ISSN: 0974-5823 Vol. 7 No. 2 February, 2022

International Journal of Mechanical Engineering

A Proposal of Traffic Control Method for Cargo Vehicles Using Public Data Traffic Accident Analysis

Seungyeon Hwang[1] and Jeongjoon Kim[2]

[1]Novacos Co., Ltd., Anyang-si, 14057, South Korea

[2]Anyang University, Department of Software, Anyang-si, 14028, South Korea

Abstract - The number of traffic accidents in Korea is more than 200,000 each year, and traffic accidents of freight vehicles cause more casualties than other vehicles. Public data showed that most of the freight vehicle accidents were unsafe driving, and there were a lot of violations such as overload, speed violation, drowsy driving, and illegal dismantling of speed limit devices. Until now, the accident rate of freight cars has not decreased, and the government and police have made a lot of effort for national security, but it has not been improved yet. Therefore, in this paper, we are going to analyze big data on the number of traffic accidents, violations of laws, traffic accidents by road type or form, and propose new ways to improve traffic.

Index Terms - traffic accidents, truck, big data, public data, vehicle, person, hadoop, R

INTRODUCTION

With more than 200,000 traffic accidents occurring in South Korea every year, cargo vehicles are the ones with more injuries and deaths compared to the total number of traffic accidents. Once an accident occurs, a truck becomes a major accident and can lead to a second or third accident. In addition to direct accidents, it is considered a trouble because it can cause problems in the flow of road traffic or consume a large amount of accident road authority in the case of an accident road, which can result in consumption of the budget.[1][2]

Among the cargo truck accident records provided by the cargo deduction, 22,814 traffic accidents occurred in 2016 based on 179,760 vehicles contracted for the deduction, 13% of which were killed, 36,085 were injured, and a total of 92,716 were reported. It should be noted that most of the cases were 91,177 cases of non-compliance with safe driving. Freight vehicles are violating quite a lot of laws, such as overloading, speed violations, drowsy driving, and illegal dismantling of speed limiting devices. The damage was found to be great. According to public data in 2016, in the case of speeding, which accounted for 0.3% of the causes of traffic accidents, the number of fines imposed by the police was a whopping 75%, with conflicting records.

Therefore, there is a need to improve on-site traffic control and unmanned camera control, which are far from traffic accident prevention, and to reduce such formal and ineffective crackdowns and find more effective traffic control methods based on records.

This study seeks to statistically identify the number of accidents, violations, casualties, etc. through big data analysis of cargo vehicle accident data in public data and propose new methods of traffic control that have not been improved.

RELATED WORKS

In this study, public data on public data portals and traffic accident analysis systems were collected for traffic accident analysis. Hadoop is a representative software framework for big data systems, and data processing was performed using Hadoop and PIG in this paper. Subsequently, the preprocessed data were implemented in the R programming language for analysis and visualization.

I. Bigdata Processing Technology

It is the process of refining and processing necessary data to be used for analysis from collected public data. A typical solution used in processing techniques is MapReduce, which uses PIG in Hadoop in this work.

MapReduce is a method of processing big data by utilizing distributed multi-node and consists of map functions and redox functions[3]. The map function receives the data and transforms it into a key-value form, groups the relevant data, and passes it to the redox function, which handles it[4,5]. Hadoop is a Java-based framework for distributing large amounts of data, storing data in the distributed system HDFS (Hadoop Distributed File System), and processing data in map-reduce[6,7].

Pig is a representative platform used in big data analysis systems and is a high-level scripting language for using MapReduce and consists of infrastructure for it.

The pig's infrastructure layer is composed of compilers and creates a data defense or sequence of MapReduce for large-scale parallelism[8].

II. Analysis and Visualization of Bigdata

Data collection, processing, and analysis based on processed data should be performed to suit the purpose of the system. It should produce meaningful results subsequently analyzed, using R, a programming language for statistical calculations and graphical visualization used in this process[9]. R is a powerful interpreter language that supports statistical and graphical functions and has recently been used primarily in big data and data science. This paper deals with vehicle traffic accident information using data analysis and graphically visualization.

ANALYSIS

The crackdown methods in this paper analyze items that show no significant reduction or continuous heavy casualties based on monthly, road types, and types of violations in four ways, and suggest items that strengthen penalties for safe road traffic.

The contents of the four analyses below were analyzed based on the items of public data provided by the Road Traffic Authority, the Public Data Portal, and the Traffic Accident Analysis System (TAAS), and are somewhat different from the information of contract vehicles released in the introduction. In the main body, accident and casualties records from 2013 to 2016 are analyzed and visualized in four categories through R.

Since general on-site crackdowns and intensive crackdowns except for unmanned camera crackdowns have restrictions on police personnel and patrol vehicles, the location and timing of major on-site crackdowns should be identified in order to reduce casualties and ensure effective road traffic safety.

I. Monthly Incident Data Analysis

Monthly accident status was analyzed from 2013 to 2016. Monthly data based on the number of accidents, deaths, and casualties per year were visualized as linear graphs using the ggplot2 library in the R language.

The monthly analysis graphs, including Fig. 1, are represented by the X-axis as monthly and 12 months, m1 to m12. The Y-axis refers to the number of occurrences, followed by orange, green, blue, and purple, respectively, in 2013, 2014, 2015, and 2016.

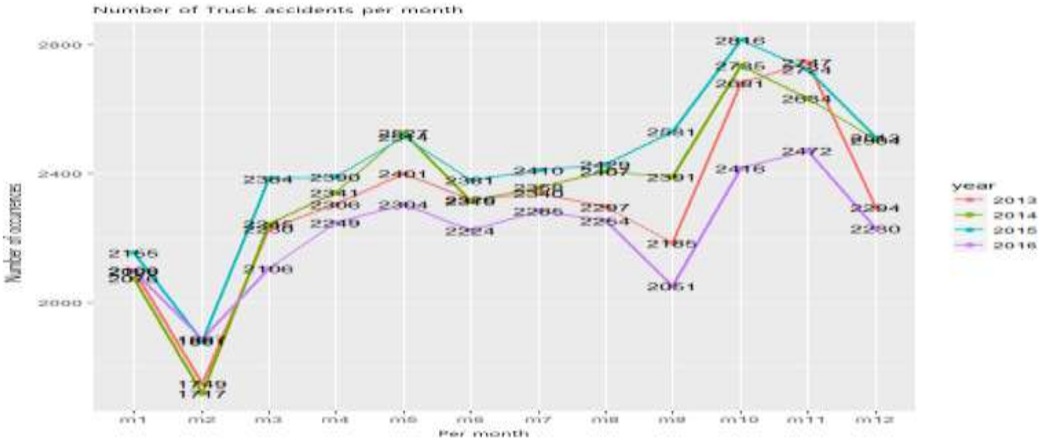


Fig.1. The number of Truck accidents per month

According to the graph above, April and May will rise slightly and the number of occurrences will increase sharply from September to November. The maximum annual accident periods are shown in Table 1 below.

Table 1. The month of the upper truck vehicle accident

Year	month	Occurrence
2013	11	2747
	10	2681
	5	2401
	7	2340
2014	10	2735
	11	2634
	5	2527
	12	2504
2015	10	2816
	11	2724

	9	2531
	5	2514
2016	11	2472
	10	2416
	5	2304
	6	2285

Fig. 2 shows the number of casualties, including serious injuries, current injuries, and reported injuries monthly. The number of casualties was the same trend as the number of occurrences in Figure 1, with the high number of casualties recorded during spring and fall. The maximum number of casualties per year is shown in Table 2 below.

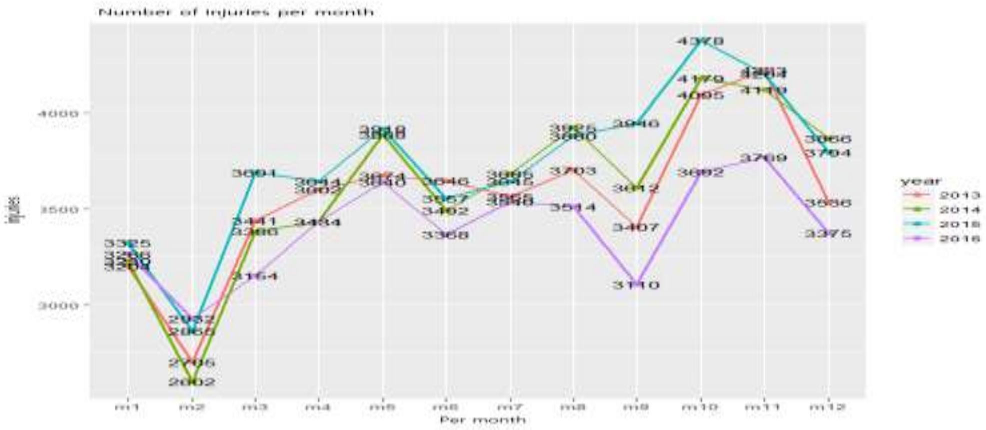


Fig.2. The number of injuries per month

Table 2. The month of the upper truck injuries

Year	month	Occurrence
2013	11	4223
	10	4095
	8	3703
	5	3674
2014	10	4179
	11	4119
	9	3925
	5	3888
2015	10	4378
	11	4204
	8	3946
	5	3918
2016	11	3769
	10	3692
	5	3640
	7	3540

The death toll could play an important role in how to crack down and improving road traffic laws. In addition, the number of deaths can be determined to estimate the seriousness and extent of the damage caused by cargo vehicle accidents, indicating the need to improve the crackdown.

In Fig. 3, the death toll shows a different type of trend, unlike Figures 1 and 2. The maximum annual death toll is shown in Table 3 below.

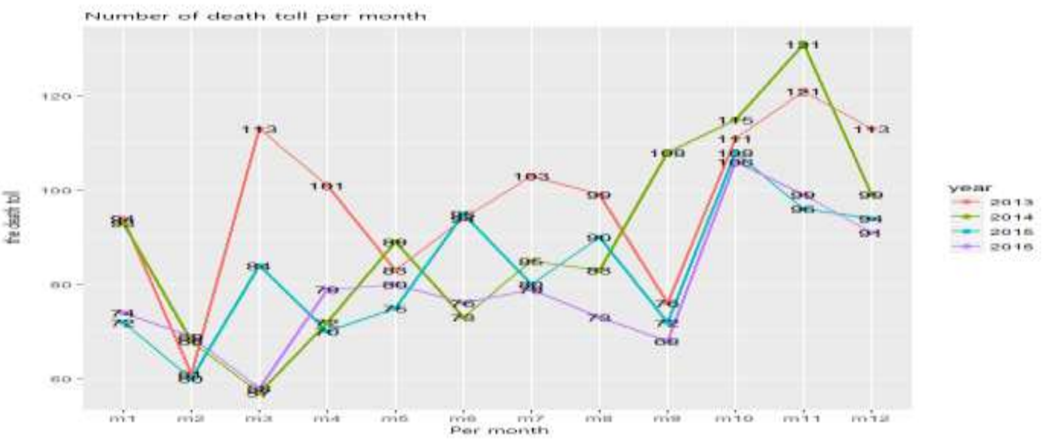


Fig.3. The number of death tolls per month

Table 3. The month of the upper death toll

Year	month	Occurrence
2013	11	121
	12	113
	3	113
	10	111
2014	11	131
	10	115
	9	108
	12	99
2015	10	108
	11	96
	6	95
	12	94
2016	10	106
	11	99
	12	91
	5	80

Analysis of monthly data for four years from 2013 showed that many cargo vehicle accidents occurred in October and November, with the number of casualties and deaths proportionally high. Next, in terms of seasons such as December, May, and June, there are many accidents in autumn and then spring.

II. Analysis of Accident Data by Road Type

The current status of accidents by road type was analyzed from 2013 to 2016. Based on the number of accidents and deaths per year, the total number of vehicles by use and freight cars were divided by type of road and visualized as a bar graph using the pplot2 library of R.

In Figures 4-7, the X-axis is represented by each type of road, followed by city, county, other, general national roads, high-speed national roads, local roads, and special metropolitan cities. The Y-axis refers to the number of traffic accidents and compares the total number of vehicles by use and the cargo trucks by year (2013, 2014, 2015, 2016).

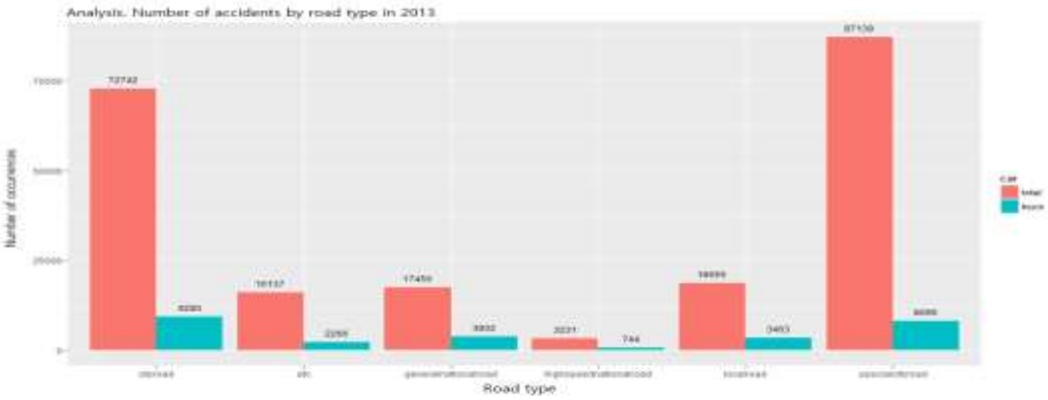


Fig.4. Number of accidents by road type in 2013

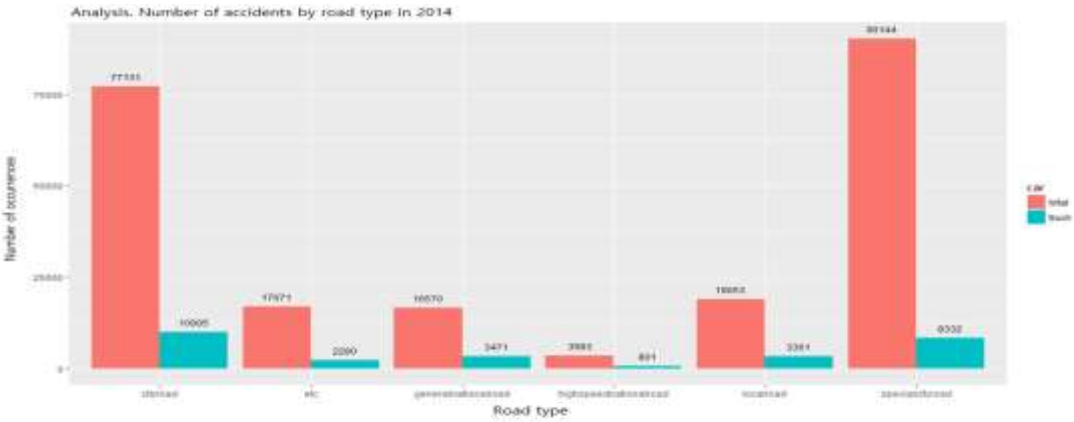


Fig.5. Number of accidents by road type in 2014

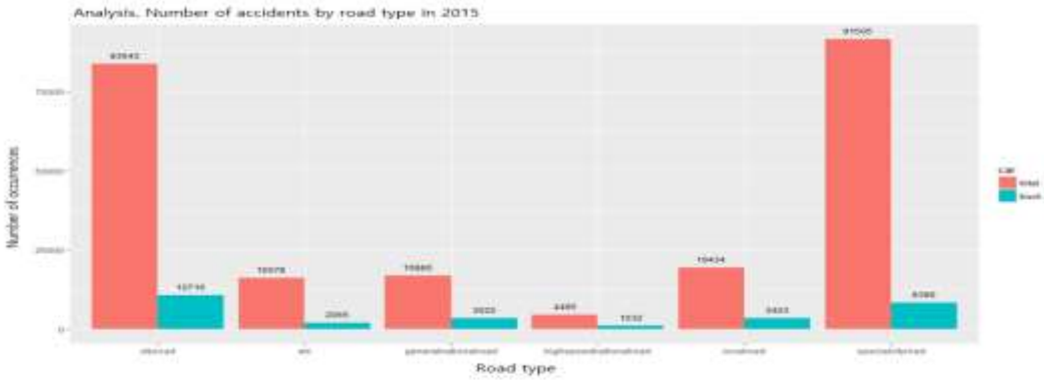


Fig.6. Number of accidents by road type in 2015

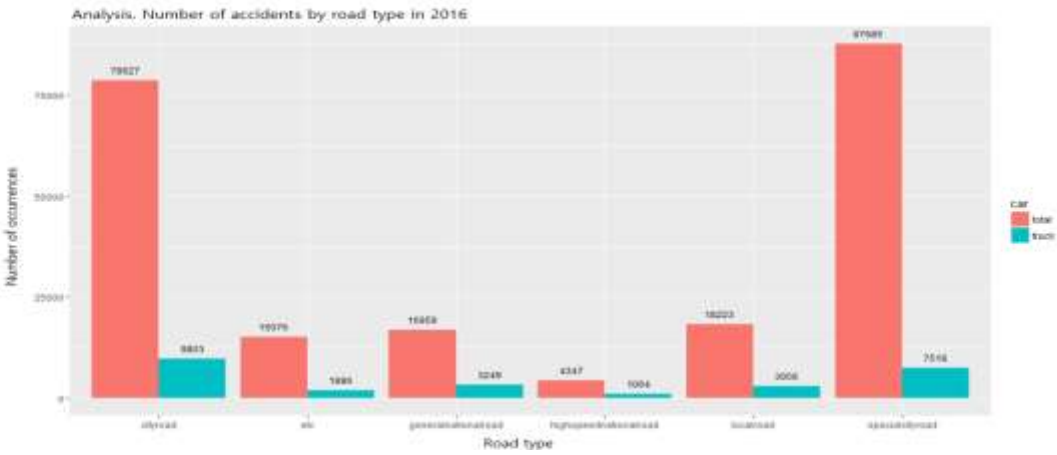


Fig.7. Number of accidents by road type in 2016

If you look at the graph above, you can see that the total number of traffic accidents by use accounts for a considerable number of traffic accidents in cities and counties and special metropolitan cities. However, if you look at the number of traffic accidents in cargo vehicles, you can see that bar graphs by road type are relatively evenly distributed compared to the total number of vehicles by use. Tables 4 and 5 below show the total number of vehicles by use and the number of traffic accidents involving cargo vehicles.

Table 4. Total number of traffic accidents by usage

Year	City and Country	Etc.	General National Road	Highway	Province	Special metropolitan city
2013	72742	16139	17450	3231	18655	78139
2014	77131	17071	16570	3583	19053	90144
2015	83643	16078	16880	4495	19434	91505
2016	78627	15076	16959	4347	18223	87685

Table 5. The number of truck accidents

Year	City and Country	Etc.	General National Road	Highway	Province	Special metropolitan city
2013	9280	2266	3802	744	3463	8095
2014	10005	2280	3471	801	3361	8332
2015	10716	2065	3522	1032	3403	8390
2016	9803	1886	3249	1064	3056	7518

The following Fig. 8-11 show the breakdown of traffic accident fatalities by road type and compare the total number of vehicles by use with cargo vehicles. Looking at the graph, when comparing the number of deaths on general national roads, express highways, and local roads with city-gun and special metropolitan cities, the number of deaths is quite high.

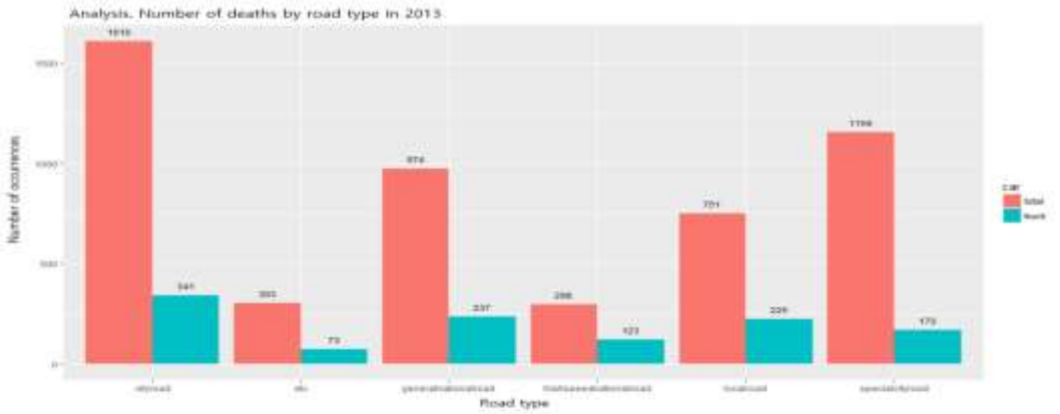


Fig.8. Number of deaths by road type in 2013

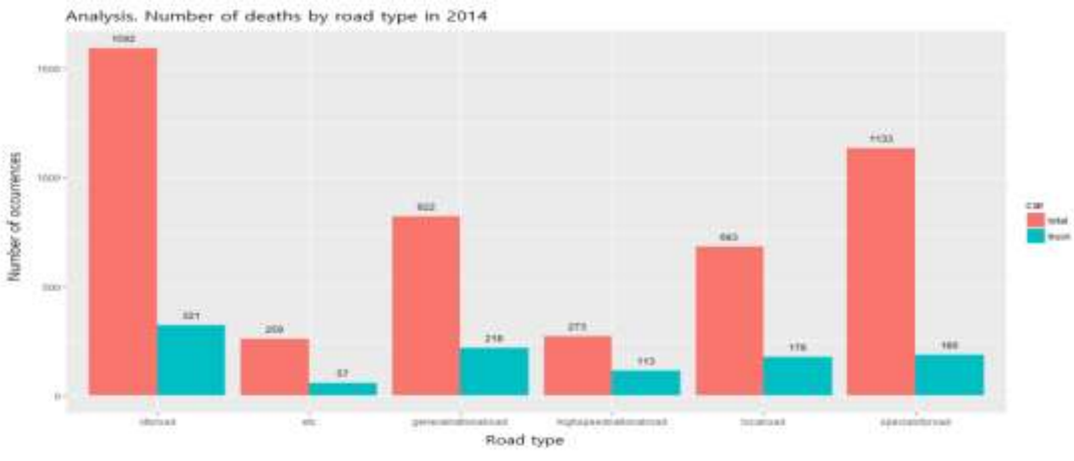


Fig.9. Number of deaths by road type in 2014

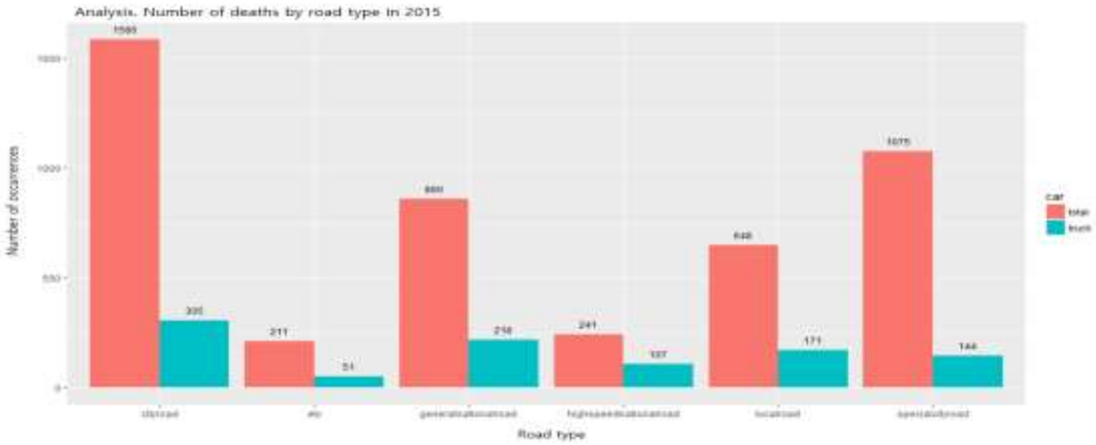


Fig.10. Number of deaths by road type in 2015

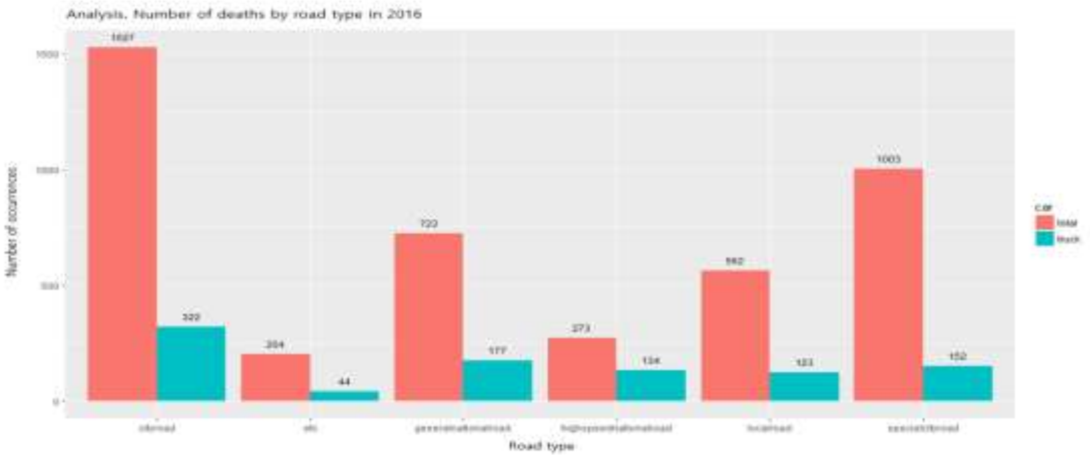


Fig.11. Number of deaths by road type in 2016

Table 6. Total number of traffic accident deaths by total usage

Year	City and Country	Etc.	General National Road	Highway	Province	Special metropolitan city
2013	1610	303	974	298	751	1156
2014	1592	259	822	273	683	1133
2015	1586	211	860	241	648	1075
2016	1527	204	723	273	562	1003

Table 7. Number of truck accident deaths

Year	City and Country	Etc.	General National Road	Highway	Province	Special metropolitan city
2013	341	73	237	123	225	170
2014	321	57	218	113	178	186
2015	305	51	218	107	171	144
2016	322	44	177	134	123	152

III. Analysis of Accident Data by Road Shape

From 2013 to 2016, the current status of accidents by road type was analyzed. Based on the number of accidents and deaths per year, the total number of vehicles by use and cargo vehicles were divided by road type and visualized as a bar graph using the ggplot2 library of R.

The X-axis in Fig. 12-15 is expressed in order of road type and other single roads, intersections, tunnels, underground roads, intersections, crosswalks, bridges, overpasses, other/unknown, and railroad crossing. The Y-axis refers to the number of traffic accidents and compares the total number of vehicles by use and cargo vehicles by year (2013, 2014, 2015, and 2016).

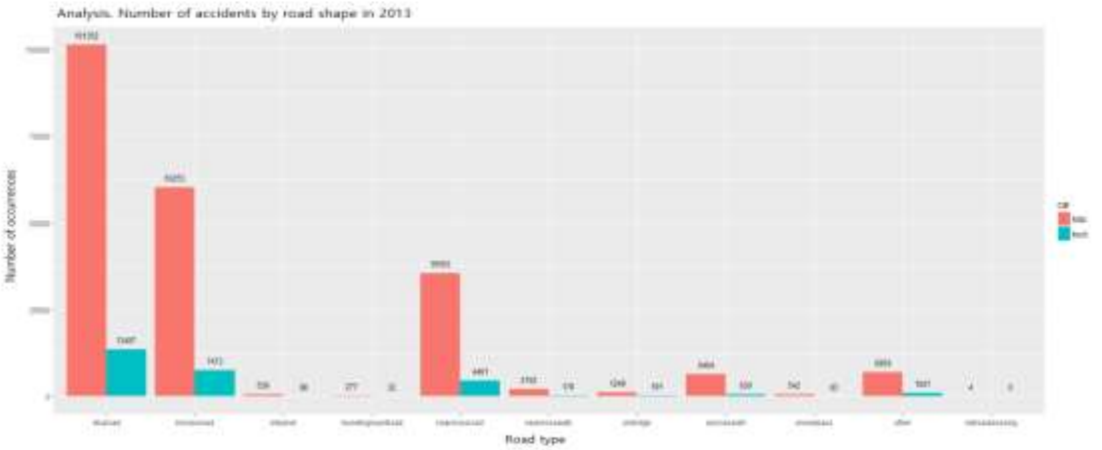


Fig.12. Number of accidents by road shape in 2013

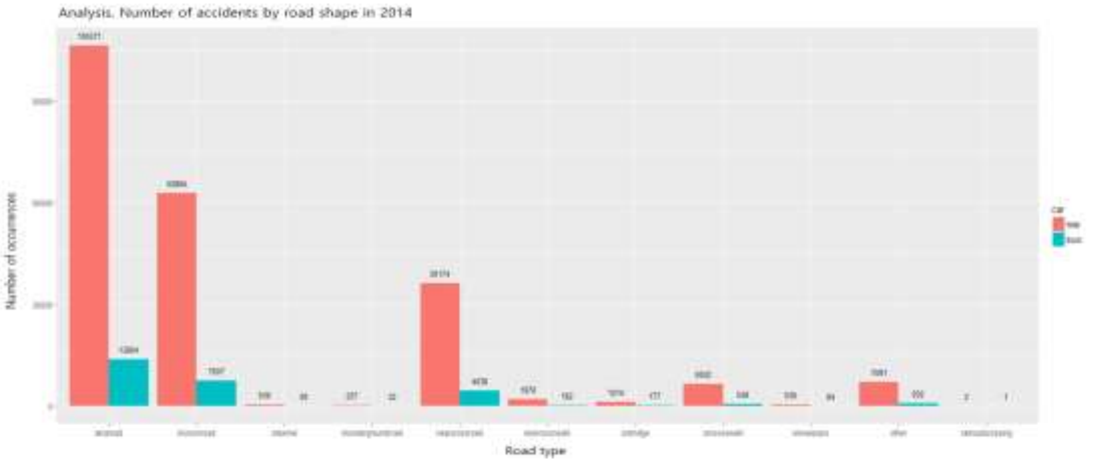


Fig.13. Number of accidents by road shape in 2014

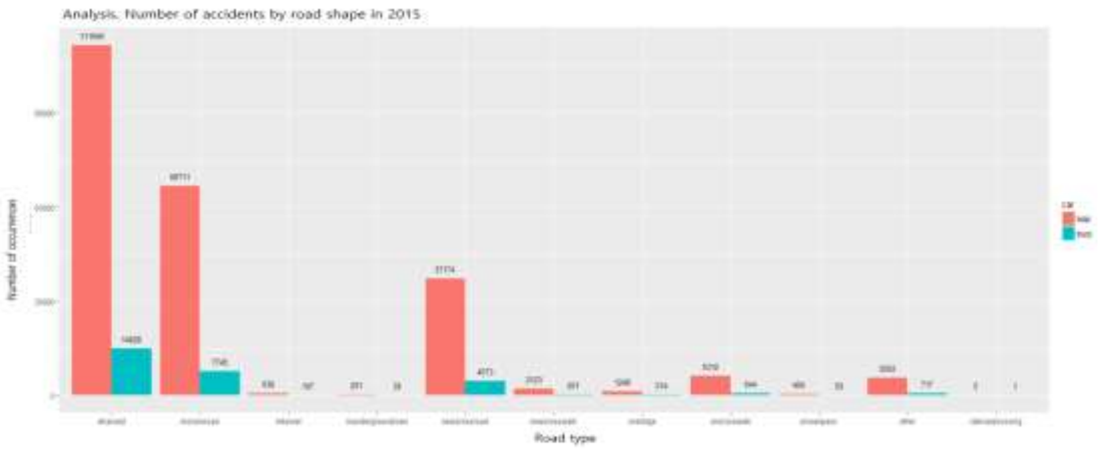


Fig.14. Number of accidents by road shape in 2015

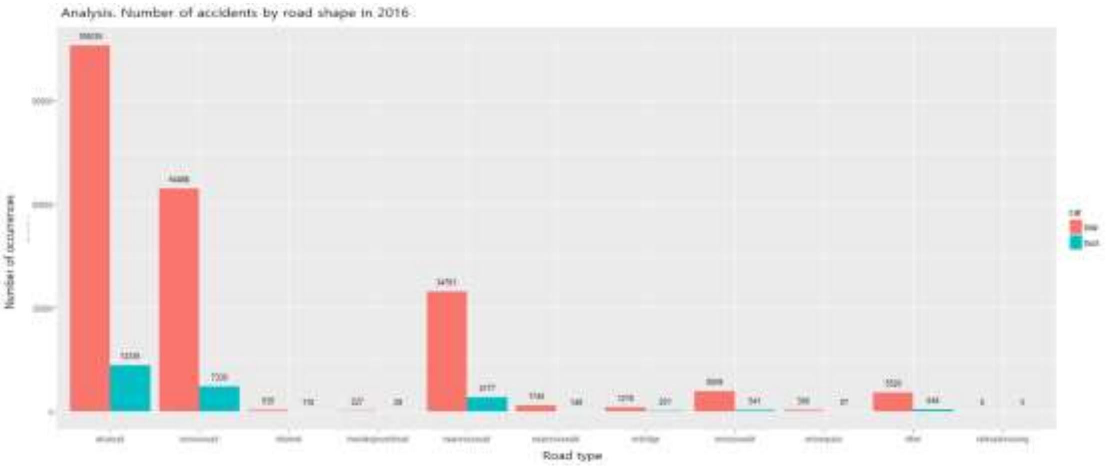


Fig.15. Number of accidents by road shape in 2016

If you look at the graph above, you can see that these three types of roads are significantly different from other types of roads, in and near intersections. Depending on the type of road, the ratio of the total number of vehicles by use to the number of cargo vehicles generated is relatively small.

Table 8. Total number of traffic accidents by usage by upper road type

Year	Other single road	In the intersection	Near the intersection
2013	101352	60253	35563
2014	106371	62894	36174
2015	111565	66711	37174
2016	106036	64488	34761

Table 9. Total number of traffic accidents by usage by upper road type

Year	Other single road	In the intersection	Near the intersection
2013	13497	7473	4461
2014	13864	7607	4638
2015	14828	7745	4573
2016	13335	7330	4177

Fig. 16-19 below shows the total number of deaths from traffic accidents by road type, compared with the total number of vehicles by use and cargo vehicles. If you look at the graph, you can see that the death toll was high, in order of other single roads, in intersections, and near intersections.

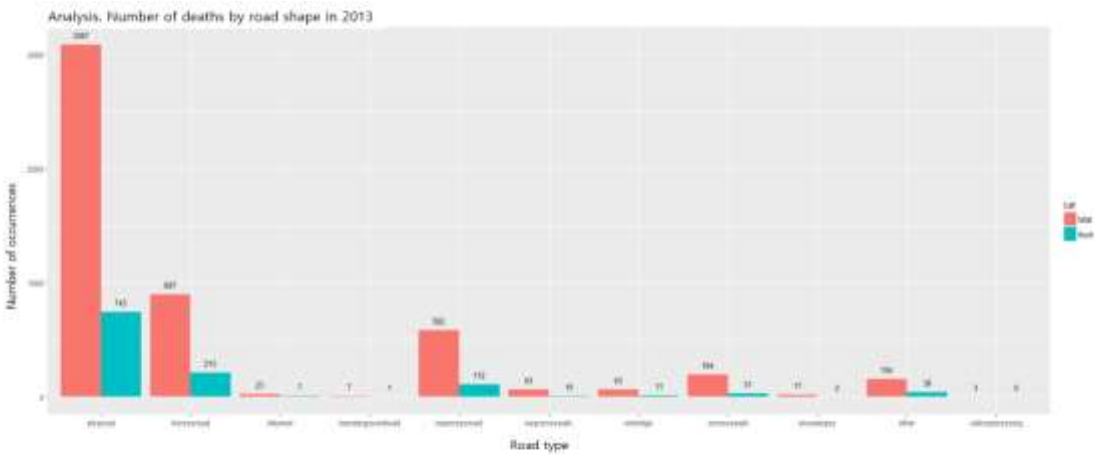


Fig.16. Number of accidents by road shape in 2013

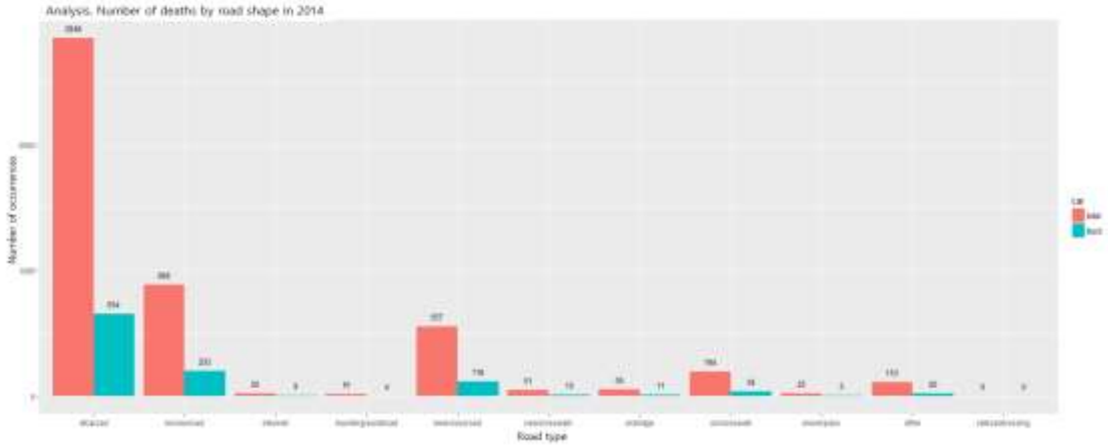
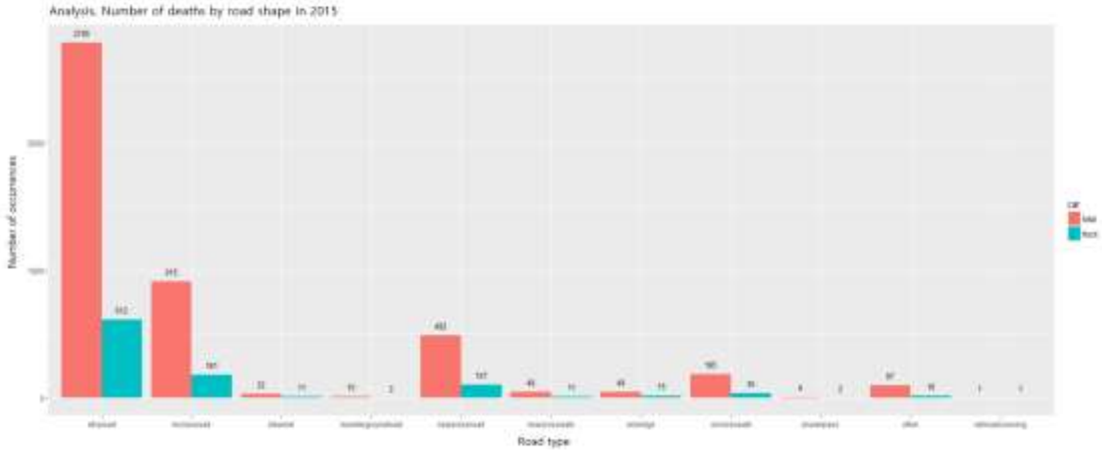


Fig.17. Number of accidents by road shape in 2014



t

Fig.18. Number of accidents by road shape in 2015

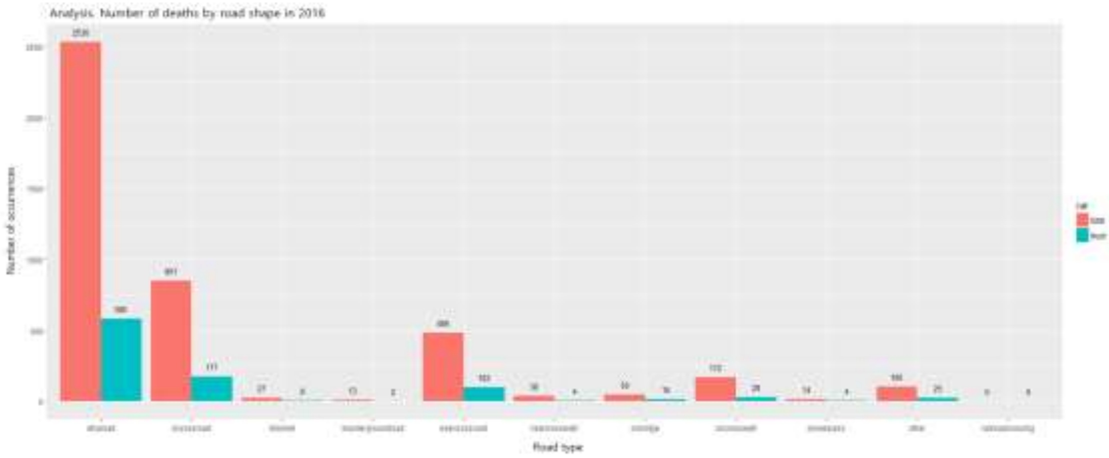


Fig.19. Number of accidents by road shape in 2016

The other three types of roads, within, and near the intersection, can see higher mortality than other types of roads. Tables 10 and 11 below compare the high death toll on major road types with the total number of vehicles by use and the number of deaths on cargo vehicles.

Table 10. Total number of deaths from traffic accidents by usage by upper road type

Year	Other single road	In the intersection	Near the intersection
2013	3087	897	582
2014	2848	886	557
2015	2785	915	492
2016	2536	851	485

Table 11. Number of people killed in traffic accidents by upper road type

Year	Other single road	In the intersection	Near the intersection
2013	743	213	112
2014	654	203	118
2015	612	181	107
2016	585	177	102

By comparing the total number of vehicles and the percentage of deaths, the number of cargo vehicles, and the percentage of deaths by use, the table 12-13 shows that the difference in the number of deaths of cargo vehicles is more than doubled.

Table 12. Total number of car accidents and death rate by upper use

Year	Other single road	In the intersection	Near the intersection
2013	3.04%	1.48%	1.63%
2014	2.67%	1.40%	1.53%
2015	2.49%	1.37%	1.32%
2016	2.39%	1.31%	1.39%

Table 13. Number of accidents involving upper freight vehicles and percentage of deaths

Year	Other single road	In the intersection	Near the intersection
2013	5.50%	2.85%	2.51%
2014	4.71%	2.66%	2.54%
2015	4.12%	2.33%	2.33%
2016	4.38%	2.41%	2.44%

IV. Analysis of Accident Data by Type of Violation of the Law

From 2013 to 2016, the company analyzed the status of accidents that violated laws and conditions, and these are data on non-compliance of safe driving, which is the direct cause of most cargo vehicle accidents. Based on the number of accidents and deaths per year.

The items of violation of the code are the same as the classification criteria for public data, and the first division is divided into driver code violations, poor maintenance, and pedestrian negligence.

First of all, the group of driver violations: overwork, speeding, overtaking methods, prohibition of overtaking, centerline violations, signal violations, safety distance violations, unfair rotation, traffic priority violations, noncompliance, safety driving obligations, pedestrian protection violations.

Secondly, there is one subcategory of poor maintenance, which violates the prohibition of re-driving the amount of maintenance.

Among the many violations in the group, the analysis was conducted with only the top 10 items that had the most accidents.

The X-axis of the analysis graphs by law violation, including Fig. 20, is expressed as il1 to il10 as major illegal items, and the contents are speeding, centerline violation, signal violation, not securing a safe distance, non-compliance with safe driving obligation, violation of the crossing method, and pedestrian protection. Violation of duty, violation of lane (violation of course change), obstruction of the passage of cars going straight and right turn, and others (violation of driver's laws). The Y-axis represents the number of occurrences and deaths and compares the total number of vehicles with the accident records of freight vehicles.

In 2013, the highest number of accidents and fatalities was due to non-compliance with safe driving obligations (il5), with 14582 accidents and 817 deaths. The next more than a thousand cases are violations of the centerline (il2), signal violations (il3), not securing a safe distance (il4), and violations of the crossing method (il6).

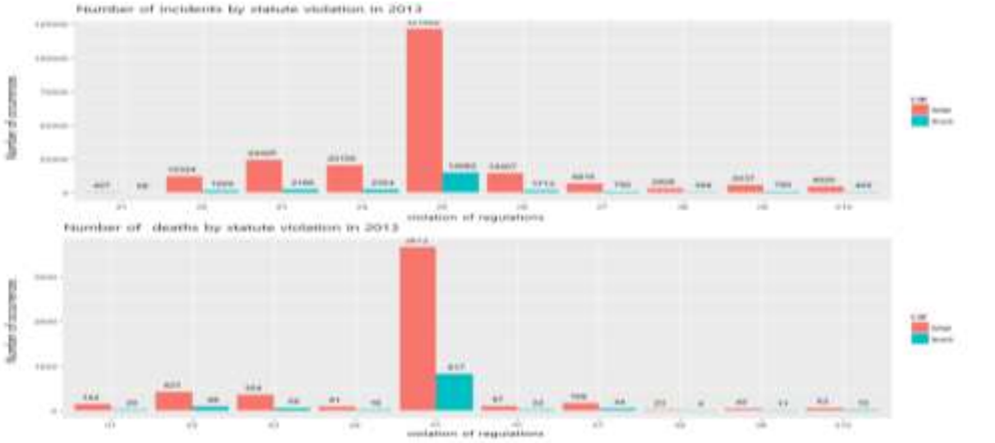


Fig.20. Number of incidents and deaths by statute violation in 2013

In Fig. 21, The highest number of accidents and violations of the death toll in 2014 was non-compliance with the mandatory safe driving (il5), with 14725 accidents and 734 deaths. More than a thousand cases will be a violation of the median line (il2), violation of the signal (il3), unsecured safety distance (il4), and violation of the intersection method (il6).

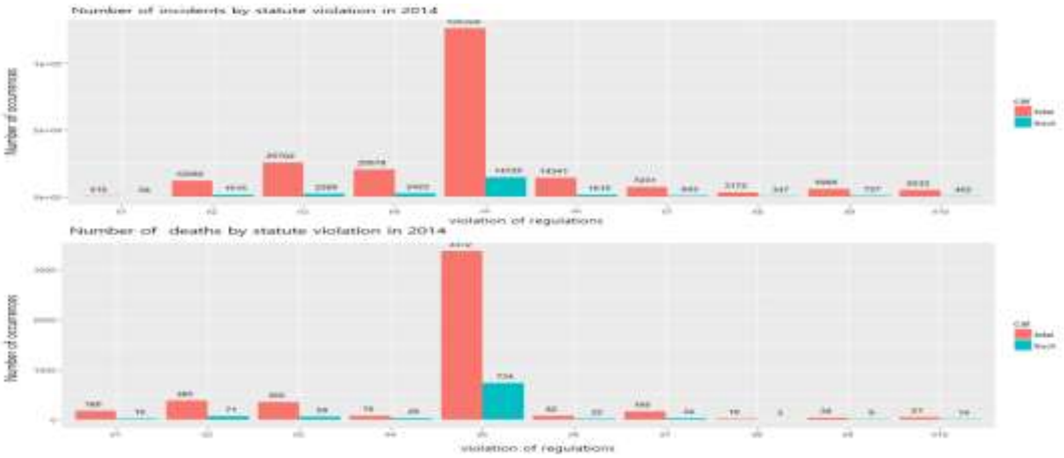


Fig.21. Number of incidents and deaths by statute violation in 2014

In Fig. 22, the highest number of accidents and violations of the death toll in 2015 was non-compliance with safe driving obligations (il5), with 1,5488 accidents and 676 deaths. More than a thousand cases will be a violation of the median line (il2), violation of the signal (il3), unsecured safety distance (il4), and violation of the intersection method (il6).

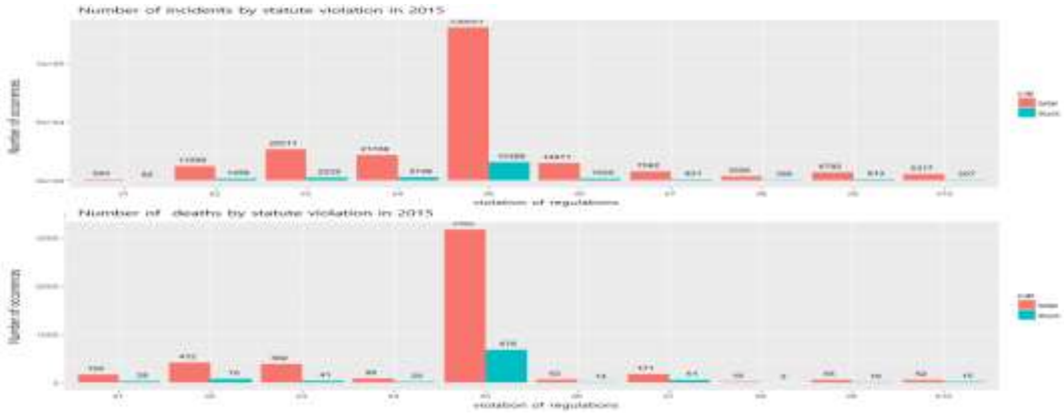


Fig.22. Number of incidents and deaths by statute violation in 2015

In Fig. 23, the highest number of accidents and violations of the death toll in 2016 was non-compliance with safe driving obligations (il5), with 13,644 accidents and 625 deaths. More than 1,000 cases will be violations of the median line (il2), signal violations (il3), unsecured safe distance (il4), and violations of intersection methods (il6).

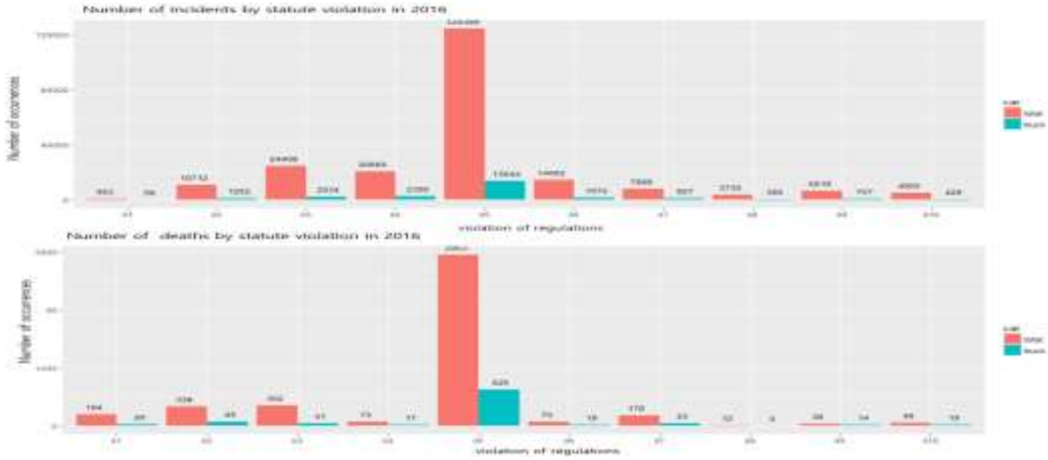


Fig.23. Number of incidents and deaths by statute violation in 2016

The number of accidents and deaths over the four years from 2013 to 2016 for non-compliance with safe driving obligations is uniquely high, with the same violations occurring every year, including more than 1,000 violations.

ANALYSIS RESULTS

I. *Suggestions for Effective On-site Crackdown*

Analysis of this study suggests major sections of roads handled by major regulators and police stations based on monthly, road type, number of occurrences, and number of casualties and deaths.

First of all, as traffic accidents continue to rise every April and September, and the number of accidents and casualties continues to rise in October, November, and then May, the peak of accidents, intensive crackdowns are needed to reduce damage to the country and personal property. In order to carry out such crackdown duties, on-site inputs shall be made to roads where accidents under the jurisdiction are frequent.

Over the four years, 10,000 road accidents and more than 300 deaths are recorded, followed by a special metropolitan road with more than 150 deaths, and 3,500 road accidents, and about 200 general countries, which can be used for major roads considering the type of roads in the jurisdiction.

Roads with more accidents than regular single roads, such as tunnel sections, intersections, and overpasses, are also a way to improve crackdowns on roads in cities and counties, special metropolitan roads, and general national roads. From 2013 to 2016, the highest number of accidents and deaths occurred at intersections, near intersections, and other single roads with ambiguous distinctions, and the number of deaths compared to the total number of vehicles also showed a double ratio. Similarly, it is proposed to increase the frequency of crackdowns at intersections in the jurisdiction and at several forms of single-route in site crackdowns.

Here, the types and types of roads can be identified in cargo vehicle accidents with higher deaths and injuries than other vehicles, thus improving as a crackdown method to reduce property damage and casualties.

II. *Proposals to Improve and Strengthen Traffic Laws for Violations of Regulations*

Legal punishment or improvement is needed for violations contrary to safe driving, which shows a large number of figures each year for road traffic laws that will be enforced on site crackdowns. The findings of this study suggest that such changes are necessary.

For non-compliance with safe driving obligations, which record the highest number of accidents and deaths each year, a statistical basis can be sought to strengthen punishment and sentence. Next, there are over 1,000 cases of violation of the centerline, signal violation, not securing a safe distance, and violation of the crossing method, All of these are parts that show the certainty of punishment as a result of the rational choice theory for intentional violations of truck drivers.

These offenses are caused by several fundamental causes of cargo vehicle accidents, and due to the deliberate negligence of drivers, the illegal release of the speed limit device, overcapacity of cargo and insufficient control over overload, and drunk driving leads to serious accidents.

Therefore, it is possible to improve and strengthen penalties for violations of laws analyzed to weaken the on-site monitoring power of truck drivers during actual crackdowns, frequent violation intentions, and change perceptions of traffic accidents and insensitivity to safety.

CONCLUSION

Most of the traffic accidents in the cargo vehicle start with the driver's failure to drive safely, leading to serious casualties in a one-car accident, unlike other models. The range of victims of truck accidents can cause damage not only to ordinary drivers but also to road traffic officials and police officers, which can constantly cause casualties, so it should be prevented by safe road traffic and preliminary traffic crackdowns to reduce casualties. While improving or newly implementing the Road Traffic Act through various cases and methods, this study proposed statistical-based efficient road traffic control methods by analyzing and visualizing open traffic accident data using statistical data, big data analysis systems, and statistical programming language R.

Individual police officers' experiences and existing methods of crackdowns can be used to identify monthly, road types and shapes, and analyze violations, including non-compliance, which is fundamentally responsible for cargo accidents, to expand effective crackdowns and safe driving.

Even after the proposed method in this study, statistical-based enforcement methods can be continuously proposed based on the information of various traffic accidents that are disclosed as the basis. It will have a major impact on improving and advancing various enforcement methods, and the revision method based on data analysis will ensure safety for many drivers and road users.

ACKNOWLEDGMENT

This research was supported by a grant(Development of technology for detecting and recognizing road signs and road surface marks based on deep learning for intelligent vehicles) from Gyeonggi-do researcher-centered R&D support project funded by Gyeonggi Province.

REFERENCES

[1] Yoon, H. J. and Yoo, J. E., "A Retrospective Analysis of Inpatients who Received Traditional Korean Medicine Treatment after a Traffic Accident during Pregnancy", The Journal of Korean Obstetrics and Gynecology, Vol. 34, No.2, 2021, pp.31-47.

[2] Kim, H. J., Jin, E. J., Oh, S. J. and Park, J. J., "A Study on the Analysis of Falling Objects from Vehicle Accidents through AI Learning-Based Keyword Analysis", Journal of Korean Society of Transportation, Vol. 39, No. 1, 2021, pp. 78-86.

[3] Kang, M. M., Kim, S. R., and Park, S. M., "Analysis and utilization of big data", Communications of the Korean Institute of Information Scientists and Engineers, Vol. 30, No. 6, 2012, pp.25-32.

[4] Kim, Y. K. and Hong, C. S., "A Study on Homomorphic Encryption Scheme using MapReduce", Journal of the Korea Institute of Information and Communication Engineering, 2016, pp.749-751.

[5] Lee, H. G., Kim, Y. W. and Kim, K. Y., "Implement of MapReduce-based Big Data Processing Scheme for Reducing Big Data Processing Delay Time and Store Data", Journal of the Korea Convergence Society, Vol. 9, No. 10, 2018, pp. 13-19.

[6] Lee, H. J., "Big Data Leverages the Hadoop Platform", The Korean Institute of Communications and Information Sciences, Vol. 29, No. 11, 2012, pp. 43-47.

[7] Kim, H. J., "Design and Implementation of an Efficient Web Services Data Processing Using Hadoop-Based Big Data Processing Technique", Journal of the Korea Academia-Industrial Cooperation Society, Vol. 16, No. 1, 2015, pp. 726-734.

[8] Shin, D. J., Lee, H. G., Eun, J. M., Kim, J. J. and Park, J. M., "Bigdata-based Anomaly Detection Technology in Web Services Environment", Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology, Vol. 8, No. 4, 2018, pp. 231-250.

[9] Yoon, C. H. and Kim, S. H., "A Tutorial on PLS Structural Equating Modeling using R: (Centering on) Exemplified Research Model and Data", Information Systems Review. Korea Society of Management Information Systems, Vol. 16, No. 3, 2014, pp. 89-112