Thesis for the Degree of Master of Engineering

Automatic Incident Detection of the Metropolitan City by Adopting Deep Learning Algorithm of Spatial-Temporal Traffic Data

by

Yujin Jin

Department of Spatial Information Engineering

The Graduate School

Pukyong National University

February 2019

Automatic Incident Detection of the Metropolitan City by Adopting Deep Learning Algorithm of Spatial-Temporal Traffic Data
(시공간 교통데이터의 심화학습 알고리즘을 적용한 대도시 도심부 자동사고검지)

Advisor: Prof. Sang-Hoon Bae

by
Yujin Jin

A thesis submitted in partial fulfillment of the requirements
for the degree of
Master of Engineering
in Department of Spatial Information Engineering, The Graduate School,
Pukyong National University

February 2019

Automatic Incident Detection of the Metropolitan City by Adopting Deep Learning Algorithm of Spatial-Temporal Traffic Data

A dissertation

by

Yujin Jin

Approved by:

崔 哲 雄

(Chairman) Prof. Chul-Uong Choi

(Member) Prof. Han-Lim Lee

(Member) Prof. Sang-Hoon Bae

February, 2019

CONTENTS

LIST OF FIGURES

LIST OF TABLES

시공간 교통데이터의 심화학습 알고리즘을 적용한
대도시 도심부 자동사고검지

진 유 진

부 경 대 학 교 대 학 원 공 간 정 보 시 스 템 공 학 과

요 약

도심부 정체는 교통혼잡비용의 주요 원인이 된다. 특히 비반복 정체인 유고는 병목현상, 2 차사고 등을 유발시켜 막대한 사회경제적 손실을 초래할 수 있다. 따라서 빠르게 검지하고 대처하는 것이 중요하다. 기존 연구들에서는 연속류 유고검지가 주를 이루어왔으며 단속류 유고검지는 연속류 유고검지 알고리즘의 적용으로 인해 낮은 검지율과 높은 오검지율을 산출하였다. 또한 도심부 간선도로는 주·정차 및 신호등 등의 다양한 변수가 존재하기에 검지가 잘 이루어지지 않았다. 최근 다양하고 복잡한 문제들을 해결할 수 있는 인공지능 방법론이 주목받고 있다. 이는 기존에 해결하지 못한 한계점들을 해결해 다양한 분야에서 접목되고 있다. 이에 본 연구에서는 GPS 차량 궤적 데이터를 활용하여 심층신경망을 적용한 대도시 도심부 유고검지를 실시하고자 하였다. 먼저, 수집된 데이터들의 오류데이터 및 이상치를 제거해주었다. 이후 결측처리 및 데이터 평활화 작업을 거쳐 링크별/시간대별 패턴데이터를 구축하였다. 패턴데이터는 유고 시 발생하는 교통류의 변화 특성을 파악할 수 있다. 인공신경망 입력 변수를 선정하기 위해 유고 시에 나타나는 상류부, 중류부, 하류부의 교통흐름을 분석해보았다. 최종적으로 속도 변수, 프로브 차량대수 변수 등 총 8 가지로 선정되었다. 신경망 학습 전 최적 심층신경망 구조를 선정하였다. 3 가지의 조건을 변경해가며 구조를 채택하였다. 학습 후 TEST SET 을 통해 돌발상황 검지율, 오검지율 및 교통상황 검지율을 산출하였다. 이를 CONVENTIONAL ARTIFICIAL NEURAL NETWORK 및 200*200 DNN 과 비교·분석하였다. 비교 결과, 최적구조 선정 후 학습된 DNN 의 돌발상황 검지율이 79.5%로 가장 높게 나타났으며, 전체적인 교통상황 검지율이 가장 높은 경우는 200*200 으로 설정한 DNN 으로 나타났다.

1. Introduction

1.1. Background

Traffic congestion costs have been rising steadily since 2010, according to a report ‘A Study on the Estimation and Trend of Traffic Congestion Cost’ by the Korea Transportation Institute. Traffic congestion costs reached 33.3 trillion won in 2015. This increases 2.8% from 2014 to 2015, accounting for 2.2% of GDP.

The urban congestion cost is estimated to be about 1.7 times larger than those incurred on cross-regional roads. The main cause is congestion. Congestion is two part. The first is repeat congestion in predictable. The second is non-repeat congestion in unpredictable. Repeated congestion occurs periodically at peak times and in specific spaces, such as on the way to work. This is easy to cope with because it can be predicted in advance through past history data. On the other hand, non-repetitive congestion can cause secondary accidents as well as social and economic losses. Therefore, it is important to detect and respond quickly.

In general, incident detection that uses traffic data obtained from fixed and interval detectors. These detectors have limitation. First, existing detectors have many old ones. Second, traffic jams occur during installation and maintenance. Third, there are data that are not collected depending on the road surface conditions. Fourth, image detectors are a problem with one fixed index frame. Fifth, the detection rate is low and the false alarm rate is high.

Also, existing studies on congestion have been the main focus in the uninterrupted flow incident detection. In the uninterrupted flow, incident detection derived high detection rates through existing incident detection algorithms. In contrast to the uninterrupted flow, the interrupted flow has various situation variables such as shoulder parking and stop, traffic light, intersection, etc. In this situation, it is more difficult to detect the incident automatically.

Along with the fourth industrial revolution, big data and artificial intelligence (AI) methods that can solve various and complicated problems are receiving spotlight. This is drawing attention by solving problems such as nonlinear and complex formulas that were difficult to solve. AI methodology is applied to various fields to solve limitations in different fields. It is also used in various areas such as calculation of traffic information, adjustment of signal phase, and prediction of traffic demand. The results of the study also show high accuracy and reliability.

In this study, we use GPS probe vehicle data to automatically incident detection using deep neural network in the metropolitan city.

1.2. Study Area

The main road in downtown Seoul, where congestion occurs frequently, is selected as a spatial area. The area designated as a 7.6km-long section of Gangnam-daero, Seoul. The selected area is known to have the largest floating population and the largest daily traffic. It is also an important transportation route by serving as a link between downtown Seoul and the suburbs.

The collected data time is set from 1 January 2016 to 30 June 2016. Weekdays probe vehicle data is used except for weekends and holidays.

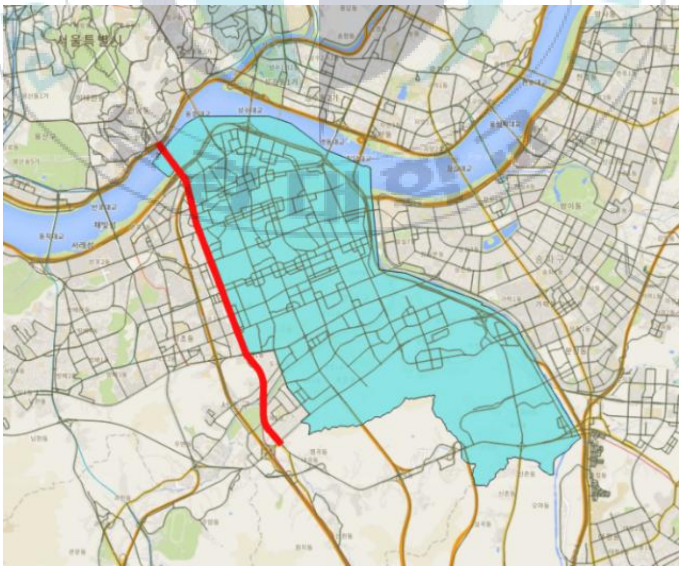

Fig. 1. The study area

1.3. Research Flow Chart

The flow chart of this study is shown in Figure 2.

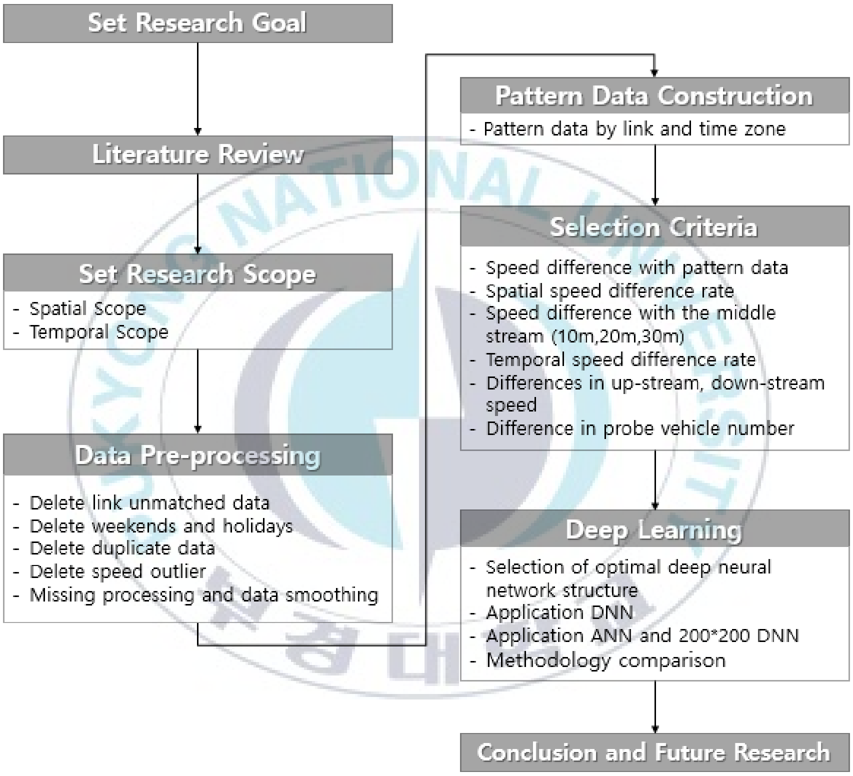

Fig. 2. The study flow chart

2. Literature Review

2.1. Existing Incident Detection Algorithm

There are 5 main types of algorithm for detection. They are distinguished by comparison or pattern recognition algorithms, statistical algorithms, time series and filtering algorithms, transport models and theoretical algorithms, and other algorithms, as shown in Table 1.

The comparison or pattern recognition algorithm determines the incident by comparing it against an established threshold. Statistical algorithms use statistical methods, such as standard deviations, to determine their contribution after comparing them with established thresholds. Time series and filtering algorithms determine the incident by time series analysis. Traffic flow model and theory algorithms are based on traffic theory using variables such as traffic volume and density. Other algorithms are characterized by determining inventory without setting a threshold value.

Table 1. Incident detection algorithm

Category	Algorithm		Occupancy	Volume	Speed
Comparison or Pattern Recognition Algorithm	California	Basic	●		
		Ver. #7	●		
		Ver. #8	●		
		APID	●	●	
	PATREG		●		
	Monica		●		
	Wave Analysis		●		
Statistical Algorithm	Bayesian		●		
	SND		●	●	
Time series and Filtering Algorithm	Time Series ARIMA		●	●	
	Exponential Smoothing		●		
	Low-pass Filter		●		
	Dutch		●		
Traffic Flow Model and Theory Algorithm	Dynamic		●	●	
	McMaster		●	●	
Other Algorithms	Neural Network		●	●	●
	Fuzzy Set		●	●	●

2.1.1. Comparison or pattern recognition algorithm

The California algorithm is the most well-known and features the simplest architecture. Using the shares of the up-stream and down-stream detector. Differences in absolute values of occupancy rates and differences in relative values are collected. This is then compared to the predefined threshold. If the threshold is exceeded, it is being determined as an incident. In addition, the algorithm has 10 variants, of which the 7th and 8th algorithms are known to perform well.

There is also an all purpose incident detection (APID) algorithm. This is a slightly modified version of the California algorithm. Both compression and persistence tests were performed on California fundamental variables. It is also characterized by the use of individual algorithms depending on the high and low traffic volume. The algorithms were divided into 3 stages and each stage of the algorithm determines their incident.

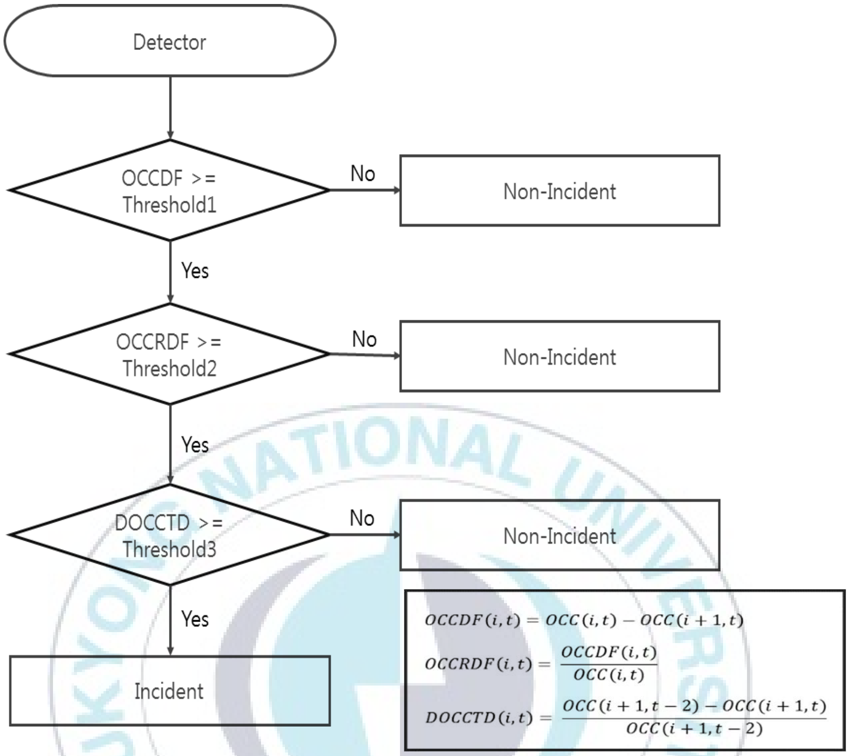

Fig. 3. All purpose incident detection algorithm

2.1.2. Statistical algorithm

Statistical methods determine the incident by comparing the observed and predicted values. Typical are Bayesian and standard normal deviation (SND) algorithms.

Bayesian algorithm applied Bayesian theory. Similar to the California algorithm used the difference in occupancy time between up-stream and down-stream detector.

The SND algorithm is based on the assumption that if an incident occurs, it will cause a sudden change in occupancy and traffic volume. The mean and SND are calculated every 3 or 5 minutes.

2.1.3. Time series and filtering algorithm

Time series and filtering algorithms compare errors between observed traffic conditions and predicted values by time series analysis. There is a typical auto-regressive integrated moving average (ARIMA), exponential smoothing, low-pass filter, dutch method.

ARIMA method is a type of time series analysis technique. Unlike the automatic moving average (ARMA) method, which uses historical data, it is a method that takes into account historical trends. To reflect short-term traffic conditions, observe recent changes in traffic parameters and set forecasts in advance. If the difference between the set prediction and the observed value is greater than the critical value, it is judged that the situation is incident.

Double exponential smoothing (DES) uses techniques to weight past and current occupancy observations to predict future traffic conditions. The value of the observed down-stream detector is compared with the calculated down-stream detector value to calculate the degree of imbalance to determine the incident situation.

2.1.4. Traffic flow model and theory algorithm

This method is to analyze incidents by setting up a function that can analyze traffic conditions. Typically, there are dynamic Algorithm, McMaster Algorithm.

Dynamic algorithm is based on the relationship between speed-density and traffic-density. It is characterized by using variables that appear in the uninterrupted flow.

McMaster Algorithm uses traffic theory. When traffic changes from normal to unusual, it is assumed that the speed shows the most sensitive change. Dividing the area into 4 areas according to the traffic volume-occupancy chart. Traffic flow areas are determined based on the collected data from each point detector. The 4 areas are divided by the boundary between congested and non-congested traffic, minimum traffic volume when be solved waiting queue traffic, and the critical occupancy rate. The algorithm takes 2 steps to determine an incident. The first is to determine the current traffic conditions by judging these 4 areas. Then, obtain the down-stream point detector data. Through this process, it can identify causes of identity.

2.2. Related Research

Yong-Kul Ki et al. (2018) proposed an algorithm to detect traffic accidents using TPCS' traffic situation information. It had set up 5 areas in Seoul as a spatial area and had used data for 2 months. Using artificial neural network (ANN) methodology, 29 out of 40 accidents were detected accurately and the detection rate was 72.5%.

Ho-Yong Kim et al. (2017) conducted a study to predict traffic accident points through an ANN. The research spatial area was set in Daegu City. Accident data, the day of the week, weather etc. were collected. In this study, it predicted the accident point through the feed-forward neural network (FFNN). The results were then compared with the Support Vector Machine (SVM), Random Forest (RF). The comparison showed the accuracies of SVM model, RF model, and FFNN model were 0.7898, 0.8227, and 0.871 respectively.

Hong-Suk Yi et al. (2016) studied the prediction of traffic congestion in the urban center. A typical auto-encoder methodology was used as a method of unsupervised. To accelerate the system, the simulation was also performed by an Intel Zion Pie processor. The result was about 50% of the theoretical performance value, but it was considered satisfactory given that the existing performance was 10%.

Unlike previous studies, Ji-yeon Hong et al. (2016) developed a predictive model for regional traffic accidents by dividing them into smaller units. Data such as the National Police Agency, Seoul Statistics, Seoul GIS Portal, and the road name management system were used. The data were collected in units of the Seoul Metropolitan Government. It was carried out by applying a multi-linear regression methodology. This allowed us to find variables related to accident-positive or accident-negative relationships.

Tae-Uk Kim et al. (2014) utilized GPS - based vehicle trajectory data after establishing an interrupted flow road in the urban center as a spatial range. Incident detection was performed by entering speed differences and geometry variables through a multilayer neural network. Verification results showed that the first day selected as a test set has an incident detection rate of 46.15%. On the second day, the detection rate for incident situations was 38.46 %.

Baozhen Yao et al. (2014) optimized the SVM parameters via tabu search algorithm to detect accidents on highways and use them in the incident. The weather, time zone, interval, and upper/lower traffic were entered as variables. Two highways in China were tested and the proposed SVM showed a lower prediction error than the existing ANNs.

Faisal Ahmed et al. (2012) intended to detect the incident status by time interval of the incident generated from the signal network link. To this end, the city center's incident detection model was developed based on critical values. Through NETSIM simulation, a virtual signal intersection was created. The regression formula was derived by setting the number of lanes, road length, traffic volume, and incident time differently. Verification results showed that the average high-altitude detection rate was 51% and 42% respectively for each lane.

Huying WEN et al. (2011) identified road-induced traffic accidents by introducing a learning vector quantification (LVQ) neural network model. A total of 9 variables were selected for input variables: speed, occupancy, and time required by up-stream, middle-stream, and down-stream. This was simulated using the VISSIM program. According to the model, LVQ was a detection rate of 96.67% and a false alarm rate of 3.45%.

Dae-Hyon Kim (2005) presented a real-time incident detection algorithm using SVM. The study used data collected on the Seoul inner highway. SVM and Backpropagation (BP) methodologies were applied. The comparison and review of the 2 methodologies showed that the SVM model was superior.

3. Data and Methodology

3.1. Data

3.1.1. National standard node/link

The government is establishing a standard transport network for smooth exchange of traffic information. National transport networks are identified by ID, which consists of nodes and links. The point at which the vehicle's speed varies rapidly is referred to as a node. The road in reality is called a link, linking nodes to nodes. Link information includes link ID, start-node ID, end-node ID, lane number, road grade, and road type. Node information includes node type, intersection name, and rotation limit. The link to Gangnam-daero that we are going to address in this study is made up of 31 one-way links.

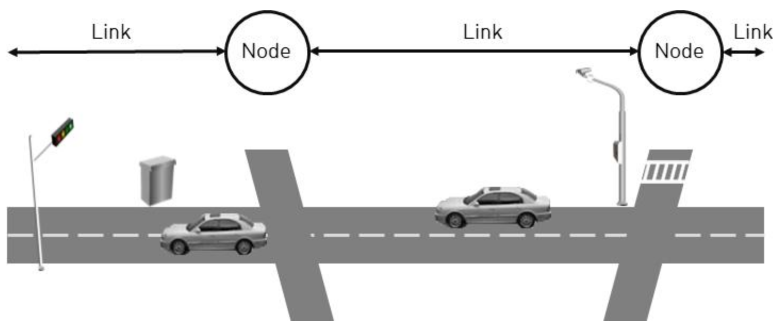

Fig. 4. National standard link and node concept

3.1.2. Smart traffic (SMT) data

SMT data means the vehicle trajectory data obtained from the probe vehicle as an abbreviation for 'SMart Traffic'. The probe vehicle which is based on GPS sends traffic information to related organizations through GPS receivers and maps that are installed in vehicles. The institution processes these data to generate segment communication information.

SMT data has started and arrived node and time. In addition, the link length allows the calculation of cross sectional speed. In this study, actual vehicle trajectory data were obtained Gangnam-daero for 6 months in 2016.

3.1.3. Incident data

In this study, construction/accident data is provided by the Transport's Traffic Information Disclosure Service of the Ministry of Land, Infrastructure and Transport. Date of generation, accident ID, time of occurrence, status, coordinates, etc. can be obtained from the data. It can be downloaded every hour or every day when utilized. Although construction/accident data have causes of construction, accidents, controls, events, and weather, only accident data have been considered. The sample valid column format for the data is shown in Table 2.

Table 2. Sample incident data format

Occurrence Time	Classification Name	Message	code X	code Y
2016.02.15. 07:55	Accident	passenger cars and buses	127.0187	37.51955
2016.02.22. 16:23	Accident	passenger cars and van	127.0175	37.5219
2016.03.29. 11:50	Accident	passenger cars and passenger cars	127.0316	37.48972

3.2. Methodology

3.2.1. An overview of artificial neural network

Artificial intelligent is the way computers and systems to think and act like people. The methodology, modeled after neurons in the human brain, was named after the university of Dartmouth in 1956. Many studies focused on exploratory reasoning, natural language processing, and simple artificial situations were conducted. However, problems such as the XOR problem and the computer's processing speed problems have led to the dark. After the first dark ages, the form of an artificial intelligence program was called the 'Expert System' which emerged in the 1980s and focused on it. Due to below-expected performance, it was in its second darkest hour. Since then, the IBM supercomputer has developed into the current AI field, winning the world chess championship. In the AI field, the 1980s started machines learning and deep learning in the 2010s.

Machine learning is a field of artificial intelligence that means that computers or systems can derive new outputs through learning based on input data or perform them for purpose. For example, suppose you give a computer a huge amount of data. Then computers and systems solve problems as if they classifying or solving math problems. This is copied from the artificial neural network structure as if the human brain is thinking and solving problems. Computers also solve problems by reducing errors through repeated learning.

The types of machine learning are largely divided into supervised learning, unsupervised learning, and reinforcement learning, such as clustering, regression, classification, etc. This is often combined in a variety of fields to show high accuracy. However, machine learning has problems such as vanishing gradient, local minima converge, overfitting, and slow learning speed that are suitable for training data.

Deep Learning is an advanced form of machine learning that allows a computer or system with broader and deeper structure. It also solves various limitations of machine learning and is being applied in many areas. These are distinguished by algorithms such as DNN, CNN, RNN, and DBN, depending on purpose of use and type of data.

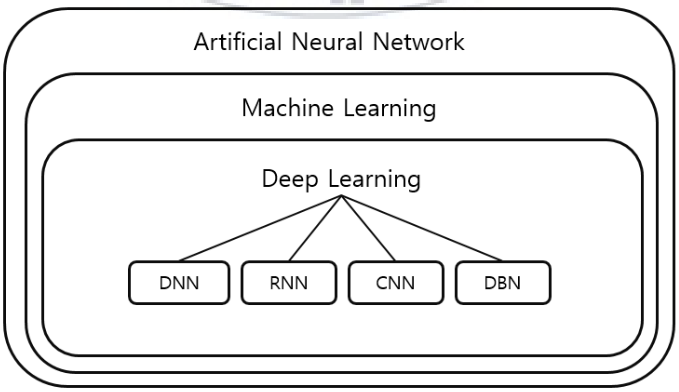

Fig. 5. Artificial intelligence

3.2.2. Deep neural network

Deep neural network (DNN) is defined as a combination of machine learning algorithms consisting of multiple layers of processing and abstracting them through a combination of nonlinear functions. DNN consists of 2 or more hidden layers, unlike ANN in terms of structure. It is getting attention by solving the limitations of machine learning that have not been solved in terms of function.

The vanishing gradient problem is solved by applying various activation functions such as Relu and Leaky Relu, which are limited to existing Sigmoid. Second, the converging phenomenon to the local minima is optimized through adjustment of parameters such as learning rate and momentum. This adjustment also helps speed up the search for the global minima. Third, overfitting's problems are solved through the application of regularization and the increase in the number of learning data. Finally, slow learning speed is solved by loading GPU. These problems are currently being applied in a variety of fields, and the strength of this DNN methodology is used in this study to detect the metropolitan city incident.

4. Selection of Deep Neural Network Variables

4.1. Data Pre-processing

Pattern data are developed to analyze traffic characteristics. For this purpose, the data are pre-processing.

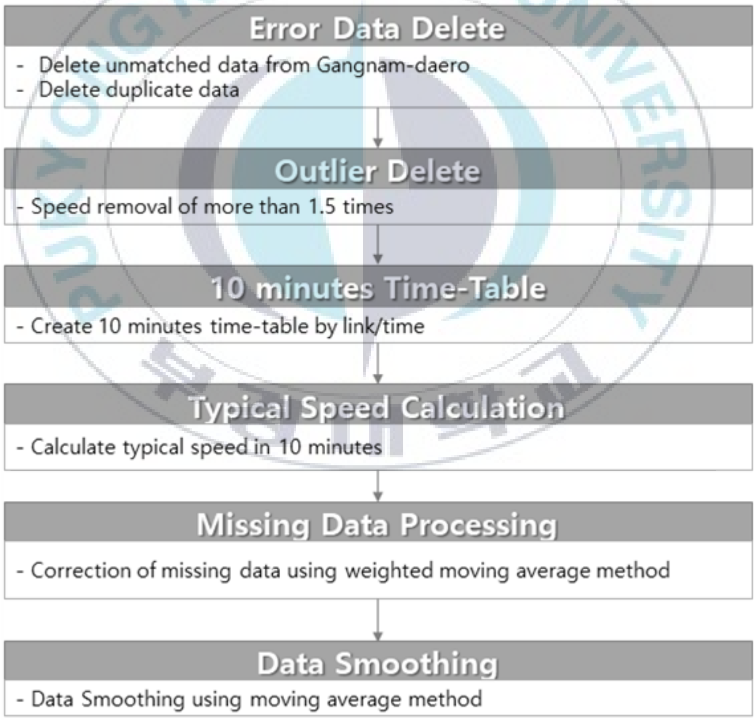

Fig. 6. Pre-processing flow chart

4.1.1. Delete error data

First, error data was removed after collecting raw data. Second, it was deleted links to Gangnam-daero and unmatched link. Third, duplicate links were deleted.

4.1.2. Delete outlier

The actual speed limit for the spatial range set in this study is 60 km/h. A few overspeeding vehicles have a large impact on average speed calculations. Generally, the upper and lower limits of the collected data are adjusted to determine the normal range of data. Therefore, a speed greater than 1.5 times the speed limit, 90 km/h, was considered to be an outlier value and deleted.

In addition, low-speed vehicle has a significant impact on average speed calculation. However, the lower limit was not determined in this study because it may be the effect of an incident detection. Therefore, only the upper limit was set and deleted.

4.1.3. 10 minutes time-table

After the pre-processing, the time-table was written every 10 minutes. It was divided by link/time and consists of the following format, as shown in Table 3. The time interval was set to 10 minutes to be more reliable and representative of the probe vehicle. The 10 minutes time-table was written in 2 directions. Individual probe vehicle data was entered by link/time.

Table 3. 10 minutes time-table

	1	2	3	4	5
1210007702	52*12	31*12	37*12	38*12	35*12
1210007703	87*12	52*12	59*12	59*12	59*12
1210008500	80*12	42*12	62*12	56*12	60*12
1210009500	89*12	47*12	66*12	58*12	71*12
1210010300	90*12	41*12	71*12	48*12	82*12
1210011300	78*12	66*12	49*12	66*12	63*12
1210013701	101*12	81*12	56*12	57*12	89*12
1210013702	103*12	79*12	55*12	64*12	91*12
1210013703	113*12	76*12	59*12	70*12	89*12
1220011902	70*12	109*12	61*12	63*12	91*12

4.1.4. Calculation of typical speed

The typical speed by link/time was calculated through the configured time-table. The 10 minutes typical speed was shown as equation (1). The 10 minutes probe vehicle's individual interval traffic speed was added and divided by the number of probe vehicles.

$$V_i = \frac{\sum_{j=1}^{N} v_i}{N} \tag{1}$$

4.1.5. Missing data processing

In some cases, it is not easy to collect information about operating and collecting transportation systems. In this case, the missing section will occur. Missing data processing consists largely of spatial and temporal trending methods. It using spatial trending method when a connection is made due to the absence or failure of the detector. The temporal trending method is used when the application of spatial trending methods is not possible or the time flow is interrupted.

Methods for correcting missing data include moving average or pattern data. It can estimate missing data by averaging the data from the previous date or using historical data of the past. In this study, the weighted average of movements was applied to treat the missing data. This is a way to give more weight to the near past in time.

$$V_t = 0.5V_{t-1} + 0.3V_{t-2} + 0.2V_{t-3} \tag{2}$$

4.1.6. Data smoothing

Data smoothing is carried out to compensate for the instability of traffic data obtained from the metropolitan city. This can obtain reliable data. The moving average method was used in this study which is a method that is calculated from the sliding window of the length k with the data to be smoothed centered.

$$V_{smooth} = \frac{v_{t-3} + v_{t-2} + v_{t-1} + v + v_{t+1} + v_{t+2} + v_{t+3}}{k} \quad (3)$$

4.2. Pattern Data Construction

In this study, pattern data of upper/lower links to Gangnam-daero were constructed by using 6 months probe vehicle data. Pattern data was constructed from 67,386 probe vehicles collected at Gangnam-daero from January to June 2016. In addition, the number of raw data is 2,956,967. The number of valid data after error removal data is 1,778,202. Since then, the number of data has 1,655,916. Pattern data was constructed through data after pre-processing. The constructed pattern data can be used to identify areas where congestion occurs. It is also possible to identify changes in traffic flow that occur in the event of an incident.

4.3. Traffic Flow Change Analysis

Based on pattern data, it was examining the traffic characteristics of up-stream, middle-stream, and down-stream that occur in case of an incident.

The first characteristic is that the speed of the middle stream decreases sharply. Fig. 6 shows an accident analysis graph that occurred in front of Nonhyeon station in Gangnam-daero at 14:53 on February 18, 2016. The x-axis of the graph is the time index, and the Y-axis of the graph represents the speed of the middle-stream. Generally, it can be seen that the speed value is between 30 and 45 km/h. However, the speed at the time of the accident was sharply reduced to 27 km/h.

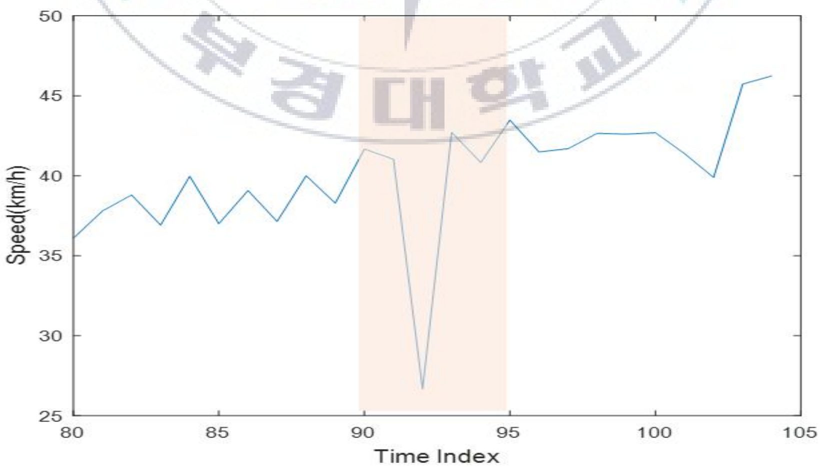

Fig. 7. Middle-stream speed

The second characteristic is that the up-stream speed is similar to middle-stream speed or the up-stream speed is more decrease.

The third characteristic is that the speed of the down-stream increase because passing cars are blocked by the flow of the middle-stream. Below Fig. 7 shows an accident that occurred on 27 April 2016. This graph shows the characteristics of the up-stream, middle-stream, and down-stream. The up-stream, middle-stream, and down-stream parts are both figures showing the difference between the existing pattern data and the data at the time of the accident. The larger the value of the negative, the greater the speed at the moment of the accident, which can be identified by the characteristics of the lower part.

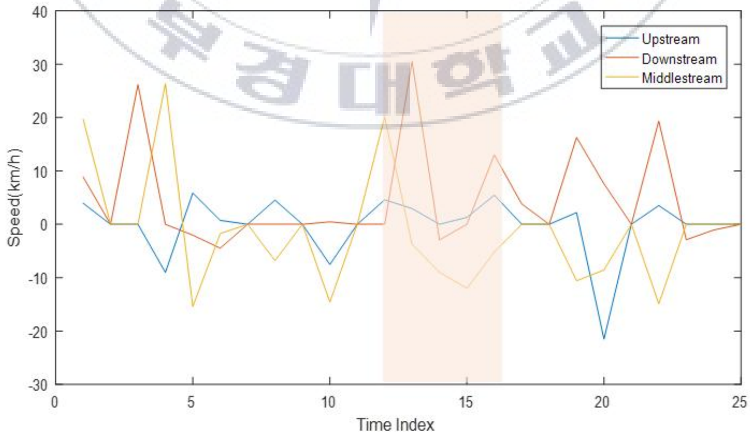

Fig. 8. Difference of speed with pattern on up-stream, down-stream, middle-stream

The fourth characteristic is the difference in the number of probe vehicles. The number of up-stream probe vehicles increases, and the number of down-stream probe vehicles decreases. Like speed, the down-stream probe counts are smaller due to vehicles that fail to descend from the up-stream.

4.4. Variable Selection

In the literature review, previous papers selected speed, the geometric structure of roads, and weather as additional variables. In this study, it has been shown that due to the nature of the selected spatial range, the incident detection will be possible only through changes in speed. Therefore, the difference in speed between up-stream links, down-stream links, and middle-stream links was selected as variables. In addition to speed, the difference in the number of probe vehicles with significant changes was also chosen as a variable. The variables selected are shown in Table 4.

Table 4. Variables for deep neural network

Definition	Calculation
Difference of Observed speed data and Pattern data	$\mathrm{SPD(i,t)} - PATTERN(i,t)$
Fraction of variation of average speed depending on Space	$\frac{SPD(i+1,t) - SPD(i,t)}{SPD(i+1,t)}$
Fraction of variation of average speed depending on Time	$\frac{SPD(i,t-1) - SPD(i,t)}{SPD(i,t-1)}$
Difference of Speed (before 10minute)	$SPD(i,t) - SPD(i,t-1)$
Difference of Speed (before 20minute)	$SPD(i,t) - SPD(i,t-2)$
Difference of Speed (before 30minute)	$SPD(i,t) - SPD(i,t-3)$
Difference value between errors of Up-stream and Down-stream compared to pattern data	$\big(SPD(i-1,t) - PTTN(i-1,t)\big)$ $-(SPD(i+1,t) - PTTN(i+1,t))$
Difference of number of Probe vehicle	$Count(i-1,t) - Count(i+1,t)$

5. Application Deep Neural Network

5.1. Application Deep Neural Network

Deep neural network was applied through the previously selected variables. The architectural features of DNN methodology are the multilayer of the hidden layer. Prior to applying DNN learning, the optimal deep neural network structure was explored.

The optimal deep neural network structure depends on the researcher's experience and outlook. Therefore, the experiment was chosen under various conditions in this study. The test conditions are shown in Table 5. The scenario was adjusted from 1 to 5, hidden layer from 1 to 5, and neurons of hidden layer from 1 to 20.

Table 5. Deep neural network 3 condition

Scenario	1~5
Hidden Layer	1~5
Neuron of Hidden Layer	1~20

In the preceding study, DNN was applied to all of the structures below 5 hidden layer. Therefore, the author chooses from 1 to 5 floor based on the preceding study.

Because of the relatively small amount of data to be used in this study, 5 scenarios were constructed to complement it. Scenario organization is displayed in Fig. 9. For scenario 1, 80% of the data was set to train data from the beginning and 20% was set to test data. Starting with scenario 2, based on the location of the test data in scenario 1, the train data and test data were set gradually.

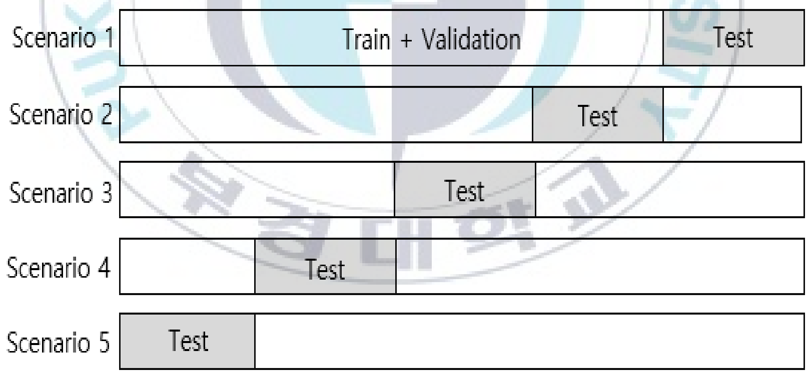

Fig. 9. Deep neural network 5 scenarios

In order to select the optimal deep neural network structure, the data was divided into training, validation, and test. For the optimal structure, the structure with the lowest value of validation error was selected. The result of the selection was that the hidden layer has 5 and the neuron number of hidden layer has 9.

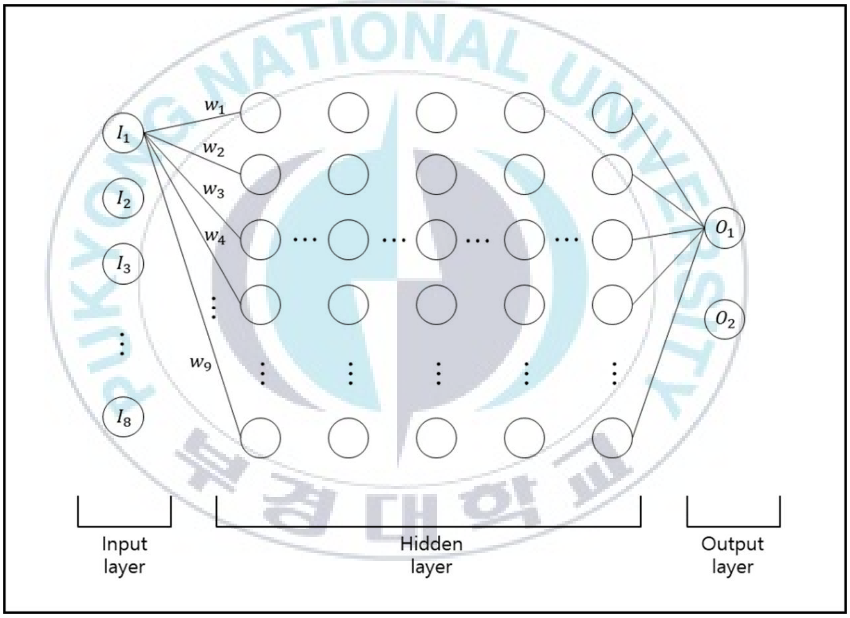

Fig. 10. Optimal deep neural network

Learning was conducted through the best structure. Scaled conjugate gradient (SCG) was selected as learning algorithm when applying DNN. SCG uses step size scaling mechanism to save time when learning. It is known to be faster than commonly used BP.

Relu activation function was used as a transfer function between the hidden layers. Relu function is converged with zero for negative numbers, but the original value is maintained for positive numbers. This is a common practice these days because it solves the vanishing gradient, a problem with the sigmoid function used in existing ANN. The softmax function was used as the output function of the output layer. The softmax function is characterized by returning all output values between 0 and 1. It is also a function that makes the sum of the output values equal to 1. It can determine which output is more likely.

In addition, regularization was performed to prevent and optimize overfitting, and the option to early stop was set. The early stop option is validation check. Validation errors increase when the network begins to continuously match data. This can cause overfitting. To avoid this, if validation errors continue to increase more than a specific number of times, the training is stopped, and the minimum value of the inspection error is found and the weight and bias are set to be returned.

5.2. Incident Detection Algorithm Performance Evaluation Based Deep Learning

As a measure of performance evaluation of automatic incident detection algorithms, the detection rate, false alarm rate, and traffic condition detection rate are typically used. The detection rate is the case where the deep neural network is also judged to incident in the actual incident conditions. False alarm rate is the rate of non-incident detection in the situation where the deep neural network determines that it is an incident. The traffic condition detection rate means that the incident is detected as incident and non-incident as non-incident. The cases are shown in Table 6.

Table 6. 4 Cases of result

Target \ Predicted	Incident	Non-Incident
Incident	A	B
Non-Incident	C	D

$$Detection\ \ Rate\ \ (DR) = \frac{A}{A+B} * 100 \tag{4}$$

$$False\ \ Alarm\ \ Rate\ \ (FAR) = \frac{C}{A+C} * 100 \tag{5}$$

$$Traffic\ \ Condition\ \ Detection\ \ Rate\ \ (TDR) = \frac{A+D}{A+B+C+D} * 100 \tag{6}$$

The results of analysis by scenario learned from the previously selected optimal deep neural network structure are as follows.

Table 7. DR, FAR, and TDR per scenario

	DR(%)	FAR(%)	TDR(%)
Scenario 1	75.0	50.0	51.5
Scenario 2	88.2	25.0	78.8
Scenario 3	81.3	48.0	54.5
Scenario 4	88.2	50.0	48.5
Scenario 5	64.7	45.0	54.5
Average	79.5	43.6	57.6

The study found that the detection rate for incident is 79.5%, the false alarm rate is 43.6, and traffic condition detection rate is 57.6%. The highest detection scenario was scenario 2, and the lowest detection rate was scenario 5.

It compared and analyzed the training data and test data of these 2 data. As a result of the analysis, scenario 2 showed a pattern similar to the difference of observed speed data and pattern data, fraction of variation of average speed depending on space, difference of speed before 20 minutes, difference of speed before 30 minutes in the training set. However, in the case of scenario 5, there were many variables that showed a pattern similar to the train set's non-incident situation.

5.3. Comparison and Verification

In this study, the conventional artificial neural network and 200*200 deep neural network were applied for comparative verification by DNN. ANN also choose the optimal structure. Scenarios have 1 to 5, and neurons have 1 to 20 in the hidden layer. For the optimal structure, the number of neurons in the hidden layer was 7. 200*200 DNN pre-selected a fixed structure. The author was set 200 neurons in hidden layer 1 and 200 neurons in hidden layer 2. The structure was established wider than the optimal neural network structure selected in this study. Learning was conducted after the adoption of the 2 structures. In both cases, the learning algorithm was equal with DNN.

Table 8. Artificial neural network 2 condition

Scenario	1~5
Neuron of Hidden Layer	1~20

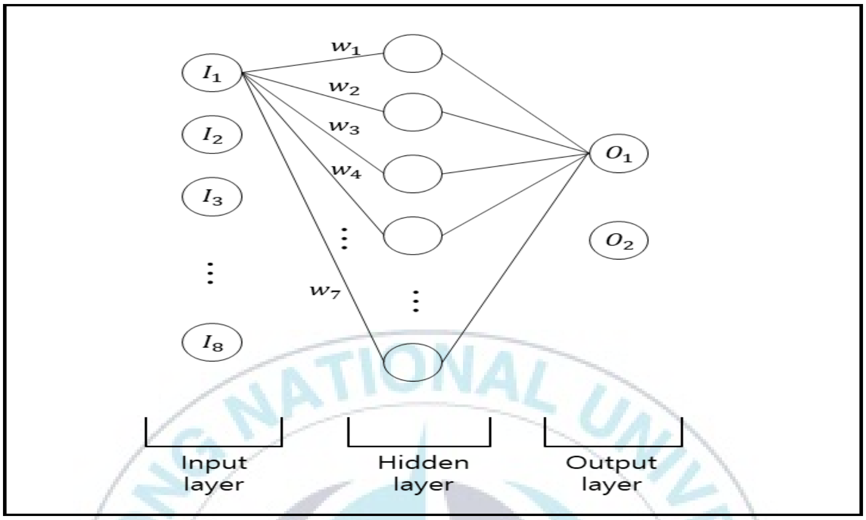

Fig. 11. Optimal artificial neural network

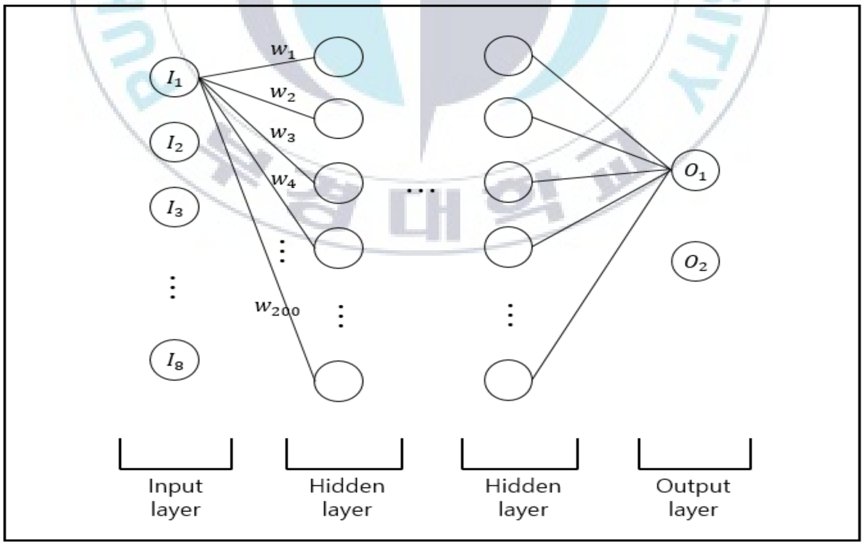

Fig. 12. 200*200 neural network

According to the incident detection by the conventional neural network, the detection rate was 63.0%, the false alarm rate was 39.3%, and the traffic condition detection rate was 58.8%. According to the 200*200 DNN, the detection rate of incident situations was 67.4%, and the false alarm rate was 39.7%. The traffic condition detection rate was 61.6%. DNN’s final comparison results are shown in Table 9.

Table 9. DNN, conventional ANN, 200*200 DNN analysis

	DR(%)	FAR(%)	TDR(%)
DNN	79.5	43.6	57.6
Conventional ANN	63.0	39.3	58.8
200*200 DNN	67.4	39.7	61.6

Comparisons show that when DNN was applied after selecting the optimal structure, it has the highest detection rate at 79.5%. The traffic condition detection rate was the highest when the structure was widely adopted. The false alarm rates were all similar. In order to lower the false alarm rate, the incident detection rate is also lowered. In this study, the author selects the one with high detection rate because we aimed to incident detection. As a result, the false alarm rate also increased. Both DNNs had higher detection rates than ANN, which is deemed to be due to DNN's resolution of the ANN's limitation.

6. Conclusions

In this study, DNN was applied to automatically incident detection on the main road in the metropolitan city. It was set up as Gangnam-daero, which have a large number of daily traffic and large floating population. The study used the 6 months period of GPS probe vehicle trajectory data from January 2016 to June 2016. Pattern data by links on weekdays were deployed except for weekends and holidays. The time was set at a 10 minutes interval for more reliable results. DNN's input data were selected through recognition of changes in traffic flow at the points where traffic accidents occur at the up-stream, middle-stream, and down-stream. Finally, it was selected as 8 variables.

Before learning about DNN, we explored the optimal deep neural network structure. We compared the validation error values with 5 scenarios, 1 to 5 hidden layers, and 1 to 20 neuron of hidden layers. The last selected structure was the number of hidden layers on the 5 hidden layers, and the number of neurons on the hidden layer is 9. Training was carried out in this structure. SCG was applied as a learning algorithm and Relu was applied as activation function. The average of the 5 scenarios detection rate was 79.5%, false alarm rate was 43.6%, and traffic condition detection rate was 57.6%. Comparing scenario 2 with scenario 5, scenario

2 was higher. In the case of scenario 2, there were many variables that show a pattern similar to the training set, but in the case of scenario 5, fewer variables were analyzed that show a pattern similar to the training set. DNN was compared to conventional artificial neural network and 200*200 DNN. According to the comparison results, DNN that learns after selecting the optimal structure and find that the detection rate was the highest. However, the traffic condition detection rate was 200*200, the highest for a widely chosen structure. The false alarm rate was found to be a small difference. The main implications of this study are as follows.

First, this study also allows automatic incident detection on the main road in the metropolitan city. This is confirmed to have an accuracy of about 79.5%. Given that the main traffic stream of studies related to the existing incident detection is uninterrupted flow, this study has a high index.

Second, the results are realistic by obtaining data processing and analysis using GPS vehicle data that drive on-road. Unlike previous studies that analyzed detection rates using virtual result values obtained through simulation, it is a meaningful result.

Due to the nature of the incident, there is a duration between the formation of a queue and its disappearance. In the future, it is deemed possible to study not only the determination of the incident but also the determination of the duration by reflecting it. Finally, as in this study, it is necessary to detect automatic incident not only on roads with many probe vehicles but also on areas with low traffic. To do this, it is necessary to carry out the research on automatic detection by linking with other detection systems as well as probe vehicles.

7. References

Ki, Y. K., Heo, N. W., Choi, J. W., Ahn, G. H., & Park, K. S., (2018), "An incident detection algorithm using artificial neural networks and traffic information", In Cybernetics&Informatics(K&I), IEEE, pp.1-5.

Lee, Y. J., (2018), "Development of an Estimation Model of Vehicle Queue Length Using Link Travel Time based on Deep Learning", Ajou University PhD Thesis.

Park, S. M., Lee, H. P., Kwon, C. W., Park, S. H., & Yun, I. S., (2018), "Study on the Development Methodology for a Crash Prediction Model in Expressway Using a Deep Neural Network (Case of the Major Expressway), The Journal of The Korea Institute of Intelligent Transport Systems, vol.17, no.4, pp.324-326.

Seo, S. H., (2018), "A Study on Weights Quantization for Improving Efficiency of Deep Neural Networks", Dongguk University Master's Thesis.

Kim, H. Y., Kim, J. J., Park, S. M., Jeon, C. N., & On, B. W., (2017), "Artificial Neural Network Model for Prediction of Traffic Accident Point", Proceedings of the Korea Intelligent Information Systems Society, vol.8, pp.33-34.

Park, H. S., & Ju, K. B, (2017), "Development direction of ITS multi function detector for climate change", The conference journal of the Korea Institute of Intelligent Transport System, pp.222-224

Lee, H. S., & Jung, H. J., (2016), "Performance of Deep Learning for Traffic Flow Prediction Using Supercomputer", Korea Information Science Society 2016 Korea Computer Engineering Conference.

Lee, J. S., (2016), "Development process and understanding of deep neural network", Journal of the Korean Institute of Communication Sciences (information and communication), vol.33, no.10, pp.40-48.

Lee, K. S., & Shin, C. H., (2016), "Development of an AIDA(Automatic Incident Detection Algorithm) for Uninterrupted Flow Based on the Concept of Short-term Displaced Flow", The Journal of The Korea Institute of Intelligent Transport Systems, vol.15, no.2, pp.13-23.

Korean Statistical Information Service, (2015), National traffic congestion cost calculation and trend analysis, http://kostat.go.kr

LeCun, Y., Bengio, Y., & Hinton, G.,(2015) "Deep learning", Nature, vol.521, no.7553, pp.436-444.

Hong, J. Y., Lee, S. B., & Kim, J. H., (2014), "Development of Traffic Accident Frequency Prediction Model by Administrative zone : A Case of Seoul", The 40th Annual Meeting of the Korean Society of Civil Engineers, pp.673-674.

Korea Road Traffic Authority Traffic Science Institute, (2014), A Study on Improvement of Traffic Information Quality in Central Traffic Information Center.

Kim, T. U., & Bae, S. H., (2014), “Incident Detection of Urban Arterial Road using by Artificial Neural Network”, Pukyong National University Master’s Thesis.

Yao, B., Hu, P., Zhang, M., & Jin, M., (2014), “A support vector machine with the tabu search algorithm for freeway incident detection”, International Journal of Applied Mathematics and Computer Science, vol. 24, no. 2, pp. 397-404

Ahmed, F., & Yaser, E. H., (2012), “A threshold-based real-time incident detection system for urban traffic networks”, Procedia-Social and Behavioral Sciences, vol.48, pp.1713-1722.

Wen, H., & Jun, L., (2011), “Traffic incident detection for urban arterial road based on data fusion and learning vector quantization”, ICTIS 2011:Multimodal Approach to Sustained Transportation System Development: Information, Technology, Implementation, pp.971-978.

Journal of the Korea Institute of Intelligent Transport System, (2008), “Traffic Information Engineering”, Gyomoon Publishers.

Kim, D. H., (2005), “Automatic Incident Detection Using Machine Learning”, The study of Seoul City, vol.6, no.1, pp.71-80.

Kim, J. T., & Choi, J. T., (2004), "Design and Implementation for Incident Detection Algorithm in Intelligent Transportation System", Journal of the Korean Society for Industrial and Applied Mathematics IT Series, vol.8, no.1, pp.79-90.

Møller, M. F., (1993), "A scaled conjugate gradient algorithm for fast supervised learning". Neural networks, vol. 15, no. 4.

Wikipedia, Google, "http://wikipedia.org/wiki/"