



Tourism demand forecasting of multi-attractions with spatiotemporal grid: a convolutional block attention module model

Haodong Sun¹ · Yang Yang¹ · Yanyan Chen¹ · Xiaoming Liu¹ · Jiachen Wang²

Received: 29 August 2022 / Revised: 11 March 2023 / Accepted: 20 March 2023 /

Published online: 5 April 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Effective tourist demand forecasting is crucial for company operations and destination management. Furthermore, tourists may plan better personalized multi-attraction itineraries based on demand forecasting to avoid travel peaks and improve the enjoyment of their vacation. This study developed a unique deep learning model called the convolution block attention module (CBAM) that is built on convolutional blocks and attention modules to estimate tourism demand precisely. Then, the passenger flow grid map was extracted from mobile phone signaling data. To forecast the subsequent period of the passenger flow grid map, the CBAM model uses the multi-channel spatial-temporal grid graph that is built by multiple successive passenger flow grid maps. Finally, the forecasted passenger flow grid map was used to derive the tourist demand for multi-attractions for the next period. The analysis of mobile phone signaling data from Beijing and Xiamen using the proposed model reveals that its mean absolute percentage error (MAPE) is 8.11%, which is lower than other benchmark deep learning models.

Keywords Tourism demand forecasting · Spatial-temporal grids · Convolution block · Attention module

1 Introduction

With the improvement of people's living standards and the rapid development of transportation, tourism will also usher in rapid development. Meanwhile, in the past few decades, tourism has become one of the most important promoter of global economic growth (Hu et al. 2021). As an important prerequisite for tourism-related decision-making, tourism forecasting has gradually attracted the attention

✉ Yanyan Chen
cdyan@bjut.edu.cn

Extended author information available on the last page of the article

of scholars (Palmer et al. 2006; Wu et al. 2019). The application topics of tourism forecasting are generally divided into three categories: event forecasting, future trends/market conditions forecasting and tourism demand forecasting (Lin and Song 2015). The aim of event forecasting is to predict the occurrence of specific long-term or short-term events or trends and to estimate the probability of these events occurring within a specific time period (Ng 1984). Researches on forecasting future trends/market conditions focus on predicting future trends/patterns and assessing the potential impact of certain events or future changes on tourism, such as changes in social values and changes in the structure of tourism (Kaynak and Marandu 2006). Literatures on forecasting tourism demand mainly focus on the arrival of tourists (Sun et al. 2019; Volchek et al. 2019), tourist flow in destination (Huang et al. 2017; Li et al. 2018; Park et al. 2017), hotel accommodation demand (Pan et al. 2012; Yang et al. 2014), tourists' characteristics (Alén et al. 2017; Hernández-López and Cáceres-Hernández 2007; Pomfret and Bramwell 2016) and so on. With the rapid development of international tourism in developed and developing countries, tourism demand forecasting has not only become an important task for national and local planners and decision makers, but also aroused great interest in finding appropriate tourism demand modeling techniques to improve the forecasting accuracy (Song et al. 2003). Therefore, tourism demand forecasting is considered to be the most popular application topic in these three categories.

Given the sensitivity and uncertainty of tourism to many factors, accurate tourism demand forecasting has become particularly important. For example, due to the impact of the novel coronavirus pneumonia, the number of domestic tourists in China in 2020 was 2.879 billion, a decrease of 52.1% from 2019. China's Ministry of Culture and Tourism is trying to revive the tourism industry with various policies. Therefore, with accurate forecasts of future demand, government agencies can make effective decisions on issues such as tourism policy design and implementation, transportation facilities and infrastructure construction. Managers and industry participants need forecasts to make tactical and operational decisions to better allocate resources, such as pricing adjustments, financial investments, staffing and scheduling (Fan et al. 2021; Li et al. 2022). Therefore, developing an effective demand forecasting model is one of the important tasks of tourism research (Bi et al. 2021; Li et al. 2022; Shahrabi et al. 2013).

Various models are used to forecast nonlinear, non-stationary, and complex tourism demand, including traditional time series, econometric models, artificial intelligence, and hybrid models (Li et al. 2022). Time series models are simple and effective, and consider the lag of search engine data as explanatory variables, among which AR, ARMA and ARIMA are the most commonly used time series models (Li et al. 2021). Increasingly advanced econometric models have been incorporated into forecasting tasks to further improve accuracy, including ADL, TVP, DFM, FAAR, Bayesian-FAVAR, and GM (Li et al. 2021). With the development of computer performance, artificial intelligence models including artificial neural networks and support vector regression models are gradually applied to demand forecasting, which are also called "black boxes" in forecasting (Hu and Song 2019; Li et al. 2018, 2019; Zhang et al. 2017a). Although many tourism demand prediction models have been proposed, most of them can only establish a one-dimensional mapping relationship

between each observation within the lag order and the observation to be predicted (Bi et al. 2021). Therefore, there are some important features embedded in the tourism demand data that cannot be fully utilized, which may lead to the loss of some important information in the forecasting process (Bi et al. 2021). As potential exploratory features, spatial features have attracted increasing research interest in recent years (Day et al. 2013; Goh and Law 2003; Li et al. 2022; Yang et al. 2015, 2022). Spatial features mainly include spatial dependence and spatial heterogeneity, which may be triggered by factors such as multi-destination tourism, limited tourism resource sharing, regional cooperation and competition. Due to spatial features, nearby tourism destinations may share (different) similar tourism demand patterns. Therefore, the combination of the spatial effect can improve the prediction accuracy (Chhetri et al. 2013; Fan et al. 2022; Li et al. 2022; Yang and Zhang 2019).

To the authors' knowledge, only seven studies have considered spatial information for tourism demand forecasting. Five of these studies developed spatiotemporal autoregressive models by including spatial lag terms in time series models or econometric models (Jiao et al. 2020; Yang and Zhang 2019). Although their models have achieved good forecasting results, these traditional models have limited nonlinear processing capabilities and hardly avoid error accumulation (Jiao et al. 2021; Long et al. 2018). To overcome the shortcomings of the above studies, Zheng et al. (2021) improved a long short-term memory (LSTM) model to extract temporal dependencies and spatial effects to accurately predict the hourly demand of some attractions in Beijing. Li et al. (2022) proposed a new spatial-temporal fused graph convolutional network (ST-FGCN) model. The model generates forecasts based on the spatial effects extracted by the graph convolutional networks and the temporal dependencies captured through long short-term memory.

There are still certain limitations, despite the fact that their works improve the relevant literature by providing deep learning-based spatiotemporal models. First, LSTM is inherently built to capture temporal dependencies. It is difficult to extract deep and hidden spatial information using the LSTM model even if the input variables have been weighted by a given spatial relationship (Li et al. 2022). Secondly, the spatial Euclidean distance between attractions cannot fully reflect the real spatial effects. On the one hand, within a region, two attractions with the similar contextual environment and close cooperation in business but far apart can also show a strong spatial correlation. On the other hand, the tourism demand of an attraction is not only spatially correlated with other attractions, but also spatially correlated with surrounding areas, such as hotels and restaurants. Therefore, more spatial information should be used to better characterize the spatial effects.

This study proposes a new convolutional block attention module (CBAM) model to forecast the tourism demand of multi-attractions based on the spatial-temporal grid passenger flow map. First, divides the research city into fixed-size grids, and extract each grid passenger flow based on mobile phone signaling data. Second, each grid map can be regarded as a channel graph, and the passenger flow between adjacent grids will have a certain spatial relationship, the grid map of several consecutive periods are stacked together to form a multi-channel spatial-temporal grid graph. The passenger flow between adjacent channels will have a certain temporal relationship, and the multi-channel spatial-temporal grid graph is used as the input

of the deep learning models. Third, deep learning models such as CNN, ConvLSTM, ConvGRU, and CBAM are used to extract the spatial-temporal features and optimize the forecasting models. Two experimental studies demonstrate the effectiveness of the deep learning models. One is to forecast multi-attractions tourism demand in Beijing and the other is in Xiamen. Since Beijing and Xiamen are famous tourist cities and attract a large number of tourists every year. Beijing is the capital of China and has more world heritage sites than any other city in the world. There are more than 200 attractions open to the public, among which there are many very famous sites, such as the Forbidden City, Temple of Heaven Park, Beihai Park, Summer Palace and Old Summer Palace. In addition, it attracts a large number of tourists, including a total of 260 million visits in 2021. Xiamen is affiliated to Fujian Province and an important tourist city on the southeast coast. It has many famous tourist attractions such as Gulangyu Island, Zengcuoan, Xiamen University, Nanputuo Temple, Hulishan Fort, etc., attracting hundreds of millions of domestic and foreign tourists every year. The results suggest that CBAM shows high forecasting accuracy and outperforms all benchmark models.

This study provides some contributions to the tourism forecasting literature.

- First, a simple yet effective attention module based on the convolutional block (CBAM) is proposed which can be widely applied to the real-time forecast of tourism demand.
- Second, the performance of the proposed model is verified that it is greatly superior to other multiple benchmarks (CNN, hybrid CNN, ConvLSTM, ConvGRU) on high precision and calculation speed.
- Third, the passenger flow map data of Beijing and Xiamen in different periods is used to validate the proposed model. The high prediction accuracy and fast calculation speed illustrates the universality and versatility of the model.

2 Literature review

2.1 Tourism demand forecasting models

Developing an accurate tourism demand forecasting model is a critical issue. Many models have been proposed to pursue better forecasting results (Bi et al. 2021; Gunter and Zekan 2021; Li et al. 2022; Song et al. 2019). As mentioned above, they can be roughly divided into three categories: time series models, econometric models, and artificial intelligence models (Bi et al. 2021; Song and Li 2008).

Time series models have been the traditional and most widely used to forecast tourism demand (Chan et al. 2005; Shahrabi et al. 2013; Yang et al. 2015, 2022). These models may recognize patterns, seasonality, and cycles using data from regularly spaced observations. They are often straightforward and efficient (Yang and Zhang 2019). They mainly include the autoregressive moving average model, exponential smoothing model, and their improved versions (Fildes et al. 2011; Goh and Law 2003). These models forecasted future tourism demand based on historical

trends and patterns, and were widely used as benchmark models in related research (Long et al. 2018).

Unlike time series models, econometric models have been used to capture the causal relationship between the number of tourists and various explanatory factors, such as income, tourism price, and some other variables that potentially influence tourists' travel motivation (Gunter and Zekan 2021). Popular econometric models include the vector autoregressive model (Gunter and Önder 2016), the error correction model (Moore 2010), the autoregressive distributed lag model (Wan and Song 2018), and the time varying parameter model (Song et al. 2011). Many researchers have compared time-series and econometric models and found that the forecasting results can vary contextually (Peng et al. 2014).

These traditional models are subject to the stability of historical patterns and economic structures in practical applications, making it difficult to deal with complex and nonlinear tourism demand data (Zhang et al. 2017a, 2020a, c). For better nonlinear modelling, AI-based models have attracted increasing focus and have shown impressive forecasting accuracy (Li et al. 2022). AI models do not need to make assumptions about the stationarity and distribution of data, and have dynamic adaptability and powerful capabilities for nonlinear time series. Multiple AI models have been developed for tourism demand forecasting (Li et al. 2018). Given the advantages of these models, this study will focus on this type of model.

According to the different "depth" of the artificial intelligence tourism demand forecasting models, they can be divided into two categories: shallow learning models and deep learning models (Bi et al. 2020, 2021). The former adopts an intelligent algorithm with a simple structure and few hidden layers, but usually requires manual construction of data features (Bi et al. 2021). Due to the good nonlinear fitting ability of shallow learning models, they have been widely used in tourism demand forecasting since the 1990s. They include radial basis function (Claveria et al. 2015), multi-layer perceptron, and support vector machines (Li et al. 2020).

Deep learning models have more complex structures and more hidden layers than shallow learning models (Kulshrestha et al. 2020; Schmidhuber 2015). They are end-to-end models that automatically extract features from dataset. These models have been successfully applied in many fields including image recognition, natural language processing, and more recently in tourism demand forecasting (Bi et al. 2020; Law et al. 2019).

There are two main types of deep learning models: convolutional neural networks (CNN) and recurrent neural networks (Schmidhuber 2015). The deep learning models used in tourism demand forecasting are almost recurrent neural networks, of which the most commonly used is the LSTM network (Bi et al. 2021; Law et al. 2019) generated forecasts for tourism arrivals in Macau based on the LSTM framework and found that the LSTM model outperformed various benchmark models, such as support vector regression and artificial neural networks. Similarly, Bi et al. (2020), Zhang et al. (2020a) incorporated many explanatory variables into the LSTM model, thereby strengthening the forecasting performance of their models. In addition to the single LSTM model, some scholars have extended the LSTM model to avoid their inherent drawbacks, such as model overfitting (Zhang et al. 2020b), numerous parameters, and slow convergence. For example, Kulshrestha et al.

(2020) used a Bayesian bidirectional LSTM to forecast tourism demand in Singapore from several major source countries. Laaroussi et al. (2020) adopted a simplified LSTM model, the Gated Recurrent Unit (GRU), to predict tourist arrivals in Morocco. Dan Xiong (2020) proposed a novel LSTM-based model to predict node trajectories in crowd scenarios simultaneously. Zheng et al. (2021) developed a correlated time series oriented Long Short-Term Memory with attention mechanism, to forecast tourism demand for multiple tourist attractions on an hourly. Zhang et al. (2020a, 2020b, 2020c) proposed a novel group-pooling-based deep learning model (GP-DLM) to improve tourism demand forecasting accuracy. They defined a novel dynamic time warping (DTW) clustering quantitative approach and revealed cross-region factors that influence travel demands of the studied regions. Shaolong Sun et al. (2022) proposed an improved machine learning paradigm, introducing valuable additional information into the training phase, which can significantly enhance the multi-step ahead forecasting performance from the view of both error calculation and statistical tests.

On the contrary, few tourism demand forecasting models are based on CNN. This may be because most of the tourism demand data is one-dimensional time series data, and CNNs cannot fully extract data features from one-dimensional data, which means that the advantages of CNNs cannot be fully utilized. However, in many fields, especially in image processing and computer vision, CNNs perform better than recurrent neural networks (Radenovic et al. 2019; Zhang et al. 2017b). For example, Bi et al. (2021) proposed a model based on deep learning with time series imaging. In this research, they transformed the time series data into images and then used the CNN model to extract features. Li et al. (2022) introduced a novel deep learning model based on a graph convolution network and LSTM for tourism demand forecasting. They applied the graph convolution network to characterize the spatial correlations among tourism destinations and LSTM to capture the temporal dependency of each destination itself.

From the above relevant literature summary, it is found that econometric models are still dominant, and artificial intelligence models are becoming more and more popular. Time series models, such as ARMA, are considered benchmark models in forecasting. Deep learning methods are used to improve forecasting accuracy, especially when using massive spatiotemporal data with nonlinear information (Wu et al. 2019).

2.2 Tourism demand forecasting with big data

With the increasing penetration of modern information and communication technologies, there is a large amount of data available to predict customer behavior, and viewed as a new driving factor of tourism demand forecasting research (Yang et al. 2022). Considering the differences in big data types, the research on tourism demand forecasting using big data is divided into three types according to their data sources: search engine data, web traffic data, and social media data.

Search engine data refers to the daily, weekly, and monthly real-time search queries entered by users in search engines, such as Google, Baidu, etc. These data

provide new data sources for tourism forecasting, such as tourist destination arrival prediction (Li et al. 2017, 2018), scenic spots (Li et al. 2019), and hotels (Zhang et al. 2019). This type of data usually uses keyword searches to select search engine data. And the review articles mainly rely on domain knowledge related to tourism demand and search query indexes under specific categories. A previous study by Hyunyoung Choi (2011) used a sub-category of Hong Kong holiday destinations for prediction in Google Trends. Subsequent research refined and expanded the number of keywords to reflect the comprehensive aspects of tourists' activities. Li et al. (2017) used 46 keywords to collect Baidu search data to predict tourism demand in Beijing, China. Law et al. (2019) used 211 Google and 45 Baidu keywords to obtain search engine data to predict Macau's tourism demand. At the same time, considering the different frequencies and period of data obtained, most articles use monthly search query data for tourism demand modelling and forecasting, and only a few articles directly use weekly search engine data for forecasting, including Pan et al. (2012) and Bangwayo-Skeete and Skeete (2015). Therefore, scholars are trying to obtain high-frequency search engine data to improve the efficiency of modeling and forecasting. Mingming Hu et al. (2022) presented mixed data sampling (MIDAS) models to forecast international tourist arrivals to Hong Kong from seven English-speaking countries based on tourists' online review data.

Web traffic data usually represents the raw number of visits to a website, indicates the potential interest of tourists, and can be considered as an explanatory variable for predicting the arrival of tourists. Yang et al. (2013) pointed out that website visits were the step that follows searching for information through search engines, and these data had stronger predictive power than search engine data. When tourists plan their vacation, they often use search engines such as Google and Baidu to retrieve relevant information (Gunter and Önder 2016). Therefore, web traffic data can predict travel and hotel demands. Yang et al. (2013) applied two types of website traffic data, visitors and visits, to predict the accuracy of the hotel room and occupancy rates in Charleston. The results showed that the ARMA model based on web traffic data could significantly reduce short-term prediction error. Gunter and Önder (2016) used 10 web traffic data series from Vienna websites to predict tourist arrivals in Vienna. There are still limitations in the research on tourism prediction using web traffic data, such as the difficulty of data acquisition. However, these data can effectively improve the forecasting accuracy of time series and econometric models.

Social media provides abundant sharing channels for travelers through forums, blogs, microblogs, social networks, image and video sharing websites, and has a significant impact on the tourism system (Li et al. 2021). In general, the social media data includes text and images that appear in questions and answers on forums, user-provided comments, and geotagged photos from various social media sites, such as Twitter, TripAdvisor, Priceline, Hotels.com, Expedia, Flickr and Yelp. Toral et al. (2017) identified unique attributes of tourist destinations through comments collected from online communities and employed a machine learning approach to training a classifier. Their findings suggested that unique attributes were the best predictors of travel destinations. Bigné et al. (2019) used text mining methods to extract important information from tweets on Twitter and study how Twitter activities affect hotel occupancy prediction. Starosta et al. (2018) computed the sentiment indices for

popular destinations in Europe, reflecting the positive or negative attitudes towards these destinations by online media. Miah et al. (2017) adopted a density-based clustering algorithm to identify tourists' photos on Flickr and predict Melbourne's future tourism demand. In contrast to search engine and web traffic data, researchers should use techniques such as text and image mining to extract useful structured data from social media data.

Big data may generally be thought of as a new driver, expressing the attention, interest, and emotions of tourists that affect tourism forecasting (Li et al. 2017; Song et al. 2019). The adoption of big data in tourism forecasting will have managerial implications and enable destination management to better understand tourist behavior, allocate resources, and form timely decisions to improve tourism demand (Ma et al. 2018; Song et al. 2019). However, there are still limitations in data quality due to search keyword selection, noise, or irrelevant information in social media data, etc. (Song et al. 2019). In particular, existing studies on social media and multi-source data are still limited, and researchers need to explore which data sources can be integrated from a rigorous and robust perspective.

Compared to the internet data, the trajectory data recorded by mobile phone take the advantages of low collection cost, high update frequency, wide spatial, and temporal coverage. Such data is a promising source for monitoring urban immigration. The most common use of mobile signaling data is to analyze the distribution of population, spatial movement, and actual distribution of social networks (Sun et al. 2021). Thus, this study aims to introduce a novel spatial-temporal model on the attention mechanism of deep learning approaches, which adopted a more scientific and effective spatial-temporal feature extraction method to forecast the demands of multiple attractions based on the spatial-temporal demand features obtained from mobile phone signaling data.

3 Methodology

In this section, we introduce the framework of deep learning models. First, we introduce the most basic convolutional neural network (CNN), then we introduce the ConvLSTM which combined CNN and LSTM, and finally we present the model of the convolutional block attention module (CBAM).

3.1 CNN

Capturing the hidden complex spatial correlations between adjacent destinations is beneficial to reduce tourism demand forecasting errors. Convolutional neural networks (CNNs) are widely used methods in spatial feature learning, especially for Euclidean-structured data such as images, videos, and regular grids (Li et al. 2022). In this paper, we adopt the CNNs to extract deeper features from the obtained spatiotemporal grid map. In particular, grid passenger flow maps at different times are regarded as different channels of the graph.

The process of feature extraction using convolutional layers is shown in Fig. 1, where the spatial grid is treated as a two-dimensional matrix. The specific steps of image convolution are: sliding the convolution kernel on the image, multiplying the pixel value on the image point with the value on the corresponding convolution kernel, and then adding all the multiplied values. This process is repeated until the convolutional kernel slides over the entire image. Let v_{ij}^{xy} denote the values v at the position (x, y) in the j th feature map of the i th layer can be calculated by:

$$v_{ij}^{xy} = b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} w_{ijm}^{pq} v_{(i-1)m}^{(x+p)(y+q)} \tag{1}$$

where m denotes the index of the feature map in the $(i - 1)$ th layer, w_{ijm}^{pq} is the (p, q) th value of the kernel connected to the m th feature map in the layer, (P, Q) denotes the kernel dimension.

3.2 ConvLSTM

ConvLSTM combines CNN and LSTM and is widely used in various spatiotemporal prediction tasks, such as traffic accident prediction, crowd flow prediction, precipitation prediction, etc. (Miao et al. 2021). The input and hidden state of the ConvLSTM in a timestamp are all 3D tensors, and the convolution operations are performed on the input-to-state and state-to-state connections, and then, memory cells have been introduced in the hidden layer to control the transmission of information, as shown in Fig. 2. More specifically, ConvLSTM first performs a convolution operation on the data in each timestamp, and then passes them along the period $[t - k + 1, \dots, t]$ through the LSTM module, which can be formulated as:

$$i_t = \sigma(W_{X_i} * X^t + W_{H_i} * H_{t-1} + W_{C_i} \cdot C_{t-1} + b_i) \tag{2}$$

$$f_t = \sigma(W_{X_f} * X^t + W_{H_f} * H_{t-1} + W_{C_f} \cdot C_{t-1} + b_f) \tag{3}$$

Fig. 1 The process of feature extraction using convolution layers

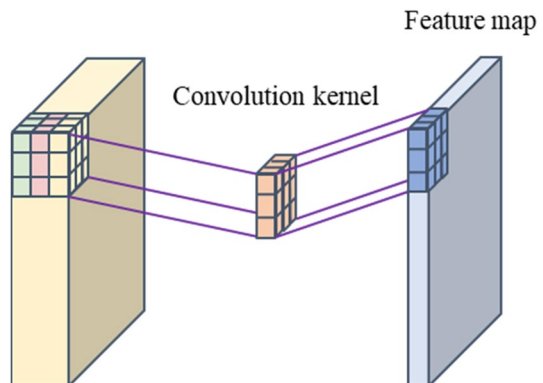


Fig. 2 Architecture of the ConvLSTM cell

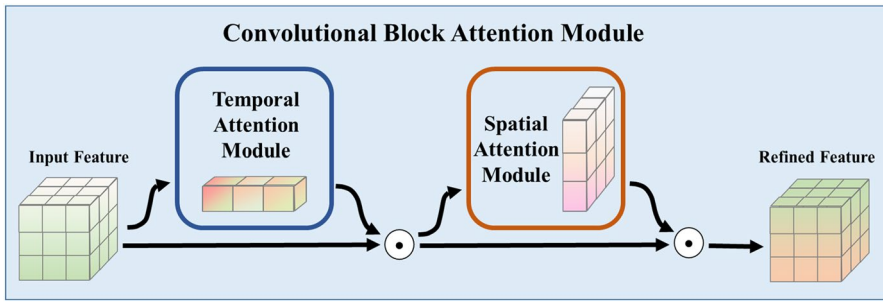
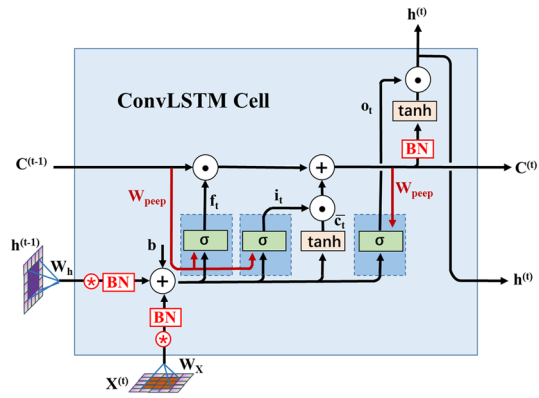


Fig. 3 The architecture of CBAM

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W_{X_c} * X^t + W_{h_c} * H_{t-1} + b_c) \tag{4}$$

$$o_t = \sigma(W_{X_o} * X^t + W_{h_o} * H_{t-1} + W_{c_o} \cdot C_t + b_o) \tag{5}$$

$$H_t = o_t \cdot \tanh(C_t) \tag{6}$$

where ‘*’ denotes the convolution operator, ‘·’ denotes the hadamard product, σ is the logistic sigmoid function, i_t, f_t, C_t, o_t , and H_t are input gate, forget gate, memory cell, output gate and hidden state, $W_{\alpha\beta} (\alpha \in \{X, h, c\}, \beta \in \{i, f, o, c\})$ and b_β are the weight matrixes and bias vectors for the memory cell state in ConvLSTM.

3.3 CBAM

CBAM has two sequential sub-modules: temporal and spatial attention modules (Sanghyun Woo et al. 2018). Given a spatial-temporal feature map $F = R^{T \times H \times W}$ as input, CBAM sequentially infers a 1D temporal attention map $M_T = R^{C \times 1 \times 1}$ and a 2D spatial attention map $M_S = R^{1 \times H \times W}$ as illustrated in Fig. 3. The overall attention process can be formulated as:

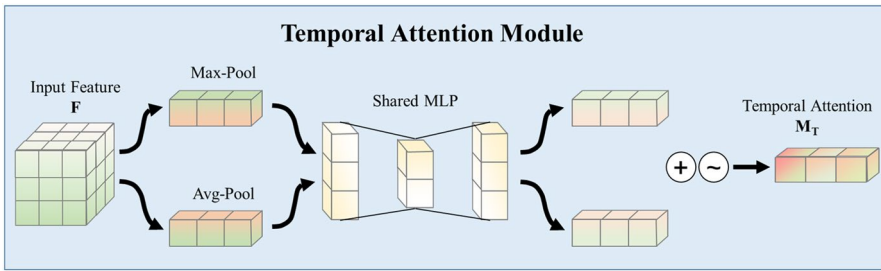


Fig. 4 Diagram of the temporal attention module

$$F' = M_T(F) \cdot F \tag{7}$$

$$F'' = M_S(F') \cdot F' \tag{8}$$

where ‘ \cdot ’ denotes hadamard product. F'' is the final refined output.

3.3.1 Temporal attention module

We generate a temporal attention map by utilizing the inter-temporal relationship of features. As each timestamp of a feature map is considered as a feature detector, the temporal attention module focuses on “when” is meaningful given an input image. To compute the temporal attention efficiently, we squeeze the spatial dimension of the input feature map. For aggregating spatial information, we use average-pooling and max-pooling features simultaneously to generate two different spatial context descriptors: F_{avg}^T and F_{max}^T . Then both descriptors are forwarded to a shared network to produce our temporal attention map $M_T \in R^{T \times 1 \times 1}$. The shared network consists of multiple-layers perceptron (MLP), with a hidden layer. The computation process of this module as shown in Fig. 4. In short, the temporal attention is computed as:

$$\begin{aligned}
 M_T(F) &= \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\
 &= \sigma\left(W_1\left(W_0\left(F_{avg}^T\right)\right) + W_1\left(W_0\left(F_{max}^T\right)\right)\right)
 \end{aligned} \tag{9}$$

where σ is the sigmoid function, W_0 and W_1 denote the weights of input layer and ReLU activation layer in MLP respectively.

3.3.2 Spatial attention module

We produce a spatial attention map by exploiting the inter-spatial relationship of features. Different from the temporal attention module, the spatial attention module focuses on “where” is an informative part, which is complementary to the temporal attention module. To compute the spatial attention, we also adopt average-pooling and max-pooling along the temporal axis to generate two 2D maps: F_{avg}^S and F_{max}^S . After then, they are concatenated to generate an efficient feature descriptor. And a

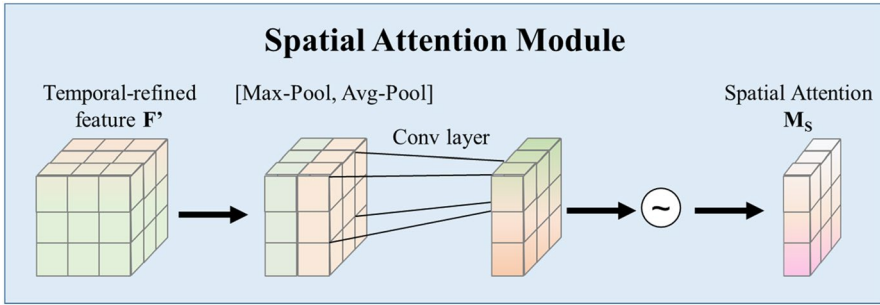


Fig. 5 Diagram of the spatial attention module

convolution layer is applied to produce a spatial attention map $M_S(F) \in R^{H \times W}$. The computation process of this module is shown in Fig. 5. In short, the spatial attention is computed as:

$$M_S(F) = \sigma\left(f^{3 \times 3}\left([AvgPool(F); MaxPool(F)]\right)\right) = \sigma\left(f^{3 \times 3}\left(\left[F_{avg}^S; F_{max}^S\right]\right)\right) \quad (10)$$

where σ denotes the sigmoid function and $f^{3 \times 3}$ represents a convolution operation with the filter size of 3×3 .

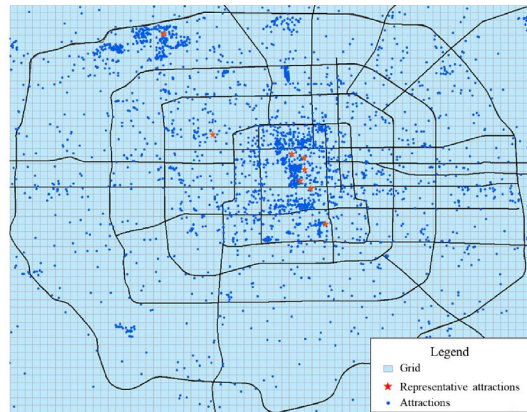
In short, given an input image, two attention modules, temporal and spatial, compute complementary attention, focusing on “when” and “where” respectively.

4 Empirical study

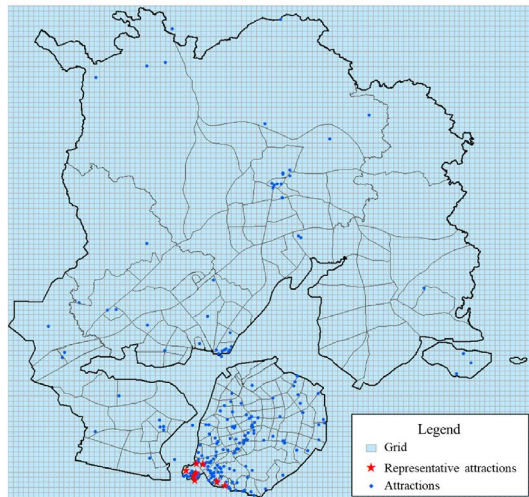
4.1 Data collection

To verify the effectiveness of the proposed model, an empirical analysis is conducted using spatial-temporal grid of tourism demand extracted from mobile signaling data at two well-known tourism cities in China, Beijing and Xiamen. The mobile signaling data were obtained from China Mobile and China Unicom operators, which contain the spatiotemporal location information of all users, such as residents and tourists, as well as other groups. Additional data were gathered from online, such as the city geographic information map and location information of city attractions. First, the city geographic information map is divided into a $500 \text{ m} * 500 \text{ m}$ grid map and the corresponding attractions grids are found by matching the city attractions. Second, taking advantage of the fact that mobile phone signaling data records the temporal and spatial locus points of each user, the flow of passengers in each time slice of each city grid is calculated. The passenger flow includes non-local tourists and local tourists (residents visit scenic spots). In this paper, we uniformly adopted 30 min intervals for a time slice. Third, the flow of passengers in all time slices of the attractions grids are extracted and the temporal regularity are analyzed. Lastly, we stacked several continuous time slice grid flow maps into multi-channel spatial-temporal grid map, and each time slice grid map is regarded as a channel graph.

Fig. 6 City grids and attractions distribution map



(a) Beijing grids map



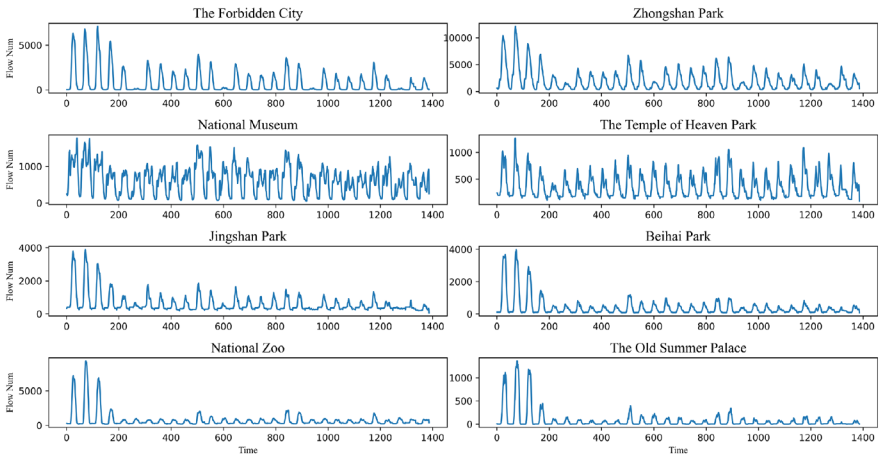
(b) Xiamen grids map

The city map of Beijing was divided into 3726 (54 rows * 69 columns) grids map. And the period is from May 1, 2019 to May 29, 2019, with 1386 time slices. The city map of Xiamen was divided into 9991 (97 rows * 103 columns) grids map and the period is from June 1, 2015 to June 30, 2015, with 1440 time slices (Fig. 6). Meanwhile, we extracted the passenger flow data of 8 representative attractions grids in the two cities respectively (Fig. 7). In addition, to fully verify the performance of different models, each dataset was randomly divided into 10 equal subsamples, and then K-fold cross-validation method was adopted to retain each subsample as a test set to validate the models, the other 9 subsamples were used as a training set to train the models. The details about the dataset are shown in Table 1.

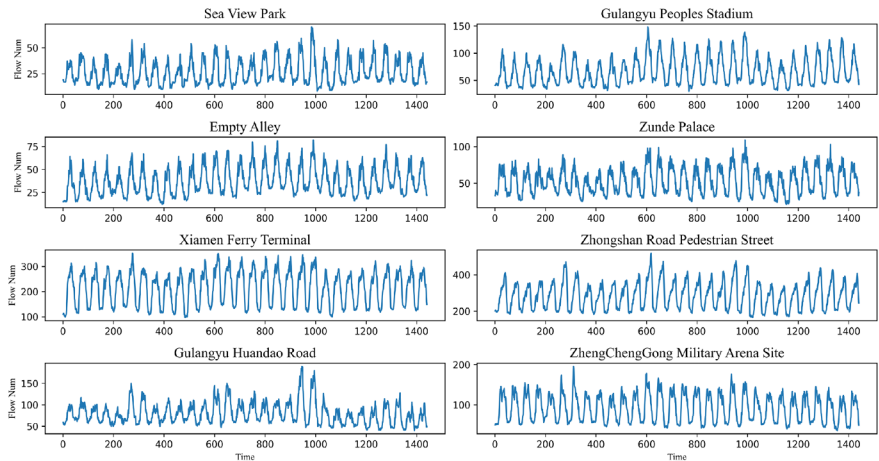
Figure 7 shows the demand for tourist flow of selected attractions in two cities. All attractions tourism demand time series show nonlinear and periodic patterns.

Table 1 Detailed division of each dataset

Dataset	Period	Grids	Time slides	Training set	Test set
Beijing	May 1 to May 29, 2019	3726	1386	1242	139
Xiamen	June 1 to June 30, 2015	9991	1440	1291	144



(a) Selected attractions in Beijing



(b) Selected attractions in Xiamen

Fig. 7 Selected attractions' tourism demand distribution

In particular, the dataset of Beijing includes Labor Day. During holidays, the tourist flow is much higher than that on weekdays.

4.2 Performance evaluations

Five measures of performance are adopted: run time, mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE) and coefficient of determination (R^2). Among them, the run time refers to the time it takes for the model to complete running on the same server, which does not need a specific formula to calculate. The others are calculated as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{12}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \tag{13}$$

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \tag{14}$$

where n is the number of forecasting samples and y_i , \hat{y}_i and \bar{y}_i represent the actual tourism demand observations, forecasted tourism demand and average of the actual tourism demand observations, respectively.

4.3 Benchmark models and parameters setting

Four benchmark models—the ConvLSTM, ConvGRU, hybrid CNN, and basic CNN models—are chosen to properly illustrate the usefulness of the proposed technique. The CNN model is the most commonly used deep learning algorithm to deal with graph convolution. In this study, the spatial grid passenger flow map of several consecutive time slices is regarded as the different channels of the graph stacked into a multi-channel spatiotemporal grid map as the input of the CNN model. Considering the influence of convolution kernel size on model accuracy, we introduce a hybrid CNN model and use different convolution kernel size combinations to test the model accuracy. Compared with CNN and hybrid CNN, ConvLSTM and ConvGRU combine the spatial graph convolution process of CNN and the time series recursion process of LSTM/GRU, in which the time recursion process of LSTM/GRU learns the regularity between time series between the input multi-channel spatiotemporal grid maps.

Table 2 The optimal parameters for each model

Models	Parameters setting
CNN	Epoch=50, layers=4, kernel size=3, learning rate=0.005
Hybrid CNN	Epoch=50, layers=8, kernel_1 size=3, Kernel_2 size=5, learning rate=0.005
ConvLSTM	Epoch=50, layers=3, kernel size=3, learning rate=0.01
ConvGRU	Epoch=50, layers=3, kernel size=3, larning rate=0.01
CBAM	Epoch=50, layers=4, kernel size=3, learning rate=0.01

In addition, the appropriate parameter setting is the key to a better forecasting performance of these models. The main parameters of these models include the following: number of time slices, kernel size, learning rate, training epochs, batch size, optimization function and loss function. In this study, we set the number of the time slice as 5, the batch size as 10, the loss function uses MSE Loss function and the optimization function uses the Adam optimizer in the torch. The other parameters setting results are shown in Table 2.

4.4 Forecasting results and analysis

The proposed CBAM model and various benchmarks are first trained based on the optimal parameters combination and 10 different training sets. Then, forecasting for different test sets is generated separately. Lastly, we calculate the average error between the true value and the forecasting value of the tourist flow of 8 attraction grids in each of the two cities under different test sets respectively. The results are shown in Table 3.

It can be seen that the proposed model significantly outperforms the benchmark models in terms of run time in both two cities' dataset. For Xiamen dataset, the run time of the proposed model is 896 s, and the CNN model, hybrid CNN, ConvLSTM and ConvGRU are 1491 s, 2375 s, 11,877 and 14,391 s respectively; for Beijing dataset, the run time of the proposed model is 330 s, and the CNN model, hybrid CNN, ConvLSTM and ConvGRU are 559 s, 845 s, 5329 and 6069 s respectively. The results show that the proposed model is 1.68 times faster than the CNN model, 14.71 times faster than the ConvLSTM and 17.23 times faster than the ConvGRU. ConvGRU and ConvLSTM run slowly, mainly because the two models need to learn the regular patterns between the input multi-channel spatial-temporal maps, on the contrary, the proposed model uses an attention mechanism to effectively learn important information in spatial and temporal information, reducing the complexity of the model and making the model relatively simple and effective. In addition, for Xiamen dataset, the MAE, RMSE, MAPE and R^2 of the proposed model are 5.31, 7.21, 6.27 and 91.97% respectively, which are better than the best MAE, RMSE, MAPE and R^2 obtained by the benchmark models (5.38, 7.26, 6.43 and 91.20%); for Beijing dataset, the MAE, RMSE, MAPE and R^2 of the proposed model are 38.05, 60.48, 9.96 and 93.37% respectively, which are superior to the best MAE,

Table 3 The results of forecasting from different models for two cities' attractions grids

Models	Xiamen dataset					Beijing dataset				
	Run time	MAE	RMSE	MAPE	R ²	Run time	MAE	RMSE	MAPE	R ²
	CNN	1491	6.57	8.59	8.34%	89.18%	559	90.71	115.75	38.28%
Hybrid_CNN	2375	7.35	9.57	7.84%	86.50%	845	47.10	69.22	24.32%	89.72%
ConvLSTM	11,877	5.38	7.26	6.61%	91.20%	5329	42.87	61.76	24.64%	90.79%
ConvGRU	14,391	5.40	7.40	6.43%	91.05%	6069	41.97	65.04	15.39%	91.44%
CBAM	896	5.31	7.21	6.27%	91.97%	330	38.05	60.48	9.96%	93.37%

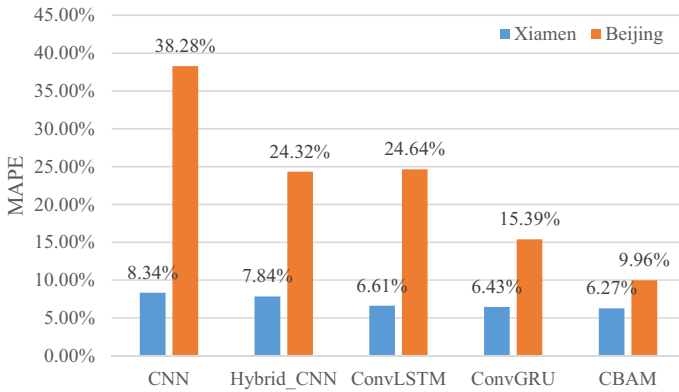


Fig. 8 Comparison of prediction accuracy of different models

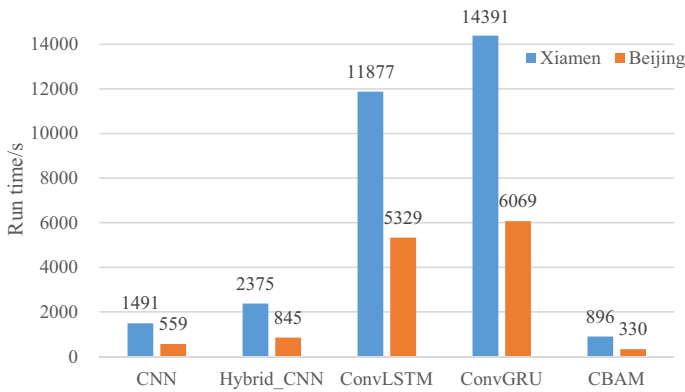


Fig. 9 Comparison running time of different models

RMSE, MAPE and R2 obtained by the benchmark models (41.97, 61.76, 15.39 and 91.44%).

On the whole, the prediction accuracy of the basic CNN model is not high. Hybrid CNN adds a 5 * 5 convolution kernel compared with basic CNN, which slightly improves the complexity of the model and the prediction accuracy of the model. Compared with basic CNN and Hybrid CNN, ConvLSTM and ConvGRU have greatly improved the accuracy of the two models. The two models can learn the relationship between the mining channel graph, and the complexity of the model increases a lot, and so does the running time of the model. The proposed CBAM model has the highest prediction accuracy and the fastest operating efficiency. The model uses the attention mechanism to mine the important information of the temporal channel and the important information of the spatial position of the spatial-temporal passenger flow graph, ignores the interference of the unimportant information, reduces the complexity of the model, and improves operational efficiency

Table 4 Forecasting results of each model in terms of the MAPE value

	Attractions	ARIMA	CNN	Hybrid CNN	ConvLSTM	ConvGRU	CBAM
Xiamen	Sea view park	16.84%	5.69%	5.66%	4.68%	4.96%	4.58%
	Gulangyu Peoples stadium	8.51%	8.46%	8.54%	8.15%	7.33%	7.23%
	Empty alley	19.67%	11.46%	12.31%	10.18%	10.07%	9.80%
	Zunde Palace	9.40%	15.14%	9.07%	7.99%	8.16%	7.93%
	Xiamen ferry terminal	5.30%	5.37%	5.32%	4.00%	4.22%	3.95%
	Zhongshan road pedestrian street	4.76%	5.50%	4.64%	3.16%	3.08%	3.08%
	Gulangyu huandao road	22.29%	7.07%	9.56%	7.40%	6.76%	6.72%
	Zhengchenggong military arena site	12.39%	8.00%	7.59%	7.31%	6.90%	6.85%
Beijing	The forbidden city	56.53%	60.57%	44.79%	43.20%	24.54%	13.69%
	Zhongshan park	24.54%	22.61%	10.00%	9.46%	9.15%	9.08%
	National Museum	29.16%	24.30%	15.41%	13.80%	14.16%	11.58%
	The temple of heaven park	19.98%	19.03%	10.38%	10.80%	9.60%	8.53%
	Jingshan park	33.04%	30.59%	8.24%	7.27%	7.09%	6.44%
	Beihai park	40.53%	31.64%	12.20%	10.53%	9.43%	8.71%
	National zoo	21.27%	19.47%	11.36%	8.47%	7.85%	7.27%
	The old summer palace	49.79%	98.05%	82.17%	93.55%	41.31%	14.36%
AVG	23.37%	23.31%	16.08%	15.62%	10.92%	8.11%	

and accuracy. The comparative analysis of prediction accuracy and running time of different models can be seen in Figs. 8 and 9.

To better describe the forecasting accuracy clearly of tourist flow under different models in each attraction grid, Table 4 shows the MAPE of each attraction grid under different models. To compare the traditional time series prediction algorithm at the same time, we also used the ARIMA algorithm to predict the passenger flow of 16 scenic spots in two cities. It can be found from the results that, on the whole, the MAPE of the CBAM model is lower than that of the optimal benchmark model in all attractions. The average MAPE of the proposed model is 8.11%, which is much lower than that of ARIMA, CNN, Hybrid CNN, ConvLSTM and ConvGRU (23.37, 23.31, 16.08, 15.62 and 10.92%, respectively). Especially for the Beijing data set, the CBAM model is significantly better than other benchmark models. The average MAPE of the proposed model for the 8 representative scenic spots in Beijing is 9.96%. The average MAPE of ARIMA, CNN, Hybrid CNN, ConvLSTM and ConvGRU was 34.36, 38.28, 24.32, 24.64 and 15.39%, respectively. For all scenic spots in Xiamen, the MAPE of the CBAM model is lower than that of other benchmark models. The prediction error of the proposed models is within 10%, and the average MAPE is 6.27%. The prediction accuracy of multi-scenic spots tourism demand is higher.

Meanwhile, to better reflect the passenger flow prediction results of the proposed model in various scenic spots of Beijing and Xiamen, the predicted value and the

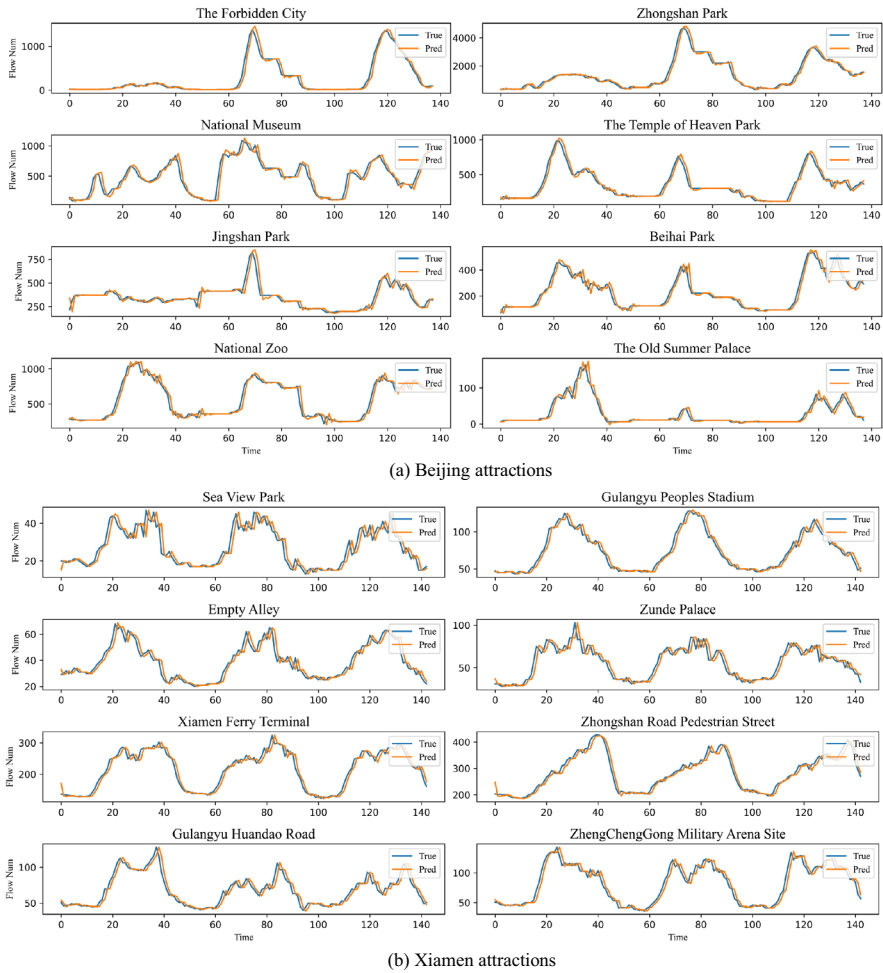


Fig. 10 Comparison of predicted and actual flows of different attractions under the CBAM model

true value of the proposed model are placed in the same figure for comparative analysis, as shown in Fig. 10. It can be seen from the figure that the predicted value of passenger flow in scenic spots of CBAM model is in good agreement with the true value. The average fitting coefficient of this model is 93.37% for passenger flow in scenic spots of Beijing and 91.97% for passenger flow in scenic spots of Xiamen.

It is particularly noteworthy that for the grid of the Old Summer Palace scenic spot, the MAPE of other benchmark models are greater than 40%, while the MAPE of the proposed model is only 14.36%, which is 2.88 times higher in accuracy. In this paper, the prediction results of different models in the scenic spot grid of the Old Summer Palace were extracted and compared with the true values, as shown in Fig. 11. It can be seen from the figure that the prediction accuracy of ARIMA and basic CNN model are the worst, and the prediction trend has a large deviation

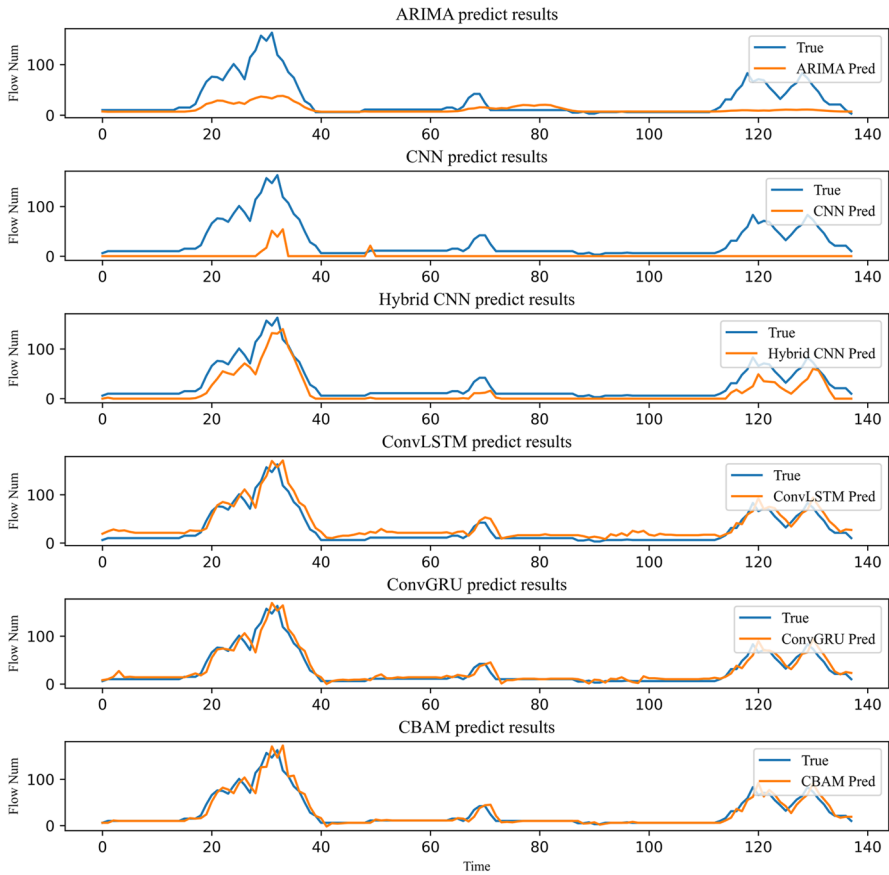


Fig. 11 Comparative analysis of grid passenger flow prediction of different models in the Old Summer Palace

from the true value, and the peak value also has a large deviation. Hybrid CNN has a significant improvement compared with CNN, and the prediction trend of ConvGRU and ConvLSTM models is consistent with the true value trend. However, in some low-peak periods of passenger flow, there is a slight deviation between the predicted value and the true value, but the calculated MAPE will be larger than that of other scenic spot grids, and the proposed model can reduce this error, because the proposed model uses maximum pooling and average pooling to mine the temporal importance of the grid map, and reduces the influence and interference of unimportant factors, which results in the high prediction accuracy of the model in the low peak period. Therefore, this fully demonstrates the universality of the proposed model.

To verify the universality of the proposed model, we also extracted three time slices (morning peak, evening peak and flat peak) of the predicted passenger flow values and the spatial distribution of errors of each grid in two cities, as shown in

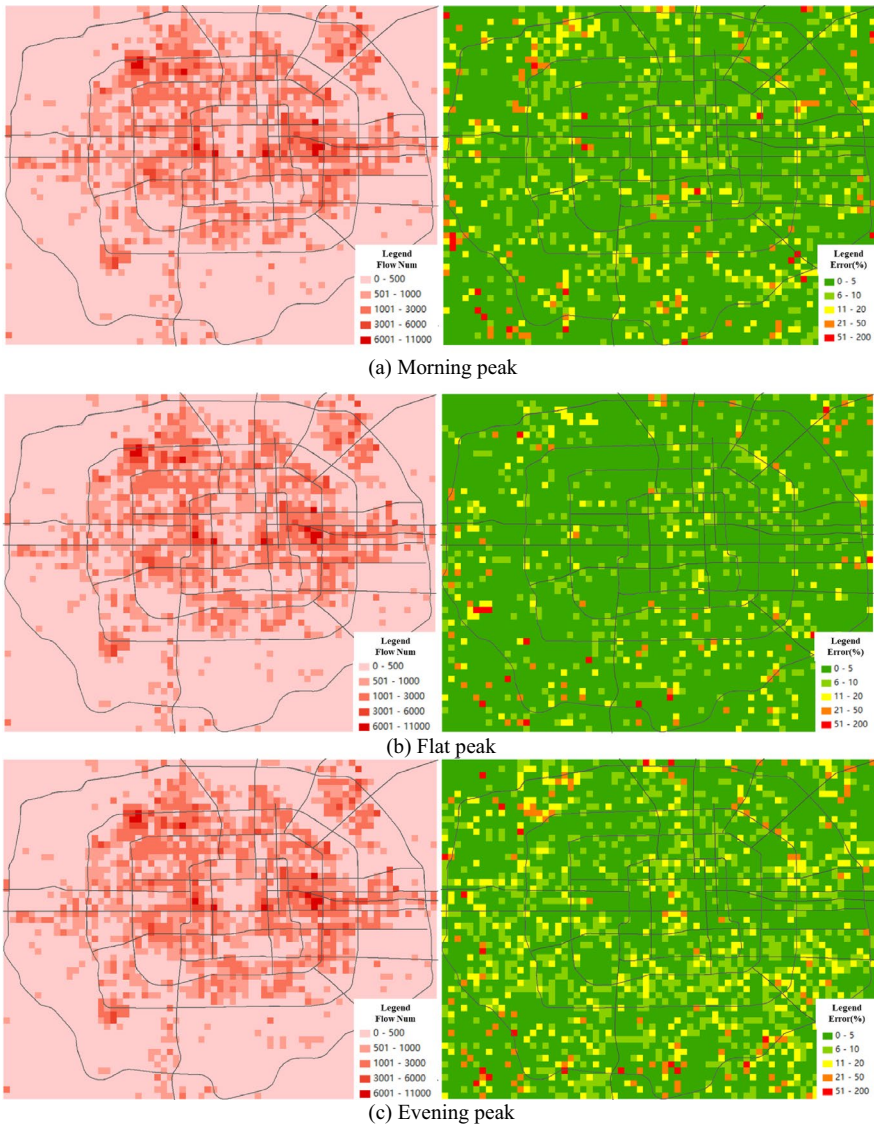


Fig. 12 Prediction results and error distribution of the grids in the Fifth Ring Road of Beijing based on CBAM model

Figs. 12 and 13. As can be seen from Fig. 12, the grids with high passenger flow density in the Fifth Ring Road of Beijing are mainly distributed along Chang ‘an Avenue, the west end of Chang ‘an Avenue is the Xidan and Financial Street area, the east end of Chang ‘an Avenue is the Dongdan and CBD area, and the grids around Peking University in the northwest Fourth Ring Road. The prediction error is basically within 10%. The grids with large prediction error are mainly distributed

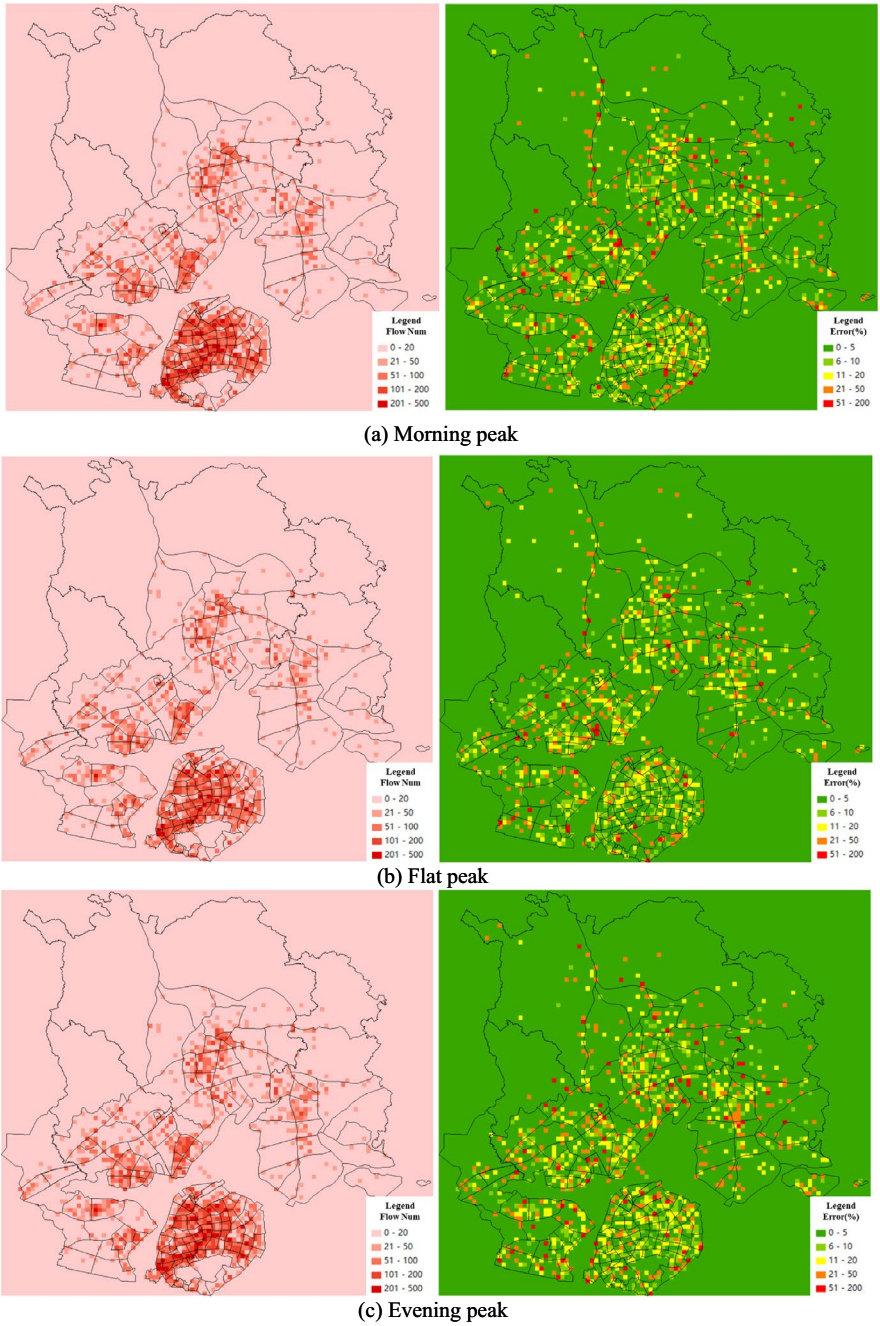


Fig. 13 Prediction results and error distribution of the grids in the Xiamen based on the CBAM model

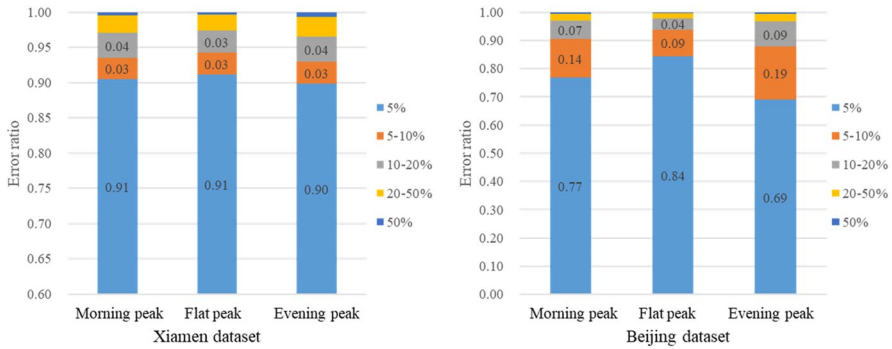


Fig. 14 CBAM model predicted accuracy distribution in different periods

around the Fifth Ring Road and the peripheral area of the research object. The analysis may be due to the low learning rate of the spatial relationship of the peripheral grids, which leads to the large prediction error of the model. From Fig. 13, the areas with high passenger flow density in Xiamen are mainly distributed in Xiamen Island, and the prediction error of the model is basically within 10%. The grids with large prediction errors are mainly distributed in the boundaries of suburbs or traffic districts. The results above demonstrate the universality of the proposed model for passenger flow prediction on all grids.

To further analyze the universality of all grids in different periods for the prediction accuracy of the model, the prediction error distributions of the grids in Xiamen City (9,991 grids) and the grids in Beijing Fifth Ring Road (3726 grids) in different periods were collected in this paper. The specific statistical distribution results of the errors are shown in Fig. 14. As can be seen from Fig. 14, the prediction error of Xiamen grids are within 5%, accounting for 90%, and the prediction errors are within 10%, accounting for 94%. The prediction error of the grids within the Beijing Fifth Ring Road are less than 5%, accounting for 77%, and the prediction errors are less than 10%, accounting for about 91%. Through the error analysis of the two research objects, it can be seen that the prediction accuracy of the proposed model is high and universal. At the same time, the more grids of the research objects in a certain range, the more spatial relations the model can learn, which will improve the prediction accuracy of the model. In addition, by comparing the proportion of error distribution in the morning and evening peak and flat peak of the two cities, it can be found that the prediction accuracy of flat peak is slightly higher than the morning and evening peak hours. The analysis may be caused by the high aggregation of passenger flow in the morning and evening peak hours, the strong mobility of population travel, and the prediction error of the grid with obvious traffic attribute characteristics (such as urban roads, hub terminals, etc.) will be larger in the morning and evening peak hours.

5 Conclusion and discussion

In this study, we forecasted short term tourism demand across two cities in different periods. A deep learning model was employed to estimate the forecasting model based on a convolution block combined with attention module. In particular, we assessed the effectiveness of the proposed model in tourism forecasting by comparing forecasting errors and calculated speed among several benchmark models such as CNN, ConvLSTM and ConvGRU. The results indicated that, the calculated speed of the proposed model is 14.71 times faster than the ConvLSTM and 17.23 times faster than the ConvGRU. Besides, the MAPE of the proposed model is superior to CNN, ConvLSTM and ConvGRU. Especially, the proposed model is more stable in forecasting performance, for example, the MAPE of the old summer palace is 14.36%, on the contrary, the MAPE of other benchmark models are greater than 40%. Therefore, in summary, the proposed model has high accuracy and fast calculation speed in predicting the passenger flow of multiple scenic spots simultaneously, which is very suitable for the real-time prediction of the passenger flow of urban scenic spots and the collaborative linkage organization among multiple scenic spots.

Tourism demand forecasting has continuously attracted researchers' attention because it can provide a valuable aid for efficient business operations and effective destination management. In this study, we proposed a novel deep learning model based on a convolutional block and attention module to make precise tourism demand forecasting. Two markedly different forecasting experiments were investigated to verify the forecasting performance of our proposed model. Empirical results demonstrate the proposed model can handle tourism demand data with spatial-temporal grids, and provide accurate tourism demand forecasts for multi-attractions. To the best of our knowledge, this is the first time that the convolution block combined with the attention module is applied to spatial-temporal tourist flow forecasting in tourism demand. The main findings and contributions of the study are summarized as follows:

- (1) A novel model are proposed based on deep learning approaches to accurately forecasting tourism demands. The CBAM model investigated forecasts based on temporal dependency and spatial effects from spatial-temporal passenger flow maps, which were rarely introduced in previous studies.
- (2) The proposed model is simple yet effective based on a convolutional block combined with an attention module that can be widely applied to the real-time forecast of tourism demand. The experiment results indicates that the performance of our proposed model is greatly superior to other several benchmarks (CNN, Hybrid CNN, ConvLSTM, ConvGRU) on high precision and calculation speed.
- (3) We used the passenger flow map data of Beijing and Xiamen in different periods to validate the proposed model, and obtain high prediction accuracy, which illustrates the universality and versatility of the model.

However, the superior predictive performance of our model can't cover the limitations of this study. First of all, the proposed model is proved in the forecast case of tourism demand under normal circumstances. Unfortunately, data from emergencies such as the outbreak of COVID-2019 cannot be obtained at present, so it is impossible to verify the accuracy of the proposed model's tourism demand prediction under emergencies. Secondly, this model is a forecasting model of a spatial-temporal passenger flow map, and without considering other exogenous variables. However, introducing exogenous explanatory variables that may affect tourism demand is considered an effective strategy to enhance the performance of forecasting models. It is worth noting how much improvement can be achieved in the proposed model by integrating factors such as GDP, points of interest geolocation data and weather. Third, parameterize the number of selected demand sequences in the demand stacking graph and the number of lags of the predicted results, to extract more accurate spatial correlation and obtain more passenger flow maps in future periods. The proposed model can only predict the passenger flow data of one period in the future, and cannot simultaneously evaluate the prediction accuracy of multiple periods of passenger flow. In future research, spatial maps based on economic factors, cultural factors, geotagged data, and time-varying parameters should be constructed and applied.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40558-023-00247-y>.

Data availability The datasets generated and analyzed during the current study are not publicly available due to the fact that they constitute an excerpt of research in progress but are available from the corresponding author on reasonable request.

References

- Alén E, Losada N, de Carlos P (2017) Profiling the segments of senior tourists throughout motivation and travel characteristics. *Curr Issues Tour* 20(14):1454–1469
- Bangwayo-Skeete PF, Skeete RW (2015) Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tour Manag* 46:454–464
- Bi J-W, Liu Y, Fan Z-P, Zhang J (2020) Exploring asymmetric effects of attribute performance on customer satisfaction in the hotel industry. *Tour Manag* 77:104006
- Bi J-W, Li H, Fan Z-P (2021) Tourism demand forecasting with time series imaging: a deep learning model. *Ann Tour Res* 90:103255
- Bigné E, Oltra E, Andreu L (2019) Harnessing stakeholder input on Twitter: a case study of short breaks in spanish tourist cities. *Tour Manag* 71:490–503
- Chan F, Lim C, McAleer M (2005) Modelling multivariate international tourism demand and volatility. *Tour Manag* 26(3):459–471
- Chhetri A, Arrowsmith C, Chhetri P, Corcoran J (2013) Mapping spatial tourism and hospitality employment clusters: an application of spatial autocorrelation. *Tour Anal* 18(5):559–573
- Claveria O, Monte E, Torra S (2015) Tourism demand forecasting with neural network models: different ways of treating information. *Int J Tour Res* 17(5):492–500
- Dan Xiong FL, Nie Y (2020) Spatial-temporal block and LSTM network for pedestrian trajectories prediction. arXiv preprint [arXiv:2009.10468](https://arxiv.org/abs/2009.10468)
- Day J, Chin N, Sydnor S, Cherkauer K (2013) Weather, climate, and tourism performance: a quantitative analysis. *Tour Manag Perspect* 5:51–56
- Fan B, Wu Y, He ZB, Chen YY, Quek TQS, Xu CZ (2021) Digital twin empowered mobile edge computing for intelligent vehicular lane-changing. *IEEE Netw* 35(6):194–201

- Fan B, Su Z, Chen Y, Wu Y, Xu C, Quek TQS (2022) Ubiquitous control over heterogeneous vehicles: a digital twin empowered edge AI approach. *IEEE Wirel Commun* 30(1):166–173
- Fildes R, Wei Y, Ismail S (2011) Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures. *Int J Forecast* 27(3):902–922
- Goh C, Law R (2003) Incorporating the rough sets theory into travel demand analysis. *Tour Manag* 24(5):511–517
- Gunter U, Önder I (2016) Forecasting city arrivals with Google analytics. *Ann Tour Res* 61:199–212
- Gunter U, Zekan B (2021) Forecasting air passenger numbers with a GVAR model. *Ann Tour Res* 89:103252
- Hernández-López M, Cáceres-Hernández JJ (2007) Forecasting tourists' characteristics by a genetic algorithm with a transition matrix. *Tour Manag* 28(1):290–297
- Hu M, Song H (2019) Data source combination for tourism demand forecasting. *Tour Econ* 26(7):1248–1265
- Hu M, Qiu RTR, Wu DC, Song H (2021) Hierarchical pattern recognition for tourism demand forecasting. *Tour Manag* 84:104263
- Huang X, Zhang L, Ding Y (2017) The Baidu Index: uses in predicting tourism flows—a case study of the forbidden city. *Tour Manag* 58:301–306
- Hyunyoung C, Varian H (2012) Predicting the present with Google trends. *Econ Rec* 88:2–9
- Jiao X, Li G, Chen JL (2020) Forecasting international tourism demand: a local spatiotemporal model. *Ann Tour Res* 83:102937
- Jiao X, Chen JL, Li G (2021) Forecasting tourism demand: developing a general nesting spatiotemporal model. *Ann Tour Res* 90:103277
- Kaynak E, Marandu EE (2006) Tourism market potential analysis in Botswana: a Delphi study. *J Travel Res* 45(2):227–237
- Kulshrestha A, Krishnaswamy V, Sharma M (2020) Bayesian BILSTM approach for tourism demand forecasting. *Ann Tour Res* 83:102925
- Laaroussi H, Guerouate F, Sbihi M (2020) Deep learning framework for forecasting tourism demand. In: 2020 IEEE international conference on technology management, operations and decisions (ICT-MOD), pp 1–4
- Law R, Li G, Fong DKC, Han X (2019) Tourism demand forecasting: a deep learning approach. *Ann Tour Res* 75:410–423
- Li X, Pan B, Law R, Huang X (2017) Forecasting tourism demand with composite search index. *Tour Manag* 59:57–66
- Li S, Chen T, Wang L, Ming C (2018) Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu index. *Tour Manag* 68:116–126
- Li K, Lu W, Liang C, Wang B (2019) Intelligence in tourism management: a hybrid FOA-BP method on daily tourism demand forecasting with web search data. *Mathematics* 7(6):531
- Li C, Ge P, Liu Z, Zheng W (2020) Forecasting tourist arrivals using denoising and potential factors. *Ann Tour Res* 83:102943
- Li X, Law R, Xie G, Wang S (2021) Review of tourism forecasting research with internet data. *Tour Manag* 83:104245
- Li C, Zheng W, Ge P (2022) Tourism demand forecasting with spatiotemporal features. *Ann Tour Res* 94:103384
- Lin VS, Song H (2015) A review of Delphi forecasting research in tourism. *Curr Issues Tour* 18(12):1099–1131
- Long W, Liu C, Song H (2018) Pooling in tourism demand forecasting. *J Travel Res* 58(7):1161–1174
- Ma Y, Xiang Z, Du Q, Fan W (2018) Effects of user-provided photos on hotel review helpfulness: an analytical approach with deep learning. *Int J Hosp Manag* 71:120–131
- Miah SJ, Vu HQ, Gammack J, McGrath M (2017) A big data analytics method for tourist behaviour analysis. *Inf Manag* 54(6):771–785
- Miao H, Fei Y, Wang S, Wang F, Wen D (2021) Deep learning based origin-destination prediction via contextual information fusion. *Multimed Tools Appl* 81(9):12029–12045
- Hu M, Li H, Song H, Li X, Law R (2022) Tourism demand forecasting using tourist-generated online review data. *Tour Manag* 90:104490
- Moore WR (2010) The impact of climate change on caribbean tourism demand. *Curr Issues Tour* 13(5):495–505

- Ng D (1984) A model estimating the demand for leisure services manpower: forecasting with multiple regression, the subjective/objective qualitative forecasting, and the Delphi methodology. *World Leisure Recreat* 26(5):45–49
- Palmer A, José Montaña J, Sesé A (2006) Designing an artificial neural network for forecasting tourism time series. *Tour Manag* 27(5):781–790
- Pan B, Wu C, Song D (2012) Forecasting hotel room demand using search engine data. *J Hosp Tour Technol* 3(3):196–210
- Park S, Lee J, Song W (2017) Short-term forecasting of Japanese tourist inflow to South Korea using Google trends data. *J Travel Tour Mark* 34(3):357–368
- Peng B, Song H, Crouch GI (2014) A meta-analysis of international tourism demand forecasting and implications for practice. *Tour Manag* 45:181–193
- Pomfret G, Bramwell B (2016) The characteristics and motivational decisions of outdoor adventure tourists: a review and analysis. *Curr Issues Tour* 19(14):1447–1478
- Radenovic F, Toliás G, Chum O (2019) Fine-tuning CNN image retrieval with no human annotation. *IEEE Trans Pattern Anal Mach Intell* 41(7):1655–1668
- Sanghyun Woo JP, Lee J-Y, Kweon IS (2018) CBAM: Convolutional block attention module. In: *Proceedings of the European conference on computer vision (ECCV)*, pp 3–19
- Schmidhuber J (2015) Deep learning in neural networks: an overview. *Neural Netw* 61:85–117
- Shahrabi J, Hadavandi E, Asadi S (2013) Developing a hybrid intelligent model for forecasting problems: case study of tourism demand time series. *Knowl Based Syst* 43:112–122
- Sun S, Li M, Wang S, Zhang C (2022) Multi-step ahead tourism demand forecasting: the perspective of the learning using privileged information paradigm. *Expert Syst Appl* 210:118502
- Song H, Li G (2008) Tourism demand modelling and forecasting—a review of recent research. *Tour Manag* 29(2):203–220
- Song H, Witt SF, Jensen TC (2003) Tourism forecasting: accuracy of alternative econometric models. *Int J Forecast* 19(1):123–141
- Song H, Li G, Witt SF, Athanasopoulos G (2011) Forecasting tourist arrivals using time-varying parameter structural time series models. *Int J Forecast* 27(3):855–869
- Song H, Qiu RTR, Park J (2019) A review of research on tourism demand forecasting: launching the Annals of Tourism research curated collection on tourism demand forecasting. *Ann Tour Res* 75:338–362
- Starosta K, Budz S, Krutwig M (2018) The impact of German-speaking online media on tourist arrivals in popular tourist destinations for Europeans. *Appl Econ* 51(14):1558–1573
- Sun S, Wei Y, Tsui K-L, Wang S (2019) Forecasting tourist arrivals with machine learning and internet search index. *Tour Manag* 70:1–10
- Sun H, Chen Y, Lai J, Wang Y, Liu X (2021) Identifying tourists and locals by K-means clustering method from mobile phone signaling data. *J Transp Eng Part A Syst* 147(10):04021070
- Toral SL, Martínez-Torres MR, Gonzalez-Rodriguez MR (2017) Identification of the unique attributes of tourist destinations from online reviews. *J Travel Res* 57(7):908–919
- Volchek K, Liu A, Song H, Buhalis D (2019) Forecasting tourist arrivals at attractions: search engine empowered methodologies. *Tour Econ* 25(3):425–447
- Wan SK, Song H (2018) Forecasting turning points in tourism growth. *Ann Tour Res* 72:156–167
- Wu J, Cheng X, Liao SS (2019) Tourism forecast combination using the stochastic frontier analysis technique. *Tour Econ* 26(7):1086–1107
- Yang Y, Zhang H (2019) Spatial-temporal forecasting of tourism demand. *Ann Tour Res* 75:106–119
- Yang Y, Pan B, Song H (2013) Predicting hotel demand using Destination Marketing Organization's web traffic data. *J Travel Res* 53(4):433–447
- Yang Y, Pan B, Song H (2014) Predicting hotel demand using destination marketing organization's web traffic data. *J Travel Res* 53(4):433–447
- Yang X, Pan B, Evans JA, Lv B (2015) Forecasting chinese tourist volume with search engine data. *Tour Manag* 46:386–397
- Yang Y, Fan Y, Jiang L, Liu X (2022) Search query and tourism forecasting during the pandemic: when and where can digital footprints be helpful as predictors? *Ann Tour Res* 93:103365
- Zhang B, Huang X, Li N, Law R (2017) A novel hybrid model for tourist volume forecasting incorporating search engine data. *Asia Pac J Tour Res* 22(3):245–254
- Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Trans Image Process* 26(7):3142–3155

- Zhang B, Pu Y, Wang Y, Li J (2019) Forecasting hotel accommodation demand based on LSTM model incorporating internet search index. *Sustainability* 11(17):4708
- Zhang B, Li N, Shi F, Law R (2020a) A deep learning approach for daily tourist flow forecasting with consumer search data. *Asia Pac J Tour Res* 25(3):323–339
- Zhang Y, Li G, Muskat B, Law R (2020b) Tourism demand forecasting: a decomposed deep learning approach. *J Travel Res* 60(5):981–997
- Zhang Y, Muskat GLB, Law R, Yang Y (2020) Group pooling for deep tourism demand forecasting. *Ann Tour Res* 82:102899
- Zheng W, Huang L, Lin Z (2021) Multi-attraction, hourly tourism demand forecasting. *Ann Tour Res* 90:103271

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Haodong Sun¹ · Yang Yang¹ · Yanyan Chen¹ · Xiaoming Liu¹ · Jiachen Wang²

Haodong Sun
sunhaodong@emails.bjut.edu.cn

Yang Yang
yyangscholar@bjut.edu.cn

Xiaoming Liu
liuxm@mot.gov.cn

Jiachen Wang
w1462435605@qq.com

¹ Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, No. 100 Pingleyuan, Chaoyang District, Beijing 100124, China

² Faculty of Information Technology, Beijing University of Technology, No. 100 Pingleyuan, Chaoyang District, Beijing 100124, China