Parking Prediction in Smart Cities: A Survey

Xiao Xiao, *Member, IEEE*, Ziyan Peng, Yunqing Lin, Zhiling Jin, Wei Shao, Rui Chen, *Member, IEEE*, Nan Cheng, *Member, IEEE*, and Guoqiang Mao, *Fellow, IEEE*

***Abstract*— With the growing number of cars in cities, smart parking is gradually becoming a strategic issue in building a smart city. As the precondition in smart parking, accurate parking prediction can reduce the time drivers spend searching for parking spaces and relieve traffic congestion. Meanwhile, VANET and the Internet-of-things (IoT) are the key elements of the current intelligent transportation system. With the IoT devices based on VANET becoming more extensively employed, a large amount of parking data is generated every day, and various methods are proposed for parking prediction, therefore, it is time to systematically summarize the parking prediction issues and the state-of-the-art prediction methods. In this survey, we first provide a comprehensive review of the existing methods used for parking prediction ranging from conventional statistical methods to the latest graph neural network methods. Then, we classify a variety of parking problems such as parking availability prediction, parking behavior prediction, and parking demand prediction. We also compile all the evaluation metrics, open data, and open-source code of the surveyed literature. Finally, we present the challenges and future directions of the parking prediction technique. As far as we know, this is the first survey exploring parking prediction methods, which will be of interest to both researchers and practitioners engaging in intelligent transportation systems (ITS) and smart cities.**

***Index Terms*— Machine learning, deep learning, parking prediction, smart parking, smart city.**

I. Introduction

PARKING problems in big cities have become more and more serious in recent years. With the fast increase in the number of vehicles on the road, drivers are finding it increasingly difficult to find sufficient parking spaces, resulting in frustration, pollution, and traffic congestion [1]. It is reported that it usually costs people around 20 minutes to find an available parking space [2]. Fortunately, with the advancement of intelligent transportation systems (ITS) in smart cities, parking status can be detected by on-ground sensors installed near parking spaces. These records can be stored and uploaded to the web using the Internet of things (IoT) techniques. Parking predictions based on these stored historical data combined with the existing parking guidance and information systems (PGIS) [3] can be an effective way to guide drivers to free parking spaces in a shorter time, therefore, as the precondition of realizing smart parking, it is necessary to perform accurate and stable parking predictions. To study the relevant models of parking prediction, we collected a large number of literature from web of science, IEEE, ACM, and other publishers.

The prediction methods can be generally divided into two categories: knowledge-driven methods and data-driven methods. The former does not need a big quantity of data, but parking data is changing with time and there are a number of non-linear features for predictions, so it is not appropriate to employ knowledge-driven methods for parking prediction. Some studies [4], [5] apply knowledge-driven methods but their focus is still on data-driven methods to increase the prediction performance. Hence, we concentrate mostly on data-driven approaches, which rely on previous data to produce accurate forecasts. We divide the data-driven models into four generations based on general real-world prediction performance, i.e., statistical methods, machine learning methods, deep learning methods, and graph neural network models. Each of these methods has its own pros and cons, for instance, the statistical method is easy to model, nevertheless, this method fails to model complex parking scenarios with various features like weather, season, etc. Deep learning methods can cope well with the nonlinearity of parking problems but consume significant computational resources. Comparing and summarizing these methods allows researchers to understand the implications of choosing various models in order to rationally use these methods in real-world parking prediction problems.

Apart from the methods for prediction, it is also important to categorize different kinds of parking prediction problems, since a single problem does not provide a comprehensive scene of the needs and status of the entire parking forecasting field. In parking availability prediction problems [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], the occupancy rate of parking spaces in the future can be predicted based on the traffic flow of the lanes around the parking lot and other information, thus drivers can find free parking spaces in a shorter time. In parking behavior prediction problems [27], [28], [29], [30], [31], governors and drivers make sound judgments about what to do next based on accurate behavior predictions. In parking

Manuscript received 7 January 2022; revised 4 October 2022 and 25 February 2023; accepted 8 May 2023. Date of publication 2 June 2023; date of current version 4 October 2023. This work was supported in part by the China Post-Doctoral Science Foundation under Grant 2021TQ0260 and Grant 2021M700105, in part by the Guangzhou Science and Technology Program under Grant 202201010870, and in part by the Key Project of the Joint Fund of the National Natural Science Foundation of China under Grant U21A20446. The Associate Editor for this article was Y. Chen. *(Xiao Xiao and Zhiling Jin contributed equally to this work.) (Corresponding authors: Ziyan Peng; Zhiling Jin.)*

Xiao Xiao, Ziyan Peng, Yunqing Lin, Zhiling Jin, Rui Chen, Nan Cheng, and Guoqiang Mao are with the School of Telecommunications Engineering, Xidian University, Xi'an 710071, China (e-mail: xiaoxiao@xidian.edu.cn; ziyanpeng@stu.xidian.edu.cn; yunqing@stu.xidian.edu.cn; jinzhiling@stu.xidian.edu.cn; rchen@xidian.edu.cn; nancheng@xidian.edu.cn; guoqiangmao@xidian.edu.cn).

Wei Shao is with the School of Computing Technologies, RMIT University, Melbourne, VIC 3000, Australia (e-mail: wei.shao@rmit.edu.au).

Digital Object Identifier 10.1109/TITS.2023.3279024

demand prediction [27], [28], governors can make new plans rationally when they know the parking demands of an area.

Previous surveys [32], [33], [34] have considered traffic forecasting, including traffic flow prediction, traffic speed prediction, and pedestrian flow prediction, rather than parking prediction problems. A number of surveys [35], [36], [37], [38], [39] have discussed the parking prediction problem from different perspectives, which will be detailed in the next section. However, these surveys do not comprehensively investigate existing parking prediction methods nor do they categorize and summarize these methods. Towards this end, we summarize the parking prediction methods according to the evolution, which will be detailed in Section IV. Compared with the above surveys, our aim is to provide a comprehensive summary of parking forecasting schemes, which is useful for both the new researchers who want to catch up with the progress of parking forecasting and the experienced researchers who are unfamiliar with these latest graph-based solutions. The contributions of this paper are listed as follows:

- We present the most comprehensive summary of the parking prediction methods in generation order.
- We categorize the parking prediction problems and provide useful resources including evaluation metrics, open datasets, and open-source codes.
- We discuss several potential challenges and directions for future research in this field.

The rest of this survey is laid out as follows. In Section II, we compare the existing surveys with our work. In Section III, we present the main methods employed for parking prediction in generation order. In Section IV, we categorize the problems of parking prediction. In Section V, we present useful information for the parking prediction issues. In Section VI, we provide several potential challenges and directions for future research. In Section VII, we conclude our work.

II. Related Work

We give a brief introduction of the related surveys with this paper in this section. The differences in comparison to the surveys are listed in Table II.

A comprehensive survey about traffic prediction using graph neural networks (GNN) is given by [32]. This survey provides lots of traffic prediction problems: traffic speed prediction, traffic flow prediction, traffic demand prediction, etc. The parking availability prediction problems are also considered. A large number of datasets and codes are also provided, thus helping researchers replicate relevant works and further explore more complex prediction tasks, therefore, only three works [40], [41], [42] about parking availability prediction are mentioned, and the approach is also limited to graph neural networks. Statistical methods, machine learning, and deep learning are all described in our research.

Traffic prediction problems with deep neural networks (DNN) are discussed in [34]. This survey provides a large amount of traffic prediction problems like crowd flow prediction, crowd demand prediction, traffic flow prediction, etc. Many deep learning methods for prediction are also presented in detail, including multi-layer perceptron (MLP), convolutional neural networks (CNN), recurrent neural networks (RNN), and graph convolutional networks (GCN). However, this survey only considers deep learning methods without conventional statistical and machine learning methods. In addition, parking prediction problems are not considered in this survey. Long-term prediction problems are also ignored. We not only discussed the parking forecast but also involved the long-term forecast.

The survey [35] has given sight to the lens of market-based allocation of goods and services in smart parking. It pays attention to the economy of parking in smart cities. Some parking availability prediction issues are also mentioned in this survey. However, these issues are merely listed and not analyzed and categorized in detail, and in addition, there is a scarcity of open datasets and comparisons of alternative methods. For most researchers, the review does not help effectively explore the parking prediction problem in depth. We analyzed the parking availability prediction in detail and compared the open data sets and methods.

Smart parking systems (SPS) are classified based on the soft and hard design factors in [36]. This review focuses on the relevance of data dependability, security, privacy, and other critical design elements in SPS. The topic of parking availability prediction is also addressed in this review, but solely as part of the software systems in smart parking. The advantages of parking availability prediction are mentioned by listing some relevant studies. However, like the survey [35], this survey just lists the related parking prediction studies without categorizing the problems or comparing prediction methods. However, we classified parking problems.

In [37], SPS are introduced and smart parking solutions are provided. Literature from 2000 to 2016 is considered in this review, including three main topics: information collection, system deployment, and service dissemination. When introducing the parking system deployment, parking prediction is incidentally presented. Some typical parking vacancy prediction examples and sources of related datasets are listed in this part. However, these parking prediction problems are considered to show the importance of parking system deployment, while the prediction methods and evaluation metrics are simply mentioned. SPS are also surveyed in [38]. This survey focuses on many types of vehicle detection methods (VDT), such as in-roadway sensors, over-roadway sensors, crowdsensing, prediction models, and so on. They consider parking prediction models as detection methods and do not discuss the methods for parking forecasting in detail. However, we mainly discuss the parking problem.

The behavior of drivers for searching parking spaces and the demand for parking is studied in [44]. Some parking concerns are examined, as well as parking characteristics and their applications, driver parking decision behavior, and the construction of demand models that take into account various factors. When introducing the parking demand issues, the authors surveyed some parking demand forecasting models using different methods such as linear regression (LR), least square regression (LSR), gaussian mixture model, and so on. However, parking demand prediction is simply a part of parking prediction problems. Besides, the latest deep learning

TABLE I

THE ABBREVIATIONS USED IN THIS ARTICLE. (ABBREVIATIONS HAVE BEEN SORTED ALPHABETICALLY)

Abbreviation	Type	Description	Abbreviation	Type	Description
ANN	Method	artificial neural network	MAPE	Metric	mean absolute percentage error
AR	Method	autoregressive model	MASE	Metric	mean absolute scaled error
ARIMA	Method	autoregressive integrated moving average model	MLP	Method	multi-layer perceptron
ARMA	Method	autoregressive integrated average model	MRD	Metric	mean relative deviation
AVR	Concept	availability rate	MSE	Metric	mean square error
BPNN	Method	back propagation neural network	NN	Method	neural network
CNN	Method	convolutional neural network	PGIS	Concept	parking guidance and information systems
COVID-19	Concept	coronavirus disease 2019	PIS	Concept	public information services
DNN	Method	deep neural network	POR	Concept	parking occupancy rate
DT	Method	decision trees	R^2	Metric	the coefficient of determination
FFNN	Method	feed-forward neural network	RAE	Metric	relative absolute error
FNR	Metric	false negative rate	RBF	Concept	radial basis function
FPR	Metric	false positive rate	RF	Method	random forest
GAT	Method	graph attention networks	RMSE	Metric	root mean square error
GCN	Method	graph convolutional network	RNN	Method	recurrent neural network
GNN	Method	graph neural network	RRSE	Metric	root relative squared error
GraphSage	Method	graph sample and aggregate	RSE	Metric	relative squared error
GRU	Method	gated recurrent unit	SAE	Method	stacked autoencoder
HA	Method	historical average	SARIMA	Method	seasonal autoregressive integrated moving average model
IoT	Concept	Internet of things	SPS	Concept	smart parking Systems
ITS	Concept	intelligent transportation systems	SSE	Metric	sum of square error
LASSO	Method	least absolute shrinkage and selection operator	SVM	Method	support vector machine
LR	Method	linear regression	SVR	Method	support vector regression
LSR	Method	least square regression	VDT	Concept	vehicle detection techniques
LSTM	Method	long short-term memory	WMAPE	Metric	weighted mean absolute percentage error

TABLE II

SUMMARY OF RELATED SURVEYS. "✓" INDICATES A COMPREHENSIVE ANALYSIS, "✓*" INDICATES A GENERAL ANALYSIS, "✓†" INDICATES LITTLE ANALYSIS, AND "–" INDICATES NOT CONSIDERED

Survey	Year	Prediction Methods	Parking Prediction Problems	Open Resources
Lin *et al.* [37]	2017	✓†	✓*	-
Kotb *et al.* [43]	2017	-	-	-
Al-Turjman *et al.* [36]	2019	✓†	✓†	-
Diaz *et al.* [38]	2020	✓†	✓†	-
Parmar *et al.* [44]	2020	✓†	✓†	-
Yin *et al.* [45]	2020	✓	-	✓
Boukerche *et al.* [46]	2020	✓*	-	✓*
Jiang *et al.* [32]	2021	✓*	✓*	✓
Lee *et al.* [34]	2021	✓*	-	✓
Aljohani *et al.* [35]	2021	✓†	✓*	-
ours	-	✓	✓	✓

prediction methods are not listed, which cannot fully reflect recent developments in parking prediction issues. We have listed the latest deep learning and graph neural networks, reflecting the latest progress in parking prediction.

Additional reviews which do not mention parking prediction problems also conclude other important information that may be useful for parking prediction. For example, [43] introduces the PGIS and the parking reservations systems (PRS). The former can provide plenty of sensor data for parking prediction, while the latter can provide information about vacant parking lots. Methods for traffic prediction are surveyed in [45], such as machine learning methods, CNN, GCN, etc. By reproducing the relevant methods, researchers can also try to use them for parking prediction tasks. A large amount of data and open-source codes are also collated in this survey, but evaluation metrics are not collated in detail. Review [46] then summarize the machine learning methods for traffic prediction problems, nevertheless, the open-source codes are not present in this literature.

III. PARKING PREDICTION METHODS

In this section, we classify the parking prediction models into four generations according to their performance. The models include statistical models, machine learning models, deep learning models, and graph neural networks (GNN) models. For each generation, we focus on detailing the most commonly used prediction models.

A. Statistical Methods (First Generation)

1) Historical Average: Historical average (HA), which treats parking time series data as a seasonal process and utilizes the average of prior seasons as forecasts, is one of the simplest statistical prediction models. For example, suppose we need to forecast the future parking occupancy rate (POR); if we consider one week as a season, we can predict the future POR at 8:00 am on Friday using the average of all the historical POR at 8:00 am on Friday. In recent years, HA has been used as a baseline in many studies [2], [12], [47], [48] due to its simplicity. However, it is important to note that, unlike most prediction models, the prediction results of HA are not affected by the prediction horizons.

2) Autoregressive Integrated Moving Average: Based on classical statistics, the autoregressive integrated moving average (ARIMA) model [49] is among the most consolidated approaches. This model has been widely employed for parking prediction problems [2], [13], [17], [50]. The non-seasonal ARIMA model is generally denoted ARIMA (p, d, q), where p is the order (number of time lags) of the model, d is the degree of differencing (the number of times the data has had prior values removed), and q is the moving-average order.

Jose et al. [17] employed the seasonal autoregressive integrated moving average (SARIMA) model for parking availability prediction. The SARIMA model can obtain higher prediction accuracy since it aggregates seasonal information. Both ARIMA and SARIMA are simple and do not rely on external variables, nevertheless they are not suitable for non-linear time-series data and their parameters are difficult to estimate [51].

3) Other Statistical Methods: Xiao et al. [26] proposed a predictive framework based on the continuous-time Markov M\M\C\C queue approach. Due to the time-dependent analytical features of the underlying queuing model, this approach is computationally efficient and outperforms the baselines, including some contemporary machine learning methods. Rong et al. [52] applied linear interpolation as a baseline. This model is based on a distance-weighted interpolation algorithm, which means that the closer the parking lots are to one other, the more comparable the parking availability data will be [53]. Bock et al. [1] employed the Kalman filters for on-street parking availability predictions and observed that the use of Kalman filters did not reach a significant statistical improvement.

The statistical methods are often simple and do not spend too much time in obtaining the prediction results. However, in the real world, several factors impact parking behavior, including weather, season, weekdays, and so on. Thus, simple statistical models are often not suitable for modeling complex parking time-series data.

B. Machine Learning Methods (Second Generation)

Machine learning methods can process more complex data and achieve higher accuracy [46]. In this section, we mainly introduce four categories of machine learning methods for parking prediction: regression models, kernel-based models, example-based models, and neural network-based models.

1) Regression Models: The regression models can be regarded as the parametric methods, where the mathematical model and relevant input-output parameters have been specified in advance, and the connections between the input and each parameter have also been determined relatively [54]. The main regression models are listed as follows:

- **Linear regression** (LR) is the most basic machine learning algorithm, and it is commonly employed in parking prediction [23], [28], [55], [56]. It primarily uses the least squares technique to describe the connection between a scalar response and one or more explanatory factors. However, LR does not fit nonlinear data well.
- **Decision tree**, which represents the relationship between object properties and values in terms of mapping. Each leaf node corresponds to the object value represented by the path from the root node to the leaf node, and each forked path corresponds to a potential attribute value. Due to their comprehensibility and simplicity, decision tree (DT) methods have been used in various parking prediction situations in recent years. [7], [9], [10], [18], [28], [52], [56], [57], [58], [59]. However, complex tree structures may cause overfitting problem, i.e., poor generalization from the training data. Besides, the tree structures are very non-robust, a tiny change of training data may lead to a large change in tree structures [60]. The random forests (RF) algorithm [61] is another regression method constructed by a large number of decision trees. This method can deal with a large number of inputs and it is more robust than decision trees, which have attracted the attention of many researchers [10], [18], [28], [30], [58], [59], [62].

  The random forest algorithm is a method proposed by Breiman [63] to gather multiple classification regression trees for making voting decisions. Aiming at the problem of parking behavior detection, Feng et al. investigated linear regression, ridge regression, lasso regression, decision tree, and random forest for parking behavior prediction in the context of weather, and discovered that random forest performs the best.
- **Least absolute shrinkage and selection operator** (LASSO) is a regression analysis machine learning approach that improves prediction accuracy by performing both variable selection and regularization [64]. This method can capture the non-linear features of parking events and has been employed in some studies [12], [48].

However, these regression models just require little parking time-series data for prediction. When faced with a vast quantity of data, it struggles to capture the seasonal fluctuations and dynamics of the data and is prone to overfitting.

2) Kernel-Based Models: For the kernel-based models, we mainly consider the support vector machine (SVM) [65]. It is also known as a support vector regression (SVR) when it is utilized for prediction or regression tasks. When compared to other nonlinear prediction models, it can avoid slipping into a local optimum [46], and has been widely adopted for parking prediction [48], [50], [56], [58], [66], [67]. The main principle behind SVM learning is to find the separated hyperplane that partitions the training data set appropriately and optimizes geometric separation. All sample points should be near to the hyperplane when performing prediction or regression tasks, and the overall variance between the sample points and the hyperplane should be minimized.

The estimated continuous-valued function of a one-dimensional example can be expressed as:

$$y = f(x) = \begin{bmatrix} w \\ b \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} x \\ 1 \end{bmatrix} = w^{\mathrm{T}} x + b, \tag{1}$$

where x is the input vector, w is the normal vector to the hyperplane, and b is a bias. To conduct prediction or regression tasks, SVR often approximates Eq. (1) as an

TABLE III
THE COMMON KERNEL FUNCTIONS OF SVR. γ IS A ADJUSTABLE PARAMETER

Kernel Name	Function	Relevant Works
Linear	$K(x_i, x_j) = x_i^{\mathrm{T}} x_j$	[48], [50], [67]
RBF	$K(x_i, x_j) = -\exp(\gamma\|\|x_i - x_j\|\|_2^2)$	[7], [23], [55]
Sigmoid	$K(x_i, x_j) = \tanh(\gamma x_i^{\mathrm{T}} x_j + b)$	[69]

optimization problem. Considering the linearly non-separable case in SVR [68], the optimization problem for SVR is derived after introducing the relaxation variables. Applying the Lagrange multiplier, the above formula can be expressed as:

$$f(x) = w^T x + b = \sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right)(x, x_i) + b, \tag{2}$$

where α_i and α_i^* are Lagrange multipliers. x is the input vector, e.g., the time series of historical parking occupancy rate and x_i is the i-th element of x.

As we indicated before, SVR is a typical kernel-based model, which applies a kernel function to make the prediction problem linearizable [68]. The kernel functions can be different, the most widely-used kernel functions are shown in Table III. Using the kernel function, Eq. (2) can be rewritten as:

$$f(x) = w^T x + b = \sum_{i=1}^{n} \left(\alpha_i - \alpha_i^*\right) K(x_i, x) + b, \tag{3}$$

where $K(\cdot)$ is the kernel function.

3) Example-Based Methods: In this survey, the example-based methods mainly indicate the K-nearest neighbors (KNN) method. Since the KNN method can capture the extra spatial correlations from traffic networks, many researchers have employed it for parking prediction [57], [67].

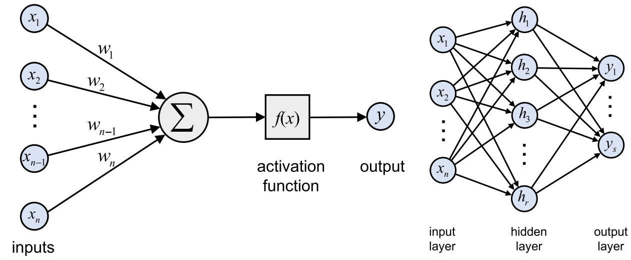
The KNN model is a data-based non-parametric regression approach that predicts the label $\mathbf{y} \in \mathbb{R}^d$ of the pattern $\mathbf{x} \in \mathbb{R}^q$ based on a set of N observations, which has two steps: K-means clustering [21], [22] and non-parametric regression forecasting based on the clustering analysis [70]. In K-means clustering, the number of clusters is set to K, and the data is then converged to the center of each cluster. The state vector defining the current state selects the neighbor data in the forecasting step, and the neighbor points are then utilized to anticipate the value of the following time step. The distance metric is very important in order to capture the neighbors correctly. The distance metric mainly includes the Manhattan, Minkowski, and Euclidean metrics [71].

KNN applies the regression function KNN$(\cdot)$ to calculate the mean of the function values of its K-nearest neighbors:

$$\mathrm{KNN}(\mathbf{x}) = \frac{1}{K} \sum_{i \in N_k(\mathbf{x})} \mathbf{y}_i, \tag{4}$$

where $N_k(\mathbf{x})$ is a set containing the indices of $\mathbf{x}$.

The KNN approach is straightforward, easy to comprehend, and apply. It also does not require parameter estimation. However, when the samples are imbalanced, such as when the sample size of one class is large and the sample size of the other is small, it is possible that when a new sample is an input, the large-capacity class's samples will predominate among the K neighbors of that sample.



(a) The structure of an artificial neuron. (b) The artificial neural networks.

Fig. 1. The illustration of the artificial neural networks and the structure of the artificial neuron [74].

4) Neural Network-Based Models: The neural network-based models learn parameters by training the historical observations [70], which show strong robustness and fault tolerance and are thus suitable for complex and dynamic traffic prediction. When compared to other traditional approaches, neural network-based models can obtain better prediction performance with a large amount of data. However, compared with other machine learning models, this model requires more computing resources and larger datasets to conduct accurate predictions. With the improvement of hardware computing performance, neural network-based models have been widely used in the field of parking prediction [7], [10], [20], [26], [50], [72].

The models for the neural network-based machine learning methods usually include the basic artificial neural networks (ANN) or the neural networks (NN). The basic unit of ANN is the artificial neuron, the structure of which is shown in Figure 1(a). In this figure, $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ is the input vector, $\mathbf{w} = \{w_1, w_2, \ldots, w_n\}$ is the corresponding weight vector, and the function $f(x)$ is the activation function. The output of a neuron is obtained as follows:

$$y = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right), \tag{5}$$

where b refers to the bias of the neuron and $\sigma(\cdot)$ refers to the activation function, which has many properties to enhance or simplify the network containing the neuron [73].

There are connections among the neurons. The output of one neuron can be used as the input of the next neuron. Many connected neurons can form ANN, which is shown in Figure 1(b). During training, the weight vector $\mathbf{w}$ and the bias b are continuously modified to fit the real model and minimize the loss function [75]. When the loss is small enough, the training process is finished and a final prediction output $\hat{y}$ can be obtained. The capacity to fully simulate complicated nonlinear connections, associative memory, and resilience and defect tolerance to noisy nerves may all be found in ANN models. However, ANN models still have some drawbacks including requiring large numbers of parameters, spending a long time on training, and lack of interpretability, etc.

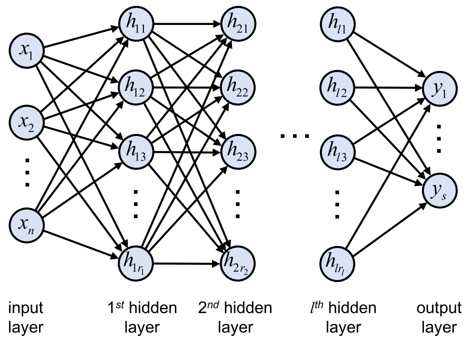


Fig. 2. The structure of multi-layer perceptron (MLP) model [13].

C. Deep Learning Methods

Generally speaking, deep learning is a subclass of machine learning whose technology is based on the ANN model [74]. Compared with classical machine learning methods, deep learning methods require large amounts of data, learn on their own from the environment and past mistakes rather than human interventions, and usually spend a long time on training and achieving higher prediction accuracy. In this section, we mainly introduce deep neural networks (DNN), including feed-forward neural networks (FFNN), recurrent neural networks (RNN), and convolutional networks (CNN).

1) Feed-Forward Neural Networks: The simplest structure of FFNN is a structure with one-hidden layer, i.e., the hidden layer of the FFNN is single. This structure can be regarded as an ANN model, which is easy to implement but lacks robustness. To solve the problems, some researchers added an extra hidden layer to make more accurate predictions [9]. Sometimes, this deep FFNN model can be also called the multi-layer perception (MLP) [13], [14], [76], the structure of which is shown in Figure 2.

The parameters of back propagation neural network (BPNN) [77] models are adjusted during training to improve the accuracy of prediction outputs., which makes it also used in parking prediction problems [48], [67], [78].

However, on the whole, compared with other latest deep learning methods, the structure of FFNN models is simple and the spatial or temporal dependencies are often ignored, therefore, the FFNN models are usually adopted as deep learning baselines in recent studies.

2) Recurrent Neural Networks: For parking prediction problems, apart from the current information, historical information is also important for the outputs. For example, if an area is short of parking spaces in the morning, drivers are more likely to avoid this time of day for parking.

The RNN models [79], [80] update the current state depending on previous states and current inputs [81], which can model more complex temporal dependencies and be applicable for parking prediction tasks. In this survey, we mainly focus on three kinds of RNN models: the vanilla RNN models, the long-short term memory (LSTM) networks models, and the gated recurrent unit (GRU) models.

- **Vanilla recurrent neural networks**. The structure of vanilla recurrent neural networks is shown in Figure 3. We can express the calculation of RNN by the following formula:

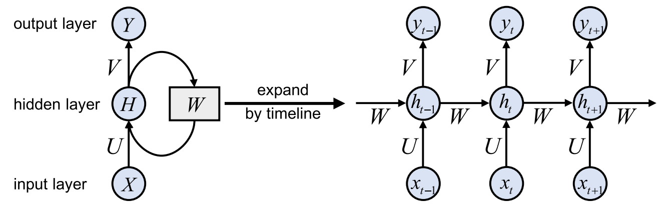


Fig. 3. The structure of vanilla recurrent neural networks [79], [80].

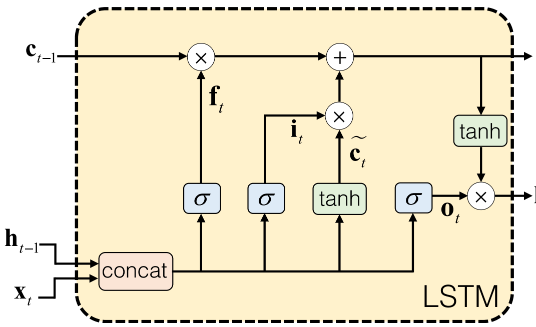
$$\mathbf{h}_t = \tanh\left(U\mathbf{x}_t + W\mathbf{h}_{t-1} + \mathbf{b}_h\right),$$
$$\mathbf{y}_t = \text{softmax}\left(V\mathbf{h}_t + \mathbf{b}_y\right), \tag{6}$$

where W, U, and V are the weight matrix. $\tanh(\cdot)$ and softmax$(\cdot)$ are the activation functions and $\mathbf{b}_h$ and $\mathbf{b}_y$ are the bias vectors. The parameters W and $\mathbf{b}$ used in each time step are the same. At time t, after this network receives the input x_t, the value of the hidden layer is h_t and the output value is y_t. The key point is that the value of h_t does not only depend on the value of x_t but also h_{t-1}, which makes RNN capture temporal dependencies and be suitable for parking prediction [17], [21], [27], [31], [40], [41], [50], [62], [67], [78], [82], [83], [84], [85].
However, the vanilla RNN models are unable to connect the relevant input information with a large gap [81], i.e., it is hard for vanilla RNN models to extract long-term dependencies. In addition, the vanilla RNN models face the problems of exploding and vanishing gradients [86].

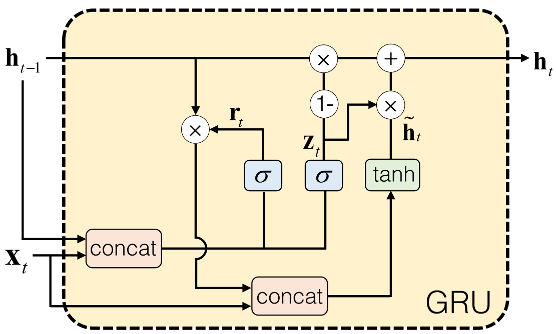
- **Long Short-term Memory Networks**. To handle the "long-term dependencies" problems, long-short term memory networks were proposed by Hochreiter et al. [87]. Compared to the vanilla RNN, LSTM adopts the gating mechanism to regulate the transmission of information. The gates in LSTM include the input gate, output gate, and forget gate [88]. The structure of LSTM is shown in Figure 4(a). The calculation of LSTM can be shown as

$$\begin{aligned}
\mathbf{i}_t &= \text{sigmoid}\left(\mathbf{W}_{xi}\mathbf{x}_t + \mathbf{W}_{hi}\mathbf{h}_{t-1}\right),\\
\mathbf{f}_t &= \text{sigmoid}\left(\mathbf{W}_{xf}\mathbf{x}_t + \mathbf{W}_{hf}\mathbf{h}_{t-1}\right),\\
\mathbf{o}_t &= \text{sigmoid}\left(\mathbf{W}_{xo}\mathbf{x}_t + \mathbf{W}_{ho}\mathbf{h}_{t-1}\right),\\
\tilde{\mathbf{c}}_t &= \tanh\left(\mathbf{W}_{xc}\mathbf{x}_t + \mathbf{W}_{hc}\mathbf{h}_{t-1}\right),\\
\mathbf{c}_t &= \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t,\\
\mathbf{h}_t &= \mathbf{o}_t \odot \tanh\left(\mathbf{c}_t\right),
\end{aligned} \tag{7}$$

where $\mathbf{i}_t$, $\mathbf{f}_t$, $\mathbf{o}_t$, $\tilde{\mathbf{c}}_t$, $\mathbf{c}_t$, and $\mathbf{h}_t$ are the input gate, forget gate, output gate, candidate gate, cell gate and the hidden status at time t, respectively. $\odot$ is the element-wise Hadamard product [89]. $\mathbf{f}_t$ determines the importance of $\mathbf{c}_{t-1}$ to calculate $\mathbf{c}_t$. When the switch of $\mathbf{f}_t$ is on, the network will disregard $\mathbf{c}_{t-1}$.
Since this structure can capture long-term dependencies of time series data and can avoid exploding and vanishing gradients [90], the LSTM models are widely employed for parking prediction [2], [12], [14], [17], [24], [27],



(a) The structure of LSTM [88].



(b) The structure of GRU [99].

Fig. 4. The illustration of the long-short term memory (LSTM) networks and the gated recurrent units (GRU). σ indicates the sigmoid activation function.

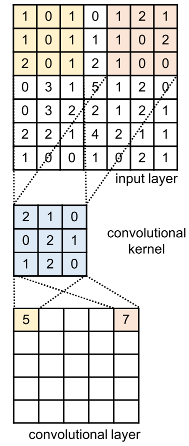
[31], [48], [52], [67], [78], [82], [83], [84], [91], [92], [93], [94], [95], [96], [97].

- **Gated Recurrent Units**. The GRU model proposed by Kyunghyun et al. [98] is another RNN variant to solve the exploding and vanishing gradients problems. The structure of GRU is shown in Figure 4. The gates value, hidden states of GRU can be calculated as

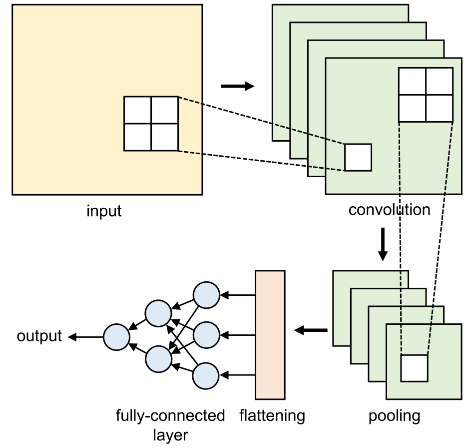
$$
\begin{aligned}
\mathbf{z}_t &= \text{sigmoid}\left(\mathbf{W}_{xz}\mathbf{x}_t + \mathbf{W}_{hz}\mathbf{h}_{t-1}\right) \\
\mathbf{r}_t &= \text{sigmoid}\left(\mathbf{W}_{xr}\mathbf{x}_t + \mathbf{W}_{hr}\mathbf{h}_{t-1}\right) \\
\tilde{\mathbf{h}}_t &= \tanh\left(\mathbf{W}_{xh}\mathbf{x}_t + \mathbf{r}_t \odot \left(\mathbf{W}_{hh}\mathbf{h}_{t-1}\right)\right) \\
\mathbf{h}_t &= (\mathbf{1} - \mathbf{z}_t) \odot \tilde{\mathbf{h}}_t + \mathbf{z}_t \odot \mathbf{h}_{t-1}
\end{aligned}
\tag{8}
$$

where $\odot$ is the element-wise Hadamard product. We can see from the figure that at time t the reset gate $\mathbf{r}_t$ is used to control the effect of the previous moment hidden status $\mathbf{h}_{t-1}$ on the current input $\mathbf{x}_t$. If $\mathbf{h}_{t-1}$ is not important to $\mathbf{x}_t$, the reset gate $\mathbf{r}_t$ can be open to eliminate the effect of $\mathbf{h}_{t-1}$ on $\mathbf{x}_t$. update date $\mathbf{z}_t$ is used to decide whether to ignore $\mathbf{x}_t$ and can avoid exploding and vanishing gradients like the forget gate of LSTM [99].
The GRU models aggregate the forget and input gates in the LSTM into a single update gate and eliminate the cell state, resulting in fewer parameters and a simpler calculation structure, therefore, compared with the LSTM models, the GRU models usually spend less time on training and have been applied by [29], [40], [41], [48], [67], and [84] for parking prediction.

3) Convolutional Neural Networks: The parking forecasting problem not only has time dependence but also has a strong spatial correlation. For example, an area for parking can show an impact on the adjacent areas. The close distance between the two areas will lead to strong relationships. To capture the spatial correlations, some studies have adopted the CNN based models for parking prediction [10], [12], [24], which show better performance than the methods ignoring the spatial dependencies.



(a)



(b)

Fig. 5. The illustration of convolutional operation (a) and the structure of CNN (b). The moving step of the convolutional operation is 1 [100].

The structure of CNN models proposed by LeCun et al. [100] is shown in Figure 5(b). The main components of CNN include the convolutional layer, the pooling layer, and the fully-connected layer. The convolutional layers can extract key information from input (such as the images) to a feature map through convolutional operations (shown in Figure 5(a)). The pooling layer is used for minimizing the feature dimension. The fully-connected layers produce classifications by connecting all of the neurons in one layer to those in another.

The convolutional kernel is the key component of the convolutional layer. As Figure 5(a) shows, the convolutional layer is created by shifting the convolution kernel to the given step on the preceding layer. The weight of the kernel is shared in convolutional layers, which can reduce the network's complexity and boost its generalization ability by controlling the number of free parameters [46].

To leverage these advantages, some researchers model the parking networks into grid structures. For example, Provoost et al. [10] created an 8×8 kernel for the traffic inputs and a 4×4 kernel for the look-back window to capture the spatial correlations. Ghosal et al. [12] modeled the block-level parking lots into grid structures for parking occupancy prediction. After using CNN models for capturing the spatial correlations, to further capture the temporal dependencies of parking prediction problems, the CNN models are often combined with RNN based models [12], [24]. Under these circumstances, the spatio-temporal correlations are captured, which is more possible to achieve higher prediction compared with methods without taking spatial dependencies into account.

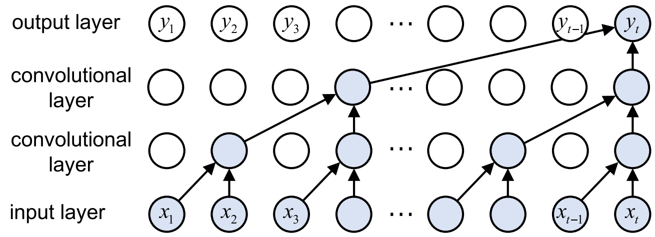
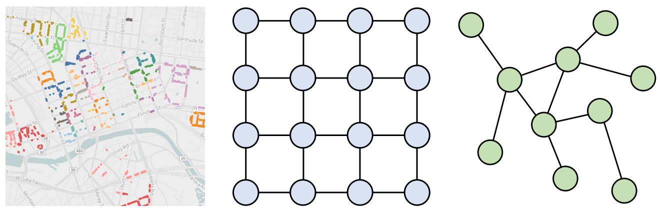


Fig. 6. The illustration of the structure of one-dimension CNN [101].



(a) (b) (c)

Fig. 7. The distribution of parking bays in Melbourne (a) [2], [14] and the illustrations of Euclidean structure data (b) and the non-Euclidean structure data (c).

Apart from capturing the spatial correlations, the CNN models, e.g., the one-dimension CNN [102], can also be employed for time series prediction. The illustration of the CNN model used for time series prediction is shown in Figure 6. The one-dimension CNN can learn time series dependences [101] and can avoid vanishing or exploding gradients [103]. In RNN, the prediction of the subsequent time step must wait for the completion of the previous time step, nevertheless the convolution in CNN can be performed in parallel because each layer uses the same filter, which leads to a high training efficiency [2]. However, the CNN model requires the use of raw sequences up to a valid historical record length and therefore may require more memory during the evaluation. Besides, the potential parameter changes should also be considered [104].

D. Graph Neural Network Models

The CNN models are generally suitable for Euclidean structure data, which is shown in Figure 7(b). However, the parking spaces are often presented as a non-Euclidean structure (shown in Figure 7(c)). Figure 7(a) shows the distribution of parking bays in Melbourne. Each point indicates a parking bay and different colors indicate different areas, therefore, modeling the spatial dependencies using CNN models is not quite appropriate. To solve the problems, researchers have used the graph to model the parking space networks and capture the spatial dependencies.

The graph of a parking network with N nodes at time t can be denoted as $\mathcal{G}_t = (\mathcal{V}_t, \mathcal{E}, W)$, where $\mathcal{V}_t$ is the set of vertices at time t, $\mathcal{E}$ is the set of edges that connect the vertices, and $W \in \mathbb{R}^{N\times N}$ is the weighted adjacent matrix of $\mathcal{G}_t$, the element of which w_{ij} indicates the relationships of vertex v_i and v_j. The weight matrix is regarded as the key to capturing spatial dependency [105], which is often calculated by the distance of nodes. Xiao et al. [2] adopted the Gaussian kernel [106] to calculate the weight matrix, which is expressed as

$$w_{ij} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), & i \neq j \text{ and} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \geq \varepsilon, \\ 0, & \text{otherwise}. \end{cases} \tag{9}$$

where d_{ij} is the distance between vertex v_i and v_j, ε is the gate to decide on the density of W, and σ^2 is for controlling the distribution of W.

To process the graph data, graph neural networks (GNN) are proposed. The GNN models can be divided into two categories: the spectral methods and the spatial methods. The spectral methods transfer graph data to the spectral domain for convolution and the most commonly used approach of the spectral methods is graph convolutional networks (GCN). The GCN is first proposed by Bruna et al. [107]. They first calculate the symmetric normalized Laplacian matrix L by $L = I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$, where $I \in \mathbb{R}^{N\times N}$ is the identity matrix and $D \in \mathbb{R}^{N\times N}$ is the degree matrix. Then the graph convolution operation is defined as

$$\Theta *_{\mathcal{G}} x = \Theta(L)x = \Theta\left(U\Lambda U^T\right)x = U\Theta(\Lambda)U^T x, \tag{10}$$

where x is the input, Θ is the graph kernel, Λ is the diagonal matrix of eigenvalues of L, and U is the matrix made up of eigenvectors of L. To reduce the complexity of the convolution operations, Defferrard et al. [108] proposed a localized spectral filtering method. Using these methods, a polynomial filter is expressed as

$$\Theta(\Lambda) = \sum_{k=0}^{K-1} \theta_k \Lambda^k, \tag{11}$$

where θ_k is the corresponding coefficient of Λ^k and K is the size of the graph convolution kernel. In this way, the graph convolution is K-localized and the complexity has reduced from $\mathcal{O}(N^2)$ to $\mathcal{O}(K|\mathcal{E}|)$. The subsequent Chebyshev approximation [108] and first-order approximation [109] can further minimize computing complexity even more and prevent gradient explosion. In conclusion, the spectral-based GNN models have sufficient theoretical knowledge, and some studies have used GCNs to extract spatial correlations and then used them for parking prediction [2], [29], [48], [110].

Different from the spectral-based GNN models, the spatial-based GNN models employ to use the message-passing mechanism and define the graph convolution based on the spatial relationship of the nodes. The central node representation and the neighbor node representation are combined in the graph convolution of spatial-based GNN models to create a new representation of that node. The two typical representatives of the spatial-based GNN models are graph sample and aggregate (GraphSage) [111] and graph attention networks (GAT) [112]. The former extends traditional GCN into an inductive learning task by training a function (convolutional layer) of the neighbors of the aggregated nodes, which acts as a generalization to unknown nodes. The latter then applied a masked self-attention layer to solve the problems of previous models based on graph convolution, which does not need to use pre-built graphs.

Compared with spectral-based GNN models, the spatial-based GNN models can be used for directed graphs and non-fixed graphs [113]. In recent years, these spatial-based GNN have been widely adopted for prediction like traffic prediction [114], [115], [116]. Unfortunately, we have only found two related works [40], [41] using spatial-based GNN models for parking prediction.

In summary, different forecasting methods have their own characteristics, and we summarize all the methods used for parking forecasting in Table IV.

IV. Problems

There have been various works for parking prediction in recent years. In this section, we discuss the state-of-the-art works and categorize the different types of parking prediction problems considered in this article. Table VI shows a complete list of the parking prediction issues explored in this study.

A. Parking Availability Prediction

Parking availability shows strong seasonal characteristics, such as the time-of-the-day repetitive pattern and the day-of-the-week repetitive pattern, which can be dug out from historical observations [128]. The prediction problem definition is predicting the data of the next H time steps based on the previous M historical observations, which can be defined as

$$\begin{aligned} &\hat{a}_{t+1}, \ldots, \hat{a}_{t+H} \\ &\quad = \underset{a_{t+1},\ldots,a_{t+H}}{\arg\max} \left\{ \log \Pr\left(a_{t+1}, \ldots, a_{t+H} \mid a_{t-M+1}, \ldots, a_t\right) \right\}, \end{aligned} \tag{12}$$

where a_t is the historical observation at time t, $\hat{a}_{t+1}$ is the predicted value as time $t+1$, and $\Pr(\cdot|\cdot)$ is the conditional probability function. The historical observations change with different prediction tasks, for example, when using the parking occupancy rate (POR) to evaluate parking availability, the historical observations become the historical POR. Generally speaking, parking availability prediction is the most common of the parking prediction problems, including predicting the parking occupancy rate (POR), and the number of free parking slots, and the car parking count. The main prediction task of some works is parking space prediction, which is ambiguous. Towards this end, we have carefully analyzed these works to classify them into the above three categories of parking availability forecasting.

1) Parking Occupancy Rate: Among the three parking availability prediction problems mentioned above, the POR prediction has been studied in many works [2], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [76], [110]. Compared with other two prediction tasks, the POR prediction can indicate the parking availability directly without providing detailed parking lot information. When the POR is close to 1, the driver will know that there are not enough parking spaces left in the area and will choose another area to park, which will effectively alleviate the problem of spending a lot of time looking for parking spaces.

The calculation of POR is important, which relates to the accuracy of the time series used for prediction. In previous studies, this problem has not been studied yet. In this survey, we divide the POR calculation into two categories, the average POR and the instantaneous POR. Suppose there are N parking bays in an area a, the average POR of this area at time t is calculated as follows:

$$\mathrm{POR_{av}} = \frac{\sum_{i=1}^{N} \tau_i}{NT}, \tag{13}$$

where τ_i is the parking duration time of i-th parking bay in the time interval $[t, t+T)$ and T is the sample time interval to generate POR time series. Correspondingly, the instantaneous POR of this area at time t is calculated by:

$$\mathrm{POR_{inst}} = \frac{n_t}{N}, \tag{14}$$

where n_t is the number of occupied parking bays in area a at time t. It is obvious that the instantaneous POR is more appropriate to respect the real parking availability since it brings POR to a precise time step. This calculation method has been adopted by several studies [2]. However, the average occupancy is more commonly used since it can be calculated by knowing only the arrival time and the departure time of the parking bay, which helps pre-processing parking data quickly. This method has been employed by [14].

2) The Number of Free Parking Slots: This problem is easy to explain, i.e., predicting the number of free parking slots in the future based on historical studies. Compared with POR prediction, this prediction problem has its own advantages, i.e., when a parking area is large and has a large number of vehicles parked, using POR to indicate parking availability may cause drivers to choose to park in other areas, even though there are a certain number of free parking bays in the area. Imagine a scenario where a parking lot has 1,000 parking bays, one driver knows it has a POR of 0.95, and another driver knows there are 50 free parking bays in the parking lot. Then often the first driver will go to another parking lot and the second driver will continue to that parking lot. Considering the role of this prediction task, there are also several works using the number of free parking slots for parking availability prediction [47], [50].

3) Car Parking Count: Some studies also use the car parking count [19], [28], [48], [67], [92], i.e., the number of occupied parking bays, to predict the parking availability. This is similar to predicting the number of free parking slots. However, if the number of occupied parking bays is used to predict parking spaces, drivers need to know additional information about the number of parking spaces in the parking area., therefore, we do not recommend this approach for predicting parking availability.

4) Availability Rate: The availability rate (AVR) was first proposed by Tekouabou and Abdellaoui et al. [59], [62]. The AVR is defined as the ratio of the capacity minus the occupation at time t, on the capacity of the parking lot, which can be calculated by

$$\mathrm{AVR}(t) = \frac{\mathrm{Capacity} - \mathrm{Occupancy}(t)}{\mathrm{Capacity}}, \tag{15}$$

TABLE IV
PARKING PREDICTION METHODS IN THE SURVEYED STUDIES

Methods		Relevant Works	Strength	Limitation
statistical methods	HA	[12], [2], [47], [48]	The algorithm is simple, and the parameters can be estimated online by the least square method (LS).	Without considering the influence of the current state change.
	ARIMA	[50], [13], [2], [17], [20], [23], [15], [47]	Short-term prediction works well.	Can only capture linear relationships, not nonlinear relationships.
	linear interpolation	[52]	Guaranteed convergence.	Only the continuity of each small curve at the connection point can be ensured.
	Kalman filters	[1]	Only the current measured value and the predicted value of the previous cycle are needed for state estimation.	When the moving target is blocked for a long time, the tracking of the target will be lost.
	continuous-time Markov model	[11], [26]	The probability of a system with maintenance capability and multiple degraded states can be calculated.	The predicted sequence may have some parts that don't actually happen.
machine learning methods	regression models	LR: [55], [28], [23], [56]; DT and RF: [57], [9], [52], [56], [58], [7], [10], [18], [59], [117], [62], [30]; LASSO: [12], [48]	It is simpler and more convenient to analyze multi-factor models.	The algorithm is relatively simple and low-level.
	kernel-based models	SVM/SVR: [66] , [56], [58], [7], [50], [55], [48], [67], [23]	Kernel function can be used to map to high-dimensional space, and nonlinear classification can be solved.	It is difficult to train large-scale data and cannot directly support multi-classification.
	example-based methods	KNN: [67], [57]	It can be used for classification and regression, and the training time complexity is lower than that of algorithms such as support vector machines.	A large amount of calculation, especially when there are many features, leads to a slower prediction speed than logistic regression and other algorithms. Compared with the decision tree model, the KNN model is not interpretable.
	neural network-based models	ANN: [7], [50], [10], [72], [26], [20]	Strong parallel processing ability, strong robustness and fault tolerance to noisy nerves, and associative memory.	A large number of parameters are needed, the learning process cannot be observed, the output results are difficult to explain, and the learning time is too long.
deep learning methods	FFNN	MLP: [9], [23], [76], [13], [14]; BPNN: [78], [48], [67]; SAE: [48], [67]	The high degree of parallel processing, high degree of nonlinear global effect, and very strong self-adaptive and self-learning functions.	It is very difficult to select the number of hidden nodes in the network. The stopping threshold, learning rate, and momentum constant need to adopt the "trial-and-error" method, which is extremely time-consuming and easy to fall into local extremum.
	RNN	vanilla RNN: [50], [40], [41], [82], [17], [78], [27], [67], [62], [21], [31] [83], [84], [85]; LSTM: [91], [12], [14], [52], [2], [82], [17], [78], [27], [92], [48], [67], [24], [31], [83], [93], [94], [84], [95], [96], [97]; GRU: [40], [41] [48], [67], [29], [84]	Faced with time series-sensitive problems and tasks, it is usually more appropriate and has a certain memory effect. The variant solves the problem of the vanishing gradient caused by a gradual reduction in gradient backpropagation.	There are disadvantages in parallel processing, and the gradient problem has been solved to some extent in its variants, but it is still not enough, and the calculation takes time.
	CNN	[10], [12], [24], [2]	Sharing convolution kernel, there is no pressure on high-dimensional data processing, and it is necessary to manually select features and train the weights, so that the feature classification effect is good.	Insufficient biological support, no memory function. Gradient descent algorithm can easily make the training result converge to the local minimum instead of the global minimum.
graph neural networks models	spectral models	[29], [2], [48], [42], [110]	Compared with the method of decomposing classes, GNN can learn on large graphs. It naturally fuses the attribute information of graphs for learning.	The calculation cost of the model increases sharply with the increase of the graph size, and can only work on undirected graphs. Given a fixed graph, it is difficult to add new nodes.
	spatial models	[40], [41]	It can be implemented in some nodes and node sampling can be introduced to improve efficiency. Any kind of diagram will do. Local convolution is performed at each node and weights can be shared between different locations and structures	In the calculation process, the mapping in the formula must be compression mapping, which seriously limits the modeling ability. Since many iterations are required between gradient descent steps, the computational cost of GNN is expensive.

TABLE V

SUMMARY OF THE SURVEYED LITERATURE

Reference	Year	Generation	Model	Prediction Problems	Predition Horizons	Region	Datasets	Other Factors
Caliskan *et al.* [11]	2007	1	Markov model	parking availability	15 minutes	off-street	-	-
Chen *et al.* [47]	2013	1	ARMA	bike parking availability	5 minutes, 1 hour, and 24 hours	on-street	Dublinbikes	weather; weekends
Rajabioun *et al.* [16]	2015	1	AR	parking availability	20 minutes	on-street; off-street	SFpark	-
Monteiro *et al.* [6]	2018	1	Poisson processes	parking availability	0 - 60 minutes	on-street	-	-
Xiao *et al.* [26]	2018	1	Markov model	parking availability	12 minutes	on-street	-	-
Liu *et al.* [15]	2018	1	AR	parking availability	-	on-street	Melbourne data	-
Bock *et al.* [1]	2019	1	Kalman filters	taxi parking availability	-	on-street	SFpark	-
Ismail *et al.* [13]	2021	1	MLP	parking availability	1, 5, 15, and 30 minutes	-	KLCC data	-
Schneble *et al.* [118]	2021	1	Markov model	parking availability	10 and 30 minutes	on-street	Melbourne data	-
Richter *et al.* [119]	2014	2	clustering	parking availability	15 minutes	on-street	SFpark	-
Zheng *et al.* [7]	2015	2	ANN	parking availability	300 minutes	on-street	SFpark; Melbourne data	-
Leu *et al.* [55]	2015	2	SVR	bike parking availability	30 minutes	on-street	Ubike	weather
Ziat *et al.* [20]	2016	2	ANN	parking availability	-	off-street	Lyon data	-
Stolfi *et al.* [22]	2017	2	KNN	parking availability	-	off-street	Birmingham data	-
Lu *et al.* [120]	2018	2	DT	parking availability	2 - 5 hours	on-street	SFpark	-
Feng *et al.* [28]	2019	2	DT	parking availability; parking behavior	-	off-street	-	weather; holidays
Kim *et al.* [18]	2019	2	RF	parking availability	10, 20, and 30 minutes	on-street	Seattle data	-
Muntean *et al.* [121]	2019	2	KNN	parking availability	-	off-street	Birmingham data	-
Chawathe *et al.* [122]	2019	2	SVR	parking availability	-	off-street	-	-
Awan *et al.* [57]	2020	2	KNN	parking availability	10 minutes	on-street	WISE-IoT	-
Abdellaoui *et al.* [59]	2020	2	RF	parking availability	-	off-street	Birmingham data	-
Zhao *et al.* [123]	2020	2	SVR	parking availability	-	off-street	-	-
He *et al.* [25]	2020	2	ANN	parking availability	30, 60, and 90 minutes	on-street	Seattle data	weather
De *et al.* [30]	2021	2	SVR	parking behavior	-	-	-	-
Parmar *et al.* [39]	2021	2	ANN	parking availability	-	on-street	-	-
Eider *et al.* [56]	2021	2	SVR	parking availability	-	-	-	-
Pozo *et al.* [124]	2021	2	DT	parking availability	10, 30, 45, and 60 minutes	on-street	-	congestion
Sonntag *et al.* [58]	2021	2	SVR	parking availability	-	on-street	Seattle data	-
Chen *et al.* [23]	2014	3	FFNN	parking availability	15 minutes	on-street	SFpark	-
Ji *et al.* [72]	2015	3	BPNN	parking availability	-	off-street	-	weekends
Vlahogianni *et al.* [76]	2016	3	MLP	parking availability; the probability of a free space to continue being free	1 - 30 minutes	on-street	-	weekends
Badii *et al.* [50]	2018	3	RNN	parking availability		off-street	-	weather
Shao *et al.* [14]	2018	3	LSTM	parking availability	1, 5, 15, and 30 minutes	on-street	Melbourne data	weekends
Rong *et al.* [52]	2018	3	LSTM	parking availability	-	off-street	-	meteorological data; POI; map mobility trace data; navigation data
Li *et al.* [78]	2018	3	LSTM	parking availability	2 hours	off-street	-	-
Camero *et al.* [21]	2018	3	RNN	parking availability	-	off-street	Birmingham data	-
Fang *et al.* [125]	2018	3	FFNN	parking availability	-	-	-	missing data

TABLE VI
(Continued.) SUMMARY OF THE SURVEYED LITERATURE

Reference	Year	Generation	Model	Prediction Problems	Predition Horizons	Region	Datasets	Other Factors
Sun *et al.* [96]	2018	3	LSTM	parking availability	-	-	-	-
Qiu *et al.* [126]	2018	3	GRU	parking availability	-	off-street	-	-
Ghosal *et al.* [12]	2019	3	CNN + LSTM	parking availability	-	on-street	SFpark	weather
Wang *et al.* [93]	2019	3	LSTM	parking availability	-	off-street	Birmingham data	-
Chen *et al.* [94]	2019	3	LSTM	parking availability	-	-	-	transfer solutions
Lv *et al.* [85]	2019	3	RNN	parking availability	-	-	-	-
Jomaa *et al.* [95]	2019	3	LSTM	parking availability	5, 15, 30, 60, and 120 minutes	on-street; off-street	BANES	-
Mishra *et al.* [91]	2020	3	LSTM	parking availability	-	on-street	-	-
Kamilaris *et al.* [9]	2020	3	MLP	parking availability; parking behavior	up to 90 minutes	off-street	Arnhem data; NDW data	weather
Provoost *et al.* [10]	2020	3	CNN	parking availability	up to 60 minutes	off-street	Arnhem data; NDW data	weather; holidays; weekends
Zhang *et al.* [27]	2020	3	LSTM	parking availability	1, 2, and 3 hours	off-street	PewLSTM	weather; holidays
Anagnostopoulos *et al.* [92]	2020	3	LSTM	parking availability	-	-	Aarhus data	-
Fan *et al.* [67]	2020	3	LSTM	parking availability	5, 15, 30, 45, and 60 minutes	off-street	Santa Monica data	-
Ttekouabou *et al.* [62]	2020	3	RNN	parking availability	-	off-street	Birmingham data	-
Ali *et al.* [127]	2020	3	LSTM	parking availability	-	off-street	Birmingham data	-
Arjona *et al.* [84]	2020	3	GRU	parking availability	-	on-street	-	weather; holidays
Zhang *et al.* [82]	2021	3	LSTM	Parking Behavior	1, 2, and 3 hours	off-street	PewLSTM	weather; holidays
Jose *et al.* [17]	2021	3	LSTM	parking availability	-	off-street	Birmingham data	weekends
Barraco *et al.* [24]	2021	3	CNN + LSTM	parking availability	1 hour and 8 hours	off-street	Birmingham data	-
Ling *et al.* [31]	2021	3	LSTM	parking behavior	-	on-street	-	weather; holidays
Yang *et al.* [83]	2021	3	RNN	truck parking availability	2, 4, 8, and 16 minutes	off-street	-	weather
Rahaman *et al.* [117]	2021	3	LSTM	parking availability	15 and 30 minutes	on-street	Melbourne data	-
Yang *et al.* [42]	2019	4	GNN	parking availability	30 minutes	on-street	-	weather, traffic speed
Zhang *et al.* [40]	2020	4	GNN	parking availability	15, 30, and 45 minutes	off-street	-	-
Zhang *et al.* [41]	2020	4	GNN	parking availability	15, 30, and 45 minutes	off-street	-	-
Xiao *et al.* [2]	2021	4	GNN	parking availability	15, 30, and 60 minutes	on-street	Melbourne data	parking duration distribution
Zhao *et al.* [48]	2021	4	GNN	parking availability	10 minutes	on-street	Shenzhen data	weather; holidays
Xiao *et al.* [110]	2021	4	GNN	parking availability	15, 30, and 45 minutes	on-street	Melbourne data	-
Gong *et al.* [29]	2021	4	GNN	parking behavior	-	off-street	-	COVID-19; weekends

where Occupancy(t) is the number of occupied parking bays at time t and Capacity is the number of all parking bays. It is easy to find out that the sum of AVR and the POR at time t is 1, therefore, we consider the two to be equivalent to some degree when measuring parking availability.

Although different studies have taken different approaches to parking occupancy prediction, the purpose of their predictions is essentially the same. All studies try to accurately predict parking occupancy to provide appropriate guidance to decision-makers or drivers. Besides, when some information is known, e.g., different approaches can be converted to each other. The main difference among these forecasting approaches is how they help the drivers to judge if a parking area is available. To better show the parking occupancy of an area, we suggest using POR and AVR for the larger parking areas and using the number of free parking slots for smaller parking areas if possible. As for the car parking count, since drivers may be not familiar with the parking area, it may mislead the drivers sometimes.

B. Parking Behavior Prediction

Parking behavior respects a series of actions and manners pertaining to parking [28]. These actions and manners may include the arrival and departure of vehicles. The purposes of car parking are also included such as shopping, working, etc. These behaviors are impacted by many factors, such as the weather, holidays, time of the day, etc. Feng et al. [28] explored different methods including LR, DT, and RF for parking behavior prediction. The weather and holiday features were also considered in their work. However, they only considered the arrivals and did not consider the departures of vehicles and their models failed to capture long-term temporal patterns. To this end, Zhang et al. [27] developed a periodic weather-aware LSTM model based on historical data, weather, environments, weekdays, and events to forecast parking behavior. This model was more applicable since it considered the departures of vehicles. This model could also achieve higher prediction accuracy since it could capture long-term dependencies.

Parking behavior prediction can provide information on the arrivals and departures of vehicles, which is also an important role of PGIS. Several works have explored this problem and confirmed the importance of parking behavior prediction in finding an available parking space [8], [9], [27], [28], [29], [30], [31]. When modeling time series prediction problems, Parking occupancy prediction and parking behavior prediction are similar and be modeled using the approaches listed in Section III. However, there are still some differences. Compared with parking occupancy prediction, parking behavior prediction focuses more on how drivers behave when looking for parking spaces. The results of parking behaviors can include some potential information. For instance, when the departure rate is high, though the POR of a parking area is 100%, it is worth waiting until a car leave. It is possible to calculate the growth of parking space usage for a parking lot by subtracting the number of arrivals from the number of departures, which is especially helpful for smart parking systems. Besides, parking behavior prediction often uses both arrival and departure datasets. Though we do not find there is research using the relationships between arrivals and departures, we think it is a potential research direction to dig out the relationships using some methods worth trying. e.g., attention mechanism.

C. Parking Demand Prediction

This problem is similar to the parking prediction problem, nevertheless it tends to be more of a macro analysis of overall parking needs. In addition to the traffic congestion and finding vacant parking spaces, the parking demand may have an effect on the parking price, i.e., high parking demand tends to result in a higher parking price [44], therefore, it is also meaningful to study the problem of parking demand prediction. Zheng et al. [129] created a parking demand prediction model based on the distribution of parking arrival and departure patterns, proposed a parking trend-based method for determining the prediction interval, and adopted utilizing curve fitting and undetermined coefficient methods to calibrate the parameters and predict the parking demand. Liu et al. [130] employed the CNN and LSTM models to capture spatial and temporal dependencies, respectively. Compared with other baselines, their model achieved the highest performance. The above researches and some other related researches [131], [132] show that accurate parking demand prediction provides technical support for parking planning, which is a direction worth studying in depth.

D. Factors of Parking Prediction

Parking prediction is generally a time series prediction problem [14], the prediction results of which may be influenced by many factors, thus we add the factors that are worth considering here.

1) On-Street or Off-Street: Off-street facilities (such as parking garages) are usually sparsely scattered in cities, so drivers are required to be more proactive in finding suitable parking places [133]. On-street parking is more widely distributed and easier to find. Furthermore, on-street parking is a city-managed public resource, whereas off-street parking is privately owned. In addition, on-street parking bays are distributed similarly to strips, while off-street parking bays are often scattered in a grid pattern. Thus, on-street parking and off-street parking should be considered separately.

2) Temporal Pattern: Temporal patterns, such as the time of the day, the day of the week, and holidays, are other factors that may influence parking prediction. For example, the POR is high at noon and low in the morning and evening [14]. On weekdays, cars are more likely to be parked in the working places while on weekends and holidays, people are more likely to park near their homes. Some studies [76], [134] also show that the distribution of duration changes with the temporal patterns like weekdays and weekends, morning and evenings., therefore, temporal patterns are important points that should not be easily ignored.

3) Other Factors: There are also other factors that can be considered. For example, the weather conditions can greatly influence parking behaviors [10], [50], [82]. Xiao et al. [2] considered the effect of parking duration distribution on POR. Gong et al. [29] took coronavirus disease 2019 (COVID-19) into account and revealed the impact of the pandemic on parking behavior. Provoost et al. [10] proved that the sensor information can be processed and help describe and predict the spatial and temporal situations.

E. Other Problems

Apart from the three kinds of parking prediction problems listed above, there are some interesting problems of parking prediction, which are listed as follows.

- **The probability of a free space to continue being free**. Vlahogianni et al. [76] applied survival analysis to forecast the likelihood that a parking space will be available in the following time periods based on the assumption that the duration of a free parking spot follows a Weibull distribution. Caicedo et al. [8] carried out a similar work. They estimated future departures according to the arrivals.
- **In-flux and out-flux**. Kamilaris et al. [9] developed a real-time parking area state model. In addition to the conventional parking occupancy rate, they also studied the in-flux and out-flux of parking, which is a more difficult

problem, since it was not based exclusively on occupancy rate.

- **Parking duration time estimation**. This problem is similar to the departure estimation since the arrival time for parking is generally known in advance. Through studying the parking duration, Shao et al. [14] found that the parking duration generally followed the exponential distribution. This property was later exploited by Xiao et al. to design an attention mechanism to improve the prediction performance [2].
- **Successful parking probability**. Wu et al. [135] proposed a successful parking probability prediction algorithm that may be used to enhance parking spot prediction after a short or long drive. Rizvi et al. [136] created a smart parking suggestion system for smart cities that was agent-oriented. In addition to the advantages of the model proposed by Wu et al. [135]. Their system took average waiting time, average parking time, and average walking time into consideration, which is more useful for determining the best parking space.
- **Bike parking availability**. Compared with the conventional parking availability prediction, this problem uses the number of bicycles as a prediction target. Leu et al. [55] used an SVR model to predict the number of bicycles in Ubike stations. A proportional selection method was also adopted for increasing the prediction performance. Chen et al. [47] then created a generalized additive model to calculate the distribution of the time spent waiting for the next available bike and the pace at which bikes arrive. These bike parking availability prediction methods can increase the utilization rate of non-motorized vehicles, thereby reducing air pollution.
- **Truck parking availability**. The prediction target converts from the general vehicles' parking availability to the truck parking availability [83].
- **Taxi parking availability**. The prediction target converts from the general vehicles' parking availability to the taxi parking availability [1].

All the parking prediction problems as well as the considered factors surveyed by this article are summarized in Table VI.

V. Evaluation Metrics, Open Datasets, and Open-Source Codes

In this section, we mainly introduce some additional information that is useful for parking prediction. The information includes evaluation metrics, open datasets, and open-source codes.

A. Evaluation Metrics

To evaluate the accuracy of the prediction, it is important to choose appropriate evaluation metrics. All the evaluation metrics used for parking prediction are listed as follows.

- **Precision**. Precision is defined as the proportion of all samples labeled as positive that are genuinely positive [57], which is expressed as

$$\text{Precision} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}}.$$

- **Recall**. Recall is defined as the fraction of all the positive samples, which is expressed as

$$\text{Recall} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}}.$$

- **F1-Score**. F1-Score is defined as the harmonic mean of precision and recall, which is expressed as

$$\text{F1-Score} = \frac{2 \times (\text{Recall} \times \text{Precision})}{\text{Recall} + \text{Precision}}.$$

- **Accuracy**. Accuracy is konwn as the ratio of the number of correctly predicted samples to the total number of samples, which is expressed as

$$\text{Accuracy} = \frac{\text{CorrectPredictions}}{\text{TotalSamples}}.$$

- **Mean Absolute Error**. Mean absolute error (MAE) is defined as

$$\text{MAE} = \frac{1}{N} \times \sum_{n=1}^{N} |\hat{y}_n - y_n|,$$

where y_n is the true value, $\hat{y}_n$ is the predicted value, and N is the number of samples.

- **Mean Absolute Percentage Error**. Mean absolute percentage error (MAPE) is defined as

$$\text{MAPE} = \frac{1}{N} \times \sum_{n=1}^{N} \left| \frac{\hat{y}_n - y_n}{y_n} \right| \times 100\%,$$

where y_n, $\hat{y}_n$, and N are defined in the same way as in MAE. However, this method can not deal with the situation when the truth values contain 0. This problem can be solved by adopting the weighted mean absolute percentage error (WMAPE) [137].

- **Mean Square Error**. Mean square error (MSE) is defined as

$$\text{MSE} = \frac{1}{N} \times \sum_{n=1}^{N} (\hat{y}_n - y_n)^2$$

where y_n, $\hat{y}_n$, and N are defined in the same way as in MAE.

- **Root Mean Square Error**. Root mean square error (RMSE) is defined as

$$\text{RMSE} = \sqrt{\frac{\sum_{n=1}^{N} (\hat{y}_n - y_n)^2}{N}},$$

where y_n, $\hat{y}_n$, and N are defined in the same way as in MAE.

- **Sum of Square Error**. Sum of square error (SSE) is defined as

$$\text{SSE} = \sum_{n=1}^{N} (\hat{y}_n - y_n)^2$$

where y_n, $\hat{y}_n$, and N are defined in the same way as in MAE.

- **Relative Absolute Error**. Relative absolute error (RAE) is defined as

$$\mathrm{RAE} = \frac{\sum_{n=1}^{N} |\hat{y}_n - y_n|}{\sum_{n=1}^{N} |\hat{y}_n - \bar{y}|},$$

where y_n, $\hat{y}_n$, and N are defined in the same way as in MAE. $\bar{y}$ is calculated by $\bar{y} = \frac{1}{N}\sum_{n=1}^{N} y_n$.

- **Root Relative Squared Error**. Root relative squared error (RRSE) is defines as

$$\mathrm{RRSE} = \sqrt{\frac{\sum_{n=1}^{N} (\hat{y}_n - y_n)^2}{\sum_{n=1}^{N} (\hat{y}_n - \bar{y})^2}},$$

where y_n, $\hat{y}_n$, and N are defined in the same way as in MAE. $\bar{y}$ is calculated by $\bar{y} = \frac{1}{N}\sum_{n=1}^{N} y_n$.

- **False Positive Rate**. False positive rate (FPR) is first adopted by [6], which is defined as

$$\mathrm{FPR} = \frac{\sum_{b=1}^{B}\sum_{n=1}^{N} \mathbb{I}\left[r_b(n) = 1 \cap a_b(n) = 0\right]}{\sum_{b=1}^{B}\sum_{n=1}^{N} \mathbb{I}\left[a_b(n) = 0\right]},$$

where B is the number of blocks, $r_b(t)$ is the parking recommendation of the n-th parking bay of the b-th block, $a_b(t)$ is the parking availability of the n-th parking bay of the b-th block, and $\mathbb{I}[x]$ equals 1 if x is true and 0 if x is false.

- **False Negative Rate**. False negative rate (FNR) is defined as

$$\mathrm{FNR} = \frac{\sum_{b=1}^{B}\sum_{n=1}^{N} \mathbb{I}\left[r_b(n) = 0 \cap a_b(n) > 0\right]}{\sum_{b=1}^{B}\sum_{n=1}^{N} \mathbb{I}\left[a_b(n) > 0\right]},$$

where the meaning of the symbols is the same as that of FPR.

- **The Coefficient of Determination**. The coefficient of determination (R^2) is defined as

$$\mathrm{R}^2 = 1 - \sum_{n=1}^{N} \frac{(\hat{y}_n - y_n)^2}{(\hat{y}_n - \bar{y})^2},$$

where the meaning of the symbols is the same as that of RRSE.

- **Mean Absolute Scaled Error**. Mean absolute scaled error (MASE) was first proposed by [138] ad is defined as

$$\mathrm{MASE} = \mathrm{mean}\left(\frac{|\hat{y}_t - y_t|}{\frac{1}{N-1}\sum_{n=2}^{N} |y_t - y_{t-1}|}\right),$$

where y_t is the real value at time t and $\hat{y}_t$ is the predicted value.

- **Mean Relative Deviation**. Mean relative deviation (MRD) was proposed by [11] for POR prediction, which is defined as

$$\mathrm{MRD} = \sum_{n=1}^{N} \pi_n \cdot \frac{|\hat{O}_n - O_n|}{N},$$

where π_n denotes the probability assigned for occupancy state $\hat{O}_n$ and O_n is the true POR.

TABLE VII

THE EVALUATION METRICS WITH RELATED STUDIES

Metrics	Relevant Studies
Precision	[57], [52], [95], [30]
Recall	[57], [52], [95], [30]
F1-Score	[57], [95], [30], [58]
Accuracy	[57], [82], [27], [92], [24], [31], [125]
MAE	[76], [6], [7], [9], [10], [12], [13], [14], [40], [41], [2], [17], [55], [18], [48], [67], [62], [21], [59], [127], [29], [30], [31], [123], [83], [93], [42], [84], [25], [85], [97], [126], [117], [110]
MAPE	[76], [12], [14], [16], [2], [28], [48], [67], [23], [29], [123], [83], [42], [97], [126], [110]
MSE	[7], [9], [10], [72], [15], [78], [22], [23], [127], [94], [84], [97], [126], [96]
RMSE	[76], [10], [13], [40], [41], [2], [82], [17], [78], [27], [55], [28], [47], [1], [48], [67], [20], [62], [59], [127], [29], [31], [56], [123], [83], [93], [84], [85], [122], [126], [120], [117], [110]
SSE	[121], [126]
RAE	[55]
RRSE	[76], [14], [55]
FPR	[6]
FNR	[6]
R^2	[7], [10], [26], [28], [62], [59], [129]
MASE	[50], [9]
MRD	[11]
CRPS	[48]

- **Continuous Ranked Probability Score**. Continuous ranked probability score (CRPS) was proposed by [139] and was adopted by [48] for probabilistic prediction of parking duration. CRPS is defined as

$$\mathrm{CRPS} = \frac{1}{|B_m|}\sum_{i=1}^{|B_m|}\left(\frac{1}{R_i}\sum_{r=1}^{R_i}\int_0^h \left(F_i(x) - H\left(x - \xi_{i,r}\right)\right)^2 dx\right),$$

where R_i is the number of parking events of the i-th block, h is the upper bound of parking duration, $F_i(x)$ is the empirical cumulative distribution function of the i-th block, $\xi_{i,r}$ is the true observation of each parking event, and $H(\cdot)$ is the Heaviside step function [48].

The evaluation metrics with related studies are summarized in Table VII. The MAE and RMSE metrics are the most two popular metrics since they are easy to calculate and compare the performance. Besides, there are also no obvious restrictions on the data used for the calculations.

B. Datasets

The open datasets are summarized in Table VIII and the details of these datasets are listed as follows. For presentation purposes, we have named some of the datasets.

- **SFpark**. This subset uses over 8,000 sensors on a total of 420 parking segments in San Francisco. This dataset spans the dates of June 13th and July 24th, 2013.

TABLE VIII
OPEN DATASETS FOR PARKING PREDICTIONS

Datasets	Related Studies	City or Country	Number of Sensors	Parking Regions	Time Range	Recording Objects	Temporal Granularity	Contents	Link
SFpark	[7], [12], [16], [1], [23]	San Francisco	over 8,000	on-street	2013-06-13 to 2013-07-24	420 parking segments	1 minute	parking events	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/YLWCSU
Melbourne data	[7], [14], [2], [118]	Melbourne	5,500	on-street	2011-01 to 2020-05	24,715 parking bays	1 second	parking events	http://www.melbourne.vic.gov.au/
Arnhem data	[9]	Arnhem		off-street	2017-08 to 2019-04	125 parking bays	1 second	parking events	https://parkeerdata.nl/opendata/arnhem/http://opendata.arnhem.nl/datasets/Arnhem::dynamische-parkeerdata
KLCC data	[13]	Kuala Lumpur		off-street	2016-06-01 to 2017-11-15	5,500 parking spots	1 second	parking status	https://www.kaggle.com/mypapit/klccparking/version/1
PewLSTM	[27]	China		off-street	2017-10-16 to 2019-08-30	10 parking lots	1 minute	parking status	https://github.com/NingxuanFeng/PewLSTM
Dublinbikes	[47]	Dublin		on-street	2012-02-26 to 2012-05-31	44 stations	5 minutes	parking status	http://www.dublinbikes.ie/
Seattle data	[18], [58]	Seattle		on-street	2019-04-25 to 2019-05-26	442 parking bays	1 second	parking status	https://data.seattle.gov/Transportation/Paid-Parking-Occupancy-Last-30-Days/rke9-rsvs
Aarhus data	[92], [25]	Aarhus		off-street	2014-05-22 to 2014-11-04	8 parking lots	1 minute	parking status	http://iot.ee.surrey.ac.uk:8080/datasets.html#parking
Shenzhen data	[48]	Shenzhen	1,735	on-street	2018-09-01 to 2019-01-01	76 street blocks	1 second	parking events	https://opendata.sz.gov.cn/
Birmingham data	[17], [22], [62], [21], [24], [59], [121], [93]	Birmingham		off-street	2016-10-04 to 2016-12-19	2,937 parking bays	1 second	parking status	https://data.birmingham.gov.uk/dataset/birmingham-parking
BANES	[95]	London		off-street	2016-08-31 to 2020-02-10	8 parking lots	1 second	parking status	https://github.com/BathHacked/documentation/wiki/Bath-Car-Park-Dat

- **Melbourne data**. This dataset uses 5,500 on-street parking sensors to record the arrivals and departures of vehicles. This dataset spans the months of January, 2011 and May, 2020.
- **Arnhem data**. This dataset contains a real-time data feed of the POR, which is updated every 11 minutes. This dataset spans the months of August 2017 and April 2019.
- **KLCC data**. This dataset records 47,603 POR data from 5,500 parking spaces. The time period for this dataset is from June, 2016 to November, 2017.
- **PewLSTM**. This datasets are composed of parking records from 452,480 parking bays of 10 parking lots in China, including shopping malls, hotels, communities, etc. The dataset ranges from October 16th, 2017 to August 30th, 2019.
- **Dublinbikes**. This dataset has 44 bike stations. Each station records 26,458 observations. The time period for this dataset is from February 26th, 2012 to May 31st, 2012 (Data of March 12th and 13th are missing).
- **Seattle data**. Parking transactions from 442 on-street parking lots are included in this dataset. The dataset ranges from April 25th to May 26th, 2019.
- **Aarhus data**. This records 55,264 parking data points from 8 parking lots. The time period for this dataset is from May 22nd, 2014 to November 4th, 2014.
- **Shenzhen data**. This dataset records 624,464 parking event by 1735 sensors over 76 street blocks in Nanshan District of Shenzhen, China. The time period for this dataset is from Sep 1st, 2018 to Jan 1st, 2019.

TABLE IX
OPEN-SOURCE CODE FOR PARKING PREDICTION

References	Problem	Link
Zhang *et al.* [27]	Parking behavior prediction	https://github.com/NingxuanFeng/PewLSTM
Camero *et al.* [21]	Parking availability prediction	https://github.com/acamero/dlopt

- **Birmingham data**. This dataset records the POR of 29 parking lots. The time period for this dataset is from October 4th, 2016 to December 19th, 2016.
- **BANES**. This dataset includes POR of 8 different off-street parking houses in London. The time period for this dataset is from August 31st, 2016 to February 10th, 2020.

C. Open-Source Codes

Until now, we have found two open-source codes for parking prediction. These codes are summarized in Table IX.

VI. CHALLENGES AND DIRECTIONS FOR FUTURE RESEARCH

In this section, we present some challenges and directions for future research on parking prediction.

A. Challenges

There is great interest and effort in parking prediction solutions based on deep learning or graph neural networks, and the performance of results is good. However, in the real world, there are still many significant challenges. In this section, we will discuss the major challenges of parking forecasting in the real world.

1) Multi-Source Data: It has been proved that multiple data sources are beneficial to enhance prediction accuracy [140]. When studying the road congestion level (CL), explicit sensing systems (e.g., the vehicle system and its GPS data, as well as the ring sensor and its log data) have direct CL measurement values, which are accurate in measurement but low in permeability and space-time coverage. implicit sensing systems (e.g., cellular networks and their signaling data, and sensing systems for buildings and their occupancy data) have indirect CL measurement values, which have high space-time coverage but low measurement accuracy. The two methods have their own advantages and disadvantages. When studying the parking problem, most existing works have considered various features, such as spatiotemporal correlations, weather, holidays, etc., for parking prediction. However, the formats of the extrinsic data may be different and they have different non-linear correlations, which can not be fused with parking data directly [141]. Besides, different feature data have different impacts on parking prediction problems. Unreasonable importance assignments may reduce prediction performance. In addition, data quality is another challenge such as missing data caused by human factors, which are not recorded, omitted, or lost, sparse data due to improper investigation or natural limitation, false data caused by measurement error, and noise data deviating from expected values. Finally, a large amount of multi-source data will increase the computing complexity and the rational fusion and usage of this multi-Source data remain a great challenge.

2) Multiple Tasks: For ITS of public information services (PIS), parallel multiple prediction tasks for presentation to provide more information are necessary. People can make the right decisions by analyzing a wide range of predicted information. For supervised deep learning and machine learning, the basic process is to input data, and the model predicts the results according to the input data. In the training stage, the parameters of the model are corrected according to the difference (loss) between the prediction and the supervised signals, so that the model parameters can match the data distribution as much as possible. The framework of multi-task learning can put together several basic training processes, so the parking problem is suitable for this framework. Existing studies have conducted multiple prediction tasks. For example, Vlahogianni et al. [76] predicted both the availability of parking and the likelihood that a vacant spot would remain free in the following time periods. However, most models can only be employed for a single task at one time, which remains to be solved in future research studies.

3) Real-Time Services: Existing studies have proposed plenty of state-of-the-art models for parking prediction and these models have shown satisfactory prediction performance. However, when these models are used in real-world problems, there are some new challenges. On one hand, since the application needs to be updated online, how to mine, acquire, process, predict, and update data in real time is not easy. At present, the data access of city-level parking platforms still lacks the spontaneous behavior of the market, and it is basically realized by administrative means by governors' functional management departments. Even if some dynamic real-time data can be accessed, it can only temporarily solve the problem of access quantity, and it is difficult to guarantee the data quality and the sustainability of data channels in the later period. On the other hand, with the rapid urban development, the parking or traffic networks are always changing, New projects and reconstruction projects of urban municipal facilities and road construction will affect the parking lot, parking space, and the number of parking spaces at any time, which may reduce the prediction performance of existing methods. Equipment maintenance is not in place or information update is not timely, which will lead to inaccurate data.

B. Future Directions

We briefly discuss the future research direction of parking prediction and provide some suggestions. These suggestions are helpful for future research in the field of parking, in order to alleviate the limitations of existing literature and solve some challenges in the real world.

1) Prediction Methods: In this section, we focus on potential state-of-the-art prediction algorithms that can be used for parking prediction problems, including graph based methods, transfer learning, meta learning, multimodal learning, and transformer.

- **Graph Based Methods:** The parking networks are often expressed as non-Euclidean spaces, which are suitable to be modeled by a graph structure. The graph based models can extract the complex spatial dependencies and inner contextual features, which is beneficial for achieving higher prediction accuracy. However, as shown in Section III-D, there are few works that have been done on using GNN models for parking prediction problems. In traffic prediction problems, GNN models have been widely used recently [115], [142], [143], [144], [145]. These methods either combine GNN and RNN or GNN and CNN to capture spatio-temporal correlations. The on-street parking networks are similar to the traffic networks, both of which are based on the roads, streets, or avenues., therefore, it is feasible to use graph neural network models for parking prediction. It should be noted that when using GNN models to solve the parking forecasting problem, geographically close nodes may not be the most influential, and the underlying graphic information may be incorrect or untimely (the data lags behind the real data).

  Ye et al. [105] pointed out that the weight matrix is seen as the key to capturing spatial dependency. Existing studies for parking prediction only use the distance of nodes to construct the weight matrix. Dai et al. [146] showed that the correlation coefficients of nodes are more important than the spatial features. The correlation coefficients are used for constructing the weight matrix of graphs and can be calculated with function similarities [147] and temporal pattern similarity [148]. For the parking occupancy prediction tasks, these correlation coefficients have also been applied by some researchers to obtain more detailed contextual information [149]. Apart from the temporal pattern similarity and function similarities, we think the diffusion process is also a variable worth considering since it contains potential information on the arrivals and departures of the parking process.

Different types of vehicles travel at different distances and travel for different lengths of time. Existing studies for parking prediction only express the spatio-temporal characteristics on a single scale, and cannot extract global information and detailed information at the same time. Multi-scale learning can construct different time scales [150] and spatial scales [151], and improve the prediction performance in complex scenes.

Existing studies for parking prediction simply adopted a single graph, which does not fully reflect the variety of information on which the parking forecast problems rely. Multiple graphs [147], [148], [152] can combine various contextual information and multiple weight matrices to model the parking prediction problem, which can lead to better prediction results. In addition, hypergraphs [153] can also be used for parking prediction problems. The hypergraph is a generalization of the graph in which one edge can link any number of vertices and can inflect the inner connections of nodes connected to the same edge. For example, vehicles and bikes are two kinds of nodes that can be connected to construct a hypergraph. Researchers can dig out the inner relationships of the parking patterns of vehicles and bikes and predict the parking availability of vehicles and bikes.

- **Transfer Learning:** Transfer learning means that a pre-trained model is reused in another task, especially for those scenarios that lack sufficient data. In the real world, for reasons such as confidentiality or an insufficient number of sensors, in some cities, there are few parking data for training. Therefore, it is a solution to transfer the semantic dictionary learned from the source city with rich parking data to the target city, enrich the feature representation of the target city, and then predict the parking of the target city. In addition, it is also a solution to use the idea of the multi-source domain to transfer knowledge from multiple related but different domains, so as to obtain more comprehensive knowledge covering the target domain and reduce the risk of negative transfer. There have been some works based on transfer learning in the parking prediction problem. Chen et al. [94] proposed an optimized transfer strategy based on minimum mean square error and adopted the LSTM models for prediction. This model could find the most optimized transfer solution actively and directly and achieve high performance for transfer prediction. Actually, the parking network is essentially a graph structure. In order to capture the spatial information more effectively, it is more appropriate to express the road network as a graph. Previous studies have applied transfer learning with GNNs for traffic flow prediction [154]. We believe the pretrained results of a large graph can be applied to smaller graphs since the structure of the parking graph is similar to the traffic flow graph. At present, the transfer learning method of the parking prediction model based on graph data remains to be studied.
- **Meta Learning:** Meta learning means a model can learn how to learn. A model that adopts the meta learning mechanism can learn from a task and generalize it to another unseen task. This "learning-to-learn" paradigm has lots of advantages such as computing efficiency and the learning strategies improve both on lifetime and evolutionary timescales [155]. Meta learning has been applied in many fields for prediction, such as traffic prediction [156], water quality prediction [157], bankruptcy prediction [158], etc. Parking prediction is a complex spatio-temporal prediction problem, which is also well-suitable for adopting the meta-learning mechanism.

Meta-learning can be applied to different scenarios. Pretrain-Finetune is a common way in transfer learning. The problem with this approach is that it is often necessary to try different migration strategies to achieve the best results. For example, whether the parameters of a certain layer are migrated or randomly initialized; When the model structures of pretrained stage model and finetune stage model are inconsistent, the parameters of a certain layer of pretrained model should be migrated to which layer of finetune model; How strong is the migration of each layer, etc. Trying these strategies will take a lot of time. Meta-learning-based transfer learning method [159], using meta-learning to learn what kind of transfer strategy can achieve the best effect. In the Domain Adaptation scenario, the traditional idea is to align the feature representations of the source domain and target domain by using confrontational learning. However, this method does not consider the downstream task at all, and the generated representation may not perform well in the downstream task. Meta-learning can be used to generate the aligned representation related to the target task [160]. In graph learning scenarios, a common paradigm is to pre-train a large number of unlabeled data, and then finetune the labeled data on the downstream target task. The problem with this two-stage method is that the optimization objectives of the two stages are different, so the model obtained in pretrained stage may not be the most suitable for downstream tasks. The idea of meta-learning [161] is introduced in the pretrained stage of the graph, and the finetune process is simulated in the pretrain process of the graph, and the effect of finetune is used as the outer loop update model parameter by Model-Agnostic Meta-Learning (MAML). A common method in the Knowledge Distillation scenario is the teacher-student architecture. The teacher model (generally a larger model) is used to train the data, and then the student model is used to transfer the knowledge learned by the teacher model. Meta-learning [162] is used to combine the two processes of teacher and student, and its core idea is to make the pseudo label produced by the teacher behave in the student and reverse it to the teacher.

It can be seen that meta-learning is a tool, which can establish the relationship between two optimization objectives through the inner loop and outer loop. Meta-learning is often used to improve the original two-stage training model, make the two-stage training more unified, and build a bridge between the two-stage training through meta-learning.

- **Multimodal Learning:** Multimodal learning is the analysis of multi-source heterogeneous data. Multimodal data have three forms: multiple data types, different data structures, or the combination of information from different databases, which are closer to the form of human understanding of the world. Multimodal learning has been applied in many fields, such as Human activity recognition [163], [164], medical application [165], [166], autonomous system [167], [168]. Parking data are always highly nonlinear and non-stationary, and they are affected by different components in different conditions (such as weather), therefore, parking prediction is also suitable for multimodal learning.

  Parking prediction tasks based on multimodal learning can include the following aspects. Multi-source data classification: Single-modal problem only focuses on the analysis and processing of a specific kind of data. Compared with a single channel, multi-modal data is closer to the real form of parking information flow under the background of big data, which is comprehensive and complex. Multi-modal semantic computing: semantic analysis is a higher-level processing of parking data. Ideally, the computer can process the conceptual relationship and logical structure of different data in a specific scene, and then understand the high-level semantics implied in different data; Understanding these high-level semantics is the premise of effective reasoning decisions. Multi-modal information fusion: Multi-modal fusion requires comprehensive and effective screening and utilization of multi-source parking data to achieve the purpose of integrated perception and decision-making. Common information fusion methods include physical layer fusion, feature layer fusion, and decision layer fusion. Physical layer fusion refers to the fusion processing of collected data at the sensor level in the first stage of perception; Feature layer fusion refers to the fusion of information at the level of feature extraction and expression, such as adopting the same feature expression form for data collected by different sensors in the same parking lot, and then carrying out corresponding superposition calculation; Decision-level fusion refers to the fusion of the output results of perception models of different modes. This fusion method has good anti-interference performance and has relatively low requirements for the performance and types of sensors, but it has a large information loss.
- **Transformer:** Transformer [169] is a sequence-to-sequence model based on the self-Attention mechanism, which consists of an encoder and a decoder. For the parking problem, the spatial transformer can capture the real-time parking situation and the directionality of traffic flow by dynamically modeling the directional spatial dependence with the self-attention mechanism. Different spatial dependency patterns can be jointly modeled by a multi-head attention mechanism to consider various relationships with different factors (such as similarity, connectivity, and covariance). On the other hand, the temporal transformer is used to model the bidirectional time dependence across multiple time steps. Because parking is highly nonlinear and dynamic in time and space, real-time and accurate parking prediction, especially long-term prediction, is still a problem to be solved. Transformer can directly model the longer-distance dependency between input sequence units and model long sequences.

2) Parking Technologies: With the rapid development of IoT technology, more and more parking technologies are being applied. These new technologies are also essential for studying parking forecasts. This section focuses on the potential applications of these technologies to the parking forecasting problem in the future.

- **Parking Guidance and Information Systems:** Sensor technology has been fully developed in recent years. On-ground sensors help record parking events and thus the parking occupancy or behavior can be predicted. When a sufficient number of sensors are deployed, Parking guidance and information systems (PGIS) can transmit parking space occupancy information to drivers in real time via the Internet and direct the drivers to an available parking space. However, the cost of deployment and maintenance of sensors is an issue that should be considered. Besides, though PGIS with sensors can improve the possibility of finding available parking bays, it also changes the behaviors of drivers. Drivers become to compete for parking rather than searching for them [170]. In addition, these systems may be of little use to drivers with more local familiarity [43]. It is an interesting question about how to increase the parking prediction performance under the circumstances that the cost of sensors is limited. Another problem worth trying is how to balance the competitiveness of drivers and the average time to find a parking space using PGIS and how to obtain accurate prediction results based on this condition.
- **Reservation and Waiting Time:** With the development of wireless communication and Internet of Things technology, smart mobile apps have widely entered our lives. Drivers can reserve parking spots ahead of time using these apps [171]. The reserved parking spaces are always vacant until the corresponding vehicle arrives, but it is actually occupied, so the information recorded by the sensor may be not true. Besides, an intelligent parking reservation system can also provide information on parking rates, directions, and available payment options, which may change the behavior of drivers. Influenced by the environment, cost, or familiarity with parking spaces, some drivers may prefer to go directly to a parking space instead of making a reservation. It is an interesting problem to predict the behaviors of drivers that use reservation services. When the parking departure rate is high, some drivers are willing to wait, which can directly affect

the results of parking behavior and parking occupancy. Departure rate is relative to the parking time [14] and the whole environment can be modeled to a queueing system. Therefore, this situation of parking may be modeled and thus help increase the performance of parking perdition in the real world. In brief, parking prediction that takes into account reservations and waiting times is also an important future research direction.

- **Autonomous Driving:** In autonomous driving, the vehicle's sensors and computer systems may handle the parking tasks, which will allow parking tasks to be carried out more efficiently. Although the parking technology regarding autonomous driving is not yet mature, there are already some cases of parking using autonomous driving technology. For example, the automated valet parking developed by Bosch [172] can help drivers driver to park their cars with autonomous driving technology. When the driver leaves their vehicle in a predetermined area, they can use an application on their mobile devices, such as smartphones, to lead their autonomous vehicle to the parking lot. In the future, autonomous driving vehicles may no longer need to be assigned a parking space by humans but may automatically select a suitable parking space. Compared to human driving, autonomous driving has fewer human factors, which is more favorable for simulation analysis. We can use algorithms or theories such as reinforcement learning [173], routing algorithms [174], queuing theory [175], etc. to model driverless parking scenes and ensure efficient inflow and outflow of the parking system and record a large number of parking events based on long-term observations. Finally, the above scenario is further optimized by the predicted parking occupancy rate, parking waiting time, parking outflow rate, and other metrics. It is worth noting that the existing reinforcement learning methods for automatic parking mainly focus on the parking behavior of a single vehicle, such as parking space availability detection, motion planning, and path tracking [176], [177], [178], etc. There is less research on the situation where a large number of autonomous driving vehicles need to park at the same time, which may be a direction to try to break through in the future when there are more and more driverless vehicles. Smart routing algorithms provide parking guidance [179] and path finding [180] benefits. Yet, we have not found routing algorithms or parking prediction applications for scenarios that contain a large number of both autonomous and manual driven vehicles, which may be a potential research direction in the future.
- **Crowdsourcing Parking:** Crowdsourcing parking [181] refers to a process of using the collective efforts of a group of people to gather and share information about parking availability and related issues in a particular area. This can be done through various means, such as mobile apps, websites, or social media platforms. The idea behind crowdsourcing parking is to leverage the power of a large community to gather real-time information about parking availability and other related data, such as the cost of parking, parking restrictions, and more. This information can then be shared with other users who are looking for parking in the same area, allowing them to make more informed decisions about where to park.

  Parking forecasting problems may suffer from a lack of data, and crowdsourcing parking can solve this problem to some degree. For example, Nie et al. [182] created a CNN-based spatial-temporal classifier with LSTM to learn from the parking experiences uploaded by devices using crowdsourcing technology. The vehicles can be guided to suitable places based on the suggestions. For parking prediction tasks, since crowdsourced parking technology has access to the most recent data, the prediction models can be updated in real time, which can be of great help in predicting the true parking status. However, when the crowdsourcing system is first built, there may not be enough data for accurate parking prediction, i.e., the cold start problem. It may be a good idea to consider this problem in conjunction with a recommendation system [183]. The mobile app developers may select some experts to test the crowdsourcing parking system and upload enough information or recommend the most popular parking areas or parking spaces to solve the problem. After obtaining enough data, the developers can apply appropriate prediction algorithms for parking predictions and then the apps can guide drivers to suitable parking bays based on the prediction results. With real-time training methods [184], the apps can be stronger and increase their prediction performance gradually. In short, parking prediction with crowdsourcing parking systems is a promising research direction.

VII. Conclusion

In this paper, we present a comprehensive review of parking prediction methods in smart cities. We divide the models into four generations according to the advancement of the prediction models, including statistical methods, machine learning methods, deep learning methods, and graph neural network models. The existing research has been able to predict the parking state at the next moment very accurately through advanced computer technologies such as big data and deep learning. However, the inherent factors such as the complexity and chaos of the system limit the function of parking flow prediction to a small range of short-term predictions. We categorize the parking prediction problems, such as parking availability prediction, parking behavior prediction, parking demand prediction, etc. In addition, we give the evaluation metrics, open data, and open-source code of the surveyed literature. Finally, we discuss the potential challenges and directions for future research.

References

[1] F. Bock, S. Di Martino, and A. Origlia, "Smart parking: Using a crowd of taxis to sense on-street parking space availability," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 2, pp. 496–508, Feb. 2020.

[2] X. Xiao, Z. Jin, Y. Hui, Y. Xu, and W. Shao, "Hybrid spatial–temporal graph convolutional networks for on-street parking availability prediction," *Remote Sens.*, vol. 13, no. 16, p. 3338, Aug. 2021.

[3] A. Sakai, K. Mizuno, T. Sugimoto, and T. Okuda, "Parking guidance and information systems," in *Proc. Pacific Rim TransTech Conf. Vehicle Navigat. Inf. Syst. Conf. 6th Int. VNIS. Ride Future*, Aug. 1995, pp. 478–485.

[4] J. Zhu et al., "KST-GCN: A knowledge-driven spatial–temporal graph convolutional network for traffic forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 15055–15065, Sep. 2022.

[5] S. Wang, P. Zhao, B. Yu, W. Huang, and H. Liang, "Vehicle trajectory prediction by knowledge-driven LSTM network in urban environments," *J. Adv. Transp.*, vol. 2020, pp. 1–20, Nov. 2020.

[6] F. V. Monteiro and P. Ioannou, "On-street parking prediction using real-time data," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2478–2483.

[7] Y. Zheng, S. Rajasegarar, and C. Leckie, "Parking availability prediction for sensor-enabled car parks in smart cities," in *Proc. IEEE 10th Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process. (ISSNIP)*, Apr. 2015, pp. 1–6.

[8] F. Caicedo, C. Blazquez, and P. Miranda, "Prediction of parking space availability in real time," *Exp. Syst. Appl.*, vol. 39, no. 8, pp. 7281–7290, Jun. 2012.

[9] A. Kamilaris et al., "Short term prediction of parking area states using real time data and machine learning techniques," in *Proc. Transp. Res. Board 99th Annu. Meeting (TRB)*, 2020, pp. 1–11.

[10] J. C. Provoost, A. Kamilaris, L. J. J. Wismans, S. J. Van Der Drift, and M. Van Keulen, "Predicting parking occupancy via machine learning in the web of things," *Internet Things*, vol. 12, Dec. 2020, Art. no. 100301.

[11] M. Caliskan, A. Barthels, B. Scheuermann, and M. Mauve, "Predicting parking lot occupancy in vehicular ad hoc networks," in *Proc. IEEE 65th Veh. Technol. Conf. VTC-Spring*, Apr. 2007, pp. 277–281.

[12] S. S. Ghosal, A. Bani, A. Amrouss, and I. El Hallaoui, "A deep learning approach to predict parking occupancy using cluster augmented learning method," in *Proc. Int. Conf. Data Mining Workshops (ICDMW)*, Nov. 2019, pp. 581–586.

[13] M. Ismail, T. Razak, R. Gining, S. Fauzi, and A. Abdul-Aziz, "Predicting vehicle parking space availability using multilayer perceptron neural network," in *Proc. IOP Conf. Mater. Sci. Eng.*, vol. 1176, no. 1. Bristol, U.K.: IOP Publishing, 2021, Art. no. 012035.

[14] W. Shao, Y. Zhang, B. Guo, K. Qin, J. Chan, and F. D. Salim, "Parking availability prediction with long short term memory model," in *Proc. Int. Conf. Green, Pervas., Cloud Comput.* Cham, Switzerland: Springer, 2018, pp. 124–137.

[15] K. S. Liu, J. Gao, X. Wu, and S. Lin, "On-street parking guidance with real-time sensing data for smart cities," in *Proc. 15th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON)*, Jun. 2018, pp. 1–9.

[16] T. Rajabioun and P. A. Ioannou, "On-street and off-street parking availability prediction using multivariate spatiotemporal models," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2913–2924, Oct. 2015.

[17] A. Jose and V. Vidya, "A stacked long short-term memory neural networks for parking occupancy rate prediction," in *Proc. 10th IEEE Int. Conf. Commun. Syst. Netw. Technol. (CSNT)*, Jun. 2021, pp. 522–525.

[18] K. Kim and N. Koshizuka, "Data-driven parking decisions: Proposal of parking availability prediction model," in *Proc. IEEE 16th Int. Conf. Smart Cities, Improving Quality Life Using ICT IoT AI (HONET-ICT)*, Oct. 2019, pp. 161–165.

[19] S. Arif, S. Olariu, J. Wang, G. Yan, W. Yang, and I. Khalil, "Datacenter at the airport: Reasoning about time-dependent parking lot occupancy," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 11, pp. 2067–2080, Nov. 2012.

[20] A. Ziat, B. Leroy, N. Baskiotis, and L. Denoyer, "Joint prediction of road-traffic and parking occupancy over a city with representation learning," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 725–730.

[21] A. Camero, J. Toutouh, D. H. Stolfi, and E. Alba, "Evolutionary deep learning for car park occupancy prediction in smart cities," in *Proc. Int. Conf. Learn. Intell. Optim.* Cham, Switzerland: Springer, 2018, pp. 386–401.

[22] D. H. Stolfi, E. Alba, and X. Yao, "Predicting car park occupancy rates in smart cities," in *Proc. Int. Conf. Smart Cities*. Cham, Switzerland: Springer, 2017, pp. 107–117.

[23] X. Chen, "Parking occupancy prediction and pattern analysis," Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, Tech. Rep., CS229-2014, 2014.

[24] M. Barraco, N. Bicocchi, M. Mamei, and F. Zambonelli, "Forecasting parking lots availability: Analysis from a real-world deployment," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops Other Affiliated Events (PerCom Workshops)*, Mar. 2021, pp. 299–304.

[25] J. He and J. Bai, "Prediction technology for parking occupancy based on multi-dimensional spatial–temporal causality and ann algorithm," in *Proc. Int. Conf. Green, Pervas., Cloud Comput.* Cham, Switzerland: Springer, 2020, pp. 244–256.

[26] J. Xiao, Y. Lou, and J. Frisby, "How likely am i to find parking?—A practical model-based framework for predicting parking availability," *Transp. Res. B, Methodol.*, vol. 112, pp. 19–39, Jun. 2018.

[27] F. Zhang et al., "PewLSTM: Periodic LSTM with weather-aware gating mechanism for parking behavior prediction," in *Proc. 29th Int. Joint Conf. Artif. Intell.*, Jul. 2020, pp. 4424–4430.

[28] N. Feng, F. Zhang, J. Lin, J. Zhai, and X. Du, "Statistical analysis and prediction of parking behavior," in *Proc. IFIP Int. Conf. Netw. Parallel Comput.* Springer, 2019, pp. 93–104.

[29] S. Gong, X. Mo, R. Cao, Y. Liu, W. Tu, and R. Bai, "Spatio-temporal parking behaviour forecasting and analysis before and during COVID-19," 2021, *arXiv:2108.07731*.

[30] V. M. de Lira, F. Pallonetto, L. Gabrielli, and C. Renso, "Predicting vehicles parking behaviour in shared premises for aggregated EV electricity demand response programs," 2021, *arXiv:2109.09666*.

[31] T. Ling, X. Zhu, X. Zhou, and S. Wang, "ParkLSTM: Periodic parking behavior prediction based on LSTM with multi-source data for contract parking spaces," in *Proc. Int. Conf. Wireless Algorithms, Syst., Appl.* Cham, Switzerland: Springer, 2021, pp. 262–274.

[32] W. Jiang and J. Luo, "Graph neural network for traffic forecasting: A survey," 2021, *arXiv:2101.11174*.

[33] A. M. Nagy and V. Simon, "Survey on traffic prediction in smart cities," *Pervas. Mobile Comput.*, vol. 50, pp. 148–163, Oct. 2018.

[34] K. Lee, M. Eo, E. Jung, Y. Yoon, and W. Rhee, "Short-term traffic prediction with deep neural networks: A survey," *IEEE Access*, vol. 9, pp. 54739–54756, 2021.

[35] M. Aljohani, S. Olariu, A. Alali, and S. Jain, "A survey of parking solutions for smart cities," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10012–10029, Aug. 2022.

[36] F. Al-Turjman and A. Malekloo, "Smart parking in IoT-enabled cities: A survey," *Sustain. Cities Soc.*, vol. 49, Jun. 2019, Art. no. 101608.

[37] T. Lin, H. Rivano, and F. Le Mouël, "A survey of smart parking solutions," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 12, pp. 3229–3253, Dec. 2017.

[38] M. G. Diaz Ogás, R. Fabregat, and S. Aciar, "Survey of smart parking systems," *Appl. Sci.*, vol. 10, no. 11, p. 3872, Jun. 2020.

[39] J. Parmar, P. Das, and S. M. Dave, "A machine learning approach for modelling parking duration in urban land-use," *Phys. A, Stat. Mech. Appl.*, vol. 572, Jun. 2021, Art. no. 125873.

[40] W. Zhang, H. Liu, Y. Liu, J. Zhou, and H. Xiong, "Semi-supervised hierarchical recurrent graph neural network for city-wide parking availability prediction," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, 2020, pp. 1186–1193.

[41] W. Zhang, H. Liu, Y. Liu, J. Zhou, T. Xu, and H. Xiong, "Semi-supervised city-wide parking availability prediction via hierarchical recurrent graph neural network," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 8, pp. 3984–3996, Aug. 2022.

[42] S. Yang, W. Ma, X. Pi, and S. Qian, "A deep learning approach to real-time parking occupancy prediction in transportation networks incorporating multiple spatio-temporal data sources," *Transp. Res. C, Emerg. Technol.*, vol. 107, pp. 248–265, Oct. 2019.

[43] A. O. Kotb, Y. Shen, and Y. Huang, "Smart parking guidance, monitoring and reservations: A review," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 2, pp. 6–16, Summer 2017.

[44] J. Parmar, P. Das, and S. M. Dave, "Study on demand and characteristics of parking system in urban areas: A review," *J. Traffic Transp. Eng.*, vol. 7, no. 1, pp. 111–124, Feb. 2020.

[45] X. Yin, G. Wu, J. Wei, Y. Shen, H. Qi, and B. Yin, "Deep learning on traffic prediction: Methods, analysis and future directions," 2020, *arXiv:2004.08555*.

[46] A. Boukerche and J. Wang, "Machine learning-based traffic prediction models for intelligent transportation systems," *Comput. Netw.*, vol. 181, Nov. 2020, Art. no. 107530.

[47] B. Chen, F. Pinelli, M. Sinn, A. Botea, and F. Calabrese, "Uncertainty in urban mobility: Predicting waiting times for shared bicycles and parking lots," in *Proc. 16th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2013, pp. 53–58.

[48] D. Zhao et al., "MePark: Using meters as sensors for citywide on-street parking availability prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 7244–7257, Jul. 2022.

[49] G. E. P. Box and D. A. Pierce, "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," *J. Amer. Stat. Assoc.*, vol. 65, no. 332, pp. 1509–1526, Dec. 1970.
[50] C. Badii, P. Nesi, and I. Paoli, "Predicting available parking slots on critical and regular services by exploiting a range of open data," *IEEE Access*, vol. 6, pp. 44059–44071, 2018.
[51] M. Alquraish, K. Ali. Abuhasel, A. S. Alqahtani, and M. Khadr, "SPI-based hybrid hidden Markov–GA, ARIMA–GA, and ARIMA–GA–ANN models for meteorological drought forecasting," *Sustainability*, vol. 13, no. 22, p. 12576, Nov. 2021.
[52] Y. Rong, Z. Xu, R. Yan, and X. Ma, "Du-parking: Spatio-temporal big data tells you realtime parking availability," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 646–654.
[53] F. Bock and M. Sester, "Improving parking availability maps using information from nearby roads," *Transp. Res. Proc.*, vol. 19, pp. 207–214, Jan. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352146516308675
[54] J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, Ny, USA: Springer, 2017.
[55] J. Leu and Z. Zhu, "Regression-based parking space availability prediction for the Ubike system," *IET Intell. Transp. Syst.*, vol. 9, no. 3, pp. 323–332, Apr. 2015.
[56] M. Eider, N. Bodenschatz, and A. Berl, "Evaluation of machine learning algorithms for the prediction of simulated company parking space occupancy," in *Proc. 11th Int. Conf. Adv. Comput. Inf. Technol. (ACIT)*, Sep. 2021, pp. 739–743.
[57] F. M. Awan, Y. Saleem, R. Minerva, and N. Crespi, "A comparative analysis of machine/deep learning models for parking space availability prediction," *Sensors*, vol. 20, no. 1, p. 322, Jan. 2020.
[58] J. Sonntag, M. Engel, and L. Schmidt-Thieme, "Predicting parking availability from mobile payment transactions with positive unlabeled learning," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 17, 2021, pp. 15408–15415.
[59] E. A. A.Alaoui and S. C. K. Tekouabou, "Parking availability prediction in smart city," in *Proc. 3rd Int. Conf. Smart City Appl.* Cham, Switzerland: Springer, 2020, pp. 450–462.
[60] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, vol. 112. Montreal, QC, Canada: Springer, 2013.
[61] T. K. Ho, "Random decision forests," in *Proc. 3rd Int. Conf. Document Anal. Recognit.*, 1995, pp. 278–282.
[62] S. C. K. Tekouabou, E. A. A. Alaoui, W. Cherif, and H. Silkan, "Improving parking availability prediction in smart cities with IoT and ensemble-based model," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 3, pp. 687–697, Mar. 2022.
[63] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[64] F. Santosa and W. W. Symes, "Linear inversion of band-limited reflection seismograms," *SIAM J. Sci. Stat. Comput.*, vol. 7, no. 4, pp. 1307–1330, Oct. 1986.
[65] X. Wang and Y. Zhong, "Statistical learning theory and state of the art in SVM," in *Proc. 2nd IEEE Int. Conf. Cognit. Informat.*, Aug. 2003, pp. 55–59.
[66] K. Gkolias and E. I. Vlahogianni, "Convolutional neural networks for on-street parking space detection in urban networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 12, pp. 4318–4327, Dec. 2019.
[67] J. Fan, Q. Hu, Y. Xu, and Z. Tang, "Predicting the vacant parking space availability: An LSTM approach," *IEEE Intell. Transp. Syst. Mag.*, vol. 14, no. 2, pp. 129–143, Oct. 2020.
[68] M. Awad and R. Khanna, *Support Vector Regression*. Berkeley, CA, USA: Apress, 2015, pp. 67–80, doi: 10.1007/978-1-4302-5990-9_4.
[69] J. Jebadurai and J. D. Peter, "SK-SVR: Sigmoid kernel support vector regression based in-scale single image super-resolution," *Pattern Recognit. Lett.*, vol. 94, pp. 144–153, Jul. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865517301241
[70] B. L. Smith, B. M. Williams, and R. K. Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transp. Res. C, Emerg. Technol.*, vol. 10, no. 4, pp. 303–321, Aug. 2002.
[71] A. Singh, A. Yadav, and A. Rana, "K-means with three different distance metrics," *Int. J. Comput. Appl.*, vol. 67, no. 10, pp. 13–17, Apr. 2013.
[72] Y. Ji, D. Tang, P. Blythe, W. Guo, and W. Wang, "Short-term forecasting of available parking space using wavelet neural network model," *IET Intell. Transp. Syst.*, vol. 9, no. 2, pp. 202–209, Mar. 2015.
[73] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," 2017, *arXiv:1710.05941*.
[74] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, "A survey of deep learning and its applications: A new paradigm to machine learning," *Arch. Comput. Methods Eng.*, vol. 27, no. 4, pp. 1071–1092, Sep. 2020.
[75] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
[76] E. I. Vlahogianni, K. Kepaptsoglou, V. Tsetsos, and M. G. Karlaftis, "A real-time parking prediction system for smart cities," *J. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 192–204, Mar. 2016.
[77] M. Buscema, "Back propagation neural networks," *Substance Use Misuse*, vol. 33, no. 2, pp. 233–270, Jan. 1998.
[78] J. Li, J. Li, and H. Zhang, "Deep learning based parking prediction on cloud platform," in *Proc. 4th Int. Conf. Big Data Comput. Commun. (BIGCOM)*, Aug. 2018, pp. 132–137.
[79] B. A. Pearlmutter, "Learning state space trajectories in recurrent neural networks," *Neural Comput.*, vol. 1, no. 2, pp. 263–269, Jun. 1989.
[80] A. Cleeremans, D. Servan-Schreiber, and J. L. McClelland, "Finite state automata and simple recurrent networks," *Neural Comput.*, vol. 1, no. 3, pp. 372–381, Sep. 1989.
[81] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019.
[82] F. Zhang et al., "Periodic weather-aware LSTM with event mechanism for parking behavior prediction," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 12, pp. 5896–5909, Dec. 2022.
[83] H. Yang et al., "Truck parking pattern aggregation and availability prediction by deep learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 12778–12789, Aug. 2022.
[84] J. Arjona, M. Linares, J. Casanovas-Garcia, and J. J. Vázquez, "Improving parking availability information using deep learning techniques," *Transp. Res. Proc.*, vol. 47, pp. 385–392, Jan. 2020.
[85] K. Lv, H. Chen, and Y. Lv, "Parking space prediction algorithm based on recurrent neural network and ensemble learning algorithm," in *Proc. IEEE 6th Int. Conf. Cloud Comput. Intell. Syst. (CCIS)*, Dec. 2019, pp. 114–119.
[86] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.
[87] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[88] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput.*, vol. 12, no. 10, pp. 2451–2471, Oct. 2000.
[89] R. Liu, S. Li, J. Liu, L. Ma, X. Fan, and Z. Luo, "Learning Hadamard-product-propagation for image dehazing and beyond," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 4, pp. 1366–1379, Apr. 2021.
[90] F. Landi, L. Baraldi, M. Cornia, and R. Cucchiara, "Working memory connections for LSTM," *Neural Netw.*, vol. 144, pp. 334–341, Dec. 2021.
[91] A. Mishra and S. Deshpande, "Deep learning based parking prediction using LSTM approach," in *Proc. Int. Conf. Inf. Commun. Technol. Intell. Syst* Cham, Switzerland: Springer, 2020, pp. 589–597.
[92] T. Anagnostopoulos, P. Fedchenkov, N. Tsotsolas, K. Ntalianis, A. Zaslavsky, and I. Salmon, "Distributed modeling of smart parking system using LSTM with stochastic periodic predictions," *Neural Comput. Appl.*, vol. 32, no. 14, pp. 10783–10796, Jul. 2020.
[93] Z. Wang, W. Xue, and X. Hei, "Research on parking space prediction based on long short-term memory," in *Proc. Int. Conf. Intell. Comput., Autom. Syst. (ICICAS)*, Dec. 2019, pp. 392–396.
[94] H. Chen, G. Chen, Q. Lu, and L. Peng, "MMSE-based optimized transfer strategy for transfer prediction of parking data," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 407–412.
[95] H. S. Jomaa, J. Grabocka, L. Schmidt-Thieme, and A. Borek, "A hybrid convolutional approach for parking availability prediction," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–8.
[96] M. Sun, Z. Li, L. Peng, H. Li, and X. Fang, "FLOPS: An efficient and high-precision prediction on available parking spaces in a long time-span," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2937–2942.
[97] L. Xiangdong, C. Yuefeng, C. Gang, and X. Zengwei, "Prediction of short-term available parking space using LSTM model," in *Proc. 14th Int. Conf. Comput. Sci. Educ. (ICCSE)*, Aug. 2019, pp. 631–635.

[98] K. Cho et al., "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, A. Moschitti, B. Pang, and W. Daelemans, Eds. Doha, Qatar: ACL, 2014, pp. 1724–1734, doi: 10.3115/v1/d14-1179.

[99] G.-B. Zhou, J. Wu, C.-L. Zhang, and Z.-H. Zhou, "Minimal gated unit for recurrent neural networks," *Int. J. Autom. Comput.*, vol. 13, no. 3, pp. 226–234, Jun. 2016, doi: 10.1007/s11633-016-1006-2.

[100] Y. LeCun et al., "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.

[101] A. Borovykh, S. Bohte, and C. W. Oosterlee, "Conditional time series forecasting with convolutional neural networks," 2018, *arXiv:1703.04691*.

[102] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: A survey," *Mech. Syst. Signal Process.*, vol. 151, Apr. 2021, Art. no. 107398.

[103] R. Mittelman, "Time-series modeling with undecimated fully convolutional neural networks," 2015, *arXiv:1508.00317*.

[104] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," 2018, *arXiv:1803.01271*.

[105] J. Ye, J. Zhao, K. Ye, and C. Xu, "How to build a graph-based deep learning architecture in traffic domain: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 5, pp. 1–21, Dec. 2020.

[106] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.

[107] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and locally connected networks on graphs," 2013, *arXiv:1312.6203*.

[108] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 3844–3852.

[109] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.

[110] X. Xiao, Z. Jin, Y. Hui, N. Cheng, and T. H. Luan, "Spatial–temporal graph convolutional networks for parking space prediction in smart cities," in *Proc. IEEE 94th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2021, pp. 1–5.

[111] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.

[112] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *Proc. 6th Int. Conf. Learn. Represent. (ICLR)*, Vancouver, BC, Canada, Apr. 2018, pp. 1–12. [Online]. Available: https://openreview.net/forum?id=rJXMpikCZ

[113] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.

[114] J. Liu, G. P. Ong, and X. Chen, "GraphSAGE-based traffic speed forecasting for segment network with sparse data," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 1755–1766, Mar. 2022.

[115] C. Zheng, X. Fan, C. Wang, and J. Qi, "GMAN: A graph multi-attention network for traffic prediction," in *Proc. 34th AAAI Conf. Artif. Intell. (AAAI), 32nd Innov. Appl. Artif. Intell. Conf. (IAAI), 10th AAAI Symp. Educ. Adv. Artif. Intell. (EAAI)*, New York, NY, USA: AAAI Press, Feb. 2020, pp. 1234–1241. [Online]. Available: https://aaai.org/ojs/index.php/AAAI/article/view/5477

[116] Z. Pan et al., "Spatio-temporal meta learning for urban traffic prediction," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 3, pp. 1462–1476, Mar. 2020.

[117] M. S. Rahaman et al., "MoParkeR : Multi-objective parking recommendation," 2021, *arXiv:2106.07384*.

[118] M. Schneble and G. Kauermann, "Statistical modeling of on-street parking lot occupancy in smart cities," 2021, *arXiv:2106.06197*.

[119] F. Richter, S. Di Martino, and D. C. Mattfeld, "Temporal and spatial clustering for a parking prediction service," in *Proc. IEEE 26th Int. Conf. Tools Artif. Intell.*, Nov. 2014, pp. 278–282.

[120] E. H.-C. Lu and C.-H. Liao, "A parking occupancy prediction approach based on spatial and temporal analysis," in *Proc. Asian Conf. Intell. Inf. Database Syst.* Cham, Switzerland: Springer, 2018, pp. 500–509.

[121] M. V. Muntean, "Car park occupancy rates forecasting based on cluster analysis and kNN in smart cities," in *Proc. 11th Int. Conf. Electron., Comput. Artif. Intell. (ECAI)*, Jun. 2019, pp. 1–4.

[122] S. S. Chawathe, "Using historical data to predict parking occupancy," in *Proc. IEEE 10th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Oct. 2019, pp. 0534–0540.

[123] Z. Zhao, Y. Zhang, and Y. Zhang, "A comparative study of parking occupancy prediction methods considering parking type and parking scale," *J. Adv. Transp.*, vol. 2020, pp. 1–12, Feb. 2020.

[124] R. F. Pozo, A. B. R. González, M. R. Wilby, J. J. Vinagre Díaz, and M. V. Matesanz, "Prediction of on-street parking level of service based on random undersampling decision trees," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 8327–8336, Jul. 2022.

[125] X. Fang, R. Xiang, L. Peng, H. Li, and Y. Sun, "SAW: A hybrid prediction model for parking occupancy under the environment of lacking real-time data," in *Proc. 44th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2018, pp. 3134–3137.

[126] J. Qiu, J. Tian, H. Chen, and X. Lu, "Prediction method of parking space based on genetic algorithm and RNN," in *Proc. Pacific Rim Conf. Multimedia*. Cham, Switzerland: Springer, 2018, pp. 865–876.

[127] G. Ali et al., "IoT based smart parking system using deep long short memory network," *Electronics*, vol. 9, no. 10, p. 1696, Oct. 2020.

[128] F. Bock, K. Xia, and M. Sester, "Mapping similarities in temporal parking occupancy behavior based on city-wide parking meter data," *Proc. Integr. Comput.-Aided Eng.*, vol. 1, pp. 1–5, Dec. 2018.

[129] L. Zheng, X. Xiao, B. Sun, D. Mei, and B. Peng, "Short-term parking demand prediction method based on variable prediction interval," *IEEE Access*, vol. 8, pp. 58594–58602, 2020.

[130] Y. Liu, C. Liu, and X. Luo, "Spatiotemporal deep-learning networks for shared-parking demand prediction," *J. Transp. Eng., A, Syst.*, vol. 147, no. 6, Jun. 2021, Art. no. 04021026.

[131] Y. Xue, C. Shao, and Y. Chen, "Study on application for parking demand prediction with scenario analysis," in *Proc. CICTP*, Jan. 2018, pp. 4110–4120.

[132] A. R. Alho and J. D. A. E. Silva, "Freight-trip generation model: Predicting urban freight weekly parking demand from retail establishment characteristics," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2411, no. 1, pp. 45–54, Jan. 2014.

[133] S. Biswas, S. Chandra, and I. Ghosh, "Effects of on-street parking in urban context: A critical review," *Transp. Developing Economies*, vol. 3, no. 1, p. 10, Apr. 2017.

[134] W. Shao, F. D. Salim, A. Song, and A. Bouguettaya, "Clustering big spatiotemporal-interval data," *IEEE Trans. Big Data*, vol. 2, no. 3, pp. 190–203, Sep. 2016.

[135] E. H.-K. Wu, J. Sahoo, C.-Y. Liu, M.-H. Jin, and S.-H. Lin, "Agile urban parking recommendation service for intelligent vehicular guiding system," *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 1, pp. 35–49, Jun. 2014.

[136] S. R. Rizvi, S. Zehra, and S. Olariu, "ASPIRE: An agent-oriented smart parking recommendation system for smart cities," *IEEE Intell. Transp. Syst. Mag.*, vol. 11, no. 4, pp. 48–61, Winter 2019.

[137] S. Kolassa et al., "Advantages of the MAD/MEAN ratio over the MAPE," *Foresight, Int. J. Appl. Forecasting*, vol. 2007, no. 6, pp. 40–43, 2007.

[138] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecasting*, vol. 22, no. 4, pp. 679–688, Oct. 2006.

[139] H. Hersbach, "Decomposition of the continuous ranked probability score for ensemble prediction systems," *Weather Forecasting*, vol. 15, no. 5, pp. 559–570, Oct. 2000.

[140] X. Fan, C. Xiang, L. Gong, X. He, C. Chen, and X. Huang, "UrbanEdge: Deep learning empowered edge computing for urban IoT time series prediction," in *Proc. ACM Turing Celebration Conf. China*, May 2019, pp. 1–6.

[141] Z. Qin, Z. Fang, Y. Liu, C. Tan, W. Chang, and D. Zhang, "EXIMIUS: A measurement framework for explicit and implicit urban traffic sensing," in *Proc. 16th ACM Conf. Embedded Networked Sensor Syst.*, Nov. 2018, pp. 1–14.

[142] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *Proc. 6th Int. Conf. Learn. Represent. (ICLR)*, Vancouver, BC, Canada, Apr. 32018, 2018, pp. 1–16. [Online]. Available: https://openreview.net/forum?id=SJiHXGWAZ

[143] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, J. Lang, Ed., Stockholm, Sweden, Jul. 2018, pp. 3634–3640, doi: 10.24963/ijcai.2018/505.

[144] L. Zhao et al., “T-GCN: A temporal graph convolutional network for traffic prediction,” *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 9, pp. 3848–3858, Sep. 2020, doi: 10.1109/TITS.2019.2935152.

[145] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, “Attention based spatial–temporal graph convolutional networks for traffic flow forecasting,” in *Proc. 33rd AAAI Conf. Artif. Intell. (AAAI), 31st Innov. Appl. Artif. Intell. Conf.* Washington, DC, USA: AAAI Press, 2019, pp. 922–929, doi: 10.1609/aaai.v33i01.3301922.

[146] X. Dai et al., “DeepTrend 2.0: A light-weighted multi-scale traffic prediction model using detrending,” *Transp. Res. C, Emerg. Technol.*, vol. 103, pp. 142–157, Jun. 2019.

[147] X. Geng et al., “Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting,” in *Proc. 33rd AAAI Conf. Artif. Intell. (AAAI), 31st Innov. Appl. Artif. Intell. Conf.* Washington, DC, USA: AAAI Press, 2019, pp. 3656–3663, doi: 10.1609/aaai.v33i01.33013656.

[148] X. Zhang, R. Cao, Z. Zhang, and Y. Xia, “Crowd flow forecasting with multi-graph neural networks,” in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Glasgow, U.K., Jul. 2020, pp. 1–7, doi: 10.1109/IJCNN48605.2020.9207457.

[149] W. Shao et al., “Long-term spatio-temporal forecasting via dynamic multiple-graph attention,” in *Proc. 31st Int. Joint Conf. Artif. Intell.*, Jul. 2022, pp. 2225–2232.

[150] Z. Tian, “Chaotic characteristic analysis of network traffic time series at different time scales,” *Chaos, Solitons Fractals*, vol. 130, Jan. 2020, Art. no. 109412.

[151] H. Yang, X. Zhang, Z. Li, and J. Cui, “Region-level traffic prediction based on temporal multi-spatial dependence graph convolutional network from GPS data,” *Remote Sens.*, vol. 14, no. 2, p. 303, Jan. 2022.

[152] D. Chai, L. Wang, and Q. Yang, “Bike flow prediction with multi-graph convolutional networks,” in *Proc. 26th ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, F. B. Kashani, E. G. Hoel, R. H. Güting, R. Tamassia, and L. Xiong, Eds. Seattle, WA, USA: ACM Nov. 2018, pp. 397–400, doi: 10.1145/3274895.3274896.

[153] L. Xia et al., “Spatial–temporal sequential hypergraph network for crime prediction with dynamic multiplex relation learning,” in *Proc. 30th Int. Joint Conf. Artif. Intell.*, Z. Zhou, Ed., Montreal, QC, Canada, Aug. 2021, pp. 1631–1637, doi: 10.24963/ijcai.2021/225.

[154] T. Mallick, P. Balaprakash, E. Rask, and J. Macfarlane, “Transfer learning with graph neural networks for short-term highway traffic forecasting,” in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 10367–10374.

[155] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, “Meta-learning in neural networks: A survey,” 2020, *arXiv:2004.05439*.

[156] Z. Pan, Y. Liang, W. Wang, Y. Yu, Y. Zheng, and J. Zhang, “Urban traffic prediction from spatio-temporal data using deep meta learning,” in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, A. Teredesai, V. Kumar, Y. Li, R. Rosales, E. Terzi, and G. Karypis, Eds. Anchorage, AK, USA: ACM, Jul. 2019, pp. 1720–1730, doi: 10.1145/3292500.3330884.

[157] H. Yao, Y. Liu, Y. Wei, X. Tang, and Z. Li, “Learning from multiple cities: A meta-learning approach for spatial–temporal prediction,” in *Proc. World Wide Web Conf.*, L. Liu, R. W. White, A. Mantrach, F. Silvestri, J. J. McAuley, R. Baeza-Yates, and L. Zia, Eds. San Francisco, CA, USA: ACM, May 2019, pp. 2181–2191, doi: 10.1145/3308558.3313577.

[158] C.-F. Tsai and Y.-F. Hsu, “A meta-learning framework for bankruptcy prediction,” *J. Forecasting*, vol. 32, no. 2, pp. 167–179, Mar. 2013.

[159] Y. Jang, H. Lee, S. J. Hwang, and J. Shin, “Learning what and where to transfer,” in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 3030–3039.

[160] G. Wei, C. Lan, W. Zeng, and Z. Chen, “MetaAlign: Coordinating domain alignment and classification for unsupervised domain adaptation,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 16643–16653.

[161] Y. Lu, X. Jiang, Y. Fang, and C. Shi, “Learning to pre-train graph neural networks,” in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 5, 2021, pp. 4276–4284.

[162] H. Pham, Z. Dai, Q. Xie, and Q. V. Le, “Meta pseudo labels,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 11557–11568.

[163] N. Neverova, C. Wolf, G. Taylor, and F. Nebout, “ModDrop: Adaptive multi-modal gesture recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 8, pp. 1692–1706, Aug. 2016.

[164] A. Valada, G. L. Oliveira, T. Brox, and W. Burgard, “Deep multispectral semantic scene understanding of forested environments using multimodal fusion,” in *Proc. Int. Symp. Exp. Robot.* Cham, Switzerland: Springer, 2016, pp. 465–477.

[165] R. Kiros, K. Popuri, D. Cobzas, and M. Jagersand, “Stacked multiscale feature learning for domain independent medical image segmentation,” in *Proc. Int. Workshop Mach. Learn. Med. Imag.* Cham, Switzerland: Springer, 2014, pp. 25–32.

[166] P. Mamoshina, A. Vieira, E. Putin, and A. Zhavoronkov, “Applications of deep learning in biomedicine,” *Mol. Pharmaceutics*, vol. 13, no. 5, pp. 1445–1454, May 2016.

[167] C. Chen, A. Seff, A. Kornhauser, and J. Xiao, “DeepDriving: Learning affordance for direct perception in autonomous driving,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 2722–2730.

[168] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[169] A. Vaswani et al., “Attention is all you need,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.

[170] Y. F. Geng and C. G. Cassandras, “New ‘smart parking’ system based on resource allocation and reservations,” *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1129–1139, Sep. 2013.

[171] M. Karthi and P. Harris, “Smart parking with reservation in cloud based environment,” in *Proc. IEEE Int. Conf. Cloud Comput. Emerg. Markets (CCEM)*, Oct. 2016, pp. 164–167.

[172] *Automated Valet Parking*. Bosch Global. Accessed: Dec. 12, 2021. [Online]. Available: https://www.bosch.com/stories/automated-valet-parking/

[173] Y. Tian, X. Cao, K. Huang, C. Fei, Z. Zheng, and X. Ji, “Learning to drive like human beings: A method based on deep reinforcement learning,” *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6357–6367, Jul. 2022.

[174] J. Yang, T. Chen, B. Payne, P. Guo, Y. Zhang, and J. Guo, “Generating routes for autonomous driving in vehicle-to-infrastructure communications,” *Digit. Commun. Netw.*, vol. 6, no. 4, pp. 444–451, Nov. 2020.

[175] H. Tavafoghi, K. Poolla, and P. Varaiya, “A queuing approach to parking: Modeling, verification, and prediction,” 2019, *arXiv:1908.11479*.

[176] W. Y. B. Lim et al., “Towards federated learning in UAV-enabled Internet of Vehicles: A multi-dimensional contract-matching approach,” *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5140–5154, Aug. 2021.

[177] P. Zhang et al., “Reinforcement learning-based end-to-end parking for automatic parking system,” *Sensors*, vol. 19, no. 18, p. 3996, Sep. 2019.

[178] J. Zhang, H. Chen, S. Song, and F. Hu, “Reinforcement learning-based motion planning for automatic parking system,” *IEEE Access*, vol. 8, pp. 154485–154501, 2020.

[179] H. Chai, R. Ma, and H. M. Zhang, “Search for parking: A dynamic parking and route guidance system for efficient parking and traffic management,” *J. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 541–556, Nov. 2019.

[180] J. Li, H. Lu, K. Xue, and Y. Zhang, “Temporal netgrid model-based dynamic routing in large-scale small satellite networks,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6009–6021, Jun. 2019.

[181] C. Zhu, A. Mehrabi, Y. Xiao, and Y. Wen, “CrowdParking: Crowdsourcing based parking navigation in autonomous driving era,” in *Proc. Int. Conf. Electromagn. Adv. Appl. (ICEAA)*, Sep. 2019, pp. 1401–1405.

[182] Y. Nie, K. Xu, H. Chen, and L. Peng, “Crowd-parking: A new idea of parking guidance based on crowdsourcing of parking location information from automobiles,” in *Proc. 45th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, vol. 1, Oct. 2019, pp. 2779–2784.

[183] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, “Recommendation systems: Principles, methods and evaluation,” *Egyptian Informat. J.*, vol. 16, no. 3, pp. 261–273, Nov. 2015.

[184] L. Pellegrini, V. Lomonaco, G. Graffieti, and D. Maltoni, “Continual learning at the edge: Real-time training on smartphone devices,” 2021, *arXiv:2105.13127*.

Xiao Xiao (Member, IEEE) received the bachelor’s degree in control technology and instrumentation from Xidian University, China, in 2004, and the Ph.D. degree in measuring and testing technologies and instruments from Zhejiang University, Hangzhou, China, in 2009. He is currently a Lecturer with the School of Telecommunications Engineering, Xidian University. He joined the Min H. Kao Department of Electrical Engineering and Computer Science, The University of Tennessee, as a Research Fellow, in 2019. His research interests include wireless network architecture, intelligent transportation systems, computer vision, and deep learning.

Ziyan Peng received the bachelor's degree from Xidian University, where she is currently pursuing the master's degree with the School of Telecommunications Engineering. Her research interests include deep learning, graph neural networks, and time series prediction.

Yunqing Lin received the bachelor's degree from the Harbin Institute of Technology. She is currently pursuing the master's degree with the School of Telecommunications Engineering, Xidian University. Her research interests include deep learning, object detection, and object tracking.

Zhiling Jin received the bachelor's degree from Xidian University, where he is currently pursuing the master's degree with the School of Telecommunications Engineering. His research interests include deep learning, graph neural networks, intelligent transportation systems, big data, and spatio-temporal data mining.

Wei Shao received the Master of Science (Hons.) in computer science (CS) from The University of Hong Kong, and the Ph.D. degree from RMIT University in 2018. He worked as a Researcher Assistant at the Computer Science Department, University of Hong Kong, from 2012 to 2014. He was carrying out an ITF project. His current work is focusing on investigating AI algorithms in mix-road autonomy car. His research interest includes fairness issue in machine learning, wireless sensing, the Internet of Things, and smart cities. He was awarded the scholarship from SUPP project.

Rui Chen (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in communications and information systems from Xidian University, Xi'an, China, in 2005, 2007, and 2011, respectively. From 2014 to 2015, he was a Visiting Scholar with Columbia University, New York City, NY, USA. He is currently an Associate Professor and a Ph.D. Supervisor of the School of Telecommunications Engineering, Xidian University. He has published about 50 papers in international journals and conferences and holds ten patents. His research interests include broadband wireless communication systems, array signal processing, and intelligent transportation systems. He is also an Associate Editor of the *International Journal of Electronics*, *Communications*, and *Measurement Engineering* (IGI Global), and a reviewer of a number of international journals and conferences, including the IEEE, Springer, and Elsevier.

Nan Cheng (Member, IEEE) received the B.E. and M.S. degrees from the Department of Electronics and Information Engineering, Tongji University, Shanghai, China, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. He is currently a joint Professor with the School of Telecommunication, Xidian University, Xi'an, China. He is also a joint Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, and Department of Electrical and Computer Engineering, University of Waterloo. His research interests include performance analysis, MAC, opportunistic communication for vehicular networks, unmanned aerial vehicles, and the application of artificial intelligence for wireless networks.

Guoqiang Mao (Fellow, IEEE) has authored or coauthored more than 200 papers in international conferences and journals, which have been cited more than 9000 times. His research interests include intelligent transport systems, applied graph theory and its applications in telecommunications, the Internet of Things, wireless sensor networks, wireless localization techniques, and network modeling and performance analysis. He received the Top Editor Award for outstanding contributions to the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in 2011, 2014, and 2015, respectively. He was the Co-Chair of IEEE Intelligent Transport Systems Society Technical Committee on Communication Networks. He has been an Editor of the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS since 2018, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2014–2019), and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY since 2010.