

Multi-Lane Traffic Pattern Learning and Forecasting Using Convolutional Neural Network

Ruimin Ke¹, Wan Li², Zhiyong Cui³, Yinhai Wang^{4,*}

ABSTRACT

Recently, the emergence of deep learning has facilitated many research fields including transportation, especially traffic pattern recognition and traffic forecasting. While many efforts have been made in the exploration of new models for higher accuracy and larger scale, few existing studies focus on learning higher-resolution traffic patterns. The most representative example is the lack of research in multi-lane pattern mining and forecasting. To this end, this paper proposes a deep learning framework that can learn multi-lane traffic patterns and forecast lane-level short-term traffic conditions with high accuracy. Multi-lane traffic dynamics are modeled as a multi-channel spatial-temporal image in which each channel corresponds to a traffic lane. The constructed multi-channel image is then learned by a convolutional neural network, which can capture key traffic patterns and forecast multi-lane traffic flow parameters. One-year loop detector data for a freeway segment in Seattle are used for model validation. The results and analyses demonstrate the promising performance of the proposed method.

¹ Ph.D. Candidate, Department of Civil & Environmental Engineering, University of Washington, Seattle, WA 98195. E-mail: ker27@uw.edu

² Ph.D. Candidate, Department of Civil & Environmental Engineering, University of Washington, Seattle, WA 98195. E-mail: wan5@uw.edu

³ Ph.D. Candidate, Department of Civil & Environmental Engineering, University of Washington, Seattle, WA 98195. E-mail: zhiyongc@uw.edu

⁴ (Corresponding author) Professor, Department of Civil & Environmental Engineering, University of Washington, Seattle, WA 98195. E-mail: yinhai@uw.edu

1. INTRODUCTION

With the increasing amount of population and urban traffic, traffic congestion has become a severe problem in many places. It is believed a better understanding of traffic patterns can benefit both traffic management agencies and road users[1]. However, this requires not only the state-of-the-art technologies for accuracy improvement but also a better formulation of traffic problems for traffic information with larger scale and higher resolution. The fast development of deep learning has been generating huge influences on traffic operation and stimulating the emergences of data-driven models for traffic analysis. Achieving high accuracies, the pioneer implementations laying the foundation for future studies were mainly on a single or just a few roadway points[2][3]. Later on, researchers started to explore the possibility of large-scale traffic learning and achieved great results by designing new deep neural networks with the consideration of spatial and temporal correlations[4]–[6]. Having extended from small scale to large scale, existing traffic learning and forecasting studies normally average traffic speeds or volumes at each milepost, thus losing the lane-level information. However, without the lane-level data, many modern and future intelligent transportation systems would not be able to work properly given the increasing needs of high-resolution traffic data[7][8]. In fact, multi-lane traffic modeling has a long history, and researchers did find out that there were differences and correlations between each lane[9]–[12]. However, most previous studies on multi-lane traffic were based on mathematical modeling with assumptions that could not be fully validated due to the lack of either enough data or appropriate data mining methods. Also, most traditional multi-lane studies focused on the modeling part, which made them good tools for the description of traffic characteristics rather than learning and forecasting.

This paper targets the multi-lane traffic pattern learning and forecasting problem by proposing a deep convolutional neural network (CNN) structure. The proposed structure learns multi-lane traffic as a multi-channel image in a way that each lane is added as a channel to the traditional 2D spatial-temporal traffic flow image. This modeling idea comes from CNN’s superiority to capture features in both single-channel grayscale images and multi-channel RGB images. In RGB images, each color channel has correlations yet differences with the other two, which is similar to the traffic flow parameter on a single traffic lane. Thus, averaging traffic flow parameters at a certain milepost and time is like doing a weighted average of three color-channels on an RGB image pixel to get its grayscale value. In this sense, previous deep learning methods for traffic pattern forecasting learn the “grayscale images” or even just its image column, yet the proposed method models traffic flow as “RGB images”. In this way, our model can capture the traffic flow features in three dimensions: the temporal dimension, the spatial dimension along the corridor, and the spatial dimension across lanes. With this way of modeling, the proposed method learns traffic as multi-channel images and can forecast the future multi-lane traffic flow conditions. In the rest of the paper, we first introduce the methodology, and describe the data used in this study, then present our numerical results, and end up with the conclusion and future work.

2. METHODOLOGY

2.1 Modeling Multi-Lane Traffic as Multi-Channel Image

The first step of our methodology is modeling the multi-lane traffic flow as a multi-channel image. In this study, we use loop detector data as the input to our model. We consider loop detector data being a natural fit for this study: Loop detectors are stationary point detectors installed on every lane with the distance interval of about half a mile, and each lane has loop detectors installed at the same mileposts. From the perspective of comparing traffic flow to an image, the loop detector is the sampling of the traffic flow just like the image is the sampling of the what we see in the real world. For a pixel in an RGB image, the three channels are the red, green, and blue values at a certain spot; and in traffic flow, loop detectors that are installed at the same milepost for each lane take the “R, G, and B” values at that spot of the corridor. Other stationary traffic detectors such as surveillance camera and microwave radar detectors could also provide the data for the proposed model if they were evenly distributed at a corridor and able to detect lane-level traffic flow parameters. However, probe-vehicle data, or in other words, GPS-based data are not appropriate for our model, because the GPS localization accuracy of most existing devices are not likely to satisfy the lane-level data collection requirement.

This modeling process is shown in Figure 1. The assumption is that loop detectors are evenly distributed on a three-lane freeway segment. There are loop detectors installed at k different locations along this segment, which means there are in total $3 \times k$ loops. Suppose the number of time steps used for learning and forecasting is n , single-lane traffic would be represented by a $k \times n$ spatial-temporal image. Thus, to represent the three-lane traffic, a $k \times n$ image with three channels are constructed. In this image, each pixel is a three-unit vector representing three lanes’ traffic condition at a given time and location. The image could be filled by speed values, volume values, or other traffic flow parameter values.

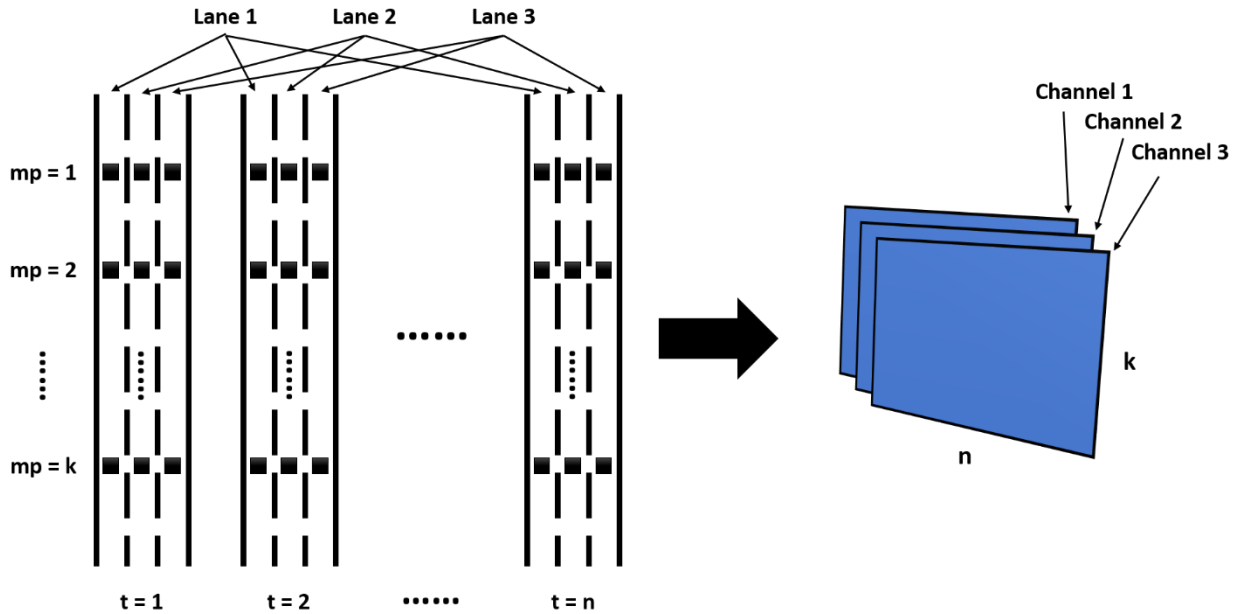


Fig.1 The modeling process of converting multi-lane traffic flow to multi-channel image

2.2 The Convolutional Neural Network Structure

To learn the multi-channel image constructed in the first step, a CNN structure is designed, and it is shown in Figure 2. In this structure, the input is our multi-channel image with the dimension of $k \times n \times c$, where k is the number of locations (mileposts) having loops installed on the corridor, and n is the time steps selected for feature learning, and c is the number of lanes or channels. In this structure, three convolutional layers are added sequentially. We select three as the number of hidden convolutional layers based on a trial-and-error process, during which we observe that three convolutional layers constantly outperform just having one or two convolutional layers, yet no improvement is observed with more than three of them. The filter size we use is all $2 \times 2 \times c$ in order to better capture the relationships between each pair of adjacent loops as well as time steps. The number of filters for each convolutional layer is chosen based on experience and the consideration to balance efficiency and accuracy. We do not insert any pooling layer in the structure because our images are relatively small compared to regular images thus we do not want to lose much information by pooling. Regular input images to a CNN normally have at least hundreds of columns and rows while the spatial-temporal image for a corridor could be in very small size. For example, 100 time steps could be at least 100 minutes even the data time resolution is 1 minute, however, a couple of hours of historical data are likely too much for short-term traffic forecasting. The third convolutional layer is flattened and connected to a fully-connected (FC) layer with 1024 nodes. This FC layer is fully connected with the output layer as well. The output has to have the same dimension with the traffic flow parameter values in a future time step, thus it has a dimension of $1 \times (k \times c)$.

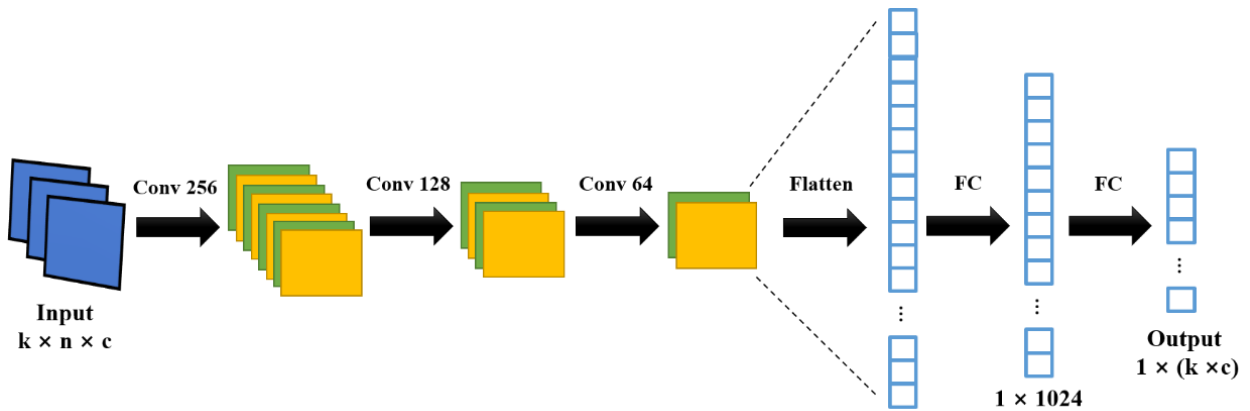


Fig. 2 The proposed convolutional neural network structure

3. DATA DESCRIPTION

For each location, loop detector has speed, volume, and occupancy data. In this paper, our model learns and forecasts speed and volume since they are more frequently used than occupancy; also, there are linear relationships among these three parameters thereby two is good to describe the macroscopic traffic flow in most cases. Our loop detector data covers a four-lane freeway segment on I-5 in Seattle area. The numerical experiment will include 40 detectors which cover about 5 miles of freeway on I-5 southbound, from milepost 170 to 165. This corridor is one of the busiest freeway segments in Washington State which connects the

University of Washington with Seattle downtown. The data is from a traffic big data analytics platform called the Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net) system[13]. The data interval is 5 minutes considering the balance between a good time resolution and the robustness to outliers. One full year of speed and volume data for the year of 2016 is used for training and testing the proposed method.

4. NUMERICAL RESULTS

The proposed method is implemented in python with the Keras deep learning library. The Keras library is built on the TensorFlow backend, and the training is operated using a GTX 1080 GPU. The loss is selected to be the mean squared error, and the optimizer is RMSprop. Given the 5-minute time interval, about 105,000 samples are generated from the one-year period. We use 80,000 samples as the training samples and the remaining 25,000 samples for testing.

Figure 3 displays the heat maps of the ground truth speeds and the forecasted speeds for all lanes in a given time range. The settings for this one is that training time step = 18 and forecasting time step = 2. From this figure, it can be seen that traffic on each lane has similar yet slightly different patterns. For example, in the morning peak, lane 1 has severest traffic congestion among the four lanes, especially at the downstream (bottom) part of this lane. We examined the map and it turned out to be the morning peak of going downtown for work, and there were some on-ramps connected to lane 1 experiencing a great amount of traffic in the morning. Also, at around noon between the morning and afternoon peaks, we can see lane 2 and 3 have slightly lower speeds than the other two. Note that the proposed method captured most key patterns, but at the same time, its predicted speed results are smoother than the ground truth. This is explainable because lane-level data contain more noises and outliers than aggregated data. Hence, some predictions would be difficult due to the randomness even if the method itself learns well.

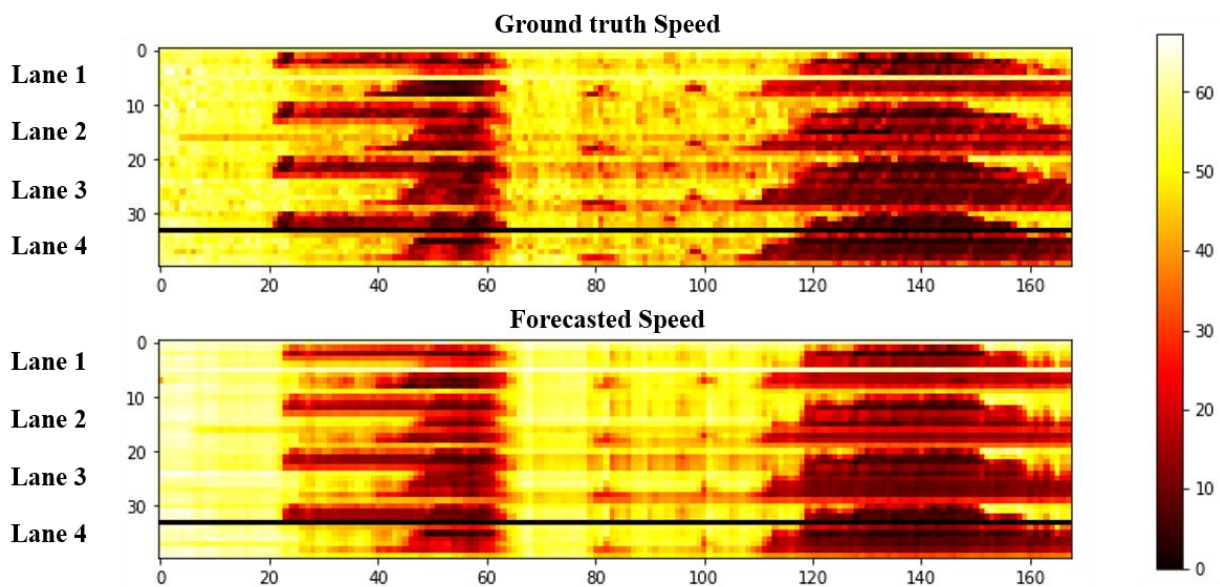


Fig.3 Heat maps showing the ground truth speeds and our forecasted speeds for all four lanes in a given time range from 6 am to 20 pm.

Figure 4 shows the volume predictions at Milepost 166.4 on I-5 southbound, where the orange curves are the predicted volumes and the blue curves are the ground truths. Using this figure, we can have an intuition of how the traffic flows vary from lane to lane, and where some of the errors are generated. It can be observed in this figure that lane 2 and lane 3 have very similar volume patterns for the day, and all lanes share similar volume patterns in the morning. However, during the afternoon and evening, lane 1 and lane 4 have quite different volume patterns, where the volume in lane 1 is the lowest among the four while that in lane 4 is the highest. We checked the map and seemed this was caused by the frequent lane changing behaviors near lane 1 and the congestion on the exit-ramp which were connected to lane 1.

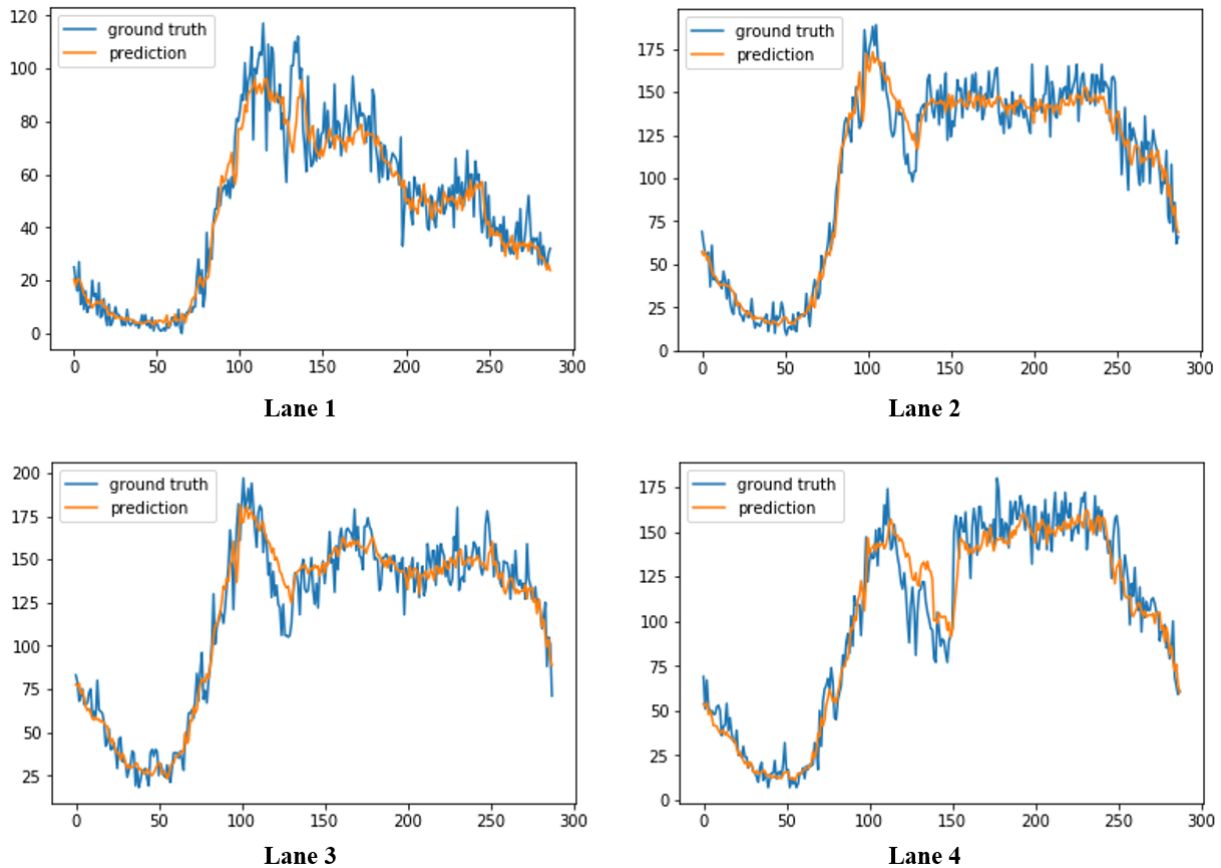


Fig. 4 The multi-lane volume predictions at Milepost 166.4 for a single day.

Table 1 summarizes the accuracies and training times for combinations of the training time step and the forecasting time step. By looking at each row, it can be seen that for a given training time step, the accuracy decreases as the forecasting time step increases, which is consistent with previous studies; and for a given training time step, the training time remains the same no matter how large the forecasting time step is, which makes sense because the number of computations is actually the same. For every single column of either speed or volume, it can be seen the forecasting accuracies stay pretty much the same as the training time step going from 6 to 18 (30 mins to 90 mins of data). This is an interesting finding that short-term traffic conditions are probably not impacted by historical traffic conditions more than thirty minutes ago. But when

the training step becomes larger, the training time also increases due to the larger input size. In general, both speed and volume predictions end up being quite efficient and accurate, and the speed forecasting has better accuracy than volume forecasting. By examining the raw data, it is found that volume data has larger variations and more vibrations in the time series.

Table 1 Sensitivity analysis with respect to training time step and forecasting time step (accuracy/training time)

			Forecasting Time Step		
			2 (10 mins)	4 (20 mins)	6 (30 mins)
Training Time Step	Speed	6 (30 mins)	88% / 700s	85% / 700s	81% / 700s
		12 (60 mins)	89% / 1200s	85% / 1200s	81% / 1200s
		18 (90 mins)	88% / 1700s	84% / 1700s	81% / 1700s
	Volume	6 (30 mins)	81% / 700s	76% / 700s	73% / 700s
		12 (60 mins)	80% / 1200s	77% / 1200s	74% / 1200s
18 (90 mins)		80% / 1700s	76% / 1700s	74% / 1700s	

5. CONCLUSION AND FUTURE WORK

In this paper, we aimed to solve the multi-lane traffic learning and forecasting problem by applying a deep convolutional neural network structure. We first developed a modeling procedure for converting multi-lane traffic into a multi-channel image. We illustrated the reasonableness of such a procedure by comparing it to the generation of digital images. Then, a convolutional neural network was designed to deal exclusively with the multi-lane traffic problem. A 5-mile corridor in I-5 interstate freeway with 40 loop detectors was selected as the test segment, and one-year speed and volume data in the whole year of 2016 were used for training and validating the proposed method. It turned out that the proposed method could capture the key patterns of traffic flows and achieved good performances regarding accuracy and efficiency.

Future work will be focusing on three aspects. First, we will test the proposed methods in more scenarios in order to validate it and enhance its robustness; second, we will explore the influences of loss functions and optimizers on the forecasting results, and develop a new loss function and a new optimizer for performance improvement; third, since traffic volume and speed often have correlations, the authors would like to conduct further study on whether it is possible to integrate their correlation into our model.

REFERENCES

- [1] R. Ke, Z. Zeng, Z. Pu, and Y. Wang, "New framework for automatic identification and quantification of freeway bottlenecks based on wavelet analysis," *J. Transp. Eng. Part A Syst.*, vol. 144, no. 9, pp. 1–10, 2018.
- [2] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transp. Res. Part C Emerg. Technol.*, 2015.
- [3] J. Tang, F. Liu, Y. Zou, W. Zhang, and Y. Wang, "An Improved Fuzzy Neural Network for Traffic Speed Prediction Considering Periodic Characteristic," *IEEE Trans. Intell. Transp. Syst.*, 2017.
- [4] Z. Cui, R. Ke, and Y. Wang, "Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction," Jan. 2018.
- [5] Z. Cui, K. Henrickson, R. Ke, and Y. Wang, "High-Order Graph Convolutional Recurrent Neural Network: A Deep Learning Framework for Network-Scale Traffic Learning and Forecasting," Feb. 2018.
- [6] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction," *Sensors (Switzerland)*, 2017.
- [7] A. Chen, A. Ramanandan, and J. A. Farrell, "High-precision lane-level road map building for vehicle navigation," in *Record - IEEE PLANS, Position Location and Navigation Symposium*, 2010.
- [8] Y. Gu, L. T. Hsu, and S. Kamijo, "Towards lane-level traffic monitoring in urban environment using precise probe vehicle data derived from three-dimensional map aided differential GNSS," *IATSS Research*, 2018.
- [9] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transp. Res. Part B*, 1994.
- [10] C. F. Daganzo, "A behavioral theory of multi-lane traffic flow. Part I: Long homogeneous freeway sections," *Transp. Res. Part B Methodol.*, 2002.
- [11] I. Bonzani and L. M. Gramani Cumin, "Modelling and simulations of multilane traffic flow by kinetic theory methods," *Comput. Math. with Appl.*, 2008.
- [12] C. Roncoli, M. Papageorgiou, and I. Papamichail, "Motorway traffic flow optimisation in presence of Vehicle Automation and Communication Systems," in *Computational Methods in Applied Sciences*, 2015.
- [13] Y. Wang *et al.*, "Digital Roadway Interactive Visualization and Evaluation Network Applications to WSDOT Operational Data Usage," 2016.