Vol 18, Spring 1404

Information Technology in Engineering Design

https://sanad.iau.ir/journal/ited

Learning Identification Strategies for Traffic Flow Model: A Review Study

Mahboubeh Zare Feizabadi(1) Seyyed Abed Hosseini(2)* Mahboobeh Houshmand(3)

(1) Department of Electrical Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran

(2) Department of Electrical Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran*

(3) Department of Computer Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran

(Date received: 1403/06/19 Date accepted: 1403/09/04)

Abstract

Travel traffic and its results, which include air pollution and the chaotic economic situation, are among the factors that limit the development of healthy and sustainable cities. Traffic flow model parameters are important for urban road network management. Learning the identification strategy should be such that modeling the traffic network guarantees the model's simplicity, accuracy, and validation. In this paper, different traffic flow system identification methods, including traffic flow modeling and its prediction, have been reviewed and analyzed. Ultimately, the advantages and disadvantages of different methods in this field are categorized.

Keywords: Traffic flow model prediction, ARIMA model, Hybrid model, Deep learning, Non-linear macroscopic traffic model.

* Corresponding author:

Seyyed Abed Hosseini

Address: Department of Electrical Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran

Email: abed_hosseyni@yahoo.com

1- Introduction

Today, the high increase in vehicles has caused severe traffic problems, which makes people's trips difficult [1]. Fortunately, the progress and development of new technologies have given researchers the possibility of more research to reduce the problems of traffic complications by implementing effective and practical algorithms [2]. To address this issue, intelligent transportation systems (ITS) [3] have been extensively researched over the years and have proven to be an effective approach for enhancing the efficiency of urban transportation. Reducing traffic congestion is possible with advanced traffic signal control strategies [4]. Dynamic traffic route planning and control with short-term traffic forecasting models is known as an important application in the ITS traffic network [5], signal optimization [6], and real-time traffic management [7]. Better use of public transportation is one of the benefits of more accurate urban traffic forecasting [8]. Forecasting the traffic flow (TF) in the future moments in a specific traffic area based on the past TF of that area is one of the goals of traffic control [9]. Integrated moving traffic forecasting and linear regression (LR) models performed a temporal analysis of TFs [10].

Recently, neural networks (NNs) have demonstrated remarkable effectiveness in traffic forecasting tasks due to their exceptional skills in pattern recognition and data handling [9]. NN-based traffic forecasting models that can analyze hidden temporal patterns in traffic data include recurrent NN (RNN), gated recurrent uni (GRU), and long-short-term memory (LSTM) [11-13]. Convolutional NN (CNN) and graph convolutional network (GCN) are typical examples of these methods [14-16]. Due to the ability to analyze graph structure, graph NNs (GNN) can be used in various applications of intelligent transportation networks [2]. A new method for combining two-dimensional (2D) time series patterns is proposed by formulating an intra-day ordered square matrix that expands in both intra-day and inter-day dimensions, and then convolution and maximum integration are used to extract information [17]. The post-processing of the machine algorithm (ML) residuals is done completely and independently by a new approach to the traffic prediction model. It combines an ML algorithm with a basic statistical time series model. A basic NN model is combined with an autoregressive integrated moving average (ARIMA) called NN-ARIMA [18]. The TF prediction model is based on the multifaceted deep learning (DL) theory [19]. Some TF forecasting problems have been solved by presenting an adaptive multifaceted DL model based on CNN and GRU.

Since TFs are affected by temporal and spatial conditions, researchers are trying to expand spatial-temporal methods [9]. It is usually difficult to analyze the distinct influence degrees without manual training. Real traffic conditions do not proceed like this, decreasing forecast accuracy. The spatial-temporal TF forecasting model named LSTM is described in [9] to improve the previous problems. A completely new traffic forecasting model for short-term TFs with time series analysis and an improved LSTM is presented in [20]. The improved LSTM combines the traditional LSTM with the bidirectional LSTM (Bi-LSTM). One of the advantages of LSTM is its impact on the processing of longer time series data. As a result, the processing and prediction of TF data are obtained on a larger scale. A DL hybrid model is also presented in [21] to predict short-term TF. An optimization method for abnormal network traffic detection based on semi-supervised double deep Q-network (SSDDQN) is proposed in [22], which is based on the representative algorithm of DDQN [23] in deep reinforcement learning (DRL). This method, with the combination of two unsupervised learning algorithms, autoencoder (AE) [24] and K-means [25], is implemented. Deep NN is used to improve traffic engineering in an application-oriented scenario where dynamic flow information of a real-world campus network is predicted [26].

Currently, the quality and accuracy of the model are important in TF prediction control strategies [27]. Various qualitative and quantitative indicators are used in specific time and place conditions to determine the traffic congestion situation [28]. In [29], an estimation of the traffic situation is presented online using the Bayesian framework, particle filter techniques, and first-order macroscopically. The benefits of combining the effect of rain in predicting the traffic situation have been seen experimentally. The technology of using artificial intelligence (AI) in vehicles to increase the accuracy of estimating and predicting TF is presented in [30]. With the development of technology and the use of vast TF data [31, 32], ML methods [33, 34] and DL [35] have been used To identify the traffic situation and achieve excellent results [36]. Selective ensemble learning (SEL), which is dedicated to identifying the TF situation on a highway, using experiments to verify the performance. Real-time and accurate highway traffic conditions are accessible through research documentation for planning purposes. Also, the reduction of model storage overheads will result from this approach [28].

Mathematical relationships between three important factors, including vehicle speed, density, and TF, make a macroscopic model based on hydrodynamic theory, which is the TF control model [37]. For example, solving the parameter estimation problem of the TF model macroscopically has been used in [38]. Due to the randomness of the TF system in reality, the Monte Carlo method has been implemented to generalize the random variable [39]. To estimate the parameters of a highway's rotation ratio, saturation flow, and free flow speed, a combination of random approximation, online self-approximation, and expectation-maximization algorithms have been used [40, 41].

The remainder of this paper is structured as follows: Section 2 introduces different methods of traffic model estimation to predict future TF based on previous data, including time series models, traffic prediction methods based on AI, and hybrid models. Section 3 explains learning network traffic detection models. Section 4 analyzes each method's advantages and disadvantages and suggests improving the quality of future works.

2- TF model prediction methods

Many methods for predicting macroscopic traffic situations have been proposed in the last three decades. These methods are related to different fields of statistics, control theory, AI, applied mathematics, etc. The classification of these methods in terms of literature is done differently [42]. This section briefly and generally discusses the three main types of research studies based on different methods of predicting the TF of the traffic network.

2-1- Statistical and mathematical time series model

In the preliminary studies, researchers have used methods based on statistics and mathematics to predict TF [43, 44]. The time series model is used for prediction based on the collected TF data. In the literature, several methods use time series models (e.g., ARIMA [45] and Kalman filter [46]) to evaluate TF in each route. Autoregressive (AR), autoregressive moving-average (ARMA), and ARIMA methods are among the primary models for predicting TF that work based on the assumption of fixed traffic data [47-49]. By developing univariate and multivariate ARIMA models in real-time and checking their results, better performance of the multivariate method has been achieved [50]. Considering that the temporal analysis of these models is performed on long-term traffic data, they can be tested for routes with little traffic and do not apply to complex traffic conditions [9]. Most of these statistical and mathematical models are based on ideal traffic conditions and do not apply to real and complex TF conditions [51, 52]. In addition, some researchers have used the combination of spatial information in predicting TF [53, 54]. However, using these traditional methods has resulted in not using the real spatial-temporal characteristics of the complex traffic network [55].

2-1-1- ARIMA model

ARIMA is a linear and univariate model. The widespread use of this method in urban areas shows its lack of accuracy in considering the non-linear characteristics of TF time series. However, ARIMAs' advantages include their simple mathematical form and high capacity to clearly identify temporal correlation in TF time series. The mathematical form of ARIMA (p, d, q) can be written as Equation (1) [56].

$$\varphi(B)(1-B)^d X_t = \theta(B)e_t, \tag{1}$$

where X_t is time series variable, $\{e_t\} \sim WN(0, \sigma^2)$, $\varphi(B) = 1 - \varphi_1 B - \cdots - \varphi_p B^p$ and $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$ are p th and q th-degree polynomials, $\varphi_1, \dots, \varphi_p$ and $\theta_1, \dots, \theta_q$ are coefficients, d denotes the order of difference, B is the backward shift operator, p and q represents AR and moving average (MA) order, respectively.

2-2- Application of AI in TF prediction modeling

Today, many AI models are widely used in the transportation industry. In [57] argues that an AI can use past data to predict the future traffic network to advance network management and automation. Using AI networks, [58, 59] and [60] focus on considering complex traffic characteristics for intelligent traffic routing and improving traffic network analysis. Urban TF forecasting using methods based on deep NNs and considering spatial-temporal features has recently made significant progress [61-64]. Initially, RNNs were used to analyze the changing patterns. However, long-term time dependencies are not analyzed with RNN-based methods [9]. For this purpose, LSTM and GRU have recently been proposed.

Cui et al. [65] used two-way LSTM or bi-LSTM to achieve higher accuracy in traffic prediction. In the improved LSTM proposed by Zhao et al., time-space independent correlations are used [66]. To communicate more and better with traffic data, Gu et al. [67] used entropy-based gray relationship analysis and investigation of the effects between different lanes to propose a lane-level traffic forecasting model by combining GRU and LSTM methods. The particle swarm algorithm optimization was done by Chan et al., which improved the effectiveness and accuracy of the prediction model [68]. Applying traffic complications and using a lot of data was done with the emergence of DL in TF prediction models [69]. Due to the importance of spatial information in predicting TF, spatial models analyze spatial correlations between adjacent locations on roads [9]. First, the method based on CNN has been used to analyze the communication of several adjacent places in the traffic network [70]. CNN does not make efficient evaluation of non-Euclidean structured data possible, and the GCN method is applied to the TF prediction model to connect with the graphical structured data [71]. As the road network in GCN cannot dynamically change over time, graph attention (GAT) employs an attention mechanism to represent the relationships between nodes in various road segments and to analyze the traffic data within the graph structure [72, 73]. Convolutional

LSTMs [74] and graph convolutional gating recursive units [75] are used to calculate time relations to obtain spatial-temporal TF data in the traffic transportation network. LSTM was presented by Tian et al. to examine deleted traffic data and irregular sampling [76]. However, AI-based models have disadvantages such as parameter selection and excessive fitting.

2-2-1- NN

Various NN-based models have been extensively utilized for short-term traffic forecasting. Research indicates that a single-layer feed-forward NN can approximate any continuous multivariate function to a specified level of accuracy. Networks with multiple hidden layers can be converted into networks with a single hidden layer [77, 78].

2-2-2- CNN

CNN is a forward NN whose neurons can respond to a part of the covered area. CNN has performed well in image processing. The CNN analysis includes three layers: convolution, activation, and integration [79]. A special method is needed to extract complex intra-day and inter-day traffic information patterns. Complex intra-day and inter-day traffic patterns are not seen in conventional time series or may be lost [17]. In similar image processing applications, the input image data and their pixels depend on neighboring pixels in all dimensions. The TF data is also highly dependent on the neighboring data information in each route in the previous and following days. Therefore, CNN can be a tool for extracting patterns so that by applying it to the input traffic data and fitting the output structure, the final reconstruction is achieved.

The CNN convolution operation extracts and learns the local features of the traffic network due to the periodicity and temporal displacement of the traffic data, bringing a deeper representation. Combining CNN with LSTM in computer vision has been very extensive and has made significant progress in areas such as video prediction [80-82].

2-2-3- RNN

RNN is a dynamic and suitable method for predicting traffic time series or speech recognition systems [19]. RNNs can be considered a special class of NN for traffic time sequences [55]. RNN can analyze time-hidden properties in traffic time sequences [83]. The structure of RNN can maintain past traffic information. NN can be considered a deep NN structure with infinite layers whose detection is in the time domain [19].

The combination of LSTM and RNN allows long-term traffic information storage and exploitation of their time dependencies. In PredRNN, multiple layers of spatial-temporal

LSTM are used to make changes in the framework of traffic inputs [82]. However, the implementation of these models is used to predict TF in the time domain and cannot be applied to dynamic spatial dependencies. Many researchers have used RNNs to predict TF [84-86]. Research shows that LSTM helps extract traffic conditions in these models, but they are incapable of fully modeling the complex spatial-temporal properties of traffic. Therefore, proposing a complete model for training and considering all spatial-temporal traffic properties effectively improves TF forecasting.

2-2-4- LSTM

The design of LSTM has been done to solve long-term dependencies. One of the advantages of LSTM is its long-term memory for maintaining traffic information, which does not mean paying less to obtain more information [87-90]. One of the most widely used methods in temporal analysis of TF prediction is the method of LSTM, which has been used in projects such as natural language processing [91] and speech recognition [92]. In a predictive NN that uses several layers based on LSTM, local predictions are made in each layer, then changes are transferred to the next layer [81]. To analyze, maintain, and extract long-term temporal dependencies, LSTM uses a cellular mode and three gates, including the input, forget, and output. LSTM lacks a comprehensive structure, as it relies solely on future information from time series and has limited access to past data, both of which are essential in practical traffic applications [93].

2-2-5- Bi-LSTM

Due to the necessity of including TF forecasts based on past and future information, a two-way circular network called Bi-LSTM was designed [20]. Bi-LSTM consists of a combination of two LSTM methods, forward and inverse LSTM. In this method, the output of the next moment is predicted based on the time information of the previous moment and combining the future situation with historical information. In the forward layer, forward calculations and in the backward layer, reverse calculations are performed and processed, and finally, the output is calculated [93].

2-2-6- GCN

Effective communication with the structured data of the traffic network and dealing with them is done by GCN, which processes these structured data in the form of a diagram by the nodes, and the performance of the traditional GCN in a traffic network is as follows [9]:

$$H_l = \begin{cases} X, & l = 0 \\ (A+E)H_{l-1}W_l, & l \geq 1 \end{cases} \tag{2}$$

where $A \in R^{N*N}$ is the adjacency matrix, $X \in R^{N*F}$ denotes the F features´ input data of N nodes, $H_l \in R^{N*F'}$ $(l \geq 1)$ means the convolution result of F' feathers after l th convolution, E is the identity matrix and W presents the trainable matrix. Figure 1 shows the complexity of the graph network.

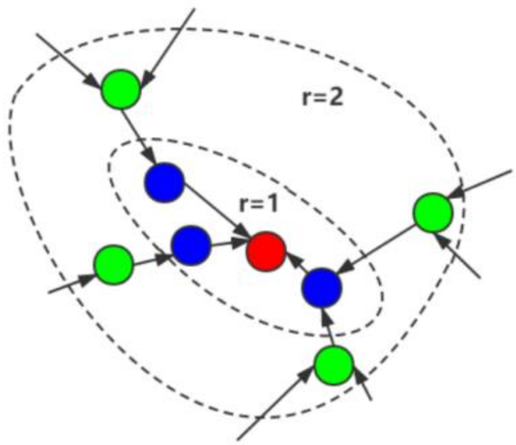

Figure 1. Layer-by-layer diffusion of convolutional graph algorithm [9]

The conventional GCN mechanism represents graph information using $A + E$, where the adjacency matrix A and the identity matrix E contain only 0s and 1s. A target node with more than one upstream traffic data node increases the complexity of the graph. Therefore, the inverse of the digraph degree matrix is used to normalize the data after each convolution operation. The equation can be written as follows:

$$H_l = \begin{cases} X, & l = 0 \\ D^{-1}(A+E)H_{l-1}W_l, & l \geq 1 \end{cases} \tag{3}$$

2-3- Hybrid models

Limitations include considering dynamic fluctuations and time-space characteristics in individual models [94, 95]. To compensate for these limitations, the researchers suggested combining and improving single prediction models [96]. DCRNN [97], Temporal GCN (T-GCN) [98], Attentive TF machine (ATFM) [55], and sequence-to-sequence model based on graph complexity [99] are four representative methods that directly combine spatial layers with temporal layers. Although hybrid models improve the limitations of single models, they still have disadvantages in terms of data mining.

Therefore, most researchers combine state prediction with data feature analysis [93]. For example, Lu et al. used a KNN and an LSTM to train and feed TF data [100]. A combined spatial-temporal analysis with alternating gate units was performed to improve the accuracy of short-term TF prediction by Dai et al. [101]. The most important advantage of hybrid models is their accuracy in predicting TF, but the disadvantages of these models include the high number of parameters and high computing power to adjust and optimize the model. In addition, most existing hybrid models have problems in data integration for model prediction, and little research has been done on this issue [93].

2-3-1- Multi-dimensional support vector regression (MSVR)

To check the effectiveness of the models, MSVR has replaced the role of NN [42]. In terms of literature, the application of SVR is in predicting traffic time series [102-106]. SVR is an algorithm based on ML in which input-output data pairs are mapped to a high-dimensional feature space through a non-linear transformation function, and their performance dependence is checked through an optimization function. Since it is usually difficult to determine this non-linear function, Mercer's theorem and the corresponding kernel function are used [107]. The linear regression function with feature space is as Equation (4).

$$f(x) = \langle w, x \rangle + b, \tag{4}$$

where $w \in X$, $b \in R$ are hyperplane regression parameters to be determined. Since Mercer's theorem represents a symmetric positive-definite function on a square as the sum of a convergent sequence of product functions, the symbol $\langle \quad \rangle$ is used to indicate non-negative eigenvalues. The solution for w and b can refer to [108-111]. Nevertheless, SVR is restricted to a one-dimensional (1D) output. Perez-Cruz et al. extended support vector machines to address multi-output, multi-input issues for regression estimation and function approximation [112]. The MSVR solution for the multi-dimensional regression parameters w and b is obtained from an iterative method under Karosh-Cohen-Tucker conditions [42].

2-3-2- NN-ARIMA and MSVR-ARIMA model

This model is important in predicting network traffic by combining ML algorithms with ARIMA. Maintaining TF characteristics on a network scale and common movement patterns is one of the goals of ML algorithms, and analyzing and recording the remaining traffic characteristics is one of the goals of ARIMA [42].

The time series of TFs observed in the network define ML algorithms as input. During a supervised learning process, the ML model is trained and validated. The spatial-temporal properties of TFs are stored in a weighted matrix. The trained model can predict the network for new inputs. ARIMA model fitting is also used for specific spatial predictions and white noise generation. Finally, the summary of the investigation's predictions and the final result is stated [42].

Finally, by combining two ML methods and the statistical method of ARIMA, researchers seek to improve the complex patterns of TFs, highly fluctuating non-linearities, and network correlations. In Equation (5), the prediction accuracy of the combined model is expressed as a mathematical equation.

$$Z_i(t) = Y_i(t) + X_i(t) + e_i(t), \tag{5}$$

where $Z_i(t)$ represents the observation of a time series at location i at time t, $Y_i(t)$ is the forecast of the ML algorithm, $X_i(t)$ is the forecast of ARIMA, and $e_i(t)$ is white noise. In short, in research, ML algorithms include NNs and SVRs, which make combinations of NN-ARIMA and MSVR-ARIMA. However, MSVR-ARIMA is not recommended by researchers for predicting the network traffic level due to its less competitive performance [42].

2-3-3- Variational mode decomposition (VMD)

VMD is a signal-smoothing method. This method decomposes the original non-linear traffic stream into several stable data streams, which is an important practice in TF prediction. TF prediction models are improved by smoothing and stabilizing non-stationary data [93].

2-3-4- Bat algorithm (BA)

BA is a swarm intelligence optimization algorithm [113]. BA measures the low or high position of the bat as the fitness value of the objective function of the problem. Changing the height and frequency of the pulse can be to some extent to reach the optimal state and it is like bat bait. In BA, the search frequency, position, and velocity are updated in the following manner [93]:

$$\begin{gathered} P_i = P_{min} + (P_{max} - P_{min})^{\partial}, \\ V_i^{t+1} = V_i^t + (S_i^t - S^*)P_i, \\ S_i^{t+1} = S_i^t + V_i^{t+1}, \end{gathered} \tag{6}$$

where P_i, P_{max} and P_{min} respectively represent the sound frequency emitted by the first bat at the current moment and the values of the maximum and minimum sound frequency. A random number of $\partial \in [0,1]$, and the current optimal position.

$$X_{new} = X_{old} + \beta A^t, \tag{7}$$

where X_{old} is a solution chosen from the current optimal solution; A^t is the average of all bat loudnesses at time t; β is a random number of $[-1,1]$.

2-3-5- Conv-LSTM

The combined Conv-LSTM method consists of CNN and LSTM to extract the spatial-temporal features of the traffic network. Conv-LSTM is the spatial-temporal TF matrix that expresses the predictable past TF and the neighbors' TF and is expressed in Equation (8). The convolution operation is performed on the TF data at each time step. Finally, a filter is applied to obtain the local domain.

$$X_t^s = \begin{bmatrix} X_{t-n}^s \\ X_{t-(n-1)}^s \\ \vdots \\ X_t^s \end{bmatrix}^T = \begin{bmatrix} f_{t-n}^1 & \cdots & f_t^1 \\ \vdots & \ddots & \vdots \\ f_{t-n}^m & \cdots & f_t^m \end{bmatrix}. \tag{8}$$

The convolution kernel filter process can be expressed as follows [21]:

$$Y_t^s = \sigma(W_s * X_t^s + b_s), \tag{9}$$

where W_s is the weights of the filter, b_s is biased, X_t^s is the input TF at time t, the symbol $*$ denotes the convolution operation, σ is the activation function and Y_t^s is the output of the convolutional layer. The extraction of spatial features from neighboring traffic locations using such a favorable method is stated.

2-3-6- Hybrid multi-modal DL framework

Hybrid models of TF prediction using DL have been presented in research [19]. According to research, it is difficult to express a model with a few hidden layers when considering the complex statistical characteristics of traffic. DL-based models address the local characteristics and long-term dependencies of TF data. A hybrid structure of DL, using CNN and GRU and considering the spatial-temporal characteristics of TFs, has been investigated in the research, which has been investigated firstly with 1D characteristics and then with multi-dimensional characteristics.

3- Learning network traffic detection models

Researchers currently control traffic signals using TF models [26], and their effectiveness relies on the model's quality and accuracy. Employing suitable mathematical models that capture steady-state and dynamic traffic characteristics proves more effective.

3-1- Detection model based on semi-supervised DRL

Models based on DRL methods have excellent performance in detecting abnormal TFs, but their strong dependence on network traffic characteristics is a fundamental flaw. As the quality of the traffic data label increases, the accuracy of the classification results will increase. Researchers have used large datasets of unbalanced classes, and traffic attacks have not been tested and have caused specific challenges in data classification [22]. Semi-supervised algorithms that combine DRL with unsupervised learning are proposed to improve the accuracy of abnormal traffic and reduce the cost of labeling.

The main work of the SSDDQN semi-supervised learning model is expressed in 4 steps [22]:

1. Deploy in busy and high-demand environments using simple and fast policy and Q functions.
2. Applying the reward function optimization process in complex and flexible traffic problems.
3. Reducing the cost of tagging and identifying unknown traffic attacks using the K-mean clustering algorithm and applying next-step traffic features.
4. Fast access to traffic detection with auto-encoder training and testing.

3-2- Detection based on non-linear macroscopic TF model

Traffic signal control systems can be expressed based on a macroscopic model as well as mathematical relationships between vehicle speed, route density, and TF and related to hydrodynamic theory [114]. The time-varying non-linear macroscopic TF model has been applied to steady-state and dynamic traffic characteristics to describe the traffic network more accurately and realistically [27]. This model is designed based on multi-parameter timing and repetitive inherent features. The transformation of the parameter identification model into the optimal tracking control model is done by the iterative learning identification strategy so that the traffic density is close to the real value and the accuracy is improved. The convergence and effectiveness of the model are checked by simulation analysis and testing and confirmed by improving the control method [27].

4- Conclusion and future works

This article reviews the methods of identifying traffic flow models, including traffic prediction and detection models. Each statistical method, based on artificial intelligence and hybrid methods in traffic forecasting and extensive models of traffic data detection, includes advantages and disadvantages that are briefly stated.

Using ARIMA [50], it was found that the multivariate model performs better. By analyzing the research done on the short-term prediction of TF, the spatial-temporal characteristics of traffic data have been shown, so the statistical models that are completely ideal based on the initial assumptions do not apply to the complex urban traffic network [51, 52]. ARIMA cannot handle non-linear traffic data [21].

CNN has performed well in the applications of image processing and temporal traffic analysis [19]. CNN also extracts traffic spatial features [21]. RNN is a dynamic system used in traffic time series prediction or speech recognition [19]. One of the main disadvantages of this method is the limitation of traffic information available to them [93].

Processing the learning sequence is one of the tasks of LSTM [19]. LSTM is the most efficient and widely used temporal analysis method in TF prediction [9]. However, one of the disadvantages of LSTM and CNN is that they do not take into account all the features of TF, which limits their performance improvement [21]. LSTM is applied to extract temporal features [21].

Bi-LSTM has been able to resolve the long-term dependence on LSTM and use past and future traffic information [93]. Extracting periodic traffic features and removing spatial dependencies are the advantages of Bi-LSTM [21]. The limitation of using the complex and dynamic characteristics of TF is one of the disadvantages of this method [21]. The extraction of spatial-temporal traffic characteristics is feasible according to Conv-LSTM [21]. The decomposition of non-linear TF data into relatively stable data is done by VMD [55]. Predicting TF with structured data and dealing with traffic nodes can be implemented in GCN. It is recommended to use the GAT diagram to describe the correlations of traffic nodes and analyze them instead of GCN, which does not cover the dynamic change in time [9].

According to the researchers, SSDDQN performed well in detecting unusual complex traffic and classifying data. This method's advantage is its accuracy, and its disadvantages are its low training time and prediction. However, the limitation of the optimization effect and the inability to detect unusual traffic attacks are the major disadvantages of this method [22].

A non-linear macroscopic TF model is presented to more accurately describe the actual performance of TF in the urban road network [27]. A precise theoretical derivation has proven the convergence of the presented algorithm with this time-varying parametric method. It has improved road traffic efficiency and reduced traffic congestion. However, the presented TF model does not consider the phenomenon of episodic congestion. In summary, the advantages and disadvantages of different methods are given in Table 1.

Table 1: Advantages and disadvantages of different methods

Ref.	Method	Advantages	Disadvantages
[50] [51] [52]	ARIMA	• Better performance of the multivariate model	• Not deal with complex models
[19] [21]	CNN	• Good image processing performance • Good for time series analysis • Utilized to extract spatial features	• Not capture the complex characteristics of TF
[93]	RNN	• Popular for handling sequence tasks	• Limited range of available background information
[21] [9]	LSTM	• Capable of processing sequence learning tasks • Applied to extract temporal feature • The most efficient, widely used, and representative mechanism	• Not capture the complex characteristics of TF • Long-term dependence
[93] [21]	Bi-LSTM	• Used to extract periodic features • Capture time dependence • Make full use of the forward and backward information	• Failure to use complex and dynamic features of TF
[21]	Conv-LSTM	• Extract the spatial-temporal feature of TF	• Need to improve forecast accuracy in daily and weekly periodic features
[93]	VMD	• Smooth the original non-linear historical TF data	• Necessity of combination with LSTM model to predict traffic flow data
[9]	GCN	• Deal with graph-structured data	• Not able to change dynamically
[22]	SSDDQN	• Good results in abnormal traffic detection • Good accuracy • Low training time • Low prediction time	• Limit optimization effect
[27]	Non-linear macroscopic TF model	• Describe more accurately the actual performance of TF • Convergence of the algorithm • Improve road traffic efficiency	• Does not consider the phenomenon of episodic congestion

For future work, the proposed method Conv-LSTM is a suitable option for predicting traffic data, considering the extent of spatial-temporal properties in traffic data and improving the accuracy of daily and weekly forecasts. In addition, the diagnostic model SSDDQN will be met by applying a non-linear macroscopic model and considering episodic densities and traffic network goals, including accuracy, high productivity, and reduction of training and prediction time.

References

1. Abbas, N., Sharafeddine, S., Hajj, H., & Dawy, Z. (2020). Price-aware traffic splitting in D2D HetNets with cost-energy-QoE tradeoffs. Computer Networks, 172, 107169.

2. Saeed Rahmani, A.B., Nizar Bouguila, & Zachary Patterson. (2023). Graph Neural Networks for Intelligent Transportation Systems: A Survey. IEEE Transactions on intelligent transportation systems, 24, 8846-8885.

3. Zhang, J., Wang, F. Y., Wang, K., Lin, W. H., Xu, X., & Chen, C. (2011). Data-driven intelligent transportation systems: A survey. IEEE Transactions on Intelligent Transportation Systems, 12(4), 1624-1639.

4. Genders, W., & Razavi, S. (2019). Asynchronous n-step Q-learning adaptive traffic signal control. Journal of Intelligent Transportation Systems, 23(4), 319-331.

5. Husein, D. M., Rizk, S., & Naim, H. Y. (2020). Differential effects of sucrase-isomaltase mutants on its trafficking and function in irritable bowel syndrome: similarities to congenital sucrase-isomaltase deficiency. Nutrients, 13(1), 9.

6. Smith, B. L., Williams, B. M., & Oswald, R. K. (2002). Comparison of parametric and nonparametric models for traffic flow forecasting. Transportation Research Part C: Emerging Technologies, 10(4), 303-321.

7. Rahman, S. A., Mourad, A., & El Barachi, M. (2019). An infrastructure-assisted crowdsensing approach for on-demand traffic condition estimation. IEEE Access, 7, 163323-163340.

8. Jiang, J., Chen, Q., Xue, J., Wang, H., & Chen, Z. (2020). A novel method about the representation and discrimination of traffic state. Sensors, 20(18), 5039.

9. Chen, Z., Lu, Z., Chen, Q., Zhong, H., Zhang, Y., Xue, J., & Wu, C. (2022). Spatial–temporal short-term traffic flow prediction model based on dynamical-learning graph convolution mechanism. Information Sciences, 611, 522-539.

10. Dai, L., Qin, W., Xu, H., Chen, T., & Qian, C. (2014, October). Urban traffic flow prediction: A MapReduce based parallel multivariate linear regression approach. In 17th

International IEEE Conference on Intelligent Transportation Systems (ITSC) (pp. 2823-2827). IEEE.

11. Song, Y., & Lu, J. (2018). RNN-based traffic flow prediction for dynamic reversible lane control decision. In Data Science and Knowledge Engineering for Sensing Decision Support: Proceedings of the 13th International FLINS Conference (FLINS 2018) (pp. 323-330).

12. Zhang, D., & Kabuka, M. R. (2018). Combining weather condition data to predict traffic flow: a GRU-based deep learning approach. IET Intelligent Transport Systems, 12(7), 578-585.

13. Cao, S., & Liu, W. (2019). Lstm network based traffic flow prediction for cellular networks. In Simulation Tools and Techniques: 11th International Conference, SIMUtools 2019, Chengdu, China, July 8–10, 2019, Proceedings 11 (pp. 643-653). Springer International Publishing.

14. Yang, D., Li, S., Peng, Z., Wang, P., Wang, J., & Yang, H. (2019). MF-CNN: traffic flow prediction using convolutional neural network and multi-features fusion. IEICE TRANSACTIONS on Information and Systems, 102(8), 1526-1536.

15. Chen, Y., Zou, X., Li, K., Li, K., Yang, X., & Chen, C. (2021). Multiple local 3D CNNs for region-based prediction in smart cities. Information Sciences, 542, 476-491.

16. Chen, Z., Zhao, B., Wang, Y., Duan, Z., & Zhao, X. (2020). Multitask learning and GCN-based taxi demand prediction for a traffic road network. Sensors, 20(13), 3776.

17. Ma, D., Song, X., & Li, P. (2020). Daily traffic flow forecasting through a contextual convolutional recurrent neural network modeling inter-and intra-day traffic patterns. IEEE Transactions on Intelligent Transportation Systems, 22(5), 2627-2636.

18. Ma, T., Antoniou, C., & Toledo, T. (2020). Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast. Transportation Research Part C: Emerging Technologies, 111, 352-372.

19. Du, S., Li, T., Gong, X., & Horng, S. J. (2020). A hybrid method for traffic flow forecasting using multimodal deep learning. International journal of computational intelligence systems, 13(1), 85-97.

20. Ma, C., Dai, G., & Zhou, J. (2021). Short-term traffic flow prediction for urban road sections based on time series analysis and LSTM_BILSTM method. IEEE Transactions on Intelligent Transportation Systems, 23(6), 5615-5624.

21. Zheng, H., Lin, F., Feng, X., & Chen, Y. (2020). A hybrid deep learning model with attention-based conv-LSTM networks for short-term traffic flow prediction. IEEE Transactions on Intelligent Transportation Systems, 22(11), 6910-6920.

22. Dong, S., Xia, Y., & Peng, T. (2021). Network abnormal traffic detection model based on semi-supervised deep reinforcement learning. IEEE Transactions on Network and Service Management, 18(4), 4197-4212.

23. Van Hasselt, H., Guez, A., & Silver, D. (2016, March). Deep reinforcement learning with double q-learning. In Proceedings of the AAAI conference on artificial intelligence (Vol. 30, No. 1).

24. Choi, H., Kim, M., Lee, G., & Kim, W. (2019). Unsupervised learning approach for network intrusion detection system using autoencoders. The Journal of Supercomputing, 75, 5597-5621.

25. Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. Electronics, 9(8), 1295.

26. Hardegen, C., Pfülb, B., Rieger, S., & Gepperth, A. (2020). Predicting network flow characteristics using deep learning and real-world network traffic. IEEE Transactions on Network and Service Management, 17(4), 2662-2676.

27. Yan, F., Qiu, J., & Tian, J. (2022). An iterative learning identification strategy for nonlinear macroscopic traffic flow model. Physica A: Statistical Mechanics and its Applications, 604, 127901.

28. Wang, Z., Chu, R., Zhang, M., Wang, X., & Luan, S. (2020). An improved selective ensemble learning method for highway traffic flow state identification. IEEE Access, 8, 212623-212634.

29. Billot, R., Faouzi, N. E. E., Sau, J., & De Vuyst, F. (2010). Integrating the impact of rain into traffic management: online traffic state estimation using sequential Monte Carlo techniques. Transportation research record, 2169(1), 141-149.

30. Khan, S. M., Dey, K. C., & Chowdhury, M. (2017). Real-time traffic state estimation with connected vehicles. IEEE Transactions on Intelligent Transportation Systems, 18(7), 1687-1699.

31. Wang, Z., Chu, R., Wu, W., Li, Q., Cai, Z., Cao, N., & Gu, M. (2018). Identification and optimization models for a freight-integrated transportation corridor with line importance and freight communication capability. Ieee Access, 7, 11114-11126.

32. Tian, G., Ren, Y., Feng, Y., Zhou, M., Zhang, H., & Tan, J. (2018). Modeling and planning for dual-objective selective disassembly using AND/OR graph and discrete artificial bee colony. IEEE Transactions on Industrial Informatics, 15(4), 2456-2468.

33. Li, L., He, S., Zhang, J., & Ran, B. (2016). Short-term highway traffic flow prediction based on a hybrid strategy considering temporal–spatial information. Journal of Advanced Transportation, 50(8), 2029-2040.

34. An, J., Fu, L., Hu, M., Chen, W., & Zhan, J. (2019). A novel fuzzy-based convolutional neural network method to traffic flow prediction with uncertain traffic accident information. Ieee Access, 7, 20708-20722.

35. Mou, L., Zhao, P., Xie, H., & Chen, Y. (2019). T-LSTM: A long short-term memory neural network enhanced by temporal information for traffic flow prediction. Ieee Access, 7, 98053-98060.

36. Chen, X., Cai, X., Liang, J., & Liu, Q. (2018). Ensemble learning multiple LSSVR with improved harmony search algorithm for short-term traffic flow forecasting. Ieee Access, 6, 9347-9357.

37. Ngoduy, D., & Liu, R. (2007). Multiclass first-order simulation model to explain non-linear traffic phenomena. Physica A: Statistical Mechanics and its Applications, 385(2), 667-682.

38. Acosta, C. D., Bürger, R., & Mejía, C. E. (2015). Efficient parameter estimation in a macroscopic traffic flow model by discrete mollification. Transportmetrica A: Transport Science, 11(8), 702-715.

39. Thonhofer, E., & Jakubek, S. (2018). Investigation of stochastic variation of parameters for a macroscopic traffic model. Journal of Intelligent Transportation Systems, 22(6), 547-564.

40. Zammit, L. C., Fabri, S. G., & Scerri, K. (2019). Real-time parametric modeling and estimation of urban traffic junctions. IEEE Transactions on Intelligent Transportation Systems, 20(12), 4579-4589.

41. Zammit, L. C., Fabri, S. G., & Scerri, K. (2017, July). Online state and parameter estimation for a 4-arm traffic junction. In 2017 25th Mediterranean Conference on Control and Automation (MED) (pp. 737-742). IEEE.

42. Ma, T., Antoniou, C., & Toledo, T. (2020). Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast. Transportation Research Part C: Emerging Technologies, 111, 352-372.

43. Lu, S., Zhang, Q., Chen, G., & Seng, D. (2021). A combined method for short-term traffic flow prediction based on recurrent neural network. Alexandria Engineering Journal, 60(1), 87-94.

44. Aljuaydi, F., Wiwatanapataphee, B., & Wu, Y. H. (2023). Multivariate machine learning-based prediction models of freeway traffic flow under non-recurrent events. Alexandria engineering journal, 65, 151-162.

45. Shekhar, S., & Williams, B. M. (2007). Adaptive seasonal time series models for forecasting short-term traffic flow. Transportation Research Record, 2024(1), 116-125.

46. Guo, J., Huang, W., & Williams, B. M. (2014). Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification. Transportation Research Part C: Emerging Technologies, 43, 50-64.

47. Zhou, B., He, D., & Sun, Z. (2006). Traffic modeling and prediction using ARIMA/GARCH model. In Modeling and Simulation Tools for Emerging Telecommunication Networks: Needs, Trends, Challenges and Solutions (pp. 101-121). Springer US.

48. Xian, Z. B., & Qiang, L. (2002). Arma-based traffic prediction and overload detection of network [j]. Journal of Computer Research and Development, 12.

49. Williams, B. M., & Hoel, L. A. (2003). Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. Journal of transportation engineering, 129(6), 664-672.

50. Thomas, T., Weijermars, W., & Van Berkum, E. (2009). Predictions of urban volumes in single time series. IEEE Transactions on Intelligent Transportation Systems, 11(1), 71-80.

51. Ma, C., Dai, G., & Zhou, J. (2021). Short-term traffic flow prediction for urban road sections based on time series analysis and LSTM_BILSTM method. IEEE Transactions on Intelligent Transportation Systems, 23(6), 5615-5624.

52. Gu, Y., Lu, W., Xu, X., Qin, L., Shao, Z., & Zhang, H. (2019). An improved Bayesian combination model for short-term traffic prediction with deep learning. IEEE Transactions on Intelligent Transportation Systems, 21(3), 1332-1342.

53. Zheng, J., & Ni, L. M. (2013, June). Time-dependent trajectory regression on road networks via multi-task learning. In Proceedings of the AAAI conference on artificial intelligence (Vol. 27, No. 1, pp. 1048-1055).

54. Deng, D., Shahabi, C., Demiryurek, U., Zhu, L., Yu, R., & Liu, Y. (2016, August). Latent space model for road networks to predict time-varying traffic. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1525-1534).

55. Liu, L., Zhen, J., Li, G., Zhan, G., He, Z., Du, B., & Lin, L. (2020). Dynamic spatial-temporal representation learning for traffic flow prediction. IEEE Transactions on Intelligent Transportation Systems, 22(11), 7169-7183.

56. Brockwell, P. J., & Davis, R. A. (Eds.). (2002). Introduction to time series and forecasting. New York, NY: Springer New York.

57. Zhang, J., Huang, T., Wang, S., & Liu, Y. J. (2019). Future Internet: trends and challenges. Frontiers of Information Technology & Electronic Engineering, 20(9), 1185-1194.

58. Yao, H., Mai, T., Jiang, C., Kuang, L., & Guo, S. (2019). AI routers & network mind: A hybrid machine learning paradigm for packet routing. IEEE computational intelligence magazine, 14(4), 21-30.

59. Rusek, K., Suárez-Varela, J., Mestres, A., Barlet-Ros, P., & Cabellos-Aparicio, A. (2019, April). Unveiling the potential of graph neural networks for network modeling and optimization in SDN. In Proceedings of the 2019 ACM Symposium on SDN Research (pp. 140-151).

60. Zhuang, Z., Wang, J., Qi, Q., Sun, H., & Liao, J. (2019). Toward greater intelligence in route planning: A graph-aware deep learning approach. IEEE Systems Journal, 14(2), 1658-1669.

61. Zhang, J., Zheng, Y., & Qi, D. (2017, February). Deep spatio-temporal residual networks for citywide crowd flows prediction. In Proceedings of the AAAI conference on artificial intelligence (Vol. 31, No. 1).

62. Zhang, J., Zheng, Y., Qi, D., Li, R., & Yi, X. (2016, October). DNN-based prediction model for spatio-temporal data. In Proceedings of the 24th ACM SIGSPATIAL international conference on advances in geographic information systems (pp. 1-4).

63. Xu, Z., Wang, Y., Long, M., Wang, J., & KLiss, M. (2018, July). PredCNN: Predictive Learning with Cascade Convolutions. In IJCAI (pp. 2940-2947).

64. Zhang, J., Zheng, Y., Sun, J., & Qi, D. (2019). Flow prediction in spatio-temporal networks based on multitask deep learning. IEEE Transactions on Knowledge and Data Engineering, 32(3), 468-478.

65. Cui, Z., Ke, R., Pu, Z., & Wang, Y. (2018). Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction. arXiv preprint arXiv:1801.02143.

66. Zhao, Z., Chen, W., Wu, X., Chen, P. C., & Liu, J. (2017). LSTM network: a deep learning approach for short-term traffic forecast. IET intelligent transport systems, 11(2), 68-75.

67. Gu, Y., Lu, W., Qin, L., Li, M., & Shao, Z. (2019). Short-term prediction of lane-level traffic speeds: A fusion deep learning model. Transportation research part C: emerging technologies, 106, 1-16.

68. Bao, X., Jiang, D., Yang, X., & Wang, H. (2021). An improved deep belief network for traffic prediction considering weather factors. Alexandria Engineering Journal, 60(1), 413-420.

69. Qiao, W., Haghani, A., & Hamedi, M. (2012). Short-term travel time prediction considering the effects of weather. Transportation research record, 2308(1), 61-72.

70. Yu, F., Wei, D., Zhang, S., & Shao, Y. (2019, September). 3D CNN-based accurate prediction for large-scale traffic flow. In 2019 4th International Conference on Intelligent Transportation Engineering (ICITE) (pp. 99-103). IEEE.

71. Chen, Z., Zhao, B., Wang, Y., Duan, Z., & Zhao, X. (2020). Multitask learning and GCN-based taxi demand prediction for a traffic road network. Sensors, 20(13), 3776.

72. Guo, S., Lin, Y., Feng, N., Song, C., & Wan, H. (2019, July). Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In Proceedings of the AAAI conference on artificial intelligence (Vol. 33, No. 01, pp. 922-929).

73. He, K., Huang, Y., Chen, X., Zhou, Z., & Yu, S. (2019, December). Graph attention spatial-temporal network for deep learning based mobile traffic prediction. In 2019 IEEE Global Communications Conference (GLOBECOM) (pp. 1-6). IEEE.

74. Cheng, X., Zhang, R., Zhou, J., & Xu, W. (2018, July). Deeptransport: Learning spatial-temporal dependency for traffic condition forecasting. In 2018 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.

75. Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2017). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. arXiv preprint arXiv:1707.01926.

76. Tian, Y., Zhang, K., Li, J., Lin, X., & Yang, B. (2018). LSTM-based traffic flow prediction with missing data. Neurocomputing, 318, 297-305.

77. Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. Mathematics of control, signals and systems, 2(4), 303-314.

78. Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. Neural networks, 2(5), 359-366.

79. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.

80. Kalchbrenner, N., Oord, A., Simonyan, K., Danihelka, I., Vinyals, O., Graves, A., & Kavukcuoglu, K. (2017, July). Video pixel networks. In International Conference on Machine Learning (pp. 1771-1779). PMLR.

81. Lotter, W., Kreiman, G., & Cox, D. (2016). Deep predictive coding networks for video prediction and unsupervised learning. arXiv preprint arXiv:1605.08104.

82. Wang, Y., Long, M., Wang, J., Gao, Z., & Yu, P. S. (2017). Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms. Advances in neural information processing systems, 30.

83. Funahashi, K. I., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. Neural networks, 6(6), 801-806.

84. Fu, R., Zhang, Z., & Li, L. (2016, November). Using LSTM and GRU neural network methods for traffic flow prediction. In 2016 31st Youth academic annual conference of Chinese association of automation (YAC) (pp. 324-328). IEEE.

85. Mackenzie, J., Roddick, J. F., & Zito, R. (2018). An evaluation of HTM and LSTM for short-term arterial traffic flow prediction. IEEE Transactions on Intelligent Transportation Systems, 20(5), 1847-1857.

86. Polson, N. G., & Sokolov, V. O. (2017). Deep learning for short-term traffic flow prediction. Transportation Research Part C: Emerging Technologies, 79, 1-17.

87. Zhao, J., Gao, Y., Yang, Z., Li, J., Feng, Y., Qin, Z., & Bai, Z. (2019). Truck traffic speed prediction under non-recurrent congestion: Based on optimized deep learning algorithms and GPS data. IEEE Access, 7, 9116-9127.

88. Lee, S., Ngoduy, D., & Keyvan-Ekbatani, M. (2019). Integrated deep learning and stochastic car-following model for traffic dynamics on multi-lane freeways. Transportation research part C: emerging technologies, 106, 360-377.

89. Dabiri, S., & Heaslip, K. (2019). Developing a Twitter-based traffic event detection model using deep learning architectures. Expert systems with applications, 118, 425-439.

90. Luong, M. T. (2015). Effective approaches to attention-based neural machine translation. arXiv preprint arXiv:1508.04025.

91. Graves, A., Mohamed, A. R., & Hinton, G. (2013, May). Speech recognition with deep recurrent neural networks. In 2013 IEEE international conference on acoustics, speech and signal processing (pp. 6645-6649). Ieee.

92. Dai, G., Tang, J., & Luo, W. (2023). Short-term traffic flow prediction: an ensemble machine learning approach. Alexandria Engineering Journal, 74, 467-480.

93. Li, L., Qu, X., Zhang, J., Li, H., & Ran, B. (2018). Travel time prediction for highway network based on the ensemble empirical mode decomposition and random vector functional link network. Applied Soft Computing, 73, 921-932.

94. Wu, Y., Tan, H., Qin, L., Ran, B., & Jiang, Z. (2018). A hybrid deep learning based traffic flow prediction method and its understanding. Transportation Research Part C: Emerging Technologies, 90, 166-180.

95. Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2005). Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach. Transportation Research Part C: Emerging Technologies, 13(3), 211-234.

96. Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2017). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. arXiv preprint arXiv:1707.01926.

97. Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., & Li, H. (2019). T-GCN: A temporal graph convolutional network for traffic prediction. IEEE transactions on intelligent transportation systems, 21(9), 3848-3858.

98. Guo, J., Song, C., & Wang, H. (2019, November). A multi-step traffic speed forecasting model based on graph convolutional LSTM. In 2019 Chinese Automation Congress (CAC) (pp. 2466-2471). IEEE.

99. Luo, X., Li, D., Yang, Y., & Zhang, S. (2019). Spatiotemporal traffic flow prediction with KNN and LSTM. Journal of Advanced Transportation, 2019(1), 4145353.

100. Dai, G., Ma, C., & Xu, X. (2019). Short-term traffic flow prediction method for urban road sections based on space–time analysis and GRU. IEEE Access, 7, 143025-143035.

101. Cortes, C. (1995). Support-Vector Networks. Machine Learning.

102. Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. Advances in neural information processing systems, 9.

103. Schölkopf, B., Burges, C. J., & Smola, A. J. (Eds.). (1999). Advances in kernel methods: support vector learning. MIT press.

104. Müller, K. R., Smola, A. J., Rätsch, G., Schölkopf, B., Kohlmorgen, J., & Vapnik, V. (1997, October). Predicting time series with support vector machines. In International conference on artificial neural networks (pp. 999-1004). Berlin, Heidelberg: Springer Berlin Heidelberg.

105. Yang, W., Yang, D., Zhao, Y., & Gong, J. (2010, January). Traffic flow prediction based on wavelet transform and radial basis function network. In 2010 International Conference on Logistics Systems and Intelligent Management (ICLSIM) (Vol. 2, pp. 969-972). IEEE.

106. Mercer, J. (1909). Xvi. functions of positive and negative type, and their connection the theory of integral equations. Philosophical transactions of the royal society of London. Series A, containing papers of a mathematical or physical character, 209(441-458), 415-446.

107. Bertsekas, D. P. (1997). Nonlinear programming. Journal of the Operational Research Society, 48(3), 334-334.

108. Fletcher, R. (2000). Practical methods of optimization. John Wiley & Sons.

109. Herbrich, R. (2001). Learning kernel classifiers: theory and algorithms. MIT press.

110. Scholkopf, B., & Smola, A. J. (2018). Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press.

111. Pérez-Cruz, F., Camps-Valls, G., Soria-Olivas, E., Pérez-Ruixo, J. J., Figueiras-Vidal, A. R., & Artés-Rodríguez, A. (2002, August). Multi-dimensional function approximation and regression estimation. In International conference on artificial neural networks (pp. 757-762). Berlin, Heidelberg: Springer Berlin Heidelberg.

112. Yang, X. S. (2010). A new metaheuristic bat-inspired algorithm. In Nature inspired cooperative strategies for optimization (NICSO 2010) (pp. 65-74). Berlin, Heidelberg: Springer Berlin Heidelberg.

113. Ngoduy, D., & Liu, R. (2007). Multiclass first-order simulation model to explain non-linear traffic phenomena. Physica A: Statistical Mechanics and its Applications, 385(2), 667-682.