Master's Thesis of Engineering

Spatiotemporal Imputation of Traffic Speed Using Generative Adversarial Network

Generative Adversarial Network를 활용한 시공간적 통행 속도 결측 대체에 대한 연구

February 2021

Department of Civil & Environmental Engineering
Seoul National University
Civil and Environmental Engineering Major

Garyoung Lee

Spatiotemporal Imputation of Traffic Speed Using Generative Adversarial Network

Dong-Kyu Kim

Submitting a master's thesis of Civil & Environmental Engineering

February 2021

Department of Civil & Environmental Engineering
Seoul National University
Civil & Environmental Engineering Major

Garyoung Lee

Confirming the master's thesis written by
Garyoung Lee
February 2021

Chair Seung-Young Kho (Seal)

Vice Chair Chungwon Lee (Seal)

Examiner Dong-Kyu Kim (Seal)

Abstract

Data that are complete and accurate are the most important premises of providing reliable traffic information because they are required by most statistical analyses. However, the problem of missing data is unavoidable since the data collection system is not free of errors. Recently, deep learning approaches, which are capable of capturing the inherent features and interactions in the data, have been proposed to deal with the problem of missing data. Spatio-temporal dependencies are key for the imputation of traffic data, and color-coded traffic speed images in time-space diagrams can represent them. In this paper, we propose a multi-input deep-convolutional generative adversarial imputation network (MI-DC GAIN) to impute the network-wide traffic speed on an urban expressway in the form of speed images. The proposed method uses a convolutional neural network (CNN) to deal with spatio-temporal patterns in the speed images and GAIN to focus on the data imputation. To facilitate the training DC-GAIN, speed images reconstructed by the traffic adaptive smoothing method (TASM) were used in the multi-input structure as additional information. Findings from the experiment showed that applying CNN to the structure of GAIN can enhance the model capability of learning traffic speed images, which are enhanced further by the multi-input structure with the additional reconstructed speed images. The MI-DC GAIN achieved much better performance than benchmark models in terms of accuracy and robustness to the level-of-congestion and the missing rate.

Keyword : Traffic Data Imputation, Generative Adversarial Network, Deep Convolutional Network, Adaptive Smoothing Method, Spatio-temporal Dependency
Student Number : 2019-21695

Table of Contents

Chapter 1. Introduction

Providing information on traffic state is the most important task in the application of many traffic management approaches, such as vehicle routing and travel time prediction. Such applications are particularly important in congested traffic conditions, which are induced by recurrent bottlenecks or traffic incidents. The benefits of traffic information are determined by its completeness and its accuracy (Li et al. 2013). However, missing data are practically unavoidable due to failures in communication or the malfunctions of detectors. Since most of the methods that provide traffic information require complete datasets, the missing data should be considered first.

Several methods for imputing data have been proposed to deal with the problem of missing information, and these methods include prediction, interpolation, and statistical learning (Li et al. 2014). The prediction methods impute the missing values by modeling the historical pattern of the data collected from target sites, and the methods that are used extensively in this approach are the support vector machine (Wu et al. 2004, Kim et al. 2019a) and the k-nearest neighbors (Myung et al. 2011). However, these approaches have difficulty dealing with continuous missing values or transitions in the traffic state since they fail to capture the spatial correlation of the data. The interpolation methods found the optimal imputation value based on the traffic information of spatiotemporal neighborhoods or the similarity of patterns in the historical data (Treiber et al. 2011, Yin et al. 2012). Such approaches depend on the assumption of a repetitive traffic pattern at the target sites, but sometimes non-recurrent patterns occur in practice (Li et al. 2013). Statistical learning methods take advantage of the statistical characteristics of traffic flow to capture its meaningful variations based on assumed probability distribution (Li et al. 2013, Bae et al. 2018). However, such assumptions sometimes fail when dealing with real data, and they do not work well in most cases of missing data (Huang et al. 2020)

Recently, deep learning methods that achieve great success in

various fields of transportation research (Kim et al. 2019b, Polson and Sokolov 2017, Liang et al. 2018) have been proposed to deal with the traffic data imputation problem. The deep learning method efficiently represents the inherent features and interactions in the data (Duan et al. 2016). Therefore, complex spatial and temporal information collected from multiple data points can be considered in the structure of the model without the statistical assumptions or the domain knowledge of the researcher (Duan et al. 2016, Asadi et al. 2019). Deep learning methods outperform the conventional prediction method, and the improvement in performance was achieved by the advanced structure of the model, e.g., convolutional neural networks (CNNs) and long short-term memory (LSTM) (Liang et al. 2018, Duan et al. 2016, Asadi et al. 2019).

The generative adversarial network (GAN) is a generative deep-learning method that learns the conditional probability distribution of the input and output data and generates data that have the approximated distribution of the training data (Goodfellow et al. 2014). Modeling the conditional probability distribution, rather than predicting the expected value with given inputs, allows the user to address abrupt changes caused by spatio-temporal attributes in the congested traffic, such as phase transition, and oscillation marking deceleration and acceleration. However, learning the conditional distribution rather than predicting the expected value is considered to be a much more difficult task in deep learning. Specifically, GAN has a non-convergence issue since it is trained by finding an equilibrium of two competing networks at the same time rather than by finding a minimum of the loss function (Goodfellow et al. 2014). Although several advanced processes for training GAN have been proposed (Arjovsky et al. 2017, Chen et al. 2016), heuristics empirically demonstrated by domain knowledge worked well in practice (Salimans et al. 2016). In the transportation research field, previous studies have reported the use of GAN models to predict traffic flow (Liang et al. 2018, Lv et al. 2018, Lin et al. 2018) and to impute missing traffic data (Huang et al. 2020, Chen et al. 2016). Lv et al. introduced the potential of GAN for traffic data generation by learning conditional probability distribution (Lv et al. 2018). Liang et

al. proposed GAN to predict traffic flow, which consisted of two modifications, i.e., LSTM and the custom loss function based on the relation of traffic flow and density (Liang et al. 2018). Huang et al. proposed time-dependent encoding to represent the time dependency of the traffic data as images, and those images are trained using CNN (Huang et al. 2020). The above studies showed strong promise of GAN in modeling traffic flow, and the models that tailored to the problem of traffic data outperformed the conventional prediction model and the original GAN with a fully connected network (FCN).

In this study, we propose a multi-input deep convolutional generative adversarial imputation network (MI-DC GAIN) that focuses on the imputation of missing spatio-temporal traffic data. We use the image-based approach that allows a structure of deep learning model to automatically recognize spatio-temporal patterns represented in the images (Ma et al. 2017, Jo et al. 2018). The images of speed contour in the urban expressway that represent spatio-temporal patterns of traffic state are used as the input and output data of the model, and those traffic speed images are trained using a convolutional structure. To facilitate the training of the model, we used the speed contour reconstructed by the traffic adaptive smoothing method (TASM) (Treiber et al. 2011) as an additional input (i.e., multi-input). The performance of the proposed model was evaluated based on different missing rates and the level of congestion. The key contributions of this study are as follows:

- An image-based approach using GAIN and CNN were applied to traffic data imputation, which requires capturing the spatio-temporal patterns of traffic data.
- Multi-input structures that use traffic speed images reconstructed by TASM were proposed to enhance the applicability of DC GAIN to the traffic data.
- MI-DC GAIN achieved a much higher and robust performance for all cases of the missing data ratio and level of congestion than the benchmark models, including temporal moving average, TASM, single-input deep-convolutional GAIN (SI-DC GAIN), and multi-input fully-connected GAIN (MI-FC GAIN).

The rest of this paper is organized as follows. First, we describe the study site and the MI-DC GAIN and the TASM in detail. Then, we discuss the specific implementation of the model and provide the experimental results to validate the proposed method. Last, conclusions and potential future research are presented.

Chapter 2. Study Site and Data

2.1. Study Site

The loop detector data are collected by the inductive double-loop detectors, which are commonly used to allow for direct speed measurements (Kessler et al. 2018). This research focuses on the detectors located in the urban expressway of the Seoul metropolitan area in Korea. The study site is a 9.36 km section from Pangyo JC to Songpa IC in the counterclockwise direction in the Seoul Outer Circular Expressway (Figure 2.1).

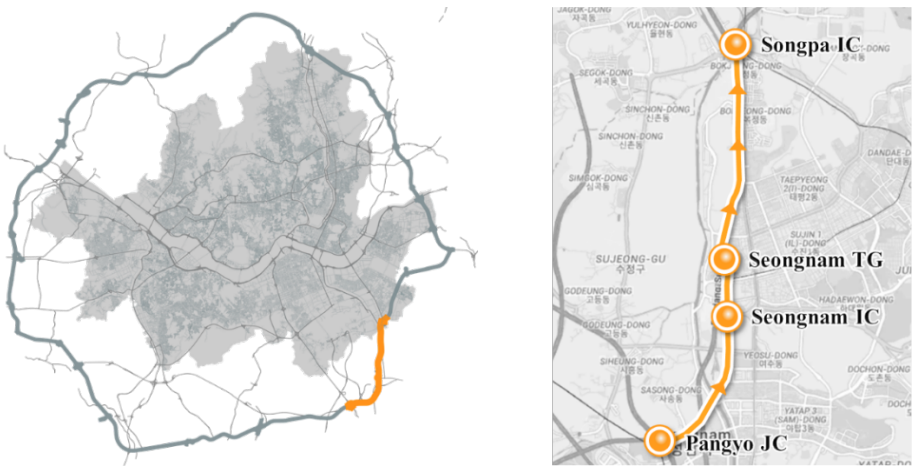


Figure 2.1 The study site from Pangyo JC to Songpa IC on the Seoul Outer Circular Expressway

2.2. Data Descriptions

From Pangyo JC to Songpa IC, there are nine loop detectors with kilometer posts of 1.6, 2.5, 3.5, 4.3, 5.5, 6.5, 7.2, 8.0, and 9.0, respectively. The average distance between the loop detectors is 0.925 km, with a standard deviation of 0.148 km. We used 5-minute aggregated traffic speed data collected from the loop detectors for 245 days, i.e., from March 2019 through October 2019. The speed limit on this route is 100 km/h, and the average traffic speed and standard deviation of the traffic speed for each detector are shown in Table 2.1. The low average speed of the detectors at the 3.5 and 5.5 kilometer posts suggest the occurrences of recurrent traffic congestion.

Table 2.1 Descriptive statistics of traffic speed data at the study site

	Pangyo JC ~ Seongnam IC			Seongnam IC ~ Seongnam TG	
Kilometer Post of Detector	1.6	2.5	3.5	4.3	
Average Speed (km/h)	84.5	85.8	81.2	96.9	
Standard Deviation of Average Speed (km/h)	17.9	19.4	19.2	15.1	
Seongnam TG ~ Songpa IC					
Kilometer Post of Detector	5.5	6.5	7.2	8.0	9.0
Average Speed (km/h)	76.5	86.5	83.6	94.8	84.2
Standard Deviation of Average Speed (km/h)	14.7	17.0	20.1	19.9	18.4

Figure 2.2 shows the average speed measured in each detector over time with the kilometer post. The average traffic speed shows that the peak traffic hours in the study site were approximately 11 A.M. and 5 P.M., and those patterns commonly were observed at all of the detectors.

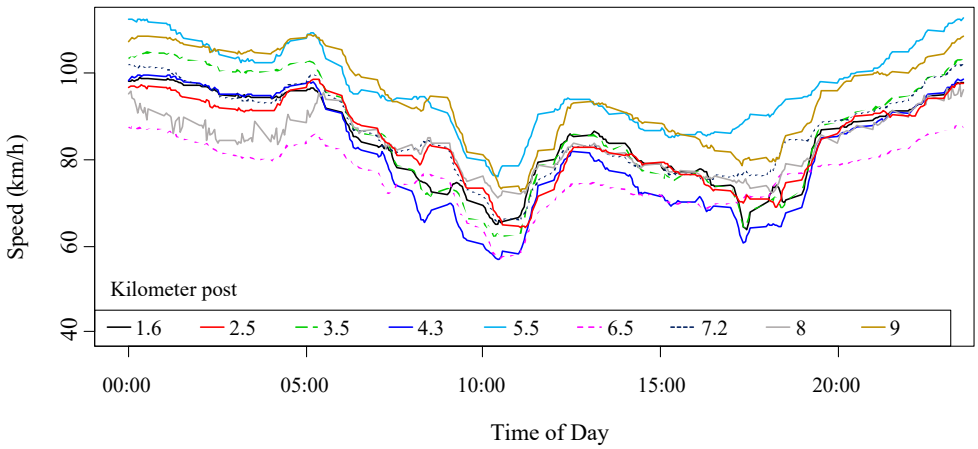


Figure 2.2 Average speed recorded by each detector over time

Figure 2.3(a) shows an example of the image of speed contour on the time-space diagram. Each pixel of the images expresses the traffic speed, i.e., the lower the speed, the redder the color, and the higher the speed, the bluer the color. Figure 2.3(b) shows an image that transformed Figure 2.3(a) into an image matrix that was used as input data for training the models. This traffic speed image naturally represents the spatio-temporal correlations (e.g., propagation of traffic congestion) of the daily traffic. The dimensions of the input image matrix were 75 x 108, of which the section length of 7.5 km was divided into 0.1-km units (i.e., from the 1.6 kilometer post to the 9.0 kilometer post). The 9 hours of data, which were from 10:00 A.M. to 7:00 P.M., were divided into 5-minute intervals. To represent the discrete loop detector data at the 9 locations as a filled image, we defined "influence areas" as those areas that are bounded by the midpoints between nearby detectors (Zheng et al. 2010) based on the assumption that the traffic state within the influence area remains the same. This assumption can be reasonable since the purpose of this study is to impute the missing data in certain locations where detectors are equipped.

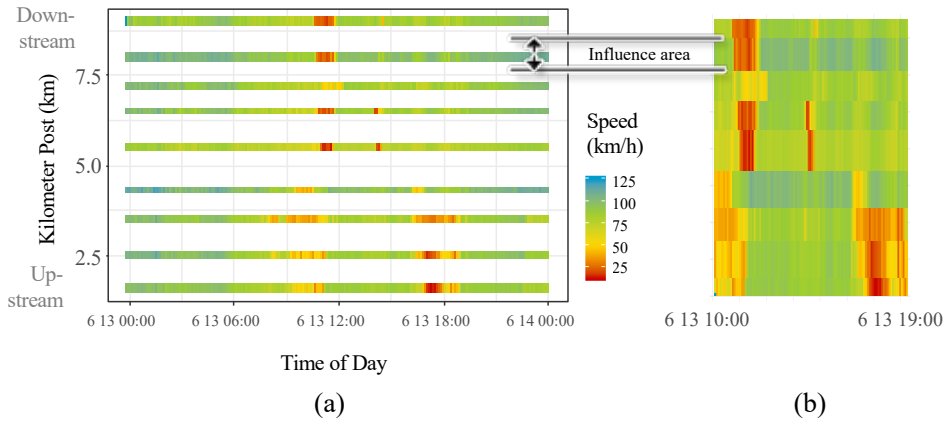


Figure 2.3 Speed images: (a) spatio-temporal image of traffic speed in the time-space diagram; (b) transformed speed images which are input for the training model

There is no consistent method to measure the level-of-congestion of daily traffic speed. Several studies have identified that the traffic congestion exists where the duration time of speed under the threshold is sufficiently long (Kim et al. 2019a, Lorenz and Elefteriadou 2001, Kim et al. 2010). Based on that concept, we measured the level-of-congestion using the mean and standard deviation of daily speed. Figure 4 shows the scatter plot of the mean and standard deviation of the daily traffic speed obtained from 9 detectors in the study site. The mean and standard deviation of the traffic speed were calculated using the speed of all the vehicles that passed through the target section during a single day (i.e., the speed value of the image matrix constructed with the nine detectors as shown in Figure 2.3(a)). The detector data showed that there was a tendency of inverse linear relationship between the mean and the standard deviation of daily traffic speed (Figure 2.4). We defined high level-of-congestion as being related to the lower the mean of daily traffic speed and the greater standard deviation of daily traffic speed.

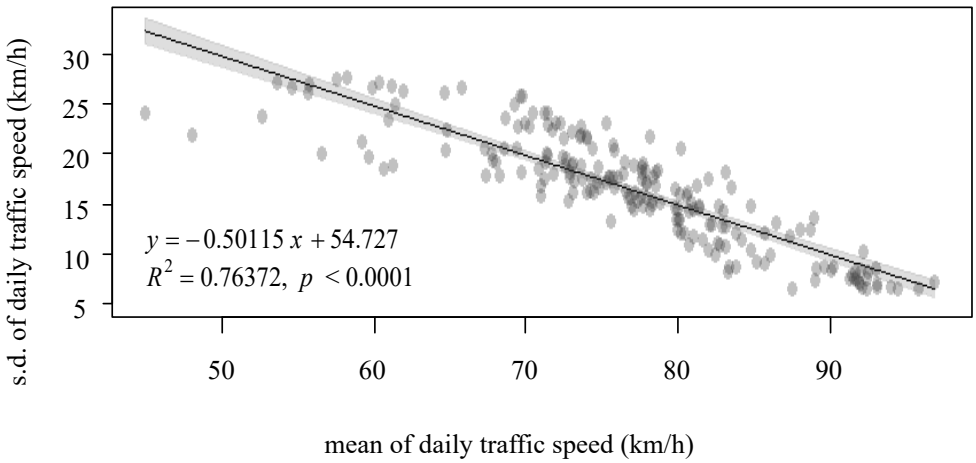


Figure 2.4 Distribution of the mean and standard deviation of the daily speed at the study site

Chapter 3. Methods

3.1. Generative Adversarial Imputation Network (GAIN)

The structure of GAIN has the same basic architecture as the widely-known GAN, which consisted of two neural networks, i.e., a generator and a discriminator (Yoon et al. 2018). The generator aims to generate realistic samples so that the discriminator classifies those generated samples as a real sample. In contrast, the discriminator has the adversarial role of distinguishing the real samples from the generated samples. That adversarial competition converges to an equilibrium where the discriminator cannot distinguish between the generated samples and the real data. In other words, at equilibrium, the generator can approximate the distribution of the training data (Goodfellow et al. 2014, Lin et al. 2018). The global optimality of the training process in equilibrium was proven by Goodfellow et al.

The main difference between GAN and GAIN in the structure of the model is the existence of the mask vector and the hint vector as inputs for the generator and the discriminator each. The mask vector, $\mathbf{M} = (M_1, \ldots, M_d)$, takes a value in $\{0,1\}^d$ that indicates whether the

components of the data vector, $\mathbf{X} = (X_1, \ldots, X_d)$, are missed or observed. In the problem setup purposed at imputation, the input data are formulated as follows. A random variable, $\tilde{\mathbf{X}} = (\tilde{X}_1, \cdots, \tilde{X}_d) \in \tilde{\chi}$, is defined in a d-dimensional space $\chi = \chi_1 \times \cdots \times \chi_d$. For each $i \in \{1, \ldots, d\}$, $\tilde{X}_i$ has the value of X_{I} where the value of M_i is 1, otherwise NaN. If the copy of $\tilde{X}$ is denoted as $\tilde{x}^1$, $\tilde{x}^2$, ..., $\tilde{x}^n$, the goal of the GAIN is to impute the missing value in each $\tilde{x}^i$ by estimating the conditional distribution of X given $\tilde{X} = \tilde{x}^i$. A random variable, H, takes a value that depends on M in a space $\mathcal{H}$, and it is used as an additional input to the discriminator, called the hint vector (Yoon et al. 2018). The hint vector provides the discriminator some partial information about the probability that the vector components were observed (i.e., the vector components were missing). Let $\mathcal{H} = \{0, 0.5, 1\}^d$ and given M, hint vector is defined as $\mathbf{H} = \mathbf{B} \odot \mathbf{M} + 0.5(1 - \mathbf{B})$ where the random variable $\mathbf{B} = (B_1, \ldots, B_d) \in \{0,1\}^d$ is defined by sampling k from $\{1, \ldots, d\}$ uniformly at random in each batch and B_j is 1 if $j \neq k$, otherwise 0. An appropriate proportion of hints can make the discriminator reasonably smart, thereby making the generator and the discriminator learn effectively (Yoon et al. 2018).

The input process of the generator is described in Equation 1 and Equation 2. The Generator, $G: \tilde{\chi} \times \{0,1\}^d \times [0,1]^d \rightarrow \chi$, emits output of imputed vector, $\bar{\mathbf{X}}$, with the input taking with $\tilde{\mathbf{X}}$, $\mathbf{M}$, and $\mathbf{Z}$. The random noise variable, $\mathbf{Z} = (Z_1, \ldots, Z_d)$, is independent of all other variables. The imputed vector $\tilde{\mathbf{X}}$ is used to make the completed data vector $\hat{\mathbf{X}}$.

$$\bar{\mathbf{X}} = G\left(\tilde{\mathbf{X}}, \mathbf{M}, (1 - \mathbf{M}) \odot \mathbf{Z}\right) \quad (3.1)$$

$$\hat{\mathbf{X}} = \mathbf{M} \odot \tilde{\mathbf{X}} + (1 - \mathbf{M}) \odot \bar{\mathbf{X}} \quad (3.2)$$

where $\odot$ is an element-wise multiplication.

Another structural difference between GAN and GAIN is the dimension of the discriminator's output. Whereas GAN's discriminator classifies whether the data vector is the real vector or the generated vector, GAIN's discriminator classifies whether each component of the data vector is a real value or an imputed value. The discriminator, D,

takes both the completed data vector $\hat{\mathbf{X}}$ and the hint matrix $\mathbf{H}$ as input. The discriminator, D, emits estimated mask matrix, $\hat{\mathbf{M}} = D(\hat{\mathbf{X}}, \mathbf{H})$, as output in which every component of the matrix is the probability that the value is observed. When the discriminator is a function of $D : \chi \rightarrow [0,1]^d$ with the i-th component of $D(\hat{x})$ is the probability that the ith component of $\hat{x}$ is observed.

The two fully-connected neural networks (or the other structure such as CNN), D and G, are trained by optimizing the following minimax problem for the quantity $V(D,G)$ in Equation 3. The D is trained to maximize the probability of correctly predicting $\mathbf{M}$, while the G is trained to minimize the probability of D correctly predicting $\mathbf{M}$ (Yoon et al. 2018). The objective function of the problem can be expressed in Equation 4 and Equation 5, according to the goals of the generator and the discriminator, respectively. As G is trained to minimize the weighted sum of the losses, the generator loss is affected significantly by the value of the hyperparameter, α. D and G are updated iteratively by the stochastic gradient descent method by incorporating each other.

$$\min_G \max_D V(D,G) = \mathbb{E}_{\hat{\mathbf{X}},\mathbf{M},\mathbf{H}}[\mathbf{M}^{\mathrm{T}} \log D(\hat{\mathbf{X}}, \mathbf{H}) + (1 - \mathbf{M}^{\mathrm{T}}) \log (1 - D(\hat{\mathbf{X}}, \mathbf{H}))] \quad (3.3)$$

$$\min_D -[\textstyle\sum_{j=1}^{k_D} L_D(\mathbf{m}(j), \hat{\mathbf{m}}(j), \mathbf{b}(j))] \quad (3.4)$$

$$\min_G [\textstyle\sum_{j=1}^{k_G} L_G(\mathbf{m}(j), \hat{\mathbf{m}}(j), \mathbf{b}(j)) + \alpha L_M(\tilde{\mathbf{x}}(j), \hat{\mathbf{x}}(j))]$$

$$L_D(\mathbf{m}, \hat{\mathbf{m}}, \mathbf{b}) = \textstyle\sum_{i:b_i=0} [m_i \log(\hat{m}_i) + (1 - m_i) \log(1 - \hat{m}_i)] \quad (3.5)$$

$$L_G(\mathbf{m}, \hat{\mathbf{m}}, \mathbf{b}) = -\textstyle\sum_{i:b_i=0} (1 - m_i) \log(1 - \hat{m}_i)$$

$$L_M(\tilde{\mathbf{x}}, \hat{\mathbf{x}}) = \textstyle\sum_{i=1}^{d} m_i (\hat{x}_i - \tilde{x}_i)^2$$

where $\log$ is an element-wise logarithm; k_D and k_G are the sizes of mini-batch of discriminator and generator; $\tilde{x}_i$, $\hat{x}_i$, b_i, $\hat{m}_i$ and m_i are i-th sample of the corresponding $\tilde{\mathbf{X}}, \hat{\mathbf{X}}, \mathbf{B}, \hat{\mathbf{M}}$, and $\mathbf{M}$; $\mathbf{m}(j), \mathbf{b}(j)$ and $\mathbf{x}(j)$ are j-th sample of the mini-batch. $L_D : \{0,1\}^d \times [0,1]^d \times \{0,1\}^d \rightarrow \mathbb{R}$ is the loss function of the discriminator, $L_G : \{0,1\}^d \times [0,1]^d \times \{0,1\}^d \rightarrow \mathbb{R}$ is the loss function of the generator, and $L_M : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is the loss function of the mask vector. More details about the theoretical background of GAIN are discussed in Yoon et al. (Yoon et al. 2018).

3.2. Traffic Adaptive Smoothing Method (TASM)

The adaptive smoothing method (ASM) is an interpolation method that applies the nonlinear spatio-temporal low-pass filter to the data (Treiber et al. 2011). Treiber et al. proposed the traffic-specific ASM (TASM) that considers the propagation patterns of the traffic flow according to the traffic state (i.e., congested state and free-flow state) (Treiber et al. 2011). The parameters for TASM are determined based on the general attributes of traffic flow propagation, such as propagation speed and critical speed dividing congested and free-flow state. TASM can filter out random fluctuations by considering the spatio-temporal characteristics of the traffic pattern.

The nonlinear filter transforms the discrete input detector speed data, v_i, into smooth spatio-temporal functions $V(x,t)$. The filter is expressed in Equation 3.6 as the weighted sum of the congested state function $V_{cong}(x,t)$ and the free-flow state function $V_{free}(x,t)$. Each function is defined as the sum of the multiplications of the smoothing kernel $\phi(x,t)$ and discrete speed data v_i divided by normalization factor $N(x,t)$ (Equation 3.8 and Equation 3.10). The smoothing kernel $\phi(x,t)$ is a localized function that decreases when the distance between the points increases (Equation 3.7). The normalization factor $N(x,t)$ is defined as the sum of the smoothing kernels $\phi(x,t)$ (Equation 3.9 and Equation 3.11).

$$V(x,t) = W(x,t) \cdot V_{cong}(x,t) + [1 - W(x,t)] \cdot V_{free}(x,t) \quad (3.6)$$

$$\phi(x,t) = \exp\left(-\frac{|x|}{\sigma} - \frac{|t|}{\tau}\right) \quad (3.7)$$

$$V_{cong}(x,t) = \frac{1}{N_{cong}(x,t)} \sum_{i=1}^{n} \phi\left(x - x_i, t - t_j - \frac{x - x_i}{c_{cong}}\right) v_i \quad (3.8)$$

$$N_{cong}(x,t) = \sum_{i=1}^{n} \phi\left(x - x_i, t - t_j - \frac{x - x_i}{c_{cong}}\right) \quad (3.9)$$

$$V_{free}(x,t) = \frac{1}{N_{free}(x,t)} \sum_{i=1}^{n} \phi\left(x - x_i, t - t_j - \frac{x - x_i}{c_{free}}\right) v_i \quad (3.10)$$

$$N_{free}(x,t) = \sum_{i=1}^{n} \phi\left(x - x_i, t - t_j - \frac{x - x_i}{c_{free}}\right) \quad (3.11)$$

$$W(x,t) = \frac{1}{2}\left[1 + \tanh\left(\frac{V_c - \min\left(V_{cong}(x,t),\ V_{free}(x,t)\right)}{\Delta V}\right)\right] \quad (3.12)$$

where σ is a unit range of spatial smoothing, and τ is a unit range of temporal smoothing, which is recommended to use half the value of the aggregation interval; c_{cong} is a propagation speed of perturbations in congested traffic; c_{free} is a propagation speed of perturbations in free traffic; V_c is crossover speed from free to congested traffic; ΔV is the width of the transition region. Based on the observation and recommended values in Treiber et al., this study set c_{cong}, c_{free}, V_c, and ΔV as -15km/h, 80km/h, 60km/h, and 20km/h, respectively (Treiber et al. 2011).

Figure 3.1 shows an example of traffic speed images reconstructed by TASM. Figure 3.1(a) shows the loop detector data without filling the influence areas, and Figure 3.1(b) is the 75 x 108 image that reconstructed Figure 3.1(a) with TASM. More details about TASM are discussed in Treiber et al. (Treiber et al. 2011).

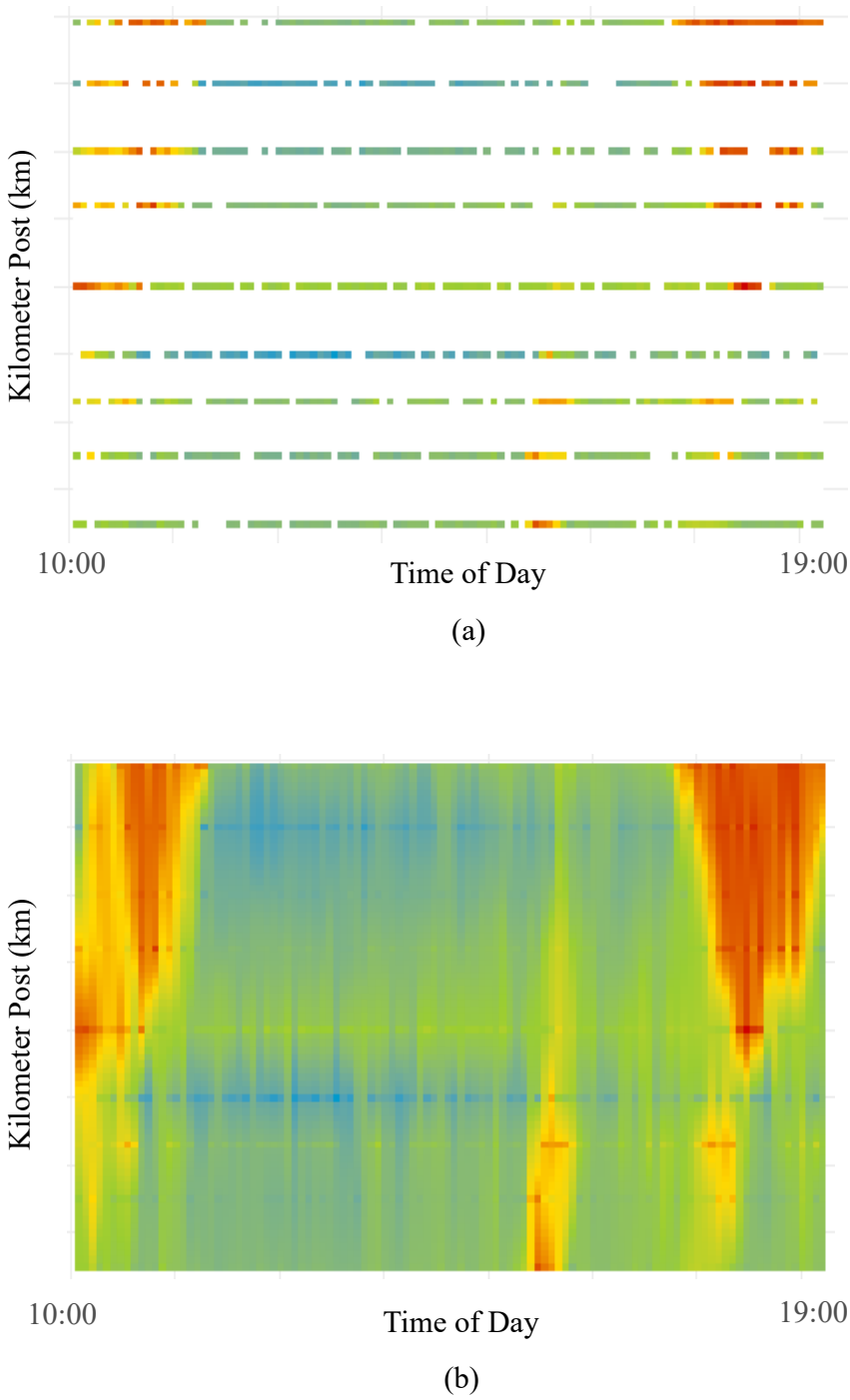


Figure 3.1 Speed images: (a) Traffic speed data collected from 9 loop detectors; (b) traffic speed image reconstructed by TASM

3.3. Multi-input GAIN

In machine learning, multiple inputs from different aspects often are used to improve learning performance, and those also can be used to improve the imputation performance of GAIN. The appropriate combination of data to process multi-input will determine whether the performance improves or not. Providing data which are completely irrelevant to each other as multi-input would only confuse a model to be learned. Based on the original GAIN architecture, a multi-input approach is applied to use the data from different aspects. Figure 3.2 shows the full model of the architecture. The input to the network consists of two different aspects, i.e., the loop detector data filled with influence areas and the data reconstructed by TASM. The logic behind this design is to train GAIN more efficiently with the aid of imputation results of TASM that consider general spatio-temporal traffic patterns.

3.4. Deep Convolutional GAIN

The original GAIN structure consists only of fully-connected layers. The original generator and discriminator pass through two hidden layers in a fully-connected form. As a fully-connected neural network cannot exploit local connectivity of the input data, we should introduce CNN to impute data taking into consideration spatio-temporal correlation in the data. CNN is effective for handling image format of which adjacent pixels of the input are correlated to each other (Jarrett et al. 2009). We followed the DCGAN architecture (Radford et al. 2015) for the convolutional generator and discriminator of GAIN. Pooling layers are replaced with strided convolutions in the discriminator and fractional-strided convolutions in the generator. Batch normalizations are used in both the generator and the discriminator after the convolutional layers. In the generator, all layers except the output layer used the ReLU activation function, and the output layer used the Tanh function. LeakyReLU was used in all layers in the activation function of the discriminator.

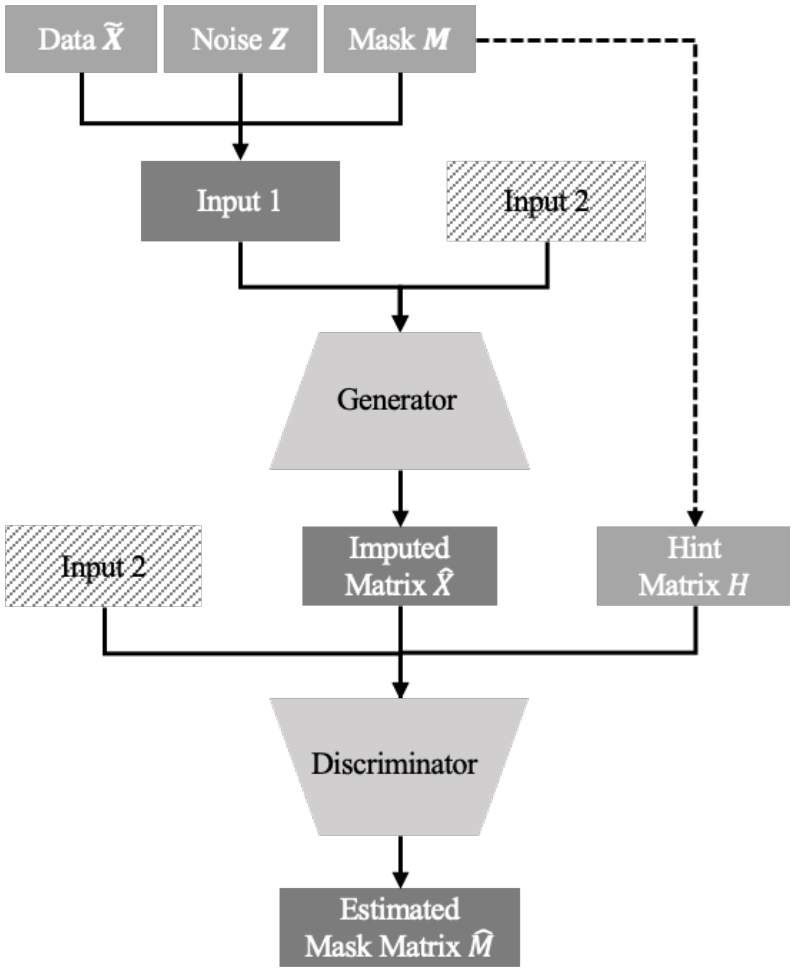
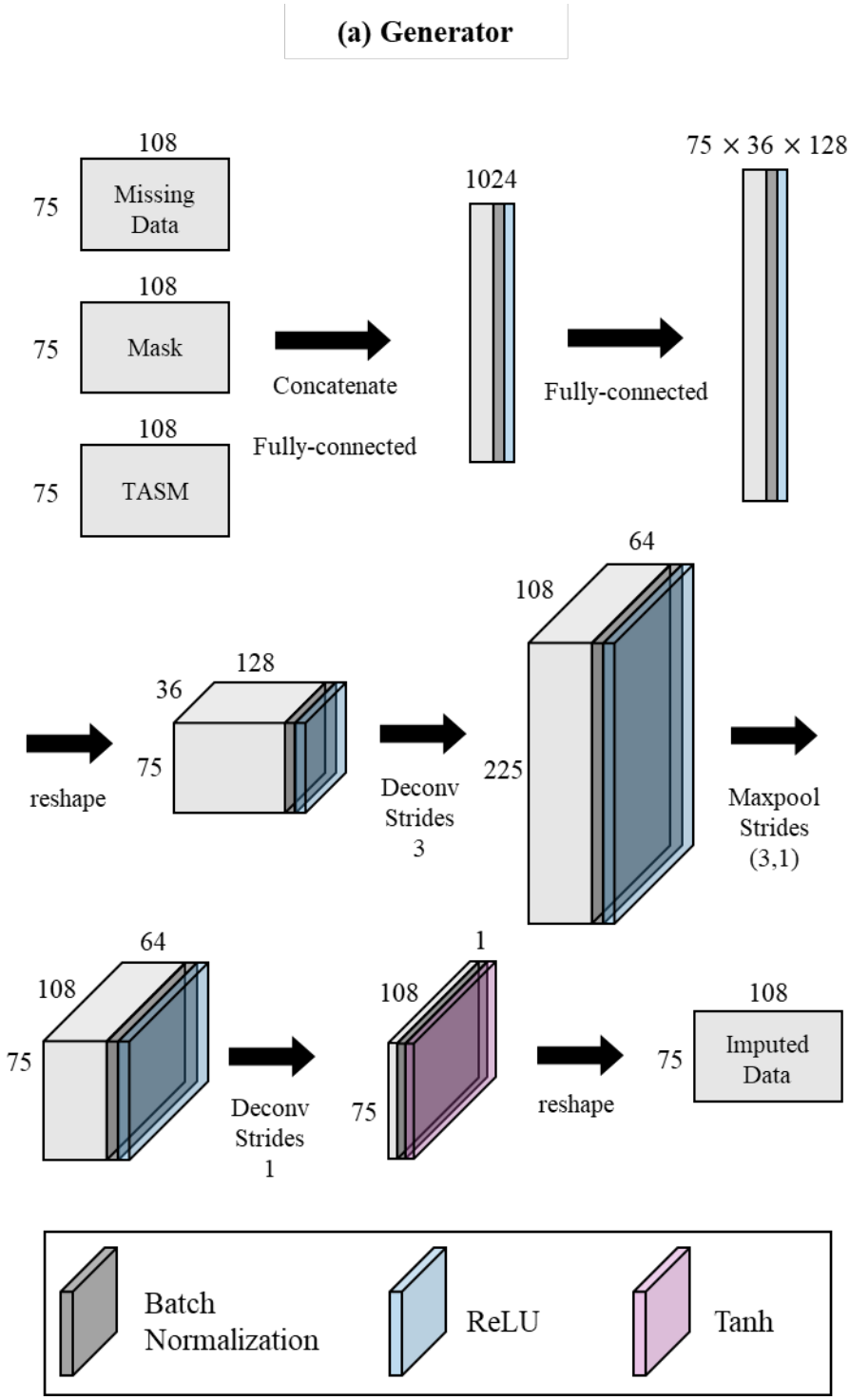


Figure 3.2 Structure of Multi-Input GAIN

3.5. MI-DC GAIN Framework

This section specifies the architecture of the proposed model using convolutional layers and multi-input. 1) The loop detector data, 2) its mask matrix, and 3) the reconstructed data with TASM were set to the identical dimensions of 75 x 108. These data were normalized with min-max normalization and then concatenated as the multi-input of the generator. The input was stretched to a fully-connected layer and passed through the ReLU activation function and batch normalization. After it had passed through another fully-connected layer with ReLU and batch normalization, then the reshaped input goes through the transposed convolutional layer and the max-pooling process. The final layer also is a transposed convolutional layer that uses the Tanh activation function. The generator finally emits the imputed matrix. Receiving the imputed matrix as an input of discriminator, the hint matrix and reconstructed data with TASM also are used as multi-inputs to the discriminator. The concatenated multi-input follows the convolutional layer and the max-pooling process twice each. After the last max pooling, the reshaped input is flattened to a fully-connected layer and reformed as an estimated mask matrix with dimensions of 75 x 108. The activation functions of all layers are set as leaky ReLU. Figure 3.3 shows the overall process explained above.

After the series of experiments to specify the model framework was completed, the hint rate was set as 40%, the training rate was set as 90%, the hyper-parameter, α, was set as 500, and the size of the mini-batch was set as 200. During the iterations, the model calculates the training loss and the test loss. The training loss is computed as an error between the generated values and the observed values, and the test loss is computed as an error between the imputed values and the missing values. The final objective of this model is to minimize the test loss that indicates the imputation performance for missing data.


(a) Generator
108
75
Missing Data
108
75
Mask
108
75
TASM
Concatenate
Fully-connected
1024
Fully-connected
75 × 36 × 128
reshape
128
36
75
Deconv Strides 3
64
108
225
Maxpool Strides (3,1)
64
108
75
Deconv Strides 1
1
108
75
reshape
108
75
Imputed Data
Batch Normalization
ReLU
Tanh

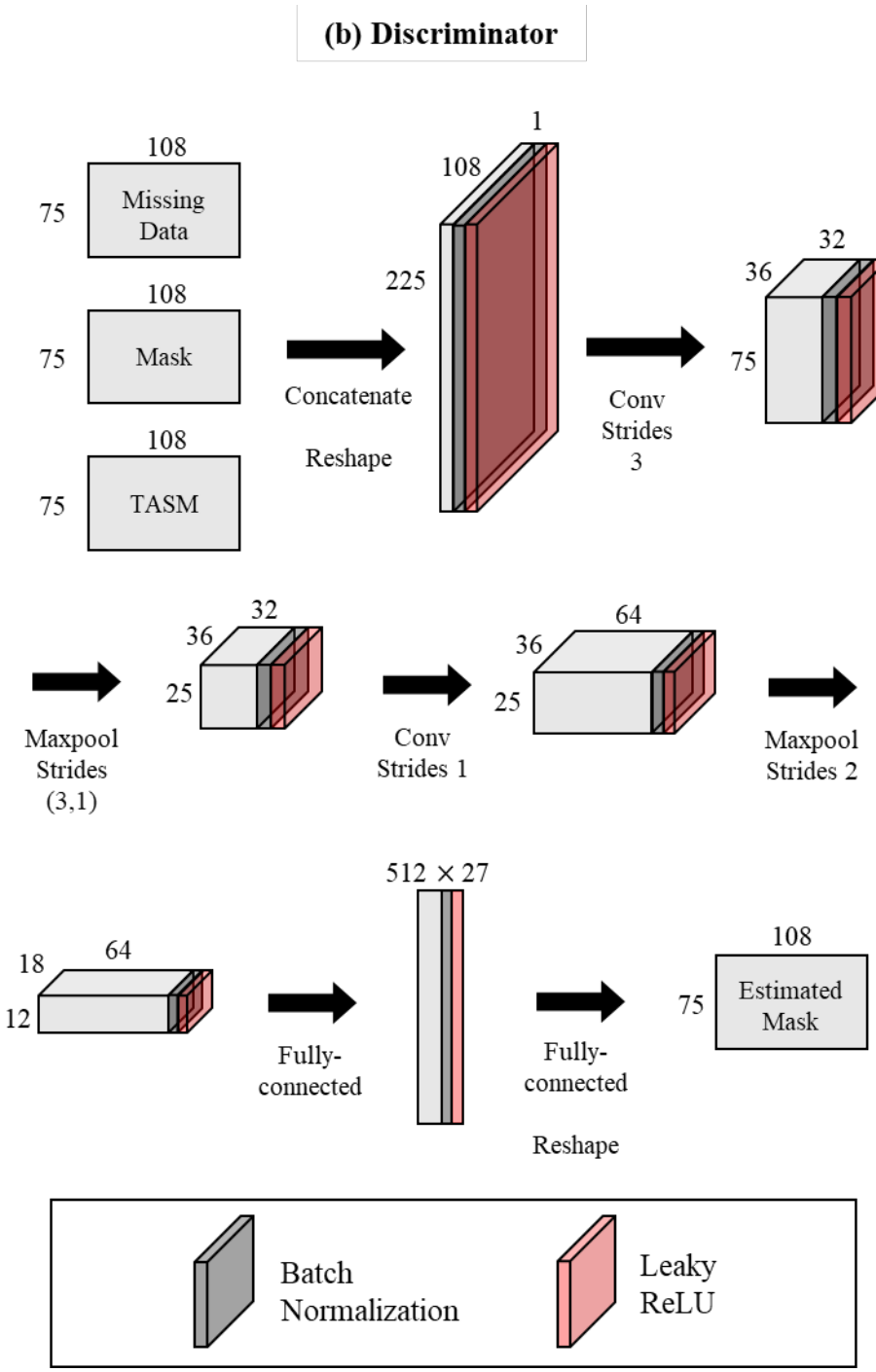


Figure 3.3 The proposed MIDC-GAIN framework: (a) Generator; (b) Discriminator

Chapter 4. Results of Application and Discussion

4.1. Imputation Performance Comparison

The imputation performance of the proposed MI-DC GAIN was compared with those of the four benchmark models. First, single-input DC GAIN (SI-DC GAIN) was constructed to evaluate whether multi-input approach can improve the imputation performance of DC GAIN. Second, we constructed another multi-input GAIN with only fully-connected layers (MI-FC GAIN) to evaluate the contribution of considering local connectivity to the performance of MI-DC GAIN. Third, we also compared MI-DC GAIN to a baseline model that imputes missing data by interpolating via a spatio-temporal nonlinear low-pass filter, i.e., TASM. Fourth, another baseline interpolation method, moving-average (MA) considering temporal continuity, is applied to impute missing data. The imputation performance was evaluated by the root mean squared error (RMSE), given by the following Equation 4.1.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n_{miss}} (x - \hat{x})^2}{n_{miss}}} \qquad (4.1)$$

where x is the observed traffic speed data, $\hat{x}$ is the imputed traffic speed data, and n_{miss} is the number of missing components.

Table 4.1 shows the RMSE statistics of each model by the missing rate. We randomly generated missings with a ratio of 5%, 10%, 20%, and 30% on ground-truth data and evaluated imputation performance. The missings include spatial and temporal continuous missing data which are located continuously in time and space. The robustness of the models according to level-of-congestion is discussed in Figure 4.1. The mean and standard deviation of the daily traffic speed are proxy measures for the level-of-congestion on that day (i.e., the number of congestion occurrences, durations of the occurrences, and changes in

traffic speed due to the congestion). The lower mean of daily traffic speed and higher standard deviation of daily traffic speed indicate higher level-of-congestion. As we identified a linear inverse relationship between the mean and the standard deviation of daily traffic speed in the study site (See Figure 2.4), a low mean of daily traffic speed can represent the high level-of-congestion itself. Figure 4.1(a) shows examples of spatio-temporal traffic speed images by the mean of daily traffic speed. The red-colored areas indicating the congested traffic state occur more at the lower mean of daily traffic speed. Figure 4.1(b) shows the RMSE distribution of five models by the mean of daily traffic speed.

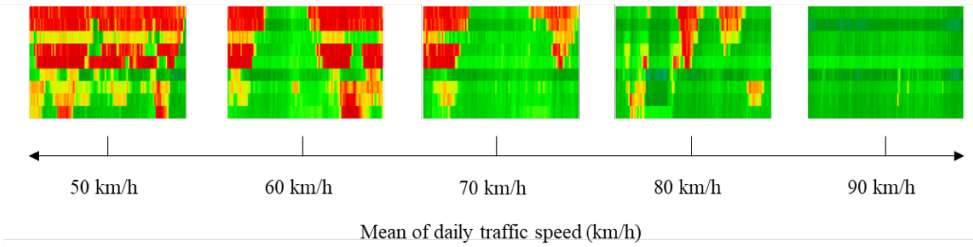
MI-DC GAIN showed the best performance compared to the four benchmark models, followed by SI-DC GAIN, moving-average, TASM, and MI-FC GAIN. The GAIN-based models showed robust performance in general against the missing rate, while TASM and MA showed a slightly lower performance as the missing rate increased. This result indicates the capability of GAIN that learn the distribution of the data rather than the expected value of the data, allowing the model to deal with multiple and continuous missing data by capturing the uncertainty of the imputed values (Yoon et al. 2018).

Table 4.1 Comparison of Imputation Performance between the Models

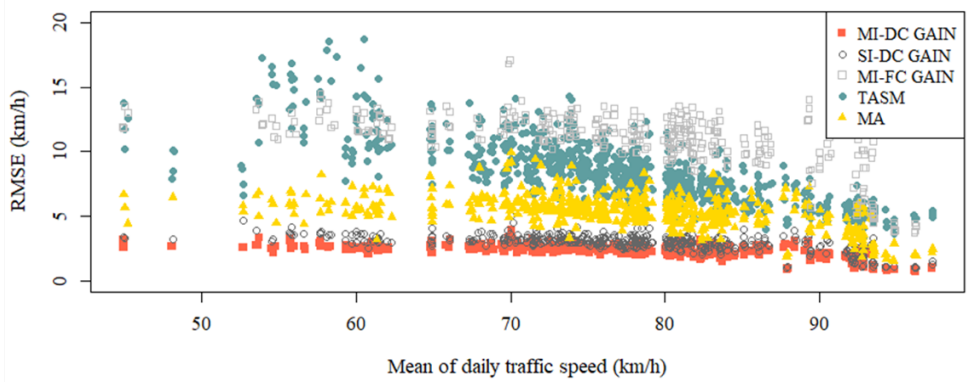
RMSE (km/h)	MI-DC GAIN			SI-DC GAIN			MI-FC GAIN		
Missing rate	mean	min	s.d	mean	min	s.d	mean	min	s.d
5%	2.42	0.71	0.54	2.94	0.98	0.67	11.83	4.47	2.09
10%	2.36	0.75	0.50	2.95	1.06	0.65	11.26	4.24	2.03
20%	2.32	0.74	0.50	3.00	1.01	0.62	10.40	3.73	2.03
30%	2.35	0.75	0.56	3.00	0.96	0.64	10.36	3.88	2.02
Overall	2.35	0.71	0.52	2.97	0.96	0.64	10.96	3.73	2.12

RMSE (km/h)	TASM			MA		
Missing rate	mean	min	s.d.	mean	min	s.d.
5%	8.51	3.64	2.67	5.05	1.36	1.73
10%	8.53	3.39	2.40	5.08	1.26	1.77
20%	8.77	4.17	2.42	5.52	1.68	1.73
30%	8.96	4.21	2.45	5.70	1.42	1.54
Overall	8.69	3.39	2.46	5.32	1.26	1.72

Note: s.d. means standard deviation.



(a)



(b)

Figure 4.1 (a) spatio-temporal images according to the mean of daily travel speed; (b) The imputation performance of models according to the mean of daily travel speed

4.2. Discussions

4.2.1. Advantages of Multi-Input (MI-DC GAIN vs. SI-DC GAIN)

In this application, we aim to evaluate the contribution of traffic images reconstructed by TASM to missing data imputation. Although MI-DC GAIN and SI-DC GAIN showed consistent performance regardless of the level-of-congestion (Figure 4.1), MI-DC GAIN had better overall performance than SI-DC GAIN. Although the SI-DC GAIN was well-performed, it has been confirmed that the multi-input approach using a reconstructed image can further improve the imputation performance. In other words, the reconstructed image representing different aspects of the data can contribute to identifying the distribution of target data. Therefore, beyond the scope of this study, the other data sources such as floating car and automatic vehicle identification could also be used as multi-input to enhance the proposed method.

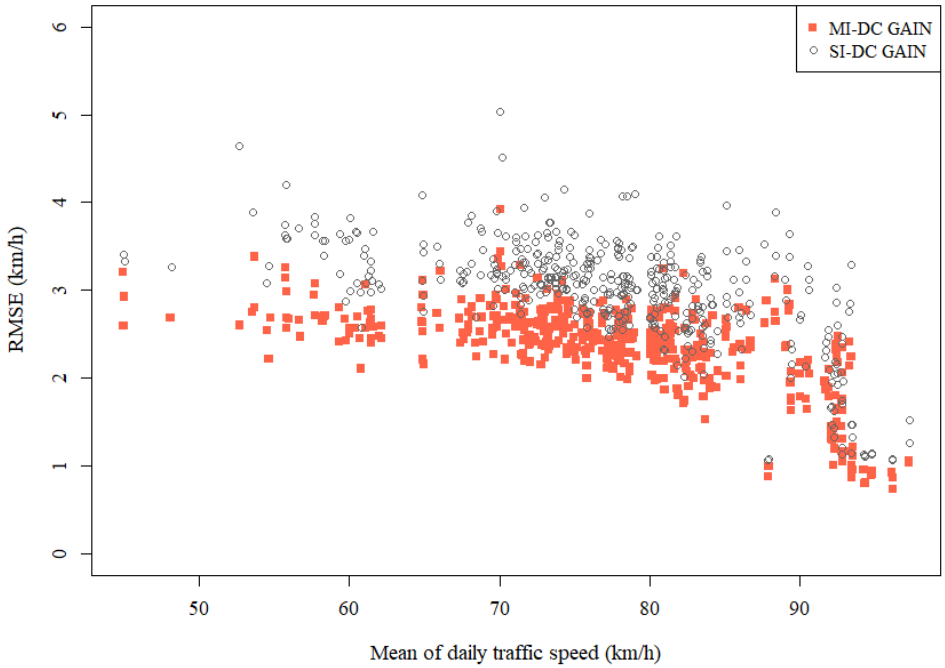


Figure 4.2 The imputation performance of models according to the mean of daily travel speed: MI-DC GAIN and SI-DC GAIN

4.2.2. Advantages of CNN (MI-DC GAIN vs. MI-FC GAIN)

According to Table 4.1 and Figure 4.1, GAIN with CNN structure showed much better imputation performance than that with FCN, and the performance difference was even more prominent when the level-of-congestion is higher. This result suggests that the CNN aid in identifying spatio-temporal dependency represented in traffic speed images, which is more pronounced in the congested traffic state. In other words, FCN failed to consider the locality of spatiotemporal data since it flattens an image to a one-dimensional array (Chen et al. 2017, Shamsolmoali et al. 2018, Shabbeer Basha et al. 2020). As RMSE of MI-FC GAIN is 4.66 times larger than that of MI-DC GAIN, it is apparent that considering the spatio-temporal correlation is very important in traffic information imputation.

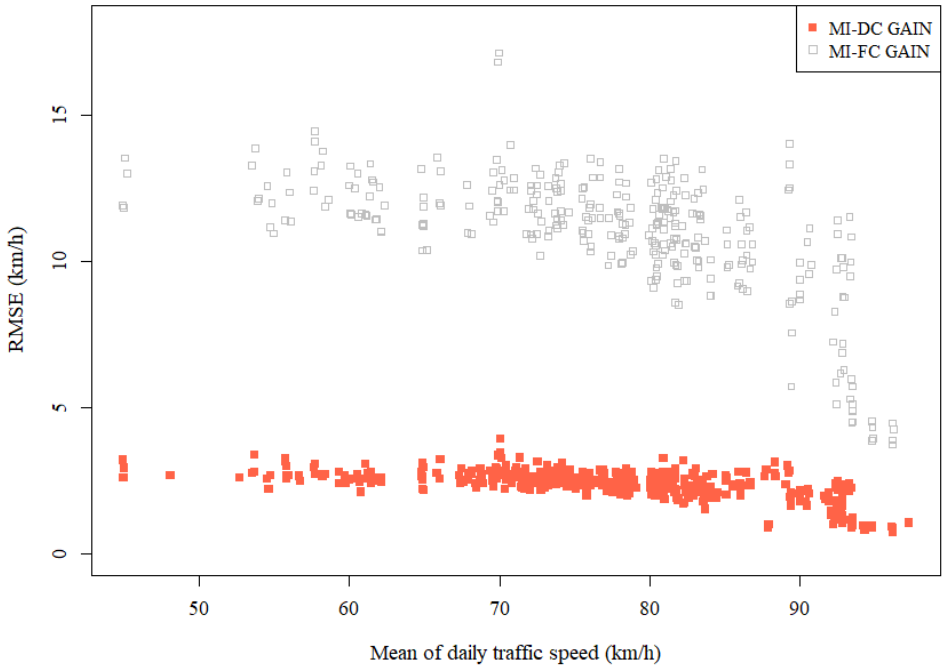


Figure 4.3 The imputation performance of models according to the mean of daily travel speed: MI-DC GAIN and MI-FC GAIN

4.2.3. Advantages of the Deep-learning Model (MI-DC GAIN vs. TASM and MA)

As mentioned in the Introduction, imputation methods can categorize into three different methods, i.e., the prediction, interpolation, and statistical learning methods. According to the results, the performance of MI-DC GAIN was better than the two baseline methods: TASM and MA. According to Figure 4.1, TASM and MA showed a high variation of RMSE by the mean of daily traffic speed. It may indicate that those imputation methods, which only consider general traffic patterns or temporal continuity, cannot respond well to highly congested traffic that has prominent uncertainty. Therefore, these two methods have less robustness on level-of-congestion.

MA is a simple interpolation method commonly used with time-series data to capture short-term transitions. It imputes missing data by calculating the average of temporal nearest values. MA is a very

simple method to impute missing values, but it is sometimes more effective for missing data imputation than complex methods when the temporal continuity is critical. The imputation results showed that MA was better than TASM in our case (i.e., imputing random continuous missings). Although MA performed quite well in general, the large variation depending on the test data suggests that it is sensitive to the traffic dynamics of each detector data and the distribution of random missings.

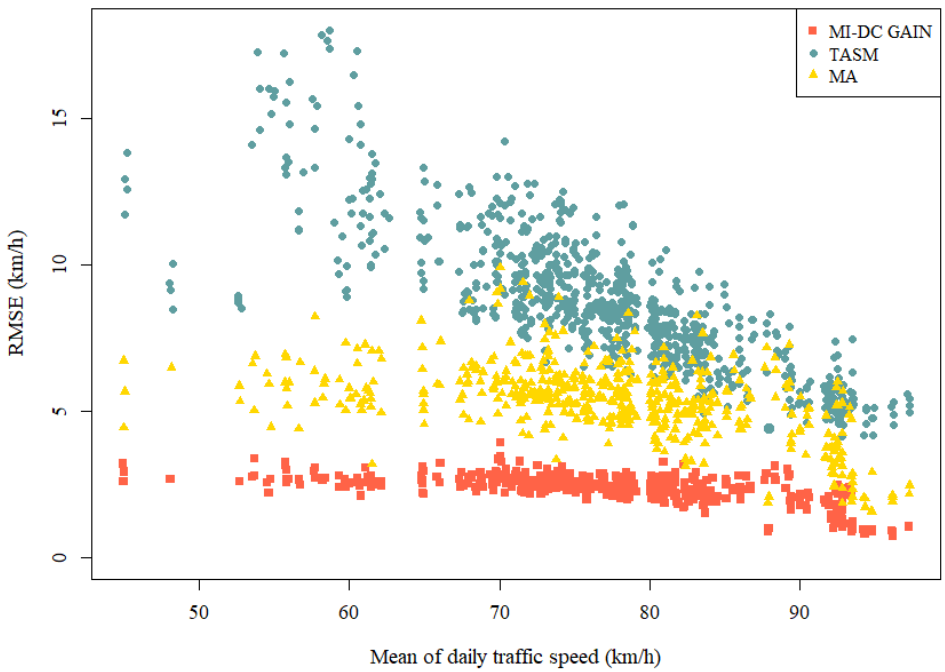


Figure 4.4 The imputation performance of models according to the mean of daily travel speed: MI-DC GAIN, TASM, and MA

TASM imputes traffic information considering spatio-temporal correlations. TASM is an effective imputation method in revealing propagation patterns, and its robustness has been proven irrespective of parameter changes in freeways (Treiber et al. 2011). However, depending on the traffic situation, the performance of this method has been found to be highly sensitive. As a result of imputing partially-missed loop detector data, TASM was found to be less accurate and

less robust than MI-DC GAIN. TASM uses the calibrated parameter that expresses the general propagation property c_{cong} and c_{free}. However, compared with freeways, traffic speed in urban expressway is more complicated due to intrinsic uncertainty associated with short-distance on-/off-ramps and frequent merging and diverging (*4*). Therefore, the TASM's parameters representing the propagation of traffic state would not be perfectly suitable for every traffic condition in the urban expressway (Chen et al. 2019).

The accuracy and robustness of TASM is disrupted in two situations, i.e., 1) when congestion occurs often and 2) when the transition occurs often. The results of high imputation errors imply that TASM could be a good reconstruction method but not that good imputation method. The more frequent the occurrence and dissipation of congestion, the greater the probability that the distributed traffic information will be uncertain, so there is a limit to generating missing traffic information using a parametric approach like TASM. Figure 4.5 shows a comparison of the traffic speed of ground-truth, TASM, and MI-DC GAIN. Several types of errors that are harder to be correctly imputed by TASM than by MI-DC GAIN are observed in this partial speed contour, as shown in Figure 4.5(c). The gray boxes on upstream show that TASM cannot capture the onset of traffic congestion, and the boxes on the midstream and downstream show that the failure to capture the dissipation of congestion. In addition, since the propagation property of traffic state set by TASM does not exactly fit those of the study site, the high speed of the downstream excessively influence the imputed speed of TASM in the midstream, causing higher imputed speed than the ground-truth speed. Likewise, the missing part of the downstream is excessively affected by the low speed of the midstream. Compared to TASM, MI-DC GAIN well performed for the missings occurred in transition and congested states as shown in Figure 4.5(d).

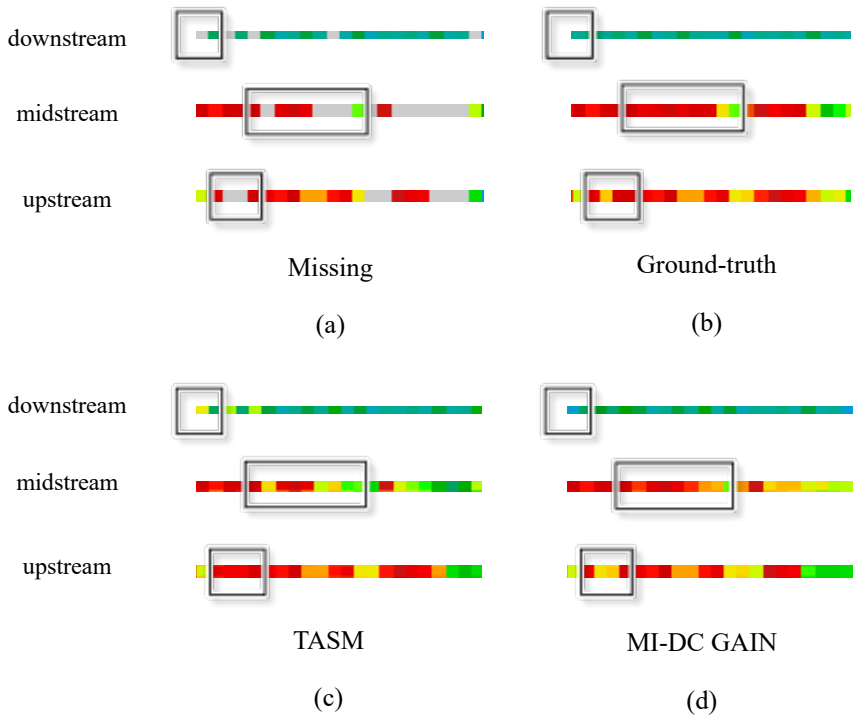


Figure 4.5 Comparison with an example of speed contour: (a) Missing; (b) Ground-truth; (c) TASM; (d) MI-DC GAIN

Chapter 5. Conclusion

In this study, we proposed an MI-DC GAIN for missing data imputation of traffic speed collected from the loop detector equipped in the Seoul outer circular expressway. To overcome the limitations reported in previous studies (i.e., prior assumptions for data distribution, low performance capturing spatio-temporal correlation, not well addressing transition of traffic state, and non-convergence of training), we used deep learning, convolutional neural network, generative modeling, and multi-input structure, respectively. Random missings, including spatial and temporal continuous missings, were generated from the data in various traffic conditions. By evaluating the performances, we found that the proposed MI-DC GAIN outperformed the benchmark models such as SI-DC GAIN, MI-FC GAIN, TASM, and

MA in terms of accuracy and robustness. The findings from the experiments are summarized as follows: 1) Learning traffic images by applying convolutional network to GAIN can improve the accuracy and robustness of missing data imputation even in congested and transition traffic states; 2) the multi-input structure with the reconstructed images can lead to further improvement of DC GAIN.

Despite the promising results obtained from this study, there are still several ways the study can be improved. First, considering that the missings of the loop detector and other sensors sometimes does not occur randomly, the performance for those non-random missings should be validated. Second, although we applied the multi-input structure to reconstructed images, it can be used for the other sources of data such as GPS trajectory data from probe car and travel time data from automatic vehicle identification. Since the proposed method uses the traffic speed images on the time-space diagram as input, it is easy to unify these multi-source data into a single image format. Third, predicting traffic speed on urban roads including signal intersections is a valuable task in traffic management, but challenging subject due to its uncertainty stemming from traffic signals, geometric conditions, and other exogenous factors (Kim et al. 2019a). Therefore, it is necessary to ensure that the outstanding performance of the MI-DC GAIN in urban freeway is also valid in those urban roads. Lastly, a breakthrough could result from high-resolution data since the spatial and temporal data aggregation blur the significant features of the traffic dynamics. The microscopic traffic data generated from connected vehicles in the future are promising candidates for those data.

Bibliography

Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein GAN. http://arxiv.org/abs/1701.07875

Asadi, R., & Regan, A. (2019). A convolution recurrent autoencoder for spatio-temporal missing data imputation. 1–8. http://arxiv.org/abs/1904.12413

Bae, B., Kim, H., Lim, H., Liu, Y., Han, L. D., & Freeze, P. B. (2018). Missing data imputation for traffic flow speed using spatio-temporal cokriging. Transportation Research Part C: Emerging Technologies, 88(February), 124–139. https://doi.org/10.1016/j.trc.2018.01.015

Basha, S. H. S., Dubey, S. R., Pulabaigari, V., & Mukherjee, S. (2020). Impact of fully connected layers on performance of convolutional neural networks for image classification. Neurocomputing, 378, 112–119. https://doi.org/10.1016/j.neucom.2019.10.008

Chen, X., Duan, Y., Houthooft, R., Schulman, J., Sutskever, I., & Abbeel, P. (2016). InfoGAN: Interpretable representation learning by information maximizing generative adversarial nets. Advances in Neural Information Processing Systems, 2180–2188.

Chen, X., Yang, X., Wang, M., & Zou, J. (2017). Convolution neural network for automatic facial expression recognition. Proceedings of the 2017 IEEE International Conference on Applied System Innovation: Applied System Innovation for Modern Technology, ICASI 2017, 814–817. https://doi.org/10.1109/ICASI.2017.7988558

Chen, X., Zhang, S., Li, L., & Li, L. (2019). Data for Traffic State Estimation and Prediction. 20(4), 1247–1258.

Duan, Y., Lv, Y., Liu, Y. L., & Wang, F. Y. (2016). An efficient realization of deep learning for traffic data imputation. Transportation Research Part C: Emerging Technologies, 72, 168–181. https://doi.org/10.1016/j.trc.2016.09.015

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 3(January), 2672–2680.

Huang, T., Chakraborty, P., & Sharma, A. (2020). Deep convolutional generative adversarial networks for traffic data imputation encoding time series as images. May. http://arxiv.org/abs/2005.04188

Jarrett, K., Kavukcuoglu, K., Ranzato, M., & LeCun, Y. (2009). What is the best multi-stage architecture for object recognition? Proceedings of the IEEE International Conference on Computer Vision, 2146–2153. https://doi.org/10.1109/ICCV.2009.5459469

Jo, D., Yu, B., Jeon, H., & Sohn, K. (2019). Image-to-image learning to predict traffic speeds by considering area-wide spatio-temporal dependencies. IEEE Transactions on Vehicular Technology, 68(2), 1188–1197. https://doi.org/10.1109/TVT.2018.2885366

Kessler, L., Huber, G., Kesting, A., & Bogenberger, K. (2018). Comparing Speed Data from Stationary Detectors Against Floating-Car Data*. IFAC-PapersOnLine, 51(9), 299–304. https://doi.org/10.1016/j.ifacol.2018.07.049

Kim, E. J., Park, H. C., Ham, S. W., Kho, S. Y., Kim, D. K., & Hassan, Y. (2019). Extracting Vehicle Trajectories Using Unmanned Aerial Vehicles in Congested Traffic Conditions. Journal of Advanced Transportation, 2019.

https://doi.org/10.1155/2019/9060797

Kim, E. J., Park, H. C., Kho, S. Y., & Kim, D. K. (2019). A Hybrid Approach Based on Variational Mode Decomposition for Analyzing and Predicting Urban Travel Speed. Journal of Advanced Transportation, 2019(December). https://doi.org/10.1155/2019/3958127

Kim, J., Mahmassani, H. S., & Dong, J. (2010). Likelihood and duration of flow breakdown: Modeling the effect of weather. Transportation Research Record, 2188, 19–28. https://doi.org/10.3141/2188-03

Kunfeng, W., Xuan, L., Lan, Y., & Fei-Yue, W. (2017). Generative Adversarial Networks for Parallel Vision. Proceedings - 2017 Chinese Automation Congress, CAC 2017, 2017-January(June 2018), 7670–7675. https://doi.org/10.1109/CAC.2017.8244166

Li, L., Li, Y., & Li, Z. (2013). Efficient missing data imputing for traffic flow by considering temporal and spatial dependence. Transportation Research Part C: Emerging Technologies, 34, 108–120. https://doi.org/10.1016/j.trc.2013.05.008

Li, Y., Li, Z., & Li, L. (2014). Missing traffic data: Comparison of imputation methods. IET Intelligent Transport Systems, 8(1), 51–57. https://doi.org/10.1049/iet-its.2013.0052

Liang, Y., Cui, Z., Tian, Y., Chen, H., & Wang, Y. (2018). A Deep Generative Adversarial Architecture for Network-Wide Spatial-Temporal Traffic-State Estimation. Transportation Research Record, 2672(45), 87–105. https://doi.org/10.1177/0361198118798737

Lin, Y., Dai, X., Li, L., & Wang, F. Y. (2019). Pattern Sensitive Prediction of Traffic Flow Based on Generative Adversarial Framework. IEEE Transactions on Intelligent Transportation

Systems, 20(6), 2395–2400. https://doi.org/10.1109/TITS.2018.2857224

Lorenz, M. R., & Elefteriadou, L. (2001). Defining freeway capacity as function of breakdown probability. Transportation Research Record, 1776, 43–51. https://doi.org/10.3141/1776-06

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 (Switzerland), 17(4), 1–16. https://doi.org/10.3390/s17040818

Myung, J., Kim, D. K., Kho, S. Y., & Park, C. H. (2011). Travel time prediction using k nearest neighbor method with combined data from vehicle detector system and automatic toll collection system. Transportation Research Record, 2256, 51–59. https://doi.org/10.3141/2256-07

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

Radford, A., Metz, L., & Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks. 4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings, 1–16.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training GANs. Advances in Neural Information Processing Systems, 2234–2242.

Shamsolmoali, P., Zareapoor, M., & Yang, J. (2019). Convolutional neural network in network (CNNiN): Hyperspectral image classification and dimensionality reduction. IET Image

Processing, 13(2), 246–253. https://doi.org/10.1049/iet-ipr.2017.1375

Treiber, M., Kesting, A., & Wilson, R. E. (2011). Reconstructing the Traffic State by Fusion of Heterogeneous Data. Computer-Aided Civil and Infrastructure Engineering, 26(6), 408–419. https://doi.org/10.1111/j.1467-8667.2010.00698.x

Wu, C. H., Ho, J. M., & Lee, D. T. (2004). Travel-time prediction with support vector regression. IEEE Transactions on Intelligent Transportation Systems, 5(4), 276–281. https://doi.org/10.1109/TITS.2004.837813

Yin, W., Murray-Tuite, P., & Rakha, H. (2012). Imputing erroneous data of single-station loop detectors for nonincident conditions: Comparison between temporal and spatial methods. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, 16(3), 159–176. https://doi.org/10.1080/15472450.2012.694788

Yoon, J., Jordon, J., & Van Der Schaar, M. (2018). GAIN: Missing data imputation using generative adversarial nets. 35th International Conference on Machine Learning, ICML 2018, 13, 9042–9051.

Zheng, Z., Ahn, S., & Monsere, C. M. (2010). Impact of traffic oscillations on freeway crash occurrences. Accident Analysis and Prevention, 42(2), 626–636. https://doi.org/10.1016/j.aap.2009.10.009

Abstract

신뢰성 있는 교통정보를 제공하기 위한 전제는 데이터의 완전무결성이다. 하지만 교통 정보 수집 시스템은 오류로부터 자유롭지 않기 때문에 결측이 필연적으로 발생한다. 최근에는 교통정보의 결측을 대체하기 위한 방법론으로 데이터의 내재적 특징과 상호작용을 포착할 수 있는 딥러닝 접근법이 활용되고 있다. 교통정보 결측의 핵심은 시공간적 상관성을 고려하는 것인데, 이는 통행속도를 시공도의 2차원 형태로 표현한 이미지를 활용함으로써 고려할 수 있다. 이 연구에서는 변형된 생성적 적대 신경망을 활용하여 지점검지기로부터 수집되는 도시부 고속도로의 통행속도 결측을 대체한다. 제안된 방법론은 신경망 구조에서 교통정보의 시공간적 패턴을 통행속도 이미지의 형태로 고려하기 위해 합성곱 신경망을 차용한다. 또한 학습을 용이하게 하기 위해 교통류 적합 스무딩 기법으로 교통정보를 재구축한 이미지를 멀티 입력자료로 활용한다. 합성곱 신경망 구조와 멀티 입력자료의 활용을 통해 생성적 적대 신경망의 통행속도 이미지 학습 성능을 향상시킬 수 있다.

주요어: 교통 데이터 결측대체, 생성적 적대 신경망, 합성곱신경망, 교통류적합 스무딩 기법, 시공간적 상관성
학번: 2019-21695