kim; Investigation, Gang Sun Kim; Resources, Choong-Ki Kim; Data curation, Gang Sun Kim; Writing—original draft preparation, Gang Sun Kim; Writing—review and editing, Gang Sun Kim, Choong-Ki Kim, and Woo-Kyun Lee; Visualization, Gang Sun Kim; Supervision, Choong-Ki Kim and Woo-Kyun Lee; Project administration, Choong-Ki Kim; Funding acquisition, Choong-Ki Kim. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.
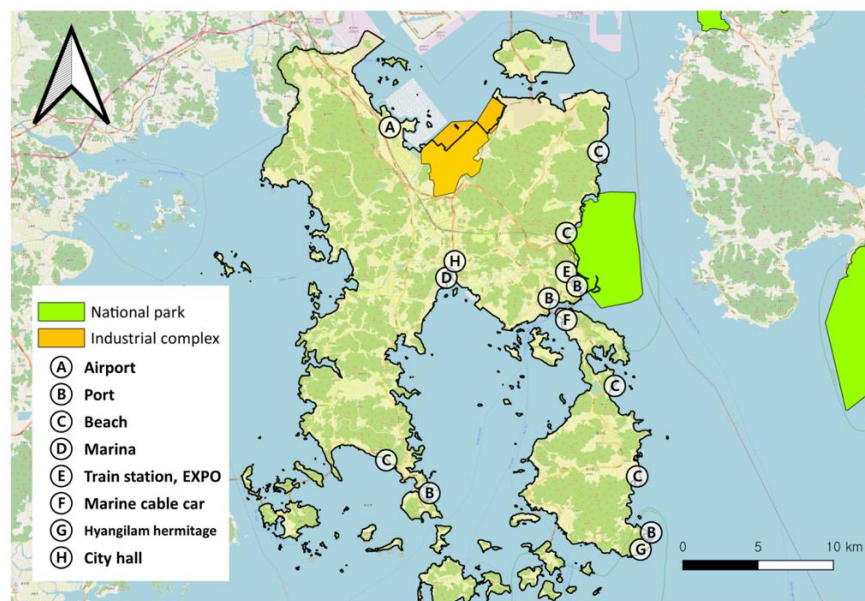
Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A
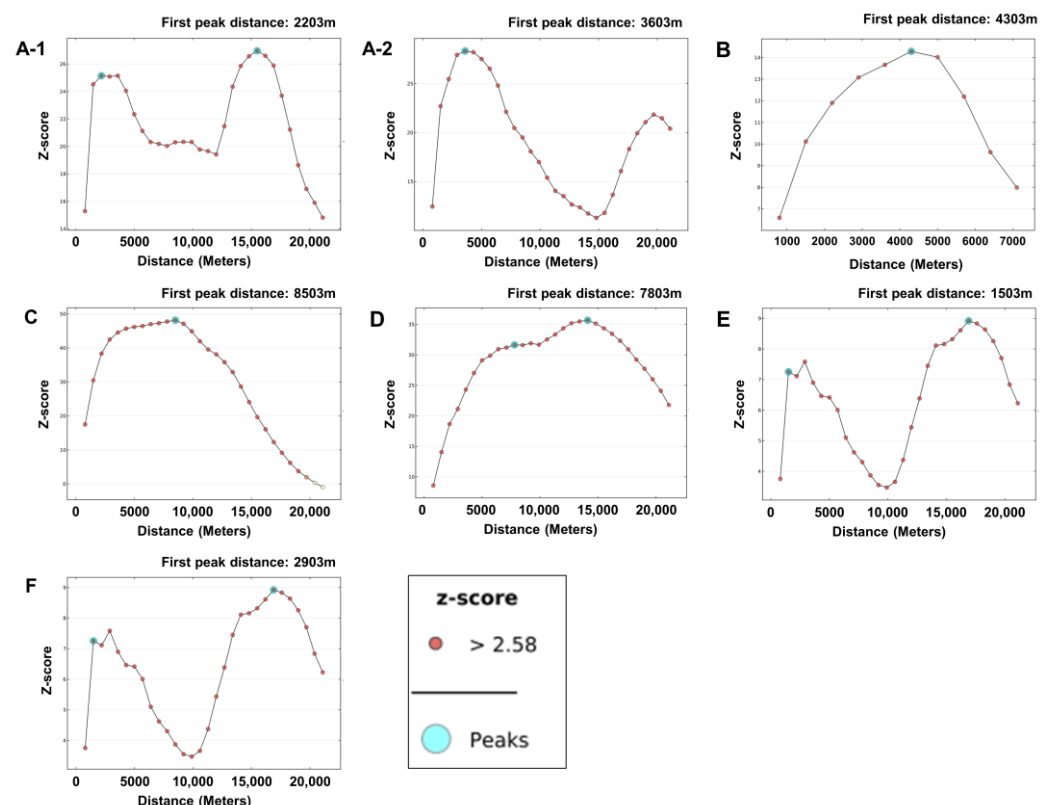
Figure A1. Z-score of Moran's I and first peak distance according to the increase in distance for each hotspot target ((**A-1**) is the north side of Incheon City; (**A-2**) is the north side of Incheon City; (**B**) is Mokpo City; (**C**) is Yeosu City; (**D**) is Busan City; (**E**) is Gangeung City; and (**F**) is Jeju Island).

Figure A2. Locations of national parks, industrial complexes, tourist attractions, and transportation facilities in Yeosu.

References

1. Arkema, K.K.; Verutes, G.M.; Wood, S.A.; Clarke-Samuels, C.; Rosado, S.; Canto, M.; Rosenthal, A.; Ruckelshaus, M.; Guannel, G.; Toft, J.; et al. Embedding ecosystem services in coastal planning leads to better outcomes for people and nature. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7390–7395. [CrossRef] [PubMed]
2. Hall, C.M. Trends in ocean and coastal tourism: The end of the last frontier? *Ocean Coast. Manag.* **2001**, *44*, 601–618. [CrossRef]
3. Papageorgiou, M. Coastal and marine tourism: A challenging factor in Marine Spatial Planning. *Ocean Coast. Manag.* **2016**, *129*, 44–48. [CrossRef]
4. Priskin, J. Assessment of natural resources for nature-based tourism: The case of the Central Coast Region of Western Australia. *Tour. Manag.* **2001**, *22*, 637–648. [CrossRef]
5. Wardana, I.M.; Utama, I.W.M.; Astawa, I.P. Model of local population perception in supporting coastal tourism development and planning in Bali. *GeoJournal Tour. Geosites* **2018**, *23*, 873.
6. Assessment, M.E. *Ecosystems and Human Well-Being: A Framework for Assessment*; World Resources Institute: Washington, DC, USA, 2003.
7. United Nations. *Transforming Our World: The 2030 Agenda for Sustainable Development*; General Assembly, 70 Session; United Nations: New York, NY, USA, 2015.
8. Lubchenco, J.; Haugan, P.M. Coastal development: Resilience, restoration and infrastructure requirements. In *The Blue Compendium: From Knowledge to Action for a Sustainable Ocean Economy*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 213–277.
9. Smith, H.D.; Maes, F.; Stojanovic, T.A.; Ballinger, R.C. The integration of land and marine spatial planning. *J. Coast. Conserv.* **2011**, *15*, 291–303. [CrossRef]
10. Yamagata, Y.; Yang, P.P.J. *Urban Systems Design: Creating Sustainable Smart Cities in the Internet of Things Era*; Elsevier: Amsterdam, The Netherlands, 2020.
11. Zaucha, J.; Gee, K. *Maritime Spatial Planning: Past, Present, Future*; Springer Nature: Berlin/Heidelberg, Germany, 2019.
12. Chua, A.; Servillo, L.; Marcheggiani, E.; Moere, A.V. Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy. *Tour. Manag.* **2016**, *57*, 295–310. [CrossRef]
13. Deville, P.; Linard, C.; Martin, S.; Gilbert, M.; Stevens, F.R.; Gaughan, A.E.; Blondel, V.D.; Tatem, A.J. Dynamic population mapping using mobile phone data. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15888–15893. [CrossRef]
14. Sonter, L.J.; Watson, K.B.; Wood, S.A.; Ricketts, T.H. Spatial and temporal dynamics and value of nature-based recreation, estimated via social media. *PLoS ONE* **2016**, *11*, e0162372. [CrossRef]
15. Wood, S.A.; Guerry, A.D.; Silver, J.M.; Lacayo, M. Using social media to quantify nature-based tourism and recreation. *Sci. Rep.* **2013**, *3*, 2976. [CrossRef]
16. Li, J.; Xu, L.; Tang, L.; Wang, S.; Li, L. Big data in tourism research: A literature review. *Tour. Manag.* **2018**, *68*, 301–323. [CrossRef]
17. Kim, G.S.; Chun, J.; Kim, Y.; Kim, C.K. Coastal tourism spatial planning at the regional unit: Identifying coastal tourism hotspots based on social media data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 167. [CrossRef]
18. Li, H.; Calder, C.A.; Cressie, N. Beyond Moran's I: Testing for spatial dependence based on the spatial autoregressive model. *Geogr. Anal.* **2007**, *39*, 357–375. [CrossRef]

19. Cho, N.; Kang, Y.; Yoon, J.; Park, S.; Kim, J. Classifying tourists' photos and exploring tourism destination image using a deep learning model. *J. Qual. Assur. Hosp. Tour.* **2022**, *23*, 1480–1508. [CrossRef]

20. Hong, T.; Choi, J.A.; Lim, K.; Kim, P. Enhancing personalized ads using interest category classification of SNS users based on deep neural networks. *Sensors* **2020**, *21*, 199. [CrossRef]

21. Kang, Y.; Cho, N.; Yoon, J.; Park, S.; Kim, J. Transfer learning of a deep learning model for exploring tourists' urban image using geotagged photos. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 137. [CrossRef]

22. Heikinheimo, V.; Minin, E.D.; Di Tenkanen, H.; Hausmann, A.; Erkkonen, J.; Toivonen, T. User-generated geographic information for visitor monitoring in a national park: A comparison of social media data and visitor survey. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 85. [CrossRef]

23. Christou, E.; Chatzigeorgiou, C. Adoption of social media as distribution channels in tourism marketing: A qualitative analysis of consumers' experiences. *J. Tour. Herit. Serv. Mark.* **2020**, *6*, 25–32.

24. Almeida, J.; Costa, C.; Nunes da Silva, F. A framework for conflict analysis in spatial planning for tourism. *Tour. Manag. Perspect.* **2017**, *24*, 94–106. [CrossRef]

25. Drius, M.; Bongiorni, L.; Depellegrin, D.; Menegon, S.; Pugnetti, A.; Stifter, S. Tackling challenges for Mediterranean sustainable coastal tourism: An ecosystem service perspective. *Sci. Total Environ.* **2019**, *652*, 1302–1317. [CrossRef]

26. Giglio, S.; Bertacchini, F.; Bilotta, E.; Pantano, P. Using social media to identify tourism attractiveness in six Italian cities. *Tour. Manag.* **2019**, *72*, 306–312. [CrossRef]

27. Richards, D.R.; Tunçer, B.; Tunçer, B. Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosyst. Serv.* **2018**, *31*, 318–325. [CrossRef]

28. Song, X.P.; Richards, D.R.; Tan, P.Y. Using social media user attributes to understand human–environment interactions at urban parks. *Sci. Rep.* **2020**, *10*, 808. [CrossRef]

29. Bondarenko, M.; Kerr, D.; Sorichetta, A.; Tatem, A. *Census/Projection-Disaggregated Gridded Population Datasets, Adjusted to Match the Corresponding UNPD 2020 Estimates, for 183 Countries in 2020 Using Built-Settlement Growth Model (BSGM) Outputs*; WorldPop, University of Southampton: Southampton, UK, 2020.

30. DCGISgroup. Esri World Imagery. Available online: https://www.arcgis.com/home/item.html?id=50c23e4987a44de4ab163e1baeab4a46 (accessed on 30 July 2024).

31. Haklay, M.; Weber, P. OpenStreetMap: User-generated street maps. *IEEE Pervasive Comput.* **2008**, *7*, 12–18. [CrossRef]

32. Ministry of Oceans and Fisheries. Coastal Portal. Available online: https://coast.mof.go.kr/main.do (accessed on 30 July 2024).

33. GEOSERVICE. Download South Korea's Latest Administrative Divisions (SHP). GIS Developer. Available online: http://www.gisdeveloper.co.kr/?p=2332 (accessed on 30 July 2024).

34. Hawelka, B.; Sitko, I.; Beinat, E.; Sobolevsky, S.; Kazakopoulos, P.; Ratti, C. Geo-located Twitter as proxy for global mobility patterns. *Cartogr. Geogr. Inf. Sci.* **2014**, *41*, 260–271. [CrossRef] [PubMed]

35. Kim, Y.H. *SNS (Social Network Service) Usage Trend and Usage Behavior Analysis*; Korea Information Society Development Institute: Jincheon, Republic of Korea, 2019.

36. Nam, J.; Khim, J.S.; Kim, C.-K.; Ryu, J.; Yoo, S.-H.; Jung, S.; Song, Y.S. *Marine Ecosystem-Based Analysis and Decision-Making Support System Development for Marine Spatial Planning*; Korea Institute of Marine Science & Technology Promotion: Seoul, Republic of Korea, 2022.

37. Fox, N.; August, T.; Mancini, F.; Parks, K.E.; Eigenbrod, F.; Bullock, J.M.; Sutter, L.; Graham, L.J. "photosearcher" package in R: An accessible and reproducible method for harvesting large datasets from Flickr. *SoftwareX* **2020**, *12*, 100624. [CrossRef]

38. Selenium. The Selenium Browser Automation Project. *Selenium*. Available online: https://www.selenium.dev/documentation/ (accessed on 30 July 2024).

39. Stevens, F.R.; Gaughan, A.E.; Linard, C.; Tatem, A.J. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PLoS ONE* **2015**, *10*, e0107042. [CrossRef]

40. Kim, G.S. A Geospatial Approach to Quantifying Coastal Tourism Activity and Scale Using Social Media Data. Doctoral Dissertation, Korea University, Seoul, Republic of Korea, 25 August 2023.

41. Harris, N.L.; Goldman, E.; Gabris, C.; Nordling, J.; Minnemeyer, S.; Ansari, S.; Lippmann, M.; Bennett, L.; Raad, M.; Hansen, M.; et al. Using spatial statistics to identify emerging hot spots of forest loss. *Environ. Res. Lett.* **2017**, *12*, 024012. [CrossRef]

42. Isobe, A.; Uchida, K.; Tokai, T.; Iwasaki, S. East Asian seas: A hot spot of pelagic microplastics. *Mar. Pollut. Bull.* **2015**, *101*, 618–623. [CrossRef]

43. Lepers, E.; Lambin, E.F.; Janetos, A.C.; DeFRIES, R.; Achard, F.; Ramankutty, N.; Scholes, R.J. A synthesis of information on rapid land-cover change for the period 1981–2000. *BioScience* **2005**, *55*, 115. [CrossRef]

44. Esri. How Hot Spot Analysis (Getis-Ord Gi*) Works. ArcGIS pro Geoprocessing Tool Reference. Available online: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm (accessed on 30 July 2024).

45. Getis, A.; Ord, J.K. The analysis of spatial association by Use of Distance Statistics. *Geogr. Anal.* **1992**, *24*, 189–206. [CrossRef]

46. Moran, P.A.P. Notes on continuous stochastic phenomena. *Biometrika* **1950**, *37*, 17–23. [CrossRef]

47. Song, M.; Peng, J.; Wang, J.; Zhao, J. Environmental efficiency and economic growth of China: A Ray slack-based model analysis. *Eur. J. Oper. Res.* **2018**, *269*, 51–63. [CrossRef]

48. Esri. How Spatial Autocorrelation (Global Moran's I) Works. ArcGIS pro Geoprocessing Tool Reference. Available online: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm (accessed on 30 July 2024).
49. Seok-Joong, J.; Mi-Hye, L. *The Theory of Ocean Tourism*; Daewangsa: Seoul, Republic of Korea, 2004.
50. Sung-gwi, K. *The Theory of Ocean Tourism*; Hyunhaksa: Seoul, Republic of Korea, 2007.
51. ul Hassan, M.; Neurohive. VGG16—Convolutional Network for Classification and Detection. Available online: https://neurohive.io/en/popular-networks/vgg16/ (accessed on 30 July 2024).
52. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:2014:1409.1556.
53. François, C. Keras. GIThub. Available online: https://github.com/fchollet/keras (accessed on 30 July 2024).
54. Fisher, D.M.; Wood, S.A.; Roh, Y.-H.; Kim, C.-K. The geographic spread and preferences of tourists revealed by user-generated information on Jeju island, South Korea. *Land* **2019**, *8*, 73. [CrossRef]
55. García-Palomares, J.C.; Gutiérrez, J.; Mínguez, C. Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS. *Appl. Geogr.* **2015**, *63*, 408–417. [CrossRef]
56. 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]