Recently, with the rapid development of deep learning models, traffic analysis using image datasets has attracted more attention. Specifically, the network traffic can be represented as images as the input for deep learning models to provide various applications (e.g., Spatio-Temporal traffic forecasting). In this study, we propose a new image generation approach for traffic density classification in terms of large-scale road network. Particularly, traffic volume and speed are at certain areas able to be measured by using surveillance systems (e.g., loop detectors). However, measuring the density is difficult which depends on the spatial correlation from the perspective of the network. Consequently, an effective image generation approach, based on information arrival and departure time of vehicles, is proposed to deal with this problem. Regarding the experiment, traffic density classification using a convolutional neural network is executed on roadside equipment data of 11 continuous intersections for evaluating the effectiveness of the proposed approach.

1. Introduction

In recent years, Deep Learning (DL) models have achieved superior performance in a variety of application domains (Alom et al., 2019). In case of Intelligent Transportation Systems (ITS), researchers are interested in applying DL models to analyze traffic flow (e.g., traffic classifications and predictions). Specifically, several well-known DL models such as Convolutional Neural Network (CNN) (Chen et al., 2018), Recurrent Neural Network (RNN), Deep Neural Network (DNN), and Deep Reinforcement Learning (DQN), have successfully applied for traffic flow analysis. In particular, it depends on the assumption of traffic datasets to apply appropriate models for the training process. For instance, CNN-based methods model the road network as an image which is able to capture the spatial dependency of traffic flow. On the other hand, RNN-based methods (i.e., LSTM network) are suitable for learning the temporal dependence. Recently, trend research focuses on combining two aforementioned models to deal with the spatial and temporal dependencies for traffic flow analysis (Wu et al., 2019).
However, conventional models are not able to explore the hidden patterns, for example, the topology of the road network and the connection between them. Therefore, it is difficult to analyse the traffic condition in terms of the large-scale network. Recently, Graph Neural Networks (GNN) has proposed as a promising concept for improving the ability to learn hidden patterns from spatial-temporal graphs, which is an important issue for various applications such as human action recognition and traffic prediction (Wu et al., 2019). Nonetheless, the dynamicity of graphs becomes an emergent issue in which we need to consider how to perform graph convolution with the dynamic spatial relation problem, especially, in the case of transportation domain, where the spatial dependencies change over time with the road condition (e.g. traffic congestion) (Diao et al., 2019).

Based on insights from previous works, we find out that representing the road network to static images as the input datasets play an important step to apply DL models for traffic analysis (Li et al., 2017). Particularly, traffic datasets, that have collected from connected devices on the road (e.g. VDS, RSE, and GPS), can be converted to images with both spatial and temporal relations (Yao et al., 2018). Consequently, the network traffic can be modeled with grid (Ma et al., 2017) or graph structures (Yu et al., 2018). Specifically, Figure 1 depicts the main steps for traffic analysis from traffic datasets by using image generation approach.

In this study, we propose a new approach for converting network traffic to images. Specifically, we focus on how to classify traffic density in terms of large-scale road network for spatial-temporal traffic flow analysis. Generally, the contributions of this paper can be summarized as follows:

- We propose a traffic density classification approach for classifying traffic condition which is an emergent issue to provide solution for improving the traffic flow in terms of large-scale road network (e.g. congestion detection, dynamic traffic light control system).
- A new image generation method is presented for covering the traffic information in terms of both spatial and temporal correlations. Specifically, difference with previous works, in this paper, the images are generated based on the average travel speed values which is able to specify the traffic condition in each location (i.e. sensor’s location) following spatial and temporal relation.
- The evaluation is executed on a real traffic dataset, which is collected and pre-processed from continuous intersections in an urban area. Therefore, we are able to evaluate the performance of the proposed method by estimating the traffic condition in terms of large-scale road networks.

Figure 1. Traffic analysis using image generation approach.
The rest of this paper is organized as follows: In Section 2, we present some literature reviews of DL models for traffic analysis. Moreover, some previous works of representing traffic road network to images and applications of image classification are presented. In Section 3, we present our proposed approach for traffic density classification problem at the large-scale road network. The experiment results, which evaluate on real traffic datasets at an urban area, are shown in Section 4. Section 5 includes some discussions and future works of this study.

2. Literature review

2.1. Deep learning models for traffic analysis

The rapid growth of traffic data becomes an emergent challenge in Intelligent Transportation System (ITS) that the traditional processing systems are not able to deal with the data analytics requirements. Recently, DL has introduced as a promising approach to deal with characteristic problems in traffic data such as highly nonlinear, time-varying, and randomness (Zhu et al., 2019). Specifically, different DL models enable different data representations to be learned for different applications. Particularly, Figure 2 depicts the applications of DL models for different fundamental tasks in ITS (Wang et al., 2019). The preliminary background of several well-known DL models are sequentially described as follows:

- **Deep Neural Network (DNN):** is an artificial neural network (ANN) with multiple layers between input and output layers. Several variants of DNN includes Multilayer Perceptron (MLP), Deep Belief Network (DBN), and Stacked Auto-Encoder (SAE), which the main difference is the design of hidden layers, have been widely applied for traffic prediction (Lv et al., 2015).

- **Recurrent Neural Network (RNN):** are designed to model the sequence data. Specifically, Long-Short Term Memory (LSTM) is a special kind of RNN, which is capable of learning long-term dependencies to address the limitation of the conventional RNN model (e.g. gradient vanishing or exploding problems). Therefore, this model has proved the capability for training time series data in ITS (Yi & Bui, 2020; Yi et al., 2019).

- **Convolution Neural Network (CNN):** starts with a convolutional layer to extract common patterns of the training instances which has achieved great success for the image classification in the competition of ImageNet (Krizhevsky et al., 2012). The model most commonly applied to analysing visual imagery. Specifically, CNN has used for

![Figure 2. The applications of deep learning models in intelligent transportation system.](image-url)
visual recognition in traffic analysis (e.g. vehicle counting). Furthermore, the hybrid model, which combines CNN and LSTM, has recently become an emergent issue for traffic analysis that incorporates both spatial and temporal dependency in the traffic flow (Yu et al., 2017).

- Deep Reinforcement Learning (DQN): A DNN is set up to learn the Q-function of reinforcement learning from inputs such as internal states and set of actions. The objective is to maximize the future rewards through a sequence of actions. This concept has been applied successfully for applying advanced technologies to improve traffic flow at urban areas (Bui & Jung, 2019).

### 2.2. Representing network traffic as images

Time series analysis based on DL models has exploited more complex architectures and can achieve better results than traditional methods. However, in the case of applying DL models for transportation domain, there are still remain challenges that need to take into account:

- Traffic time series consider temporal correlations. It is difficult to long term traffic analysis, especially in case of recurring incidents (e.g. rush hours or accidents) which can cause non-stationarity.
- The models mainly focus on training the dataset of a single road location or a small network region (e.g. intersection) due to the traffic data from sensors that do not consider the spatial correlations from the perspective of the network.

Recent studies apply Graph Neural Network (GNN) (Zhu et al., 2019) to deal with the spatial-temporal correlation problem. Specifically, several state-of-the-art GNN models such as Spatiotemporal Recurrent Convolutional Networks (SRCN) (Yu et al., 2017), Spatial-temporal Graph Neural Networks (STGNN) (Li et al., 2018), Spatio-Temporal Graph Convolutional Networks (Yu et al., 2018), and Graph Multi-Attention Network (GMAN) (Zheng et al., 2020) have been proposed as emergent models for analysing the temporal evolution and spatial dependencies in large-scale road network. However, this method mainly focus on traffic forecasting problem (Yin et al., 2020), which analyse and predict traffic flow in each certain point (i.e. sensor’s location).

For estimating the traffic condition (e.g. congestion estimation) in large-scale areas, representing the road network as image for the input data become a promising method (Ma et al., 2017). Specifically, traffic road network is convert into a time-space matrix which is constructed using time and space dimension information (e.g. vehicle position or average speed) as shown in Figure 3.

However, the drawback of this approach is that the spatial structure of traffic road network is represented with Euclidean space among road sections. Therefore, it is difficult to specify the traffic condition among road sections. In this regard, this paper propose a new method for generating image from traffic road network which based on the average traffic features (e.g. traffic volume, traffic speeds). To the best of our knowledge, this is the first paper which focus on converting network traffic to images in order to classify the traffic condition in large-scale road network.
2.3. Image classification using CNN model

CNNs is the most popular neural network model for the image classification problem. Specifically, the model follows a hierarchical architecture to build a network. Subsequently, the output is processed based on a fully-connected layer where all the neurons are connected to each other. Figure 4 depicts a CNN architecture to classify handwritten digits datasets (i.e. MNIST datasets). Consequently, CNN-based image classification has been applied in various applications. For instance, Li et al. (2014) present a medical image classification approach using CNN for classifying lung image patches with interstitial lung disease. Authors in Hershey et al. (2017) adopt various CNN architectures (e.g. AlexNet, VGG, Inception, and ResNet) to classify the soundtracks for the large-scale audio datasets. Seo and Shin (2019) propose a hierarchical convolutional neural network for fashion image datasets. Regarding the transportation domain, our previous works apply CNN architectures for the feature extraction to analyse traffic flow from CCTV and road sound datasets, respectively Bui, Yi et al. (2020) and Bui, Oh et al. (2020).

3. Proposed approach

Traffic information, that have collected from connected devices on the road network, are considered with time and space dimensions in order to classify the traffic congestion. In this regard, we first present the large-scale traffic density classification problem. Then, an image generation approach is proposed for representing the traffic conditions.

Figure 3. Traffic-to-image conversion on a road network.

Figure 4. A CNN architecture to classify handwritten digits.
3.1. Image-based traffic density classification

The main focus of traffic analysis regarding the proposed image generation approach is to classify the traffic density in the large-scale road network. Specifically, the traffic conditions are represented by three conventional parameters such as traffic volume, speed, and density. Surveillance systems (e.g. loop detectors) are able to measure the traffic volume and speed. However, measuring the density is difficult which have recently attracted more attention (Chung & Sohn, 2018; Kurniawan et al., 2018). Normally, the traffic densities are able to classify different conditions following the time of a day. For instance, Figure 5 shows the traffic volume of a symmetric road during one day which have collected from Vehicle Detection System (VDS) (Yi & Bui, 2019). Consequently, traffic conditions can be defined into three groups which are low, normal, and high densities following the time of a day.

In this regard, supporting $L = \{l_1, l_2, \ldots, l_n\}$ denotes a set of traffic condition labels, the traffic density classification problem is regarded as a supervised learning problem in which given a training set $D = \{(x_i, Y_i | 1 \leq i \leq m)\}$, where $Y_i \subset L$, the objective is to learn a multi-label classifier from $D$ to predict labels of new images.

3.2. Image generation approach

The traffic information can be represented in the spatio-temporal correlation as a graph $G(V, E)$, which $V$ is the set of Node (e.g. intersection) and $E$ denotes the set of links which are the possible directions among Nodes. The time interval depends on the sampling resolution of connected devices. Normally, data might be aggregated in several minutes (e.g. 5 minutes) for analysing the traffic condition. In this regard, the research question is that how to represent the weight (traffic density) among nodes for generating the image datasets. Figure 6 illustrates the image generation process of our proposed approach. Specifically, traffic datasets from surveillance systems (e.g. loop detector) collect traffic information such as speed and traffic volume at a certain location. Since we consider the density of traffic flow in a large road network, speeds of vehicles are taken into account. However, the average speed at a certain location can not explain in more detail the traffic condition (e.g.congestion or non-congestion). In this regard, we calculated the traffic travel time based on the arrival and departure times of vehicles in

![Figure 5. Traffic volumes of a symmetric road with two sensors.](image_url)
3.2.1. Traffic condition representation

The average travel speed is calculated based on the travel time and the distance between two nodes/intersections as follows:

\[ \bar{v}_{ij} = \frac{d_{ij}}{\hat{t}_{ij}} \]  

where \( d_{ij} \) is distance from intersection \( i \) (start point) to \( j \) (end point) and \( \hat{t}_{ij} \) is the approximated travel time between two intersections, which can be calculated as follows:

\[ \hat{t}_{ij} = \frac{\sum_{v=1}^{n} t_{v}^j - t_{v}^i}{n} \]  

where the \( t_{v}^i \) and \( t_{v}^j \) are the arrival and departure time of vehicle \( v \) at intersections \( i \) and \( j \), respectively. Since the distance is constant, the average travel speed depends on travel time values. In this research, the travel time would be used to represent the traffic conditions (e.g. congestion or non-congestion).

3.2.2. Image representation at single intersection

Since an intersection consists of more than three legs, we use polygons to present the traffic conditions at intersections. Specifically, the polygon can be a triangle, square or pentagon shapes in cases of the intersection with 3 legs, 4 legs, or 5 legs, respectively. Furthermore, since we represent the traffic condition using travel time of vehicles, the radius of the polygons is maximum travel time of vehicles, which is formulated as follows:

\[ I_{\text{max}} = \frac{t_{\text{max}}}{t_{\text{min}}} \]  

where \( I_{\text{max}} \) denote the maximum normalized travel time at a certain intersection. \( t_{\text{max}} \) and \( t_{\text{min}} \) are calculated based on the maximum and minimum speeds as follows:

\[ \frac{t_{\text{max}}}{t_{\text{min}}} = \frac{d}{v_{\text{min}} / \text{max}} \]  

Normally, in the field of transportation, we are able to measure the maximum and minimum speeds based on the posted speed limits values \( v_{\text{lim}} \) depending on the type
of road. In this regard, Equation (3) can be re-formulated as follows:

$$I_{\text{max}} = \alpha v_{\text{lim}},$$

(5)

where $\alpha$ is a certain value which defines the ratio between the maximum speed and posted limited speed. Consequently, Figure 7 illustrates the concept of converting traffic conditions at an intersection with 4 legs to the image of polygons at a certain time. Specifically, the vertexes of the polygon are computed based on the values of average travel time corresponding with different directions (e.g. North, East, West, and South), which are computed based on the approximated travel time $\tilde{t}_{ij}$ (Equation (2)).

In particular, the area of the polygon represents the traffic condition of the area around the intersection. Specifically, if the area is close to the maximum value ($I_{\text{max}}$), the traffic is the congested condition. For instance, the example in Figure 6 shows that the congestion occurs in the East direction of the intersection, other directions are normal conditions of the traffic flow.

3.2.3. Converting network traffic to images
Supporting the network traffic includes multiple intersections, we can represent the network as a graph $G(V, E)$ where $V$ is a set of intersections and $E$ denotes the set of possible links among intersections. Consequently, the traffic conditions of the road network is represented based on densities of traffic flow in each intersection, which is illustrated in Figure 8.

Figure 7. An example of converting traffic condition at an intersection with 4 legs.

Figure 8. An illustration of converting traffic conditions to images of a road network.
4. Empirical study and applications

4.1. Data description and experiment setup

In this study, data from roadside equipment (RSE) is taken into account. Specifically, RSE devices have installed on major intersections, which are able to connect to on-board equipment (OBE). Consequently, when vehicles arrives the intersections, OBE automatically connects to the RSE through communication system based on the dedicated short range communication (DSRC) technology. Therefore, we collect and pre-process RSE data of a test base area during 5 months, which includes 11 continuous intersections in order to analyse traffic conditions as shown in Figure 9. Specifically, the traffic density images are generated based on the value of maximum speed, posted limited speed, and the approximated travel time of each node/intersection. In particular, the posted limited speed $v_{lim}$ is set around 10 KPH (6.2 mph), and the desired maximum speed equals 1.5 times of the posted speed.

Furthermore, we consider the scenario of traffic flow in the morning (6AM to 10AM) at the test base for evaluating our proposed method and the time interval for each recorded/image is set 5 min. Consequently, the traffic image dataset includes 5040 samples for the training, 720 samples for the validation, and 1440 for the testing. Specifically, the set of traffic condition labels $\mathcal{L}$ are described as follows:

- **Label 1**: from 6AM to 7AM, traffic condition is low density.
- **Label 2**: from 7AM to 8AM, traffic condition is normal density.
- **Label 3**: from 8AM to 9AM, traffic condition is normal density.
- **Label 4**: from 9AM to 10AM, traffic condition is high density.

Regarding the training model, the architecture of the proposed CNN is determined based on the hyperparameter optimization using Tree-structured Parzen Estimator (TPE) (Bergstra et al., 2011). Particularly, the hyperparameters of the model are shown in Table 1.
4.2. Experiment result

Figure 10 demonstrates the accuracy of the trained model on the new image dataset. Specifically, we take a comparison following the number of labels. Particularly, we are able to achieve over 0.95 of the training accuracy to classify the low density and high density. In case of the 3 and 4 labels, the training accuracies are 0.86 and 0.74, respectively.

For more detail, Table 2 shows the results of all cases for the classification of the generated image datasets. The experimented results indicate that the generated image dataset is able to classify the traffic density of the large-scale road network, especially in the case of distinguishing congested and non-congestion conditions. Consequently, the traffic condition can be estimated at certain times to provide smart traffic control for improving the traffic flow.

Moreover, we are able to improve the classification problem by increasing the layers of CNN architecture. For instance, Figure 11 demonstrates the training results by adopting two well-know CNN architecture which are Resnet50 (He et al., 2016) and Inception-v3 (Szegedy et al., 2016) (without pre-trained model). However, we need to deal with several challenging issues by applying the aforementioned CNN models such as time-consuming and transfer learning problems. Therefore, we will take this issue as a future work regarding this study.

Table 1. Hyperparameters of the proposed CNN model.

| Layer | Name          | Values               |
|-------|---------------|----------------------|
| Input | Image Size    | 330 × 330 (pixels)  |
| Layer 1 | Convolution  | Filter (128,3,3)     |
| Layer 1 | Pooling     | Pooling(2,2)         |
| Layer 2 | Convolution  | Filter (64,2,2)      |
| Layer 2 | Pooling     | Pooling(2,2)         |
| Layer 3 | Flatten     | –                    |
| Layer 3 | Fully-connected | –                  |

Figure 10. Results of the trained model with different values of labels.
5. Conclusion and future work

Recently, text-to-image generation is an emergent approach that provides many applications using Deep Learning models, and transportation system is no exception. In this study, we propose a new image generation approach to deal with the traffic density classification in terms of large-scale network traffic. Specifically, traffic density at a certain area (e.g. intersection) is able to represent based on the average travel time of vehicles. Consequently, we are able to classify the traffic conditions based on the generated image by using CNN network. Particularly, RSE data at an urban area with traffic flow from 11 continuous intersection have collected and pre-processed to evaluate our proposed approach. Specifically, the experiment shows promising results for classifying the traffic density with large-scale network traffic.

From our point of view, there are several issues to extend the trained model in order to improve performance: (i) Designing a custom trained model which focuses on specific traffic datasets. Specifically, compared with other applications, the context of traffic image datasets have several different characteristics such as one channel value (e.g. traffic speed) and abstract features input images; (ii) Increasing the depth of CNN architecture, which is able to learn more complex relations and maintain the acceptable

| Labels | Description       | Loss | Val-Loss | Test (%) |
|--------|-------------------|------|----------|----------|
| 02     | low-normal        | 0.107| 0.084    | 91.25    |
| 02     | low-high          | 0.07 | 0.061    | 94.86    |
| 02     | normal-normal     | 0.411| 0.644    | 78.06    |
| 02     | normal-high       | 0.101| 0.081    | 91.66    |
| 03     | low-normal-normal | 0.513| 0.659    | 76.11    |
| 03     | low-normal-high   | 0.320| 0.463    | 86.48    |
| 03     | normal-normal-high| 0.596| 0.690    | 73.64    |
| 04     | low-normal-normal-high | 0.655 | 0.724 | 70.51 |

Figure 11. Results of the training model comparing with other CNN architectures.
time-consuming tasks. The aforementioned problems are interesting issues that we will take into account in the future work regarding this study.

**Acknowledgements**

This work was partly supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government (MSIT) (No.2018-0-00494, Development of deep learning-based urban traffic congestion prediction and signal control solution system) and Korea Institute of Science and Technology Information(KISTI) grant funded by the Korea government (MSIT) (K-20-L02-C09-S01).

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was partly supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korean Ministry of Science and ICT (MSIT) (No. 2018-0-00494, Development of deep learning-based urban traffic congestion prediction and signal control solution system) and Korea Institute of Science and Technology Information (KISTI) grant funded by the Korea government (MSIT) [K-20-L02-C09-S01].

**Notes on contributors**

**Jiho Cho** is a Postdoctoral Researcher in Korea Institute of Science and Technology Information (KISTI) since 2019. He received the Ph.D. degree in computer engineering at the Hannam University, Daejeon, Korea in 2019. His research interests are the intelligent transportation systems based on transportation big data and artificial intelligence, machine learning, network security and online privacy. Recently, he focuses on developing intelligent traffic management systems by using deep learning methodologies.

**Hongsuk Yi** is a principal researcher in Korea Institute of Science and Technology Information (KISTI), Korea, since 2000. He received the Ph.D. degree Sogang University, Seoul, Korea in 1997. His research interests are traffic congestion problem, smart cities, supercomputing, and heterogeneous computing by using deep learning techniques. Recently, he focuses on developing traffic signal control solutions by using deep learning methodologies.

**Heejin Jung** is a senior researcher in Korea Institute of Science and Technology Information (KISTI) since 2016. He received his Ph.D. degree in transportation engineering at the Virginia Tech., U.S. in 2012. His research area are the intelligent transportation systems based on transportation big data and artificial intelligence, intelligent intersection, autonomous vehicle and data-driven driving behavior. Especially, he focuses on an intelligent traffic management systems using convergence methodology of big data and AI.

**Khac-Hoai Nam Bui** received the M.S. degree in computer science and information engineering from Aletheia University, New Taipei, Taiwan, in 2014, and the Ph.D. degree in computer science and engineering from Chung-Ang University, Seoul, South Korea, in 2018. He has been a Postdoctoral Researcher with Korea Institute of Science and Technology Information (KISTI), Daejeon, Korea, since February 2019. His research interests are Internet of things and ambient intelligence by using AI methodologies such as data mining, machine learning, and logical reasoning.
References

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., & Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics*, 8(3), 292. doi: https://doi.org/10.3390/electronics8030292

Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011, December). Algorithms for hyper-parameter optimization. In *Neurips 2011. Proceedings of the 5th annual conference on neural information processing systems* (pp. 2546–2554).

Bui, K. H. N., & Jung, J. J. (2019). ACO-based dynamic decision making for connected vehicles in IoT system. *IEEE Transactions on Industrial Informatics*, 15(10), 5648–5655. doi: https://doi.org/10.1109/TII.9424

Bui, K. H. N., Oh, H., & Yi, H. (2020). Traffic density classification using sound datasets: An empirical study on traffic flow at asymmetric roads. *IEEE Access*, 8, 125671–125679. doi: https://doi.org/10.1109/Access.6287639

Chen, M., Yu, X., & Liu, Y. (2018). PCNN: Deep convolutional networks for short-term traffic congestion prediction. *IEEE Transactions on Intelligent Transportation Systems*, 19(11), 3550–3559. doi: https://doi.org/10.1109/TITS.2018.2835523

Chung, J., & Sohn, K. (2018). Image-based learning to measure traffic density using a deep convolutional neural network. *IEEE Transactions on Intelligent Transportation Systems*, 19(5), 1670–1675. doi: https://doi.org/10.1109/TITS.2017.2732029

Diao, Z., Wang, X., Zhang, D., Liu, Y., Xie, K., & He, S. (2019, January). Dynamic spatial-temporal graph convolutional neural networks for traffic forecasting. In *AAAI 2019. Proceedings of the 33rd AAAI conference on artificial intelligence* (pp. 890–897).

He, K., Zhang, X., Ren, S., & Sun, J. (2016, June). Deep residual learning for image recognition. In *CVPR 2016. Proceedings of the 29th IEEE conference on computer vision and pattern recognition* (pp. 770–778).

Hershey, S., Chaudhuri, S., Ellis, D. P. W., Gemmeke, J. F., Jansen, A., Moore, R. C., & Wilson, K. W. (2017, March). CNN architectures for large-scale audio classification. In *ICASSP 2017. Proceedings of the 42nd international conference on acoustics, speech and signal processing* (pp. 131–135).

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012, December). ImageNet classification with deep convolutional neural networks. In Nips 2012. *Proceedings of the 26th annual conference on neural information processing systems* (pp. 1106–1114).

Kurniawan, J., Syahra, S. G., & Dewa, C. K. (2018). Traffic congestion detection: Learning from CCTV monitoring images using convolutional neural network. *Procedia Computer Science*, 144, 291–297. doi: https://doi.org/10.1016/j.procs.2018.10.530

Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D. D., & Chen, M. (2014, December). Medical image classification with convolutional neural network. In *ICARCV 2014. Proceedings of the 13th international conference on control automation robotics & vision*.

Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2017). Graph convolutional recurrent neural network: Data-driven traffic forecasting. *CoRR* abs/1707.01926. http://arxiv.org/abs/1707.01926

Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2018, April). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In *ICLR 2018. Proceedings of the 6th the international conference on learning representations*.

Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F. (2015). Traffic flow prediction with big data: A deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, 16(2), 865–873. https://doi.org/10.1109/TITS.2014.2345663
Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., & Wang, Y. (2017). Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction. Sensors, 17(4), 818. doi: https://doi.org/10.3390/s17040818

Seo, Y., & Shin, K. (2019, February). Hierarchical convolutional neural networks for fashion image classification. Expert Systems with Applications, 116, 328–339. doi: https://doi.org/10.1016/j.eswa.2018.09.022

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016, June). Rethinking the inception architecture for computer vision. In CVPR 2016. Proceedings of the 29th IEEE conference on computer vision and pattern recognition (pp. 2818–2826).

Wang, Y., Zhang, D., Liu, Y., Dai, B., & Lee, L. H. (2019). Enhancing transportation systems via deep learning: A survey. Transportation Research Part C: Emerging Technologies, 99, 144–163. doi: https://doi.org/10.1016/j.trc.2018.12.004

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2019). A comprehensive survey on graph neural networks. CoRR abs/1901.00596. http://arxiv.org/abs/1901.00596

Wu, Y., Tan, H., Qin, L., Ran, B., & Jiang, Z. (2018, May). A hybrid deep learning based traffic flow prediction method and its understanding. Transportation Research Part C: Emerging Technologies, 90, 166–180. doi: https://doi.org/10.1016/j.trc.2018.03.001

Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., & Li, Z. (2018, February). Deep multi-view spatial-temporal network for taxi demand prediction. In AAAI 2018. Proceedings of the 32nd AAAI conference on artificial intelligence (pp. 2588–2595).

Yi, H., & Bui, K. H. N. (2019, September). VDS data-based deep learning approach for traffic forecasting using LSTM network. In EPIA 2019. Proceedings of the 19th epia conference on artificial intelligence (pp. 547–558).

Yi, H., & Bui, K. H. N. (2020). An automated hyperparameter search-based deep learning model for highway traffic prediction. IEEE Transaction on Intelligent Transportation Systems, 1–10. https://doi.org/10.1109/TITS.2020.2987614

Yu, B., Yin, H., & Zhu, Z. (2018, July). Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In IJCAI 2018. Proceedings of the 27th international joint conferences on artificial intelligence (pp. 3634–3640).

Zhu, L., Yu, F. R., Wang, Y., Ning, B., & Tang, T. (2019, January). Big data analytics in intelligent transportation systems: A survey. IEEE Transactions on Intelligent Transportation Systems, 20(1), 383–398. doi: https://doi.org/10.1109/TITS.2018.2815678