PAPER

MF-CNN: Traffic Flow Prediction Using Convolutional Neural Network and Multi-Features Fusion

Di YANG[†], *Student Member*, Songjiang LI[†], Zhou PENG[†], Peng WANG[†], Junhui WANG^{††},
and Huamin YANG^{†a)}, *Nonmembers*

SUMMARY Accurate traffic flow prediction is the precondition for many applications in Intelligent Transportation Systems, such as traffic control and route guidance. Traditional data driven traffic flow prediction models tend to ignore traffic self-features (e.g., periodicities), and commonly suffer from the shifts brought by various complex factors (e.g., weather and holidays). These would reduce the precision and robustness of the prediction models. To tackle this problem, in this paper, we propose a CNN-based multi-feature predictive model (MF-CNN) that collectively predicts network-scale traffic flow with multiple spatiotemporal features and external factors (weather and holidays). Specifically, we classify traffic self-features into temporal continuity as short-term feature, daily periodicity and weekly periodicity as long-term features, then map them to three two-dimensional spaces, which each one is composed of time and space, represented by two-dimensional matrices. The high-level spatiotemporal features learned by CNNs from the matrices with different time lags are further fused with external factors by a logistic regression layer to derive the final prediction. Experimental results indicate that the MF-CNN model considering multi-features improves the predictive performance compared to five baseline models, and achieves the trade-off between accuracy and efficiency.

key words: Intelligent Transportation Systems, traffic flow prediction, convolutional neural networks, multiple spatiotemporal features, external factors

1. Introduction

With the rapid urban growth, traffic problems, such as traffic congestion, traffic accidents, and air pollution, have become the constraints of economic development and the bottleneck of social progress in the world. Fortunately, Intelligent Transportation Systems (ITS) can help people to solve these issues by processing large amounts of traffic data to provide traffic control, route guidance, road condition warning, etc. Among these applications, accurate traffic flow prediction is the prerequisite of the implementation, which can provide convenience for traffic management, improve public travel quality, relieve road congestion, and reduce the accident rate.

Traffic flow is defined as the total number of vehicles passing over a given section or point of a road during a given time interval. The prediction of traffic flow has

a long history in ITS, attracting numerous attention from worldwide researchers. In general, there are two very effective models in traffic flow prediction: (1) statistical models, which focus on analyzing and forecasting at a single point based on temporal and spatial series from historical data of traffic flow, such as autoregressive integrated moving average model (ARIMA) [1]–[3], Kalman filter [4], grey theory [5], and stochastic model [6]. The statistical models are characterized by simple computing and high efficiency, however, they show weakness in dealing with non-linearity within large-scale network data and the accuracy decreases in the face of complex road conditions; (2) intelligent computational models, represented by neural networks [7], Bayesian networks [8], fuzzy and evolutionary techniques [9], have obtained remarkable achievements in traffic prediction. Compared with statistical models, intelligent computational models have superior advantages in prediction accuracy and nonlinear processing ability [10]. However, with rich amount of traffic data, most of them are not entirely satisfying due to the shallow structure.

In recent year, deep learning [11] has been paid more attention to establish traffic flow prediction models by using deep neural networks, which can perform better multiple representations and abstractions with less prior knowledge. For example, Lv et al. utilized a stacked autoencoders (SAE) model for traffic flow prediction [12]. Polson et al. improved the structure of the deep neural network with a linear model to forecast traffic flow [13]. Ma et al. proposed a novel traffic forecast model based on long short-term memory (LSTM) network for traffic speed prediction [14]. Unfortunately, the majority of existing models aim at the traffic prediction of a single road. It might ignore part spatial features and temporal features in the traffic network. Network-scale traffic, which is more effective and practical for understanding traffic regularity, should be paid more attention. How to extract spatial and temporal features efficiently is the primary consideration. Convolutional neural network (CNN) is one of the effective methods to fill this gap.

CNN, in the community of deep learning, has achieved great success in computer vision tasks, image processing, and natural language processing (NLP) tasks with the natural advantage of preserving neighborhood relations and spatial locality [15], [16]. Motivated by the remarkable performance of CNN, several attempts have been made to apply CNN in traffic prediction tasks. Zhang et al. applied CNN

Manuscript received September 15, 2018.

Manuscript revised February 7, 2019.

Manuscript publicized May 20, 2019.

[†]The authors are with School of Computer Science and Technology, Changchun University of Science and Technology, Changchun, 130022 China.

^{††}The author is with Jilin Provincial Expressway Administration, Changchun, 130022 China.

a) E-mail: yhm.cust@gmail.com (Corresponding author)

DOI: 10.1587/transinf.2018EDP7330

and deep residual network to predict citywide crowd flows, achieving satisfying results [17]. Ma et al. represented continuous road speed by a time-space matrix, which learned by CNN to predict traffic speed [18]. Yu et al. employed spatiotemporal recurrent convolutional networks to predict the traffic speed of urban roads, which utilized CNN to mine the space features in traffic network, whereas LSTM to learn the temporal features of traffic congestion evolution [19]. The previous researches indicate that CNN has shown promising performance in learning traffic data.

However, in most of researches, the prediction models suffer from the shifts brought by complex factors. The self-features of traffic flow and external factors that directly affect observed values are not fully considered. In time dimension, traffic flow not only exhibits temporal continuity, but also observed periodicities in one day and one week. Periodicities can help better understand traffic phenomena, and find out the travel laws. In space dimension, traffic flow of one road is not only related to its neighbors but also far away roads according to the theory of propagation. Meanwhile, adverse weather and holidays are two main factors that increase both travel time and crash risk, thus affecting traffic flow. Without a comprehensive understanding of different traffic features, it will be difficult to prevent the prediction models from the shifts brought by complex factors. How to involve more realistic conditions to improve prediction accuracy and robustness of the traffic flow model is an important issue worthy of investigation.

Inspired by these successful researches, in this paper, we propose a deep learning model, named MF-CNN, that collectively considers spatiotemporal features and external factors in the short-term traffic flow prediction, which aims at preventing the prediction model from the complex factors' shifts, utilizing the available traffic features to improve the prediction accuracy. There are two stages in the MF-CNN model. In the first stage, besides temporal continuity as short-term feature, we consider two periodicities, namely daily periodicity and weekly periodicity observed in traffic flow as two long-term features. Then we map the combination of temporal features and spatial features to three two-dimensional spaces, each one is composed of time and space, represented by two-dimensional matrices. In the second stage, a comprehensive prediction model that incorporates three CNNs and a logistic regression layer for high-level spatiotemporal features extraction and multi-features fusion to derive the final prediction results. The contributions of the paper can be summarized as follows:

- The MF-CNN model coordinates both short-term and long-term spatiotemporal features rather than a single spatiotemporal feature, as well as external factors in one end-to-end learning structure for short-term traffic flow prediction.
- The inputs of multi-features involved have great impacts on the predictive performance. With the same network structure, different combinations of features derive different performance. The findings extend and

complement the study of traffic flow features.

- The experimental results on two real traffic datasets show that the MF-CNN model is effective and promising in short-term traffic flow prediction, which can achieve the trade-off between efficiency and accuracy compared with five baseline models.

The rest of this paper is organized as follows: Sect. 2 reviews the literature relating to traffic features analysis and the CNN model. Section 3 describes the proposed model MF-CNN with multi-features. Section 4 defines the experimental settings, and two real-world traffic datasets are employed to test the accuracy and efficiency of the MF-CNN model. Finally, the last section concludes the exposition.

2. Related Work

2.1 Traffic Features Analysis

In the past decades, a large amount of researches on traffic flow prediction has been explored in need of the growing demand for ITS. Complex spatiotemporal features in traffic network can provide significant information for the accurate traffic flow prediction, which contain local salient characteristics of traffic network [20], [21]. Traffic parameters, such as flow, speed, and density, are observed strongly correlated to time and space. To show the spatiotemporal features in traffic data, the representation falls into three categories: (1)vectors. One-dimensional time vectors, space vectors, and their combinations are most commonly used in single road prediction [7], [12]; (2)matrices. The two-dimensional O-D matrices and time-space matrices are two typical representation. The former is usually employed to reflect the transfer of traffic variables from one place to another, and the latter is an intuitive representation of variables in time and space. Both of them are widely leveraged in traffic problem [18], [22]; (3)tensors. Some researchers have attempted to apply three-dimensional tensors on missing data imputation and anomaly detection in transportation system. This is a more complex representation of traffic data, and each dimension in tensors has different definitions as needed [23], [24].

Besides the temporal continuity in self-features of traffic flow, traffic flow often exhibits strong periodicities. This is due to the strong temporal regularities of urban life, for example, people need to work on time during workdays. The SARIMA model first considered the periodic cycles into the traffic flow prediction model, showing that the consideration of periodicity can improve the prediction accuracy [25]. With the continuous studies of traffic flow features, Jiang et al. proposed a hybrid wavelet-packet-ACF model for analysis of traffic flow time series, concluded that traffic flow exhibited a short-term periodicity with a period of 24 hours(1 day) and a long-term periodicity with 168 hours (1 week), and the more extended period did not display visible periodicity [26]. Zhang et al. studied traffic flow data by spectral analysis technique, indicating that

traffic flow exhibited strong day-to-day cyclical patterns, which were beneficial to generate better predictive performance [27]. These studies theoretically prove the existences of periodicities in traffic flow. In addition, some researchers point out that external factors, such as adverse weather, holidays, weekends would cause uncertainties that dramatically affect the patterns of traffic flow. Ordinarily adverse weather, such as heavy rain, snow, fog, would cause travel speed reduction and traffic flow accumulation, in serious cases, may bring traffic accidents. Adding the weather factors into the prediction model can reduce the inaccuracy of traffic prediction [28], [29]. Traffic flow patterns are different between holidays and ordinary day, as well as workdays and weekends [30]. The traffic flow increases extremely on the rest days (holidays and weekends) due to the traditional public travel habits.

In our work, we comprehensively consider the periodicities of traffic flow, weather conditions, and holiday factors, so that the prediction model can be robust to spatiotemporal and external factors' shifts.

2.2 Convolutional Neural Networks

The CNN is a typical deep neural network. The following three aspects make it suitable for traffic prediction: 1) powerful ability of local feature extraction. CNNs can implicitly learn the latent higher-level feature of the traffic flow, avoiding the extraction of explicit spatiotemporal features; 2) weight sharing mechanism. The temporal and spatial correlations of traffic flow are not only exhibited on a single road but are reflected in all roads, which share the similar properties. Weight sharing means that the same learning features can be reused for each location of the traffic network. In addition, the weight sharing reduces the number of weights, which makes the network easy to optimize and reduces the risk of overfitting; 3) high extensibility. CNNs can also use the back propagation algorithm to learn network parameters, which facilitates the combination of CNNs and traditional neural networks (e.g., BP-NN) to complete the learning of complex features (e.g., external features) for prediction purposes.

A typical CNN is composed of input layer and output layer, as well as multiple hidden layers as shown in Fig. 1. The hidden layers typically consist of convolutional layers, pooling layers, and fully connected layers. The convolutional layers with convolution filters generate feature maps from inputs through sliding window followed by nonlinear

Fig. 1 A typical structure of CNN.

activation functions, sharing the weights and reducing the number of parameters. The feature maps can be calculated as follows:

$$h_{i,j,k} = f(w_k^T x_{i,j} + b_k) \quad (1)$$

where (i,j) is the pixel index in the feature map, k is used to index the channels of the feature map, f is the activation function, $x_{i,j}$ denotes the input patch centered at location (i,j) , $h_{i,j,k}$ represents the output in the feature map of the k -th channel at location (i,j) , w_k and b_k respectively stand for the weights and biases.

Pooling layers compress the feature maps passed from convolutional layers, and further simplify computing complexity. Let *down* stand for the pooling operation. Then the feature maps can be further represented as follows:

$$S_k = \text{down}(H_k) \quad (2)$$

where H_k denotes the feature map of the k -th channel from the convolutional operation, S_k represents the pooling results.

After continuous operations of convolution and pooling, high-level features are learned from the original inputs. The final tensor of feature maps S will be flattened to be a vector. The fully connected layers are used to connect all the high-level features in the vector and perform prediction by back propagation algorithm to optimize the network structure. The process is as follows:

$$o = g(w * \text{flatten}(S) + b) \quad (3)$$

where o stands for the outputs, g is the activation function.

The CNN essentially learns a large number of mappings between input and output without the need for any previous knowledge. To our cases, the typical CNN is the basis of the proposed model MF-CNN, in which we will combine the multi-features of traffic flow to complete the prediction task.

3. Methodology

In this section, the MF-CNN model is introduced to capture the multiple spatiotemporal features and external factors in short-term traffic flow prediction. First, we present the proposed architecture of MF-CNN model in Sect. 3.1. Then two main components of the MF-CNN model are respectively exhibited in Sect. 3.2 and 3.3. Subsection 3.2 introduces the traffic representation based on multi-features in the MF-CNN model, and Sect. 3.3 illustrates the traffic features extraction and traffic flow prediction.

3.1 The MF-CNN Model

The self-features of traffic flow provide meaningful information for prediction models. The traffic flow is observed temporal continuity, as well as daily periodicity and weekly periodicity. For instance, traffic flow is observed to be continuous during a period. Meanwhile, due to the process of

Fig. 2 MF-CNN model, consisting of three CNNs and a logistic regression layer to output traffic flow prediction at time point t .

urban life, the trend of traffic flow in a whole day is usually recurring in every consecutive workday, the daily distribution of peaks and valleys of traffic flow could be generally stable. This situation is also extended to every week. People are tending to repeat their works on the same day of consecutive weeks. Besides, favorable or adverse weather and holidays will drastically influence traffic flow on the roads. Therefore, it will cause losses and shifts if only the temporal continuity is taken into consideration without periodicities and external factors.

To tackle this problem, the MF-CNN is proposed that collectively incorporates multiple spatiotemporal features and external factors to predict traffic flow in continuous roads. Figure 2 shows a graphical illustration of the MF-CNN model. First, for representing the different spatiotemporal features, the inputs of the traffic flow in the continuous roads are mapped to three two-dimensional spaces, which each one is composed of time and space, represented by two-dimensional matrices. The three spatiotemporal matrices consisting of temporal continuity, daily periodicity, and weekly periodicity share the same spatial dimension and different temporal dimension. Then, three CNNs with different structures are applied to extract the latent high-level spatiotemporal features from the original inputs, capturing both nearby and distant spatial features, as well as both near and far temporal features. Finally, a logistic regression layer connects all the flattened high-dimensional spatiotemporal features and external features to complete the traffic flow prediction. The logistic regression structure avoids the complex selection of weights for multiple features fusion, instead, learns the weights by the network itself.

3.2 Traffic Representation

In our cases, as shown in Fig. 3(a), the traffic flow on all lanes in one direction of p continuous roads $\{y_t^s\}_{s=1}^p$ at time point t is the target to predict. Specifically, traffic flow of the p roads during $\{t-n, t-n+1, \dots, t-2, t-1\}$ are organized

Fig. 3 Traffic representation: (a) traffic network; (b) traffic flow matrix.

as inputs for prediction of next time point t . Put all the historical data together, a two-dimensional time-space matrix X is generated:

$$X = \begin{bmatrix} y_{(t-n)}^1 & y_{(t-n+1)}^1 & \cdots & y_{(t-1)}^1 \\ y_{(t-n)}^2 & y_{(t-n+1)}^2 & \cdots & y_{(t-1)}^2 \\ \vdots & \vdots & \ddots & \vdots \\ y_{(t-n)}^p & y_{(t-n+1)}^p & \cdots & y_{(t-1)}^p \end{bmatrix} \quad (4)$$

Figure 3(b) shows a graphical representation. The x -axis represents time, each row stands for a continuous historical time series on a specific spatial location. The y -axis represents space, each column represents a sequence of dots of the selected neighboring roads at a specific time point. The element at time t on road p stands for the value of traffic flow associated with this specific time point and space location. For example, the value 108 at position $(t-1, p)$ means that there are total 108 vehicles passing over road p during a fixed time interval (e.g., 5min, 15min) at time point $t-1$.

To take the different spatiotemporal features of temporal continuity, daily periodicity, and weekly periodicity as inputs, the above representation is extended to three spatial

Fig. 4 The inputs of spatiotemporal features.

and temporal two-dimensional spaces as shown in Fig. 4, in which t denotes the predicted time point, p represents continuously forecasted roads, c stands for the number of time series for continuity simulation, d and w represent the number of consecutive days and weeks for periodicity simulation. l_c is the continuous time interval, l_d and l_w respectively denote daily time interval and weekly time interval, which depend on the l_c as following:

$$l_d = T_{day}/l_c \quad (5)$$

$$l_w = T_{week}/l_c \quad (6)$$

where T_{day} denotes the time unit of a whole day, T_{week} denotes the time unit of an entire week.

The spatial correlations and temporal correlations of traffic flow can be both correspondingly reflected in the three types of matrices, containing both nearby and distant spatial features, and both near and far temporal features.

3.3 Traffic Flow Prediction

The inputs of MF-CNN model consist of two parts: three types of spatiotemporal features and external factors. Significantly, the essential difference between a traffic matrix and a raw RGB image is that a traffic matrix has smaller data redundancy. Unlike raw RGB images which several pixels represent the similar feature information, each element in traffic matrices has real meaning representing traffic flow at time t on road p . Some studies have shown that subsampling may bring accuracy loss [17], [31]. To avoid that situation, we do not adopt pooling layers in MF-CNN, but only successive convolutional layers to extract each spatiotemporal features.

The two-dimensional convolutional layers are naturally utilized to capture the spatiotemporal features. A set of n two-dimensional filters $\{W_1, W_2, \dots, W_n\}$ is applied to each layer of the input matrices via convolution operation $(*)$ to get the feature maps:

$$F = \{f_1, f_2, \dots, f_n\} = O(X * \{W_1, W_2, \dots, W_n\}) \quad (7)$$

where O is a nonlinear activation function, which is the rectified linear unit (ReLU) in our case. ReLU is defined as $O(x) = \max(0, x)$, which is the most widely used activation

Fig. 5 The architecture of the fusion and output layer.

function in CNNs. A border-mode is applied to pad each input outside the border with zero to make the input and output have the same size in a convolution operation, for the purpose to retain more feature information.

By convolution operation, the model will get three sets of high-level features. As shown in Fig. 5, each set of the extracted features will be flattened to a fixed-dimensional vector. The external factors including weather conditions, temperature, wind power, holidays, and weekends are also transformed into a fixed-dimensional vector. A logistic regression layer is used to merge the three feature vectors with external factors vector together to derive the final prediction, avoiding the complex selection of weights for multiple features fusion.

During the training process of the MF-CNN model, the object is to minimize the mean squared error (MSE) between the predicted value and ground truth, through which the weights can be learned. The objective function of the architecture is as following:

$$\min_{\theta} \frac{1}{N} \sum_{i=1}^N \|X_i - \hat{X}_i\|_2^2 \quad (8)$$

where N represents the number of samples, X_i represents ground truth, $\hat{X}_i$ represents predicted value; $\theta = \{w, b\}$, w denotes the weights that connect all two adjacent layers, and b is the biases of each layer.

4. Experiments

4.1 Dataset

Two datasets are used in this study. One dataset is provided by Jilin Provincial Expressway Administration (JPEA), China. The continuous roads from Changchun to Yanbian are research objects. Another dataset is collected from Caltrans Performance Measurement System (PeMS), U.S.A [32]. The continuous road detectors from west to east on Freeway SR4-W are the targets. Traffic flow is aggregated every 15min according to the suggestion from Highway Capacity Manual [33]. The datasets and the external factors are described in Table 1. For each dataset, we select 90% of the data as training data, the remaining 10% is

Table 1 Data description.

Dataset	JPEA	PeMS
Time span	1/1/2017-7/31/2017	1/1/2018-8/5/2018
Time interval	15 minutes	15 minutes
Road sections	18 roads	20 roads
Total samples	20352	20832
External factors		
Features	Code Values	Description
Weather	0-9	10 types of weather conditions
Temperature	[-24, 34]	the range of temperature
Wind power	0-3	4 types of wind power
Holiday	0,1	0:non-holiday; 1:holiday
Weekend	0,1	0:non-weekend; 1:week-end

Fig. 6 The temporal features of traffic flow: (a) temporal continuity of July 1, 2017; (b) daily periodicity of May 15 - 21, 2017; (c) weekly periodicity of Mar. 1 - July 26, 2017.

chosen as the validation set.

4.2 Multi-Features Analysis

Figure 6(a) shows the traffic flow of the road from Changchun East to Lianhua Mountain aggregated every 15min at 12:00 pm-16:00 pm on July 1, 2017, using JPEA dataset. It can be seen that the traffic flow in the short-term time series has similar traffic status. Figure 6(b) depicts traffic flow at all-time intervals of 7 days from May 15 to May 21, 2017. The peaks and valleys of traffic flow on the same road in different days usually come at similar periods and keep generally stable. The trend shows the daily periodicity of the traffic flow. Figure 6(c) describes traffic flow at a certain time interval (11:45 am-12:00 pm) on Wednesday from March 2017 to July 2017. The traffic flow on the same day of different weeks also shows similar trends, indicating the weekly periodicity. However, the weekly trend is not always stable. In Fig. 6(c), traffic flow in May reaches a local peak as the International Workers holiday comes. Moreover, in June and July, as the climate slowly warms up, the trend of traffic flow shows fluctuated. In general, the overall trend is rising. The temporal features exhibit similar trends in each

Fig. 7 Impacts of external factors on traffic flow: (a) weather condition; (b) holiday.

link of the whole traffic network, validating the existence of traffic self-features in temporal continuity, daily periodicity, and weekly periodicity. They provide meaningful information for traffic flow prediction.

Figure 7(a) shows the differences in traffic flow trends between July 13 (stormy rain) and July 25 (sunny day). The stormy rain sharply affects the traffic flow, the peak declines significantly. Figure 7(b) depicts traffic flow on holiday (International Workers Day) and ordinary day (May 15). Traffic flow on holidays increases extremely compared to the ordinary days. These phenomena illustrate that the weather conditions and holidays both have strong impacts on traffic flow.

4.3 Experimental Setup

The proposed model is built on Theano and Keras supported by python libraries. For the input, to improve the convergence speed of the model, Min-Max normalization method is used to scale the data into the range [0, 1]. The sigmoid activation function is applied at the output layer, whose range is also [0, 1]. Then, the results are processed by anti-normalization to ground truth. For each one of the three spatiotemporal features consisting of temporal continuity, daily periodicity, and weekly periodicity, a structure with three convolutional layers is employed. In this case, the convolutions of temporal continuity and daily periodicity respectively use 64, 32, and 16 filters of size 3×3 ; weekly periodicity uses 64, 32, and 16 filters of size 2×2 . For each convolutional layer, the activation function is ReLU. The batch size is 512. Early stopping is applied to prevent the model from overfitting. The loss function is the mean squared error (MSE). For the external factors, we transform weather conditions, wind power, holidays and weekends into binary vectors by one-hot coding, and use Min-Max normalization to scale temperature into the range [0, 1]. We continue to train the model on the full training data for a fixed number of epochs (e.g., 100 epochs).

For the extra hyperparameters in the proposed model, the time interval l_c is set to 15min, the lengths of daily and weekly periodicity l_d and l_w are fixed to 96 (one day) and 672 (one week), respectively. For the temporal dimension of continuity, consecutive days and consecutive weeks in input matrices, we choose c from {4, 8, 12, 16, 20, 24, 28, 32}, d from {5, 6, 7}, and w from {1, 2, 3}. After performing the hyperparameters of each set individually, the best structures

Table 2 Comparison among different models in traffic flow prediction.

Dataset	JPEA			PeMS		
Model	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE
ANN	0.0118	15.03	0.0160	0.0323	8.98	0.0447
SAE	0.0111	14.34	0.0155	0.0292	8.65	0.0411
CNN	0.0107	14.11	0.0159	0.0280	8.61	0.0417
LSTM	0.0121	14.67	0.0165	0.0374	11.64	0.0524
CLSTM	0.0101	13.31	0.0148	0.0249	8.15	0.0389
MF-CNN	0.0096	13.27	0.0145	0.0254	8.23	0.0392

are obtained, which c is set to 8 (8 adjacent time intervals), d is set to 7 (7 consecutive days), and w is set to 2 (2 consecutive weeks). All parameters of our model are set by trial and error to yield an optimal structure.

4.4 Model Comparisons

To demonstrate the effectiveness of the proposed model, the following baseline models that include five popular intelligent computational models are compared with the MF-CNN model:

- (1) ANN, represents traditional neural network prediction analysis. Eight previous time intervals and 4 nearby roads are extracted as spatiotemporal features. The ANN contains three hidden layers with 400 hidden units in each layer training by gradient descent.
- (2) SAE [12], is a neural network created by stacking autoencoders, which the output of the autoencoder on the previous layer is the input of the next layer. SAE is formed up by three autoencoder layers with 400 hidden units in each layer.
- (3) CNN [18], is set to three convolutional layers with 64, 32, and 16 filters of size 3×3 with the consideration of only temporal continuity.
- (4) LSTM [14], is an improvement of the recurrent neural network (RNN), which is suitable for processing and forecasting time series related problems. The structure of LSTM is set to one hidden layer with 500 hidden units.
- (5) CLSTM [19], is the combination of CNN and LSTM. The structure of CNN is set to three convolutional layers with 64 filters of size 3×3 , and the LSTM is set to one hidden layer with 100 hidden units.

Three performance indexes are used to evaluate the effectiveness of the proposed model, which are mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE). They are defined as following:

$$MAE = \frac{1}{N} \sum_{i=1}^N |X_i - \hat{X}_i| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \hat{X}_i)^2} \quad (10)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|X_i - \hat{X}_i|}{X_i} \times 100\% \quad (11)$$

where X_i and $\hat{X}_i$ respectively denote the ground truth and predicted value. Each model is iterated 10 times and we count the average performance indexes of each model.

Table 2 shows the predictive performance of the MF-CNN model and baseline models. It can be observed that the MF-CNN model on the JPEA dataset outperforms the five baseline models in the three measurements of predictive performance. The maximum reductions of the MF-CNN model observed in MAE, MAPE, and RMSE are respectively 20.66%, 11.71%, and 12.12%, while the minimum are 4.95%, 0.30%, and 2.03%. On the PeMS dataset, CLSTM performs more prominently. The MF-CNN model is slightly inferior to CLSTM ranking the second with the differences of the three measurements to be respectively 1.97%, 0.97%, and 0.77%. On the both datasets, the MF-CNN model considering network-scale traffic achieves better performance than single road prediction models ANN, SAE, and LSTM. Compared with network-scale traffic prediction models CNN and CLSTM, MF-CNN is superior to CNN with the reduction observed in MAE, MAPE, and RMSE respectively to be 10.28%, 5.95%, and 8.81% on JPEA dataset, and 9.29%, 4.41%, and 6.00% on PeMS dataset. It validates the advantage of taking multi-features into consideration than one single feature, and prevents the model from the shifts brought by complex factors. In the comparison of CLSTM, MF-CNN is not always ranking the best, but superior to other four models. It indicates that the proposed model considering more potential spatiotemporal features and external factors of traffic flow is beneficial to improve predictive performance.

Figure 8 depicts the predictive performance on one road in a whole day on July 31, 2017 of JPEA dataset and Aug.4, 2018 of PeMS dataset. Each subgraph consists of two parts: the prediction curve and the error. The error is obtained from the difference between the ground truth and the predicted value. The closer the error is to the midline, the better the predictive performance is. It can be seen that the predicted value curves of the MF-CNN model and CLSTM model represent significant advantages over other four baseline models on both datasets, showing the similar predictive performance and effectiveness.

The execution efficiency of the models is measured in terms of running time. Figure 9 shows that, the MF-CNN model does not achieve an apparent advantage due to the

Fig. 8 Performance comparison among different models: (a)on JPEA dataset; (b)on PeMS dataset.

Fig. 9 Comparison of different models in execution efficiency.

computational complexity, but its running time is still in an acceptable range. The MF-CNN model spends almost half as much time as CLSTM. That is because, the MF-CNN model only depends on the feature extraction ability of CNNs, while the CLSTM model combining LSTM and CNN increases the computational complexity in training

data, and with the increase of depth, the training gets much harder. It indicates that the MF-CNN model achieves the trade-off between the prediction accuracy and efficiency.

4.5 Multi-Features Impact Analysis

To analyze the impact of different features to the traffic flow prediction, combinations of various features, namely, temporal continuity (C), daily periodicity (D), weekly periodicity (W), and external factor (E), are fixed. They are fitted into the MF-CNN model with the same network structure. The experimental results are presented in Table 3, and visual perspective in Fig. 10.

The findings can be summarized as follows:

- (1) Each type of properties helps to build better prediction results. As the number of the properties increases, the error rate decreases, showing that the model taking multi-features into account is superior to that consider-

Fig. 10 Predictions of different features: (a)on JPEA dataset; (b)on PeMS dataset.

Table 3 Comparison among different features of the MF-CNN model.

Dataset	JPEA			PeMS		
Features Fusion	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE
C	0.0107	14.11	0.0159	0.0280	8.61	0.0417
C+D	0.0100	13.51	0.0152	0.0264	8.52	0.0409
C+W	0.0102	13.58	0.0157	0.0267	8.63	0.0415
C+D+W	0.0098	13.42	0.0149	0.0257	8.28	0.0401
C+D+W+E	0.0096	13.27	0.0145	0.0254	8.23	0.0392

ing single-feature.

- (2) The combinations of properties C and D, properties C and W, yield improvements over single property C, demonstrating the consideration of different temporal periodicities is both beneficial to improve traffic flow prediction.
- (3) The model with properties C and D has lower errors than the model with properties C and W. The results

indicate that traffic flow prediction is more sensitive to short-term periodicity than long-term periodicity.

- (4) Compared with the model with properties C, D and W, the model with properties C, D, W and E not only yields a better prediction result but also discovers the relation between traffic and external factors including weather conditions and holiday events.

5. Conclusions

In this paper, we estimate traffic flow by multi-features fusion, and propose a novel deep learning model, named MF-CNN, to predict network-scale traffic flow. This study aims to overcome the problem of complex factors shifts to the robustness of prediction models, and fully utilize the available traffic features to improve the prediction accuracy. In MF-CNN model, the traffic self-features are classified into temporal continuity as short-term feature, daily periodicity and weekly periodicity as long-term features. Each one is represented by two-dimensional time-space matrices with different time lags as one of the inputs. Then the three different spatiotemporal matrices are learned by CNNs to obtain the high-dimensional features with both spatial and temporal correlations. The extracted features are further combined with external factors to fit into a logistic regression layer to complete the final prediction.

The proposed model is evaluated on real traffic data from JPEA dataset and PeMS dataset. The results show that the MF-CNN model on JPEA dataset outperforms the five baseline models with the maximum reduction observed in MAE, MAPE, and RMSE to be 20.66%, 11.71%, and 12.12%, and the minimum to be 4.95%, 0.30%, and 2.03%. On the PeMS datasets, the MF-CNN model is slightly inferior to CLSTM with the differences of MAE, MAPE, and RMSE to be 1.97%, 0.97%, and 0.77%. The execution efficiency of the MF-CNN model measured in terms of running time is in an acceptable range. The running time of the MF-CNN model is not the optimum, but superior to LSTM and CLSTM. The results of impacts of different features on the traffic flow prediction show that the spatiotemporal feature of temporal continuity is essential, daily periodicity is more influential in prediction accuracy than weekly periodicity, and the model with consideration of external factors yields more accurate prediction results than the model with only spatiotemporal correlations. These findings extend and complement the study of traffic flow features.

Consequently, the MF-CNN model considering multiple spatiotemporal features and external factors can prevent the model from complex factors' shifts. It achieves the trade-off between the prediction accuracy and efficiency, showing potential and promising capacity in traffic flow prediction.

Acknowledgments

This work was supported by the National Key Research and Development Program (Grant No.2017YFB1401800), Special Funds for Provincial Industrial Innovation in Jilin Province (Grant No.2016C090) and the Key Program for Science and Technology Development of Jilin Province (Grant No.20190302118GX). The authors would like to appreciate the anonymous reviewers for their valuable comments.

References

- [1] M.S. Ahmed and A.R. Cook, "Analysis of freeway traffic time-series data by using box-jenkins techniques," *Transportation Research Record*, no.722, 1979.
- [2] S. Lee and D.B. Fambro, "Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting," *Transportation Research Record*, vol.1678, no.1, pp.179–188, 1999.
- [3] Y. Kamarianakis and P. Prastacos, "Space-time modeling of traffic flow," *Computers & Geosciences*, vol.31, no.2, pp.119–133, March 2005.
- [4] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended kalman filter: a general approach," *Transportation Research Part B-methodological*, vol.39, no.2, pp.141–167, Feb. 2005.
- [5] A. Bezuglov and G. Comert, "Short-term freeway traffic parameter prediction," *Expert Systems With Applications*, vol.62, pp.284–292, Nov. 2016.
- [6] T. Tsekeris and A. Stathopoulos, "Short-term prediction of urban traffic variability: Stochastic volatility modeling approach," *J. Transportation Engineering-asce*, vol.136, no.7, pp.606–613, July 2010.
- [7] K.Y. Chan, T.S. Dillon, J. Singh, and E. Chang, "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and levenberg-marquardt algorithm," *IEEE Trans. Intell. Transp. Syst.*, vol.13, no.2, pp.644–654, June 2012.
- [8] S. Sun, C. Zhang, and G. Yu, "A bayesian network approach to traffic flow forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol.7, no.1, pp.124–132, March 2006.
- [9] H. Dezani, R.D.S. Bassi, N. Marranghello, L.F. dos Santos Gomes, F. Damiani, and I.N. da Silva, "Optimizing urban traffic flow using genetic algorithm with petri net analysis as fitness function," *Neurocomputing*, vol.124, pp.162–167, Jan. 2014.
- [10] E.I. Vlahogianni, M.G. Karlaftis, and J.C. Golias, "Short-term traffic forecasting: Where we are and where we're going," *Transportation Research Part C-emerging Technologies*, vol.43, pp.3–19, June 2014.
- [11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol.521, no.7553, pp.436–444, May 2015.
- [12] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol.16, no.2, pp.865–873, April 2015.
- [13] N.G. Polson and V.O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transportation Research Part C-emerging Technologies*, vol.79, pp.1–17, June 2017.
- [14] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C-emerging Technologies*, vol.54, pp.187–197, May 2015.
- [15] O. Abdel-Hamid, A. rahman Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional neural networks for speech recognition," *IEEE Trans. Audio, Speech, Language Process.*, vol.22, no.10, pp.1533–1545, Oct. 2014.
- [16] A. Tatsuma and M. Aono, "Food image recognition using covariance of convolutional layer feature maps," *IEICE Trans. Inf. & Syst.*, vol.E99-D, no.6, pp.1711–1715, 2016.
- [17] J. Zhang, Y. Zheng, and D. Qi, "Deep spatio-temporal residual networks for citywide crowd flows prediction," *National Conference on Artificial Intelligence*, pp.1655–1661, 2016.
- [18] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction," *Sensors*, vol.17, no.4, p.818, 2017.
- [19] H. Yu, Z. Wu, S. Wang, Y. Wang, and X. Ma, "Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks," *Sensors*, vol.17, no.7, p.1501, 2017.

- [20] Z. Zhang, Q. He, H. Tong, J. Gou, and X. Li, "Spatial-temporal traffic flow pattern identification and anomaly detection with dictionary-based compression theory in a large-scale urban network," *Transportation Research Part C-emerging Technologies*, vol.71, pp.284–302, Oct. 2016.
- [21] Y. Yang, Y. Xu, J. Han, E. Wang, W. Chen, and L. Yue, "Efficient traffic congestion estimation using multiple spatio-temporal properties," *Neurocomputing*, vol.267, pp.344–353, Dec. 2017.
- [22] E. Cascetta and M.N. Postorino, "Fixed point approaches to the estimation of o/d matrices using traffic counts on congested networks," *Transportation Science*, vol.35, no.2, pp.134–147, May 2001.
- [23] M.T. Asif, N. Mitrovic, J. Dauwels, and P. Jaillet, "Matrix and tensor based methods for missing data estimation in large traffic networks," *IEEE Trans. Intell. Transp. Syst.*, vol.17, no.7, pp.1816–1825, July 2016.
- [24] H. Fanaee-T and J. Gama, "Event detection from traffic tensors: A hybrid model," *Neurocomputing*, vol.203, pp.22–33, Aug. 2016.
- [25] B.M. Williams, P.K. Durvasula, and D.E. Brown, "Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models," *Transportation Research Record*, vol.1644, no.1, pp.132–141, 1998.
- [26] X. Jiang and H. Adeli, "Wavelet packet-autocorrelation function method for traffic flow pattern analysis," *Computer-aided Civil and Infrastructure Engineering*, vol.19, no.5, pp.324–337, Sept. 2004.
- [27] Y. Zhang, Y. Zhang, and A. Haghani, "A hybrid short-term traffic flow forecasting method based on spectral analysis and statistical volatility model," *Transportation Research Part C-emerging Technologies*, vol.43, pp.65–78, June 2014.
- [28] A. Koesdwiady, R. Soua, and F. Karray, "Improving traffic flow prediction with weather information in connected cars: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol.65, no.12, pp.9508–9517, Dec. 2016.
- [29] Y. Jia, J. Wu, and M. Xu, "Traffic flow prediction with rainfall impact using a deep learning method," *J. Advanced Transportation*, vol.2017, pp.1–10, 2017.
- [30] Y. Bao, F. Xiao, Z. Gao, and Z. Gao, "Investigation of the traffic congestion during public holiday and the impact of the toll-exemption policy," *Transportation Research Part B-methodological*, vol.104, pp.58–81, Oct. 2017.
- [31] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.39, no.4, pp.640–651, April 2017.
- [32] California Department of Transportation, "Caltrans Performance Measurement System," <http://pems.dot.ca.gov/>, 2019. [Online; accessed 16-January-2019].
- [33] Highway Capacity Manual, vol.1–4 (2010), Transportation Research Board, 2010.

Di Yang received the B.S. and M.S. degrees in computer science and technology from Changchun University of Science and Technology, Changchun, China, in 2013 and 2016. She is a Ph.D. candidate in Changchun University of Science and Technology, China. Her research interests include intelligent transportation, urban computing, etc.

Songjiang Li received the B.S. and M.S. degrees in computer science and technology from Changchun University of Science and Technology, Changchun, China, in 2002 and 2010, received the Ph.D. degree in computer science from Jilin University, Changchun, China, in 2015. His research interests include intelligent transportation, urban computing, etc.

Zhou Peng received the B.S. degree in computer science and technology from Harbin Engineering University, Harbin, China, in 2015. He is a M.S. candidate in Changchun University of Science and Technology, China. His research interests include intelligent transportation, deep learning, etc.

Peng Wang works as a professor in School of Computer Science and Technology, Changchun University of Science and Technology. His research interests include software engineering and information systems, and database and data mining, etc.

Junhui Wang is an engineer in Jilin Provincial Expressway Administration. His research interests include intelligent transportation, urban computing, etc.

Huamin Yang works as a professor in School of Computer Science and Technology, Changchun University of Science and Technology since 1999. He is vice chairman of CCF Jilin section. His research interests include computer vision, internet of things technology, intelligent computing and intelligent software, etc.