

Article

For Preventative Automated Driving System (PADS): Traffic Accident Context Analysis Based on Deep Neural Networks

Minhee Kang ¹, Jaein Song ² and Keeyeon Hwang ^{3,*}

¹ Department of Smartcity, Hongik University, Seoul 04066, Korea; speakingbee@hanmail.net

² Department of Urban Planning, Hongik University, Seoul 04066, Korea; wodlsthd@nate.com

³ Department of Urban Design & Planning, Hongik University, Seoul 04066, Korea

* Correspondence: keeyeonhwang@gmail.com; Tel.: +82-010-8654-7415

Received: 5 October 2020; Accepted: 30 October 2020; Published: 2 November 2020



Abstract: Automated Vehicles (AVs) are under development to reduce traffic accidents to a great extent. Therefore, safety will play a pivotal role to determine their social acceptability. Despite the fast development of AVs technologies, related accidents can occur even in an ideal environment. Therefore, measures to prevent traffic accidents in advance are essential. This study implemented a traffic accident context analysis based on the Deep Neural Network (DNNs) technique to design a Preventive Automated Driving System (PADS). The DNN-based analysis reveals that when a traffic accident occurs, the offender's injury can be predicted with 85% accuracy and the victim's case with 67%. In addition, to find out factors that decide the degree of injury to the offender and victim, a random forest analysis was implemented. The vehicle type and speed were identified as the most important factors to decide the degree of injury of the offender, while the importance for the victim is ordered by speed, time of day, vehicle type, and day of the week. The PADS proposed in this study is expected not only to contribute to improve the safety of AVs, but to prevent accidents in advance.

Keywords: preventive automated driving system; automated vehicle; traffic accidents; deep neural networks

1. Introduction

Amid an active discussion of the Fourth Industrial Revolution, Automated Vehicles (AVs) are expected to play an important role in leading the Fourth Industrial Revolution. AVs are defined as vehicles capable of navigating, controlling, and avoiding risk partly or totally without human assistance [1]. According to the Society of Automotive Engineering [2], AVs can be categorized into six levels, ranging from none auto-system (SAE level 0) to full auto-system (SAE level 5). Human intervention is minimized from SAE level 3 and driverless driving is possible at level 5. With such features as driving safety improvement, increase in convenience and mobility [3,4], AVs are highly evaluated as key future mobility of reducing traffic accidents. The benefits mentioned above will be accomplished when AVs fully take root. However, some researches have indicated that the public still expresses a low level of acceptance for AVs [5–7]. It is mainly attributed to AVs traffic accidents arisen during the test driving by Google, Uber, etc. In particular, a fatal pedestrian accident involving Uber has been at the forefront of ethical controversy over AVs. Neither of the types of traditional ethics (deontology, utilitarianism) fit in well to provide a proper answer to this accident, nor the trolley dilemma excuse is unsuitable [8–10]. In response, the AVs guidelines, including provisions of preventive design and safety, were issued in Germany and the United States [11,12]. Specifically, German AV ethics guidelines state that “Automated and connected technology should prevent accidents wherever this is practically possible” in its fifth clause. In addition, various studies emphasized that

trust in AVs is the most important determinant to accept AVs for their mobility, and that the trust is decided by perceived safety risk, compatibility, and system quality [13–15].

To respond to the prevention of potential AVs related accidents, this study proposes Preventive Automated Driving System (PADS) of using Deep Neural Networks (DNNs)-based traffic accident context analysis. The study conducts experiments to identify key features affecting traffic accidents caused by unpredictable conditions such as black ice, sink-hall, centerline crossing, and so on.

The paper is organized as follows: Section 2 reviews studies on the effects of AVs and various traffic accident cases, and examines deep learning applications used in transportation research. Section 3 introduces methods to collect and process traffic accident data. In Section 4, we introduce how to build an optimal DNNs algorithm for forecasting the severity of accident injuries and extract factors causing accidents. Section 5 validates the important factors extracted in Section 4 by using a random forest-based machine-learning algorithm. Finally, Section 6 concludes the paper with a summary of empirical findings and derives future researches and implications related to preventive AVs.

2. Literature Review

2.1. AVs Introduction Impacts

It is important to estimate AV's impacts on traffic because it is inevitable that AVs will be mixed with human driving vehicles (HVs) in road traffic. With AV's Market Penetration Rate (MPR) growing, the positive effects can become bigger in congested conditions [16]. Most studies have been conducted quantitatively using a microscopic simulation software, VISSIM: A Leksandra Deluka Tibljaš et al. [17] designed a rotary interchange and evaluated safety in a mixed traffic condition. Yan Wang et al. [18] confirmed that AVs introduction not only kept Level of Service (LOS) higher, but also improved safety at signaled intersections. However, when MPR reached over 50%, a negative impact began to appear with a growing traffic delay [19]. Lee et al. [20] also identified that the maneuvering of AVs should be properly controlled by various traffic and road conditions because the driving behavior of HVs is affected by the aggressiveness of AVs. Some studies have mentioned that AVs contribute to decrease traffic accidents (collision) and delays [21]. Kolarova et al. [22] conducted an online survey for analyzing the potential changes in the Value of Travel Time Savings (VTTS). It is shown for commuting trips that AVs reduce 41% of VTTS on average compared to HVs. For leisure or shopping trips, no significant changes in the VTTS were found. Tscharktschiew & Evangelinos [23] investigated the impact of the transition in automated driving capabilities (driving mode choice) on road congestion pricing and vice versa, accounting for the interdependencies between traffic flow, the chosen level of autonomous driving, effective road capacity and marginal travel cost. The result suggested that when inconveniences related to autonomous driving are sufficiently high, the imposition of congestion tolls may lead to a situation where drivers abandon autonomous technologies entirely and opt instead for fully manual driving, not the generally expected positive effects.

2.2. AVs Cognition Survey

Along with AVs impact studies, several AVs-related cognition surveys have been conducted. Most cognition studies used surveys, confirming that preference towards AVs was higher for men than for women [24] and higher for the younger generation than for the older [25]. According to Nordhoff et al. [26], most people think that impacts of AVs appear to be positive, but responded that AVs safety benefit needs to be experimentally verified [27]. Moreover, Im et al. [28] analyzed web articles and comments about AVs using text mining techniques, showing that the number of reports that include AVs as a keyword increased, and there are more negative views than positive views. Specifically, the articles and comments with negative views discussed AVs ethics, traffic accidents, and the problem of sudden unintended acceleration.

2.3. AVs Traffic Accidents and Derived AVs Ethics

We reviewed AVs traffic accidents researches and AVs ethics issues for building the design of preventive AVs. Hong et al. [29] & Yang et al. [30] classified types of AVs accident as follows: the negligence of the drivers, accidents due to mechanical defects, malfunction of S/W, and accidents caused by information error, hacking, weather, etc. In the case of the steadily rising trolley dilemma problem [31], doubts were brought up as to whether the trolley dilemma could apply to AVs. They suggested that deriving accident algorithms to respond to it could be misleading [32]. Bae & Lee [33] approached decision-making criteria for the protection priority in accident situations. They proposed two solutions: to enforce them by law and to leave them to the AVs Artificial Intelligence (AI) system itself. In regards to setting them on AVs AI system, Gogoll & Müller [7] discussed differences between Mandatory Ethics Setting (MES), considering society as a whole and Personal Ethics Setting (PES) considering individual interests. For the accident in an ideal condition, Goodall [34] proposed ethical collision algorithms, and Fleetwood [35] discussed Germany's AVs ethics guidelines and main ethical topics.

2.4. Deep Learning Application in Transportation Area

There exist many transportation prediction studies using deep learning technologies. These can be bisected into two categories: prediction of either traffic flow or traffic accidents. First, studies of forecasting traffic flow compare proposed model performance to traditional classical algorithms [36–39]. Particularly, the Long-Short Term Memory (LSTM) model is compared with Statistically Adjusted Engineering (SAE), Radial Basis Function (RBF), Support Vector Machine (SVM), and Auto Regressive Integrated Moving Average (ARIMA) model, and DNNs is compared with Random Forest, a kind of machine learning model. Traffic accident prediction studies were carried out using Social Network Service data (Twitter) using Deep Belief Network (DBN) & LSTM [40]. For real-time accident detection, Chen et al. [41] analyzed the accident impact using GPS-based vehicle data. They used Stack Denoise Autoencoder (SDA), which is more effective to detect accident risk than the traditional models.

Furthermore, as rapid progress is made nowadays in AVs technology, backed by advances in the areas of deep learning and AI, various studies about AVs using AI have been implemented. Especially as AVs requires an accurate perception of surrounding environments to operate reliably, most studies are related to Convolutional Neural Networks (CNNs). That is why object detection is a fundamental function of AVs systems, including camera sensor (2D), Lidar (3D), radar, GPS, etc. Among them, 3D object detection for AVs studies have been carried out recently, and a new methodology (combining or extension) has been proposed. For instance, Li, P., et al. [42] proposed the Stereo R-CNN for 3D object detection as Faster Regions with CNN (Faster R-CNN) for stereo inputs extending to detect objects simultaneously with images on the left and right. The experiments on the challenging KITTI dataset show that their method outperforms the state-of-the-art stereo-based method (Stereo R-CNN) around 30% AP on both 3D detection and 3D localization tasks. Also, Chen, S., et al. [43] developed a CNN–LSTM based on prior knowledge and temporal information for AVs driving. The proposed algorithm was found to be approximately 85% accurate in mimicking human drivers. Zeng, W., et al. [44] proposed the Deep Structured Self-Driving Networks (DSDNet), which performs object detection, motion prediction, and motion planning with a single neural network. The algorithm showed that it has outperformed the state-of-the-art method (DSDNet) on several challenging datasets in general. Existing AVs datasets are limited in the scale and variation of the environments they capture, even though generalization within and between operating regions is crucial to the overall viability of the technology. In response, Sun, P., et al. [45] conducted the study based on the vast amount of real data it owned as a leading group in the study of AVs. This study is based on actual AVs data, and presents a large-scale multimodal camera-LiDAR dataset that is significantly larger, higher quality, and more geographically diverse than any other existing similar dataset. In addition, Djuric, N., et al. [46] introduced a deep learning-based approach that takes into account a current world state and produces raster images of each actor's vicinity for presenting an effective solution to a critical part of the AVs problem. The method first rasterizes actor contexts,

followed by training CNNs to use the resulting raster images to predict the actor's short-term trajectory and the corresponding uncertainty. Also, they tested the framework (system), which was deployed to a fleet of AVs.

2.5. Summary

In summary, diverse studies had been conducted about the traffic impacts of AVs. Most studies verified that when AVs are introduced, LOS, speed, and road volume (capacity) are improved and traffic accidents decreased even in a condition of mixed traffic with HVs. Also, cognition studies on AVs social acceptability were carried out. In addition, AI is actively being applied and developed in various ways for AVs. There have been various studies such as AVs control study based on CNNs, End-to-End, traffic volume forecasting based on DNN-Based Traffic Flow prediction (DNN-BTF), and short traffic flow prediction using LSTM.

Although studies about AVs introduction and acceptance surveys have been actively delved into, it is found that there exist limitations in building the foundation for commercialization. In addition, researches on legal and ethical issues on AVs traffic accidents have been carried out from various perspectives, but opinions have been continuously raised whether it is appropriate to answer all the ethical issues related with AVs traffic accidents as pre-conditions of the commercialization. Similarly, most AVs technology researches have gradually advanced to study related to the technology supplementation of themselves, but the researches related with the prevention of AVs traffic accidents have been found insufficient. Therefore, as proposed in the German and U.S. ethics guidelines and NVIDIA report [47], this study aims to lay the foundation for the design of the PADS using AI as a way to prevent AVs traffic accidents in advance.

3. Methodology

DNNs we intend to use in this study are a kind of Artificial Neural Network, consisting of the input layer, the output layer, and the hidden layers in between. DNNs are capable of modeling complex non-linear relationships, such as common artificial neural networks, with the ability to express basic elements in hierarchical configurations and the added layers to converge the characteristics of lower layers. In addition, regardless of continuous or categorical variables, non-linear combinations between input variables are easy to analyze, and automatic feature extraction reduces the hassle of variable selection. These features are used in the study to extract ambient situation information factors, such as weather information, external factors, etc. [39].

Traffic accidents are issues directly related to human life, and it is believed that legal and ethical problems will be inevitable in the event of an accident by learning and predicting inaccurate content. Thus, the German Ethics Guidelines for AVs stated that they should be designed to prevent accidents in advance and that they allow the use of AI technology to improve safety. Therefore, it is considered that the top priority is to learn how to prevent accidents by recognizing accident situations in advance as a solution to traffic accidents of AVs, and in this study, DNN, which has higher predictability than conventional machine learning algorithms, is to be used.

3.1. Data Collecting & Pre-Process

The data used for this study was from the accident data of Seoul city collected from 2017 to 2018. The main dataset is Traffic Accident Analysis System (TAAS) [48] data provided by The Road Traffic Authority (KoROAD). The information on traffic accident conditions in TAAS can be obtained in the form of Excel data on the TAAS website. However, it was necessary to extract the location information (coordinates) of an accident to identify the traffic situation information such as several lanes, speed, etc. during the traffic accident in detail. So, we crawled location coordinates data from TAAS to merge link attributes with traffic accident condition data. The crawled data include location coordinates, accident number, date of the accident, day of the accident, the content of accident, the number of deaths, the number of severe injuries, the number of light injuries, the number of wounded, accident type,

violation of the law, road conditions, weather, road type, offender information, and victim information (see Table 1).

Table 1. Crawled Traffic Accident Analysis System (TAAS) data.

Class	Feature	Class	Feature	Class	Feature
A_NO	Accident Assignment Number	WO_NO	The number of Wounded	O_sex	Offender Sex
Date	Date of Accident(YY/MM/DD)	A_type	Accident Type	O_age	Offender Age
Day of the week	Day of Accident	V_Law	Violation of Law	O_injury	Offender injury degree
A_fact	Content of Accident	R_con	Road Conditions	V_car	Victim Car
D_NO	The number of Death	W_con	Weather	V_sex	Victim Sex
SI_NO	The number of Severe injuries	R_type	Road Type	V_age	Victim Age
LI_NO	The number of light injuries	O_car	Offender Car	V_injury	Victim injury degree

However, TAAS has a limitation in that it does not provide data on traffic environment such as the number of lanes and speed, etc. Accordingly, we used Transport Operation & Information Service (TOPIS) [49] from the Seoul city and Korea Transport DataBase (KTDB) [50] provided by MOLIT to get the traffic environment data of each node & link. The link speed data were drawn from TOPIS (See Table 2), and the data on the number of lanes were extracted from KTDB (node-link data). TOPIS and KTDB data were obtained and utilized in Excel form from the above-mentioned sites.

Table 2. Sample of TOPIS speed data.

Date	Day of the Week	Link ID	Road Type	Time		
				1	2	3
20180601	Fri.	1080012200	Minor arterial road	44.6	27.88	51.79
20180601	Fri.	1080012800	Minor arterial road	17.73	24.32	24.67
20180601	Fri.	1080012700	Minor arterial road	23.15	29.9	32.56
20180601	Fri.	1080012100	Minor arterial road	47.88	40.4	44.29
20180601	Fri.	1230024700	Other road	21.62	21.17	27.13
20180601	Fri.	1230019500	Other road	30.59	28.86	30.7

TOPIS provides link speed data only on minor arterial roads, but not on collector roads. Thus, we utilized KTDB to get the data for collector roads. The data on speed and number of lanes were merged based on the TAAS coordinate system, and we used the 1-h data for the learning.

After the collection, we refined the data for the analysis. We excluded X, Y coordinates, local area name, and accident number from the dataset because they are only useful for merging purposes. Also, the data on the number of casualties such as the number of deaths, severe injuries, etc. were excluded because they are deemed unsuitable for this study which aims to prevent accidents in advance. When checking the basic statistics of the data, it was found that the number of accidents was appeared constant in monthly and daily bases. Also, the seasonality that we wanted to check in monthly accidents seems to be well reflected in the “weather” factor; likewise for the “day” factor reflected by the “day of the week” factor. Therefore, we only used the time and the day of the week data in the analysis.

Since TOPIS data do not include speed data under the minor arterial road level, 10 to 20 km/h, which is the average speed in Seoul, was allocated for the empty data cells. Finally, a total of 77,000 pre-processed data were used, with 38,625 cases in 2017 and 38,796 cases in 2018.

3.2. Learning Process

The purpose of this study is to prevent AVs-related accidents in advance by learning from the accident analysis of HVs. The dependent variables adopted for offender and victim identically consist of 5 different degrees of injuries such as death, serious injury, injury, minor injury, and no injury. Traffic accidents do not always consist of two parties, the offender and the victim. Sometimes there exists only one party case. For instance, when only vehicle damage is occurred with no human casualties (only one party), there exist only offenders. These cases also are incorporated into the learning process. The independent variables accounting for the seriousness of the injuries consist of road conditions, weather, road shape, number of lanes, and link speed (See Table 3). The learning was implemented by adding new factors such as time (T), day of the week (D), vehicle type of offender & victim (C), and violation of the law (L) in a step-by-step manner.

Table 3. Basic Data Variable Setting for Learning.

Feature	Class	Categorical Variable
Offender injury degree	Output Layer	1~6
Victim injury degree	Output Layer	1~8
Road Condition (RC)	Input Layer	1~5
Weather (W)	Input Layer	1~5
Road shape (S)	Input Layer	1~4
Number of Lane (L)	Input Layer	0~5
Link Speed (Sp)	Input Layer	1~6

4. Design of Optimal Deep Neural Networks (DNNs)

This chapter proposes an optimal DNNs for more accurate traffic accident prediction. The process of optimizing the model consists of three steps: setting up the data range—epoch—and hidden layers.

4.1. Environment on Building Algorithm

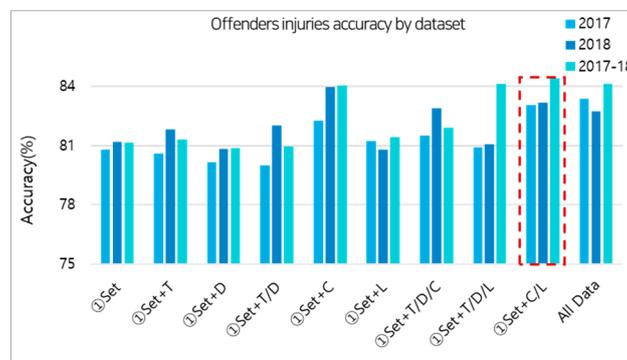
This study intends to produce an optimal model of accident prediction based on the backpropagation algorithm [51] and SGD (Stochastic Gradient Descent) optimizer of the Tensorflow and Keras libraries [52,53]. The ratio of training and testing data was set at 8:2. It used the dropout of randomly skipping a certain amount between nodes in weight update for minimizing an overfitting. We experimentally confirmed that the highest prediction was achieved where the dropout value was set at 0.2. Also, we used ReLU (Rectified Linear Unit) function [54] to prevent gradient vanishing when using the Sigmoid function [55]. The batch size was set to 64 for stable learning.

4.2. Learning Data Range for Building optimal DNNs

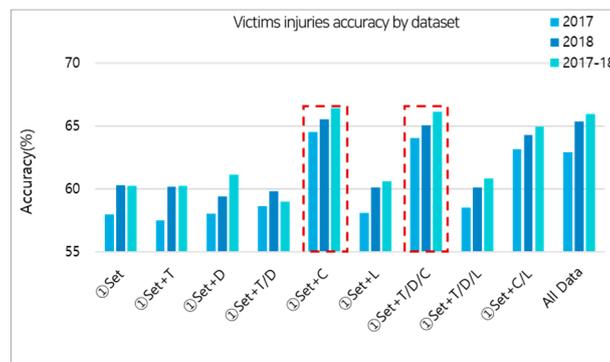
First, we performed DNNs using 2017, 2018, and 2017–2018 integrated data. For extracting optimal data range suited to the model, we set up the default model (RC, W, S, L, Sp; ① set, which are only external factors) and simulated the degree of injury to the offender and victim as output. The prediction results show that the predicted accuracy of the degree of injury to the offender is over 80%. Also, it appeared that the 2018 data's prediction accuracy was higher by about 0.1–2.0% than the 2017 data, and the 2017–2018 integrated data was more accurate by about 0.5–1.5% than the 2018 data. The degree of injury to the victim showed about 60–65% accuracy, and prediction based on the integrated data was found more accurate by 0.5–2.5% than 2017, 2018 data. Therefore, we decide to implement the simulation by using the 2017–2018 integrated data set (See Table 4 & Figure 1).

Table 4. Offender/Victim Injuries Accuracy by datasets.

Feature	DATA					
	Offender			Victim		
	2017	2018	2017–2018	2017	2018	2017–2018
①Set	80.80	81.20	81.16	57.99	59.89	61.05
①Set + T	80.60	81.80	81.32	57.22	59.41	61.21
①Set + D	80.16	80.82	80.85	58.79	59.51	60.75
①Set + T + D	79.98	82.00	80.93	58.55	59.77	60.92
①Set + C	82.26	83.96	84.02	64.12	65.36	66.46
①Set + L	81.22	80.77	81.41	58.86	59.41	60.67
①Set + T + D + C	81.52	82.88	81.89	64.16	65.12	66.43
①Set + T + D + L	80.89	81.08	84.12	58.11	60.26	62.75
①Set + C + L	83.04	83.15	84.38	64.07	65.01	65.83
All data	83.37	82.74	84.15	62.92	65.56	65.98



(a) Offender injuries accuracy



(b) Victim injuries accuracy

Figure 1. (a) Offender Injuries Accuracy by dataset; and (b) Victim Injuries Accuracy by dataset: The red dash line rectangular stands for the highest accuracy at that factors in 2017–2018

4.3. Learning Data Epoch for Building Optimal DNNs

Second, we performed DNNs by varying epoch, a state where one learning is completed for the entire data set within the network. When the epoch is set up on a large scale, training can cause overfitting, resulting in less accuracy in testing, verification, and application of new data. So, the step equally simulated one step (data range) process for extracting optimal data epoch. We set up epoch in five divisions of 100 units and simulated the degree of injury to the offender and victim as output. It is confirmed that the highest accuracy was achieved in both offender and victim at 100 epochs. The result

suggests that the higher epoch was set up, the lower accuracy was shown due to overfitting. Therefore, we will simulate the study by setting 2017–2018 integrated data and 100 epochs (See Table 5).

Table 5. Offender/Victim Injuries Accuracy by Epoch.

Feature	Epoch									
	Offender					Victim				
	100	200	300	400	500	100	200	300	400	500
①Set	81.16	81.02	80.89	81.13	81.17	61.05	60.55	60.87	90.13	60.75
①Set + T	81.32	81.20	81.22	80.98	81.25	61.21	61.13	60.75	60.81	60.92
①Set + D	80.85	80.80	80.90	80.01	80.57	60.75	60.43	61.21	60.84	61.33
①Set + T + D	80.93	80.77	80.89	80.65	80.91	60.92	60.88	60.13	60.74	60.27
①Set + C	84.02	83.70	83.52	83.78	83.97	66.60	65.43	64.79	66.60	65.98
①Set + L	81.41	81.50	81.13	80.87	80.15	60.67	59.70	60.12	60.45	60.33
①Set + T + D + C	81.89	81.73	81.51	81.27	81.55	66.43	65.47	66.11	65.87	65.90
①Set + T + D + L	84.12	83.96	83.87	83.17	84.03	62.75	61.72	62.70	62.13	62.60
①Set + C + L	84.38	84.27	84.29	84.02	84.30	65.83	64.98	65.74	65.10	65.33
All data	84.15	84.05	83.79	83.98	84.03	65.98	64.80	64.93	65.77	65.68

4.4. Setting the Hidden Layers Building Optimal DNNs

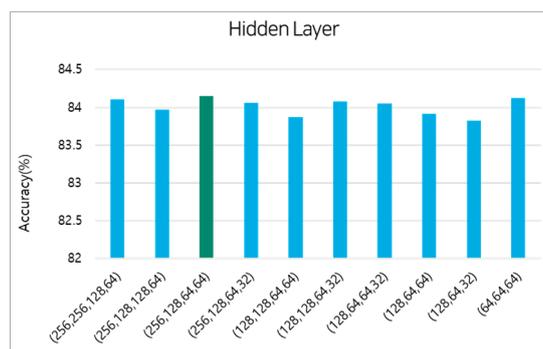
Third, we performed DNNs by setting the hidden layers and the number of nodes. The complexity of neural networks is determined by these two settings. It is important to set up the optimal node and hidden layers suitable to the model because model overfitting might occur and lead to poor learning. So, this means that the hidden layers and the number of nodes should be set up to suit this model.

For obtaining the optimal level of hidden layers and nodes, we tried to learn based on ① Set. However, the results showed that the learning was not done properly due to a shortage in the number of features in ① Set. Thus, we extracted a hidden layer based on “all data” which contained a default feature and new factors (Time, Day, Car of offender & victim, violation of Law).

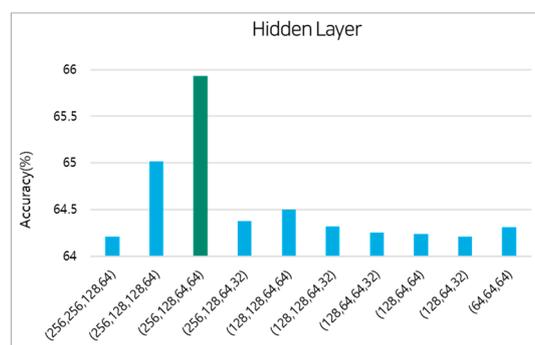
Consequently, the prediction accuracy of injury to the offender was observed to be mostly higher than 84%, and the highest accuracy was 84.15% with (256,128,64,64) nodes in the hidden layers. The prediction accuracy of injury to the victim reached mostly higher than 64%, and the highest accuracy was 65.93% with (256,128,64,64) nodes in hidden layers as well. It suggests that the optimal hidden layers about traffic accident prediction is (256,128,64,64), which deduced output value to the converging process. Therefore, we will simulate the study by setting 2017–2018 integrated data, 100 epochs, and (256,128,64,64) hidden layers(See Figure 2 & Table 6).

Table 6. Offender/Victim Injuries Accuracy by Hidden Layers.

The Number of Nodes in Hidden Layer	Accuracy	
	Offender	Victim
(256,256,128,64)	84.10	64.21
(256,128,128,64)	83.97	65.02
(256,128,64,64)	4.15	65.93
(256,128,64,32)	84.06	64.38
(128,128,64,64)	83.87	64.50
(128,128,64,32)	84.08	64.32
(128,64,64,32)	84.05	64.25
(128,64,64)	83.91	64.24
(128,64,32)	83.82	64.21
(64,64,64)	84.12	64.31



(a) Offender injuries accuracy



(b) Victim injuries accuracy

Figure 2. (a) Offender Injuries Accuracy by Hidden Layers; and (b) Victim Injuries Accuracy by Hidden Layers: Different colored bars stand for the highest accuracy is derived from the hidden layers.

5. Result

5.1. Extracting Traffic Accidents Context Features

Learning based on the optimization model was carried out to identify factors affecting the degree of offender injury. As shown in Table 7 and Figure 4, the ① set, which only uses external factors, showed 81.16% prediction accuracy. It means that more than 80% of the injury to the offender can be predicted using external factors only. In addition, 84% prediction accuracy was achieved on the degree of injuries if additional factors of the vehicle were applied to the ① set, and 84.38% accuracy was confirmed when the vehicle type and the violation of the law were taken into consideration. As with offender factors, 61.05% accuracy was confirmed for the victim's injury when using external factors only. In addition, 66.46% of accuracy was observed in predicting victim's injuries when vehicle type was added to ① Set and 64.43% accuracy when adding time and day factors along with the vehicle type.

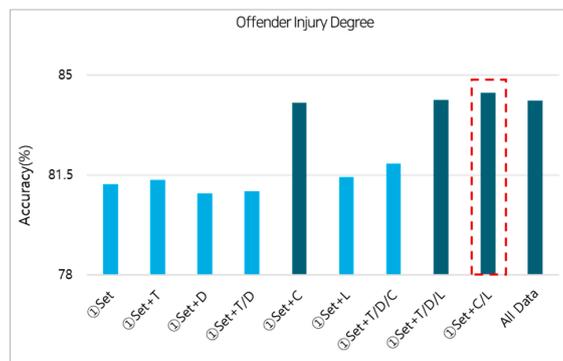
As a result of the analysis of the degree of injury of the offender, high accuracy is confirmed when the vehicle type and the violation of the law are added to the speed, the external factor. In addition, when the time, day, and vehicle types are set as additional factors in addition to external factors, high accuracy is achieved in the analysis of the victim's injury. The time series and vehicle type are considered to be the main factors in determining the victim's injury (See Table 7 & Figure 3).

Following the learning through the optimization model, we identified specific features that determine the degree of traffic accident injury. Based on the analysis through DNN learning, a random forest feature importance analysis is performed, which is a machine learning technique, to confirm factors more accurately. As a result, the main factors of injuring offenders were identified in the order of "Vehicle Type, Speed, Time, and Day of the week", and those of the victims were in the order of "Speed, Time, Victim Vehicle Type, and Day of the Week". It was found that the factors that determine the

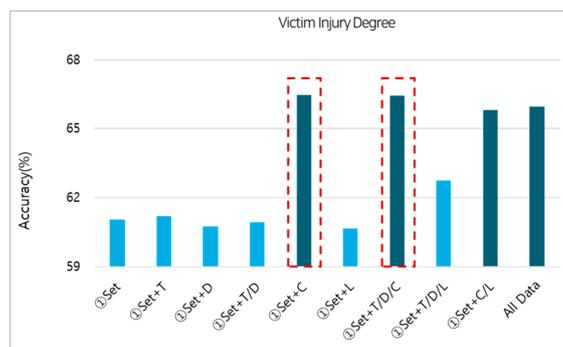
degree of injury to the offender and the victim are similar. However, for the offender, the vehicle type is found more critical, while for the victim, the link speed is identified as more important (See Table 8 & Figure 4).

Table 7. Offender/Victim Injuries Prediction Accuracy using optimal DNNs model.

Feature	Accuracy	
	Offender	Victim
①Set	81.16	61.05
①Set + T	81.32	61.21
①Set + D	80.85	60.75
①Set + T + D	80.93	60.92
①Set + C	84.02	66.46
①Set + L	81.41	60.67
①Set + T + D + C	81.89	66.43
①Set + T + D + L	84.12	62.75
①Set + C + L	84.38	65.83
All data	84.15	65.98



(a) Offender injuries accuracy

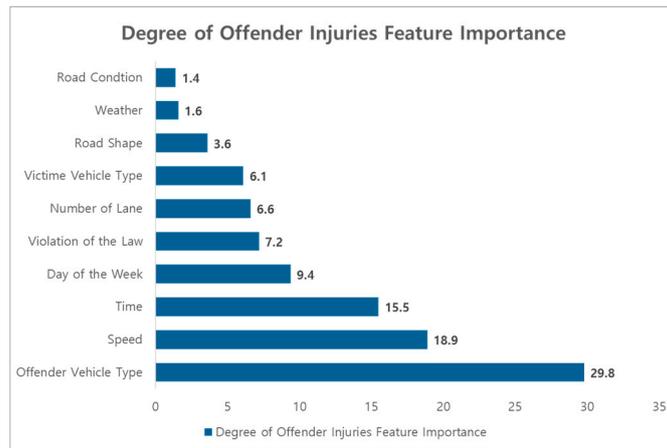


(b) Victim injuries accuracy

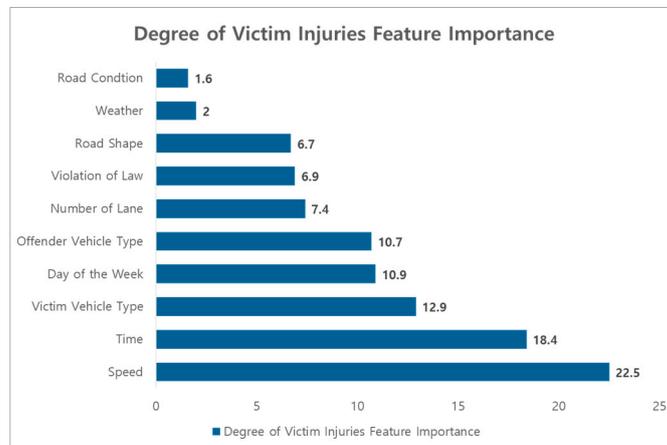
Figure 3. (a) Offender Injuries Accuracy by Optimal DNNs; (b) Victim Injuries Accuracy by Optimal DNNs: The red dash line rectangular stands for the highest accuracy at that factors.

Table 8. Result of Importance Analysis for Traffic Accident Factors.

Feature		Accuracy								
Offender Factors	Offender Vehicle Type (29.8)	Speed (18.9)	Time (15.5)	Day of the Week (9.4)	Violation of the Law (7.2)	Number of Lane (6.6)	Victim Vehicle Type (6.1)	Road Shape (3.6)	Weather (1.6)	Road Condition (1.4)
Victim Factors	Speed (22.5)	Time (18.4)	Victim Vehicle Type (12.9)	Day of the Week (10.9)	Offender Vehicle Type (29.8)	Number of Lane (6.6)	Violation of the Law (7.2)	Road Shape (3.6)	Weather (1.6)	Road Condition (1.4)



(a) Offender injuries Feature Importance



(b) Victim injuries Feature Importance

Figure 4. (a) Injuries Feature Importance; (b) Victim Injuries Feature Importance

5.2. Additional Consideration

Although the main factors determining the degree of offender and victim injury were identified with this study, it is necessary to find out why there exist the differences in accuracy between the two, notwithstanding the same dataset were used for the analysis. It is because the principal cause of injury to the offender is the vehicle type whose data was originated from the TAAS that requires no data correction, but the cause of injury to the victim (speed) is the data obtained from the TOPIS data that entail a lot of missing values (see Section 3.1). It is judged that the difference in the factors that determine the injury occurred in the original data (TAAS & TOPIS), lead to the differences in accuracy. However, the accuracy derived from this analysis is higher than that of existing statistics-based traffic accident prediction models (Poisson regression model, negative binomial regression model, etc. [56–59]). This is judged to be meaningful, as car accidents can be prevented in advance by predicting traffic accident injuries through DNN learning.

6. Conclusions

Automated Vehicles are under the spotlight as an alternative to diminishing traffic accidents. However, AVs accidents that occurred during several test drives led to worry about AVs safety. When AVs are commercialized and mixed with human driving vehicles, AVs-related accidents can occur even under ideal conditions. In order to solve this problem, studies on AVs ethics have been

conducted, but it was difficult to obtain an adequate answer. So, Germany and the United States issued AVs ethical guidelines emphasizing the prevention of traffic accidents. This study analyzed the traffic accident data based on DNNs to propose preventive measures for AVs accidents.

For the analysis of traffic accidents, we preprocessed TAAS, standard node-link data, and TOPIS data. In the DNNs analysis, the input layer consists of external factors and additional factors, and the output layer was set to the degree of injuries for both offenders & victims. Also, to construct the optimal hidden layers, we controlled the learning data range, the epoch, and the number of nodes in the hidden layers. For the analysis, we conducted learning by adding input factors incrementally. The results show that in the offender case, 81% of prediction accuracy was achieved when only external factors were considered, while the accuracy increased to 85% when factors such as violation of law and vehicle types are added to the analysis. In the victim case, the prediction accuracy remained at 61% with external factors only, but the accuracy increased to 67% when additional factors were taken into consideration. The main reason determining about 20% prediction accuracy deviation between offender and victim comes from the level of data accuracy of the key injury determinant between offender and victim. The offender's degree of injury is largely determined by the type of vehicle, which did not require the data collection. On the other hand, the speed which affected the degree of victim injury most were from the merged data of TOPIS and TAAS with many missing data. The analysis revealed that factors such as vehicle type, time, and day were found important in determining the degree of injury. In addition, a random forest importance analysis identified that the injury determinants of the offender were ordered vehicle type, speed, time, and day, while those of victims were speed, time, vehicle type, and day.

As discussed, vehicle type and speed were identified as the main factors of determining injury, respectively. Accordingly, in the future, we plan to conduct researches that can more accurately prevent accidents by recognizing vehicle types using the CNNs (YOLO, etc.) technique, and that can predict the degree of injury according to speed or suggesting an appropriate vehicle safety distance. Also, future study will apply a developed methodology (Meta AI etc.) using updated traffic accident data. However, since data from HVs were utilized to develop preventive measures for AVs accidents, there is a limit to directly apply the results to AVs. Nevertheless, since the analysis focused on external factors (such as weather, road conditions, road types, etc.) that are difficult to control even AVs, it is expected that they will be used as basic data for analyzing the impact of external factors related to traffic accidents related to AVs in the future. Future study is needed to analyze traffic accidents by using data on the causes (e.g., weather, body defects, etc.) of AVs accidents as input values.

Author Contributions: Conceptualization, K.H.; Data curation, J.S.; Formal analysis, M.K.; Investigation, M.K.; Methodology, M.K.; Project administration, K.H.; Software, M.K. and J.S.; Visualization, M.K.; Writing—original draft, M.K.; Writing—review & editing, M.K. and K.H. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported by a grant (code 20TLRP-B148970-03) from Transportation and Logistics R&D Program (TLRP) funded by Ministry of Land, Infrastructure and Transport of Korean government.

Acknowledgments: This paper is a revised and complemented content of "A Study on DNN (Deep Neural Network)-based Traffic Accident Context Analysis for The Design of Preventive Automated Driving System" Minhee Kang's Master Degree.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. SAE On-Road Automated Vehicle Standards Committee. *Taxonomy and Definitions for Terms Related to Driving Automation Systems for on-Road Motor Vehicles*; SAE International: Warrendale, PA, USA, 2018.
2. National Highway Traffic Safety Administration. *Vision for Safety 2.0 Guidance for Automated Vehicles*; US Department of Transportation: Washington, DC, USA, 2017.
3. Lee, K.; Jeon, S.; Kim, H.; Kum, D. Optimal path tracking control of autonomous vehicle: Adaptive full-state linear quadratic gaussian (lqg) control. *IEEE Access* **2019**, *7*, 109120–109133. [[CrossRef](#)]

4. Singh, S. *Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey* (No. DOT HS 812 115); US Department of Transportation: Washington, DC, USA, 2015.
5. Abraham, H.; Lee, C.; Brady, S.; Fitzgerald, C.; Mehler, B.; Reimer, B.; Coughlin, J.F. Autonomous vehicles and alternatives to driving: Trust, preferences, and effects of age. In Proceedings of the transportation research board 96th annual meeting (TRB'17), Washington, DC, USA, 8–12 January 2017.
6. Zhang, T.; Tao, D.; Qu, X.; Zhang, X.; Lin, R.; Zhang, W. The roles of initial trust and perceived risk in public's acceptance of automated vehicles. *Transp. Res. Part C Emerg. Technol.* **2019**, *98*, 207–220. [[CrossRef](#)]
7. Hartwich, F.; Witzlack, C.; Beggiato, M.; Krems, J.F. The first impression counts—A combined driving simulator and test track study on the development of trust and acceptance of highly automated driving. *Transp. Res. Part F Traffic Psychol. Behav.* **2019**, *65*, 522–535. [[CrossRef](#)]
8. Gogoll, J.; Müller, J.F. Autonomous Cars: In Favor of a Mandatory Ethics Setting. *Sci. Eng. Ethics* **2016**, *23*, 681–700. [[CrossRef](#)] [[PubMed](#)]
9. Gurney, J.K. Crashing into the unknown: An examination of crash-optimization algorithms through the two lanes of ethics and law. *Albany Law Rev.* **2015**, *79*, 183.
10. Bonnefon, J.F.; Shariff, A.; Rahwan, I. Autonomous vehicles need experimental ethics: Are we ready for utilitarian cars? *arXiv* **2015**, arXiv:1510.03346.
11. Federal ministry of transport and digital infrastructure. *Ethics Commission: Automated and Connected Driving*; BMVI: Berlin, Germany, 2017.
12. National Highway Traffic Safety Administration. *Federal Automated Vehicles Policy: Accelerating the Next Revolution in Roadway Safety*; US Department of Transportation: Washington, DC, USA, 2016.
13. Choi, J.K.; Ji, Y.G. Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *Int. J. Hum. Comput. Interact.* **2015**, *31*, 692–702. [[CrossRef](#)]
14. Xu, Z.; Zhang, K.; Min, H.; Wang, Z.; Zhao, X.; Liu, P. What drives people to accept automated vehicles? Findings from a field experiment. *Transp. Res. Part C Emerg. Technol.* **2018**, *95*, 320–334. [[CrossRef](#)]
15. Man, S.S.; Xiong, W.; Chang, F.; Chan, A.H.S. Critical Factors Influencing Acceptance of Automated Vehicles by Hong Kong Drivers. *IEEE Access* **2020**, *8*, 109845–109856. [[CrossRef](#)]
16. Ko, Y.; Rho, J.H.; Donghyung, Y. Assessing Benefits of Autonomous Vehicle System Implementation through the Network Capacity Analysis. *Korea Spat. Plan. Rev.* **2017**, *93*, 17–24. [[CrossRef](#)]
17. Tibljaš, A.D.; Giuffrè, T.; Surdonja, S.; Trubia, S. Introduction of Autonomous Vehicles: Roundabouts Design and Safety Performance Evaluation. *Sustainability* **2018**, *10*, 1060. [[CrossRef](#)]
18. Wang, Y.; Wang, L. Autonomous vehicles' performance on single lane road: A simulation under VISSIM environment. In Proceedings of the 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI) IEEE, Shanghai, China, 14–16 October 2017; pp. 1–5.
19. Kang, M.H.; Im, I.J.; Song, J.I.; Hwang, K.Y. Analyzing Traffic Impacts of Automated Vehicle on Expressway Weaving Sections—A Case Study using Seoul-Singal Ramp Area-. *J. Transp. Res.* **2019**, *26*, 59–75. (In Korean)
20. Lee, S.Y.; Oh, M.S.; Oh, C.; Jeong, E.B. Automated Driving Aggressiveness for Traffic Management in Automated Driving Environments. *J. Korean Soc. Transp.* **2018**, *36*, 38–50. (In Korean) [[CrossRef](#)]
21. Morando, M.M.; Tian, Q.; Truong, L.T.; Vu, H.L. Studying the Safety Impact of Autonomous Vehicles Using Simulation-Based Surrogate Safety Measures. *J. Adv. Transp.* **2018**, *2018*. [[CrossRef](#)]
22. Kolarova, V.; Steck, F.; Bahamonde-Birke, F.J. Assessing the effect of autonomous driving on value of travel time savings: A comparison between current and future preferences. *Transp. Res. Part A Policy Pract.* **2019**, *129*, 155–169. [[CrossRef](#)]
23. Tscharktschiew, S.; Evangelinos, C. Pigouvian road congestion pricing under autonomous driving mode choice. *Transp. Res. Part C Emerg. Technol.* **2019**, *101*, 79–95. [[CrossRef](#)]
24. König, M.; Neumayr, L. Users' resistance towards radical innovations: The case of the self-driving car. *Transp. Res. Part F Traffic Psychol. Behav.* **2017**, *44*, 42–52. [[CrossRef](#)]
25. Abraham, H.; Lee, C.; Brady, S.; Fitzgerald, C.; Mehler, B.; Reimer, B.; Coughlin, J.F. *Autonomous Vehicles, Trust, and Driving Alternatives: A Survey of Consumer Preferences*; MIT AgeLab: Cambridge, MA, USA, 2016; pp. 1–16.
26. Nordhoff, S.; Van Arem, B.; Happee, R. Conceptual Model to Explain, Predict, and Improve User Acceptance of Driverless Podlike Vehicles. *Transp. Res. Rec. J. Transp. Res. Board* **2016**, *2602*, 60–67. [[CrossRef](#)]
27. Kaan, J. User Acceptance of Autonomous Vehicles. Master's Thesis, Faculty of Technology, Policy and Management (TPM), Delft University of Technology, Delft, The Netherlands, 2017.

28. Im, I.J.; Song, J.I.; Lee, J.Y.; Hwang, K.Y. Analysis of the Perception of Autonomous Vehicles Using Text Mining Technique. *J. Korea Inst. Intell. Transp. Syst.* **2017**, *16*, 231–243.
29. Hong, T.-S.; Kwon, Y.-S.; Kuack, B.-S.; Lee, J.-C. An Exploratory Study on Criminal Responsibility of Self-driving Car Accident Type. *Leg. Theory Pract. Rev.* **2018**, *6*, 231–252. [[CrossRef](#)]
30. Yang, J.S.; Kim, K.J.; Kwon, H.B. A Study on the Differences of Perception on the Causes of Self-driving Cars in Human Distribution. *J. Distrib. Manag. Res.* **2019**, *22*, 15–22. (In Korean)
31. Thomson, J.J. The Trolley Problem. *Yale Law J.* **1985**, *94*, 1395. [[CrossRef](#)]
32. Kim, S.-R. The Ethical Approach to Traffic Accident-Algorithm of Autonomous Driving Vehicle and Its Legal Implications. *IT LAW Rev.* **2017**, *15*, 193–219. [[CrossRef](#)]
33. Bae, I.H.; Lee, K.H. Criteria of the Prior Protection in Fully Autonomous Car Accidents. *Study Future* **2018**, *3*, 25–48. (In Korean)
34. Goodall, N.J. Ethical Decision Making during Automated Vehicle Crashes. *Transp. Res. Rec. J. Transp. Res. Board* **2014**, *2424*, 58–65. [[CrossRef](#)]
35. Fleetwood, J. Public Health, Ethics, and Autonomous Vehicles. *Am. J. Public Health* **2017**, *107*, 532–537. [[CrossRef](#)] [[PubMed](#)]
36. Wu, Y.; Tan, H.; Qin, L.; Ran, B.; Jiang, Z. A hybrid deep learning based traffic flow prediction method and its understanding. *Transp. Res. Part C Emerg. Technol.* **2018**, *90*, 166–180. [[CrossRef](#)]
37. Polson, N.G.; Sokolov, V.O. Deep learning for short-term traffic flow prediction. *Transp. Res. Part C Emerg. Technol.* **2017**, *79*, 1–17. [[CrossRef](#)]
38. Zhao, Z.; Chen, W.; Wu, X.; Chen, P.C.Y.; Liu, J. LSTM network: A deep learning approach for short-term traffic forecast. *IET Intell. Transp. Syst.* **2017**, *11*, 68–75. [[CrossRef](#)]
39. Kim, D.H.; Hwang, K.Y.; Yoon, Y. Prediction of Traffic Congestion in Seoul by Deep Neural Network. *J. Korea Inst. Intell. Transp. Syst.* **2019**, *18*, 44–57. [[CrossRef](#)]
40. Zhang, Z.; He, Q.; Gao, J.; Ni, M. A deep learning approach for detecting traffic accidents from social media data. *Transp. Res. Part C Emerg. Technol.* **2018**, *86*, 580–596. [[CrossRef](#)]
41. Zhang, W.; Liu, K.; Zhang, W.; Zhang, Y.; Gu, J. Deep Neural Networks for wireless localization in indoor and outdoor environments. *Neurocomputing* **2016**, *194*, 279–287. [[CrossRef](#)]
42. Li, P.; Chen, X.; Shen, S. Stereo R-CNN Based 3D Object Detection for Autonomous Driving. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 7636–7644.
43. Chen, S.; Leng, Y.; Labi, S. A deep learning algorithm for simulating autonomous driving considering prior knowledge and temporal information. *Comput. Civ. Infrastruct. Eng.* **2019**, *35*, 305–321. [[CrossRef](#)]
44. Zeng, W.; Wang, S.; Liao, R.; Chen, Y.; Yang, B.; Urtasun, R. DSDNet: Deep Structured self-Driving Network. *arXiv* **2020**, arXiv:2008.06041.
45. Sun, P.; Kretzschmar, H.; Dotiwalla, X.; Chouard, A.; Patnaik, V.; Tsui, P.; Guo, J.; Zhou, Y.; Chai, Y.; Caine, B.; et al. Scalability in Perception for Autonomous Driving: Waymo Open Dataset. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 2443–2451.
46. Djuric, N.; Radosavljevic, V.; Cui, H.; Nguyen, T.; Chou, F.-C.; Lin, T.-H.; Singh, N.; Schneider, J. Uncertainty-aware Short-term Motion Prediction of Traffic Actors for Autonomous Driving. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision, Snowmass Village, CO, USA, 1–5 March 2020; pp. 2095–2104.
47. NVIDIA. *NVIDIA Self-Driving Safety Report*; NVIDIA: Santa Clara, CA, USA, 2018.
48. Traffic Accident Analysis System. 2019. Available online: <http://taas.koroad.or.kr> (accessed on 20 September 2020).
49. Seoul Transport Operation and Information Services. 2018. Available online: <http://topis.seoul.go.kr> (accessed on 20 September 2020).
50. Korea Transport DataBase. 2019. Available online: <https://www.ktdb.go.kr> (accessed on 20 September 2020).
51. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533. [[CrossRef](#)]
52. TensorFlow. Available online: <https://www.tensorflow.org/> (accessed on 20 September 2020).
53. Keras. Available online: <https://keras.io/> (accessed on 20 September 2020).

54. Nair, V.; Hinton, G.E. Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 21–24 June 2010; pp. 807–814.
55. Bridle, J.S. Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters. In *Advances in Neural Information Processing Systems 2*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1990; pp. 211–217.
56. Lee, S.H.; Lee, Y.D.; Do, M.S. Safety Impacts of Red Light Enforcement on Signalized Intersections. *J. Korean Soc. Transp.* **2012**, *30*, 93–102. (In Korean) [[CrossRef](#)]
57. Lee, S.H.; Park, M.H.; Woo, Y.H. A Study on Development Crash Prediction Model for Urban Intersections Considering Random Effects. *J. Korea Inst. Intell. Transp. Syst.* **2015**, *14*, 85–93. (In Korean) [[CrossRef](#)]
58. Lee, S.H.; Woo, Y.H. Safety Performance Functions for Central Business Districts Using a Zero-Inflated Model. *Int. J. Highw. Eng.* **2016**, *18*, 83–92. (In Korean) [[CrossRef](#)]
59. Lee, S.H.; Lee, Y.D.; Do, M.S. Analysis on Safety Impact of Red Light Cameras using the Empirical Bayesian Approach. *Transp. Lett. Int. J. Transp. Res.* **2016**, *8*, 241–249. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).