Deep Learning-Based Traffic Information Prediction Methods in the Internet of Vehicles

Chenguang He(✉), Bohan Zhang, Liang Ye, and Hua Tan

School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China
{hechenguang,yeliang}@hit.edu.cn, {23S005088, 1201051820}@stu.hit.edu.cn

Abstract. In recent years, the rapid development of deep learning technology has provided many methods to predict traffic information prediction. In the Internet of Vehicles (IoV), accurate and real-time traffic information prediction plays an important role in improving system performance and user experience. How to effectively capture the temporal and spatial dependencies of traffic information is a major challenge in this field. In this paper, we focus on three neural network models (GRU, TGCN and TGCN-att) for traffic information prediction and train these three models using real datasets. We analyzed the outputs of each model separately, compared the performance metrics such as mean square error (RMSE) and mean absolute error (MAE) between the predicted and real values, and calculated the accuracy of the predictions of each model. The simulation results show that since road networks generally have a complex topology, correctly capturing the spatial dependence between data is very important for improving the prediction accuracy of the models when performing traffic information prediction.

Keywords: Internet of Vehicles · Deep Learning · Neural Network · Traffic Information Prediction · Spatial and Temporal Dependence

1 Introduction

Internet of Vehicles (IoV) is a kind of intelligent transportation system based on the Internet and sensor technology, which can realize data transmission and information interaction between vehicles and vehicles (V2V) and between vehicles and infrastructure (V2I) by acquiring relevant information about vehicles, roads and their surroundings. With the continuous development of IoV technology, its requirements for real-time and accuracy of traffic information are increasing [1]. However, traditional traffic information prediction methods (e.g., historical averaging [2], etc.) are difficult to adapt to today's complex and changing traffic environment, and their accuracy is poor, which cannot meet the needs of IoV applications. In contrast, deep learning technology, as a powerful feature learning and data processing tool, is able to capture the correlations and dependencies between data from massive historical datasets, and has a significant advantage in handling complex and large-scale data. Therefore, deep learning technology can be applied to IoV system so as to realize the real-time prediction of traffic information.

H.-H. Chen and W. Meng (Eds.): WiSATS 2024, LNICST 606, pp. 53–63, 2025.
https://doi.org/10.1007/978-3-031-86203-8_5

Existing models for predicting traffic information can be categorized into parametric and nonparametric models [3]. Parametric models assume that the data follow some known distribution and that the distribution has a set of unknown parameters, and the values of these unknown parameters can be determined by fitting to the observed data. Non-parametric models, on the other hand, model the data directly without making any assumptions about the distribution of the data, and such models can be used to deal with complex nonlinear relationships. Common parametric models include linear regression models, logistic regression models, time series models, etc., which are suitable for situations with low complexity and small amounts of data. Common non-parametric models include decision trees, support vector machines, neural network models, etc. Compared with parametric models, they can be more flexible to deal with complex data patterns, and can still achieve good prediction results when the a priori information is unknown.

As early as in the twentieth century, Hamed et al. [4] proposed the use of autoregressive integral moving average model (ARIMA) to predict the traffic flow of urban arterial roads. Sun et al. [5] solved the interval prediction problem by using a local linear model. Lippi et al. [6] effectively solved the problem of traffic congestion by using a support vector machine to predict traffic information. Ojeda et al. [7] accomplished the traffic prediction task by using the Kalman filtering.

In recent years, with the rapid development of deep learning technology, neural network models have gradually been emphasized by researchers. Huang et al. [8] proposed a deep belief network (DBN) model that can capture random features of data from multiple datasets. Qi et al. [9] proposed a robust hierarchical deep learning method that can extract deep semantic features of data for the traffic congestion detection task. However, these models only consider the temporal features of the data and ignore the constraints of urban road networks on traffic information data, which leads to less accurate prediction results.

In order to fully utilize the spatial and temporal correlation of the data, Lv et al. [10] proposed a SAE model that can effectively capture the spatial and temporal characteristics of traffic data to achieve short-term traffic flow prediction. Wu and Tan [11] captured the spatial dependence and temporal dependence of the traffic flow data by using Convolutional Neural Networks (CNN) and Long Short-Term Memory Neural Networks (LSTM), respectively. Yu et al. [12] proposed a neural network model called SRCN, which combines the advantages of both deep convolutional neural networks (DCNN) and LSTM to effectively capture the spatial and temporal dependence of urban traffic data.

In this paper, we train and test three representative neural network models using real datasets, and compare and analyze the outputs of the three models. The simulation results show that, due to the constraints of the urban road network structure, it is not enough to consider only the temporal characteristics of the data when performing traffic information prediction, but also the spatial dependence of the data, which is very important to improve the prediction accuracy of the models.

2 System Model

2.1 Road Network Topology

Urban roads generally have complex topological structures, and these structures determine to some extent the traveling characteristics of vehicles, including speed and direction. Graph theory is an important branch of applied mathematics, which can provide an abstracted way to describe and analyze the network structure, so we can model the road structure according to the viewpoint of graph theory. In IoV system, roads can be regarded as nodes in the network, while the connectivity between roads can be represented as edges, so that we can describe the topology of the road network in a graphical way.

We use graph $G = (N_s, E_s)$ to represent the road network within a certain area, where N_s is the node set and E_s is the edge set. Based on this, we define two matrices $\boldsymbol{A} = [a_{ij}]_{N\times N}$ and $\boldsymbol{X} = [x_{ij}]_{T\times N}$ to characterize the road network. Here, N is the total number of nodes, i.e., the number of road strips within the range; T is the total time length.

We call the matrix $\boldsymbol{A}$ the adjacency matrix of the graph G, which is used to represent the connectivity between the roads within the region. The element a_{ij} in matrix $\boldsymbol{A}$ is binary and is defined as follows:

$$a_{ij} = \begin{cases} 1, & if\ i\ is\ connected\ to\ j \\ 0, & if\ i\ is\ not\ connected\ to\ j \end{cases} \tag{1}$$

We call the matrix $\boldsymbol{X}$ the feature matrix of the graph G, which is used to represent the traffic information on each road at different moments. In this paper, we consider the speed information of vehicles, i.e., the element x_{ij} in matrix $\boldsymbol{X}$ represents the average speed of all vehicles on the ith road at the jth time node.

If we know the information about the speed of vehicles on each road in the past period, then the average speed of vehicles on these roads in the future period can be predicted by deep learning methods. In the next subsections, we highlight several neural network models for traffic information prediction.

2.2 GRU Model

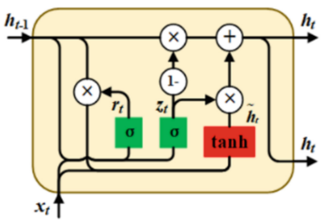

Fig. 1. Basic structure of GRU.

Gated Recurrent Unit (GRU) is a common Recurrent Neural Network (RNN) variant for processing sequential data. Compared to standard RNN, GRU introduces the concepts of update gates and reset gates. These gating mechanisms allow GRU to handle long-term dependencies between data more efficiently while mitigating the problems of gradient vanishing and gradient exploding. The basic structure of a GRU unit is shown in Fig. 1.

Assuming that the current time step is t, the input is $\boldsymbol{x}_t$, the hidden state of the previous time step is $\boldsymbol{h}_{t\text{-}1}$, and the hidden state of the current time step is $\boldsymbol{h}_t$, the formulas for the update gate, the reset gate, the candidate memory, and the hidden state of the current time step of GRU are, respectively:

$$z_t = \sigma\left(\boldsymbol{W}_z \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_z\right) \tag{2}$$

$$\boldsymbol{r}_t = \sigma\left(\boldsymbol{W}_r \cdot \left[\boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}_r\right) \tag{3}$$

$$\widetilde{\boldsymbol{h}}_t = \tanh\left(\boldsymbol{W} \cdot \left[\boldsymbol{r}_t \odot \boldsymbol{h}_{t-1}, \boldsymbol{x}_t\right] + \boldsymbol{b}\right) \tag{4}$$

$$\boldsymbol{h}_t = (1 - z_t) \odot \boldsymbol{h}_{t-1} + z_t \odot \widetilde{\boldsymbol{h}}_t \tag{5}$$

where σ denotes the sigmoid activation function, tanh denotes the hyperbolic tangent activation function, $\odot$ denotes the element-by-element multiplication of vectors, $\boldsymbol{W}$ and $\boldsymbol{b}$ are the weight matrix and bias vector, respectively.

GRU is able to effectively process sequential data with time dependencies, and therefore can be used for traffic information prediction. In addition to GRU, Long Short-Term Memory Neural Network (LSTM) also has the ability to capture long-term dependencies between data, however, LSTM has more network parameters, more complex structure, and the performance is not much improved compared to GRU, therefore, in this paper, we choose GRU to capture temporal dependencies between data.

2.3 TGCN Model

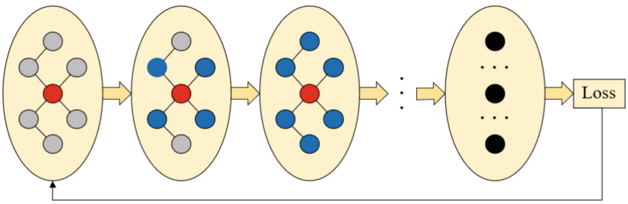

Fig. 2. Basic structure of GCN.

In IoV system, the road structure will determine the traffic information of vehicles to a certain extent, therefore, capturing and utilizing the network spatial features correctly is also particularly important for traffic information prediction. Traditional Convolutional Neural Network (CNN) can effectively capture spatial dependencies when processing

2D data (e.g., images), however, this capability exists only in Euclidean space. For urban road networks with complex topologies, Graph Convolutional Neural Network (GCN), which is capable of processing arbitrary structured data, is required to capture spatial dependencies between data. The basic structure of GCN is shown in Fig. 2.

The propagation rule for each convolutional layer is:

$$\boldsymbol{H}^{(l+1)} = \sigma\left(\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\boldsymbol{H}^{(l)}\boldsymbol{W}^{(l)}\right) \tag{6}$$

where $\tilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{I}$ is the adjacency matrix that increases the self-loop and $\tilde{\boldsymbol{D}}^{D}$ is the degree matrix of $\tilde{\boldsymbol{A}}$.

GCN can effectively capture the spatial dependence of data, and we input the output of GCN at the tth time step $\boldsymbol{H}_t$ into GRU to realize the propagation and updating of information in both time and space dimensions. Neural network with this structure is called Time-Graph Convolutional Neural Network (TGCN). Compared with GRU mentioned in the previous subsection, TGCN incorporates information in both temporal and spatial dimensions, and is able to better process and analyze complex graph structures, making it an effective tool for traffic information prediction.

2.4 TGCN-Att Model

In deep learning, the attention mechanism is often used to extract key information from the input data. The attention mechanism allows the neural network model to dynamically assign weights, with different weights representing different levels of importance, when processing structured data. Thus, this mechanism can consciously focus attention on certain important data while ignoring other unimportant components [13].

Adding the attention mechanism to the TGCN model can enhance the model's attention to the important nodes and edges, so as to capture the key information in the graph data more accurately, better analyze the complex graph structure, and then improve the model's expressive ability and predictive performance. The basic structure of TGCN based on the attention mechanism is shown in Fig. 3.

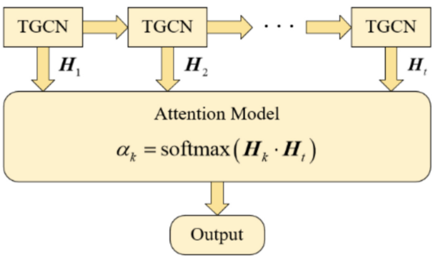

Fig. 3. Basic structure of TGCN-att.

In this paper, we refer to this model as TGCN-att model. Its output expression at the tth time step is:

$$\boldsymbol{H}_t = \sigma\left(\sum_{k=1}^{T} \alpha_k \cdot \boldsymbol{A}_k \boldsymbol{X}_k \boldsymbol{W} + \boldsymbol{b}\right) \tag{7}$$

where α_k is the attention weight of the kth time step, which is computed by the attention mechanism. The calculation method we use is dot product attention calculation method:

$$\alpha_k = \text{softmax}(\boldsymbol{H}_k \cdot \boldsymbol{H}_t) \tag{8}$$

3 Experiments

3.1 Dataset

We use a real dataset to train and test the three neural network models mentioned in Sect. 2. The dataset comes from literature [1], which records the traffic information of 156 main roads in Luohu District, Shenzhen over a period of time. In this paper, traffic information specifically refers to the traveling speed of vehicles, which is recorded every 15 min. The size of the adjacency matrix in the dataset is 156 × 156, which is used to represent the connectivity between the roads. The size of the feature matrix is 2976 × 156, so the total length of time it records is 31 days.

3.2 Parameter Settings

In our experiment, we first normalize the vehicle speed magnitude in the dataset into the interval [0,1], and then select the first 80% of data for training the models, and the last 20% of data for testing the prediction performance of the models. The hyperparameters of the neural network model mainly include learning rate, batch size, training epoch and number of hidden units. Based on the historical experience, we set the learning rate size to 0.001 and the training epoch to 1000. Considering the memory limitation of the hardware device and the generalization ability of the model, we set the batch size to 32. For hidden units, too few may cause the model to fail to capture complex patterns and structures in data, and too many may cause the model to overfit the training data and fail to generalize to other samples. Therefore, considering these factors, we set the number of hidden units to 100.

During model training, our goal is to minimize the error between the predicted and true values, and the loss function can be expressed as:

$$\text{Loss} = \left\| Y - \widehat{Y} \right\| + \lambda L_{reg} \tag{9}$$

where Y and $\hat{Y}$ are the true and predicted values of vehicle speed, respectively, L_{reg} is the L2 regularization term, and λ is a hyperparameter.

We strictly follow the principle of controlling variables during the experiments and only replace the neural network model without changing the relevant parameters of the model. In Sect. 4, we compare the prediction results of the three models and analyze their performance on the test set.

4 Simulation Results Analysis

4.1 Evaluation Metrics

We use the following three metrics to evaluate the prediction performance of the neural network models.

(1) Root Mean Square Error (RMSE):

$$\mathrm{RMSE} = \sqrt{\frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(y_{ij} - \hat{y}_{ij}\right)^2} \tag{10}$$

(2) Mean Absolute Error (MAE):

$$\mathrm{MAE} = \frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}\left|y_{ij} - \hat{y}_{ij}\right| \tag{11}$$

(3) Accuracy:

$$\mathrm{Accuracy} = 1 - \frac{\left\|Y - \hat{Y}\right\|_F}{\|Y\|_F} \tag{12}$$

where M and N are the number of time samples and the number of roads, respectively; y_{ij} and $\hat{y}_{ij}$ are the true and predicted values of vehicle speed at the jth time node on the ith road, respectively; and Y and $\hat{Y}$ denote the set of y_{ij} and $\hat{y}_{ij}$, respectively. The smaller the RMSE and the MAE, the smaller the difference between the predicted value and the true value, the higher the prediction accuracy of the model, and the better the prediction performance.

4.2 Prediction Results and Error Analysis

Table 1. Prediction error and accuracy of three neural network models.

Metric	GRU	TGCN	TGCN-att
RMSE	4.110179	4.092278	4.058892
MAE	2.685558	2.764561	2.736572
Accuracy	0.713611	0.714859	0.717185

Figures 4, 5 and 6 show the prediction results of GRU, TGCN and TGCN-att compared with the actual vehicle speed in the coming day, respectively. It can be seen that all three models are able to fit the trend of vehicle speed well. However, the prediction error of GRU is larger compared to TGCN and TGCN-att because it does not take into account the spatial dependence of data. This is better illustrated by the data in Table 1. Overall,

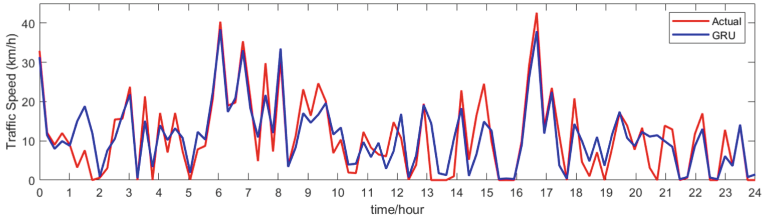

Fig. 4. Prediction results of GRU.

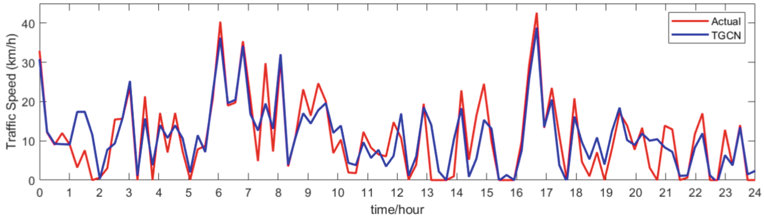

Fig. 5. Prediction results of TGCN.

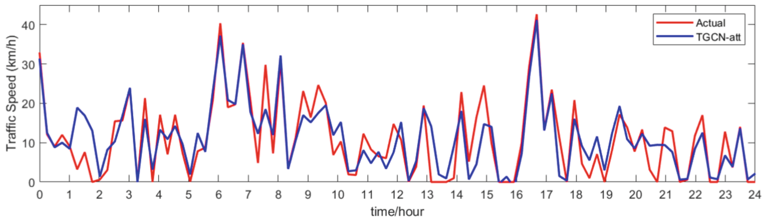

Fig. 6. Prediction results of TGCN-att.

the prediction of TGCN-att is better than that of GRU and TGCN, which indicates that TGCN-att improves the global feature capture ability of the model by introducing the attention mechanism.

Next, we will further analyze the error performance of the three neural networks.

Figure 7 shows the changes of loss function values of the three neural networks during the training process. GRU has a relatively simple structure, which makes it easier to train and optimize, so it is able to obtain a smaller loss function value at the beginning of the training, and its size can be kept stable as the training process proceeds. The loss function values of both TGCN and TGCN-att show a trend of decreasing and then stabilizing. The difference is that the convergence speed of TGCN-att is faster, because the attention mechanism can help the model better capture the important features of the input data, and reduce the burden of the model to deal with redundant data. In addition, the attention mechanism can dynamically adjust the attention weights to adapt to different input data, thus making the model more adaptive. When the number of training times is sufficiently high, the loss function value of TGCN-att is significantly smaller than that of GRU and

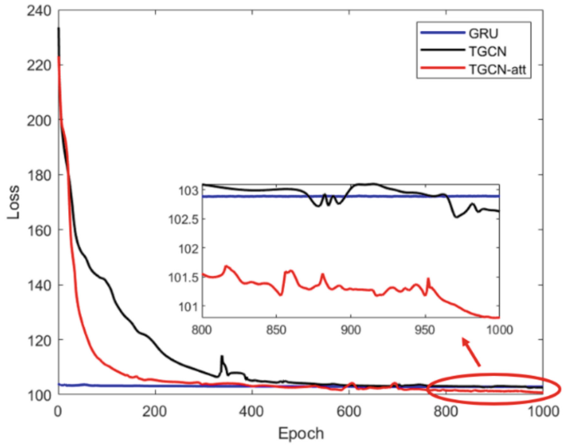

Fig. 7. Changes of loss function values of the three neural networks during the training process.

TGCN, which also illustrates from another perspective that TGCN-att can fit the training data better, thus obtaining better prediction performance.

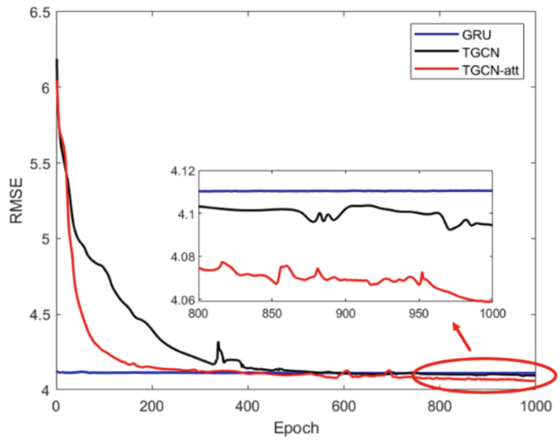

Fig. 8. Changes of RMSE of the three neural networks during the training process.

Figure 8 shows the changes of RMSE of the three neural networks during the training process. Since there are almost no outliers or anomalies in the dataset we used, the general trends of RMSE and MAE are the same, and here we choose only one of them for analysis. It can be seen that the RMSE of GRU is the smallest at the beginning of the training period, which is due to the fact that it has fewer network parameters, the simplest network structure, and that GRU can only capture the temporal correlation of data. TGCN and TGCN-att, on the other hand, contain graph convolutional layers and modeling of spatial properties, and their structures are relatively complex, thus generating a large RMSE at the beginning of the training period. As the training process proceeds, the RMSE of

TGCN and TGCN-att gradually decrease, and eventually tend to be stabilized. In the steady state, the RMSE of TGCN-att is the smallest, TGCN is the second, and GRU is the largest. This indicates that considering the spatial dependence of data can improve the prediction performance of the neural network model, in addition, since TGCN-att incorporates the attention mechanism, it can capture the key information in the graph data more effectively, which leads to a smaller prediction error of the model.

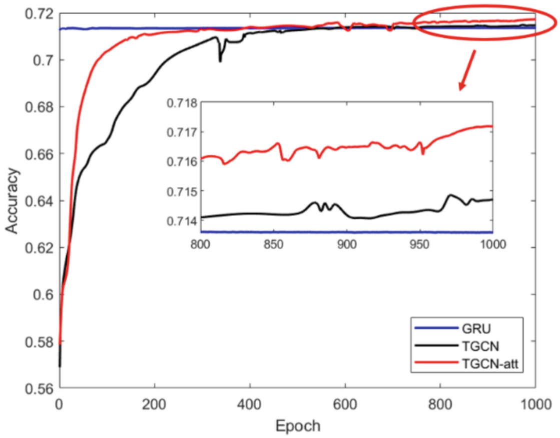

Fig. 9. Changes of prediction accuracy of the three neural networks during the training process.

Figure 9 shows the changes of prediction accuracy of the three neural networks during the training process. The prediction accuracy of TGCN and TGCN-att increases gradually with the training process, and finally exceeds GRU and reaches a stable value, which is consistent with the results of the previous error analysis. Figure 9 still illustrates that considering spatial and temporal dependencies is very important for traffic information prediction, and at the same time, the introduction of the attention mechanism can further reduce the prediction error and improve the prediction accuracy of the model.

5 Conclusion

In this paper, three neural network models, GRU, TGCN and TGCN-att, are used for traffic information prediction, and the prediction performance of the three models is compared and analyzed. Among them, GRU can only capture the temporal dependence of data, while TGCN can not only use GRU to obtain temporal dependence, but also use GCN to obtain spatial dependence. Adding the attention mechanism on the basis of TGCN can obtain the TGCN-att model, which is able to obtain a better prediction performance by dynamically adjusting the attention weights. The experimental results show that when the number of training times is sufficiently high, the prediction error of GRU, TGCN and TGCN-att decreases and the accuracy increases in turn, which demonstrates that it is necessary to consider the topology of the road network in traffic information prediction, because the complex road structure will make the traffic data have a strong spatial dependence on each other, and at the same time, this also proves that the attention mechanism can make the neural network model have better data processing ability, so it can further improve the prediction performance of the model.

Acknowledgement. This work is supported by the Key R&D Program of Heilongjiang Province under Grant JD22A001.

References

1. Zhao, L., et al.: T-GCN: a temporal graph convolutional network for traffic prediction. IEEE Trans. Intell. Transp. Syst. **21**(9), 3848–3858 (2020)
2. Liu, J., Guan, W.: A summary of traffic flow forecasting methods. J. Highway Transp. Res. Develop. **21**(3), 82–85 (2004)
3. van Lint, H., van Hinsbergen, C.: Short-term traffic and travel time prediction models. Artif. Intell. Appl. Critical Transp. **22**, 22–41 (2012)
4. Hamed, M.M., Al-Masaeid, H.R., Said, Z.M.B.: Short-term prediction of traffic volume in Urban arterials. J. Transp. Eng. **121**(3), 249–254 (1995)
5. Sun, H., Zhang, C., Ran, B.: Interval prediction for traffic time series using local linear predictor. In: Proceedings of 7th International IEEE Conference on Intelligent Transportation Systems, Washington, DC, USA, pp. 410–415 (2004)
6. Lippi, M., Bertini, M., Frasconi, P.: Short-term traffic flow forecasting: an experimental comparison of time-series analysis and supervised learning. IEEE Trans. Intell. Transp. Syst. **14**(2), 871–882 (2013)
7. Ojeda, L.L., Kibangou, A.Y., de Wit, C.C.: Adaptive Kalman filtering for multi-step ahead traffic flow prediction In: Proceedings under *American Control Conference*, Washington, DC, USA, pp. 4724–4729 (2013)
8. Huang, W., Song, G., Hong, H., Xie, K.: Deep architecture for traffic flow prediction: deep belief networks with multitask learning. IEEE Trans. Intell. Transp. Syst. **15**(5), 2191–2201 (2014)
9. Qi, W., Jia, W., Xuelong, L.: Robust hierarchical deep learning for vehicular management. IEEE Trans. Veh. Technol. **68**(5), 4148–4156 (2019)
10. Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F.-Y.: Traffic flow prediction with big data: a deep learning approach. IEEE Trans. Intell. Transp. Syst.Intell. Transp. Syst. **16**(2), 865–873 (2015)
11. Wu, Y., Tan, H.: Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework (2016). https://arxiv.org/abs/1612.01022
12. Yu, H., Wu, Z., Wang, S., Wang, Y., Ma, X.: Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks. Sensors **17**(7), 1501 (2017)
13. Bai, J., et al.: A3T-GCN: attention temporal graph convolutional network for traffic forecasting. ISPRS Int. J. Geo-Inf. **10**, 485 (2021)