Deep learning methods in transportation domain: A review

Article *in* IET Intelligent Transport Systems · July 2018

DOI: 10.1049/iet-its.2018.0064

4 authors, including:

Hoang Nguyen
The Commonwealth Scientific and Industrial Research Organisation

Minh Le Kieu
University of Auckland

IET Intelligent Transport Systems

Special Issue: Selected Papers from the 25th ITS World Congress Copenhagen 2018

Deep learning methods in transportation domain: a review

ISSN 1751-956X
Received on 26th January 2018
Revised 22nd April 2018
Accepted on 13th June 2018
E-First on 30th July 2018
doi: 10.1049/iet-its.2018.0064
www.ietdl.org

Hoang Nguyen[1] ✉, Le-Minh Kieu[1], Tao Wen[1], Chen Cai[1]
[1]*Data 61 – CSIRO, Level 5, Garden St, Eveleigh, NSW 2015, Australia*
✉ *E-mail: hoang.nguyen@data61.csiro.au*

Abstract: Recent years have seen a significant amount of transportation data collected from multiple sources including road sensors, probe, GPS, CCTV and incident reports. Similar to many other industries, transportation has entered the generation of big data. With a rich volume of traffic data, it is challenging to build reliable prediction models based on traditional shallow machine learning methods. Deep learning is a new state-of-the-art machine learning approach which has been of great interest in both academic research and industrial applications. This study reviews recent studies of deep learning for popular topics in processing traffic data including transportation network representation, traffic flow forecasting, traffic signal control, automatic vehicle detection, traffic incident processing, travel demand prediction, autonomous driving and driver behaviours. In general, the use of deep learning systems in transportation is still limited and there are potential limitations for utilising this advanced approach to improve prediction models.

1 Introduction

The ability to process a large amount of data to provide accurate traffic forecasts is important in modern transportation decision support systems. An efficient decision support system can potentially help to minimise incident response time, enhance situation awareness and reduce congestion duration.

However, traffic data processing and modelling are challenging because of the complexity of road networks and spatial-temporal dependencies among them. Furthermore, traffic patterns are heterogeneous, meaning different road segments often have distinct time-variant traffic patterns. A large amount of traffic data is recorded hourly from multiple data sources and sensors, but it is difficult to combine into features for training prediction models, due to significant differences in time, network coverage and data quality.

In the literature, there are many applications of traditional machine learning methods (e.g. support vector machines, logistic regression, decision trees, Bayesian network etc.) which have been developed to predict traffic data [1–5]. However, most of these prediction systems used shallow traffic models which were considered as unsatisfying for big data scenarios [6]. Therefore, this paper reviews deep learning architecture (DLA), a rising research interest in the machine learning field, with a special focus on the transportation domain. In the following sections, a general overview of deep learning approaches is presented, followed by their application in popular traffic data analytics topics including transportation network representation, traffic flow forecasting, traffic signal control, automatic vehicle detection, incident inference, and travel demand prediction.

2 Deep learning overview

A DLA usually consists of multiple levels of representation, constructed by composing non-linear modules that transform the representation at one level into a representation at a higher and more abstract level [7]. With sufficient numbers of these transformations, the models are able to learn complicated functions and structures. For example, in the classification task, the features that are important for discrimination are usually retained from higher layers of representation, while irrelevant variations are suppressed. The key advantage of deep learning over traditional methods is that the feature selection process is automated by a general-purpose learning procedure, without any human involvement. With their specifiable hierarchical learning depths, deep learning approaches have demonstrated high performance in discovering the structure of high-dimensional data in many domains, such as computer vision [8, 9], natural languages processing [10, 11], speech recognition [12, 13] and bioinformatics [14, 15].

Fig. 1 illustrates a traditional neural network versus a DLA. The key difference is the number of hidden layers. Simple neural networks usually contain only one hidden layer and require a feature selection process. On the other hand, a deep learning neural network contains two or more hidden layers and it can perform optimal feature selection and model tuning during the learning process [16]. There are many other deep learning structures including recurrent neural networks (RNNs) [17], deep convolutional networks (DCNs) [18], deep restricted Boltzmann machines (RBM) [19], stacked auto-encoders (SAEs) [20], deep belief networks (DBN) [21] and long short-term memory networks (LSTM) [22], which will all be reviewed and discussed along with their applications in the transportation domain in the following sections.

3 Transportation network representation

Due to their capability of modelling spatial and temporal dependencies within traffic networks, a number of deep learning methods were applied in a study by Ma *et al.* [23], to represent the traffic condition on road segments (links) in transportation networks. Traffic congestion for a network with N links within T time intervals was expressed as a two-dimensional matrix

$$\begin{bmatrix} C_1^1 & C_1^2 & C_1^3 & \dots & C_1^T \\ C_2^1 & C_2^2 & C_2^3 & \dots & C_2^T \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_N^1 & C_N^2 & C_N^3 & \dots & C_N^T \end{bmatrix} \Longrightarrow \begin{bmatrix} C_1^{T+1} \\ C_2^{T+1} \\ \vdots \\ C_N^{T+1} \end{bmatrix}$$

${C_n}^t$ represents the traffic congestion condition on the nth link at time t, as either a binary value (congested or not) or as an n-level traffic condition (e.g. quiet, light traffic, heavy traffic, congested etc.). Learning from historical data of T intervals, the model will then predict the traffic condition of all links at time $T+1$. When a prediction is performed on an individual row (link), the task is

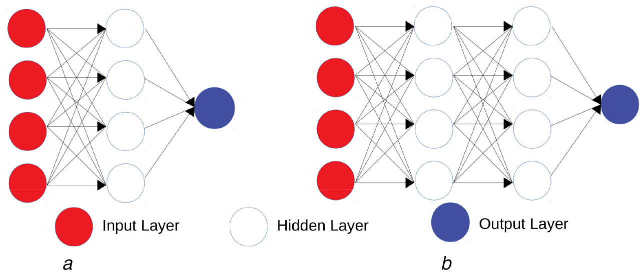


Fig. 1 *Difference between simple neural network and deep learning neural network*

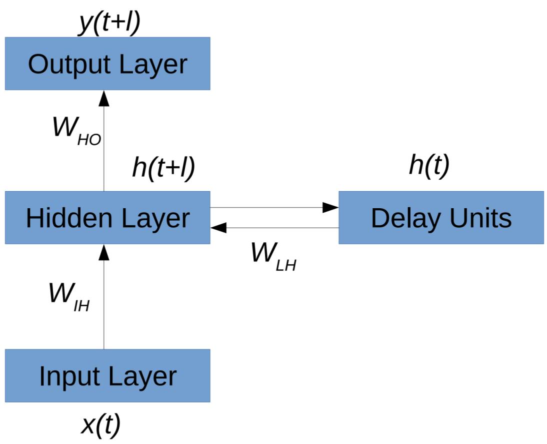


Fig. 2 *State space neural network structure*

similar to a single sequence learning problem. However, in the case where multiple rows (links) are predicted simultaneously, the task can be considered as a high-dimensional sequence learning problem, which requires a DLA with temporal processing capabilities.

The authors used a combination of a RNN and a RBM to model the traffic network. RNN is a special neural network which has at least one internal state to provide a feedback loop from intermediate outputs, back to the inputs. This permits temporal sequence learning. In particular, they leveraged a widely used structure of RNN named state space neural network structure (Fig. 2). In this structure, delay units are used to connect previous hidden unit activations with new inputs into the neural network.

An RBM model is based on energy estimation, and structurally comprises a hidden layer, along with a visible layer, where there are mutual connections between units in both the visible and hidden layers (Fig. 3). The authors used a conditional RBM, which adds an additional feedback loop between the visible and hidden layers, to support learning temporal sequences. The bias values for both the hidden layer and visible layer are updated using feedback from the previous visible units. Both RNN and conditional RBM models apply a similar approach to update bias values.

The conditional RBM and RNN were combined to build an RNN-RBM model as represented in Fig. 4. $b_v^{(t)}$ and $b_h^{(t)}$, respectively, represent the bias vectors for the visible layer and hidden layer in RBM model at time t, and are updated through the hidden units $u^{(t-1)}$ in RNN model at time t–1. Weight matrices W_{uv} and W_{uh} are provided to connect the RNN model and RBM model [21].

This network representation is simple to implement; however, its main limitation is that the model needs to automatically learn and infer the spatial dependency (e.g. related road segments) from historical data, which may result in low prediction accuracies.

Aiming to improve upon the above limitation, Fouladgar *et al.* [24] proposed a deep traffic flow model based on a convolutional neural network (CNN) that considered inflow and outflow information in addition to traffic condition on a road segment. CNN is often composed of one or more convolutional layers with non-linear activation functions and one or more fully connected layers, as in a standard neural network. This architecture was

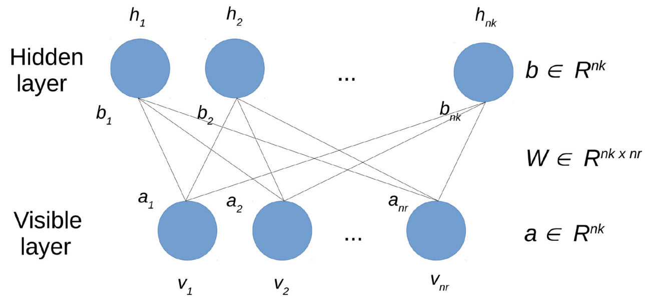


Fig. 3 *RBM structure*

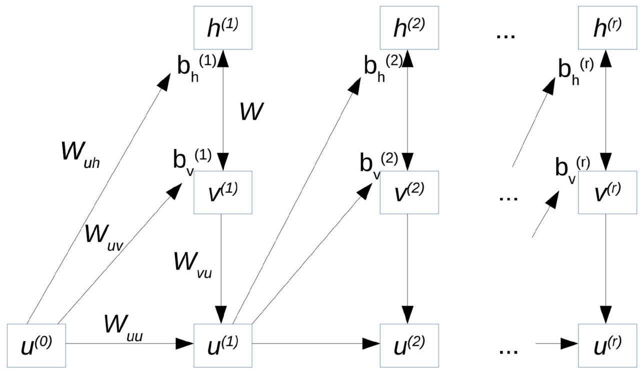


Fig. 4 *RNN-RBM model*

originally designed to exploit the 2D structure of an input image for image processing purposes. Transportation researchers take advantage of this trait to implicitly consider the spatial correlation between road segments in a traffic network. Each convolutional layer usually comes with a sub-sampling (max-pooling) layer which is responsible for reducing the size of the feature maps, hence reducing the computational requirements as well as minimising the likelihood of over-fitting. The CNN structure is optionally followed by a number of fully connected layers, which are similar to a standard multilayer neural network.

From the study [24], for each point S on the traffic network, L_1, L_2, …, and L_n are adjacent points which have flow to S, and R_1, R_2, …, and R_m are adjacent points into which S has flow. L_i and R_i are both arranged by their distance to S, and then a point snapshot at time t is defined as

$$[L_n \to L_{n-1} \to \dots L_2 \to L_1 \to S \to R_m \to R_{m-1} \to \dots R_2 \to R_1 \to S] \text{ over time series } [t-\delta, \dots, t-1, t]$$

Formally, a point snapshot S after time t is defined by a matrix with size $(n+m+1) \times t$, and a network snapshot is the union of all point snapshots (Fig. 5).

Fig. 6 illustrates the deep learning model during the training process. In this figure, the features of the model are classified into two general groups, namely, Traffic Condition and Incidents. Firstly, past traffic conditions are passed to the first filter of CNN as the training set. Then, the outcome of the first layer will be sent to the second convolutional layer. After each cycle, the size of the input will be reduced due to the filter $(\zeta \times \zeta)$ applied on it. This trend is repeated until all of the matrices are reduced to a set of one-dimensional arrays. The incidents can then be provided as input for the fully connected network. The final fully connected network will predict the values of traffic flow on N junctions at next time period $t+1$.

In the same paper, Fouladgar *et al.* [24] also introduced a LSTM network to model traffic conditions. LSTM is an extended RNN architecture which is capable of learning long-term dependencies. LSTM carries the information forward through multiple time steps, as a long-term memory. It has a similar chain architecture as in RNN, but the repeating modules have a special internal design. An RNN module only has a single neural network layer (e.g. tanh layer), while there are four different interacting

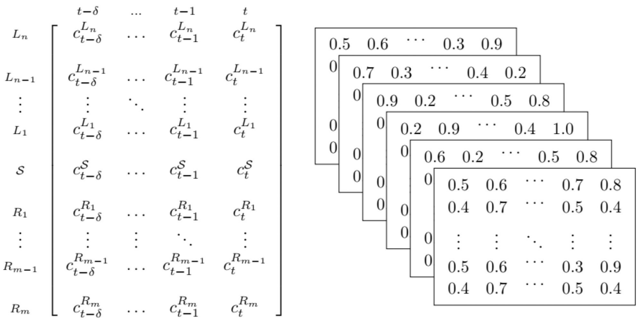


Fig. 5 *Point snapshot and network snapshot [24]*

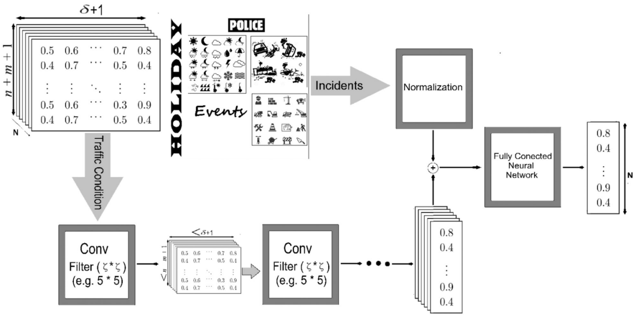


Fig. 6 *Deep traffic flow convolutional network (flow chart from [24])*

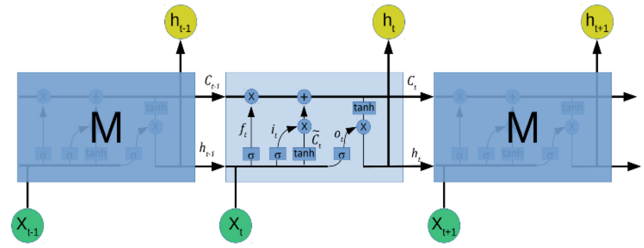


Fig. 7 *Standard LSTM module with four interacting layers*

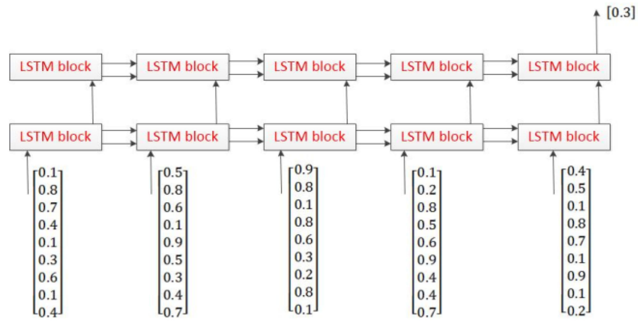


Fig. 8 *Long short-term memory traffic flow [24]*

layers inside an LSTM repeating module in general. Fig. 7 shows an instance of an LSTM repeating module.

The authors claimed that the short-term future traffic condition for a network junction can be predicted based on the traffic patterns from its neighbouring roads, even when there was no historical traffic condition for it. Considering the structure of LSTM, the traffic flow network is presented in Fig. 8. From this figure, the traffic condition of junction S and traffic condition of In(S) and Out(S) at time $t-\alpha$ are passed to the first LSTM layer. Then, in the next step the traffic condition of point S, In(S), and Out(S) at time $t-(\alpha-1)$ is passed to the second LSTM layer. This cycle is repeated until the last layer and output is the traffic condition of S at time $t+1$ (Fig. 8).

The above transportation network representation methods attempt to predict the traffic condition for all road segments in the network. However, the impact prediction for a specific incident or area may not be modelled appropriately because DLA based on a large traffic network may be too general to model smaller areas, or individual incidents at a specific location. Like many other methods, a DLA may not achieve high performance, without proper configuration and parameter selection.

4 Traffic flow prediction

Traffic flow prediction is a basic problem in transportation modelling and data analytics. It aims at estimating the number of vehicles in a road segment for several time intervals in the future [25]. Many reported methods have not been able to achieve desirable results, due to reasons including shallow architectures, manual feature selection and separated learning [26].

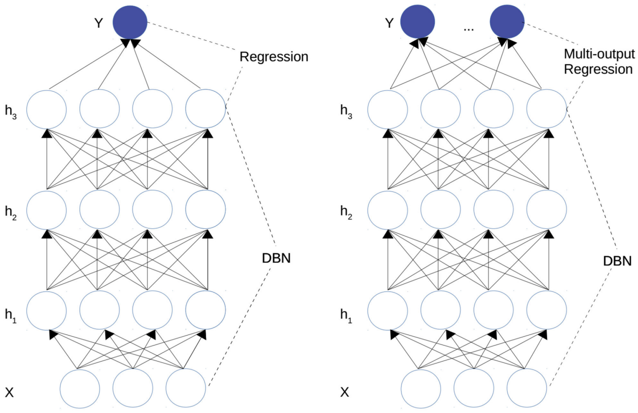
Huang *et al.* [26] introduced a two-level DLA which included a multi-output regression layer at the top and a DBN at the bottom for traffic flow prediction. The authors claimed that this is the first application of deep learning in transportation research. A DBN is constructed as a stack of RBMs, with only one hidden layer in each RBM. The activations of trained units within each RBM are passed to the next RBM in the stack. In earlier research, a fast greedy learner for DBN was proposed by Hinton *et al.* [27] which learnt only one layer at a time.

Huang *et al.*'s DLA for traffic flow prediction is presented in Fig. 9. On a single road or intersection, the input space X contains all the observation points for previous time intervals. Then the input space is connected to a DBN for unsupervised feature learning. Finally, the supervised prediction is performed by a sigmoid regression layer at the top. For predicting traffic flow for the large traffic network, a multi-output regression is executed for a group of road segments. In this case, the related roads should be grouped together, as the overall performance is only improved when jointly trained tasks are related to each other [28]. For transportation network, several grouping methods were used to find related roads in the traffic flow prediction including spatial clustering (adjacent roads and junctions) or traffic flow variation clustering [29]. In this study, the authors presented a weight clustering method which found related roads based on the weight of the top layer of the architecture. The deepest layer of the DBN usually contains only the most representative and useful features from the data.

The authors conducted experiments for both short-term traffic flow, as well as experiments for up to 60 min predictions. The results of their DLA were compared with other existing methods for traffic flow prediction, including ARIMA, Bayesian model, support vector regression (SVR), neural network model, and the multivariate non-parametric regression model [30]. However, the performance of the DLA was not found to be significantly different than the traditional methods, when tested on PeMS dataset. DLA only excelled when tested on road segments with low-to-medium levels of traffic flow.

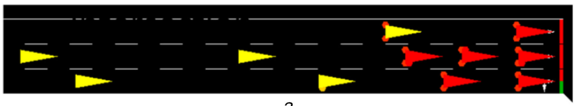
In a more recent study, Polson and Sokolov [31] introduced a DLA for short-term traffic flow prediction (up to 40 min) that combined a linear model trained using L1 regularisation and a sequence of tanh layers. This architecture is a standard and simple implementation of deep learning, but they evaluated their predictions on special events data which had sharp and sudden traffic flow regime changes. The deep learning approach achieved a mean-squared error (MSE) of 14% lower than the neural network with one hidden layer.

Zhao *et al.* [32] presented another deep learning approach for short-term traffic forecast, which represented the temporal-spatial correlations using an LSTM network. This work is similar to the LSTM structure for transportation networks presented in [24]. The main contribution of [32] is that besides traditional ML methods (e.g. ARIMA and SVM), the authors also compared their proposed LSTM network with other deep learning approaches (e.g. RNN and SAE). According to the experimental results, the performance of the LSTM network outperformed all other methods, especially when the forecast time is long. These results demonstrated the advantages of LSTM over: RNN when dealing with a long-time sequence problem; and over the classical parameterised approach when dealing with big traffic data.



a *b*

Fig. 9 *Deep learning architecture for traffic flow prediction*
(a) Single road, ***(b)*** All roads prediction tasks are jointly trained via multi-output regression [26]



a

0	0	0	0	0	0	0	0	1	0	0	1
1	0	0	0	1	0	0	0	1	0	1	1
0	1	0	0	0	0	1	0	0	1	0	1

b

0	0	0	0	0	0	0	0	0.2	0	0	0
0.8	0	0	0	0.5	0	0	0	0	0	0	0
0	0.6	0	0	0	0	0.4	0	0	0	0	0

c

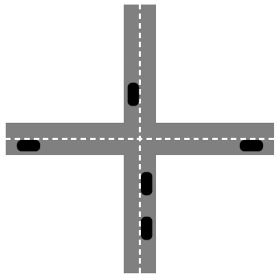
Fig. 10 *Discrete traffic state encoding example*
(a) Vehicle positions and colour-coded speeds, ***(b)*** Binary position matrix, ***(c)*** Encoded speed matrix [36]

5 Traffic signal control

Traffic congestion is one of the biggest issues in modern transportation systems. One solution is to extend road infrastructure, but this is extremely expensive and time-consuming – especially for the already established cities. Alternatively, increasing the efficiency of management systems, including traffic signal controllers (TSCs) can help to optimise vehicle flow in order to reduce congestion and emissions. Many systems have utilised reinforcement learning for TCS, which demonstrated some remarkable results [33–35]. However, most of them did not fully benefit from recent available data [36]. To take advantage of recent big data, Genders and Razavi [36] introduced a deep artificial neural network based on DCN to build an adaptive traffic signal control agent. This agent was trained using reinforcement learning to develop an optimal control policy. The method was then evaluated in the traffic micro simulator SUMO. In this research, the authors introduced a new state space as discrete traffic state encoding based on information density (Fig. 10).

The encoded traffic state was used as input to a DCN and trained using Q-learning with experience replay [37]. The deep reinforcement learning (DRL) method was then compared with a one hidden layer neural network TSC agent. DRL is an extension of traditional reinforcement learning with a deep network. With the support of deep learning, the entire process of reinforcement learning from observation to action is changed as there is no requirement to explicitly design state space or action space. On average, DRL reduced cumulative delay by 82%, queue length by 66% and travel time by 20% [36].

Van Der Pol [38] applied a similar DRL method for TSC, which combined DCN and Q-learning; however, he differentiated his work by using a binary matrix to encode the traffic state for the whole intersection, rather than just a road segment (Fig. 11). Furthermore, his experiments were executed for both single agent and coordinated multi-agent (up to 4) reinforcement learning. In



$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Fig. 11 *Binary position matrix for vehicle position modelling [38]*

general, the DRL approach outperformed the baselines and it is applicable to a multi-agent setting. However, there were several factors that could affect the stability of the algorithm and it is recommended to use a prioritised experience replay [38].

Independently, Li *et al.* [39] also set up a deep neural network (DNN), to learn the Q-function of reinforcement learning for traffic signal timing. Instead of using a DCN, this research utilised an SAE neural network for approximating Q function (Fig. 12). SAE is a DNN that comprises multiple layers of sparse auto-encoders. The SAE neural network takes the state matrix as input and will output the *Q*-value for each possible action. In this work, a greedy layer-wise approach was applied to train each layer to be stacked in the auto-encoders for initialising the weights of a deep network. It is concluded that DRL performed remarkably better than traditional methods in deciding an appropriate signal timing plan.

6 Automatic vehicle detection from image processing

Traditionally, traffic data is collected by magnetic loop detectors or piezoelectric sensors which may not produce consistent and reliable counts. Increasingly, non-intrusive recognition systems, such as unmanned aerial vehicle image system or camera-based vehicle recognition system, have become viable with the introduction of deep learning algorithms [40].

A DNNs method was proposed to classify cars, sedans and vans [41]. However, some pre-extracted features from a well-trained DNN were transferred to their model for the experiment due to lack of training datasets, so the adaptability of their model needs further clarification. The support vector machine (SVM) technique has also been used to conduct multi-class and intra-class vehicle-type classifications, including geometric-based approaches and appearance-based approaches [42]. Wang *et al.* [43], proposed a 2D deep belief network for vehicle detection. The algorithm applied second-order planes instead of first-order vectors as input, and bilinear projection for retaining discriminative information. The model achieved 96.05% accuracy on test data. Adu-Gyamfi *et al.* [44] proposed a deep convolutional neural network in their vision system to detect and classify vehicles into seven classes, and achieved average recall rates between 89 and 99% for seven classes of vehicles.

However, the accuracy of the aforementioned state-of-the-art vehicle detection and classification algorithms still needs to improve, especially under adverse weather conditions. In addition, video cameras can monitor road and traffic conditions in real-time and extract important traffic information such as volume, density, vehicle type and even congestion status; however, only a few researchers have worked on this. Finally, vehicle speed detection, which currently requires coordinate system transformations and field calibration, should also provide valuable data for transportation agencies.

7 Travel demand prediction

Travel demand prediction aims to estimate the number of road or public transport users in the future. It is one of the most fundamental problems in transportation, because most transportation models use passenger demand as an input. The majority of travel demand prediction studies in literature were for planning purposes which attempted to predict the long-term travel

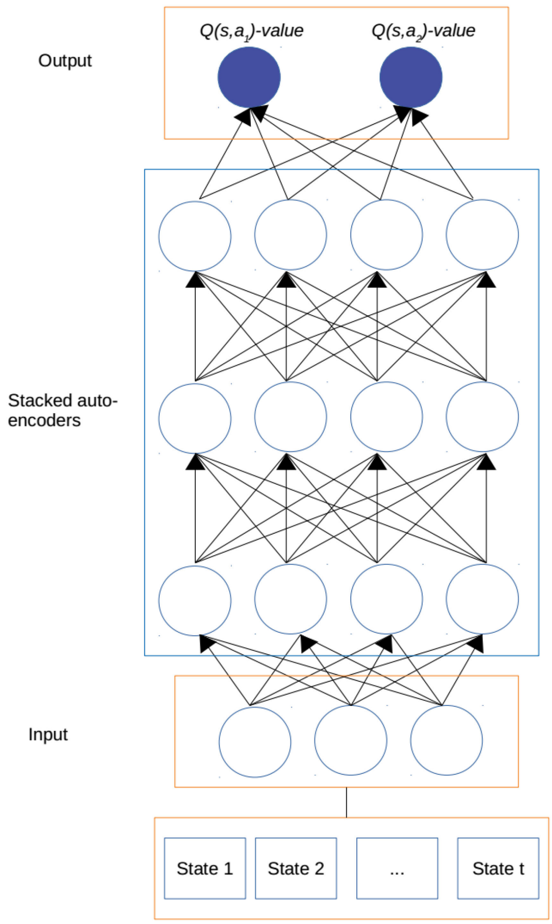


Fig. 12 *Deep SAE network for traffic signal timing*

demand. A four-step demand model (trip generation, trip distribution, modal choice and trip assignment) is the most popular approach for this purpose.

Short-term travel demand prediction is an emerging research topic, which aims to anticipate the traffic demand in the near future, based on current and historical data sources. Short-term travel demand prediction has only attracted much interest recently, with the proliferation of modern data sources such as smart cards, traffic sensors, probes, GPS and CCTV. These data are enormous in volume, rich in detail and complicated in their architecture. Deep learning techniques for short-term travel demand prediction have emerged from the need to exploit these data sources for a versatile, accurate prediction in real time or within a short computation time.

Cheng *et al.* [45] presented several deep learning models to forecast the day-to-day travel demand variations on a large-scale transport network in Florida, USA. The models aim to inherently consider both temporal and spatial correlations. Three models including a DNN, a stacked LSTM network and a feature-level data fusion model were developed and evaluated. The experiment results showed that the stacked LSTM network had the best accuracy among the proposed models.

Ke *et al.* [46] proposed a new deep learning approach to simultaneously consider the spatial, temporal and exogenous dependencies when predicting short-term passenger demand for an on-demand transport service. The proposed model, namely the fusion convolutional long short-term memory network (FCL-Net), is a fusion model of multiple convolutional LSTM (conv-LSTM) layers, standard LSTM layers and convolutional layers. The experiment results showed that the proposed FCL-Net model was able to capture the correlation of spatial-temporal features and exogenous variables to provide better predictive performance, compared to other approaches in literature.

Zhu *et al.* [47] developed a deep learning model to predict car-sharing demand, as an important input for identifying locations of car-sharing depots. Similar to [45, 46], the main benefit of the deep learning structure of the proposed model is to consider the spatial and temporal correlation of demand. In [47], the deep learning model consists of an SAE model to learn the latent spatial-temporal correlation of demand, and a logistic regression layer to enhance the prediction accuracy. In a very recent study, Yao *et al.* [48] proposed a deep multi-view spatial-temporal network, which also aimed to capture the spatial-temporal correlation of demand, to predict the taxi demand. The proposed model has three explicit views: temporal, spatial and semantic view. In earlier research, a sequence learning model based on RNN was introduced by Xu *et al.* [49] to predict future taxi requests in each area of a city, using recent demand and other relevant features (e.g. weather, time, and drop-offs). Besides taxi, the travel demand using technology platform such as Uber is growing rapidly, especially in big cities. Zhu and Laptev [50] introduced a novel end-to-end model architecture for time series prediction using Uber data. In this work, the prediction uncertainty was quantified using a Bayesian Neural Network, which was further applied for large-scale anomaly detection.

Liu and Chen [51] are one of the few recent studies which focus on prediction of passenger demand in public transport. The paper proposed a DNN to forecast the demand for the mass rapid transit system in Taipei. The model took into consideration various explanatory variables, including historical passenger flow, temporal, directional and holiday factors. Another DNN for public transport demand prediction was proposed in [52], which focuses on the stop and stop-to-stop levels of demand. The stop-to-stop level of demand makes this study similar to a dynamic OD estimation study, which is also essential in public transport-related studies.

8 Traffic incident processing

Traffic incident processing is another significant problem in intelligent transportation systems. Major incidents can cause fatal injuries to travellers and long delays on a road network. Therefore, understanding the main cause of incidents and their impact on a traffic network is crucial for a modern transportation management system.

Chen *et al.* used a deep stack denoise autoencoder [53] to model hierarchical feature representation of human mobility [54]. The model was then trained along with incident data, to generate a traffic incident risk map, given real-time input of human mobility (Fig. 13). The experiment results showed that the model is capable of predicting the traffic accident risk by observing human mobility. However, human mobility data may not be sufficient to construct a reliable model for risk prediction. Other factors could be considered, including land use and origin-destination data, to improve the presented model.

For individual car accident prediction, Chen *et al.* proposed a deep genetic algorithm (GA)-optimised neural network to predict rear-end collisions, for use in an automatic avoidance system [55]. In this model, the probability of a collision risk is estimated based on vehicle-to-infrastructure, vehicle-to-vehicle communication and GPS data. To deal with randomness and local optimisation issues, the authors decided to use GAs to optimise the coefficient array and thresholds of the neural network. Experimental results showed that a rationale estimation for collision risk could be generated by the proposed framework, in a car-following scenario. This model can also be applied in the autonomous driving domain, which is reviewed in the following section.

9 Autonomous driving

There are numerous applications of deep learning techniques in computer vision [56]. Since image and video processing have strong relationships with autonomous driving, there is natural interest in the application of deep learning computer vision to autonomous vehicles. Huval *et al.* [57] performed an empirical evaluation of deep learning on highway driving. This study demonstrates that the existing CNN model is able to perform lane and vehicle detection at the required computing speed for a real-time system.

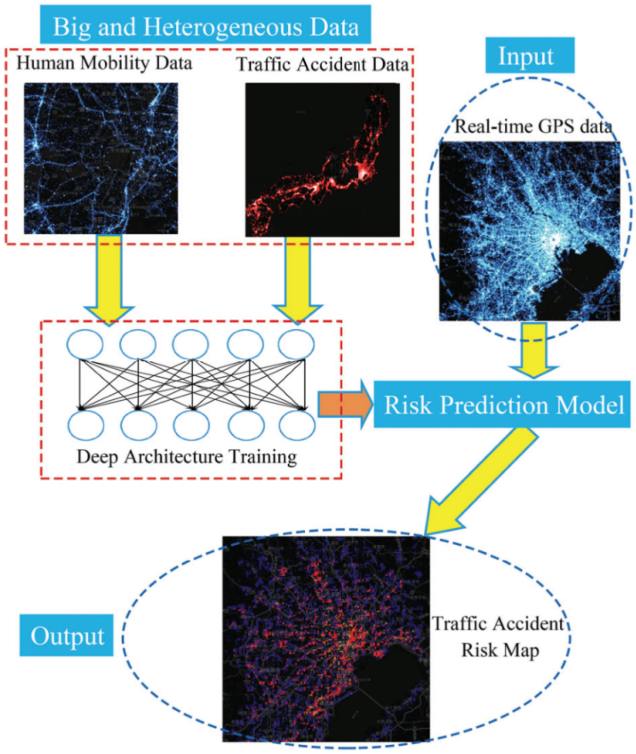


Fig. 13 *Traffic accident risk map prediction using deep learning on human mobility and traffic incident data (flow chart from [54])*

While most research in autonomous driving focused on highway or arterial road driving, an earlier work presented by Hadsell *et al.* proposed a DBN in a long-range vision system, for autonomous off-road driving [58]. In this paper, the features from an input image were extracted and trained a real-time classifier to predict traversability. The experimental outcomes showed that the model was capable of identifying trees, paths, artificial objects and ground in a smooth and accurate way, to the horizon. However, the success of the classifier depended on the size of training data, which required large context-rich image windows.

10 Driver behaviours

With the availability of in-vehicle sensors and GPS data, automatic classifying driving styles of human drivers are an interesting research problem. A high-dimensional representation of driving features is expected to bring advanced benefits to autonomous driving and auto insurance industries. One of the first attempts of applying deep learning to driving behaviour analysis using GPS data was introduced by Dong *et al.* [59]. The authors developed CNNs with 1D convolution and RNNs to study their performance. As a result, high level and interpretable features were effectively extracted by this method, which was able to describe complex driving patterns. Furthermore, deep learning algorithms significantly outperformed classical methods on identifying the driver based on GPS driving patterns.

Another application of deep learning in driver behaviour analysis is detecting a drowsy driver [60]. In this case, the main method is based on computer vision techniques. CNN was utilised to capture latent facial features and complex non-linear feature interactions. The model achieved over 92% accuracy. The system can be applied in real time, to give an early warning to drivers for drowsiness, and to avoid a traffic accident.

11 Conclusions

This paper presented a review of recent deep learning approaches in the transportation domain. In general, deep learning algorithms have been applied to many popular transportation topics, and have demonstrated promising results for traffic data analytics. Most of the presented studies are heavily application-focused, without a strong novel contribution to the theory. However, there has been some effort to integrate spatial-temporal dependencies of the traffic network into deep learning models, but the number of studies is still limited and their scopes are quite specific. For example, in incident impact prediction, very few deep learning studies have been published and it is difficult to find a novel method which can predict the spatial-temporal consequences of an incident on the road network.

12 Acknowledgments

This study was supported by Strategic Research Funding at the Data61, CSIRO, Australia and partly by the NSW Premier's Innovation Initiative project. We would like to express our gratitude to Mr. Ben Itzstein at Data61, CSIRO for proofreading the final version of the manuscript.

13 References

[1] Soysal, M., Schmidt, E.G.: 'Machine learning algorithms for accurate flow-based network traffic classification: evaluation and comparison', *Perform. Eval.*, 2010, **67**, (6), pp. 451–467

[2] Nguyen, H., Cai, C., Chen, F.: 'Automatic classification of traffic incident's severity using machine learning approaches', *IET Intell. Transp. Syst.*, 2017, **11**, (10), pp. 615–623

[3] Fusco, G., Colombaroni, C., Comelli, L., *et al.*: 'Short-term traffic predictions on large urban traffic networks: applications of network-based machine learning models and dynamic traffic assignment models'. 2015 Int. Conf. Models and Technologies for Intelligent Transportation Systems (MT-ITS), Budapest, Hungary, 3 June 2015, pp. 93–101

[4] Ellis, K., Godbole, S., Marshall, S., *et al.*: 'Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms', *Front. Public Health*, 2014, **2**, pp. 39–46

[5] Jahangiri, A., Rakha, H., Dingus, T.A.: 'Predicting red-light running violations at signalized intersections using machine learning techniques'. Transportation Research Board 94th Annual Meeting, 2015 (No. 15-2910)

[6] Lv, Y., Duan, Y., Kang, W., *et al.*: 'Traffic flow prediction with big data: a deep learning approach', *IEEE Trans. Intell. Transp. Syst.*, 2015, **16**, (2), pp. 865–873

[7] LeCun, Y., Bengio, Y., Hinton, G.: 'Deep learning', *Nature*, 2015, **521**, (7553), pp. 436–444

[8] Farabet, C., Couprie, C., Najman, L., *et al.*: 'Learning hierarchical features for scene labeling', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, **35**, (8), pp. 1915–1929

[9] Tompson, J.J., Jain, A., LeCun, Y., *et al.*: 'Joint training of a convolutional network and a graphical model for human pose estimation'. Advances in Neural Information Processing Systems, Montréal, Canada, 2014, pp. 1799–1807

[10] Collobert, R., Weston, J., Bottou, L., *et al.*: 'Natural language processing (almost) from scratch', *J. Mach. Learn. Res.*, 2011, **12**, pp. 2493–2537

[11] Bordes, A., Chopra, S., Weston, J.: 'Question answering with subgraph embeddings'. arXiv preprint arXiv:1406.3676, 14 June 2014

[12] Hinton, G., Deng, L., Yu, D., *et al.*: 'Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups', *IEEE Signal Process. Mag.*, 2012, **29**, (6), pp. 82–97

[13] Sainath, T.N., Mohamed, A.R., Kingsbury, B., *et al.*: 'Deep convolutional neural networks for LVCSR'. 2013 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), Vancouver, Canada, 26 May 2013, pp. 8614–8618

[14] Leung, M.K., Xiong, H.Y., Lee, L.J., *et al.*: 'Deep learning of the tissue-regulated splicing code', *Bioinformatics*, 2014, **30**, (12), p. i121–9

[15] Xiong, H.Y., Alipanahi, B., Lee, L.J., *et al.*: 'The human splicing code reveals new insights into the genetic determinants of disease', *Science*, 2015, **347**, (6218), p. 1254806

[16] Schmidhuber, J.: 'Deep learning in neural networks: an overview', *Neural Netw.*, 2015, **61**, pp. 85–117

[17] Graves, A., Mohamed, A.R., Hinton, G.: 'Speech recognition with deep recurrent neural networks'. 2013 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), Vancouver, Canada, 26 May 2013, pp. 6645–6649

[18] Simonyan, K., Zisserman, A.: 'Very deep convolutional networks for large-scale image recognition'. arXiv preprint arXiv:1409.1556, 4 September 2014

[19] Goodfellow, I., Mirza, M., Courville, A., *et al.*: 'Multi-prediction deep Boltzmann machines'. Advances in Neural Information Processing Systems, Lake Tahoe, CA, USA, 2013, pp. 548–556

[20] Gehring, J., Miao, Y., Metze, F., *et al.*: 'Extracting deep bottleneck features using stacked auto-encoders'. 2013 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), Vancouver, Canada, 26 May 2013, pp. 3377–3381

[21] Sarikaya, R., Hinton, G.E., Deoras, A.: 'Application of deep belief networks for natural language understanding', *IEEE/ACM Trans. Audio Speech Lang. Process. (TASLP)*, 2014, **22**, (4), pp. 778–784

[22] Sainath, T.N., Vinyals, O., Senior, A.: 'Convolutional, long short-term memory, fully connected deep neural networks'. 2015 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Queensland, Australia, 19 April 2015, pp. 4580–4584

[23] Ma, X., Yu, H., Wang, Y., *et al.*: 'Large-scale transportation network congestion evolution prediction using deep learning theory', *PLoS ONE*, 2015, **10**, (3), p. e0119044

[24] Fouladgar, M., Parchami, M., Elmasri, R., *et al.*: 'Scalable deep traffic flow neural networks for urban traffic congestion prediction'. arXiv preprint arXiv:1703.01006, 3 March 2017

[25] Lippi, M., Bertini, M., Frasconi, P.: 'Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning', *IEEE Trans. Intell. Transp. Syst.*, 2013, **14**, (2), pp. 871–882

[26] Huang, W., Song, G., Hong, H., *et al.*: 'Deep architecture for traffic flow prediction: deep belief networks with multitask learning', *IEEE Trans. Intell. Transp. Syst.*, 2014, **15**, (5), pp. 2191–2201

[27] Hinton, G.E., Osindero, S., Teh, Y.W.: 'A fast learning algorithm for deep belief nets', *Neural Comput.*, 2006, **18**, (7), pp. 1527–1554

[28] Caruana, R.: 'Multitask learning'. *Machine learning*, 1997, **28**, (1), pp. 41–75

[29] Jin, F., Sun, S.: 'Neural network multitask learning for traffic flow forecasting'. IEEE Int. Joint Conf. Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1 June 2008, pp. 1897–1901

[30] Smith, B.L., Williams, B.M., Oswald, R.K.: 'Comparison of parametric and nonparametric models for traffic flow forecasting', *Transp. Res. C Emerg. Technol.*, 2002, **10**, (4), pp. 303–321

[31] Polson, N.G., Sokolov, V.O.: 'Deep learning for short-term traffic flow prediction', *Transp. Res. C Emerg. Technol.*, 2017, **79**, pp. 1–7

[32] Zhao, Z., Chen, W., Wu, X., *et al.*: 'LSTM network: a deep learning approach for short-term traffic forecast', *IET Intell. Transp. Syst.*, 2017, **11**, (2), pp. 68–75

[33] Balaji, P.G., German, X., Srinivasan, D.: 'Urban traffic signal control using reinforcement learning agents', *IET Intell. Transp. Syst.*, 2010, **4**, (3), pp. 177–188

[34] Prashanth, L.A., Bhatnagar, S.: 'Reinforcement learning with function approximation for traffic signal control', *IEEE Trans. Intell. Transp. Syst.*, 2011, **12**, (2), pp. 412–421

[35] El-Tantawy, S., Abdulhai, B., Abdelgawad, H.: 'Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (MARLIN-ATSC): methodology and large-scale application on downtown Toronto', *IEEE Trans. Intell. Transp. Syst.*, 2013, **14**, (3), pp. 1140–1150

[36] Genders, W., Razavi, S.: 'Using a deep reinforcement learning agent for traffic signal control'. arXiv preprint arXiv:1611.01142, 3 November 2016

[37] Mnih, V., Kavukcuoglu, K., Silver, D., *et al.*: 'Human-level control through deep reinforcement learning', *Nature*, 2015, **518**, (7540), pp. 529–533

[38] van der Pol, E.: 'Deep reinforcement learning for coordination in traffic light control'. Doctoral dissertation, Master's thesis, University of Amsterdam

[39] Li, L., Lv, Y., Wang, F.Y.: 'Traffic signal timing via deep reinforcement learning', *IEEE/CAA J. Autom. Sin.*, 2016, **3**, (3), pp. 247–254

[40] Konoplich, G.V., Putin, E.O., Filchenkov, A.A.: 'Application of deep learning to the problem of vehicle detection in UAV images'. XIX IEEE Int. Conf. Soft Computing and Measurements (SCM), St. Petersburg, Russia, 2016, pp. 4–6

[41] Zhou, Y., Cheung, N.-M.: 'Vehicle classification using transferable deep neural network features'. arXiv, vol. 1601, 2016

[42] Moussa, G.S.: 'Vehicle type classification with geometric and appearance attributes', *Int. J. Civil Archit. Sci. Eng.*, 2014, **8**, (3), pp. 273–278

[43] Wang, H., Cai, Y., Chen, L.: 'A vehicle detection algorithm based on deep belief network', *Sci. World J.*, 2014, **2014**, pp. 1–7

[44] Adu-Gyamfi, Y.O., Asare, S.K., Sharma, A., *et al.*: 'Automated vehicle recognition with deep convolutional neural networks', 2017

[45] Cheng, Q., Liu, Y., Wei, W., *et al.*: 'Analysis and forecasting of the day-to-day travel demand variations for large-scale transportation networks: a deep learning approach'. Proc. Annual Meeting of the Transportation Research Board, Washington, DC, USA, January 2017

[46] Ke, J., Zheng, H., Yang, H., *et al.*: 'Short-term forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach', *Transp. Res. C Emerg. Technol.*, 2017, **85**, pp. 591–608

[47] Zhu, X., Li, J., Liu, Z., *et al.*: 'Location deployment of depots and resource relocation for connected car-sharing systems through mobile edge computing', *Int. J. Distrib. Sens. Netw.*, 2017, **13**, (6), p. 1550147717711621

[48] Yao, H., Wu, F., Ke, J., *et al.*: 'Deep multi-view spatial-temporal network for taxi demand prediction'. Proc. Thirty-Second AAAI Conf. Artificial Intelligence (AAAI 2018), New Orleans, LA, February 2018

[49] Xu, J., Rahmatizadeh, R., Bölöni, L., *et al.*: 'Real-time prediction of taxi demand using recurrent neural networks', *IEEE Trans. Intell. Transp. Syst.*, 2017, pp. 1–10

[50] Zhu, L., Laptev, N.: 'Deep and confident prediction for time series at uber'. 2017 IEEE Int. Conf. Data Mining Workshops (ICDMW), New Orleans, Louisiana, 18 November 2017, pp. 103–110

[51] Liu, L., Chen, R.C.: 'A MRT daily passenger flow prediction model with different combinations of influential factors'. 2017 31st Int. Conf. Advanced Information Networking and Applications Workshops (WAINA), Taipei, Taiwan, 27 March 2017, pp. 601–605

[52] Baek, J., Sohn, K.: 'Deep-learning architectures to forecast bus ridership at the stop and stop-to-stop levels for dense and crowded bus networks', *Appl. Artif. Intell.*, 2016, **30**, (9), pp. 861–885

[53] Bengio, Y.: 'Learning deep architectures for AI', *Trends Mach. Learn.*, 2009, **2**, (1), pp. 1–27

[54] Chen, Q., Song, X., Yamada, H., *et al.*: 'Learning deep representation from big and heterogeneous data for traffic accident inference'. AAAI, Toronto, Ontario, Canada, 12 February 2016, pp. 338–344

[55] Chen, C., Xiang, H., Qiu, T., *et al.*: 'A rear-end collision prediction scheme based on deep learning in the Internet of vehicles', *J. Parallel Distrib. Comput.*, 2017, **117**, pp. 192–204

[56] Wan, J., Wang, D., Hoi, S.C., *et al.*: 'Deep learning for content-based image retrieval: A comprehensive study'. Proc. 22nd ACM Int. Conf. Multimedia, Orlando, FL, USA, 3 November 2014, pp. 157–166

[57] Huval, B., Wang, T., Tandon, S., *et al.*: 'An empirical evaluation of deep learning on highway driving'. arXiv preprint arXiv:1504.01716, 7 April 2015

[58] Hadsell, R., Erkan, A., Sermanet, P., *et al.*: 'Deep belief net learning in a long-range vision system for autonomous off-road driving'. IEEE/RSJ Int. Conf. Intelligent Robots and Systems, 2008. IROS 2008, Nice, France, 22 September 2008, pp. 628–633

[59] Dong, W., Li, J., Yao, R., *et al.*: 'Characterizing driving styles with deep learning'. arXiv preprint arXiv:1607.03611, 13 July 2016

[60] Dwivedi, K., Biswaranjan, K., Sethi, A.: 'Drowsy driver detection using representation learning'. 2014 IEEE Int. Advance Computing Conf. (IACC), Gurgaon, India, 21 February 2014, pp. 995–999