

This is a post-peer-review, pre-copyedit version of a Journal article published in the

CCF Transactions on Pervasive Computing and Interaction

The final authenticated version is available online at:

<https://link.springer.com/article/10.1007%2Fs42486-020-00039-x>



Deep learning for intelligent traffic sensing and prediction: recent advances and future challenges

Xiaochen Fan¹ · Chaocan Xiang² · Liangyi Gong³ · Xin He² · Yuben Qu⁴ · Saeed Amirgholipour¹ · Yue Xi⁵ · Priyadarsi Nanda¹ · Xiangjian He¹

Received: 12 May 2020 / Accepted: 19 August 2020
© China Computer Federation (CCF) 2020

Abstract

With the emerging concepts of smart cities and intelligent transportation systems, accurate traffic sensing and prediction have become critically important to support urban management and traffic control. In recent years, the rapid uptake of the Internet of Vehicles and the rising pervasiveness of mobile services have produced unprecedented amounts of data to serve traffic sensing and prediction applications. However, it is significantly challenging to fulfill the computation demands by the big traffic data with ever-increasing complexity and diversity. Deep learning, with its powerful capabilities in representation learning and multi-level abstractions, has recently become the most effective approach in many intelligent sensing systems. In this paper, we present an up-to-date literature review on the most advanced research works in deep learning for intelligent traffic sensing and prediction.

Keywords Traffic prediction · Deep learning · Pervasive computing · Intelligent transportation system · Literature review

1 Introduction

The concept of smart cities (Gharaibeh et al. 2017; Silva et al. 2018) has become prevalent across different urban domains that apply information and communication technologies (ICT) to the physical world. By the term of ‘smart city’, it refers to a technology-intensive ecosystem that aims to deliver a wide range of ubiquitous services and

utility applications, such as intelligent transportation, home automation, smart grid, e-health, environment monitoring, and smart logistics (Chamoso et al. 2018; Nagy and Simon 2018). With the rapid population growth and the unprecedentedly growing number of vehicles, intelligent transportation management has become critical for the sustainability of smart cities. The emerging intelligent transportation system (ITS) (Moustaka et al. 2018) is envisioned to

✉ Chaocan Xiang
xiang.chaocan@gmail.com

✉ Xiangjian He
xiangjian.he@uts.edu.au

Xiaochen Fan
xiaochen.fan@student.uts.edu.au

Liangyi Gong
gongliangyi@gmail.com

Xin He
hexin@cqu.edu.cn

Yuben Qu
quyuben@sjtu.edu.cn

Saeed Amirgholipour
Saeed.AmirgholipourKasmani@student.uts.edu.au

Yue Xi
yuexi@mail.nwpu.edu.cn

Priyadarsi Nanda
priyadarsi.nanda@uts.edu.au

¹ School of Electrical and Data Engineering, University of Technology Sydney, Sydney, Australia

² College of Computer Science, Chongqing University, Chongqing, China

³ School of Software and BNRist, Tsinghua University, Beijing, China

⁴ Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

⁵ School of Computer Science, Northwestern Polytechnical University, Xi'an, China

revolutionize the existing transportation management system to a more advanced level. To improve traffic efficiency and alleviate traffic issues, the ITS aims to bring forth the cutting-edge technologies for traffic sensing, data communication, information processing, and intelligent computing. One of the core functions of the ITS is to enhance the accuracy and efficiency of traffic sensing and prediction (Liu et al. 2018). Accurate sensing and reliable prediction on traffic status are fundamentally essential for various urban transportation services and traffic applications. For example, with the precise information on future traffic predictions by ITS, transportation operators would have comprehensive knowledge in their decision-making for traffic dispersion and congestion management (Wang et al. 2019c).

In ITS, traffic sensing data can be obtained from diverse sources, ranging from conventional traffic monitoring infrastructures to ubiquitous mobile and IoT devices. Traditional traffic infrastructures, including loop detectors, traffic cameras, and radars, are commonly deployed at road intersections to monitor road traffic and detect the presence of passing vehicles (Nagy and Simon 2018). However, the high costs in deployment and maintenance impede their extension on a city scale, thus limiting the coverage of traffic sensing data. Thanks to the proliferation of pervasive mobile and IoT devices, more advanced traffic sensing technologies are integrated into ITS by exploiting global positioning system (GPS), automatic fare collection (AFC) system, mobile cellular stations, and social media platforms, etc. For example, with the equipped GPS sensors, smart mobile devices can generate the mobility trajectories of the onboard participants, thereby providing accurate traffic sensing data. Such emerging mobility data sources substantially break the bottleneck of data insufficiency and further make it possible to fuse information from multiple traffic sensing modalities.

To leverage the diversity and variety of traffic sensing data for fine-grained prediction, numerous research efforts have been devoted to devising sophisticated computation models. Traditional traffic prediction methodologies generally apply statistical models to analyze historical traffic data, and further use handcrafted features to conduct traffic prediction. Meanwhile, such statistical models are invariant and cannot be extended for large-scale traffic predictions, as they cannot model comprehensive features (e.g., spatial features) for the entire transportation networks. To achieve more advanced feature learning in traffic prediction, machine learning models have been applied to address the non-linearity and exploit spatiotemporal correlations in traffic sensing data. These models are typically with the advantages of data processing capacity, implementation flexibility and generalization ability. Classical machine learning models for traffic prediction include non-parametric Bayesian networks, K-nearest neighbors (KNN), support vector machine (SVM), and artificial neural networks (ANN). Nevertheless,

the amount of traffic sensing data in ITS has been growing from Trillion-byte level to Peta-byte level, which substantially calls for processing models with capabilities in feature extraction. In this regard, the classic machine learning models with shallow architectures only have limited latent spaces, which restrict their abstractive representation learning on big traffic data for prediction purposes.

In recent years, deep learning is making significant achievements with state-of-the-art performance in Artificial Intelligence (AI) community. Modern deep neural networks usually consist of tens or hundreds of successive layers (LeCun et al. 2015) to discover intricate structures from high-dimensional data, and further extract hierarchical representations in feature learning. As a consequence, the researchers in the ITS community have recognized the importance of deep learning and already started to exploit deep neural networks for intelligent traffic sensing and prediction. The integration of deep learning and ITS has been well justified by that deep learning can develop complex representations from large-scale traffic datasets in an incremental, layer-by-layer way. Moreover, the incremental intermediate representations of spatial and temporal traffic states can be jointly learned by the deep-learning models.

Scope of the survey. In this paper, we present a comprehensive, up-to-date survey of deep learning for intelligent traffic sensing and prediction. Our goal is to thoroughly cover various aspects of deep learning and outline deep-learning models that can assist different applications of ITS. We first provide an overview of deep learning for ITS, covering the preliminaries of intelligent traffic sensing and prediction with the recent advances driven by deep learning techniques. Aside from taxonomically reviewing the existing related works, we investigate the pros and cons of various deep-learning models for serving different traffic prediction applications in ITS. We further present several key insights into the future research challenges and directions of this cross-domain research field. We hope that this article can benefit the research community with some comprehensive knowledge of the up-to-date developments in deep learning for ITS.

1.1 Our contributions

We summarize our contributions in this paper as follows:

- We deliver a systematic review of deep learning, particularly for intelligent traffic sensing and prediction in ITS.
- We investigate the various types of representative deep-learning models and provide detail-oriented analysis regarding their customization to different ITS applications.
- We scrutinize the application-level aspects from hundreds of related papers that contribute to traffic sensing

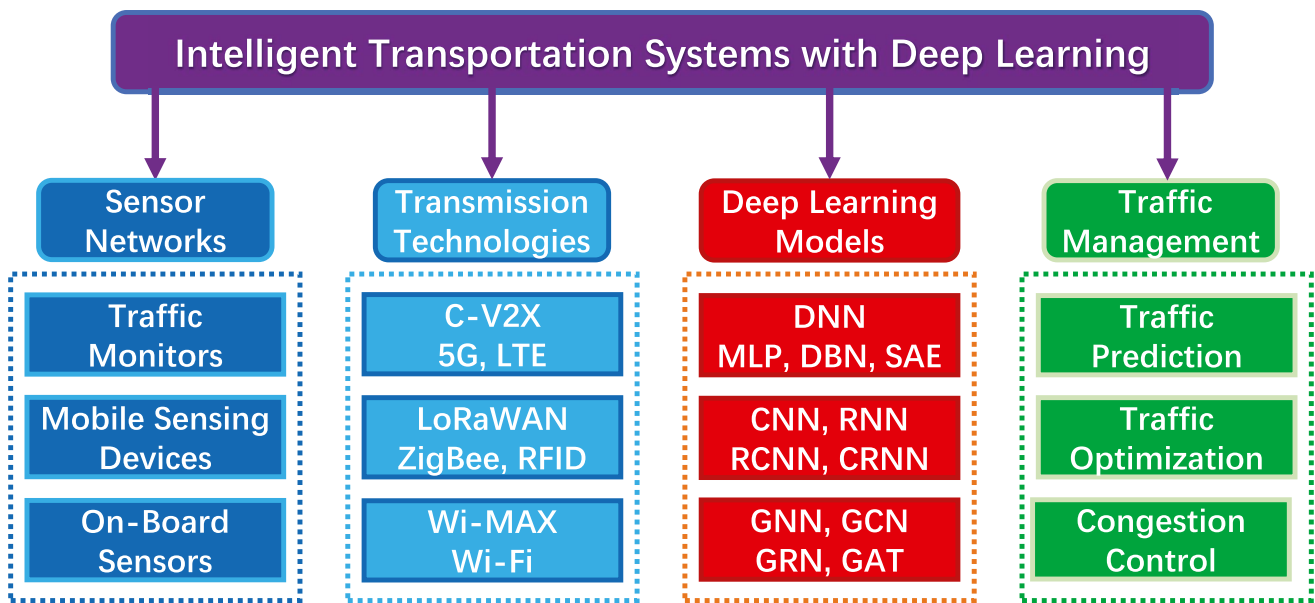


Fig. 1 Hierarchical architecture of urban ITS with deep learning

and prediction for ITS, featuring the in-depth analysis from different perspectives.

- We thoroughly discuss the emerging research challenges for deep learning in ITS over several essential areas, and we envision the future directions of this promising research field.

The rest of this paper is organized as follows. Section 2 provides an overview of ITS with a summary of existing surveys that cover machine learning and deep learning techniques. Then, Sect. 3 examines the most notable deep neural network models for intelligent traffic sensing and prediction. Section 4 reviews the categorized applications of ITS driven by deep-learning techniques. Section 5 presents open issues and future challenges in deep learning for ITS. Finally, Sect. 6 concludes the paper.

2 Traffic sensing and prediction in ITS: an overview

With the ever-expanding traffic networks and the diverse traffic sensing technologies, traffic prediction has become more daunting at present. Though deep learning and ITSs are two independent areas, the unprecedented amount of sensing data has seriously challenged existing computation methodologies of traffic data processing and traffic prediction. Particularly, traffic sensing data from various types of sources have complex correlations with non-linearity, cross-domain, and time-varying properties (Nagy and Simon

2018). As a consequence, the emerging sophisticated traffic prediction problems cannot be simply attained by the existing conventional machine-learning techniques for the following reasons.

First, the traditional machine-learning models only have shallow space for representation learning, which cannot preserve enough useful features for large traffic datasets. Second, the shallow machine-learning models rely on handcrafted features and cannot automatically extract high-dimensional representations for joint learning. Third, despite the explosive growth of input traffic sensing data, the classical machine-learning models cannot improve their performance by developing more valuable representations in traffic prediction. Therefore, deep learning-driven traffic prediction becomes inevitable, imperative, and viable. In this section, we first present an overview of the ITS architecture and its key components. Then, we introduce the related review articles and further highlight the necessity of this up-to-date survey.

2.1 Key components in ITS

As illustrated in Fig. 1, there are basically four major components in the architecture of an ITS, namely the sensor networks, transmission technologies, deep-learning models, and traffic management operations.

First, traffic sensor networks are the primary subsystem that in charge of collecting traffic information on road networks from vehicles and mobile devices (mainly via wireless sensing). Second, wireless communication technologies are critical for transmitting real-time traffic data between

Table 1 Summary of previous surveys and reviews in traffic prediction and deep learning

Publication	Years	Scopes		
		Traffic prediction	Deep learning	Other ML
Lee and Gerla (2010)	2010			
Bolshinsky and Friedman (2012)	2012	✓		✓
Li et al. (2013)	2013			
Djahel et al. (2014)	2014	✓		✓
Castillo et al. (2015)	2015	✓		✓
Nellore and Hancke (2016)	2016			
Seo et al. (2017)	2017	✓		✓
Liu et al. (2018)	2018	✓	✓	
Nagy and Simon (2018)	2018	✓		✓
Zhu et al. (2018)	2018	✓	✓	
Wang et al. (2019c)	2019	✓	✓	✓
Do et al. (2019)	2019	✓	✓	

traffic sensors and traffic monitoring systems. The above two components are out of our scope in this survey; therefore, we provide preliminary knowledge of them as follows. The detailed technical information of traffic sensor networks is provided in Table 5 of Appendix A.1. The wireless communication technologies in ITS are classified in Table 6 of Appendix A.2, based on the communication standard, data rate, and topology.

Third, deep-learning models are the core component for processing ITS information with deep neural networks. Substantially, deep learning is a subfield of machine learning (ML). With multiple successive layers of representations, deep-learning models are powerful in high-level representation learning and feature extraction (Zhang et al. 2019b). Moreover, the advanced graphics processing units (GPU) and parallel computing infrastructures of traffic data centers further accommodate deep-learning models to perform city-wide traffic prediction tasks within milliseconds (Wang et al. 2019c). We believe that deep learning will continue to revolutionize ITS by enhancing its capability, integrality, and sustainability.

Fourth, traffic management operations are the last step to put the information from traffic sensing and deep-learning models into practice. The traffic management units include traffic prediction (an essential scope in this article), traffic optimization, and congestion control.

2.2 Previous efforts of related reviews and surveys

We list the previous surveys and reviews that are related to ITS and deep learning in Table 1. Among the above works, Lee and Gerla (2010) surveyed the developments of vehicle-to-Vehicle sensing techniques for vehicular networks. Bolshinsky and Friedman (2012) reviewed

the conventional methods and initial takes of neural networks for traffic prediction. Secondly, Li et al. (2013) presented a survey on traffic control and highlighted the design philosophy of traffic control systems. Djahel et al. (2014) presented a study on different technologies in traffic management systems, ranging from information collection to service delivery. Then, Castillo et al. (2015) studied traffic sensor placement, flow observability, and flow prediction in traffic networks. More recently, Nellore and Hancke (2016) provided a taxonomy of different wireless sensor networks and wireless communication technologies for urban traffic management. Seo et al. (2017) summarized the models, datasets, and methodologies for traffic state estimation, particularly on highways. Liu et al. (2018) investigated urban traffic prediction with various mobility data using deep learning, and further compared basic deep-learning models for processing traffic indicator information. Similarly, Nagy and Simon (2018) focused on urban traffic sensing and prediction methods by covering different data sources, data models, and prediction models. Zhu et al. (2018) surveyed the ITS from the perspective of big data and discussed the issues of big data in ITS from several aspects. Moreover, Wang et al. (2019c) focused on applying deep-learning models to achieve high-accuracy visual recognition of traffic signs. At last, Do et al. (2019) presented a review of short-term traffic state prediction with neural network-based models.

Summary. The recent development in deep learning has produced hundreds of papers contributing to the applications of intelligent traffic sensing and prediction. Despite that the above articles have concluded some initial takes of machine learning techniques in ITS, there still lacks an up-to-date survey for researchers to gain sufficient knowledge on the latest advancements in deep learning for ITS. In this paper, we typically focus on the significant results of deep

Table 2 Enabling techniques for deep learning-driven traffic sensing and prediction

Enabling techniques	Typical examples	Function and scope	Performance improvement
WSNs and big data	Road sensors on-board GPS	Large-scale traffic sensing	High
Deep learning libraries	TensorFlow Keras, PyTorch	Fast building DL architectures	Medium
Neural network models	DNN, CNN, RNN LSTM, GAN, DRL	Layer-by-layer feature extraction	High
Optimization algorithms	Adam, RMSprop Adagrad/Delta	Accelerate training with convergence	Medium
Cloud-based platforms	Google Colab Azure, SageMaker	Cloud-based online environment for DL	Medium
High-performance processing units	GPU TPU	Enable highly parallel computing	High

learning for ITS in the last five years and restrict our scope to the related papers from premier conferences and top-tier journals, to provide the readers with a high-level comprehensive review.

3 Deep learning preliminaries

In this section, we give a brief introduction to deep learning and its preliminaries. Then, we present the most popular deep-learning models that can be applied for traffic data processing and prediction.

3.1 A brief introduction to deep learning

Deep learning (LeCun et al. 2015) is one of the sub-branches in machine learning, and deep-learning methods are essentially representation-learning methods with multiple levels of representations. In recent years, deep learning has achieved tremendous advances (Schmidhuber 2015) in computer vision, pattern recognition, language translation, robots, and self-driving. Deep-learning models learn representations from raw data in an incremental, layer-by-layer manner. Thereby, complex and high-dimensional representations can be developed. In particular, these representations are learned via different models of deep neural networks (Goodfellow et al. 2016), i.e., the long chains of geometric functions and operations that are structured into modules called layers. These layers are parameterized by ‘weights’, which can be learned and updated during the training process. Indeed, the knowledge of a deep-learning model is stored in its weights. During this process, the critical aspect of deep learning is the automatic feature extraction, as features are learned using a feedback signal, not handcrafted. In the following, we introduce the evolution of deep learning together with its milestones, enabling technologies and universal workflow.

Deep learning is not a relatively new subfield of machine learning, and the milestone works of its current take-off can be traced back to the late 1980s (Chollet 2017). Notably, Rumelhart et al. (1986) described a new

learning procedure, i.e., backpropagation, to efficiently train the neural networks. Subsequently, LeCun et al. (1990) further presented the first convolutional neural network (CNN) that can be trained by backpropagation. Furthermore, Hochreiter and Schmidhuber (1997) introduced another gradient-based model, long short-term memory (LSTM), which later became one of the standard deep-learning models. Despite all the above milestones, it takes nearly another two decades for deep learning to break through some major bottlenecks for its boom. To conclude, there are three driving forces, i.e., hardware, data, and algorithms, that contribute to the tremendous developments of modern deep learning, and we explicitly introduce the detailed rationale as follows.

First, the typical deep-learning models would require exceeding computational capacity that off-the-shelf CPUs cannot provide. Fortunately, since the early 2000s, some technology companies (e.g., NVIDIA and AMD) have been massively investing parallel chips (known as GPU) for empowering and rendering complex 3D scenes in video games. In 2007, NVIDIA launched CUDA (NVIDIA: Cuda 2019), a parallel computing platform and programming model for GPUs to replace CPUs in various parallel computing scenarios. As deep neural networks are highly parallelizable with matrix multiplications, the scientific research community is driven to implement and benefit from more sophisticated deep-learning models on GPUs. In 2016, the technology giant Google announced the tensor processing unit (TPU) at the Google I/O conference and revealed that TPUs had been used in their data center for years. Second, as deep learning is an engineering science, deep-learning models are strictly reliant on data. However, the Big Data becomes available till the Internet took off over the last 20 years together with the exponential growth of storage in hardware. Third, the feedback signal used for deep-learning model training can quickly fade away, particularly when the number of layers increased. Such that, a reliable way to train the complex deep neural networks is of great necessity. It was until the early 2010s, a series of critical algorithmic improvements for gradient propagation were discovered,

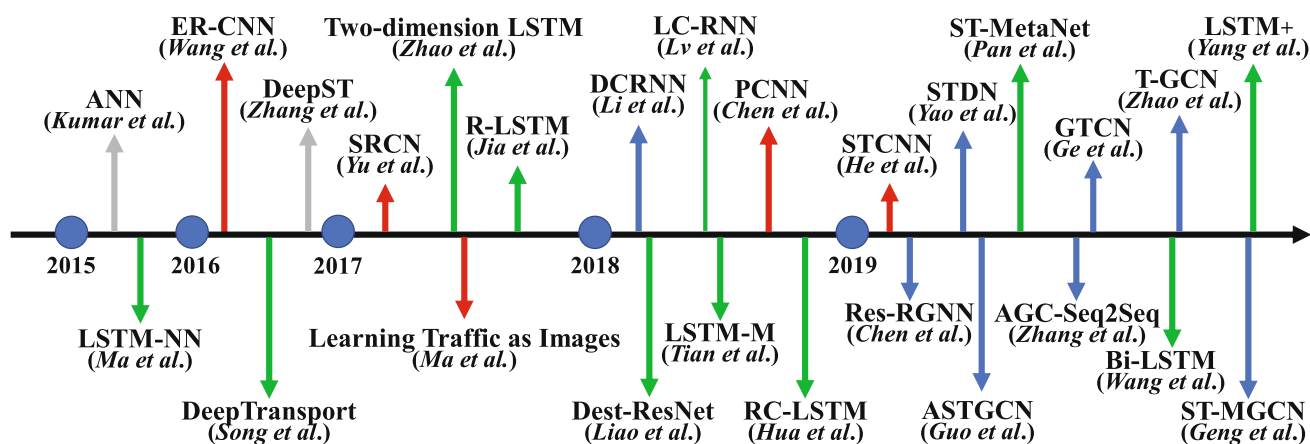


Fig. 2 Major milestones of deep learning-driven traffic prediction since 2015. Currently, the most popular models are graph neural networks (in blue) for network-wide traffic prediction. Other popular deep-learning for ITS include RNN/LSTM (in green) and CNN (in red)

including batch normalization, residual connections, and depth-wise separable convolutions (Chollet 2017).

In summary, we conclude the enabling techniques for deep learning-driven traffic sensing and prediction in Table 2, including big sensing data, integrated libraries, neural network models, optimization algorithms, online platforms, and high-performance hardware units.

3.2 Deep learning for traffic sensing and prediction: a brief chronology

Before reviewing a variety of traffic prediction related studies in Sect. 4, we summarize some significant milestones in research studies of deep learning-driven traffic prediction in terms of the temporal dimension in Fig. 2. From this timeline, we observe the research development of urban traffic prediction as follows. First, the initial takes of deep learning-driven traffic prediction are based on basic deep neural networks, such as ANN, MLP, DBN, and SAE. For example, Kumar et al. (2015) applied an ANN to incorporate historical traffic data and temporal dependencies for making traffic predictions, and they achieved better performance than conventional machine-learning methods. Nevertheless, the fully connected structure (dense layers) of ANN makes it computation-intensive to process the explosively growing traffic data and is incapable of learning long-term dependencies. Instead, researchers start to propose more efficient deep-learning models based on convolutional neural networks, recurrent neural networks, and their combinations.

CNN models are capable of extracting network-wide spatial features from traffic data that is formatted like images (matrices). For instance, Ma et al. (2017) presented a CNN model to ‘learn traffic as images’ and achieved surprising improvement in traffic speed prediction. Other examples of traffic prediction models based on CNN include ER-CNN

(Wang et al. 2016), SRCN (Yu et al. 2017b), PCNN (Chen et al. 2018), and STCNN (He et al. 2019). Regarding the LSTM models, they can preserve long-term temporal dependencies in historical data without vanishing gradients and achieve better performance in traffic prediction. Since traffic data are basically time series data, a variety of LSTM-based variants have been developed for traffic prediction, including two-dimension LSTM (Zhao et al. 2017), LC-RNN (Lv et al. 2018), ST-MetaNet (Pan et al. 2019), Bi-LSTM (Wang et al. 2019a) and LSTM+ (Yang et al. 2019a).

A newly emerging trend of deep learning-driven traffic prediction is the graph neural network (GNN). Since road networks can be modeled as graph structures, and traffic data can also be represented in the forms of graphs (Wu et al. 2020). Existing GNN driven traffic prediction models can be categorized as into three categories: (1) Recurrent GNNs [Res-RGNN (Chen et al. 2019b)]; (2) Convolutional GNNs [DCRNN (Li et al. 2017), AGC-Seq2Seq (Zhang et al. 2019a) and T-GCN (Zhao et al. 2019)]; (3) Spatial-temporal GNNs [ST-MGCN (Geng et al. 2019), GTCN (Ge et al. 2019) and ASTGCN (Guo et al. 2019)].

With respect to the deep learning-related papers in ITS to be reviewed in Sect. 4, we provide a top-down summary in Table 3 by categorizing deep-learning models and data sources.

In terms of traffic data sources, traffic infrastructures are the most reliable and sustainable sources to provide ubiquitous and direct traffic sensing data. Meanwhile, on-board GPS and smartphones have come up as two alternative data sources for traffic prediction. As both of them provide continuous location information of vehicles and passengers, researchers can convert the trajectory data into meaningful information on traffic speed and traffic volume.

As for deep-learning models, various neural networks have been employed to perform traffic prediction tasks.

Table 3 Summary of deep learning-related papers for intelligent traffic sensing and prediction in terms of data sources and deep-learning models

	DNN	CNN	RNN	RCNN	GNN
Traffic sensors	Kumar et al. (2015), Huang et al. (2014), Lv et al. (2014), Yang et al. (2016), Jia et al. (2016), He et al. (2018), Koedwiyadi et al. (2016)	Deng et al. (2019), Ma et al. (2015), Yu et al. (2017), Wang et al. (2019a), Zheng et al. (2019)	Fu et al. (2016b), Zhao et al. (2017), Kang et al. (2017), Jia et al. (2017), Tian et al. (2018), Liao et al. (2018b), Yang et al. (2019a), Duan et al. (2016), Song et al. (2016)	Li et al. (2017), Yuan et al. (2018), Di et al. (2019), Wang et al. (2016)	Chen et al. (2019b), Ge et al. (2019), Zhao et al. (2019), Diao et al. (2019)
On-board GPS	Lemieux and Ma (2015), Chen et al. (2016), He et al. (2018)	Zhang et al. (2017), Ma et al. (2017), Jo et al. (2018), Chen et al. (2018), He and Shin (2019), Wang et al. (2018)	Pan et al. (2019), Liao et al. (2018a), Duan et al. (2016)	Yao et al. (2019), Lv et al. (2018), Yu et al. (2017b), He et al. (2019)	Guo et al. (2019), Zhao et al. (2019)
Mobile phones	Chen et al. (2016), Zhang et al. (2018a)	Dabiri and Heaslip (2018)	Google (2019), Liao et al. (2018a), Song et al. (2016), Jiang et al. (2018a, b), Fan et al. (2018)	Ke et al. (2017), Yao et al. (2018)	Zhang et al. (2019a), Geng et al. (2019)

First, since traffic data is inherently sequential and exhibits temporal correlations, the recurrent neural network is frequently used to capture temporal dependencies in traffic data. Second, as road networks have specific topologies, the network-wide traffic data has spatial correlations in nature. To exploit such property in traffic data, convolutional neural networks are also employed to automatically extract non-linear features from traffic data that are transformed into 2-dimensional shapes. Third, CNNs and RNNs are further combined as spatial-temporal neural networks to jointly capture spatial and temporal correlations in more complex traffic prediction tasks. Moreover, the emerging graph neural network models, including recurrent GNNs, convolutional GNNs, and spatial-temporal GNNs, can effectively capture the hidden patterns of Euclidean data, considering that the graph structure arises naturally in traffic networks. At last, deep reinforcement learning models are further developed for traffic control and autonomous driving.

3.3 Deep-learning models for ITS

As a specific subfield of machine learning, deep learning focuses on learning successive layers of increasingly meaningful representations from raw data. In particular, deep learning has achieved near-human-level performance in image processing, speech recognition, and language translation (Goodfellow et al. 2016). In this section, we introduce the preliminaries about deep-learning models¹ and discuss how to apply them in traffic sensing and prediction of ITS.

Deep neural networks. The deep neural network (DNN) is the initial artificial neural network (ANN), including multi-layer perceptron (MLP), deep belief network (DBN), and stacked auto-encoder (SAE). Fig. 3 shows the general architectures of different DNNs, where the main differences are the connections between hidden layers. As shown in Fig. 3a, the MLP has one input layer, one or several hidden layers, and one output layer. In the MLP, each unit in a layer is densely connected to all the units in the following layer. At its hidden layer, the input vector is multiplied by the weight matrix, whose parameters are further trained in a supervised manner with backpropagation. Moreover, an activation function [e.g., sigmoid or Rectified Linear Unit (Glorot et al. 2011)] is employed to generate the output and improve the non-linearity of the model.

As a simple feedforward artificial neural network model, MLP shows promising performance (Kumar et al. 2015) in traffic prediction when there are sufficient labeled training data. However, due to the fully-connected structure, MLP could entail high computation complexity with low

¹ Technical details and implementations of deep-learning models can be referred from <https://github.com/rasbt/deeplearning-models>.

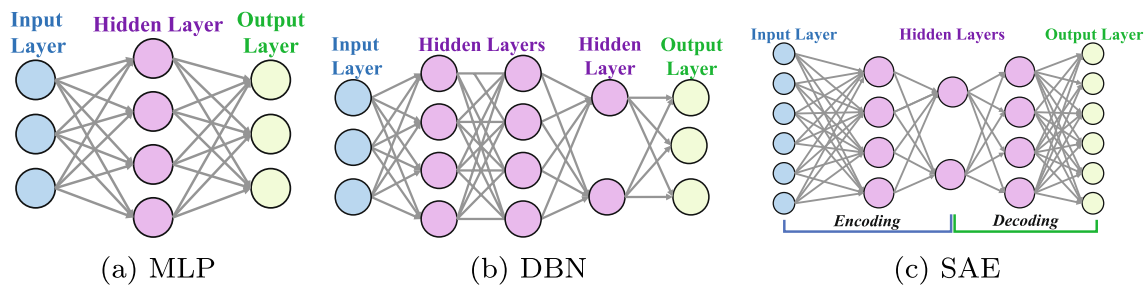


Fig. 3 The basic structures of deep neural networks (Liu et al. 2017)

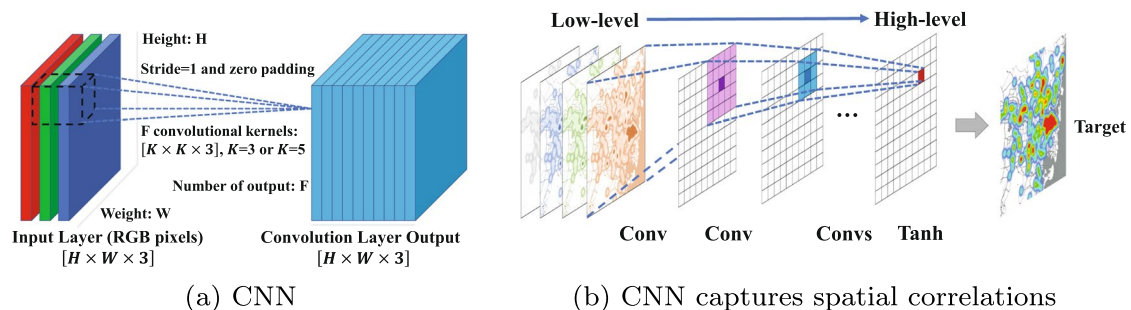


Fig. 4 CNN's structure and examples (Course CS231n 2019)

convergence efficiency. Therefore, some variants of MLP have been proposed, including DBN (Fig. 3b) and SAE (Fig. 3c). In general, the bottom layers of DBN and SAE models are stacked with hidden variables for unsupervised pre-training. For example, DBN models employ stacked modules of Restricted Boltzmann Machines (RBM) (Le Roux and Bengio 2008) as the bottom layers, where layers are connected, but units are not. The DBN models follow a layer-by-layer procedure for learning the top-down, generative weights. The successful implementations of DBN models in traffic prediction include (Koesdwiady et al. 2016; Souza et al. 2016). In terms of SAE, the hidden layers perform encoding on the input data, and the output layer reconstructs the input layer from the encoded feature representations. In traffic prediction, the objective of an SAE model is to minimize the reconstruction errors, where the encoding and decoding operations are inverse to each other in training (Yang et al. 2016).

Convolutional neural networks. The convolutional neural network is comprised of a set of learnable filters (kernels) to process images or image-like data that has multiple dimensions (e.g., width, height, and depth). As shown in Fig. 4a, the convolution operations will slide over the input image data. Each filter outputs the weighted sum of each pixel's neighbors by element-wise multiplying the filter's weights with the original pixel values. The above process

will be repeated for all pixels, and the convolution operation over the image will result in a feature map of the filter. After each convolution operation, the CNN further employs pooling layers to down-sample feature maps, normally by max-pooling operations. To induce the spatial hierarchies of representation and reduce the number of parameters, the max-pooling operations process the feature maps by outputting the max value of each channel. Particularly, CNN models have two essential properties: first, they learn representations that are translation invariant, making convolution layers highly data-efficient and modular; second, they learn spatial hierarchies of local patterns in a down-sampling manner (as shown in Fig. 4b), allowing convolution layers to extract successive spatial extent of the input data. The examples of CNN-based traffic prediction include traffic volume prediction (Yao et al. 2019; Deng et al. 2019) and traffic speed prediction (Ma et al. 2017; Jo et al. 2018).

Recurrent neural networks. The recurrent neural networks (RNN) are designed to model sequential data, especially when sequential or temporal correlations exist between data samples. As shown in Fig. 5a, an RNN processes sequential data by iterating through each sequence element and maintaining a state that contains information relative to the previous input data. The RNN model has an internal loop, and when it is unrolled, each copy of the network outputs some information to the next successor.

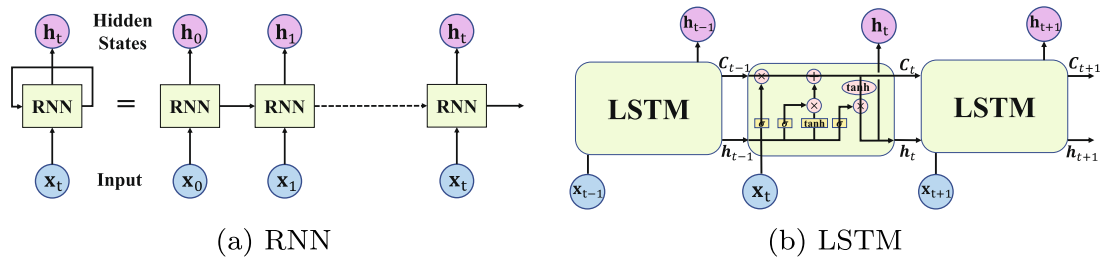


Fig. 5 The structures of RNN and LSTM (Olah 2015)

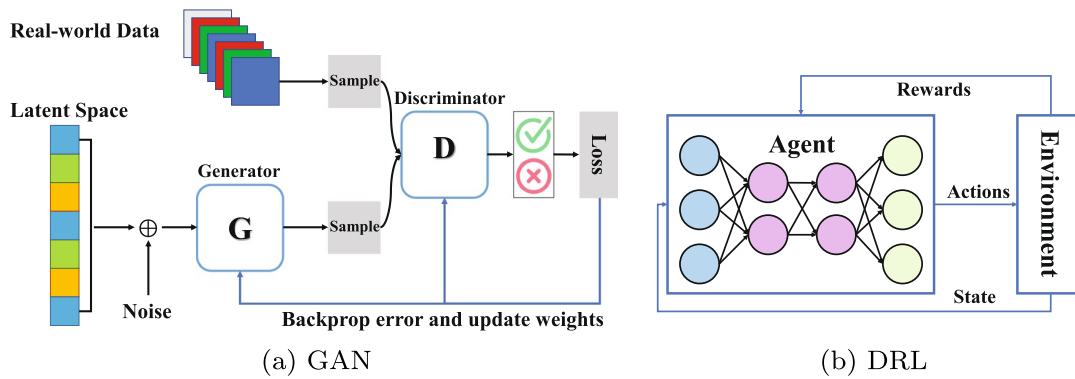


Fig. 6 The generative adversarial network and deep reinforcement learning model (Bau et al. 2019)

However, RNNs suffer from the problem of vanishing gradients, and they can hardly capture long-term dependencies in practice (Bengio et al. 1994). For this reason, different variants of RNNs have been proposed, and the long short-term memory networks (Fig. 5b) can successfully prevent vanishing in processing sequential data (Hochreiter and Schmidhuber 1997). The key idea of the LSTM is the cell state, i.e., a horizontal line running through the top of the LSTM model. Moreover, the LSTM updates information to the cell state with three different gates, including the input gate, the forget gate, and the output gate. Given a time-sequential data of $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_t, \dots, \mathbf{x}_T)$, where $\mathbf{x}_t \in \mathbb{R}^N$, the LSTM updates its cell state \mathbf{s}_t and hidden state \mathbf{h}_t at time interval t as:

$$\mathbf{s}_t = \mathbf{f}_t \odot \mathbf{s}_{t-1} + \mathbf{i}_t \odot \tanh(\mathbf{W}_s[\mathbf{h}_{t-1}; \mathbf{x}_t] + \mathbf{b}_s), \tag{1}$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{s}_t), \tag{2}$$

where $\mathbf{i}_t = \sigma(\mathbf{W}_i[\mathbf{h}_{t-1}; \mathbf{x}_t] + \mathbf{b}_i)$ is the input gate, $\mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}; \mathbf{x}_t] + \mathbf{b}_f)$ is the forget gate, $\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}; \mathbf{x}_t] + \mathbf{b}_o)$ is the output gate, $[\cdot; \cdot]$ is a concatenation operation; σ is a logistic sigmoid function, \odot is a pointwise multiplication, $\mathbf{W}_f, \mathbf{W}_i, \mathbf{W}_o, \mathbf{W}_s$ and $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_o, \mathbf{b}_s$ are the learnable parameters.

Another popular variant of RNN is the gated recurrent unit (GRU), which is a simplified LSTM that has no separate memory cells. In specific, a GRU cell has only two gates, i.e., an update gate for determining the amount of memory to retain, and a reset gate for calculating the amount of information from the previous state to preserve. As the traffic data on a road segment is essentially time series, there have been numerous traffic prediction models based on RNNs (Ma et al. 2015; Fu et al. 2016; Zhao et al. 2017; Kang et al. 2017). We will introduce the details of these works in Sect. 4.

Generative adversarial networks. The generative adversarial network (GAN) is indeed an alternative to variational auto-encoders for learning latent spaces from given data (Lv et al. 2018). GANs are capable of generating reasonably realistic synthetic data such as images, by forcing the generated data to be statistically indistinguishable from the real ones. As illustrated in Fig. 6a, a GAN model typically consists of two parts, i.e., a generator network \mathcal{G} and a discriminator network \mathcal{D} . The former seeks to approximate the target data distribution from training data, and the latter predicts or estimates the probability that a generated sample is from the training set or created by the generator network. Both \mathcal{G} and

\mathcal{D} are neural networks, and their training process is iterative to supervise each other.

Taking image generation as an example, the objective of \mathcal{G} is to be trained to fool the \mathcal{D} with increasingly realistic images during the training process. In contrast, the discriminator \mathcal{D} will continuously adapt to set a higher bar of realism for the generated images. Consequently, after training is finished, the generator \mathcal{G} is capable of turning any point in its input space into a believable image (Chollet 2017). In traffic prediction studies, different GAN models have been adopted for traffic data imputation (Chen et al. 2019c), traffic-state estimation (Liang et al. 2018), and pattern-sensitive traffic prediction (Lin et al. 2018).

Deep reinforcement learning. Deep reinforcement learning (DRL) uses deep neural networks to develop an agent for interacting with an environment and update policies to gain maximum long-term rewards over a series of time intervals. During each time step t , the agent would receive some observations o_t from the environment and must perform an action a_t that will be transmitted back to the environment. Ultimately, the agent would receive a reward r_t from the environment and start the next session. The behaviors of a DRL agent are governed by a *policy*, which is a function that maps from observations of the environment to the actions of the agent. The objective of the DRL is to produce a good policy that an agent can find the best action to perform accordingly. The general process of DRL is illustrated in Fig. 6b, and the major breakthroughs of DRL's achievement include Deep Q-network (Mnih et al. 2015) and AlphaGo (Silver et al. 2016). In traffic-related studies, DRL models are implemented for traffic prediction (Li et al. 2016a), traffic signal control (Wei et al. 2018), data recovery (Tang et al. 2019) and resource deployment (Li et al. 2020).

4 Applications of deep learning in traffic prediction for ITS

Deep learning has been widely applied to a range of traffic-related applications for smart cities. In this section, we present the state-of-the-art research works across the most critical domains of traffic prediction. Specifically, we first introduce the essential prerequisite of traffic prediction, *i.e.*, traffic data models. Then, we review all relevant studies in five categories that deep learning has been making remarkable advances.

4.1 Traffic data models

Data models characterize the dimension, granularity, and relevant features of traffic measurements. In particular, two

main characteristics of traffic data should be taken into consideration when creating data models.

(1) **Time interval.** In the real-world datasets, time intervals of traffic measurements range from seconds, minutes to hours. Meanwhile, the most commonly used time intervals are in minute-scales (e.g., 5–30 min per sample). Moreover, it also depends on the desired traffic prediction horizon of traffic prediction models that whether a specific time interval should be adopted. For example, the hourly scale can be used for predicting network-wide traffic mobility, and the minute scale can be helpful when predicting rush hour traffic jams (Nellore and Hancke 2016).

(2) **Spatial property.** Traffic data that covers a point, a road, a region, or even an urban area would have different spatial dimensions. Subsequently, different data models should be applied to traffic data with different spatial dimensionality. Typically, there are three major data models for traffic sensing data, *i.e.*, the scalar model, the vector model, and the matrix model. The scalar model is the simplest data model for traffic data on a single road segment. For example, given a position-fixed sensor p , its traffic flow measurement at time t can be denoted by $f_{p,t}$, where $t = 1, 2, \dots, T$. The scalar models can only represent basic traffic sensing data (*i.e.*, traffic volume or traffic speed) at a single and fixed position. The vector model is more advanced in describing actual traffic states over a period of time. The vector models can be categorized into the univariate type and the multivariate type. For the univariate one, a vector model denotes the current traffic state measured by a specific sensor at time interval t is denoted by $\mathbf{f}_t = \{f_{t-1}, \dots, f_t\}$, where l is the 'lag'. For the multivariate case, given traffic flow measurements from multiple traffic sensors in a road network, the overall traffic data can be denoted by $\mathbf{F}_t = \{\mathbf{f}_t^1, \mathbf{f}_t^2, \dots, \mathbf{f}_t^N\}$, where N is the total number of sensors. The multivariate vector models can be useful to identify spatial correlations for the downstream and upstream traffic in adjacent road segments. The matrix model is the most fine-grained traffic data model, as it can preserve both spatial information and temporal information. In a time-space traffic data matrix, each entry $f_{i,t}$ represents a specific measurement of traffic state from sensor i at time interval t . Correspondingly, a time-space matrix of traffic flows for N traffic sensors over T time intervals can be denoted by:

$$\mathbf{F} = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,T} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ f_{N,1} & f_{N,2} & \cdots & f_{N,T} \end{bmatrix}. \quad (3)$$

A time-space matrix data model has a similar structure to an image, which is represented by pixels arranged in rows and columns. As a result, the time-space matrix can be used as by the CNN-based traffic prediction models.

Table 4 A Summary of predicting targets, deep-learning models, wireless sensors, and data models

Deep learning-driven traffic prediction summary			
Predicting target	DL model	Wireless sensors	Data models
Traffic volume	DNN	Road sensor	Scalar value
Traffic speed	CNN	Traffic cameras	Univariate vector
Traffic demand	RNN	On-board GPS	Multivariate vector
Travel time	RNN + CNN	Smartphones	Time-space matrix
Traffic anomaly	GNN	Social media	Region matrix
Traffic pattern	DRL	AFC systems	Graph model

In the following, we review various traffic prediction applications, including traffic volume prediction, traffic speed prediction, etc. To provide a clear overview of these applications, as shown in Table 4, we classify them in terms of predicting targets, deep-learning models, wireless traffic sensors, and traffic data models.

4.2 Deep learning for traffic volume prediction

(1) **Initial efforts.** Traffic volume prediction is a problem of time series prediction in essence. Conventionally, a series of parametric models (i.e., statistical models) have been adopted to solve primitive problems in traffic prediction. For example, Williams and Hoel (2003) proposed a seasonal ARIMA to process and predict traffic volume. Moreover, by taking the spatial characteristics of a road network into consideration, Min and Wynter (2011) presented a spatial-temporal autoregressive model to achieve accurate and scalable traffic prediction at a fine granularity. Besides, Chandra and Al-Deek (2009) used a vector auto-regressive (VAR) model to address the effect of upstream and downstream on the traffic volume of a specific location. Guo and Williams (2010) investigated the Kalman filter with a time-varying process for variance adaptation in short-term traffic volume forecasting.

However, the above time series models cannot deal with non-linearity in traffic data, so that the forecast errors can be substantial with irregular variations in traffic. Therefore, non-parametric models have been further proposed, including K-nearest neighbors (KNN) and Bayesian networks (Zhang et al. 2013; Zhan et al. 2016; Zhang et al. 2016). For example, Zhang et al. (2013) presented a KNN-based short-term traffic flow prediction system. Zhan et al. (2016) predicted city-wide traffic volume by using Bayesian networks to learn the high-level features from vehicle GPS trajectories. Zhang et al. (2016) further proposed DeepST, a deep neural network model for modeling spatio-temporal closeness in traffic data to enhance prediction accuracy.

More recently, Meng et al. (2017), Zhang et al. (2018b) applied spatio-temporal semi-supervised learning with an affinity graph structure to predict city-wide traffic volume, based on loop detector data and taxi trajectories. Although the above probabilistic machine learning models can handle the non-linear and irregular variances in traffic prediction, they actually perform shallow learning in feature extraction. Consequently, they are subject to dealing with traffic data by simple transformations, i.e., using one or two successive representation spaces. Since the data volume and data dimensions of urban traffic networks have been growing explosively, the complex traffic prediction tasks that require refined feature representations cannot be attained by the above techniques.

(2) **Advanced models.** The emergence of deep learning in traffic prediction starts with multi-layer perceptron (MLP), i.e., a three-layer forward neural network. As the units in each layer of each MLP are densely connected, a substantial number of parameters need to be learned via the backpropagation process. For instance, Kumar et al. (2015) proposed an MLP model to incorporate traffic volume, speed, road density, and temporal information to predict short-term traffic volume on highways. As MLP is a straightforward model with a fully-connected structure, it would entail high complexity with low efficiency in the representation learning process. Therefore, a subsequent branch of deep learning models is proposed to reduce computation cost in traffic prediction, such as deep belief network and stacked auto-encoder. The DBN is a stack of restricted Boltzmann machines that are trained in a greedy and layer-wise manner. The key idea of using a deep belief network is to effectively capture the features of traffic data without prior knowledge by unsupervised feature learning. For example, Huang et al. (2014) proposed a deep architecture that incorporates a deep belief network and a multi-task regression layer, where the DBN at the network's bottom performs unsupervised feature learning and a top regression layer is used for supervised traffic prediction. Koesdwiady et al. (2016) incorporated deep belief networks and data fusion techniques to derive more accurate traffic flow prediction with historical traffic data and weather data. Moreover, Soua et al. (2016) proposed a deep belief network-based approach to predict traffic flow using multi-stream data (i.e., historical traffic data, weather data, and event-based data).

Similar to DBN, the stacked auto-encoder is another type of pre-trained deep neural network to learn compact representation for dimension reduction. Specifically, Lv et al. (2014) proposed an SAE model to learn generic features from historical traffic flow data. This model can discover the non-linear spatial and temporal correlations with greedy layer-wise training. To further improve the performance of SAE models on traffic prediction, Yang et al. (2016) proposed a novel stacked auto-encoder Levenberg–Marquardt

(SAE-LM) model. By introducing the LM algorithm to train the neural networks and the Taguchi method to optimize its structure, the SAE-LM model showed higher accuracy and more efficiency in traffic flow forecasting.

Recurrent neural networks and long short-term memory networks become prevalent with their outstanding performance in capturing temporal features for accurate traffic volume prediction. For instance, Fu et al. (2016b) initiatively used basic LSTM and GRU to predict traffic flow. Zhao et al. (2017) proposed a two-dimension LSTM network to capture correlations in the temporal domain and spatial domain from the original destination correlation matrix. To improve prediction accuracy with multi-source data, Kang et al. (2017) further studied the effects of various input settings (traffic flow, occupancy, and speed) on the performance of LSTM for traffic flow prediction. Meanwhile, Jia et al. (2017) introduced two models, namely R-DBN and R-LSTM, to creatively take rainfall as an impact factor in traffic prediction. Besides, Tian et al. (2018) proposed the LSTM-M model to infer traffic flow by explicitly combining the missing traffic patterns with masking vectors. To build capabilities of LSTMs and satisfy different requirements in predicting traffic volume, many research works have proposed different variants of LSTMs, which are further combined with other deep-learning models. Hua et al. (2018) proposed RC-LSTM that contains fewer parameters due to sparse neural connectivity in comparison to conventional LSTM. Liao et al. (2018b) integrated multi-source information, including crowd map queries, road intersections, and geographical/social attributes, as the input of an LSTM-based sequence to sequence learning framework.

(3) **Cutting-edge techniques.** More recently, spatiotemporal traffic forecasting has attracted massive interest from research communities, as it integrates the convolutional neural networks to enable spatial feature extraction from traffic data. For example, Yao et al. (2019) revisited spatial-temporal dynamics in traffic data and proposed STDN, which combined a local CNN model to capture the dynamic similarity of traffic flows and an LSTM model to learn the sequential information. Zhang et al. (2017) designed a deep spatio-temporal residual network (ST-ResNet) to collectively predict the inflow and outflow of traffic in every region of a city. The ST-ResNet incorporated convolutional neural networks with residual unit sequences and dynamically aggregated their outputs for crowd/traffic flow prediction. Moreover, Li et al. (2017) modeled the traffic flow as a diffusion process on a directed graph, and they proposed a diffusion convolutional recurrent neural network (DCRNN) to solve the spatiotemporal forecasting problem. DCRNN can capture the spatial and temporal dependence in traffic data by using bi-directional random walks on the graph, and model the temporal dependency using an encoder-decoder architecture with scheduled sampling. Likewise, Deng et al.

(2019) further designed a random subspace learning strategy for a deep CNN architecture. It can learn hierarchical feature representations from incomplete traffic data for prediction. Furthermore, Zheng et al. (2019) proposed DeepSTD, a two-phase end-to-end deep learning framework to leverage spatio-temporal disturbances to predict citywide traffic flow.

Inspired by the human's ability to capture the focus in a particular vision field, attention mechanisms have been integrated into sequence-to-sequence learning, including traffic prediction (Xu et al. 2015). For example, Yang et al. (2019a) proposed an improved LSTM+ solution by integrating attention mechanisms to capture high-impact historical data for feature enhancement. Guo et al. (2019) proposed an attention-based spatiotemporal graph convolutional network, where the graph convolutions can capture spatial features, and the convolutions in the temporal dimension can capture dependencies on historical data of different time intervals. In a state-of-the-art work in spatial-temporal data mining for traffic prediction, Pan et al. (2019) presented ST-MetaNet, a deep-meta-learning based model, consisting of meta-knowledge learner, meta graph attention network and meta recurrent neural network. The ST-MetaNet can learn of traffic-related embeddings from geo-graph attributes and further model both spatial and temporal correlations for high-accuracy and network-wide traffic prediction.

4.3 Deep learning for traffic speed prediction

(1) **Basic efforts.** Besides traffic volume, traffic speed is another essential indicator of traffic status that can serve many ITS applications. Intuitively, the value of vehicular speed on the road can reflect the crowdedness level (CL) of road traffic (Qin et al. 2018). For example, Google Maps Google (2019) visualize CL of road traffic with crowd sensed traffic speed data from individual mobile devices and in-vehicle sensors. In literature studies, the revolution pattern of traffic speed prediction is similar to that of traffic volume prediction. Conventional methods for traffic speed prediction include ARIMA (Lefèvre et al. 2014; Wang et al. 2014), VAR (Chandra and Al-Deek 2009), Kalman Filter (Guo and Williams 2010), SVM (Wang and Shi 2013), KNN (Rasyidi et al. 2014), and Support Vector Regression (SVR) (Asif et al. 2013). Likewise, the initial takes of applying deep-learning models for traffic speed prediction started from deep neural networks. For instance, Dia (2001) first introduced a time-lag recurrent network (TLRN) model for predicting short-term traffic speed. Vlahogianni et al. (2005) further provided a multi-layer perceptron network with a structural optimization strategy to learn representations from multivariate traffic speed data. Moreover, the stacked auto-encoder (Lemieux and Ma 2015) and deep belief network (Jia et al. 2016) have been further adopted for traffic speed prediction, respectively.

(2) **Deep-learning models.** Research studies using LSTM for traffic speed prediction have become more influential in recent years. For instance, Ma et al. (2015) proposed a long short-term memory network for traffic speed prediction by capturing long-term temporal dependency. Yu et al. (2017a) proposed a Deep LSTM architecture to unify LSTM with SAE for forecasting traffic speed in peak-hour and post-accident conditions. Liao et al. (2018a) presented a deep spatiotemporal residual network to integrate sequence learning from different modalities for hotspot traffic speed prediction. Wang et al. (2019a) used bidirectional LSTM (Bi-LSTM) to model each path in the road network, and multiple all Bi-LSTMs were further stacked to incorporate temporal information for traffic speed prediction.

CNN is another research focus for traffic speed prediction, as it is capable of extracting features from local input patches and allowing for representation modularity. As a pioneering work in the ITS community, Ma et al. (2017) advocated ‘Learning Traffic as Images’ and proposed a deep convolutional neural network for speed prediction in large-scale transportation networks. By converting network-wide traffic to the image-like data format, they constructed a time-space matrix with temporal and spatial traffic data and further employed CNNs to process the traffic images for feature extraction and network-wide traffic speed prediction. Similarly, Jo et al. (2018) proposed image-to-image learning to predict traffic speed with a novel CNN model that consists of convolutional and deconvolutional filters.

To further exploit the potential of CNN in long-term and large-scale traffic prediction, recurrent convolutional networks have been proposed to incorporate CNN and LSTM for more accurate traffic prediction. Wang et al. (2016) proposed eRCNN, an error-feedback recurrent convolutional neural network structure for continuous traffic speed prediction, by utilizing the implicit correlations among nearby road segments to improve prediction accuracy. Yu et al. (2017b) introduced spatiotemporal recurrent convolutional networks (SRCNs) that inherited the advantages of both CNN and LSTM. In SRCNs, the spatial dependencies of road network-wide traffic can be captured by its deep convolutional neural networks, while the temporal dynamics can be learned by the LSTM component. Lv et al. (2018) proposed a look-up convolutional recurrent neural network (LC-RNN) as a rational integration of RNN and CNN. LC-RNN contained several look-up convolution layers that can perform topology-aware convolution operations to capture spatial traffic dynamics of surrounding areas effectively. Additionally, different variants of recurrent convolutional neural networks, such as PCNN (CNN for periodic traffic data) (Chen et al. 2018) and STCNN (spatio-temporal CNN) (He et al. 2019) have been further proposed for traffic speed prediction on different datasets.

(3) **Graph neural network models** To capture structural features of graphic traffic networks, the state-of-the-art research studies (Wu et al. 2020; Chen et al. 2019) focused on graph convolutional networks (GCN) to learn the interactions between road links in the traffic networks. Chen et al. (2019b) first utilized multiple residual recurrent graph neural networks (Res-RGNNs) to jointly capture spatial dependencies and temporal dynamics of traffic networks. Ge et al. (2019) proposed a temporal graph convolutional network (GTCN), which was composed of spatial-temporal components and external components to achieve traffic speed prediction. Zhang et al. (2019a) further proposed a novel attention graph convolutional sequence-to-sequence model, namely AGC-Seq2Seq, addressing the multistep prediction challenge. Moreover, Zhao et al. (2019) presented T-GCN, a temporal graph convolutional network model that combined GCN and gated recurrent units to learn complex topological structures and predict traffic speed. Diao et al. (2019) constructed a dynamic Laplacian matrix to represent spatial dependencies between road segments. They further proposed a dynamic spatio-temporal graph convolutional neural network for traffic forecasting.

4.4 Deep learning for traffic prediction with miscellaneous tasks

Besides traffic volume and traffic speed, there have been numerous deep learning-driven applications in traffic prediction with other miscellaneous tasks. In the following, we briefly highlight four research directions.

(1) **Passenger demand prediction.** It also called traffic demand prediction, which is a critical component in taxi services and ride-hailing services. Accurate prediction of passenger demand would benefit the operations of ITSs to allocate available transportation resources to different urban areas. Ke et al. (2017) proposed a fusion convolutional LSTM network (FCL-Net) to address spatial, temporal, and exogenous dependencies for short-term passenger demand forecasting for the on-demand ride services platform. Moreover, Zhang et al. (2017) proposed a spatio-temporal residual network (ST-ResNet) to collectively forecast the crowd flows based on traffic trajectories. Yao et al. (2018) proposed a deep multi-view spatial-temporal network (DMVST-Net) to model traffic correlations with three different views, i.e., temporal view, spatial view and semantic view for taxi demand prediction. Furthermore, He and Shin (2019) used a spatio-temporal deep capsule network (STCapsNet) to accurately predict ride demands and driver supplies, exploiting vectorized neuron capsules to account for comprehensive spatio-temporal and external factors. Recently, Geng et al. (2019) proposed ST-MGCN, a spatiotemporal multi-graph convolution network for ride-hailing demand forecasting. They first identified non-Euclidean correlations

of ride-hailing demand in different regions and then modeled these correlations with multi-graph convolution for demand forecasting. To infer the citywide traffic volume with biased GPS trajectories, Tang et al. (2019) presented the JMIDI framework to jointly model the dense and incomplete trajectories for citywide traffic volume inference, using dense trajectory data from GPS and incomplete trajectory data from the camera surveillance system.

(2) **Travel time prediction.** Estimating and predicting travel time is crucial for both passengers and drivers in planning the commuting time and selecting fast routes, respectively. To accurately estimate travel time on highways, Duan et al. (2016) adopted the LSTM neural network to predict the travel time of vehicles based on traffic data of 66 road links provided by Highways England. Moreover, Li et al. (2016a) exploited to build a Q-function reinforcement learning with DNN by using sampled traffic state/control as the input and the corresponding performance of the traffic system as the output. Similarly, Wei et al. (2018) proposed IntelliLight, a more effective deep reinforcement learning model with offline training and online testing based on synthetic data and real-world data, respectively. Wang et al. (2018) proposed DeepTEE, an end-to-end deep learning framework for travel time estimation, for predicting the travel time of the whole path directly. The core component of DeepTEE was a spatio-temporal learning architecture that consisted of a geo-based convolutional layer and an LSTM-based RNN layer.

(3) **Traffic anomaly prediction.** Traffic anomalies, such as traffic congestions and accidents, are the major causes of traffic delay. It is of great importance to detect and predict traffic anomalies in a timely manner. For example, Chen et al. (2016) studied the relationship between traffic accident data and human mobility data. They developed a deep-learning model based on SAE to learn hierarchical feature representations and further indicate the risk level of traffic accidents. He et al. (2018) made a first attempt to detect illegal parking event by mining massive trajectories from bike traffic data. Yuan et al. (2018) proposed a deep learning framework called Hetero-ConvLSTM. They employed a convolutional LSTM neural network with a model ensemble approach to address the spatial heterogeneity in traffic data and further improved the accuracy of traffic anomaly prediction. In addition, Di et al. (2019) proposed a ConvLSTM based congestion propagation model to process spatial traffic matrix for traffic congestion prediction. Likewise, Zhang et al. (2018a) leveraged social media data of over 3 million tweets for detecting traffic accidents, by feeding these data into LSTM and DBN models to effectively mining information of possible traffic accident.

(4) **Urban mobility prediction.** Understanding how large-scale transportation networks evolve is critical for urban traffic management. Therefore, some research studies have linked traffic prediction with urban mobility modeling

and prediction (Zheng 2019). For example, Song et al. (2016) proposed DeepTransport, an intelligent system that used deep learning architectures to jointly model human mobility and transportation patterns with 1.6 million users' GPS trajectories. Jiang et al. (2018a) proposed and implemented DeepUrbanMomentum, an online deep-learning system for short-term urban mobility prediction based on real-world human mobility data. Jiang et al. (2018b) also proposed a deep Regions-of-Interests based architecture to model urban mobility sequence and predict city-scale mobility effectively. Fan et al. (2020) leveraged building sensing data (e.g., building occupancy) with cross-domain learning for nearby urban mobility prediction. More recently, Yang et al. presented VeMo (Yang et al. 2019b) system that utilized data from the electric toll collection (ETC) to transparently model and predict state-level urban mobility. Subsequently, Wang et al. (2019b) quantified dynamic city-level patterns of electric vehicles with comprehensive data analysis from spatial and temporal dimensions. Xiang et al. (2020) investigated edge computing and low-rank theory in large-scale urban mobility datasets from a real-world ITS.

5 Future directions

In this section, we envision the promising and potential research directions for future ITS with deep learning.

Multi-source data fusion for advanced traffic prediction. With the ever-increasing number of vehicles on the road, accurate predictions on traffic states should take consideration of multiple data sources that are related to traffic status (Fan et al. 2019). It has been proven that instead of using single-source traffic sensing data, jointly considering multiple data sources can enhance the accuracy, reliability, and sustainability of traffic prediction (Liu et al. 2018). Data sources, which are not directly generated from vehicles but certainly affect traffic, are called extrinsic data (Qin et al. 2018). There are a variety of extrinsic data, including the topology of road networks, weather conditions, social events, point of interest, and public holidays. However, it is extremely difficult to fuse these extrinsic data into features for traffic prediction, as they have different non-linear correlations with traffic data (Fan et al. 2020). Moreover, it is challenging to build a concrete traffic prediction model by taking traffic data and multi-source data as input. The multi-level non-linearity would make traffic modeling and prediction exceedingly computation-intensive, and this challenge remains to be tackled in the future research study.

Real-time, large-scale, and fine-grained traffic predictions with big traffic data. With the rapid development of ITSs, traffic sensing infrastructures are generating sensing data at the Trillion-byte level to the Peta-byte level. Such an unprecedented volume of data has posed considerable

difficulties for real-time and fine-grained traffic prediction. For example, a dataset of shared electric vehicle networks contains nearly 5 TB vehicular GPS trajectory data (Wang et al. 2019b). Taming such big traffic data for traffic modeling and prediction requires more advanced techniques on both deep-learning models and parallel computing hardware. First, the deep-learning models based on GNN can further extract high-level feature abstractions from a network-wide traffic dataset. Second, parallel computing infrastructures (e.g., computing clusters) with GPUs and TPUs are envisioned to boost processing traffic data for different prediction purposes. Nevertheless, it is still an open issue about how to enable and manage advanced parallel computing with the ever-increasing big traffic data.

Data privacy, data storage, and open-source data.

With the explosive amount of traffic data being generated by traffic infrastructures and on-board GPS sensors, there are rising issues concerning data privacy, data storage, and data openness in traffic-related research. First, the aggressive increase in connected autonomous vehicles makes data sharing become universal (Liu et al. 2020). Meanwhile, the ITS must guarantee the privacy of individuals who contribute to their personal traffic information. Regarding this, privacy-preserving data publishing techniques (Fan et al. 2016), privacy-aware data structures (Wu et al. 2018), and encryptions (encrypted and compressive) privacy Wu et al. (2018) have been proposed in recent years. Second, regarding the unprecedented increase in traffic sensing data, efficient and low-cost data storage becomes a vital issue and has attracted research interest (Li et al. 2016b). For example, Chen et al. (2019a) developed a novel framework called TrajCompressor to perform cost-effective online trajectory compression, by exploiting vehicle heading direction from GPS data. Third, the evaluability and verifiability of ITS-related studies are subject to the availability of corresponding traffic datasets. However, it is still challenging to develop consolidated standards for public traffic data. Consequently, most of the traffic prediction methods are based on different datasets for evaluation, making it difficult to comprehensively compare their performance (Gharaibeh et al. 2017).

6 Conclusion

In this paper, we have presented an in-depth literature review on the recent advances in deep learning for traffic sensing and prediction. First, we have provided a brief introduction to the ITS and summarized the previous survey articles related to traffic sensing and prediction. Then, we have introduced the most popular deep-learning models that can be applied for ITS. Moreover, we have presented state-of-the-art deep learning-based applications in traffic sensing and

prediction, including traffic volume prediction, traffic speed prediction, passenger demand prediction, travel time prediction, traffic anomaly prediction, and urban mobility prediction. Furthermore, we have envisioned the future directions of deep learning for ITS and discussed the emerging challenges. We hope that this survey can benefit the research community with a comprehensive knowledge of the latest developments of deep learning for intelligent traffic sensing and prediction in ITS.

Acknowledgements This research is supported by the NSF of China Projects: Grants no. 61872447, the Natural Science Foundation of Chongqing: Grant no. CSTC2018JCYJA1879, National Postdoctoral Program for Innovative Talents of China No. BX20190202, in part by China NSF grant No. 61702525, and China Scholarship Council: Grant no. 201603170125.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

A Wireless sensor networks for traffic sensing and prediction

A.1 Wireless sensing technologies for urban traffic systems

Sensors are the fundamental elements in traffic sensing, and wireless sensor networks are widely used to satisfy the requirements of real-time and accurate traffic sensing (Xiao et al. 2019). A wireless sensor node usually consists of five critical functional modules as follows (Xu et al. 2014).

- A sensing module for vehicle detection and data acquisition.
- A wireless transceiver module for wireless data transmission.
- A local data processing module for converting physicochemical signals into traffic values.
- A memory module for storing sensing data and backup of system settings.
- A power supply module that consistently provides energy for the sensor.

We categorize wireless traffic sensors in Table 5 and introduce different traffic sensing technologies as follows. **inductive loop sensors**, as the most commonly used devices, are installed in the road surface to detect the presence of vehicles by the inducing currents from the vehicle. Similarly, magnetic sensors (including Magnetic sensors and

Table 5 Wireless sensing technologies for urban traffic systems

Technologies	Presence	Count	Speed	Classification	Cost
Inductive loop	✓	✓	✓	✓	Low
Magnetometer	✓	✓	✓		Moderate
Magnetic induction coil	✓	✓	✓		Low
Microwave radar	✓	✓	✓	✓	Low
Active infrared	✓	✓	✓	✓	Moderate
Passive infrared	✓	✓	✓		Low
Laser radar	✓	✓	✓	✓	Moderate
Ultrasonic	✓	✓			Low
Acoustic array	✓	✓	✓		High
Video image camera	✓	✓	✓	✓	High
On-board GPS	✓	✓	✓		Low
Mobile signaling	✓	✓	✓		Low

Table 6 Wireless communication technologies for urban traffic systems

Technologies	Standard	Data Rate	Topology	Feature
Bluetooth	IEEE 802.15.1	3 Mb/s	P2P	Low-power
ZigBee	IEEE 802.15.4	250 kb/s	Mesh	Scalable
Z-Wave	Z-Wave Alliance	250 kb/s	Cluster	Indoor
LoRaWAN	LoRa Alliance	50 kb/s	Star	Low-cost
WiFi	IEEE 802.11	300 Mb/s	Star	Ubiquitous
WiMAX	IEEE 802.16	75 Mb/s	P-M	High-speed
LTE	3GPP	75-300 Mb/s	Star	High-coverage
LTE-A	3GPP	500-1000 Mb/s	P2P	High-speed

Magnetic induction coil) can detect the presence of a vehicle through the anomaly in the magnetic field (Gong et al. 2018). Moreover, microwave radar sensors leverage antenna beams to detect the presence, passage, volume, lane occupancy, speed, or length of a vehicle by the reflected signals. Likewise, infrared sensors (either active or passive) detect the energy reflected by or emitted from vehicles, then convert the energy into electrical signals to further determine the presence of vehicles. Besides, laser radar sensors transmit power in the near-infrared spectrum and provide traffic measurements, such as vehicle presence, traffic volume, and traffic speed. Modern laser sensors can provide precise two-dimensional or three-dimensional image data of vehicles.

As another short-range sensing technique, RFIDs (Xiao et al. 2018a, b) have been utilized for fine-grained object detection. However, they are not feasible for the scenarios of large-scale traffic sensing, due to the constraints of communication scalability and the cost of RFID tags. Ultrasonic sensors work with pulse waveforms and can detect vehicle count, presence, and occupancy information. Furthermore, acoustic arrays are passive sensors that use signal processing algorithms to measure traffic volume and traffic speed in vehicular networks. For real-time traffic surveillance, video image sensors are the most pervasive devices of roadways that transmit television

imagery to traffic operators. With the installed data processing modules, surveillance cameras can perform more advanced traffic sensing tasks, including plate recognition, driving behavior detection, and even driver facial recognition. Moreover, onboard GPS sensors can be categorized as indirect sensors that can provide city-wide trajectory data of vehicles. GPS trajectory can be utilized by speed inference models (Zhan et al. 2016) and traffic volume estimation models (Meng et al. 2017). Meanwhile, the tradeoff between incentive pricing and sensing quality on sensing data like GPS remains as a challenge, and various mechanisms have been proposed to address this issue (Qu et al. 2018; Xiang et al. 2016)

A.2 Wireless communication technologies for urban traffic systems

There are a number of wireless communication technologies that can support traffic data transmission under various requirements (e.g., transmitting distance, data volume). As shown in Table 6, we summarize the critical enabling transmission technologies for traffic sensing and prediction, including Bluetooth, ZigBee, Z-Wave, LoRaWAN, WiFi, WiMAX, LTE, and LTE-A.

To begin with, Bluetooth and ZigBee are more suitable for short-range communication between traffic sensors and road-side units, where Bluetooth is characterized for Peer-to-Peer (P2P) communications and ZigBee has higher scalability with lower transmission rate. In addition, Z-Wave has been applied for short-range communication of indoor traffic applications (Xiang et al. 2015), such as smart parking. Moreover, LoRaWAN can support wireless communication between gateways for long-range traffic monitoring scenarios (e.g., highways) and further secure bidirectional communication with moderate data load. Alternatively, WiFi with different configurations under IEEE 802.11 standards can be used for short-range, regional, and opportunistic traffic data transmission at intersections and business-intensive areas (Zhu et al. 2017; Fu et al. 2016a; Xiang et al. 2014). Moreover, WiMAX allows scalable data rates for long-range communication. Thereby, it is more desirable for video surveillance and image cameras in traffic sensing systems. At last, LTE and LTE-A are both under the 3GPP standard. Thus, they can provide portable mobile broadband connectivity across urban areas for traffic data transmission.

References

- Asif, M.T., Dauwels, J., Goh, C.Y., Oran, A., Fathi, E., Xu, M., Dhanya, M.M., Mitrovic, N., Jaillet, P.: Spatiotemporal patterns in large-scale traffic speed prediction. *IEEE Trans. Intell. Transp. Syst.* **15**(2), 794–804 (2013)
- Bau, D., Zhu, J.Y., Strobelt, H., Zhou, B., Tenenbaum, J.B., Freeman, W.T., Torralba, A.: Visualizing and understanding generative adversarial networks. *arXiv preprint arXiv:1901.09887* (2019)
- Bengio, Y., Simard, P., Frasconi, P., et al.: Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.* **5**(2), 157–166 (1994)
- Bolshinsky, E., Friedman, R.: Traffic flow forecast survey. Tech. rep, Computer Science Department, Technion (2012)
- Castillo, E., Grande, Z., Calviño, A., Szeto, W.Y., Lo, H.K.: A state-of-the-art review of the sensor location, flow observability, estimation, and prediction problems in traffic networks. *J. Sensors* **2015** (2015)
- Chamoso, P., González-Briones, A., Rodríguez, S., Corchado, J.M.: Tendencies of technologies and platforms in smart cities: a state-of-the-art review. *Wirel. Commun. Mob. Comput.* **2018** (2018)
- Chandra, S.R., Al-Deek, H.: Predictions of freeway traffic speeds and volumes using vector autoregressive models. *J. Intell. Transp. Syst.* **13**(2), 53–72 (2009)
- Chen, Q., Song, X., Yamada, H., Shibasaki, R.: Learning deep representation from big and heterogeneous data for traffic accident inference. In: Thirtieth AAAI Conference on Artificial Intelligence (2016)
- Chen, M., Yu, X., Liu, Y.: PCNN: deep convolutional networks for short-term traffic congestion prediction. *IEEE Trans. Intell. Transp. Syst.* **19**(11), 3550–3559 (2018)
- Chen, C., Ding, Y., Xie, X., Zhang, S., Wang, Z., Feng, L.: Trajcompressor: an online map-matching-based trajectory compression framework leveraging vehicle heading direction and change. *IEEE Trans. Intell. Transp. Syst.* **21**(5), 2012–2028 (2019a)
- Chen, C., Li, K., Teo, S.G., Zou, X., Wang, K., Wang, J., Zeng, Z.: Gated residual recurrent graph neural networks for traffic prediction. *Proc. AAAI Conf. Artif. Intell.* **33**, 485–492 (2019b)
- Chen, Y., Lv, Y., Wang, F.Y.: Traffic flow imputation using parallel data and generative adversarial networks. *IEEE Trans. Intell. Transp. Syst.* (2019c)
- Chollet, F.: *Deep Learning with Python*. Manning, Shelter Island (2017)
- Course CS231n, S.U.: Convolutional neural networks for visual recognition. <http://cs231n.stanford.edu/> (2019)
- Dabiri, S., Heaslip, K.: Inferring transportation modes from GPS trajectories using a convolutional neural network. *Transp. Res. Part C Emerg. Technol.* **86**, 360–371 (2018)
- Deng, S., Jia, S., Chen, J.: Exploring spatial-temporal relations via deep convolutional neural networks for traffic flow prediction with incomplete data. *Appl. Soft Comput.* **78**, 712–721 (2019)
- Di, X., Xiao, Y., Zhu, C., Deng, Y., Zhao, Q., Rao, W.: Traffic congestion prediction by spatiotemporal propagation patterns. In: 2019 20th IEEE International Conference on Mobile Data Management (MDM), pp. 298–303. IEEE (2019)
- Dia, H.: An object-oriented neural network approach to short-term traffic forecasting. *Eur. J. Oper. Res.* **131**(2), 253–261 (2001)
- Diao, Z., Wang, X., Zhang, D., Liu, Y., Xie, K., He, S.: Dynamic spatial-temporal graph convolutional neural networks for traffic forecasting. In: Thirty-Three AAAI Conference on Artificial Intelligence (2019)
- Djahel, S., Doolan, R., Muntean, G.M., Murphy, J.: A communications-oriented perspective on traffic management systems for smart cities: challenges and innovative approaches. *IEEE Commun. Surv. Tutor.* **17**(1), 125–151 (2014)
- Do, L.N., Taherifar, N., Vu, H.L.: Survey of neural network-based models for short-term traffic state prediction. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **9**(1), e1285 (2019)
- Duan, Y., Lv, Y., Wang, F.Y.: Travel time prediction with lstm neural network. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), pp. 1053–1058. IEEE (2016)
- Fan, X., Yang, P., Li, Q., Liu, D., Xiang, C., Zhao, Y.: Safe-crowd: secure task allocation for collaborative mobile social network. *Secur. Commun. Netw.* **9**(15), 2686–2695 (2016)
- Fan, Z., Song, X., Xia, T., Jiang, R., Shibasaki, R., Sakuramachi, R.: Online deep ensemble learning for predicting citywide human mobility. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2**(3), 105 (2018)
- Fan, X., Xiang, C., Gong, L., He, X., Chen, C., Huang, X.: Urbanedge: deep learning empowered edge computing for urban IOT time series prediction. In: Proceedings of the ACM Turing Celebration Conference-China, pp. 1–6 (2019)
- Fan, X., Xiang, C., Chen, C., Yang, P., Gong, L., Song, X., Nanda, P., He, X.: Buildsensys: reusing building sensing data for traffic prediction with cross-domain learning. *IEEE Trans. Mob. Comput.* (2020)
- Fu, X., Sha, C., Lei, C., Sun, L., Wang, N.: Localization algorithm for wireless sensor networks via norm regularized matrix completion. *J. Res. Dev.* **53**, 216–227 (2016a)
- Fu, R., Zhang, Z., Li, L.: Using LSTM and GRU neural network methods for traffic flow prediction. In: 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), pp. 324–328. IEEE (2016b)
- Ge, L., Li, H., Liu, J., Zhou, A.: Temporal graph convolutional networks for traffic speed prediction considering external factors. In: 2019 20th IEEE International Conference on Mobile Data Management (MDM), pp. 234–242. IEEE (2019)
- Geng, X., Li, Y., Wang, L., Zhang, L., Yang, Q., Ye, J., Liu, Y.: Spatiotemporal multi-graph convolution network for ride-hailing

- demand forecasting. In: 2019 AAAI Conference on Artificial Intelligence (2019)
- Gharaibeh, A., Salahuddin, M.A., Hussini, S.J., Khreishah, A., Khalil, I., Guizani, M., Al-Fuqaha, A.: Smart cities: a survey on data management, security, and enabling technologies. *IEEE Commun. Surv. Tutor.* **19**(4), 2456–2501 (2017)
- Glorot, X., Bordes, A., Bengio, Y.: Deep sparse rectifier neural networks. In: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, pp. 315–323 (2011)
- Gong, L., Zhao, Y., Xiang, C., Li, Z., Qian, C., Yang, P.: Robust lightweight magnetic-based door event detection with smartphones. *IEEE Trans. Mob. Comput.* **18**(11), 2631–2646 (2018)
- Goodfellow, I., Bengio, Y., Courville, A.: *Deep Learning*. MIT Press, Cambridge (2016)
- Google. Google maps. <https://www.google.com/maps/> (2019)
- Guo, J., Williams, B.M.: Real-time short-term traffic speed level forecasting and uncertainty quantification using layered Kalman filters. *Transp. Res. Rec.* **2175**(1), 28–37 (2010)
- Guo, S., Lin, Y., Feng, N., Song, C., Wan, H.: Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. *Proc. AAAI Conf. Artif. Intell.* **33**, 922–929 (2019)
- He, S., Shin, K.G.: Spatio-temporal adaptive pricing for balancing mobility-on-demand networks. *ACM Trans. Intell. Syst. Technol. (TIST)* **10**(4), 39 (2019)
- He, T., Bao, J., Li, R., Ruan, S., Li, Y., Tian, C., Zheng, Y.: Detecting vehicle illegal parking events using sharing bikes' trajectories. In: *KDD*, pp. 340–349 (2018)
- He, Z., Chow, C.Y., Zhang, J.D.: STCNN: A spatio-temporal convolutional neural network for long-term traffic prediction. In: 2019 20th IEEE International Conference on Mobile Data Management (MDM), pp. 226–233. *IEEE* (2019)
- Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
- Hua, Y., Zhao, Z., Liu, Z., Chen, X., Li, R., Zhang, H.: Traffic prediction based on random connectivity in deep learning with long short-term memory. In: 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), pp. 1–6. *IEEE* (2018)
- Huang, W., Song, G., Hong, H., Xie, K.: Deep architecture for traffic flow prediction: deep belief networks with multitask learning. *IEEE Trans. Intell. Transp. Syst.* **15**(5), 2191–2201 (2014)
- Jia, Y., Wu, J., Du, Y.: Traffic speed prediction using deep learning method. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), pp. 1217–1222. *IEEE* (2016)
- Jia, Y., Wu, J., Xu, M.: Traffic flow prediction with rainfall impact using a deep learning method. *J. Adv. Transp.* **2017** (2017)
- Jiang, R., Song, X., Fan, Z., Xia, T., Chen, Q., Miyazawa, S., Shibasaki, R.: Deepurbanmomentum: an online deep-learning system for short-term urban mobility prediction. In: *AAAI*, pp. 784–791 (2018a)
- Jiang, R., Song, X., Fan, Z., Xia, T., Chen, Q., Chen, Q., Shibasaki, R.: Deep ROI-based modeling for urban human mobility prediction. *Proc. ACM Inter. Mob. Wearable Ubiquitous Technol.* **2**(1), 14 (2018b)
- Jo, D., Yu, B., Jeon, H., Sohn, K.: Image-to-image learning to predict traffic speeds by considering area-wide spatio-temporal dependencies. *IEEE Trans. Veh. Technol.* **68**(2), 1188–1197 (2018)
- Kang, D., Lv, Y., Chen, Y.: Short-term traffic flow prediction with lstm recurrent neural network. In: 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), pp. 1–6. *IEEE* (2017)
- Ke, J., Zheng, H., Yang, H., Chen, X.M.: Short-term forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach. *Transp. Res. Part C Emerg. Technol.* **85**, 591–608 (2017)
- Koesdwiady, A., Soua, R., Karray, F.: Improving traffic flow prediction with weather information in connected cars: a deep learning approach. *IEEE Trans. Veh. Technol.* **65**(12), 9508–9517 (2016)
- Kumar, K., Parida, M., Katiyar, V.K.: Short term traffic flow prediction in heterogeneous condition using artificial neural network. *Transport* **30**(4), 397–405 (2015)
- LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**(7553), 436 (2015)
- LeCun, Y., Boser, B.E., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W.E., Jackel, L.D.: Handwritten digit recognition with a back-propagation network. In: *Advances in Neural Information Processing Systems*, pp. 396–404 (1990)
- Lee, U., Gerla, M.: A survey of urban vehicular sensing platforms. *Comput. Netw.* **54**(4), 527–544 (2010)
- Lefèvre, S., Sun, C., Bajcsy, R., Laugier, C.: Comparison of parametric and non-parametric approaches for vehicle speed prediction. In: 2014 American Control Conference, pp. 3494–3499. *IEEE* (2014)
- Lemieux, J., Ma, Y.: Vehicle speed prediction using deep learning. In: 2015 IEEE Vehicle Power and Propulsion Conference (VPPC), pp. 1–5. *IEEE* (2015)
- Le Roux, N., Bengio, Y.: Representational power of restricted Boltzmann machines and deep belief networks. *Neural Comput.* **20**(6), 1631–1649 (2008)
- Li, L., Wen, D., Yao, D.: A survey of traffic control with vehicular communications. *IEEE Trans. Intell. Transp. Syst.* **15**(1), 425–432 (2013)
- Li, L., Lv, Y., Wang, F.Y.: Traffic signal timing via deep reinforcement learning. *IEEE/CAA J. Automatica Sinica* **3**(3), 247–254 (2016a)
- Li, Z., Wang, W., Xu, T., Zhong, X., Li, X.Y., Liu, Y., Wilson, C., Zhao, B.Y.: Exploring cross-application cellular traffic optimization with baidu trafficguard. In: 13th USENIX Symposium on Networked Systems Design and Implementation NSDI, pp. 61–76 (2016b)
- Li, Y., Yu, R., Shahabi, C., Liu, Y.: Diffusion convolutional recurrent neural network: data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926* (2017)
- Li, Y., Han, Z., Zhang, Q., Li, Z., Tan, H.: Automating cloud deployment for deep learning inference of real-time online services. In: *Proc. of IEEE INFOCOM* (2020)
- Liang, Y., Cui, Z., Tian, Y., Chen, H., Wang, Y.: A deep generative adversarial architecture for network-wide spatial-temporal traffic-state estimation. *Transp. Res. Rec.* **2672**(45), 87–105 (2018)
- Liao, B., Zhang, J., Cai, M., Tang, S., Gao, Y., Wu, C., Yang, S., Zhu, W., Guo, Y., Wu, F.: Dest-resnet: a deep spatiotemporal residual network for hotspot traffic speed prediction. In: 2018 ACM Multimedia Conference on Multimedia Conference, pp. 1883–1891. *ACM* (2018a)
- Liao, B., Zhang, J., Wu, C., McIlwraith, D., Chen, T., Yang, S., Guo, Y., Wu, F.: Deep sequence learning with auxiliary information for traffic prediction. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 537–546. *ACM* (2018b)
- Lin, Y., Dai, X., Li, L., Wang, F.Y.: Pattern sensitive prediction of traffic flow based on generative adversarial framework. *IEEE Trans. Intell. Transp. Syst.* **20**(6), 2395–2400 (2018)
- Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., Alsaadi, F.E.: A survey of deep neural network architectures and their applications. *Neurocomputing* **234**, 11–26 (2017)
- Liu, Z., Li, Z., Wu, K., Li, M.: Urban traffic prediction from mobility data using deep learning. *IEEE Netw.* **32**(4), 40–46 (2018)
- Liu, K., Xiao, K., Dai, P., Lee, V., Guo, S., Cao, J.: Fog computing empowered data dissemination in software defined heterogeneous vanets. *IEEE Trans. Mob. Comput.* (2020)

- Lv, Y., Chen, Y., Li, L., Wang, F.Y.: Generative adversarial networks for parallel transportation systems. *IEEE Intell. Transp. Syst. Mag.* **10**(3), 4–10 (2018)
- Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F.Y.: Traffic flow prediction with big data: a deep learning approach. *IEEE Trans. Intell. Transp. Syst.* **16**(2), 865–873 (2014)
- Lv, Z., Xu, J., Zheng, K., Yin, H., Zhao, P., Zhou, X.: LC-RNN: a deep learning model for traffic speed prediction. In: *IJCAI*, pp. 3470–3476 (2018)
- Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., Wang, Y.: Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction. *Sensors* **17**(4), 818 (2017)
- Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y.: Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transp. Res. Part C Emerg. Technol.* **54**, 187–197 (2015)
- Meng, C., Yi, X., Su, L., Gao, J., Zheng, Y.: City-wide traffic volume inference with loop detector data and taxi trajectories. In: *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 1–10 (2017)
- Min, W., Wynter, L.: Real-time road traffic prediction with spatio-temporal correlations. *Transp. Res. Part C Emerg. Technol.* **19**(4), 606–616 (2011)
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Belle-mare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al.: Human-level control through deep reinforcement learning. *Nature* **518**(7540), 529 (2015)
- Moustaka, V., Vakali, A., Anthopoulos, L.G.: A systematic review for smart city data analytics. *ACM Comput. Surv. (CSUR)* **51**(5), 103 (2018)
- NVIDIA. Cuda. <https://developer.nvidia.com/cuda-zone/> (2019)
- Nagy, A.M., Simon, V.: Survey on traffic prediction in smart cities. *Pervas. Mob. Comput.* **50**, 148–163 (2018)
- Nellore, K., Hancke, G.P.: A survey on urban traffic management system using wireless sensor networks. *Sensors* **16**(2), 157 (2016)
- Olah, C.: Understanding LSTM networks. <https://colah.github.io/posts/2015-08-Understanding-LSTMs/> (2015)
- Pan, Z., Liang, Y., Wang, W., Yu, Y., Zheng, Y., Zhang, J.: Urban traffic prediction from spatio-temporal data using deep meta learning. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1720–1730. ACM (2019)
- Qin, Z., Fang, Z., Liu, Y., Tan, C., Chang, W., Zhang, D.: Eximius: a measurement framework for explicit and implicit urban traffic sensing. In: *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, pp. 1–14. ACM (2018)
- Qu, Y., Tang, S., Dong, C., Li, P., Guo, S., Dai, H., Wu, F.: Posted pricing for chance constrained robust crowdsensing. *IEEE Trans. Mob. Comput.* **19**(1), 188–199 (2018)
- Rasyidi, M.A., Kim, J., Ryu, K.R.: Short-term prediction of vehicle speed on main city roads using the k-nearest neighbor algorithm. *J. Intell. Inf. Syst.* **20**(1), 121–131 (2014)
- Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *Nature* **323**(6088), 533–536 (1986)
- Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Netw.* **61**, 85–117 (2015)
- Seo, T., Bayen, A.M., Kusakabe, T., Asakura, Y.: Traffic state estimation on highway: a comprehensive survey. *Annu. Rev. Control* **43**, 128–151 (2017)
- Silva, B.N., Khan, M., Han, K.: Towards sustainable smart cities: a review of trends, architectures, components, and open challenges in smart cities. *Sustain. Cities Soc.* **38**, 697–713 (2018)
- Silver, D., Huang, A., Maddison Chris, J., Guez, E.: Mastering the game of go with deep neural networks and tree search. *Nature* **529**(7587), 484 (2016)
- Song, X., Kanasugi, H., Shibasaki, R.: Deeptransport: prediction and simulation of human mobility and transportation mode at a city-wide level. *IJCAI* **16**, 2618–2624 (2016)
- Soua, R., Koesdwiady, A., Karray, F.: Big-data-generated traffic flow prediction using deep learning and Dempster–Shafer theory. In: *2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 3195–3202. IEEE (2016)
- Tang, X., Gong, B., Yu, Y., Yao, H., Li, Y., Xie, H., Wang, X.: Joint modeling of dense and incomplete trajectories for citywide traffic volume inference. In: *The World Wide Web Conference*, pp. 1806–1817. ACM (2019)
- Tian, Y., Zhang, K., Li, J., Lin, X., Yang, B.: LSTM-based traffic flow prediction with missing data. *Neurocomputing* **318**, 297–305 (2018)
- Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C.: Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach. *Transp. Res. Part C Emerg. Technol.* **13**(3), 211–234 (2005)
- Wang, J., Shi, Q.: Short-term traffic speed forecasting hybrid model based on chaos-wavelet analysis-support vector machine theory. *Transp. Res. Part C Emerg. Technol.* **27**, 219–232 (2013)
- Wang, H., Liu, L., Qian, Z., Wei, H., Dong, S.: Empirical mode decomposition-autoregressive integrated moving average: hybrid short-term traffic speed prediction model. *Transp. Res. Rec.* **2460**(1), 66–76 (2014)
- Wang, J., Gu, Q., Wu, J., Liu, G., Xiong, Z.: Traffic speed prediction and congestion source exploration: a deep learning method. In: *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pp. 499–508. IEEE (2016)
- Wang, D., Zhang, J., Cao, W., Li, J., Zheng, Y.: When will you arrive? estimating travel time based on deep neural networks. In: *Thirty-Second AAAI Conference on Artificial Intelligence* (2018)
- Wang, J., Chen, R., He, Z.: Traffic speed prediction for urban transportation network: a path based deep learning approach. *Transp. Res. Part C Emerg. Technol.* **100**, 372–385 (2019a)
- Wang, G., Chen, X., Zhang, F., Wang, Y., Zhang, D.: Experience: understanding long-term evolving patterns of shared electric vehicle networks. In: *The 25th Annual International Conference on Mobile Computing and Networking*, pp. 1–12. ACM (2019b)
- Wang, Y., Zhang, D., Liu, Y., Dai, B., Lee, L.H.: Enhancing transportation systems via deep learning: a survey. *Transp. Res. Part C Emerg. Technol.* **99**, 144–163 (2019c)
- Wei, H., Zheng, G., Yao, H., Li, Z.: Intellilight: a reinforcement learning approach for intelligent traffic light control. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 2496–2505. ACM (2018)
- Williams, B.M., Hoel, L.A.: Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. *J. Transp. Eng.* **129**(6), 664–672 (2003)
- Wu, X., Tang, S., Yang, P., Xiang, C., Zheng, X.: Cloud is safe when compressive: efficient image privacy protection via shuffling enabled compressive sensing. *Comput. Commun.* **117**, 36–45 (2018a)
- Wu, X., Yang, P., Tang, S., Zheng, X., Wang, X.: Privacy-aware data publishing against sparse estimation attack. *J. Netw. Comput. Appl.* **109**, 78–88 (2018b)
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Philip, S.Y.: A comprehensive survey on graph neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* (2020)
- Xiang, C., Yang, P., Tian, C., Cai, H., Liu, Y.: Calibrate without calibrating: an iterative approach in participatory sensing network. *IEEE Trans. Parallel Distrib. Syst.* **26**(2), 351–361 (2014)

- Xiang, C., Yang, P., Wu, X., He, H., Wang, B., Liu, Y.: istep: a step-aware sampling approach for diffusion profiling in mobile sensor networks. *IEEE Trans. Veh. Technol.* **65**(10), 8616–8628 (2015)
- Xiang, C., Yang, P., Tian, C., Zhang, L., Lin, H., Xiao, F., Zhang, M., Liu, Y.: Carm: crowd-sensing accurate outdoor RSS maps with error-prone smartphone measurements. *IEEE Trans. Mob. Comput.* **15**(11), 2669–2681 (2016)
- Xiang, C., Zhang, Z., Qu, Y., Lu, D., Fan, X., Yang, P., Wu, F.: Edge computing-empowered large-scale traffic data recovery leveraging low-rank theory. *IEEE Trans. Netw. Sci. Eng.* (2020)
- Xiao, F., Chen, L., Sha, C., Sun, L., Wang, R., Liu, A.X., Ahmed, F.: Noise tolerant localization for sensor networks. *IEEE/ACM Trans. Networ.* **26**(4), 1701–1714 (2018a)
- Xiao, F., Wang, Z., Ye, N., Wang, R., Li, X.Y.: One more tag enables fine-grained RFID localization and tracking. *IEEE/ACM Trans. Netw. (TON)* **26**(1), 161–174 (2018b)
- Xiao, F., Chen, L., Zhu, H., Hong, R., Wang, R.: Anomaly-tolerant network traffic estimation via noise-immune temporal matrix completion model. *IEEE J. Sel. Areas Commun.* **37**(6), 1192–1204 (2019)
- Xu, G., Shen, W., Wang, X.: Applications of wireless sensor networks in marine environment monitoring: a survey. *Sensors* **14**(9), 16932–16954 (2014)
- Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zemel, R., Bengio, Y.: Show, attend and tell: neural image caption generation with visual attention. In: *International Conference on Machine Learning*, pp. 2048–2057 (2015)
- Yang, H.F., Dillon, T.S., Chen, Y.P.P.: Optimized structure of the traffic flow forecasting model with a deep learning approach. *IEEE Trans. Neural Netw. Learn. Syst.* **28**(10), 2371–2381 (2016)
- Yang, B., Sun, S., Li, J., Lin, X., Tian, Y.: Traffic flow prediction using LSTM with feature enhancement. *Neurocomputing* **332**, 320–327 (2019a)
- Yang, Y., Xie, X., Fang, Z., Zhang, F., Wang, Y., Zhang, D.: Vemo: enabling transparent vehicular mobility modeling at individual levels with full penetration. In: *The 25th Annual International Conference on Mobile Computing and Networking*, pp. 1–16 (2019b)
- Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J., Li, Z.: Deep multi-view spatial-temporal network for taxi demand prediction. In: *Thirty-Second AAAI Conference on Artificial Intelligence* (2018)
- Yao, H., Tang, X., Wei, H., Zheng, G., Li, Z.: Revisiting spatial-temporal similarity: a deep learning framework for traffic prediction. In: *AAAI Conference on Artificial Intelligence* (2019)
- Yu, R., Li, Y., Shahabi, C., Demiryurek, U., Liu, Y.: Deep learning: a generic approach for extreme condition traffic forecasting. In: *Proceedings of the 2017 SIAM International Conference on Data Mining*, pp. 777–785. SIAM (2017a)
- Yu, H., Wu, Z., Wang, S., Wang, Y., Ma, X.: Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks. *Sensors* **17**(7), 1501 (2017b)
- Yuan, Z., Zhou, X., Yang, T.: Hetero-convlstm: a deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 984–992. ACM (2018)
- Zhan, X., Zheng, Y., Yi, X., Ukkusuri, S.V.: Citywide traffic volume estimation using trajectory data. *IEEE Trans. Knowl. Data Eng.* **29**(2), 272–285 (2016)
- Zhang, L., Liu, Q., Yang, W., Wei, N., Dong, D.: An improved k-nearest neighbor model for short-term traffic flow prediction. *Proc. Soc. Behav. Sci.* **96**, 653–662 (2013)
- Zhang, J., Zheng, Y., Qi, D., Li, R., Yi, X.: DNN-based prediction model for spatio-temporal data. In: *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 1–4 (2016)
- Zhang, J., Zheng, Y., Qi, D.: Deep spatio-temporal residual networks for citywide crowd flows prediction. In: *Thirty-First AAAI Conference on Artificial Intelligence* (2017)
- Zhang, Z., He, Q., Gao, J., Ni, M.: A deep learning approach for detecting traffic accidents from social media data. *Transp. Res. Part C Emerg. Technol.* **86**, 580–596 (2018a)
- Zhang, J., Zheng, Y., Qi, D., Li, R., Yi, X., Li, T.: Predicting citywide crowd flows using deep spatio-temporal residual networks. *Artif. Intell.* **259**, 147–166 (2018b)
- Zhang, Z., Li, M., Lin, X., Wang, Y., He, F.: Multistep speed prediction on traffic networks: a deep learning approach considering spatio-temporal dependencies. *Transp. Res. Part C Emerg. Technol.* **105**, 297–322 (2019a)
- Zhang, C., Patras, P., Haddadi, H.: Deep learning in mobile and wireless networking: survey. *IEEE Commun. Surv. Tutor.* (2019b)
- Zhao, Z., Chen, W., Wu, X., Chen, P.C., Liu, J.: LSTM network: a deep learning approach for short-term traffic forecast. *IET Intel. Transp. Syst.* **11**(2), 68–75 (2017)
- Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., Li, H.: T-GCN: a temporal graph convolutional network for traffic prediction. *IEEE Trans. Intell. Transp. Syst.* (2019)
- Zheng, Y.: *Urban Computing*. MIT Press, Cambridge (2019)
- Zheng, C., Fan, X., Wen, C., Chen, L., Wang, C., Li, J.: Deepstd: mining spatio-temporal disturbances of multiple context factors for citywide traffic flow prediction. *IEEE Trans. Intell. Transp. Syst.* (2019)
- Zhu, H., Xiao, F., Sun, L., Wang, R., Yang, P.: R-TTWD: robust device-free through-the-wall detection of moving human with WiFi. *IEEE J. Sel. Areas Commun.* **35**(5), 1090–1103 (2017)
- Zhu, L., Yu, F.R., Wang, Y., Ning, B., Tang, T.: Big data analytics in intelligent transportation systems: a survey. *IEEE Trans. Intell. Transp. Syst.* **20**(1), 383–398 (2018)



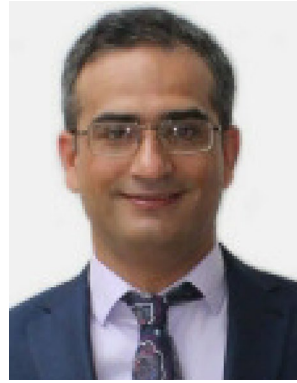
Xiaochen Fan (S'13) received his B.S. degree in computer science from the Beijing Institute of Technology, China, in 2013. He is currently a Ph.D. candidate in School of Electrical and Data Engineering at University of Technology Sydney (UTS), Australia. His research interests include mobile/pervasive computing, deep learning and IoT.



Chaocan Xiang is an associate professor at the College of Computer Science, Chongqing University, Chongqing, China. He received the BS and Ph.D. degrees in computer science and engineering from the Nanjing Institute of Communication Engineering, China, in 2009 and 2014, respectively. He studied at the University of Michigan-Ann Arbor in 2017. His current research interests include wireless sensor networks, crowd-sensing networks, and IoT.



Liangyi Gong is a Postdoc researcher at the School of Software and BNRist, Tsinghua University, Beijing, China. He received his B.E. degree in 2010 and Ph.D. degree in 2016 from Harbin Engineering University, China. His major research interests are Mobile/Pervasive Computing and Network Security.



Saeed Amirgholipour received his M.S. degree in computing sciences from Isfahan University in 2009. Between 2011 to 2017, he was a Lecturer at Azad University, Iran. Currently, he is a Ph.D. student at the University of Technology Sydney, Australia. His research interests include computer vision, deep learning video analytics, and image sentiment analysis.



Xin He is an undergraduate student in the College of Computer Science, Chongqing University, Chongqing, China. His major research interests are computer networking and crowd sensing networks.



Yue Xi received the B.S. degree from the Qingdao University of Technology, China, in 2011, the M.S. degree from Guizhou University, China, in 2014. He is currently pursuing the dual Ph.D. degrees in computer science with the University of Technology Sydney, Australia, and North western Polytechnical University, Xi'an, China. His research interests include computer vision, image processing, machine learning, and deep learning.



Yuben Qu received the B.S. degree in Mathematics and Applied Mathematics from Nanjing University, and both the M.S. degree in Communication and Information Systems and the Ph.D. degree in Computer Science and Technology from Nanjing Institute of Communications, in 2009, 2012 and 2016, respectively. He is currently a post-doc in the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China. From October 2015 to January 2016, he was a visiting

research associate in the School of Computer Science and Engineering, The University of Aizu, Japan. His research interests include mobile edge computing, air-ground integrated networks, D2D communications, and crowdsensing.



Priyadarsi Nanda earned his Ph.D. degree from the University of Technology Sydney, Australia. He is a senior lecturer in the School of Computing and Communications at University of Technology Sydney, Australia. He has more than 26 years of experience in cybersecurity, Internet of Things security, networks quality of service, assisted health care using sensor networks, and wireless sensor networks. He has published more than 70 refereed research articles.



Xiangjian He (M'99 SM'05) received the Ph.D. degree in computing sciences from the University of Technology Sydney, Australia, in 1999. Since 1999, he has been with the University of Technology Sydney. He is currently a Full Professor and the Director of the Computer Vision and Pattern Recognition Laboratory.