Pedestrian Detection Using HOG Feature-Based Cascade Classifier with Vehicle Black-Box Camera for Supporting Driver Assistance in Urban Road Environments

Jong Bae Kim

***Abstract*—In this paper, we propose a method to detect pedestrians in real time from road images obtained from a black-box, which is a vehicle image recording camera, and to automatically provide pedestrian appearance information to the driver. To detect a pedestrian in the input road image, the proposed method applies a pedestrian detector using a cascade learning device based on the histogram of oriented gradients (HOG) feature information. The pedestrian detector uses the cascade learning device to extract feature information about the pedestrian area based on the histogram description feature information for pedestrian learning. The pedestrian detector detects the candidate pedestrian areas, and the final pedestrian area is detected through the pedestrian verification process. The results of applying the proposed method to urban road images indicate that the accuracy of detection is approximately 93%.**

***Index Terms*—Pedestrian detection, HOG, ADAS, vehicle black-box camera, cascade learning.**

I. Introduction

According to the global status report on road safety by the World Health Organization [1], approximately 1.3 million people die each year on roads, and 20–50 million people are involved in traffic accidents. More than half of the pedestrian traffic accidents are caused by the careless driving of a vehicle driver than by pedestrian carelessness such as unauthorized crossing. Therefore, in order to prevent pedestrian traffic accidents, various techniques for reducing the carelessness of the driver are suggested. Recently, institutional and technical measures have been proposed for the reduction of pedestrian traffic accidents worldwide. However, owing to the increase in the number of vehicle registrations, the effect of reducing pedestrian traffic accidents is decreased.

In the case of Korea, the number of vehicle registrations continues to increase. As shown in Fig. 1, according to the Ministry of Land, Infrastructure and Transportation in Korea, the number of registered cars is continuously increasing [1]. The number of registered vehicles in 2017 is approximately 22.53 million, up 3.3% from the previous year. In addition, the number of driving license holders over 60 years of age (Fig. 2) reached 4.6 million by 2016, an increase of approximately 11% over the previous year [2]. Consequently, the number of pedestrian traffic accidents is also increasing owing to the increase in the number of vehicles and the number of elderly drivers. According to the Traffic Accident Analysis System of Korea's Road Traffic Corporation, an average of 50,000 pedestrian traffic accidents (Fig. 3) occurred from 2013 to 2017. Pedestrian traffic accidents are caused by various factors such as carelessness of road users (driver, pedestrian), violation of traffic regulations, poor road environment, and weather. However, in the event of a traffic accident, an accident investigation is carried out in order to evaluate the cause, such as traffic signal violation, violation of pedestrian protection obligation, or unauthorized crossing. Therefore, there is a limit to the identification of the cause of the accident for securing practical pedestrian safety, and accordingly the establishment of measures to reduce accidents. In other words, while efforts are being made to prevent accidents through diversified methods such as the use of various safety technologies [3]-[6] to prevent traffic accidents, expansion of road infrastructure, and driver safety and accident prevention education, pedestrian traffic accidents still account for more than 46,000 traffic accidents every year.

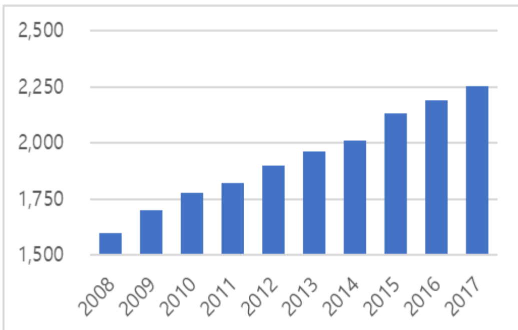


Fig. 1. Korean vehicle registration numbers (2008–2017, ten thousand) [1].

Therefore, safety support technologies for reducing pedestrian traffic accidents should be supplemented, and further, automobile pedestrian collision accident prevention technology is required. An economical method to prevent pedestrian traffic accidents is to use image processing. In recent years, there has been a tendency to install video recording devices in vehicles to identify possible accidents. Using the images captured by this device, object detection can be performed through image processing, and the results can be effectively utilized in supporting safe driving

Manuscript received April 5, 2019; revised January 17, 2020. This work was supported with the support of the "Ministry of Education, Science and Technology" (NRF-2016R1D1A1B03931986).

J. B. Kim is with the Department of Computer and Software, Sejong Cyber University, Seoul, 04992, S. Korea (e-mail: jb.kim@sjcu.ac.rk).

doi: 10.18178/ijmlc.2020.10.2.947

automatically. In other words, it is necessary to apply a method for detecting pedestrians in real time in the road images acquired from the black-box installed in the vehicle and providing the pedestrian appearance information to the driver, so as to bring them to the driver's attention.

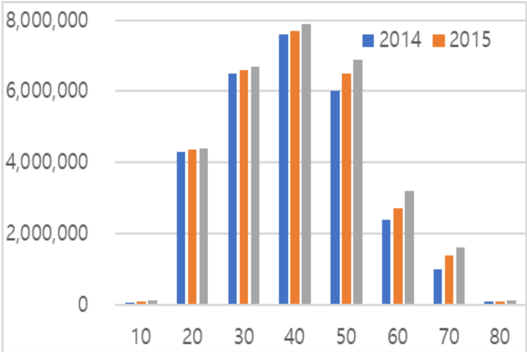


Fig. 2. Status of driving license holders by age group [7].

In order to provide safe driving support technology for vehicles, the detection of objects located in all directions is mainly performed by attaching additional hardware sensors that use radar or ultrasonic waves. Object detection using these sensors is widely applied in vehicles because of their robustness and sensitivity to environmental changes (fog, rain, dust, etc.). However, these hardware-based sensors are considered as additional vehicle options as they are expensive, and are mostly installed in high-end luxury vehicles. In addition, these omnidirectional object detection sensors are installed in the front bumper section of the vehicle, and thus have various disadvantages such as a malfunction of the sensor due to a slight collision accident between vehicles. As an alternative to sensor-based object detection during vehicle driving, vision-based object detection technology is gaining popularity [8]. Currently, most Korean drivers are installing a black box in their vehicles. Some commercial vehicles (taxi, bus, etc.) require that a video recording device be installed for safety. However, a vehicle black box is still used to identify the cause after an accident. If a black box is equipped with an image processing function to detect and recognize objects of interest in the acquired road images and provide relevant information to the driver, it can be used as a safety support system for accident prevention.

Therefore, in this paper, we propose a method to detect a pedestrian from a road image obtained from a black box. In the proposed method, the pedestrian area is manually selected in advance in order to acquire and learn the pedestrian image. Then, the histogram of oriented gradient (HOG) feature is extracted from the selected pedestrian area and decomposed into multiple resolutions to reduce the dimensions of the feature information. Then, a pedestrian detector is generated using the cascade learning proposed by Viola-Jones [9]-[12]. The feature of the cascade learning machine is to apply the binary tree classifier technique to extract characteristic feature information from learning samples. The feature information is similar to the pedestrian area among the input feature information, and the feature information is excluded from the learning object repeatedly. Thus, a robust pedestrian detector can be generated by extracting only the optimal pedestrian feature information.

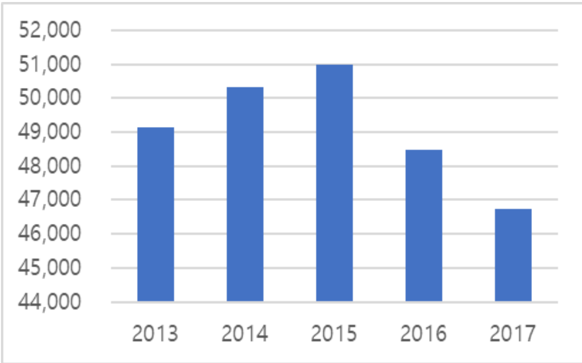


Fig. 3. Number of pedestrian traffic accidents (2013–2017) [13].

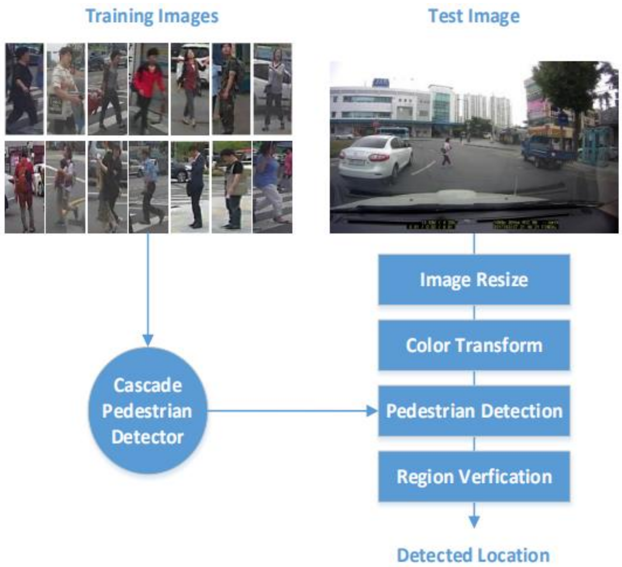


Fig. 4. Flowchart of pedestrian detection processing.

II. Proposed Method

As shown in Fig. 4, the proposed method is divided into the pre-processing step, the pedestrian detection step, and the area verification step. In the pre-processing step, the size of the HD color image obtained from the black box is reduced, the image is converted to grayscale, and the visibility of the image is improved through the histogram smoothing process of the brightness value. In the proposed method, pedestrian images were obtained from city roads. Pedestrian images for learning excluded the overlapping images. Most of the pedestrians in the experimental environment are images waiting in the crosswalk. And for real-time processing, the size of input image is reduced by using bi-linear interpolation. Feature information is extracted from pedestrian images for learning. The extracted feature information is input to the learner to generate a detector for pedestrian detection. Next, the pedestrian areas are detected using the cascade pedestrian detector (Fig. 5) generated in the learning step, and then the final pedestrian areas are detected using the prior knowledge about the pedestrian information in the pedestrian verification step. The knowledge information about the pedestrian is used to detect the final pedestrian area by using a vertically long rectangle, size information, position information, and pedestrian template matching similarity information. In the learning phase, a multi-level classifier is used to generate an optimal pedestrian classifier that can classify the HOG feature of the pedestrian and other areas.

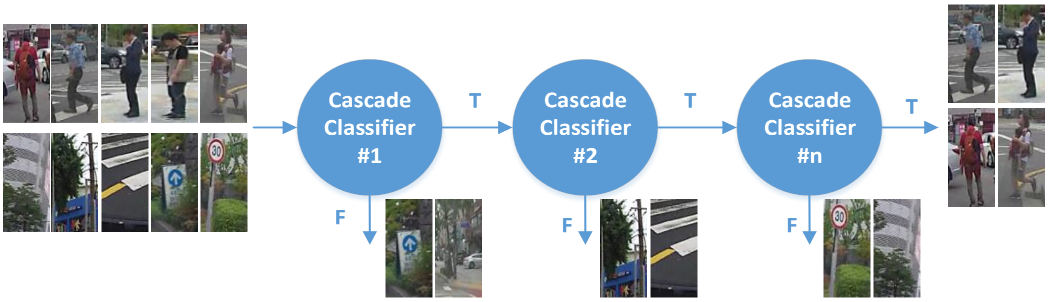


Fig. 5. Cascade pedestrian learning machine.

After reducing the size of the image from 1920 × 1080 color pixels to 900 × 505 pixels and converting to a grayscale image, the image is input to the cascade pedestrian detector. The result is then output as the value between the square position information corresponding to the pedestrian area and the similarity information [0, 1] between the learned pedestrian. Then, the positional information of the pedestrian in the black box image, the size of the detected area, and the aspect ratio information are compared with the pedestrian template in which the pedestrian image is acquired in the learning step. Thereafter, the final pedestrian region is verified by determining the pedestrian similarity value derived from the cascade learning device, the size information of the pedestrian area, the ratio information, the position information, and the similarity information through template matching.

A. *Pre-processing Step*

In this step, the size of the road image obtained from the black box is reduced and noise is removed to detect the pedestrian. In the proposed method, the nearest interpolation method is applied to reduce the image size and a denoising convolutional neural network (DnCNN) is applied to remove noise [14]. The image of the outdoor road environment includes noise such as light, weather, and irregular reflection of the object surface. Especially, the road images obtained from the black box during driving include various types of noise due to the influence of the vehicle motion, road pollution, illumination, etc., and the image quality is deteriorated. Therefore, in order to effectively detect a pedestrian in the road image, it is necessary to apply a step of removing the noise included in the road image in advance. In general, noise included in the road image is mainly composed of white noise. Therefore, it is necessary to remove the white noise included in the image obtained from the road. Fig. 6 shows the results of adding white Gaussian noise with a variance of 0.02 to the original road image and removing the noise. The upper left image in Fig. 6 is the original image, and the upper right image is Gaussian noise. And the lower image is an image from which noise is removed.

Table I shows the peak signal-to-noise ratio (PSNR) values and the structural similarity (SSIM) index values that measure the structural similarity. A large PSNR value indicates that image quality is improved owing to less noise.

In addition, the closer the SSIM value is to 1, the higher is the similarity to the original image. As a result of removing the noise included in the image by applying the DnCNN method, it can be seen that the PSNR and SSIM values are improved close to the original image.

Fig. 6. Noise removal results: original image, noise image, denoise image (top, right, down).

TABLE I: NOISE IMPROVEMENT RESULTS

Image / Measure	Original image vs. noise image	Original image vs noise removal image
PSNR	17.7774	25.6050
SSIM	0.4257	0.7723

B. *Candidate Pedestrian Region Detection Step*

After the pre-processing step, candidate pedestrian regions are detected using a pedestrian detector previously learned in order to detect pedestrians in the reduced-size road image with noise removed. Multi-step pyramid shrinking images are generated through the multi-resolution decomposition process in the input image. The reduced image generation level SL is determined using Eq. (1). In Eq. (1), depth is 8, I is the size of the original image, and p is a pedestrian region of size 100 × 41. When the size of the input image I is 505 × 900, SL is generated in 19 steps. Fig. 7 shows the result obtained by decomposing the input image into multiple stages according to the value obtained using Eq. (1). As shown in Fig. 7, multi-level image decomposition extracts feature information that uniquely expresses feature information in each of the decomposed steps.

$$SL = floor\left(\text{depth} \times \log_2\left(\min\left(\frac{I}{p}\right)\right) + 1\right) \tag{1}$$

Fig. 8 shows the feature information extracted from *SL* 1-level. The feature information includes color, gradient, gradient size, and 6-ways gradient. As a result, feature maps for the color, the edge gradient size, and the edge gradient direction are extracted at each decomposition step.

In order to detect the candidate pedestrian area from the road image, the input road image is decomposed into multi-step resolution images. After extracting feature information, as shown in Fig. 8, from the low-resolution decomposed images, we compare the similarity with the pedestrian feature information learned in advance. In the proposed method, a binary decision tree method is applied for region similarity comparison. Pedestrian candidate regions are detected in the input image. The pedestrian area is detected by the binary discrimination method that contains the most feature information about the pedestrian areas learned in the pedestrian learning phase.

Fig. 7. Multilevel decomposition result image (*SL*=19).

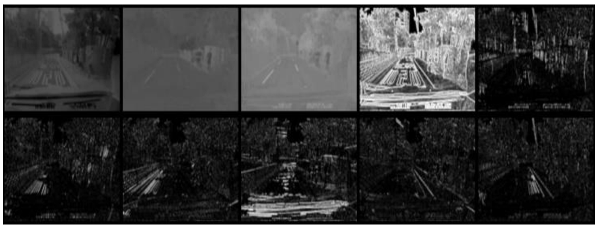


Fig. 8. Feature information extraction result.

Fig. 9. Pedestrian detection results in multilevel decomposition images.

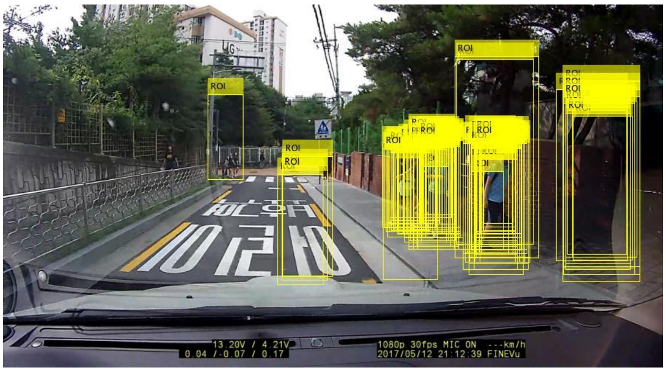


Fig. 10. Pedestrian candidate region detection results using pedestrian detector.

Fig. 9 shows the results of detecting the pedestrian region in each of the multilevel decomposed images, and Fig. 10 shows the detection results of candidate pedestrian regions. Fig. 9 shows the result of detecting the pedestrian area by decomposing the input image into 19 multi-level images and applying the pedestrian detector generated through learning from each of the decomposed images. And, Fig. 10 shows the result of projecting and displaying the pedestrian areas detected in each multi-level image through the original input image through Fig. 9.

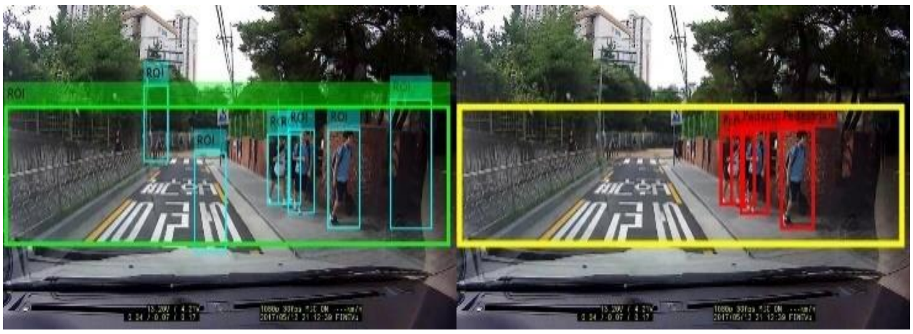


Fig. 11. Candidate pedestrian area (left) through the pedestrian detector and final detection results (right).

C. Pedestrian Region Verification Step

In order to verify whether the pedestrian is included in the detected candidate pedestrian regions, information such as position, size, color, and pattern is used. The location information used for the pedestrian verification is an average value that defines the position of the pedestrian in the learning images in advance. The size and color information are also obtained for the average horizontal / vertical size of pedestrian regions. Fig. 11 shows the results of extracting the candidate pedestrian region detected by the pedestrian detector and the final detection result, and Fig. 12 shows the results of the detected pedestrian regions.

Fig. 12. Pedestrian regions detected from pedestrian detector.

III. Experimental Results

In order to evaluate the proposed method, a road image of size 1920 × 1080 pixels obtained at various time zones and locations was used as input. The proposed pedestrian detection method was tested using Windows 10 OS (Quad Core 3.4 GHz, 32 G, GPU 1080ti) and a Matlab program. The pedestrian images used for pedestrian training were 320 normalized images of size 74 × 32 pixels. The parameter for cascade learning was set as FalseAlarmRate of 0.05 and the learning execution phase was set to 7 in total.

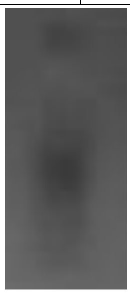
Fig. 13 shows the results of detecting the pedestrian region by applying the proposed method. Table II shows the values used in the pedestrian verification step to detect only the pedestrian regions in the candidate pedestrian regions detected by the multilevel pedestrian detector. The similarity comparison between the detected candidate pedestrian region and the pedestrian template in Table III is performed for detecting pedestrian regions from candidate pedestrian regions. When the similarity value is 0.4 or more, it is classified as a pedestrian region.

The experimental results show that the pedestrian regions are included in the candidate pedestrian regions detected by the pedestrian detector and some of them are detected as the pedestrian region by the street lights, trees, etc. The results show that most of the falsely detected pedestrian regions are filtered at the pedestrian verification step. However, in the template matching process, some non-pedestrian regions are detected as pedestrian regions. The experimental results show that the proposed method takes approximately 0.24 s per frame and the detection performance of the pedestrians is approximately 93%. The pedestrian detection was evaluated as correct when the degree of overlap between the pedestrian region defined by the ground-truth method and that detected by the proposed method was greater than 50%.

Fig. 13. Results of the proposed pedestrian detection method.

The experimental results show that most of the pedestrian regions are correctly detected. However, there was a case where non-pedestrian regions were also mis-detected as pedestrian regions. In order to resolve this problem, it is necessary to acquire more pedestrian learning images and to create a robust learning model for detecting pedestrians through the learning process. The proposed method exhibits an excellent performance of pedestrian detection in a city road image and demonstrates that real-time processing is possible. In addition, it can be seen that the proposed method is applicable to an intelligent safe driving support system. In future work, we will conduct research to generate a robust pedestrian detector to reduce pedestrian misdetection.

TABLE II: PARAMETERS FOR VERIFYING THE PEDESTRIAN REGION

Values	min	max
Horizontal size	9	49
Vertical size	17	91
H/V Ratio	0.979	19.2935
Brightness Percentage of Total / region	0.2783	2.6368
Template	 74×32 pixel size	

IV. CONCLUSION

In this paper, we proposed a pedestrian detector generated using a multi-level learning machine based on HOG feature information. We also proposed a method to detect the pedestrian region in real time using the pedestrian detector.

Experimental results showed that the detection rate is approximately 93% and the processing time is approximately 2.4 sec per frame. The proposed method can be applied to an intelligent safe driving support system. Including the proposed method in the functioning of a vehicle black-box can contribute to the prevention of pedestrian traffic accidents.

CONFLICT OF INTEREST

The author declares no conflict of interest.

AUTHOR CONTRIBUTIONS

The author read and approved the final manuscript.

ACKNOWLEDGMENT

This work was supported with the support of the "Ministry of Education, Science and Technology" (NRF-2016R1D1A1B03931986).

REFERENCES

[1] e-Nara Standard: Automobile Registration Status. [Online]. Available: http://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=1257
[2] National Police Agency, Driver's license holder. [Online]. Available: https://www.data.go.kr/dataset/15029967/fileData.do
[3] A. Brunetti, D. Buongiorno, G. F. Trotta, and V. Bevilacqua, "Computer vision and deep learning techniques for pedestrian detection and tracking: A survey," *Neurocomputing*, vol. 300, pp. 17-33, 2018.
[4] S. Bota and S. Nedesvchi, "Multi-feature walking pedestrians detection for driving assistance systems," *IET Intell. Transp. Syst.*, vol. 2, no. 2, pp. 92-0194, 2008.
[5] M. Jeong, B. C. Ko, and J. Y. Nam, "Early detection of sudden pedestrian crossing for safe driving during summer nights," *IEEE Trans. on CSVT.*, vol. 27, no. 6, pp. 1368-1380, 2017.
[6] L. Zhao and C. E. Thorps, "Stereo and neural networks-based pedestrian detection," *IEEE Trans. on Its*., vol. 1, no. 3, pp. 148-154, 2000.
[7] Status of driving license holders by age group. . [Online]. Available: https://www.dataplanet.co.kr/user/chart/view?chart_id=TDCBG
[8] J. B. Kim, "Efficient vanishing point detection for advanced driver assistance system," *Advanced Science Letter*, vol. 23, no. 5, pp. 4115-4118, 2017.
[9] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE CVPR*., vol. 1, pp. I-1, 2001.
[10] J. S. Choi and J. B. Kim, "A study on the traffic light detection in real-time for ADAS," *Proceedings of KICS*., p. 877, 2018.
[11] J. S. Choi, J. B. Kim, "A Study on the road crosswalk detection using Haar feature and Cascade object detector for advanced driver assistance system," *Proceedings of KICS*., 20B-19, 2018.
[12] M. Nieto, G. Vélez, O. Otaegui, S. Gaines, and G. V. Cutsem, "Optimising computer vision based ADAS: vehicle detection case study," *IET Intelligent Transport Systems*, vol. 10, no. 3, pp. 157-164, 2016.
[13] Traffic Accident Analysis System, Pedestrian Traffic Accident Status. [Online]. Available: http://taas.koroad.or.kr/sta/acs/gus/selectPdstrnTfcacd.do?menuId=WEB_KMP_OVT_MVT_TAS_PDT
[14] C. Cruz, A. Foi, V. Katkovnik, and K. Egiazarian, "Nonlocality-reinforced convolutional neural networks for image denoising," *IEEE Signal Processing Letters*, vol. 25, no. 8, pp. 1216-1220, 2018.

Jong Bae Kim received the Ph. D. in the Department of Computer Engineering from Kyungpook National University, S. Korea, in 2012. He is currently a professor of the Department of Computer Software, Sejong Cyber University, S. Korea. His research interests are in the area of computer vision and pattern recognition, advanced driving assistance system, ITS, artificial intelligent.