Research Article

TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Transportation Research Record
2021, Vol. 2675(8) 494–505

DOI: 10.1177/03611981211999066
journals.sagepub.com/home/trr
SAGE

Analysis of Vehicle–Pedestrian Interactive Behaviors near Unsignalized Crosswalk

Byeongjoon Noh[1], Dongho Ka[1], David Lee[1], and Hwasoo Yeo[1]

Abstract

Road traffic accidents are a leading cause of premature deaths and globally pose a severe threat to human lives. In particular, pedestrians crossing the road present a major cause of vehicle–pedestrian accidents in South Korea, but we lack dense behavioral data to understand the risk they face. This paper proposes a new analytical system for potential pedestrian risk scenes based on video footage obtained by road security cameras already deployed at unsignalized crosswalks. The system can automatically extract the behavioral features of vehicles and pedestrians, affecting the likelihood of potentially dangerous situations after detecting them in individual objects. With these features, we can analyze the movement patterns of vehicles and pedestrians at individual sites, and understand where potential traffic risk scenes occur frequently. Experiments were conducted on four selected behavioral features: vehicle velocity, pedestrian position, vehicle–pedestrian distance, and vehicle–crosswalk distance. Then, to show how they can be useful for monitoring the traffic behaviors on the road, the features are visualized and interpreted to show how they may or may not contribute to potential pedestrian risks at these crosswalks: (i) by analyzing vehicle velocity changes near the crosswalk when there are no pedestrians present; and (ii) analyzing vehicle velocities by vehicle–pedestrian distances when pedestrians are on the crosswalk. The feasibility of the proposed system is validated by applying the system to multiple unsignalized crosswalks in Osan city, South Korea.

With the proliferation of advancements in information and communication technologies, many cities in the world are turning into smart cities. This means they are creating intelligent platforms, and improving the quality of life in the overall fields such as health care, transportation, environment, and safety. Despite such technological advancements, road traffic accidents globally pose a severe threat to human lives and have become a leading cause of premature deaths (*1*). According to statistics, approximately 1.2 million people are killed each year from road traffic accidents, and up to 50 million are injured worldwide. Pedestrians are exposed to the risk of death or serious injuries on the road for a variety of reasons, such as drivers failing to yield to pedestrians on the crosswalk (*2*, *3*). In particular, the Traffic Accident Analysis System in South Korea reported that accidents involving pedestrians crossing the road account for 40% of the total number of vehicle–pedestrian accidents in 2019 (*4*). Therefore, it is necessary to alleviate the causes of deaths of vulnerable road users.

There are various ways to reduce pedestrian injuries and to protect road users from traffic accidents, such as deploying speed bumps and speed cameras which suppress drivers and pedestrians engaging in risky or illegal behaviors, and operating 24 h closed-circuit television (CCTV) surveillance centers at administrative districts. In addition, a variety of studies have reported on examples of analyzing actual accident statistics to model the high fatality or injury rates of pedestrians and investigate their factors (*5*–*8*). However, these only used historical data of traffic accidents to reinforce a safe urban environment post facto. A different strategy is to pinpoint potential pedestrian risk scenes (e.g., near-miss collision) before traffic accidents occur, by analyzing behaviors of objects near crosswalks based on vision sensors. Along with the development of deep learning-based video processing technology, the use of vision sensors makes it easier to study potential traffic risks for a long time period. It also allows various analyses, such as evaluating the behavioral factors that pose a threat to pedestrians near crosswalks based on pedestrian–automobile interactions (*9*–*11*) and investigating pedestrian behavioral patterns (*12*). Further, based on their subtle interactions, it can

[1]Department of Civil and Environmental Engineering, Korea Advanced Institute of Science and Technology, Yuseong-gu, Daejeon, Republic of Korea

Corresponding Author:
Hwasoo Yeo, hwasoo@kaist.ac.kr

provide administrators with useful information for efficiently making decisions (*13, 14*).

However, since most of the CCTVs already deployed on roads record videos in oblique views, it is difficult to extract objects' behavioral features such as vehicle velocities and pedestrian positions with precision. Current studies rely on manual inspection to extract these features from video footage, and it requires much more cost and time when extended to the urban scale. Therefore, we should address these challenges by seeking to develop automated processes that generate useful data for pedestrian safety analysis across many sites in the city.

This paper proposes a new analytical system for potential pedestrian risk scenes based on video data from CCTVs deployed on the roads, especially unsignalized crosswalks. The proposed system has three objectives: (i) to automatically extract the traffic-related objects' behavioral features affecting the likelihood of potentially dangerous situations after detecting them in individual objects; (ii) to analyze their behavioral features and relationships among them by camera location; and (iii) to support administrators in making efficient decisions to improve the safety of road environments by providing the movement patterns of objects at unsignalized crosswalks. This study follows earlier experiments using sample footage, but improves on the methods and expands to a much larger dataset covering five cameras over two weeks. Objects' behavioral features, such as vehicle and pedestrian velocities and distances between them, are analyzed to show how they can be useful for monitoring the behavior of traffic on the road. The feasibility of the proposed system is confirmed by applying the prototype to multiple crosswalks in Osan city, South Korea. To conclude, the behaviors of vehicles and pedestrians in those five areas are compared by visualizing the changes of behaviors. This research is the first step in the authors' long-term effort to develop intelligence surveillance cameras in an advanced transportation safety system that helps to identify unsafe pedestrian environments in cities.

Materials and Methods

As represented in Figure 1, the proposed system works with four modules: (i) data sources; (ii) data structuring; (iii) extraction of objects' behavioral features; and (iv) analysis of potential pedestrian risk. First, the motion scenes are extracted from the video streams by using

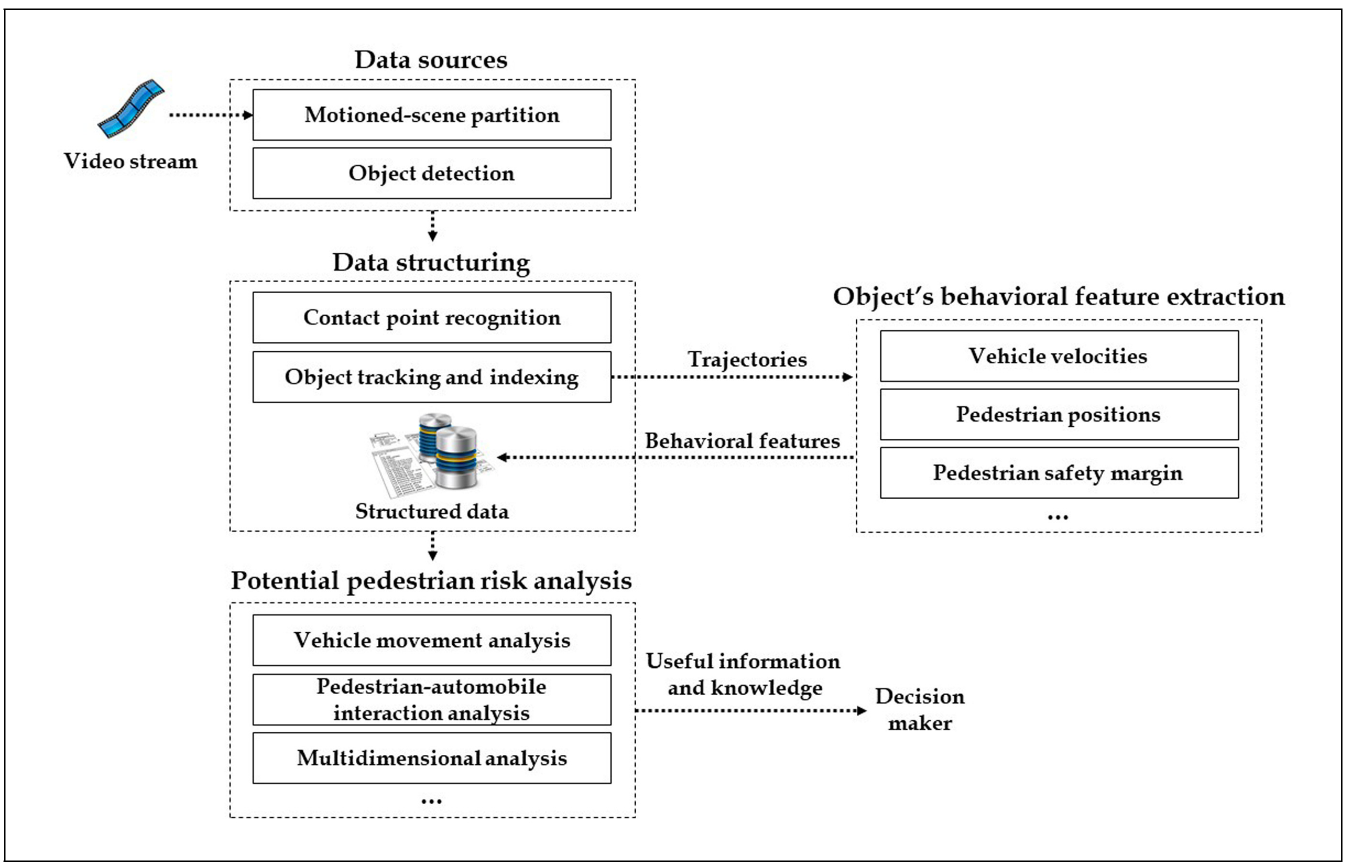

Figure 1. Overall architecture of the proposed system.

Table 1. Information of Characteristics of the Studied Spots

Spot code	Crosswalk width (m)	School zone	Speed camera	Fence	Red urethane	Raised crosswalk	No. of lanes	Speed limit (km/h)	Frame size	Frames per second
A	About 15	✓	X	✓	X	✓	2	30	1280 × 720	30
B	About 15	✓	X	✓	✓	✓	2	30	1280 × 720	30
C	About 15	X	X	X	X	X	2	30	1280 × 720	15
D	About 15	✓	X	✓	✓	X	2	30	1280 × 720	11
E	About 20	✓	X	X	✓	X	2	30	1920 × 1080	25

frame difference methods, and the traffic-related objects (vehicles and pedestrians, in this paper) are detected in the video footage by using deep learning. The data structuring module is a sequence of processes from recognizing contact points of objects to tracking and identifying objects in consecutive frames to convert the unstructured video data into structured data. Various behavioral features of objects are then extracted, and they are analyzed for potential traffic risk scenes. Finally, based on the results of these analyses with nuanced data on pedestrian–automobile interactions, decision-makers and administrators can obtain powerful and useful information to improve road environment safety.

Data Sources

In these experiments, video data were used from CCTV cameras over five unsignalized crosswalks in Osan city, South Korea (Spots A to E). In general, these CCTVs are intended to record and deter instances of street crime, but the chosen spots have different characteristics for safety facilities, such as having fences installed or the roads paved with red urethane. The information for each spot is described in Table 1, including road characteristics and recording metadata. In particular, roads designated as "school zone" are certain areas near facilities for children under 13 years, such as daycare centers, elementary schools and tutoring academies, to suppress risky behaviors and prevent traffic crashes involving vulnerable road users. In South Korea, penalties for breaking traffic rules or causing accidents in these areas are severe, with fines of up to 3,000 million won or life imprisonment (*15*).

All video frames were processed locally on a computer server deployed in Osan Smart City Integrated Operations Center, and the authors only obtained the processed trajectory data after removal of the original video data. This was to protect the privacy of anyone appearing in the footage. Future systems could employ internet-connected cameras that process images on-device in real time, and deliver only trajectory information back to servers.

Figure 2 shows the CCTV views being actually recorded at crosswalks A to E, respectively. In general, the unsignalized crosswalks have a small floating population with narrow road width. Thus, this study used video recorded on weekdays (from January 9 to January 28) during commuting hours (8:00–10:00 a.m. and 6:00–8:00 p.m.), when there is a relatively high floating population time.

Next, only view clips with moving vehicle activity, regarded as "scenes," were extracted from the whole video streams. In general, the video streams had long idle periods (when nothing had been captured), and motion scenes occasionally occurred, as described in Figure 3. To handle the huge volume of video data recorded in multiple areas for long time periods (about 280 h, 5 crosswalks × 14 days × 4 h per day), a simple and low computational complexity method was required.

In this study, a frame difference method was applied, a commonly used approach for recognizing the movements in a video (*16, 17*). As seen in Figure 4, we can observe an intensity for the pixel positions which have changed in two consecutive frames by calculating the pixel-based difference between them. Then, the "motion" is detected by comparing it with a threshold, as follows:

$$|P[I(t)] - P[I(t+1)]| > Threshold \tag{1}$$

where an image obtained at the time t denoted by $I(t)$, and $P[*]$ means the pixel value in the image.

With the motioned frames identified by frame difference, the motion scenes can be grabbed, and then the traffic-related objects (e.g., vehicles and pedestrians, in this paper) are detected and segmented by using a deep learning model to extract the video clips (scenes) as a series of frames from when the vehicle enters the scene to when it exits. When multiple vehicles are in one video clip, each vehicle's presence in the scene is treated as a separate scene. Pedestrians may or may not be present in each scene. In this experiment, a mask R-CNN model, an extension of the faster R-CNN model, was used in the Detectron 2 platform as implemented by Facebook AI Research (FAIR) (*18*). The deep learning model used is pre-trained by the Microsoft common objects in context (MS COCO) image dataset (*19*), and has a much faster processing time than those of other existing platforms (*20*). Note that the primary goal of this step is to detect only the traffic-related objects defined in this paper. Since this pre-trained model has sufficient high accuracy

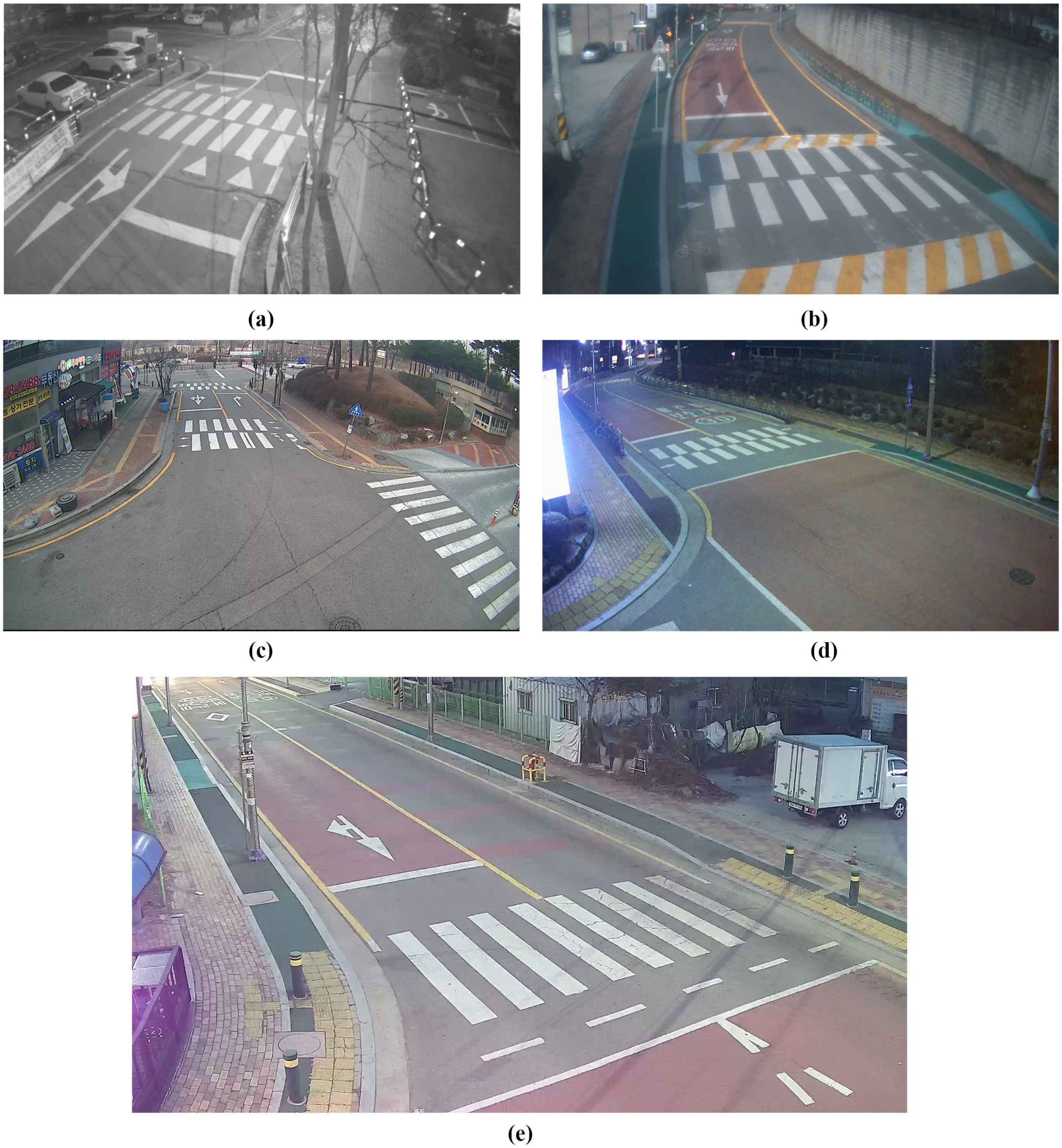

Figure 2. Actual closed-circuit television views in: (*a*) Spot A, (*b*) Spot B, (*c*) Spot C, (*d*) Spot D, and (*e*) Spot E.

for detecting these objects in the video footage, additional training was no longer required for these purposes. As a result, in total about 27,000 scenes were extracted from the entire video dataset involving traffic-related objects, as seen in Table 2. The majority of scenes captured only passing vehicles, called "car-only scenes," containing a scene where at least one vehicle was in camera view, while "interactive scenes" involved at least one pedestrian and one car in the scene at the same time.

Data Structuring

This section describes the process of converting video data (unstructured) into a structured RDB (relational

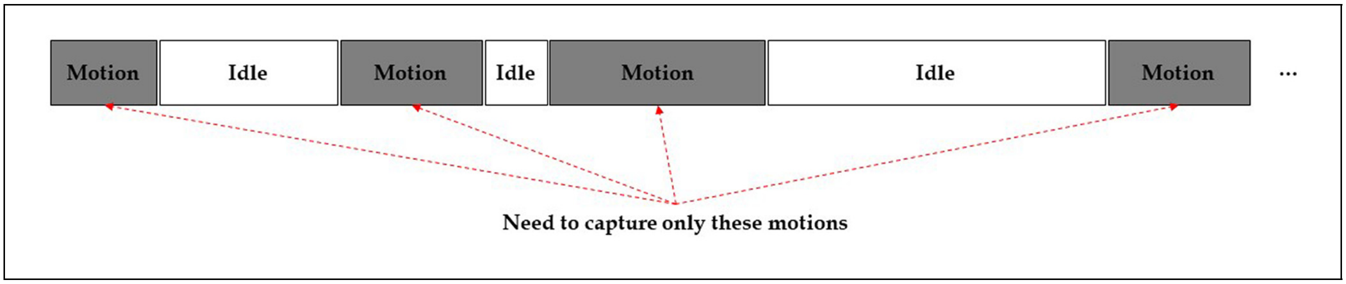

Figure 3. Composition of the actual video stream.

Figure 4. Example of frame difference between foreground and background.

Table 2. Numbers of Extracted Scenes after Preprocessing

	Number of extracted scenes (after preprocessing)	
Spot code	Car-only scenes	Interactive scenes
A	4,633	2,322
B	2,481	1,395
C	3,533	2,079
D	1,843	1,002
E	4,572	3,203

database) schema, which consists of three steps: (i) contact points recognition; (ii) perspective transform; and (iii) object tracking and indexing.

Since road-deployed CCTV cameras typically record from oblique views, it is difficult to reliably extract objects' behavioral features such as velocities, distances, and positions. To address this, reference points were devised for determining their features, called the "contact points" of each object. The contact points are obtained by perspective-transforming ground tip points into an overhead view. Ground tip means a point on the ground directly underneath the center of the object in the oblique view. In this experiment, the ground tips of a vehicle and a pedestrian are located directly under the center of the front bumper and on the ground between the pedestrian's feet, respectively. To determine ground tips for them, the object mask and central axis line of vehicle lane were identified, and then perspective transform was conducted using the "transformation matrix" function in OpenCV library. The more detailed procedures for these are explained in the authors' previous studies (*21*, *22*). This paper focuses on the object tracking and indexing process. To extract the behavioral features of a single object, it is important to distinguish each object captured in consecutive and multiple frames. Thus, the threshold and minimum distance methods based on a modified Kalman filter were used to track objects from frame to frame. A Kalman filter has been applied in a wide range of engineering applications such as computer vision, transportation, and robotics to trace and identify objects in consecutive frames, and to estimate the unknown current or future states of objects in the video (*23*, *24*). To calculate the next position of object, it performs two steps repeatedly: (i) estimating the next object's parameters by using current values such as speeds and positions; and (ii) updating these parameters based on the prior predicted values and information obtained about the next position.

The tracking and indexing algorithm used in this study consists of two parts: (i) estimating the candidate points based on smoothing; and (ii) assigning objects in the next frame by calculating and comparing distances. First, when the new objects in next frames (called origin target objects) appear, each trajectory is smoothed, including the new objects, by the Kalman filter. Then, all distances between the origin target objects and the estimated target

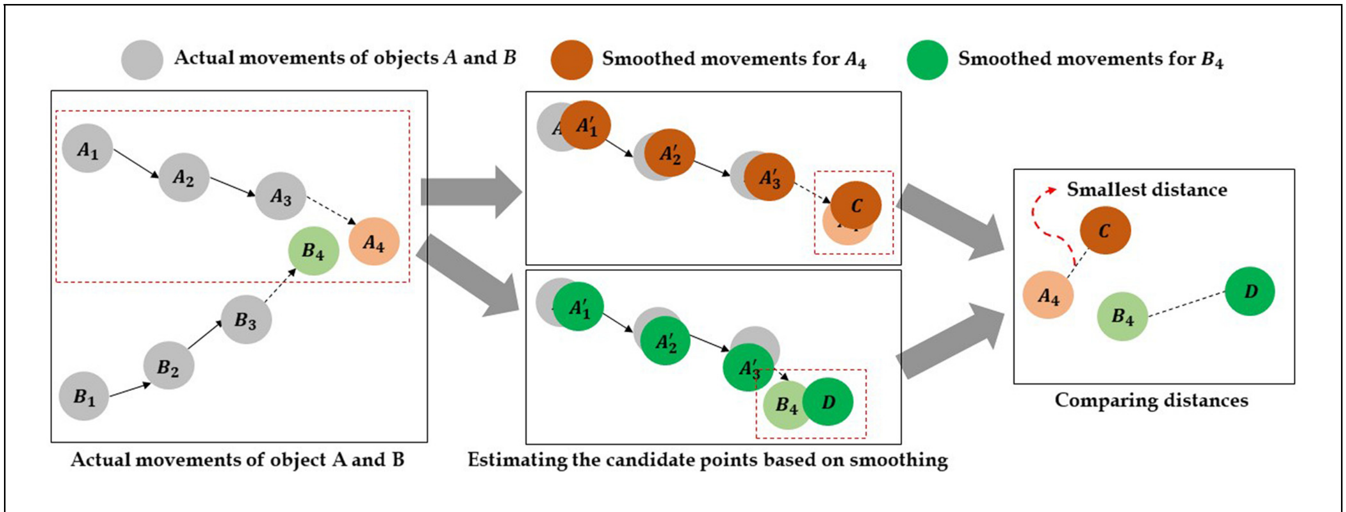

Figure 5. Example of finding correct trajectory of object A.

objects are calculated, and the closest object is chosen. For example, Figure 5 illustrates the overall process of finding the trajectory of object A. There are the actual trajectories of two objects in a scene, A (ordered by $A_1, A_2, A_3,$ and A_4) and B (ordered by $B_1, B_2, B_3,$ and B_4). Assume that we have already connected the trajectories of objects A and B from frames 1–3 so far. To correctly assign A_4 to the trajectory of object A, the trajectories through frames 1–3 were smoothed with each object in frame 4. The estimated points are denoted with apostrophes, such as A'_1, A'_2, and A'_3, and the estimated target objects are denoted C and D. Then, the distances between the origin target objects and the estimated target objects are compared, and the target object is assigned with the smallest distance from the estimated target object to the trajectory of object A. This process is repeated until the last frame in the scene.

With the trajectories of these objects, various behavioral features that affect the potential risk can be extracted from video, such as instantaneous velocity and acceleration in one frame, and the sequence of velocities and distances measured at different times of the scene. To analyze the potential risks from unstructured-type video data, the structured database schema was designed (see Figure 6), and the extracted features stored in the database. This experiment was conducted using four behavioral features for each object: vehicle velocity, pedestrian position, vehicle–pedestrian distance, and vehicle–crosswalk distance. The methods of extraction are described in the next section.

Automated Object Behavioral Feature Extraction

This section describes the behavioral feature extraction and the automated processes from the recognized trajectories. With the obtained trajectories, the various behavioral features can be extracted, such as vehicle velocity, acceleration, position, pedestrian velocity, vehicle–pedestrian relative position, and so on. However, this paper presents four selected features: vehicle velocity, pedestrian position, vehicle–pedestrian distance, and vehicle–crosswalk distance, and describes the methods of extraction in detail.

Vehicle Velocity. Vehicle velocity is a basic measurement that can signal potential risky situations, and it is a significant risk factor for crash severity in vehicle–pedestrian collisions. The speed limit in all the studied locations was 30 km/h; if there are many vehicles moving over the speed limit at any point, it contributes to high potential risk at that location. To extract the velocity from an assembled trajectory, simply divide the distance an object has moved between consecutive frames by the time interval between these frames.

The object's distance in pixel distance is converted into a real-world distance unit in meters. Pixel distance between point i in the j-th and $(j+1)$-th frames in x-y plane, $D_{pixel}\left(point_i^j, point_i^{(j+1)}\right)$, is computed by Euclidean distance, and it is converted by using the pixel-per-meter constant (P). To infer P, the pixel length of the crosswalk is divided by its actual length; the actual lengths of crosswalks were measured in field visits. For example, if the length of the crosswalk is 20 m, and the pixel length is 800 pixels, 1 m is about 40 pixels (= 800/20). Meanwhile, the time interval in the video is the frame interval and must also be converted to real-world seconds. The time conversion constant (F) is derived by dividing the skipped frames by FPS (frames per second). For example, if the video is recorded at 28 FPS, and

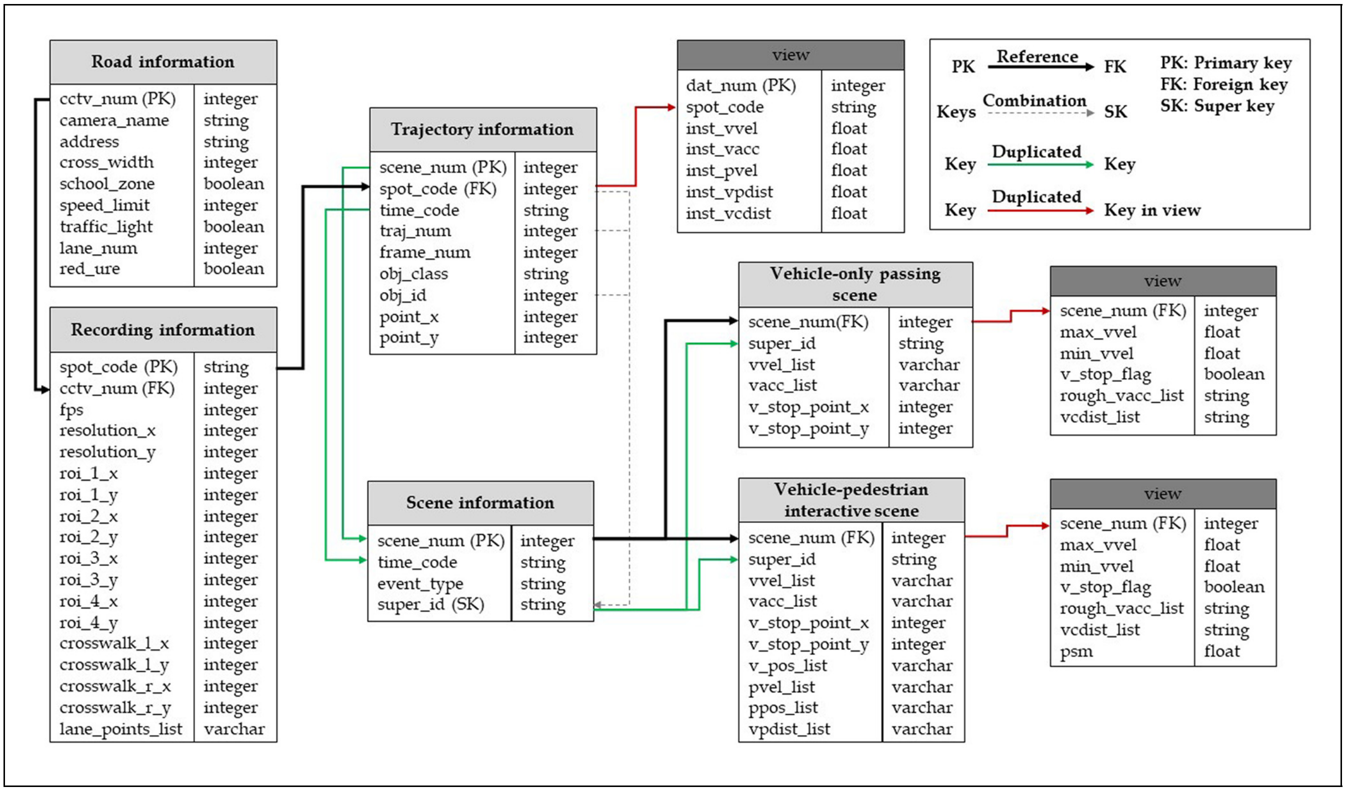

Figure 6. Database schema for extracted behavioral features.

every 5th frame is sampled, F is equal to 5/28. Finally, the i^{th} object's velocity in the j^{th} and $(j+1)^{th}$ frames can be calculated as follows:

$$velocity_i^{j,(j+1)} = \frac{D_{pixel}\left(point_i^j, point_i^{(j+1)}\right)}{F * P}(m/s) \quad (2)$$

The unit is finally converted into kilometers per hour, and applied to all the frames in the scene.

Pedestrian Position. The position of the pedestrian is also crucial information in evaluating the potential accident risks. A pedestrian on the road, even when cars are moving at slow speed, may be more at risk than a pedestrian on the sidewalk when cars are moving at high speed. In this paper, pedestrian positions are categorized into four areas: "sidewalk," "crosswalk," "crosswalk influenced area" (CIA), and "road." CIA refers to the road area adjacent to the crosswalk, which pedestrians often enter while crossing the road (*25–27*). Some people cross the street on the correct zebra-marked zone but others walk around this area, so both of these two areas (crosswalk and CIA in Figure 7) were considered as crosswalkable areas. In this paper, CIA is encompassed with a buffer of ~3 m on either side of the crosswalk, as illustrated in Figure 7.

Vehicle–Pedestrian Distance and Vehicle–Crosswalk Distance. These features refer to the physical distance between vehicle and pedestrian and vehicle and crosswalk by frame, respectively. Distance between vehicle and nearest pedestrian i and pedestrian p in the j^{th} frame is as follows:

$$V - P\ distance_{i,p}^j = \frac{D_{pixel}\left(vehicle_i^j, pedestrian_p^j\right)}{P}(m) \quad (3)$$

Meanwhile, the distance between vehicle and crosswalk is measured from the crosswalk line closest to the vehicle. When the vehicle is on the crosswalk, the distance is 0:

$$V - C\text{distance}_i^j = \begin{cases} \frac{D_{pixel}\left(\text{vehicle}_i^j, \text{crosswalk}\right)}{P}(m), \text{if a vehicle is out of crosswalk} \\ 0, \text{otherwise} \end{cases} \quad (4)$$

Experiments and Results

With the proposed system, each object's behavioral features can be extracted automatically. This system was applied to multiple unsignalized crosswalks in Osan city, South Korea. Statistical analysis was then conducted on

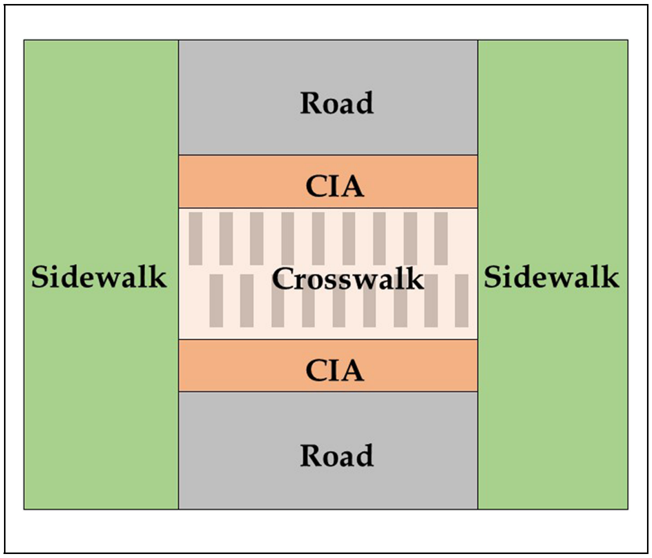

Figure 7. Categorized pedestrian positions.
Note: CIA = crosswalk influenced area.

the extracted features and behavioral characteristics of objects at these crosswalks.

Experimental Design

This section describes the experimental design for analysis. Before the main analyses, it briefly explains the results of data preprocessing and statistical analyses with histograms of average vehicle velocities. Next, in the experiment, behavioral features were analyzed with two scenarios: (i) investigating vehicle velocity changes near the crosswalk when there are no pedestrians present; and (ii) analyzing vehicle velocities by vehicle–pedestrian distances when pedestrians are on the crosswalk.

Table 3 shows the statistical values of average vehicle velocities, and Figure 8 illustrates histograms of them in each spot.

The maximum average velocities are in the range of about 59.2 to 79.7 km/h, and minimum values range from 2.2 km/h to 9.4 km/h. As illustrated in Figure 8, the overall values are skewed right since many cars pass slowly near the crosswalk. The speed limit for all areas with school zones is 30 km/h. When considering that the mean values in all spots are near or under the regulation speed, these are reasonable values.

In general, cars tend to move faster when there are no pedestrians present, and slow down when there are pedestrians. We can observe these tendencies by separating the average vehicle velocities into car-only scenes and interactive scenes, as seen in Figure 9. In all spots, the velocities in interactive scenes are lower than those in car-only scenes. In this experiment, we can observe that the vehicles traveled slowly in Spot C (no fence area). It is the authors' guess that drivers may be concerned that pedestrians will suddenly jump out into the road at such locations.

Table 3. Information on Vehicle Velocities in Each Spot

Spot code	Max. (km/h)	Min. (km/h)	Mean (km/h)
Spot A	79.7	4.1	18.1
Spot B	68.1	2.2	22.3
Spot C	63.9	9.4	14.0
Spot D	59.2	3.3	21.4
Spot E	70.2	7.4	31.8

Note: Min. = minimum; Max. = maximum.

Results

Analyzing Changes in Vehicle Velocities near Unsignalized Crosswalk. This section analyzes vehicle velocity changes near the crosswalk in each spot when there are no pedestrians present (car-only scenes). Since each scene has a different number of frames, it is difficult to compare the sequence of velocities across scenes, besides calculating average velocities. Thus, the velocities of the vehicle at certain distances from the crosswalk (vehicle–crosswalk distance) are measured, as illustrated in Figure 10.

The baselines are positioned from the starting line of the crosswalk at 10 m, 5 m, 3 m, and 0 m (starting line of the crosswalk). In the experiment, velocity changes that occur before, during, and after the vehicles pass the crosswalk are observed in each area. Figure 11 shows the velocity measurements at each area, averaged across scenes at that spot.

In these figures, the x-axes mean the distance from crosswalks; −10 m, −5 m, and −3 m are 10 m, 5 m, and 3 m before entering the crosswalks, respectively. The moment the vehicles enter the crosswalks is 0 m; and 3 m and 5 m are 3 m and 5 m after passing the crosswalks, respectively. The blue area is inside the crosswalks.

The average vehicle velocity measurement in all areas never exceeded the speed limit (30 km/h), except for Spot E. In detail, it can be observed that vehicles in spots A and B slowed down faster than those in other areas just before crosswalks, and still decrease velocity while passing the crosswalks. On the other hand, the vehicles in Spot E decreased velocity before entering crosswalks, but rapidly accelerated even as they entered the crosswalks. It is hypothesized that the raised crosswalks (speed bump type crosswalks) seen in Figure 2, *a* and *b*, were forcing drivers to alleviate the impact by rapidly decelerating. By contrast, in Spot C, the vehicles decelerated at a distance of about 3–5 m rather than just in front of the crosswalk, thereafter increasing again. It can be interpreted that when the vehicles approach the crosswalk, they slow

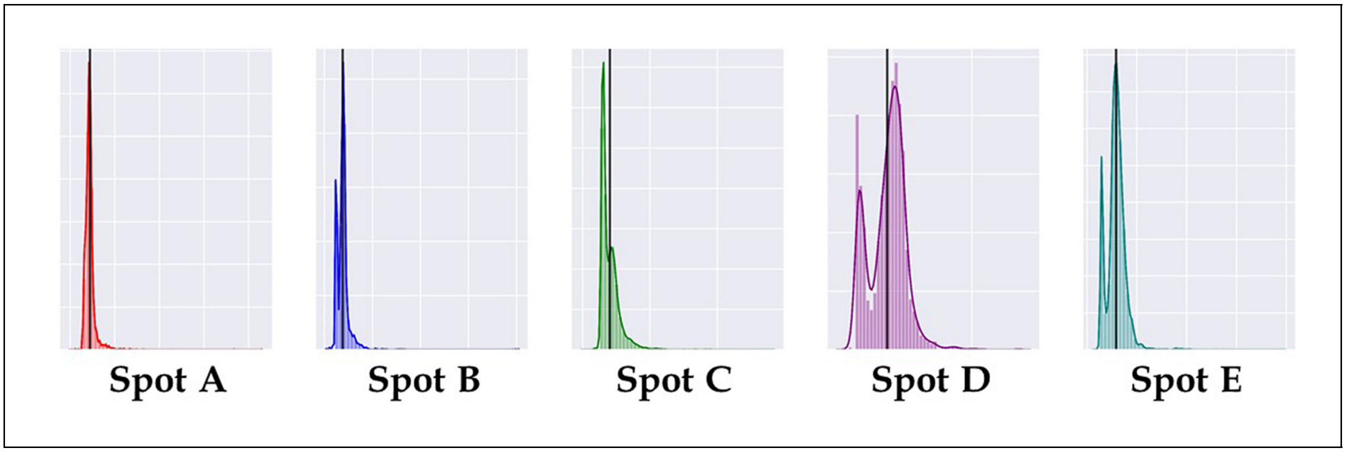

Figure 8. Distributions of average vehicle velocities in each spot.

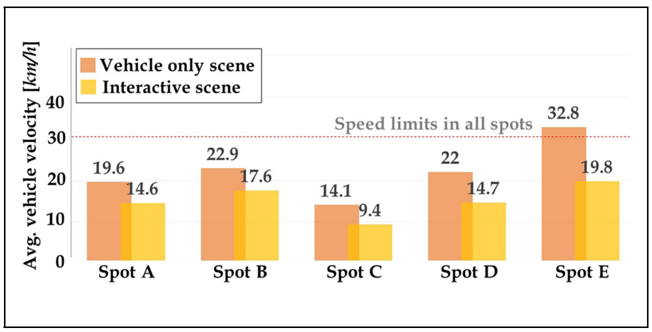

Figure 9. Average vehicle velocities in each spot by scene type.

down in advance and accelerate again after checking that there is no pedestrian.

To summarize, it can be observed that in some spots, most vehicles slowed down before entering crosswalks. In particular, the raised crosswalks appeared to force rapid deceleration when deployed, ensuring that vehicles entered those crosswalks at lower speeds. Moreover, it is the authors' guess that the ideal safe speed change pattern is to reduce the speed before entering the crosswalk, and then pass it after checking for the presence of pedestrians, as at Spot C. Therefore, it is necessary to analyze the behavior characteristics of vehicles when there are pedestrians present.

Analyzing Vehicle Velocity Changes by Vehicle–Pedestrian Distances near Unsignalized Crosswalk. This section analyzes the changes of vehicle velocity by the distance between vehicles and pedestrians. The primary goal of this experiment is to investigate how vehicles behaved when pedestrians were on the crosswalks. At signalized crosswalks, most vehicles would not enter the crosswalk when pedestrians were present because of the traffic signal. Thus, scenes at unsignalized crosswalks when pedestrians were

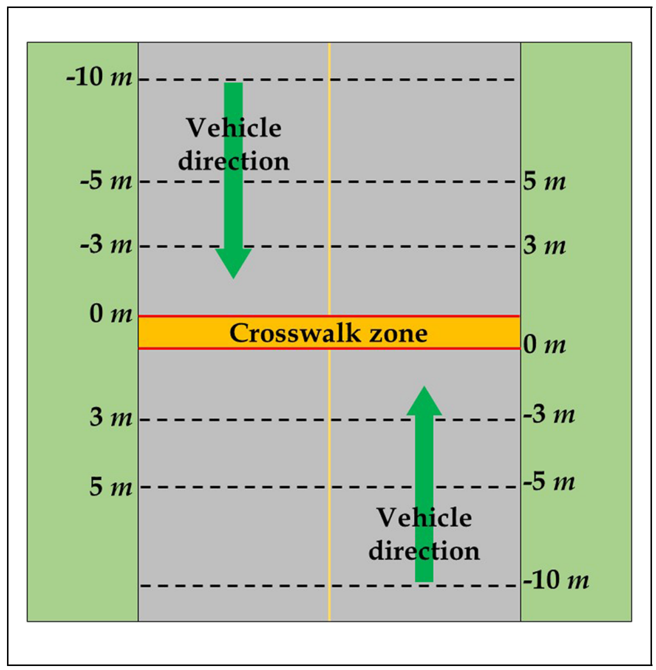

Figure 10. Baselines set at a certain distance from the crosswalk.

in the crosswalks or CIAs were targeted. Vehicle velocities were extracted from frames when the vehicle–pedestrian distances were between 2 m and 16 m, and the results are presented in Figure 12.

In these figures, we can observe a "lasso-like" pattern in all areas. Vehicles dramatically slowed down when approaching a pedestrian, then increased their speeds just before reaching their closest distance. In Spot C, although this area was not designated as a school zone, vehicles moved slowly near pedestrians. One of the reasons, the authors speculate, is that this crosswalk is near a busy residential area with a high density of pedestrians crossing during the commuting hours when the data

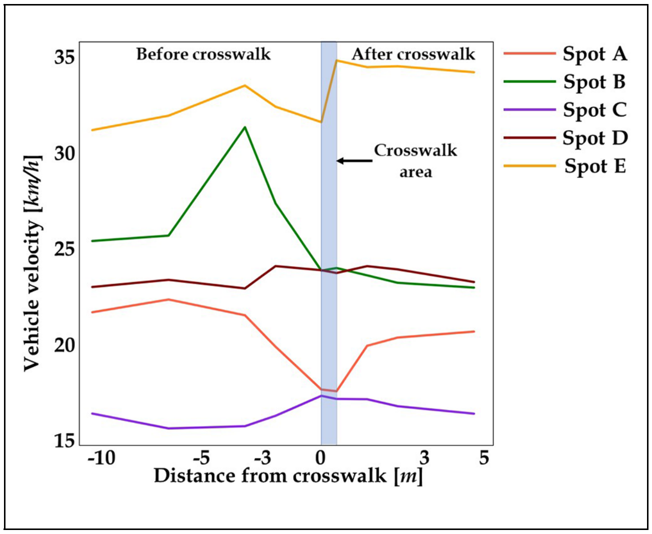

Figure 11. Results of average velocity changes by distance from the crosswalk in each spot.

were collected. Meanwhile, vehicles in Spots A and B decelerated rapidly, and did not increase speed again, unlike other areas (except Spot C), similar to their behaviors in car-only scenes. It is hypothesized that the raised crosswalks influenced this behavior, along with the presence of pedestrians. Most vehicle speeds in all spots stayed under the speed limit, but speeds in Spot E were the highest; in the scene when the vehicles were closest to pedestrians, their average speed exceeded 20 km/h.

To summarize, common patterns can be seen in all the unsignalized spots, as cars decelerated when approaching crossing pedestrians, and accelerated again just before passing the pedestrians. These patterns look like a lasso. In the raised crosswalks (Spots A and B), vehicles decelerated rapidly and did not increase their speeds again, unlike other spots. Finally, overall speeds were low in areas with a high floating population.

Discussion

The proposed analytical system in this paper had three main objectives: (i) to automatically extract the traffic-related objects' behavioral features which affect the likelihood of potentially dangerous situations after detecting

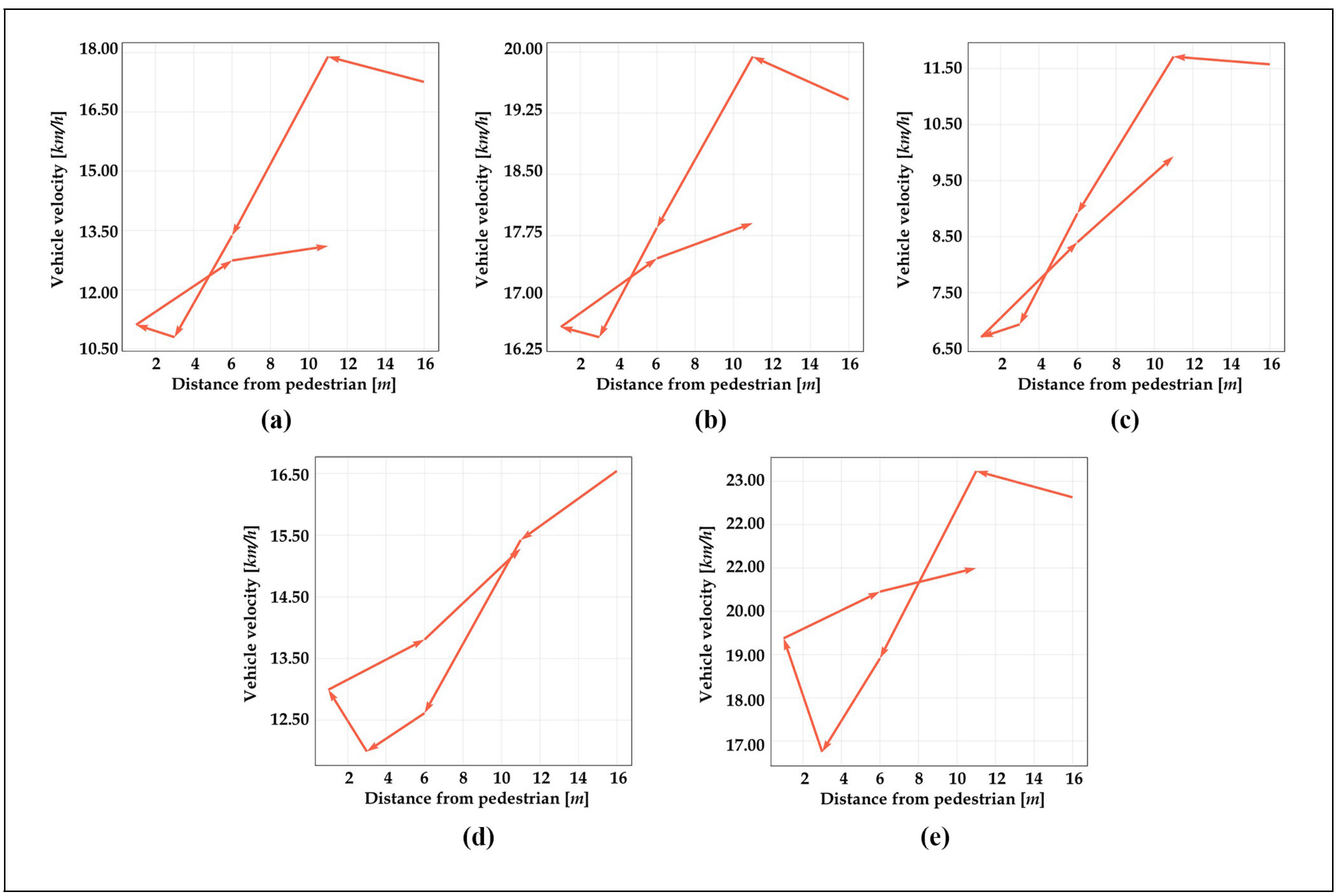

Figure 12. Results of average velocity changes by distances between vehicles and pedestrians in: (*a*) Spot A, (*b*) Spot B, (*c*) Spot C, (*d*) Spot D, and (e) Spot E.

them into individual objects; (ii) to analyze their behavioral features and relationships among them by camera location; and (iii) to support administrators making efficient decisions to improve the safety of the road environment. Based on the authors' previous studies (*21*, *22*), the scale was expanded to larger urban areas with more CCTVs available, as well as longer daily study periods, and various behavioral features which affect the potential pedestrian risks were analyzed. In addition, unstructured video footage was converted into structured-typed datasets with the design of the database schema to analyze potential traffic risks.

These experiments used only four behavioral features affecting potential traffic risks, and conducted two analyses: (i) changes of vehicle velocity near the crosswalk for car-only scenes; and (ii) vehicle velocity by vehicle–pedestrian distances when pedestrians on the crosswalk for interactive scenes. As a result, changes of vehicle velocity near unsignalized crosswalks were visualized, and compared for each area. In addition, the patterns of vehicle movements according to road environments and pedestrian presence can be obtained. With the proposed system, it is possible to handle a variety of behavioral features, and further apply it to scrutinize various potentially dangerous situations depending on the purposes. Because interactions between vehicles and pedestrians occur more frequently than actual crashes, this approach can prevent actual traffic accidents proactively with richer, denser, and more consistent perspectives on the safety of these environments.

Conclusions

This study proposed a newly designed potential risk analytical system for supporting administrators to make efficient decisions to improve the safety of road environments. The proposed system works automatically with one sequential process for extracting motion scenes, detecting objects, and extracting their features based on video footage. This paper covers only the behavioral feature extraction and parts of potential risk analyses. The core methodologies are (i) to recognize precise trajectories of each object despite the oblique view video footage; (ii) to design a database scheme for storing various behavioral features in the form of structured data converted from unstructured video; and (iii) to analyze them to determine how vehicles move in each spot and environment. The feasibility of the proposed system was confirmed by developing it with the Detectron 2 platform and OpenCV libraries, applying it to actual video data from unsignalized crosswalks in Osan city, South Korea, and visualizing and interpreting these scenes by each spot.

The proposed analytical system elicits useful information that can be obtained without additional expenditure in cameras, based on the trajectories of vehicles and pedestrians. The proposed system does not by itself identify the best control or traffic calming measures to prevent actual traffic accidents and crashes. However, it can provide traffic engineers and urban designers with clues to help further investigation using data analysis techniques, as in the authors' ongoing works. The next step is to implement a complete decision support system providing greater contexts and comprehensive insights to help guide administrators making decisions, such as deploying safety structures on roads, using objective data analysis.

Acknowledgment

The authors thank Osan Smart City Integrated Operations Center for providing the CCTV video dataset.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: BN and DK; data collection: BN; analysis and interpretation of results: BN, DL and HY; draft manuscript preparation: BN, DK and HY. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT(NRF-2017R1A2B2002329), and by Korea Ministry of Land, Infrastructure and Transport (MOLIT) as "Innovative Talent Education Program for Smart City."

References

1. Ho, G. T., Y. P. Tsang, C. H. Wu, W. H. Wong, and K. L. Choy. A Computer Vision-Based Roadside Occupation Surveillance System for Intelligent Transport in Smart Cities. *Sensors*, Vol. 19. No. 8, 2019, p. 1796.
2. Lytras, M. D., and A. Visvizi. Who Uses Smart City Services and What to Make of It: Toward Interdisciplinary Smart Cities Research. *Sustainability*, Vol. 10, No. 6, 2018, 1998.
3. Akhter, F., S. Khadivizand, H. R. Siddiquei, M. E. E. Alahi, and S. Mukhopadhyay. IoT Enabled Intelligent Sensor Node for Smart City: Pedestrian Counting and

Ambient Monitoring. *Sensors (Switzerland)*, Vol. 19, No. 15, 2019, p. 3374.
4. Traffic Accident Analysis System. http://taas.koroad.or.kr/. Accessed July 13, 2020.
5. Olszewski, P., P. Szagała, M. Wolański, and A. Zielińska. Pedestrian Fatality Risk in Accidents at Unsignalized Zebra Crosswalks in Poland. *Accident Analysis & Prevention*, Vol. 84, 2015, pp. 83–91.
6. Haleem, K., P. Alluri, and A. Gan. Analyzing Pedestrian Crash Injury Severity at Signalized and Non-Signalized Locations. *Accident Analysis & Prevention*, Vol. 81, 2015, pp. 14–23.
7. Gandhi, T., and M. M. Trivedi. Pedestrian Protection Systems: Issues, Survey, and Challenges. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 8, No. 3, 2007, pp. 413–430.
8. Gitelman, V., D. Balasha, R. Carmel, L. Hendel, and F. Pesahov. Characterization of Pedestrian Accidents and an Examination of Infrastructure Measures to Improve Pedestrian Safety in Israel. *Accident Analysis & Prevention*, Vol. 44, No. 1, 2012, pp. 63–73.
9. Ka, D., D. Lee, S. Kim, and H. Yeo. Study on the Framework of Intersection Pedestrian Collision Warning System Considering Pedestrian Characteristics. *Transportation Research Record*, 2019. 2673: 747–758.
10. Murphy, B., D. M. Levinson, and A. Owen. Evaluating the Safety in Numbers Effect for Pedestrians at Urban Intersections. *Accident Analysis & Prevention*, Vol. 106, 2017, pp. 181–190.
11. Kadali, B. R., and P. Vedagiri. Proactive Pedestrian Safety Evaluation at Unprotected Mid-Block Crosswalk Locations Under Mixed Traffic Conditions. *Safety Science*, Vol. 89, 2016, pp. 94–105.
12. Ke, R., J. Lutin, J. Spears, and Y. Wang. A Cost-Effective Framework for Automated Vehicle-Pedestrian Near-Miss Detection Through Onboard Monocular Vision. *Proc., IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Hawaii, 2017, pp. 25–32.
13. Fu, T., L. Miranda-Moreno, and N. Saunier. A Novel Framework to Evaluate Pedestrian Safety at Non-Signalized Locations. *Accident Analysis & Prevention*, Vol. 111, 2018, pp. 23–33.
14. Fu, T., W. Hu, L. Miranda-Moreno, and N. Saunier. Investigating Secondary Pedestrian-Vehicle Interactions at Non-Signalized Intersections Using Vision-Based Trajectory Data. *Transportation Research Part C: Emerging Technologies*, Vol. 105, 2019, pp. 222–240.
15. The National Law Information Center in Korea. http://www.law.go.kr/. Accessed May 20, 2020.
16. Liu, H., J. Dai, R. Wang, H. Zheng, and B. Zheng. Combining Background Subtraction and Three-Frame Difference to Detect Moving Object from Underwater Video. *OCEANS 2016*, Shanghai, April 10, 2016, pp. 1–5. IEEE.
17. Sengar, S. S., and S. Mukhopadhyay. Moving Object Detection Based on Frame Difference and W4. *Signal, Image and Video Processing*, Vol. 11, No. 7, 2017, pp. 1357–1364.
18. Facebook AI Research. https://ai.facebook.com/. Accessed April 12, 2020.
19. COCO Dataset. http://cocodataset.org/#home. Accessed September 3, 2019.
20. Github. https://github.com/facebookresearch/detectron2. Accessed January 10, 2020.
21. Noh, B., W. No, and D. Lee. Vision-Based Overhead Front Point Recognition of Vehicles for Traffic Safety Analysis. *Proc., 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, Singapore, October 8, 2018, pp. 1096–1102.
22. Noh, B., W. No, J. Lee, and D. Lee. Vision-Based Potential Pedestrian Risk Analysis on Unsignalized Crosswalk Using Data Mining Techniques. *Applied Sciences*, Vol. 10, No. 3, 2020, p. 1057.
23. Haroun, B., L. Q. Sheng, L. H. Shi, and B. Sebti. Vision Based People Tracking System. *International Journal of Computer and Information Engineering*, Vol. 13, No. 11, 2019, pp. 582–586.
24. Sun, X., H. Yao, and S. Zhang. A Refined Particle Filter Method for Contour Tracking. In *Visual Communications and Image Processing*, International Society for Optics and Photonics, Vol. 7744, Huangshan, China, 2010, p. 77441M.
25. Figliozzi, M. A., and C. Tipagornwong. Pedestrian Crosswalk Law: A Study of Traffic and Trajectory Factors that Affect Non-Compliance and Stopping Distance. *Accident Analysis & Prevention*, Vol. 96, 2016, pp. 169–179.
26. Fu, T. *A Novel Approach to Investigate Pedestrian Safety in Non-Signalized Crosswalk Environments and Related Treatments Using Trajectory Data*. Doctoral dissertation. McGill University, 2019.
27. Sisiopiku, V. P., and D. Akin. Pedestrian Behaviors at and Perceptions Towards Various Pedestrian Facilities: An Examination Based on Observation and Survey Data. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 6, No. 4, 2003, pp. 249–274.