Retraction

Retracted Article: Prediction architecture of deep learning assisted short long term neural network for advanced traffic critical prediction system using remote sensing data

Wei Wang* R. Dinesh Jackson Samuelb and Ching-Hsien Hsuc

*College of Computer Science, Xianyang Normal University, Xianyang, China; bFaculty of Technology, Design and Environment, Visual Artificial Intelligence Lab, Oxford Brookes University, Oxford, UK; cCollege of Information and Electrical Engineering, Asia University, Taichung, Taiwan

ABSTRACT

This paper presents a Neural Convolutional Network (NCN) based approach for learning traffic as images and predicting high accuracy network-wide broad traffic speed. In the recent past, images describe time and space of traffic flow, where a 2-dimensional time-space matrix is used to convert space dynamics. In recent years, neural networks have been widely used for the prediction of short term traffic, where the description is covered by an NCN in two consecutive steps: abstract data extraction and network-wide traffic forecast. This paper proposes Prediction Architecture of Neural Convolutional Short Long Term Network (PANCSLTN) for the purpose of effectively capturing dynamic nonlinear traffic systems with deep learning assistance. The PANCSLTN can resolve the problem of backdated decay error via memory blocks and shows superior prediction capacity for time series with long-time dependency. Moreover, PANCSLTN can determine the optimum time laggards automatically and the travel data from Beijing microwave traffic detectors which are used for model the training and testing to validate the effective PANCSLTN using remote sensing technique. PANCSLTN can deliver the most accurate and stable prediction performance compared to different topologies in dynamical neural resealing networks or other dominant parametric and nonparametric algorithms during experimental analysis.

Introduction about the importance of critical prediction systems

The prospect of the future is one of the most attractive issues for people and transport management, where the same applies (Ma et al., 2017) to critical data management. It is of great interest and significance to understand the creation of traffic throughout the road network rather than on a single road (Jia et al., 2017) that helps people with complete information about traffic to take better road decisions and supports road managers in systematic management and allocation of resources (Jiang & Fei, 2016) using remote sensing techniques. Here, the large-scale network critical prediction requires more challenging predictions for models, such as the ability to deal with the greater computational complexity of network topology (Cui et al., 2018), the ability to build a smarter and more effective prediction to overcome the spatial connection of the two roads and the ability to predict long-term traffic (Hsu, 2017) is considered as a significant area of research. Fortunately, the traditional models of traffic prediction that typically view traffic speeds as sequential data which do not provide them due to constraints such as hypotheses and suppositions (Zhang et al., 2018), ineptitude to deal with outliers, noisy and incomplete data and failure to cope with the dimensionality of the problem (Baskar et al., 2019) has been surveyed using remote sensing techniques.

As shown in the above Figure 1 the performance of TIS implementations depends on the quality of transport information Systems (TIS) (Yu et al., 2017). This specifically refers to ATMS and Advanced Traveler Information Systems (Yeon et al., 2019), where both transport agencies and travelers need accurate and reliable information on traffic (Polson & Sokolov, 2017). The awareness and prediction of future traffic conditions (Kong et al., 2018), for instance, travel times, travel flow and travel velocity (Xie et al., 2007), is one of the main needs of transport societies. Successful implementation of the application for traffic forecasting can not only help preplanning and rescheduling of routes by passengers (Bezuglov & Comert, 2016), it can provide guidance to transport and reduce congestion professionals to increase traffic protection (Xu & Niu, 2018) based on remote sensing techniques.

According to the stochastic flow features, the avoidance of traffic status cannot be a simple task. Nonetheless, large traffic sensors rapidly increasing availability, coverage and generate numerous studies based on different sources of information (Cai et al., 2018).
The majority of these studies use the sensor data of inductive lapse to determine whether the travel time or the volume of traffic is short, using videos or emails as the basis of the prediction model formation (Yuan et al., 2019). Neural artificial networks (NAN) are widely used to address various problems of traffic prediction because of their advantages, including their capacity to use multidimensional data (Manogaran, G., et al). NAN presented a real-time prediction algorithm for the prediction of traffic speed based on adverse weather conditions. Nevertheless, NAN data-driven method cannot illustrate appropriately for the spatial associations of a road network. However, a NAN is lower in its predictability due to its shallow architecture compared to deep learning approaches (Gu et al., 2019). More sophisticated and solid deep learning models have been recently used to predict traffic flow using deep learning architectures with remote sensing techniques. Deep learning approaches take advantage of much more complicated and more complex systems than a NAN (Boto-Giralda et al., 2010). Nevertheless, these efforts often focus primarily on forecasting road traffic or a small area of the network. Few studies have seen the whole transport system and have assessed the evolution of traffic on a large scale directly (Lv et al., 2018). Most of these models simply considered the temporal correlations of the evolution of traffic at a single location and did not take their spatial correlations into account from the network perspective. Based on the survey this paper’s contributions can be summed up as follows:

1. Time changes and spatial dependency in network traffic are taken into consideration and applied simultaneously to problems related to traffic prediction by leveraging the proposed NCN imaging approach assisted deep learning.
2. Spatiotemporal traffic characteristics can be derived with high accuracy of prediction using NCN automatically.
3. Due to the implementation of the convolution and pooling strategy, the proposed approach may be extended to large-scale traffic speed problem prediction.

In Section 2, a two-step process is introduced that involves the transformation of network traffic into images and an NCN for the estimation of network traffic. In Section 3, four prediction tests are conducted using the proposed method on two transport networks and compared to the other prediction methods. In Section 4, final conclusions are drawn with future guidelines for the study.

**Literature study on various traffic conditions**

A few recent approaches to speed predictions have been established in most algorithms focused on machine learning and statistical analysis. To analyze the future performance of 8 different models, this paper analyzed two minutes of travel data from three remote traffic microwave sensors in the Southbound 4th Ring Road segment in Peking City. In particular, authors reported five machine learning methods: such as The Neural Network Back Propagation (BPNN) (Jiang et al., 2016), the Non-Linear Autoregressive System with the NARX Network Exogenous Inputs, the Vector Support Machine with radial kernel function (SVM-NR) and the Linear Function Support Machine (SVM-LIN), as the nominee. The results of the prediction are meaningful and the prediction accuracy in pace is deteriorated when all the model predicts increase in the time steps; two conventional statistical models are clearly superior to the BPNN, NARXNN or SVM-RBF; the prediction output of ANN is higher than that of the SVM and MLR; the ST model can co-exist as time progress increases.

This paper offers a new way of building a fuzzy neural network to anticipate multifaceted journey speed based on 2-minute traveling speed data on the 3 microwaves on a fourth ring road in Beijing, Germany. To complete the fuzzy inference is used as the first-order system Takagi-Sugeno. Two learning mechanisms are suggested to train the evolving neural network (EFNN) (Tang et al., 2017). The K-means
approach is used first to partition sample entry to various clusters, and each cluster is assisted by a Gaussian fuzzy membership function that tests the membership of cluster center samples. The cluster centers are changed and the membership functions are revised based on the number of entry samples increase. Second, a weighted recursive low-square estimator is used in the Takagi-Sugeno style fuzzy law, to optimize the linear function parameters. To capture the periodic variable in raw speed data, a trigonometrically regression function is implemented. In particular, the predicted performance is compared between the proposed model and six traditional models, which are the artificial neural network, support vector machine, self-repressive, built-in average and self-repressive models.

The present article suggests a new model for the prediction of air pollutant concentration, the neural network extended (LSTME) (Li et al., 2017) under critical prediction system, that takes into account inherently a spatiotemporal correlation. Long shorter layers of memory (LSTM), used to automatically extract useful features from historical data concerning air pollution, have been incorporated into the architecture for performance enhancement with auxiliary data, including weather information and time stamp information. The tests were conducted with the STDL model, the Timeline Delay Network (TDNN), the Autoregressive Movable Average (ARMA), the SVR model and the LSTM standard NN model and comparative results show that the LSTME model is superior to those of other statistically-based models. The results were collected and the LSTME model was officially superior.

Their popularity has been based mainly on rotating networks (CNN) and repeating networks, and deep-learning methods are renowned for their artificial intelligence status. The CNN always achieves dominant performance in visual tasks through the use of fundamental spatial properties of images and video. Recurrent Networks (RNN) can successfully characterize the time correlation, with particularly long-term memory methods (LSTMs), thereby exhibiting superior time series. In this report, the CNN and LSTM integrated deep architecture to predict future traffic flow (CLTFP) (Wu & Tan, 2016) are discussed. One-dimensional CNN is used to capture traffic flow spatial features and two LSTMs are used to minimize short-term volatility and traffic flow periodicities. The CLTFP provided is in comparison with popular open dataset forecasting methods. Experimental results indicate that in traffic flow modeling, the CLTFP has significant advantages. Moreover, from Granger’s point of view, the proposed CLTFP is analyzed, and several CLTFP properties are discovered and discussed.

One of the key challenges of the intelligent transportation network is a short-term traffic forecast. Precise forecast results allow passengers to make appropriate modes of journey, routes and departures, which are significant for traffic management. A more effective approach for traffic data processing is a viable way of promoting prediction accuracy. In recent years, the availability of extensive traffic data and computing power has led to the improvement of the accuracy of short-term traffic projections using deep learning methods. A new model of long-term memory (LSTM) (Zhao et al., 2017) traffic forecast is proposed, unlike traditional predictive models, in a two-dimensional network consisting of many Memory Units, the proposed LSTM network considers temporally spatially correlation in traffic systems. The proposed LSTM network will produce improved results as compared with other representative forecast models. As discussed in the above methods in recent decades, short-term traffic projections have been attracted by various interests from foreign researchers. There have been significant work efforts to improve traffic forecasting methods Deep learning approaches use architectures that are far broader and more complex than an NCN, and they produce better results than standard methods.

Architecture of deep learning assisted-neural convolutional short long term network

As shown in the above Figure 2 determines the average speed of the floating cars in the form of time and space matrix traffic details should be considered jointly for forecasting network-wide traffic congestion as an output image. Let as considered an axis of a and b which forms a matrix time and space, the elements in the matrix are the values of the time and space-dependent traffic variables. In the way each pixel in the image shares the matrix value, where the generated matrix can be viewed as an image source. A floating car with a dedicated GPS-device records a vehicle trajectory with specific information about the speed and locations of a vehicle at some point. On the path, space-time traffic details can be calculated and further transformed into a time-space matrix, which serves for time-space images on each road section. In general, short intervals, such as 10-sec for traffic forecasting, are meaningless. Thus these data are collected for wider intervals, for instance, many minutes, if the sampling resolution is high using remote sensing access. A series of dotted areas can be simply and linearly organized in the b axis. It can lead to a large size and informational problems because of the sequences of dots are repetitive. The parts are then spatially organized and placed into the Y-axis with the relation to a predefined starting point for a lane.
Computational analysis for time and space dimensions

Algorithm 1: To Determine Time and Space Dimensions of Equation

1. Procedure: determining mathematical modeling using the time and space dimensions in a matrix format.
2. Input: set the value as $V_i D_i$.
3. Calculate the traffic average speed of equation.
4. Calculate $D^j = [x_j, x_{j+1}, \ldots, x_{j+R-1}] j \in [1, y-R+1]$
5. Update the pooling & flatten convolution.
6. Calculate $u^i_k = \text{pool}(\tau([V_i D_i^k + Y_i]), i \in [1,e_1])$
7. Calculate $u^i_k = \text{flatten}([u^i_k, u^i_{k+1}, \ldots, u^i_{e_k}]), i = e_k$
8. Update $b$
9. Finally, determine the traffic average speed of the equation.
10. End for
11. End procedure

As shown in the above algorithm (1) the time-space matrix is represented using the time and space information dimension which can be mathematically denoted in the matrix format in the following equation (1)

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1y} \\ \vdots & \vdots & \ddots & \vdots \\ X_{Cy} & X_{c2} & \cdots & X_{cy} \end{bmatrix}$$ (1)

$Y$ is the length of the time interval; where the road section distance is denoted as $C$, $i^{th}$ column vector of $X$ is the transportation traffic speed at time $j$ and the pixel $X_{ji}$ is the traffic average speed at section $j$ at the time $i$. Here the matrix $X$ forms the clear final image which has been characterized using NCN.

Characteristics of neural convolution network (NCN) with deep learning assistance

Deep learning assisted NCN transportation structure with four main components, i.e., input model, extraction of traffic characteristics, predictive and model performance. The following is clarified in each of the sections. First, the model input is the image generated from a spatiotemporal characteristic transportation network. The input and output interval lengths can be $E$ and $R$ respectively. The input model $D^j$ is represented in the following equation (2),

$$D^j = [x_j, x_{j+1}, \ldots, x_{j+R-1}] \cdot j \in [1, y-R+1]$$ (2)

Here $j$ denotes the index of the sample, time interval length is $y$, $x_j$ representing column vector traffic speed of all transportation road networks.

As shown in the above Figure 3 $V_i D_i$ are the input layer parameters denoted for the activation of the function, $u_i$ is the input layer of the image which is sent to the multiplication factor and integrated by using i/p, o/p and forget gates to produce the output layer of the image. Furthermore, the traffic features are omitted by integrating convolution and pooling, which is the core component of the NCN model. The first convolution output pooling layer can be written in the following equation (3)

$$u^i_1 = \text{pool}(\tau(V^i_1 + Y^i_1)), i \in [1, e_1]$$ (3)

Here the input, output, and parameters of the layer (Figure 4) are represented as $D^i_1$, $u^i_1$ and $(V^i_1, Y^i_1)$. $i$ is the function activation of the pooling layer (pool) and the total number of convolution of $i^{th}$ layer is denoted as $e_1$. The output in the $i^{th}$ layer ($K \neq 1, k = 1K$) has the
pooling convolution layer $u_i^k$ which is represented in the following equation (4)

$$u_i^k = \text{pool} \cdot \left( \tau \left( \sum_{l=1}^{c_{1}} V_{i,l}^l D_{i,l}^l + Y_{i,l}^l \right) \right), i \in [1, e_1] \quad (4)$$

The features acquired by traffic extraction are merged in the estimation model into a dense vector comprising the final and highest features of the input transport network. The vector in the dense form can be written as flatten convolution (flatten) $u_{k}^{\text{flatten}}$ in the following equation (5)

$$u_{k}^{\text{flatten}} = \text{flatten} \cdot ([u_1^k, u_2^k, \ldots, u_k^k]), i = i e_k \quad (5)$$

The depth of deep learning NCN is represented as $k$, concatenating procedure is discussed in the flatten way. Finally, through a fully connected layer, the vector is transformed into model output. The output $\hat{b}$ of the flatten model can be entered in the following equation (6)

$$\hat{b} = V_k u_{k}^{\text{flatten}} + y_k$$

$$= V_k (\text{flatten} \left( \tau \left( \sum_{l=1}^{c_{1}} V_{i,l}^l D_{i,l}^l + Y_{i,l}^l \right) \right)) + y_k$$

$$\quad (6)$$

Where $V_k$ and $y_k$ are the fully connected parameters, $\hat{b}$ are the network-wide predicted traffic speed. For a prediction problem, all these abstract features are essential, because of different model outputs, where the formation of objectives differ at each and every matrix of an image. Due to the constant speeds of the traffic the outputs should therefore take on continuous cost functions. Cross-entropy cost functions are typically used in the image classification problem.

**Analysis of pooling and convolution layer of NAN**

Algorithm.2. Mathematical Modeling to Determine the Characteristics of Pooling layers of NAN

1. Procedure: determining the characteristics of pooling layers using Neural Artificial Network
2. Input: set the value as $V_{k,l}^i Y_{k,l}^i$.
3. Calculate the model parameters using the back propagation algorithm.
4. Calculate $b_{\text{conv}} = \sum_{k=1}^{a} \sum_{l=1}^{b} (V_{k,l}^i Y_{k,l}^i)$
5. Update square means deviation of the equation.
6. Calculate $b_{\text{pool}} = \max(Y_{k,l}^i), k \in [1, \ldots a], l \in [1, b]$
As shown in the above algorithm (2) Convolutional layers are different from the traditional neural feeding network, where each neuron is linked with every neuron output and the network is fully connected (completely connected layer). It should be noted that each layer is activated by an activation function before discussing the specific layers. To handle the complexities of a transportation network, the advantages of using the activation function are as: (a) The activating function (a) converts the output into a manageable, sized range of data that supports the modeling training; and (b) the combination of layer-based activation function can simulate highly complex nonlinear tasks. The complexity of the transportation network for the activation function is defined in the following equation (7)

\[
f_{i(a)} = \begin{cases} 
  a, & \text{if } a > 0 \\
  0, & \text{otherwise}
\end{cases}
\] (7)

Such abstract traffic functions are combined to derive more abstract traffic features based on deep learning assistance and a higher level, the method confirms the composition of the NCN, which means that each filter composes a local path \( Y_{k}^{*} \) from the lower to the higher level. Therefore one convolution filter \( V_{li} \) is determined in the following equation (8)

\[
b_{convolu} = \sum_{k=1}^{x} \sum_{l=1}^{y} (V_{li} Y_{k}^{*})
\] (8)

The pooling layers are built to down sample and add data because only prominent numbers are collected from the surrounding region. NCN is guaranteed by pooling layers locally, where the NCN can always derive the same function from the input, irrespective of function changes, rotations or scales. The above data are used to classify the most prominent characteristics of input layers in the maximum way based on NCN network size. The maximum pooling layer operation is carried out in the following equation (9)

\[
b_{pool} = \max(Y_{ik}), k \ni [1, \ldots, a]; l \ni [1, \ldots b]
\] (9)

Here \( a \) and \( b \) are determined as the pooling size windows. The NCN predictions are traffic speeds in various sections of the lane, and the Square Mean Deviation (SMD) is used to measure the distance between predictions and speeds of ground-truth traffic. The training goal of NCN is taken by minimizing SMD, therefore SMD can be written in the following equation (10)

\[
\text{SMD} = \frac{1}{y} \sum_{j=1}^{y} (\hat{b}_j - b_j)^2
\] (10)

Therefore, the model parameters of the following equation are represented by the back propagation algorithm and denoted in the following equation (11) (12) & (13)

\[
\varphi = \arg \min \frac{1}{y} \sum_{j=1}^{y} (\hat{b}_j - b_j)^2
\] (11)

\[
\varphi = \arg \min \frac{1}{y} (V_k b_k^{flatten} + y_k - b_j)^2
\] (12)

\[
\varphi = \arg \min \frac{1}{y} (V_k (pool(r(\sum_{i=1}^{V_l} D_i^k + Y_i))) + y_k)^2
\] (13)

In the above equation the back propagation algorithm is determined the form of augmented minimizing, transportation Network spatiotemporal features can be automatically extracted by implementing NCN convolutional and pooling layering, thereby avoiding the need for manual feature selection; Deep learning NCN provides high-quality network-wide traffic data to create network-wide speed predictions; and NCN is replicated in vast transportation networks as it combines masses in convolutional layers using a pooling mechanism, NCN reflects high-level traffic information.

**Neural Artificial Network (NAN)**

NAN is considered another common traffic countermeasure due to its ability to handle multi-dimensional content, structure of versatile model, good ability of learning, and generalization form of adaptability. Utilized NAN multilayer perception is depicted in the following equation (14) & (15)

\[
b = f(\varphi_0 + \sum_{i=1}^{y} \varphi f(\varphi_0 + \sum_{j=1}^{x} \varphi_i y_i))
\] (14)

\[
b_j = \sum_{i=1}^{V_{li}} \varphi (V_{li} (pool(r(\sum_{i=1}^{V_l} D_i^k + Y_i))) + y_k)^2
\] (15)

In the above equation \( V_{li} \) represents the input and hidden vector layer, \( e \) and \( y \) are the transfer functions of the input and hidden layer. The weight values of the neuron metrics are denoted as \( \varphi_i \).

The common aim is to reduce the sum of square errors to a minimum. Errors will be truncated upon memory cell output, and after that, they will enter the linear NAN in a memory cell, where errors can always flow back, and error output is usually exponentially decaying. The following activation function is determined in the equation (16)

\[
z = \sum_{i=1}^{V_{li}} V_{li} (r(\sum_{i=1}^{V_{li}} w_i^l b_i^l + X_i))
\] (16)

Here \( Z \) is the activation function, \( w_i^l X_i \) is the input layer of source nodes, \( b_i^l \) is the hidden layer of source nodes at \( \varphi_i \). NAN is considered another common
traffic countermeasure due to its ability to handle multi-dimensional content; numerous results in the transport field are being made by Neural Artificial Network (NAN). The previous research on the use of the neural network in traffic forecasting can be traced back to the neural network principle in a freeway calculation of traffic time. Since then, there have begun to emerge ever greater neural network variants to improve NAN traffic efficiency.

Results and discussions

The two deep learning-based algorithms are selected to compare and test the efficiency of the proposed algorithm. NCN represents the traditional neural network and tries with hidden layers to learn the characteristics (Manogaran et al., 2019). Multiple layer neural networks with auto encoder model inputs are encrypted into dense or sparse representations and fed into the next layer (Zhang & Ge, 2013). PANCSLTN is an extension of the NAN and is common because of the architecture which can fix long-term memory and prevent the disappearance of gradient issues from conventional NCN.

As shown in Figure 5(a) the analysis of mission results can confirm long-term forecasts using NCN. Usually, long-term forecasting reaches higher MSEs than short-term predictions, when the input timespan is set, which means that making long-term forecasts is difficult rather than making short-term forecasts. The predicted traffic velocity has been converted into three categories: high (0–20 km/h), moderate (20–40 km/h) and free-flow (> 40 km/h). For travelers to schedule their routes this presentation is preferable.

Figure 5. (a) MSE on NCN. (b) MSE on NAN.
As shown in Figure 5(b) the capacity and efficiency of NCN can be validated by measuring and comparing MSE of NAN for the tasks involved in predicting high-level transport network speed. The results show that a further depth of the NAN model reduces MSEs in test data significantly, which means that a NAN model with depth-4 reaches the lowest MSEs for the formation and testing data. Long-term forecasts using NAN can be confirmed by comparing task outcomes typically, long-term predictions perform longer than short-time forecast predictions, because their input time is set, which means that it is more difficult to predict the long-term than to predict the short-term (Peng et al., 2018).

The training time on neural artificial network algorithms is more effectively trained in the model than the NCN, since these algorithms are simple and easy to train. Such algorithms, therefore, represent a significant compromise between the efficiency of their preparation and the accuracy of the prediction with validation based on the epoch range as shown in the Figure 6(b). Many like NAN and PANCSLTN, needless preparation than the NCN and other deep learning architectures. This is mainly due to the fact, with a view to removing comprehensive network-wide spatiotemporal traffic based on the training time as shown in the Figure 6(a), a large number of
convolution kernels apply to each picture that shows the training time to be high in NAN.

As shown in Figure 7 the data-driven process of an NCN, however, cannot explain particularly well the spatial connections between road networks. However, the predictive accuracy of an NCN is lower because of its low architecture in comparison to deep learning approaches. Network spatiotemporal traffic characteristics can automatically be extracted with high predictability using a NAN. To avoid over fitting of the model, an early stop criterion is applied. Model over fitting is the case in which model training does not
improve the NCN validation data prediction accuracy, even as it improves the NAN test data prediction precision.

As shown in Figure 8 to test PANCSLTN efficiency for a short-term travel speed prediction, PANCSLTN is not implemented in the field of transport to the best of our knowledge. The neural Long Short-Term Memory Network exploits the long-term transient dependency of short-term travel forecasts. In this analysis, the Prediction Architecture of Neural Convolutional Short Long Term Network (PANCSLTN) is proposed to estimate the urban traveled speed based on the NAN sensor traffic test data in trained network shows high range with less error rate as shown in the Figure 9 in order to address these drawbacks of conventional NCN and using NAN technique travel speed raises to a greater extent in the proposed algorithm.

As shown in the above Table 1 demonstrates the forecast output of various detector positions only with speed data is calculated based on NAN. The best performance algorithm is highlighted in the NAN using the proposed algorithm and presents the prediction results of different algorithms while combining both historical speed and length. Compared to NCN, NAN stands best because of the ability to automatically evaluate optimum input window dimension PANCSLTN is still superior to others in relation to other techniques in deep learning.

The PANCSLTN will learn time series and calculate time lags automatically according to a long time. This feature is particularly desired when traffic forecasting problems are encountered where future conditions of traffic are common for long-distance events. Efficiency of travel speeds is enhanced by the time delay. Proper time lag will improve the accuracy of the travel speed forecast. PANCSLTN is an efficient approach without prior information of time lag to short-term travel prediction. SVM is the ideal approach for the prediction of time series and for generating comparable results, although significant effort may be required to promote calibration of parameters.

**Conclusion**

This paper presents a new, Prediction Architecture of Neural Convolution Short Long Term Network (PANCSLTN) to predict the speed of travel with microwave detector data. In signal processing with satisfactory results, deep learning methods are commonly used because deeper learning architectures usually conceive of more complex non-linear functions than other Neural Networks. For further work, space and time information in PANCSLTN may be taken into consideration, implying that traffic speeds may include further inputs from adjacent detectors. Therefore, the prediction output should be tested with various data aggregation rates. Another important path for research is the introduction of several layers into PANCSLTN for a deep architecture which may improve neural networking learning capability.
Table 1. Performance of NAN.

| Number of Tasks Used | BPNN | EFNN | LSTM | CLTFF | LSTM | PANCSLTN |
|----------------------|------|------|------|-------|------|----------|
| 10                   | 65.7 | 69.5 | 70.4 | 71.3  | 67.8 | 71.4      |
| 20                   | 69.8 | 60.7 | 76.7 | 67.4  | 76.3 | 68.6      |
| 30                   | 74.5 | 68.6 | 68.4 | 65.6  | 60.4 | 79.9      |
| 40                   | 65.4 | 76.8 | 78.9 | 73.6  | 75.7 | 89.3      |
| 50                   | 76.5 | 83.6 | 86.6 | 87.3  | 88.7 | 95.8      |

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Baskar, S., Periyayagi, S., Shakeel, P. M., & Dhulipala, V. S. (2019). An energy persistent range-dependent regulated transmission communication model for vehicular network applications. *Computer Networks*, 152, 144–153. https://doi.org/10.1016/j.comnet.2019.01.027

Bezuglov, A., & Comert, G. (2016). Short-term freeway traffic parameter prediction: Application of grey system theory models. *Expert Systems with Applications*, 62, 284–292. https://doi.org/10.1016/j.eswa.2016.06.032

Boto-Giralda, D., Diaz-Pernas, F. J., Gonzalez-Ortega, D., Diez-Higuera, J. F., Antón-Rodriguez, M., Martinez-Zarzuela, M., & Torre-Diez, I. (2010). Wavelet-based denoising for traffic volume time series forecasting with self-organizing neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 25(7), 530–545. https://doi.org/10.1111/j.1467-8667.2010.00668.x

Cai, P., Wang, Y., Lu, G., Chen, P., Ding, C., & Sun, J. (2016). A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. *Transportation Research Part C: Emerging Technologies*, 62, 21–34. https://doi.org/10.1016/j.trc.2015.11.002

Cai, Z., Ke, R., & Wang, Y. (2016). Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction. *arXiv Preprint arXiv:1801.02143*

Gu, Y., Lu, W., Qin, L., Li, M., & Shao, Z. (2019). Short-term prediction of lane-level traffic speeds: A fusion deep learning model. *Transportation Research Part C: Emerging Technologies*, 106, 1–16. https://doi.org/10.1016/j.trc.2019.07.003

Hsu, D. (2017). Time series forecasting based on augmented long short-term memory. *arXiv Preprint arXiv:1707.00666*. https://arxiv.org/abs/1707.00666

Jia, Y., Wu, J., Ben-Akiva, M., Seshadri, R., & Du, Y. (2017). Rainfall-integrated traffic speed prediction using deep learning method. *IET Intelligent Transport Systems*, 11(9), 531–536. https://doi.org/10.1049/iet-its.2016.0257

Jiang, B., & Fei, Y. (2016). Vehicle speed prediction by two-level data driven models in vehicular networks. *IEEE Transactions on Intelligent Transportation Systems*, 18(7), 1793–1801. https://doi.org/10.1109/TITS.2016.2620498

Jiang, H., Zou, Y., Zhang, S., Tang, J., & Wang, Y. (2016). Short-term speed prediction using remote microwave sensor data: Machine learning versus statistical model. *Mathematical Problems in Engineering*, 2016, 1–13. https://doi.org/10.1155/2016/9236156

Kong, Y. L., Huang, Q., Wang, C., Chen, J., Chen, J., & He, D. (2018). Long short-term memory neural networks for online disturbance detection in satellite image time series. *Remote Sensing*, 10(3), 452. https://doi.org/10.3390/rs10030452

Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., & Chi, T. (2017). Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. *Environmental Pollution*, 231, 997–1004. https://doi.org/10.1016/j.envpol.2017.08.114

Lv, Z., Xu, J., Zheng, K., Yin, H., Zhao, P., & Zhou, X. (2018, January). LC-RNN: A deep learning model for traffic speed prediction. In *IJCAI* (pp. 3470–3476).

Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., & Wang, Y. (2017). Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction. *Sensors*, 17(4), 818. https://doi.org/10.3390/s17040818

Manogaran, G., Shakeel, P. M., Priyan, R. V., Chilamkurti, N., & Srivastava, A. (2019). Ant colony optimization-induced route optimization for enhancing driving range of electric vehicles. *Int J Commun Syst*, (2019). doi:10.1002/dac.3964

Peng, L., Liu, S., Liu, R., & Wang, L. (2018). Effective long short-term memory with differential evolution algorithm for electricity price prediction. *Energy*, 162, 1301–1314. https://doi.org/10.1016/j.energy.2018.05.052

Polson, N. G., & Sokolov, V. O. (2017). Deep learning for short-term traffic flow prediction. *Transportation Research Part C: Emerging Technologies*, 79, 1–17. https://doi.org/10.1016/j.trc.2017.02.024

Tang, J., Liu, F., Zou, Y., Zhang, W., & Wang, Y. (2017). An improved fuzzy neural network for traffic speed prediction considering periodic characteristic. *IEEE Transactions on Intelligent Transportation Systems*, 18(9), 2340–2350. https://doi.org/10.1109/TITS.2016.2643005

Wu, Y., & Tan, H. (2016). Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework. *arXiv Preprint arXiv:1612.01022*. https://arxiv.org/abs/1612.01022

Xie, Y., Zhang, Y., & Ye, Z. (2007). Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition. *Computer-Aided Civil and Infrastructure Engineering*, 22(5), 326–334. https://doi.org/10.1111/j.1467-8667.2007.00489.x

Xu, S., & Niu, R. (2018). Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China. *Computers & Geosciences*, 111, 87–96. https://doi.org/10.1016/j.cageo.2017.10.013

Yeon, K., Min, K., Shin, J., Sunwoo, M., & Han, M. (2019). Ego-vehicle speed prediction using a long short-term memory based recurrent neural network. *International Journal of Automotive Technology*, 20(4), 713–722. https://doi.org/10.1007/s12239-019-0067-y

Yu, H., Wu, Z., Wang, S., Wang, Y., & Ma, X. (2017). Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks. *Sensors*, 17(7), 1501. https://doi.org/10.3390/s17071501

Yuan, J., Abdel-Aty, M, Gong, Y., & Cai, Q. (2019). Real-time crash risk prediction using long short-term memory recurrent neural network. *Transportation Research Record*, 2673(4), 314–326. doi:10.1177/0361198119840611
Zhang, D., Lindholm, G., & Ratnaweera, H. (2018). Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. *Journal of Hydrology*, 556, 409–418. https://doi.org/10.1016/j.jhydrol.2017.11.018

Zhang, Y., & Ge, H. (2013). Freeway travel time prediction using Takagi–Sugeno–Kang fuzzy neural network. *Computer-Aided Civil and Infrastructure Engineering*, 28(8), 594–603. https://doi.org/10.1111/mice.12014

Zhao, Z., Chen, W., Wu, X., Chen, P. C., & Liu, J. (2017). LSTM network: A deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, 11(2), 68–75. https://doi.org/10.1049/iet-its.2016.0208