Sensors and Materials, Vol. 36, No. 7 (2024) 2867–2878
MYU Tokyo

S & M 3706

Sensor Network Based on Machine Learning in Tourist Area

Chengxiang Wang,[1,2†] Qilong Chen,[1†] Pingrong He,[1]
Wei-Ling Hsu,[3*] and Hsin-Lung Liu[4**]

[1]Institute of Land and Urban-Rural Planning, Huaiyin Normal University,
Wangyin, Huaiyin, Huai'an 223300, China
[2]Department of Urban and Regional Development, Hanyang University,
Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea
[3]School of Civil Engineering, Jiaying University,
Meigong, Meizhou 514015, China
[4]Department of Leisure Management, Minghsin University of Science and Technology,
No. 1, Xinxing Road, Xinfeng Hsinchu, Taiwan 30401, China

(Received January 3, 2024; accepted June 21, 2024)

Keywords: traffic flow, tourist area, machine learning, tourist awareness

The sensor network is an important part of the real-time monitoring system of traffic flow in tourist areas. In this study, we built a traffic flow prediction model based on machine learning to assist the construction and optimization of a sensor network. Using the sensor network, we developed a real-time monitoring system of traffic flow in tourist areas. Factors affecting traffic flow in the study area (Hongze Lake, China) were determined through interviews with 50 tourists. Using the scores of the factors, machine learning models such as random forest, decision tree, and support vector machine were constructed to predict traffic flow, and the result was used to design an appropriate sensor network. The predictions of three machine learning methods were compared to build a traffic flow prediction model. After comparing the predicted results with the tested ones, the performance of the prediction model was validated. Referring to such results, we selected the area near a square and bridges and the intersections of roads as key areas to install sensors for monitoring traffic flow. The system is being implemented in the study area to renovate the area to promote further growth of the tourism industry.

1. Introduction

The number of tourists is subject to temporal fluctuations and spatial imbalance in tourist destinations.[1,2] In tourism planning and the design of tourist areas, the spatial and temporal monitoring of traffic flow is essential to meet the needs of tourists and provide appropriate services and facilities. From the perspective of tourism management, in scenic areas, the spatial monitoring of traffic flow even prevents excessive gathering to avoid unexpected events such as stampede.[3–5] Thus, it is necessary to construct a prediction model of traffic flow to accurately

*Corresponding author: e-mail: quartback@hotmail.com
**Corresponding author: e-mail: hsinlung@must.edu.tw
†These authors contributed equally to this work and should be considered co–first authors.
https://doi.org/10.18494/SAM4887

ISSN 0914-4935
https://myukk.org/

forecast it. In running the model, a sensor network is mandatory for the effective prediction and monitoring of traffic flow.

The application of the Internet of Things (IoT) encompasses transportation, medical care, senior care, and education with smart sensors.[6–8] Traffic flow monitoring in tourist areas also demands the use of smart sensors to gather information using remote sensing or video surveillance equipment.[5,9] Because of the limited detection range of video surveillance equipment, it is necessary to set up multiple devices to form a sensor network to cover the entire tourist area.

The sensor network is the "eye" of the real-time traffic flow monitoring system in tourist areas, and the system monitors the traffic flow of all sections of tourist areas in real time by means of the sensor network. Using sensor networks, a real-time monitoring system of traffic flow in tourist areas is developed to assist the management and construction of tourist areas. When a sensor detects that the local flow of people has reached a warning level, it will trigger the alarm device of the system, enabling the tourist area manager to conduct traffic guidance in a timely manner. In addition, the traffic data collected by the sensor network can continue to train the machine learning model, which gradually increases the accuracy of the traffic flow prediction model.[10,11] With the forecast data, managers can realize the accurate delivery of public facilities in tourist areas and increase management personnel in advance in possible crowded areas.

Satellite remote sensing equipment is appropriate for urban tourist areas,[12] while in smaller scenic spots, video surveillance equipment is more suitable. For the prediction of traffic flow in such areas, regardless of the data source, the data were processed with space syntax theory to calculate the degree of crowdedness. The areas with a high degree of crowdedness can be regarded as congested areas. In these areas, sensors are operated to track people's movements and traffic flow. The processed data from such sensors can be used to prevent congestion in an area and avoid crimes and unexpected events such as stampede.[13] In previous research, road segment models were constructed using space syntax theory. Selectivity was calculated to determine traffic flow on the road. However, the models ignored the impact of ambiance.[14] Therefore, it is necessary to integrate various factors of ambiance and commercial service in a model to improve the accuracy of traffic flow prediction. In tourist areas, a sensor network should be designed for the data collection and monitoring of traffic flow.

In this study, we added the concept of the point-of-interest (POI) factor to conventional factors to predict traffic flow. The accuracy of the prediction was verified using machine learning algorithms including decision tree, random forest, and support vector machine. An optimal machine learning algorithm was then selected to predict traffic flow in tourist areas. The model and its result provide a reference for sensor network design and the construction of an efficient and intensive system for traffic flow prediction.

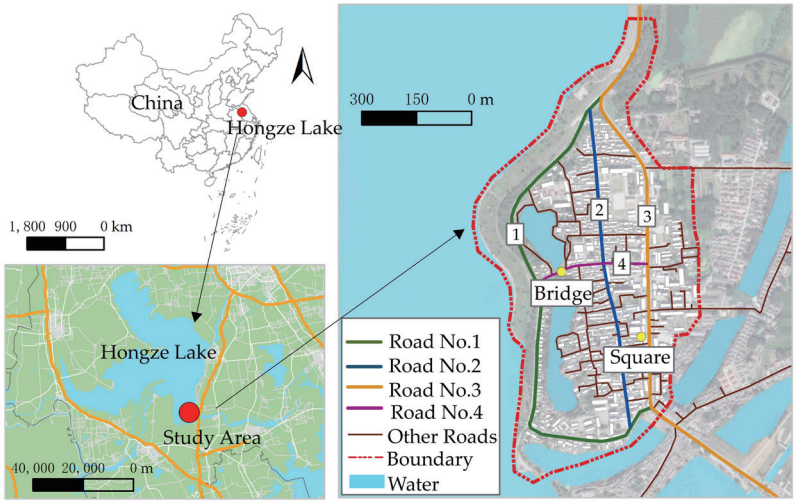
2. Method

2.1 Study area

The study area was the Jiangba Ancient Town Scenic Area in China, located along Hongze Lake, the fourth-largest freshwater lake in China (Fig. 1). On the west side, there is the Hongze Lake Embankment. The embankment was built more than 1800 years ago during the Eastern Han Dynasty. It is 67 km long and is known as the "Great Wall on the Water." Spring and autumn are the peak tourist seasons. Most tourists arrive by car and walk around the area. Facilities related to tourism include a stage square, small playgrounds, Yonghe Bridge, restaurants, and accommodations. The area has been renovated recently with the construction of new roads and tourist service facilities. As part of the renovation project, a new intelligent traffic and early warning system is planned for the better management of tourists and traffic flow. By collecting real-time data, tourist behaviors and preferences can be analyzed to transform the area to satisfy visitors.

2.2 Data sources

The data used in this study included traffic flows, POIs, and road segments around the Hongze Lake. POI was determined after assessing the current situation. The data on road segments were extracted from the topographic map provided by the government and adjusted for the current situation. Remote sensing images of Hongze Lake and pond were obtained with a resolution of 10 m from the Sentinel-2 satellite over a period of 10 days. 2023 images were collected after cloud removal by the mean synthesis method. A random forest classification method was used to extract water bodies from the images. The classification accuracy was 0.897. Since traffic flow was not observed, current traffic data were collected on the basis of the



(a) (b)

Fig. 1. (Color online) (a) Location and (b) view of study area.

ratings of 50 first-time visitors to the area. They were asked to suggest two cognitive points about tourism in the area and scored as follows: "I would not visit again" (score 1), "I may come again" (score 2), and "I will come again" (score 3). 94 cognitive points were obtained about the current traffic situation and impression of the area.

2.3 Influencing factors

The 94 cognitive points were categorized into landscape and POI factors. The factors were evaluated using a line segment model based on the space syntax theory,[15,16] which is commonly used in research on traffic flow. In addition to small-scale and large-scale selectivities, landscape factors included the distance to Hongze Lake (D-lake) and the distance to the pond (D-pond), while POI factors included outdoor leisure venues, accommodations, indoor viewing venues, landmarks, stores, and specialty in the industry. Small-scale and large-scale selectivities were evaluated using Depthmap (space syntax calculation software). The search parameter was set to 500 m for the small-scale selectivity, and *n* to assess the large-scale selectivity. The radius of the small-scale selectivity was defined by a 10-min walking distance (500 m). D-lake was the distance from the road point to the interior pond landscape, and D-pond was the distance from the road point to the pond in Hongze Lake. Kernel density values of the six independent variables in POI factors were also calculated using GIS 10.6.

2.4 Machine learning algorithms

2.4.1 Decision tree

A decision tree is a supervised machine learning model[17] that makes judgments and decisions based on a tree structure that consists of root, branch, and leaf nodes(Table 1). The

Table 1
Description of influencing factors and independent variables.

Influencing factors	Independent variable	Description
Traffic factors	Small-scale aspect	Road small-scale selectivity based on space syntax line segment model
	Large-scale aspect	Road large-scale selectivity based on space syntax line segment model
Landscape factors	D-lake	The distance between the road and the landscape of Hongze Lake, Hongze Lake embankment, and the fishing boat pier
	D-pond	Distance from road to interior pond landscape
POI factors	Outdoor leisure venues	Outdoor leisure venues, such as squares and small playgrounds
	Accommodation	Homestay, tourist accommodation facilities that have been artistically renovated
	Indoor viewing venues	Indoor viewing venue, an indoor venue with lake views, leisure seating, and dining and drinking facilities
	Landmarks	Landmarks such as water towers, sculptures, and bridges
	Stores	General stores, including clothing stores, hairdressers, and restaurants
	Specialty in the industry	Specialty industries include fishing net textiles and aquatic product sales

decision-making process is transparent, and the classification results are interpretable owing to its simple, accurate, and intuitive explanation, so the decision tree is widely used in various studies.[18] In every decision, there are multiple possibilities, each of which leads to different results. Classifying an attribute multiple times seems like a tree that continuously branches. The goal of using a decision tree is to find the best splitting method in multiple attempts. Common criteria of the best decision tree include information gain, gain rate, and Gini coefficient. We used the Gini coefficient to classify attributes and design a classification and regression tree (CART), as

$$Gini(D) = 1 - \sum_{i=1}^{k} p_k^2, \tag{1}$$

where p_k presents the proportion of samples of the k category in sample set D, $k = 1, 2, ..., m$.

For sample set D with an attribute a, the Gini coefficient of attribute a is calculated using

$$Gini_{index}(D,a) = \sum_{i=1}^{k} \frac{|D^k|}{|D|} Gini(D^k). \tag{2}$$

$Gini_{index}(D,a)$ represents the probability that the number of randomly selected samples k from sample set D is in a certain category. The higher the purity of the sample, the smaller the $Gini_{index}(D,a)$ value. When building a decision tree, the attribute with the smallest Gini coefficient is prioritized for partitioning. The decision tree classifies nodes from the root to the top. The earlier an attribute is classified, the greater the impact on the prediction result.

2.4.2 Random forest

Random forest is an ensemble learning algorithm. By training a model with sample sets in multiple tree structures, classification and regression are conducted. It has a high level of predictive performance.[19,20] The operation steps of random forest are as follows.

Step 1: Attributes are extracted from one sample at a time n times and replace former ones from the original training set to obtain a subtraining set X' composed of the n samples.

$$X' = (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n) \tag{3}$$

Here, $x_1 \in R^n$ is the input value and $y_1 \in R$ is the output value.

Step 2: m attributes are randomly selected from the h attributes of the sample to form the feature subset T'.

$$T' = \{t_1, t_2, \ldots, t_m\},\ m < h. \tag{4}$$

Step 3: On the basis of X' and T', a CART is built to select a segmentation variable $t \in T'$ and its value, and obtain the optimal t, s combination.

$$\min_{t,s}\left[\min_{c_1}\sum_{x_i\in R_{1(t,s)}}\left(y_i-c_1\right)^2+\min_{c_2}\sum_{x_i\in R_2(t,s)}\left(y_i-c_2\right)^2\right] \tag{5}$$

Here, R_i is the i subregion and c_i is the output value of the i subregion.

Step 4: The selected (t, s) area is divided into categories to calculate the corresponding output values.

$$R_1(t,s)=\{x\mid x_t\le s\}, R_2(t,s)=\{x\mid x_t\le s\}. \tag{6}$$

$$c_m=\frac{1}{N_m}\sum_{x_i\in R_m(t,s)} y_i, x\in R_m, m=1,2. \tag{7}$$

Step 5: Steps 3 and 4 are repeated until the two subregions meet the preset conditions.

Step 6: The input space is divided into M regions R_1, R_2,…, R_M to form a decision tree based on

$$f(x)=\sum_{m=1}^{M} c_m I, x\in R_m. \tag{8}$$

Step 7: Steps 1 to 6 are repeated until a K decision tree is constructed.

Step 8: The output values of each decision tree are overweight-averaged to obtain the final value of the random forest.

$$result_{weight}=\frac{1}{N}\sum_{i=1}^{K} w_i f_i(x) \tag{9}$$

2.4.3 Support vector machine

The support vector machine[21] is a supervised machine learning algorithm. The support vector machine finds the best-fitting hyperplane to build a predictive model. The calculation is carried out as follows.

Step 1: Given input data X and learning target y, the following sets are defined:

$$X=\{X_1,\ldots,X_n\}, y=\{y_1,\ldots,y_n\}. \tag{10}$$

The input X contains multiple features to form a multidimensional space, and the learning target is a binary variable.

Step 2: The linear model $f(x, w)$ is established as

$$f(x,w)=\sum_{j=1}^{m} w_j g_j(x)+b, \tag{11}$$

where w is a set of nonlinear variables and $j = 1, 2, ..., n$.

Step 3: The maximum margin hyperplane is calculated as

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$$
$$\text{subject to } \begin{cases} y_i - f(x_i, w) - b \le \varepsilon + \xi_i^*, \\ f(x_i, w) + b - y_i \le \varepsilon + \xi_i^*, \\ \xi_i \xi_i^* \ge 0, \end{cases} \tag{12}$$

where C is the regularization parameter and ξ_i is a loss function used to estimate the degree to which the sample does not meet the constraints.

2.4.4 Evaluation of model performance

Mean squared error (*MSE*) and mean absolute error (*MAE*) indicate the difference between the predicted and test values of the corresponding point, and are calculated using Eqs. (13) and (14), respectively.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} w_i (y_i - \hat{y}_i)^2, \tag{13}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \left| w_i (y_i - \hat{y}_i) \right|, \tag{14}$$

where w_i is the weight of sample i. *MSE* indicates the penalty for errors, while *MAE* represents the robustness of a model with outliers.

3. Results

3.1 Descriptive statistics

The road segment model was introduced into GIS 10.6 for raster processing. The road data were converted into 11681 rasters in 1 × 1 m^2. Dependent and independent variables affecting traffic flow were assigned to 11681 grid points. Table 2 presents descriptive statistics of the collected data. The maximum, minimum, and 1st, 2nd, and 3rd quartile values of scores of two perspectives and six POIs are presented. D-lake, D-pond, landmarks, stores, and specialty in the industry had values at the 75% quartile. All variables had maximum kernel density values above 90. This showed that the distribution of POIs around the road was uneven, which may cause the traffic flow around POIs to be significantly higher than that in other areas. The covariances of scores of D-lake and D-pond were smaller than those of other variables, which indicated that the scores of D-lake and D-pond showed less dispersion.

Table 2
Descriptive statistics.

Name	Mean	Maximum kernel density	Quartile			Minimum kernel density	Standard deviation	Covariance
			25%	50%	75%			
Traffic flow	513	0	128	415	767	2417	470.88	0.92
Small-scale aspect	1617	0	264	744	1883	11364	2310.08	1.43
Large-scale aspect	28458	0	1514	7350	36825	192097	42138.34	1.48
D-lake	140	1	62	129	206	389	90.41	0.64
D-pond	162	0	49	141	255	570	128.98	0.80
Outdoor leisure venues	4	0	0	0	0	91	13.27	3.58
Accommodation	6	0	0	0	0	94	17.46	3.07
Indoor viewing venues	5	0	0	0	0	95	18.25	3.67
Landmarks	13	0	0	0	1	163	29.14	2.24
Stores	81	0	0	0	78	636	156.56	1.94
Specialty in the industry	13	0	0	0	2	241	33.17	2.51

3.2 Spatial distribution of traffic flow and image factors

The current traffic flow of the study area based on the 94 cognitive points suggested by 50 participants is shown in Fig. 2(a). Congested areas were observed around the square and bridges, and were distributed within the intersections of Roads one and three. The 11681 road raster data samples were classified by the natural breakpoint method. The data for 10 independent variables were visualized to display the traffic flow. As shown in Fig. 2(b), large-scale and small-scale selectivities indicated that Roads two and three were mainly selected. Road three was chosen for large-scale selectivity (greater than a 500 m search radius), and Road two was selected for small-scale selectivity (less than a 500 m search radius). Road three was a densely populated area with a higher traffic flow, while Road two was a densely populated area with a lower traffic flow. There was a significant difference between the traffic selectivity calculated by the line segment model owing to the road structure and the traffic flow indicated as tourists' cognitive points. This implies that in a real situation, in addition to the road network and conditions, other factors affect traffic flow. Figure 2(c) shows that the shortest distance to reach Hongze Lake (D-lake) is Road one. There is a road around the lake where Hongze Lake and Hongze Lake Dam can be seen. The shortest distance to the inner pond (D-pond) was along an inner branch road around the pond and parts of Roads one and Road four. At several intersections on Road one, tourists can see Hongze Lake and the pond in overlapped landscapes. At the western intersection of Road four, two northern reservoirs and Hongze Lake can be observed at the same time. Figures 2(a) and 2(c) show that intersections where beautiful natural landscapes can be observed had a higher traffic flow. The kernel density values of outdoor leisure venues (POI_1) were found around the square [Fig. 2(d)]. Because the small playground was away from the road, no obvious high-kernel-density areas were identified. Accommodation (POI_2) was mainly located at the south and north ends of Road two and the north end of Road three. Indoor viewing venues (POI_3) were mainly located in the north and south parts of Road one. At the north end of Road one, tourists can see the Hongze Lake embankment, while at the south end, fishing boats near the pier are found. Landmarks (POI_4) included the water tower near Road three, the sculpture

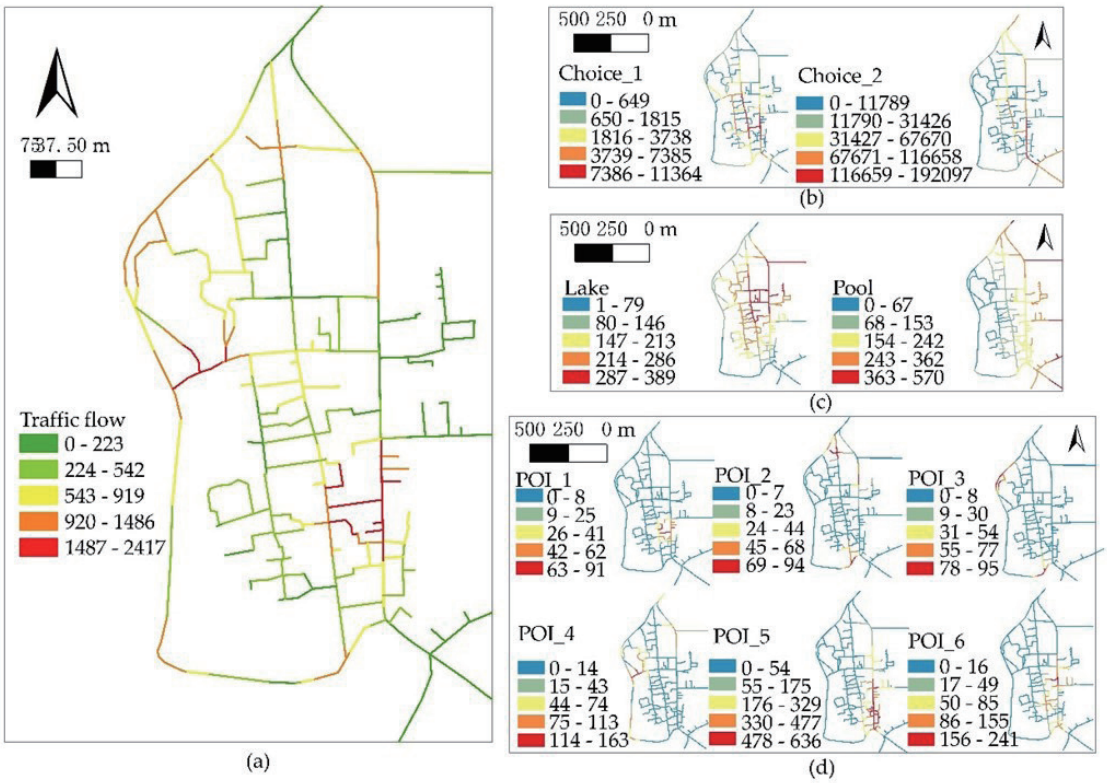


Fig. 2. (Color online) (a) Traffic flow and kernel densities (b) by large-scale and small-scale selectivities and of (c) D-lake and D-pond and (d) POIs.

of Hongze Lake, and the bridge over the internal pond. Stores (POI_5) are mainly located on both sides of Road three but are more densely distributed in the south part of Road three. Specialty in the industry (POI_6) of fishing net textiles and aquatic product stores were located in the south part of Roads two and three. A high kernel density was found for POIs in high-traffic-flow areas. Thus, at POIs, there was high traffic and human flow on the road.

3.3 Model training and selection

Using the dataset, machine learning models, including decision tree, random forest, and support vector machine, were trained to compare their prediction results. *MSE* and *MAE* were used to evaluate the model prediction accuracy. The smaller the *MSE* and *MAE* values, the smaller the error between the predicted and test values. The evaluation results are shown in Table 3. Comparing the *MSE* values of the three models reveals that the random forest model showed the smallest *MSE* followed by the decision tree model. According to Table 2, the mean of traffic flow is 513, and the *MAE* values of the decision tree and random forest models are 4.41 and 5.82, respectively, indicating that the errors of the two models are less than 1.2%. The errors of the first two models were within the acceptable range, but the *MSE* of the support vector

Table 3
Model performance evaluation.

	Decision tree	Random forest	Support vector machine
MSE	476.04	179.30	233408.37
MAE	4.41	5.82	357.24

machine model was 223408.37. Therefore, the support vector machine model was not appropriate for the prediction of traffic flow in the study area. The *MAE* values of the decision tree and random forest models were similar and significantly smaller than that of the support vector machine model. The decision tree model showed a slightly smaller *MAE* than the random forest model but not significantly. Therefore, the random forest model with fewer outliers was selected for prediction in this study.

3.4 Traffic flow forecast

The reconstruction plan of the area is presented in Fig. 3(a). The road network around the pond is mainly reconstructed. New roads are being constructed alongside the pond to increase the accessibility of tourists to the east side of the pond and increase the number of indoor viewing venues along the pond. In addition to that, stores and landmarks are being added alongside the lake. According to the prediction of our model [Fig. 3(b)], traffic and human flows

Fig. 3. (Color online) (a) Tourism area reconstruction plan and (b) traffic flow prediction after reconstruction.

are inferred to be concentrated in the square and on the roads around the bridge near the central intersection of Road one and the northern intersection of Road three. Therefore, in the sensor network design, the square and bridges are the key areas for installing sensors for monitoring (key monitoring areas). The central intersection of Road one and the northern intersection of Road three must also be important areas for installing sensors (sub-key monitoring area). Therefore, in the design of sensor network, the square and bridge are key monitoring areas, the central intersection of Road one and the northern intersection of Road three are sub-key monitoring areas, and other areas are general monitoring areas.

According to the preliminary forecast of key, sub-key, and general monitoring areas, a sensor network was constructed to detect traffic flow. Using sensor networks, a real-time monitoring system of traffic flow in tourist areas was developed to assist the management and construction of tourist areas. The machine learning model was further trained using the traffic data collected by the sensor network to improve the accuracy of our prediction model.

4. Discussion

The existing research on tourist volume in tourist areas focuses on the analysis of the tourist volume in scenic spots in time and space, most of which are from the statistical perspective.[2,22] A little attention is paid to the study of road space passenger flow in tourist areas, and traffic flow is predicted by applying spatial syntax theory. Our method can predict traffic flow from the perspective of vehicles and pedestrians.[13,16] On the basis of the results, we improved the prediction model by introducing factors of tourist selectivity, distances to scenic spots, and POIs. After machine learning, a new prediction model was formulated, and the improved model had a far higher prediction accuracy.

5. Conclusions

In this study, three machine learning methods (random forest, decision tree, and support vector machine) were trained to construct traffic flow prediction models. The results show that the error difference of traffic flow prediction models constructed using random forest and decision tree is less than 2%, which proves the feasibility of machine learning applied to traffic flow prediction in tourist areas. The traffic flow prediction model based on machine learning can assist the construction and optimization of sensor networks. After comparing the model parameters, the traffic flow prediction model constructed by random forest was found to have fewer abnormal values. This model was selected to assist the construction of the sensor network, and the random forest prediction model was trained using real-time monitoring data of human flow after completion, which improved the prediction accuracy and assisted the optimization of the sensor network. The results of the traffic flow prediction model constructed by random forest showed that squares and bridges and the central intersection of Road one and the northern intersection of Road three had higher kernel densities and traffic flows. According to traffic flow prediction, the key, sub-key, and general monitoring areas were determined to be squares, bridges, the central intersection of Road one and the northern intersection of Road three, and the

remaining areas, respectively. On the basis of these results, sensor networks were constructed. Using a sensor network, we developed a real-time monitoring system of traffic flow in tourist areas. In the follow-up study, we plan to integrate street-view photographs to further improve the prediction model and accuracy. Machine learning will be used to analyze street-view photographs, and factors such as the color richness, color saturation, and emotional atmosphere of the scene will be added to the prediction model to further improve its accuracy.

References

1 O. Saenz-de-Miera and J. Rosselló: Tourism Manage. **33** (2012) 466. https://doi.org/10.1016/j.tourman.2011.06.015
2 S. Du, A. A. Bahaddad, M. Jin, and Q. Zhang: Fractals **30** (2022) 2240103. https://doi.org/10.1142/S0218348X2240103X
3 J. Wang, X. Huang, Z. Gong, and K. Cao: J. Destination Mark. Manage. **15** (2020) 100383. https://doi.org/10.1016/j.jdmm.2019.100383
4 I. P. Albaladejo, M. I. González-Martínez, and M. P. Martínez-García: Tourism Manage. **53** (2016) 132. https://doi.org/10.1016/j.tourman.2015.09.018
5 Y. Xiao, X. Tian, and M. Xiao: Engineering **12** (2020) 194. https://doi.org/10.4236/eng.2020.123016
6 W.-L. Hsu, K. Lan, Z. Dong, H.-L. Liu, S. Y.-R. Yang, and B.-W. Jeang: Sens. Mater. **35** (2023) 2783. https://doi.org/10.18494/SAM4457
7 K. Lin, Y. Li, J. Sun, D. Zhou, and Q. Zhang: Inf. Fusion **57** (2020) 15. https://doi.org/10.1016/j.inffus.2019.11.001
8 S. Khosravi, S. G. Bailey, H. Parvizi, and R. Ghannam: Sensors **22** (2022) 7633. https://doi.org/10.3390/s22197633
9 Y. Gao and J.-D. Schmöcker: Transportation Research Part C: Emerging Technologies **160** (2024) 104531. https://doi.org/10.1016/j.trc.2024.104531
10 I. Hammad and K. El-Sankary: Sensors **19** (2019) 3491. https://doi.org/10.3390/s19163491
11 M. O. Costa, R. M. Campos, and C. G. Soares: Ocean Eng. **287** (2023) 115724. https://doi.org/10.1016/j.oceaneng.2023.115724
12 Y. Tang: IOP Conference Series: Earth and Environ. Sci. **638** (2021) 012081. https://doi.org/10.1088/1755-1315/638/1/012081
13 J. Zhou and P. Wang: Physics and Chem. Earth, Parts A/B/C **132** (2023) 103485. https://doi.org/10.1016/j.pce.2023.103485
14 Z. He, N. Deng, X. R. Li, and H. Gu: J. Travel Res. **61** (2022) 597. https://doi.org/10.1177/0047287521995134
15 Y. O. Kim and A. Penn: Environ. Behav. **36** (2004). https://doi.org/10.1177/0013916503261384
16 T. Du, H. Vejre, C. Fertner, and P. Xiang: Sustainability **12** (2020) 33. https://doi.org/10.3390/su12010033
17 J. R. Quinlan: Mach. Learn. **1** (1986) 81. https://doi.org/10.1023/A:1022643204877
18 Z. He, Z. Wu, G. Xu, Y. Liu, and Q. Zou: IEEE Trans. Knowl. Data Eng. **35** (2021) 251 https://doi.org/10.1109/tkde.2021.3075023
19 J. Sun, H. Yu, G. Zhong, J. Dong, S. Zhang, and H. Yu: IEEE Trans. Cybern. **52** (2022) 205. https://doi.org/10.1109/tcyb.2020.2972956
20 J. S. Rhodes, A. Cutler, and K. R. Moon: IEEE Trans. Pattern Anal. Mach. Intell. **45** (2023) 10947. https://doi.org/10.1109/tpami.2023.3263774
21 J. Liu, X. Huang, X. Liu, and Z. Han: Sens. Mater. **31** (2019) 2155. https://doi.org/10.18494/SAM.2019.2314
22 G. Harb and C. Bassil: Current Issues in Tourism **23** (2020) 666. https://doi.org/10.1080/13683500.2018.1544612