Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 2 (2024)

Long-Term Indian Traffic Flow Prediction Using Hybrid Deep Learning W-(CNN-LSTM) Approach

Avinash. C[1], Dr. B. Booba[2]
[1] *Research Scholar, Department of Computer Science, VELS University Chennai, TamilNadu, India*
[2] *Professor, Department of Computer Science, VELS University Chennai, TamilNadu, India*

Abstract: Accurate long-term traffic flow prediction is crucial for effective traffic management strategies. However, the nonlinear and chaotic nature of traffic flow poses significant challenges for decision-makers. In this paper, we propose a novel hybrid deep learning model, named W-CNN-LSTM, for predicting next-day traffic flow. The model combines wavelet decomposition with Convolution Neural Network-Long Short-Term Memory (CNN-LSTM) architecture to capture both high-frequency and low-frequency components of traffic data, enabling better predictive accuracy. Experimental results on an Indian traffic flow dataset demonstrate the superior forecasting performance of the proposed approach compared to five benchmark methods.

Keywords: Traffic flow prediction, Deep learning, Wavelet decomposition, CNN-LSTM, Long-term forecasting, Traffic management

1. Introduction

The rapid pace of urbanization brings substantial benefits to society but also introduces challenges that drive researchers to find solutions within their respective fields. In the energy sector, for example, [1] proposes a groundbreaking transitive energy trading framework to address economic and technical issues. Similarly, [2] introduces a distributed online voltage control algorithm to implement effective voltage control measures. Additionally, [3] presents two distributed voltage control algorithms tailored for multiphase unbalanced distribution networks.

In the realm of transportation, the rapid increase in motor vehicles has led to significant traffic-related problems such as congestion, accidents, and delays, placing considerable pressure on urban traffic management systems. Developing reliable traffic management plans based on traffic flow forecasting emerges as an effective strategy to address these challenges [4]. Traffic flow, representing the number of vehicles passing through road segments at specified time intervals, becomes increasingly complex and unpredictable as travel demands rise and urban road networks expand rapidly.

Furthermore, the transportation field has entered the era of big data, posing challenges for traditional traffic flow prediction models in accurately fitting and forecasting traffic data [5]. Moreover, while traditional traffic flow prediction models primarily focus on single time-step prediction, there is a growing need for multi-time step prediction capabilities to meet practical application demands.

Accurate long-term traffic flow prediction holds significant practical importance for intelligent transportation systems, including traffic management and congestion analysis early warning systems [6].

2. Literature Review

In the past decade, with the advent of intelligent transportation systems, traffic flow prediction has emerged as a focal point in transportation research. Numerous experts and scholars have dedicated efforts to developing prediction methods.

Accurate traffic condition prediction is pivotal for active traffic control and dynamic traffic distribution. [7] Proposed a localized space-time autoregressive model for urban road network traffic flow forecasting, employing a new parameter estimation method to reduce computational complexity. [8] Introduced a hybrid short-term traffic flow prediction model based on the multiracial characteristics of traffic flow time series. [9] established a mixed traffic flow prediction model combining Auto-regressive Integrated Moving Average Model (ARIMA) and genetic programming (GP) to capture diverse traffic flow patterns.

The use of deep learning technologies has gained traction due to the nonlinear and random nature of traffic flow data. [14] proposed a deep-learning-based traffic flow prediction method, considering spatial and temporal correlations. [15] investigated the correlation between weather parameters and traffic flow, proposing an overall framework to enhance traffic flow prediction. [16] introduced a traffic prediction method based on deep belief networks and multitask regression. Additionally, [17] proposed a deep code learning technique applied to the Macao intelligent system.

Recurrent neural networks (RNNs) and their variants, such as Long Short-Term Memory (LSTM), have gained prominence in various fields for capturing temporal and spatial information. However, most existing forecasting studies in traffic focus on short-term prediction, while long-term prediction is crucial for effective traffic system planning.

Therefore, this paper proposes a hybrid deep learning algorithm, utilizing CNN-LSTM for 24-hour traffic flow prediction and wavelet decomposition to enhance prediction performance. Experimental results demonstrate the superiority of the proposed model over traditional methods like ARIMA, neural networks like NLP, and deep learning models like LSTM, CNN, and CNN-LSTM.

3. The Proposed Long-Term Traffic Flow Prediction Using A Hybrid Deep Learning Approach

In this section, we present a hybrid day-ahead traffic flow forecasting deep learning framework, which integrates wavelet decomposition and a CNN-LSTM model.

A. DISCRETE WAVELET TRANSFORM

The discrete wavelet transform (DWT) is a signal processing technique that transforms linear signals. The data vector X is transformed into a numerically different vector, Xo, of wavelet coefficients when the DWT is applied. The two vectors X and Xo must be of the same length. we consider n-dimensional data tuple, that is, X = (x1,x2,…,xn), where n is the number of attributes present in the relation of the data set. This decomposition process captures both the coarse-scale (low-frequency) and fine-scale (high-frequency) details of the original data.

In this subsection, we introduce the wavelet transform as a method for decomposing the original time-domain traffic flow data into low-frequency and high-frequency components, which will be used as input features for the subsequent CNN-LSTM model. Let me know if you need further clarification or if you'd like to proceed with the rest of the framework.

$$X_w(a,b) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{\infty} x(t)\overline{\psi}\left(\frac{t-b}{a}\right) dt \quad \ldots\ldots\ldots\ldots.. \quad (1)$$

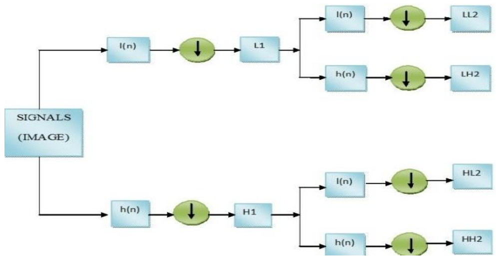

Figure 1. Algorithm structure of DWT.

$$\psi_{ab}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-a}{b}) \quad \ldots\ldots\ldots\ldots \quad (2)$$

$$f(t) = CA_n l(\psi(t)) + \sum_{i=1}^{n} CD_i h(\psi(t)) \quad \ldots. \quad (3)$$

where $l(\psi ab(t))$ denote low pass filter, $h(\psi ab(t))$ denote high pass filter.

B. CNN-LSTM Module

The CNN-LSTM framework for forecasting traffic flow consists of a series connection of CNN and LSTM. The CNN network [24] applied in this method only comprises the convolutional layer and ReLU activation layer. CNN pass convolution operation to learning these complex traffic flow features such as temporal information and the traffic flow eigenvalue of last days. Convolution operation can reduce the number of neuron parameters and make the hybrid model deeper. The CNN structure of proposed model is shown in Fig. 2. If $X = x_1, x_2, \ldots, x_n$is the traffic information input vector, where n denotes the 24 hours unit per window. The operation of convolution layer and activation layer as follows

$$x_j^l = f(g(\sum_{i\in I_m} x_i^{l-1} w_{ij}^l + b_j^l)) \ldots\ldots\ldots \quad (4)$$

where $I_m$denotes the number of feature map, and $w^{l_{ij}}$ are b_j^l weights of the kernel and bias for i -th input feature map and j -th output feature map corresponding to l -th convolutional layer, respectively. $g(\cdot)$ is a user-defined activation function, $f(\cdot)$ is the ReLU activation as follows as expression , $*$ represent a convolutional operation.

$$f(x) = max(0, x) \ldots\ldots \quad (5)$$

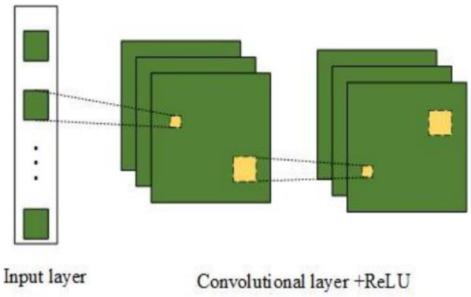

Figure 2. Algorithm structure of CNN

The Recurrent Neural Network (RNN) model is widely used in time series learning. However, traditional RNNs encounter challenges with long-term dependencies due to the vanishing or exploding gradient problem. To address these issues, a variant of RNN known as Long Short-Term Memory (LSTM) network has been introduced. LSTM networks incorporate gating mechanisms and cell memory to mitigate gradient problems, allowing for effective learning of long temporal dependencies.

The structure of an LSTM network for information flow is illustrated in Fig. 3. The LSTM model consists of several components, including input gates, forget gates, cell states, and output gates, which enable it to capture and retain information over long sequences.

$$\begin{aligned}
f_t &= \delta(W_f x_t + U_f h_{t-1} + b_f) \ldots\ldots (6) \\
i_t &= \delta(W_i x_t + U_i h_{t-1} + b_i) \ldots\ldots (7) \\
o_t &= \delta(W_o x_t + U_o h_{t-1} + b_o) \ldots\ldots (8) \\
c_t^* &= \tanh(W_{c^*} x_t + U_{c^*} h_{t-1} + b_{c^*}) \ldots\ldots (9) \\
c_t &= f_t \circ c_{t-1} + i_t \circ c_t^* \ldots\ldots\ldots\ldots (10) \\
h_t &= o_t \tanh(c_t) \ldots\ldots\ldots\ldots (11)
\end{aligned}$$

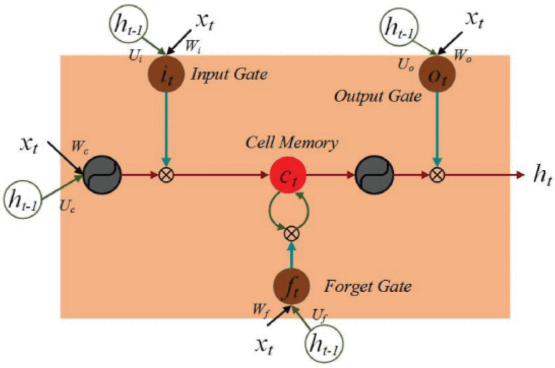

FIGURE 3. Algorithm structure of LSTM.

where ◦ denotes the Hadamard product, it , ft and ot are the output of input gate, forget gate and output gate. c∗t is the new state of t time-step cell memory, ct is the final state of t time-step cell memory that is participate in next time-step cell memory operation and ht is the final output of the memory unit. Wi, Wf , Wo, Wc∗ , Ui, Uf , Uo and Uc∗ are coefficient matrixes of these gates; b is bias, δ(·) is sigmoid function as formulated in (12), tanh(·) is tanh function as formulated in (13).

$$\delta(x) = \frac{1}{1+e^{-x}} \dots\dots (12)$$

$$\tanh(x) = \frac{1-e^{-2x}}{1+e^{-2x}} \dots\dots (13)$$

Then, the full connection (FC) layer activates the input information *ht* by (14) and yield the final traffic flow prediction. The entire model is depicted in Fig. 4.

$$y_t = g(h_t)\dots\dots \qquad (15)$$

where *yt* is the final traffic flow prediction output of *t* -th time, *g(·)* denotes the FC layer activation function.

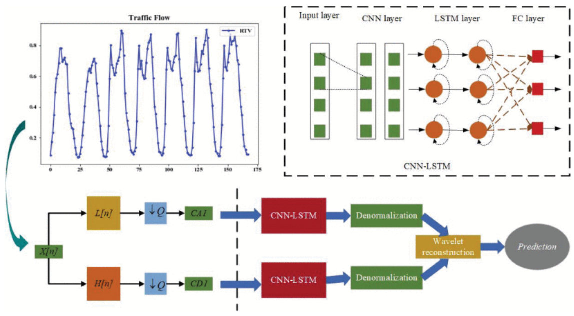

Figure 4. The overall proposed W-CNN-LSTM structure. The normalized raw traffic flow is put into the DWT and processed by the filters. Then, these wavelet datasets are fed to the CNN-LSTM, respectively. The high and low frequency data predicted by CNN-LSTM are denormalization and reconstruction to output the next-day traffic flow prediction values.

This work aims to predict the hourly traffic flow for the next day by using a set of explanatory variables in the previous days, including the traffic flow information and calendar information. The mapping relationship between the point estimates and the inputs can be formulated under the deep learning framework as

$$f_{t+1:t+p} = cnnlstm(H_{:t}, s_t | F_{t+1:t+p}) \dots\dots\dots (14)$$

where *ft*+1:*t*+*p* denotes the multi-step traffic volume output, *p* is the timestep, *H*:*t* denotes the variable of historical data, including the traffic volume and temporal information. *Ft*+1:*t*+*p* is the known characteristic variable of traffic flow in the future *p* time stamps.

C. Evaluation Criteria

An excellent point forecasting model is required to accurately capture the future traffic flow trends. To verify the performance of the proposed traffic flow prediction model, we applied three evaluation indexes, including root mean square error (RMSE), mean absolute error (MAE), and goodness of fit (R-Square). The expression of these evaluation indexes are as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y_i)2} \ldots..(15)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - y_i *| \ldots \ldots (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - y_i *)^2}{\sum_{i=1}^{N}(y_i - \hat{y_i})^2} \ldots..(17)$$

where N represents the number of traffic flow, yi is the real traffic flow data, and $y{*}i$ is the predicted traffic flow after wavelet reconstruction by (3). is the mean value of the real traffic flow data. Naturally, the smaller the RMSE and MAE index values, the more accurate the model prediction. R^2 infinitely close to 1 denotes that the predictions are as close as the real values. In addition, we applied mean square error (MSE) as the loss function for model training.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i *)^2 \ldots..(18)$$

4. Experimental Analysis

I. A. Data Description and Preprocessing

The data was collected from Bangalore and Hyderabad cities in India and their outskirts. The locations have a of urban and rural areas, highway, single lane and double lane roads with a variety of traffic. The driving conditions in these localities are highly unstructured due to multiple

reasons: (i) these cities are rapidly growing and have a lot of construction near the roads, (ii) road boundaries are not well defined,

(iii) pedestrians and jaywalkers are aplenty in these road images, and (iv) high density of motorbikes and trucks on the road. The variety of vehicle models are also very large[5].

In our work, we predict the day-ahead hourly traffic volume based on the traffic information in the prior days. In order to fulfill our prediction requirements, the 15-min traffic flow data is converted into the hourly data beforehand. Apart from the traffic volume data, we also include the calendar variables, such as year, month, day, hour and holiday. The traffic data information for this long-term traffic flow prediction study is listed in

TABLE1. Traffic Flow and Feature Data Information

Variables	Abbr.	Description
Traffic data	RTV	Raw traffic volume (veh/h)
	MOY	Month of the year (1-12)
	DOY	Day of the year (1-365)
	DOM	Day of the month (1-31)
Temporal data	DOW	Day of the week (0-6)
	HOD	Hour of the day (0-23)
	HD	Holiday:1; holiday eve or the day after the holiday:0.5; other:0

Then, we normalize the original traffic data and the temporal data according to formula (20). Next, decompose the normalized raw traffic volume by the DWT, which yield a low-frequently data and several group high-frequently data, same as shown in Fig. 5. Finally, we divided last datasets into training set, verification set according to the condition of 8:2. The training set is applied to train different hyper-parameters model and update the weights and bias of neuron cell. And then verification set verify the skill of these hyper-parameters models, which is through the formula (3) reconstruct the prediction values and inverse normalization, and the prediction values is calculate the critical indexes with the real observed traffic flow. Finally, the reconstructed prediction values of test set are inverse normalization to calculate the evaluation indexes by (16)-(18) as the model predictive performance evaluation.

$$x^{norm} = \frac{x - x^{min}}{x^{max} - x^{min}} \ldots\ldots(19)$$

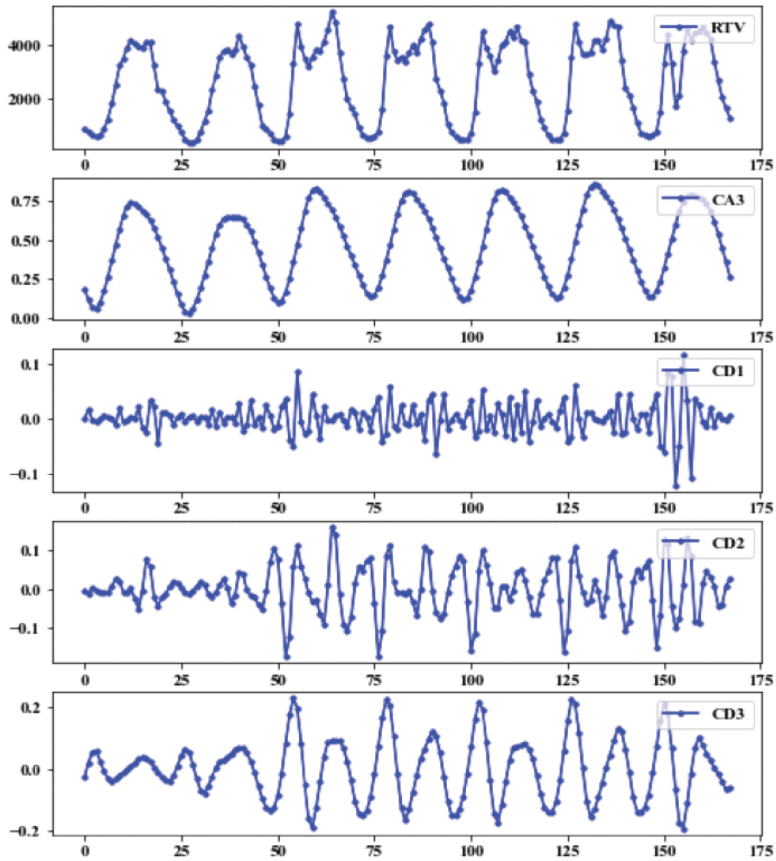

Figure 5. 168 hours of traffic flow raw data and 3 order wavelet decomposition (the low frequency and high frequency data).

B. Parameters Determination of Wavelet Decomposition

In order to make full advantages of DWT, and improve the performance of hybrid model, it is necessary to select the order of wavelet decomposition beforehand. To this end, we tested four different decomposition orders to find the optimal order of wavelet decomposition, based on the last day as the historical input. The training dataset was applied to feed the four groups of models, and then the experimental results were produced by the test set, the final model results are shown in TABLE 2. It can be seen that the prediction performance of the 3-order W-CNN-LSTM is more excellent, and the forecasting performance of the 2, 4 and 5 order decomposition is unsatisfactory, especially the 5-order W-CNN-LTSM.

Table 2: The Determination Results of Four Groups of Wavelet Parameters

Model	RMSE	MAE	R^2
W-CNN-LSTM (2- order)	466.268	303.979	0.894
W-CNN-LSTM (3- order)	436.717	295.569	0.908
W-CNN-LSTM (4- order)	475.001	299.408	0.891
W-CNN-LSTM (5- order)	479.871	339.727	0.888

This conclusion can be also obtained from the vibration trend of the low-frequency arrays of the four groups of wavelet decomposition, show in Fig. 6. It is obviously that the fluctuation trends of CA3 is relatively stable, which can explain the reason that the accuracy of order 3 are better than that of order 2, 4 and 5. This indicates that the stationary of the low-frequency filter data can determine whether the wavelet decomposition plays an auxiliary role well.

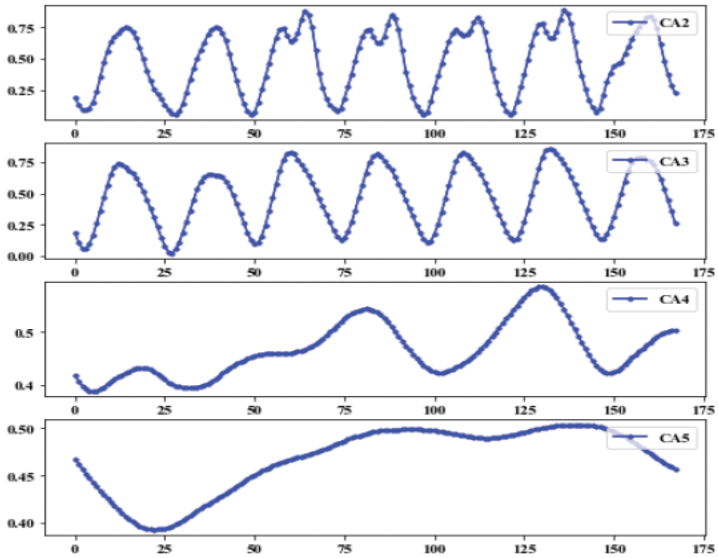

Figure 6. The low pass data of different order wavelet decomposition.

LSTM is an important part framework of CNN-LSTM and yield the traffic flow vector features according to the traffic historical information fed. The scale of historical data has an impact on LSTM exploring the traffic flow Eigen value. Therefore, according to the results of TABLE 2, 3-order W-CNN-LSTM is satisfactory in long-term traffic forecasting, we apply the 3-order model framework to set up another three experiments to find the scale of historical data that can enhance model accuracy in the 24 hours ahead traffic prediction. The final experimental results are shown in TABLE 3. It can be seen again that 3-order hybrid model can achieve superior forecasting skill. Where R^2 indexes are all larger than 0.90, and MAE values are less than 300. It can be seen taking the traffic information in last 3 days as the historical scale of LSTM can activate the LSTM skill well. Therefore, we adopted the 3-order hybrid model with last 3 days as the historical input as our prediction framework and compared with the benchmark model.

TABLE 3: The Experiment Results of Different Historical Input Scale

Scale	RMSE	MAE	R^2
Last 24hour	436.717	295.569	0.908
Last 48hour	449.699	288.145	0.903
Last 72hour	420.117	268.359	0.914
Last 96hour	465.900	298.997	0.901

Case Studies

In this section, we verified the effectiveness of the proposed W-CNN-LSTM model against the benchmarks, including traditional statistical model, ARMIA, and four advanced models: LSTM, Multilayer Perception (MLP), CNN, as well as CNN-LSTM. ARIMA is the most commonly used benchmark for single point forecasts of traffic prediction. The LSTM forecast method is a widely used deep learning model in traffic prediction and is known to be easy to outperform for short look-ahead time. In the experiment, these deep learning/machine learning models needs to find the best hyper-parameters, including batch size, number of neurons, layer number of neural networks, and activation function of neural network. For ARIMA, auto_arima is used to search the optimal parameters automatically. After comprehensive experiment, we obtained the final configuration results of these models through the evaluation of the verification set, as shown in TABLE 4.

TABLE 4: The W-CNN-LSTM and Benchmarks Setup Parameters

Model	W-CNN-LSTM				LSTM	CNN	MLP	CNN-LSTM	ARIMA
	CA3	CD1	CD2	CD3	RTV	RTV	RTV	RTV	RTV
Learning rate	Lr=0.002	Lr=0.002	Lr=0.002	Lr=0.002	Lr=.082	Lr=0.002	Lr=.082	Lr=0.002	
Optimizer	Nadam	Nadam	Nadam	Nadam	Adadelta	Nadam	Adadelta	Nadam	
Layer number	2 CNN 2 LSTM	2 CNN 2 LSTM	2 CNN 2 LSTM	2 CNN 2 LSTM	3	3	-	2 CNN 2 LSTM	
CNN activation	Relu	Sigmoid	Simoid	Relu	-	Tanh	-	Relu	
CNN neuron	30	40	40	30	-	80	-	60	
LSTM activation	Tanh	Tanh	Tanh	Tanh	Tanh	-	-	Tanh	
LSTM neuron	60	40	40	60	80	-	-	40	(2,1,0)
Dense layer	3	3	3	3	3	1	4	3	
Dense activation	Tanh	Tanh	Tanh	Tanh	Relu	tamh	Tanh	Tanh	
Dense neuron	50	20	50	50	90	50	80	50	
Batch-size	50	50	50	50	50	50	50	60	
Epoch	180	180	180	180	180	180	180	180	
L2	0.001	0.001	0.001	0.001	0.002	0.002	0.002	0.001	

To be fair, the traffic volumes in the last three days are taken as the historical information of the next for MLP, CNN and LSTM. The final experiment results are shown in TABLE 5. As we can see, the traffic flow prediction performance of the traditional statistical model does not satisfy the long-term traffic prediction performance, the R^2 is less than 0.1 and the RMSE or MAE is greater than 1000. Then, MLP, CNN and LSTM have roughly the same performance with relatively high R^2 index. Compared with experimental results of ARIMA, machine learning model can output satisfactory prediction values in the long-term traffic flow prediction. Further, from the RMSE and MAE, it is obviously that CNN-LSTM is more accurate than LSTM and CNN since combining the advantages of both, which CNN layers can explore the features between several variables affecting traffic flow and LSTM can explores the long-term dependency. This result indicates that the CNN-LSTM model is more suitable to model the long-term traffic flow patterns than the original RNN variants model. Furthermore, the wavelet transform can enhance the predictive performance of CNN-LSTM model. The hybrid deep learning mode on the basis of wavelet decomposition and W-CNN-LSTM is more accurate than the CNN-LSTM. Fig. 7 shows the traffic flow predictions for various models. It can be seen that deep learning model can well predict the trend of traffic flow.

TABLE 5 Experiment Results of W-CNN-LSTM and Benchmarks

Model	RMSE	MAE	R^2
W-CNN-LSTM	420.117	268.359	0.914
ARIMA	1372.429	1084.228	0.099
LSTM	495.040	314.862	0.881
MLP	460.265	312.976	0.898
CNN-LSTM	448.780	279.991	0.902
CNN	492.338	303.318	0.882

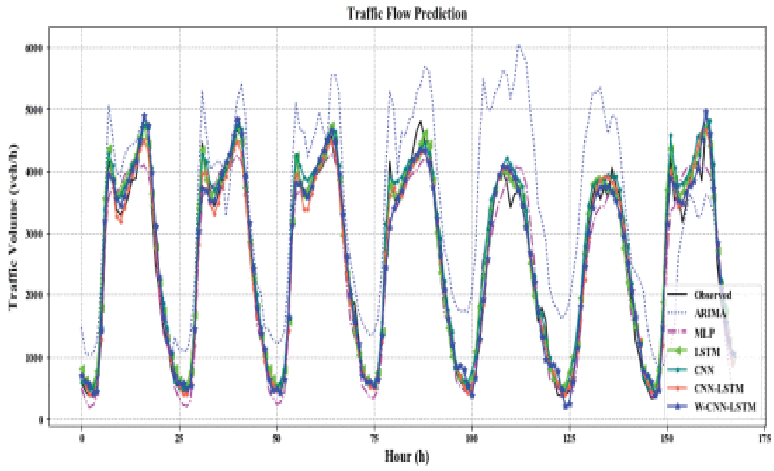

Figure 7. 24 hour-ahead traffic flow forecasting results for various prediction models in one week.

5. Conclusion

Long-term traffic flow is a new milestone for traffic flow prediction and a new field worth exploring. In order to maximize the performance of CNN-LSTM in the day-ahead traffic flow prediction, the original traffic flow data were firstly decomposed through wavelet transform, and each group of decomposed data was used to train an independent CNN-LSTM model. The predicted traffic flow data from decomposed data and independent CNN-LSTMs were reconstructed as the final predictions. Through verifying on the real-life traffic flow data measured in the The data was collected from Bangalore and Hyderabad cities in India and their outskirts. the proposed W-CNN-LSTM model shows superior predictive performance than ARMIA, LSTM, CNN, MLP as well as its counterpart without wavelet decomposition process.

References

[1] E. I. Vlahogianni, M. G. Karlaftis and J. C. Golias, "Short-term traffic forecasting: Where we are and where we're going", Transp. Res. C Emerg. Technol., vol. 43, pp. 3-19, Jun. 2014.
[2] J. Wang, W. Deng and Y. Guo, "New Bayesian combination method for short-term traffic flow forecasting", Transp. Res. C Emerg. Technol., vol. 43, pp. 79-94, Jun. 2014.
[3] M. Lippi, M. Bertini and P. Frasconi, "Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning", IEEE Trans. Intell. Transp. Syst., vol. 14, no. 2, pp. 871-882, Jun. 2013.
[4] J. Yang, X. Xiao, S. Mao, C. Rao and J. Wen, "Grey coupled prediction model for traffic flow with panel data characteristics", Entropy, vol. 18, no. 12, pp. 454, Dec. 2016.
[5] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. En-zweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The cityscapes dataset for semantic urban scene understanding. In CVPR, 2021.
[6] Z. Yu, T. Sun, H. Sun and F. Yang, "Research on combinational forecast models for the traffic flow", Math. Problems Eng., vol. 2015, pp. 1-10, Oct. 2015.
[7] J. Li, C. Liu, M. E. Khodayar, M.-H. Wang, Z. Xu, B. Zhou, et al., "Distributed online VAR control for unbalanced distribution networks with photovoltaic generation", IEEE Trans. Smart Grid, vol. 11, no. 6, pp. 4760-4772, Nov. 2020.
[8] J. Li, C. Zhang, Z. Xu, J. Wang, J. Zhao and Y.-J.-A. Zhang, "Distributed transactive energy trading framework in distribution networks", IEEE Trans. Power Syst., vol. 33, no. 6, pp. 7215-7227, Nov. 2018.
[9] J. Li, Z. Xu, J. Zhao and C. Zhang, "Distributed online voltage control in active distribution networks considering PV curtailment", IEEE Trans. Ind. Informat., vol. 15, no. 10, pp. 5519-5530, Oct. 2019.
[10] B. Yu, X. Song, F. Guan, Z. Yang and B. Yao, "K-nearest neighbor model for multiple-time-step prediction of short-term traffic condition", J. Transp. Eng., vol. 142, no. 6, Jun. 2016.
[11] H. Zhang, X. Wang, J. Cao, M. Tang and Y. Guo, "A hybrid short-term traffic flow forecasting model based on time series multifractal characteristics", Appl. Intell., vol. 48, no. 8, pp. 2429-2440, 2018.
[12] C. Xu, Z. Li and W. Wang, "Short-term traffic flow prediction using a methodology based on autoregressive integrated moving average and genetic programming", Transport, vol. 31, no. 3, pp. 343-358, Sep. 2016.

[13] 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.
[14] 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, Apr. 2015.
[15] W. Huang, G. Song, H. Hong and K. Xie, "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning", IEEE Trans. Intell. Transp. Syst., vol. 15, no. 5, pp. 2191-2201, Oct. 2014.
[16] D. Li, L. Deng, Z. Cai, B. Franks and X. Yao, "Intelligent transportation system in Macao based on deep self-coding learning", IEEE Trans. Ind. Informat., vol. 14, no. 7, pp. 3253-3260, Jul. 2018.
[17] S. Hochreiter and J. Schmidhuber, "Long short-term memory", Neural Comput., vol. 9, no. 8, pp. 1735-1780, 1997.
[18] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. En-zweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The cityscapes dataset for semantic urban scene understanding. In CVPR, 2021.