Traffic Mirror Detection and Annotation Methods from Street Images of Open Data for Preventing Accidents at Intersections by Alert

Da Li
Fukuoka University
Fukuoka, Japan
lida@fukuoka-u.ac.jp

Hikaru Hagura
Kyoto Sangyo University
Kyoto, Japan
daigakuyhgrhika@gmail.com

Taichi Miyabashira
Fukuoka University
Fukuoka, Japan
tl191319@cis.fukuoka-u.ac.jp

Yukiko Kawai
Kyoto Sangyo University
Kyoto, Japan
kawai@cc.kyoto-su.ac.jp

Shintaro Ono
Fukuoka University
Fukuoka, Japan
onoshin@fukuoka-u.ac.jp

Abstract

In recent years, research on autonomous driving has been actively pursued in the automotive industry. In Japan, a bill to revise the ***Road Traffic Act*** *law regarding level 4 autonomous driving is passed in 2022, indicating a proactive approach toward autonomous driving. In light of these trends, improving safety during car travel has become an even more important challenge than before. Especially, intersections with poor visibility have been one of the major causes of traffic accidents, and improving safety at such intersections is an essential element in enhancing safety during car/bicycle travel. In this study, we aim to develop a system capable of identifying blind spots reflected in traffic mirrors by analyzing worldwide open data such as road images (e.g., Google Street View) and road information (e.g., OpenStreetMap). By Annotating the critical points from open data, the application could provide alerts to pedestrians and vehicles, enhancing safety in the vicinity of these blind spots. Specifically, we initially investigate the most effective deep learning model for detecting traffic mirrors. Additionally, we analyze the location information of traffic mirrors from geospatial data and road image data to construct a traffic mirror distribution map. Furthermore, we intend to equip bicycles with smartphones to track and detect the trajectories of pedestrians and vehicles reflected in these traffic mirrors.*

1. Introduction

Improving vehicle driving safety is becoming an increasingly important challenge [7]. In particular, intersections and curves with poor visibility are major causes of traffic accidents due to the increased risk of collisions, making it an essential element for improving safety while driving cars and riding bicycles [10]. Various technologies and initiatives are being implemented by automobile manufacturers, local governments, and research institutions to address this issue [2] [15] [4].

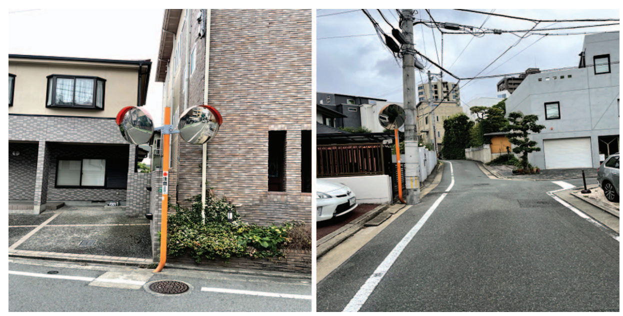

Figure 1. The image data of the road including the traffic mirror(s).

In Japan, approximately 25% of all accidents occur at intersections, with many of these accidents attributed to driver negligence and the road environment [12]. The narrow roads and numerous obstructions, such as buildings, walls, and plants, create blind spots at branch roads. To address this issue, road safety mirrors (traffic mirrors) have been installed at intersections, T-junctions, and sharp curves in Japan to enhance safety during driving [13] [16]. Such locations can have poor visibility, requiring extra caution for oncoming vehicles and pedestrians from both sides. In contrast, these traffic mirrors allow drivers to see the situation on the side road by reflecting vehicles and pedestrians in the mirror, thereby reducing the incidence of blind spots (as shown in Figure 1).

The main objective of this study is to develop a sys-

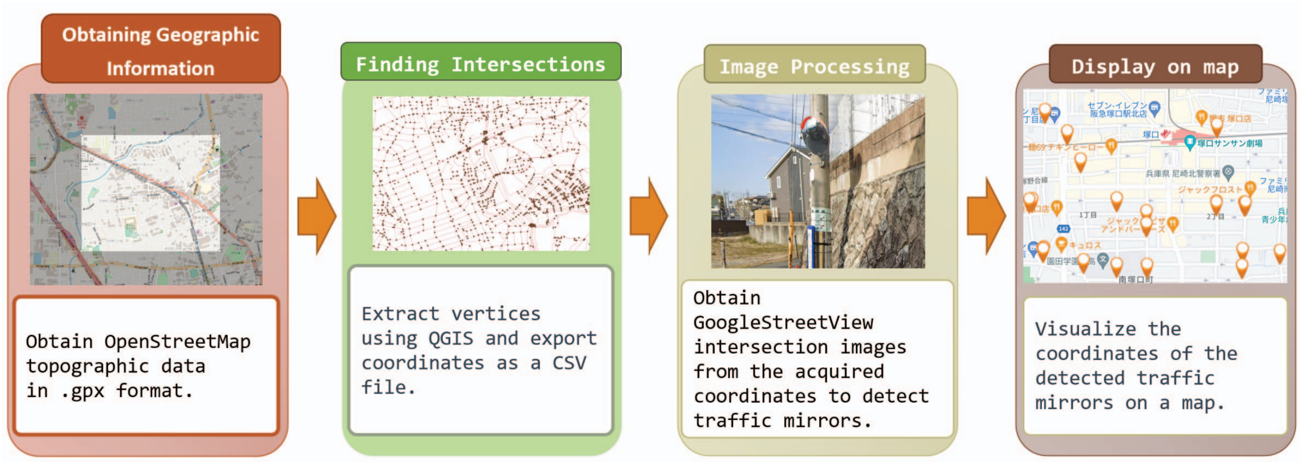

Figure 2. Overview of the traffic mirror distribution map construction method.

tem that utilizes smartphones equipped with built-in cameras and mounted on bicycles to analyze *road images* and *road information*. The goal is to identify intersections with traffic mirrors and detect pedestrians and vehicles present in blind spots on sidewalks, as reflected by these traffic mirrors. By mapping the locations of traffic mirrors, the system can emit an alert sound to warn pedestrians and vehicles, even in cases where the mobile camera fails to detect the objects in the mirrors. The ultimate aim is to enhance safety by providing timely alerts to pedestrians and vehicles regarding potential hazards.

To achieve this objective, we initially investigate the optimal deep learning model for detecting traffic mirrors. This step is crucial since there are currently no pre-trained models available for traffic mirror detection. We employ three types of object detection models and two optimization algorithms to train the models for detecting traffic mirrors on the road. The detection performance of each model is then compared to evaluate their effectiveness.

In addition, we analyze the location information of traffic mirrors from OpenStreetMap (OSM) geographic data as the "*road information*" and Google Street View (GSV) image data as the "*road images*" to generate a traffic mirror distribution map in this study. This traffic mirror distribution map will assist in identifying potential hazards at intersections and provide valuable information to drivers and pedestrians. An overview of our traffic mirror distribution map construction method is shown in Figure 2.

The rest of this paper is organized as follows. Section 2 discusses previous research conducted on traffic mirrors. Next, in Section 3, we describe our deep learning-based methods for detecting traffic mirrors. Section 4 explains our approach to constructing the traffic mirror distribution map. Finally, in Section 5, we conclude the paper and discuss future work.

2. Related Work

Traffic mirrors play an important role in enhancing vehicle driving safety, especially in Japan where many local roads are narrow and have blind spots. As a result, many researchers have focused on the development of methods for detecting and analyzing traffic mirrors, as well as utilizing them to improve driving safety. Hino et al. [6] conducted research on a system that recognizes dangerous situations using on-board cameras and traffic mirrors, while Sato et al. [16] conducted research on extracting features of moving objects to address the issue of low-resolution mirror surfaces in traffic mirror image recognition. In addition, they utilized Google Street View images for detecting the traffic mirror, and generated training data using CycleGAN [20] based on computer graphics (CG) images created in Unity. However, neither study considered the appropriate object detection model for detecting traffic mirrors, which serves as the starting point for their respective systems.

Moreover, Kojima et al. proposed a system that generates virtual mirror images from surveillance camera footage at intersections and projects them onto a head-up display, providing drivers with a visual warning of approaching vehicles [9]. However, the cost of surveillance cameras is higher compared to existing traffic mirrors. As a study to prevent accidents at intersections without using traffic mirrors, Yoshida et al. [19] developed a system using smartphones carried by bicycles and pedestrians. In this system, smartphones with a dedicated application installed communicate with each other through GPS information exchange to understand each other's position, and notify the user when they approach each other to raise awareness. However, since this system developed relies on GPS information exchanged between smartphones, there is a possibility that other users will not be able to determine their position if they are unable to transmit their own coordinates. In this

Table 1. Evaluation results for each traffic mirror detection method.

Models	Optimization Algorithms	AP_{50}	AP_{60}	AP_{70}	AP_{80}	AP_{90}
Faster R-CNN	SGDM	**0.9454**	0.8993	0.7821	0.5328	0.1233
	Adam	0.8046	0.7403	0.5540	0.2406	0.0122
SSD	SGDM	0.8325	0.8272	0.8043	0.6619	0.0949
	Adam	0.8408	0.8397	0.8128	**0.6763**	**0.1781**
YOLOv4	SGDM	0.9326	**0.9277**	**0.8891**	0.6461	0.0890
	Adam	0.9424	0.9238	0.8544	0.5931	0.0707

study, an alert is triggered by approaching a pre-established position of the traffic mirror, eliminating the need to obtain the position information of other users. Moreover, the coordinates of the traffic mirror are stored in the device, so the user does not need to communicate anything other than receiving GPS information.

3. Traffic Mirror Detection Methods

In this section, we conduct an investigation aimed at identifying the most efficient deep learning model for detecting traffic mirrors.

3.1. Data Set

We utilized images captured on roads including traffic mirrors, initially intended for detecting traffic mirrors from car-mounted camera images. As a preliminary stage, a total of 770 images were used for the dataset, including road images from Google Street View and images captured on roads located in Fukuoka, Osaka, and Kyoto, Japan that contain traffic mirrors, as shown in Figure 1. These data were preprocessed by resizing to a resolution of $500 * 500$ and were used for training and validation of each model.

3.2. Deep Learning Models

In this paper, we compare three object detection models: Faster R-CNN [14], Single Shot MultiBox Detector (SSD) [11], and YOLOv4 [1] with two optimization algorithms: Adam [8] and SGDM [17]. The Faster R-CNN model uses ResNet-50 [5] as its backbone and is trained using online learning (batch size = 1) due to limitations in the execution environment. Similarly, the SSD model also uses ResNet-50 as its backbone and is trained using mini-batch learning with a batch size of 32. The YOLOv4 model uses CSP-DarkNet53 [18] as its backbone and is trained using online learning.

3.3. Evaluation

In deep learning, even with the same model, there may be a significant difference in performance depending on the initial values of the neural network and the input data during the learning process. Therefore, in this paper, we conduct a three-fold cross-validation by dividing the dataset into three parts, using two for training and the remaining one for validation. The detection results are evaluated by a value called Intersection over Union (IoU), which represents the proportion of overlap between the ground truth and the detected regions, and is considered True Positive when it is above a certain threshold. Model evaluation is performed by calculating the Average Precision (AP) at each stage of IoU from 0.5 to 0.9 AP is an evaluation metric that approaches 1 as detection and classification become more accurate.

3.4. Discussion

Table 1 shows the performance evaluation of each object detection model. When comparing the AP for each IoU, Faster R-CNN with SGDM shows the highest performance with an AP of 0.9454 when the IoU threshold is set to 0.5 or higher. However, at a threshold of 0.6 or higher, YOLOv4 with SGDM surpasses the aforementioned Faster R-CNN by about 0.028. Similarly, when the threshold is set to 0.8 or higher, SSD with Adam shows better performance than other models.

Considering the analysis of the traffic mirror area, Faster R-CNN with SGDM is the most suitable model in terms of detection only, but it is preferable to have an IoU of 0.7 or higher. Additionally, as the AP drops sharply when the threshold is set to 0.8 or higher, currently, YOLOv4 with SGDM is considered the best model for detecting traffic mirrors. Figure 3 shows examples of the detection results for YOLOv4. The yellow rectangle represents the detected area, while the red rectangle and numerical value represent the IoU between the ground truth and the two rectangles.

When comparing the optimization algorithms used by each model, it was found that Faster R-CNN and YOLOv4, which performed online learning, showed better performance with SGDM, while SSD, which performed mini-batch learning, showed better performance with Adam. Previous research [3] suggested that Adam outperforms SGDM, but in cases where the batch size is very small, SGDM is considered superior.

After comparing and considering the combination of the three object detection models, Faster R-CNN, SSD, and YOLOv4, with the two optimization algorithms, SGDM and Adam, this study has arrived at the conclusion that YOLOv4 and SGDM are the most suitable combination for

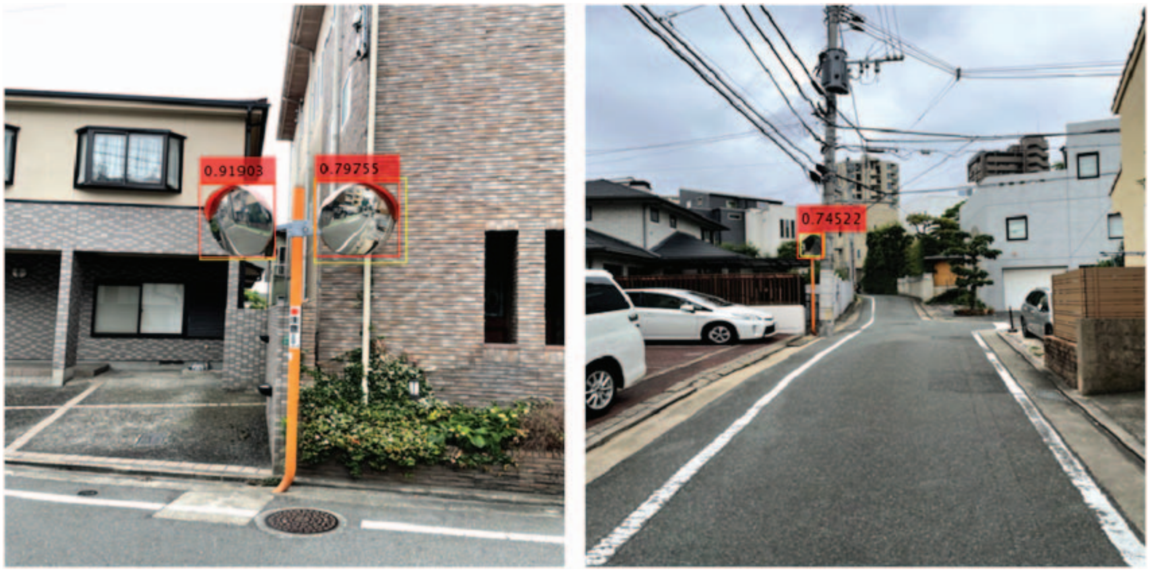

Figure 3. Examples of detecting traffic mirrors using YOLOv4.

achieving the objectives of this study at the current stage.

4. Traffic Mirror Distribution Map Construction Method

Figure 2 presents an overview of the process for creating a traffic mirror distribution map using the coordinates of intersections where traffic mirrors are installed. The steps involved in creating the map are as follows:

1. Extract roads from OpenStreetMap (OSM).
2. Extract coordinates of intersections using QGIS and obtain their latitude and longitude.
3. Acquire Google Street View (GSV) images of the intersections.
4. Use the trained object detection model to detect the presence of traffic mirrors in the images.
5. Display the identified mirrors on the map.

Firstly, we extract the road data from OSM. The extracted road data is then input into QGIS[1], a Geographic Information System software, to obtain the coordinates of intersections with crossroads or T-junctions. Subsequently, these intersection coordinates are used to retrieve corresponding road images from GSV in four directions: north, south, east, and west. Finally, the pre-trained traffic mirror detection model (as mentioned in Section 3) is applied to detect the presence of traffic mirrors, annotate their position coordinates, and generate the map. Further details will be discussed in the following subsections.

[1] https://qgis.org/ja/site

4.1. Obtaining Geographic Data from OSM and Extracting Intersections

Overpass Turbo[2] is used to obtain geographic data, which is then exported as a “gpx” file. By setting the search keyword to “highway”, information on roads that are passable by cars can be obtained. The geographic data is then loaded into QGIS and assigned geometric attributes. The intersection points of lines are obtained through the intersection, and a list of position coordinates of crossroads and T-junctions is created.

4.2. Acquiring GSV Images and Generating the Traffic Mirror Distribution Map Based on Traffic Mirror Detection

GSV images are obtained from the position coordinates obtained in the previous step. The images are taken in all four directions, with a resolution of $1000 * 1000$ pixels, and are used to detect traffic mirrors using the pre-trained model. According to the evaluation results of Section 3, the YOLOv4 model with the optimization algorithms SGDM is used for the training, and 770 training data images resized to $500 * 500$ pixels are used. Finally, the position coordinates where the model detects the presence of traffic mirrors are obtained, and pins are plotted on the map at these locations using the Google Maps API.

[2] https://overpass-turbo.eu

Figure 4. An example of the traffic mirror distribution map using traffic mirror data.

4.3. Implementation and Evaluation

This subsection explains the implementation of the traffic mirror distribution map. We generated a traffic mirror map for an area of 2.7 square kilometers in the northern part of Amagasaki city, Japan. The map was implemented using Google Maps API, and pins were plotted at locations identified as having traffic mirrors. We analyzed 1,132 intersections where Google Street View images were available within the analysis area. We constructed a traffic mirror detection model using a training dataset of 770 images with a resolution of $500 * 500$ pixels. Furthermore, in this verification, we visualized the coordinates of intersections where traffic mirrors were correctly detected, as well as those where there were false detections or no detections, to confirm the detection performance.

The method for calculating the accuracy of traffic mirror detection is as follows: True Positive (TP) refers to the intersection where one or more mirrors were detected in one or more of the four cardinal directions (north, south, east, west). False Negative (FN) refers to the intersection where no mirrors were detected in one or more of the four cardinal directions. False Positive (FP) refers to the intersection where a mirror was detected in one or more of the four cardinal directions, but there was actually no mirror present in the image. Finally, True Negative (TN) refers to the result where no detection was performed in the image where no mirror was present. Based on these criteria, we classified the intersections and calculated the recall and precision.

In the results of the evaluation, firstly, out of the 123 intersections where traffic mirrors were present and were the subject of this verification, the recall rate, which indicates the percentage of intersections where the machine learning model (YOLOv4) was able to detect the traffic mirror, was 63.4%. Furthermore, the precision rate, which indicates the percentage of intersections where the traffic mirror was actually present out of those detected by the machine learning model, was 90.0% (the specific results are shown in Table 2). Figure 4 shows the distribution of traffic mirrors in this verification, implemented using the Google Maps API. Orange pins indicate the intersections classified as TP, brown pins indicate the intersections classified as FN, and yellow pins indicate the intersections classified as FP. Additionally, three intersection images show the images where the mirrors were detected through image analysis.

Table 2. The specific results (number) of traffic mirror detection.

Intersection	TP	FN	FP	TN
1132	78	45	9	1000

4.4. Consideration

Firstly, we explain the challenges and considerations regarding image collection through Google Street View. Although we obtained GSV images by specifying coordinates and directions, there were some locations where the same image was captured from different directions, making it difficult to determine the direction of the traffic mirror in certain intersections. Additionally, we acquired incomplete GSV images, including indoor images (examples are shown in (B) of Figure 5) that were not necessary for our study. Furthermore, there were images where the traffic mirror existed in reality but was not captured, as well as locations where no corresponding image was available. To address these issues, we will improve the image acquisition process by obtaining images from multiple adjacent coordinates, rather than just one coordinate per intersection.

Furthermore, our system was able to detect about 60% of intersections where a traffic mirror exists, however, it was unable to detect the remaining 40%. We believe that the reason for this is due to the insufficient generalization performance of the learning model (as depicted in (A) of Figure 5, certain non-circular traffic mirrors cannot be detected), and we plan to improve this by increasing the amount of training data.

Finally, we implemented the generated traffic mirror distribution to be accessible on smartphones. For application development, we utilized Flutter[3], a framework for building cross-platform applications. Currently, the application is available on the Android platform. The screen is rendered using the WebView[4] component provided by the Android developer toolkit, which accesses a web page hosted

[3]https://flutter.dev

[4]https://developer.android.com/reference/android/webkit/WebView

Figure 5. Challenges regarding the image collection process through Google Street View.

Figure 6. The mobile application of the traffic mirror distribution map.

on the internet (refer to Figure 6). When the screen is displayed, the application retrieves saved coordinates from the database. An ongoing experiment involves triggering an alarm from the smartphone when approaching an intersection with a traffic mirror. There may be instances where the camera cannot fully capture the presence of a traffic mirror. However, on the contrary, if the presence of a traffic mirror is known, it is possible to emit an alert sound even when the camera fails to capture it.

4.5. Future Work

At present, we have been able to implement a system that notifies users with a warning sound when approaching an intersection with a mirror. However, there are issues with the accuracy of GPS itself, which can be affected by the location. Therefore, to accurately provide warning notifications, it is necessary to consider the GPS error. Our future plans include solving the above issues, aiming to implement a practical navigation application, and determining the optimal distance for warning notifications by investigating the accuracy of GPS.

In the future, our primary focus will revolve around addressing the challenges related to GPS accuracy and machine learning precision in order to create a functional navigation application. Additionally, in this study, GSV data is utilized as the *road image* data, and OSM data is used as the source of *road information*. This choice is primarily due to the availability and free accessibility of these datasets. However, in the future, we have plans to acquire and incorporate high-resolution street view data to enhance the accuracy of traffic mirror detection. Furthermore, we plan to conduct user evaluations to assess the usability of the finalized application.

5. Conclusion

In this paper, through the comparison and examination of the combination of three object detection models, Faster R-CNN, SSD, and YOLOv4, with two optimization algorithms, SGDM and Adam, we have derived the combination of YOLOv4 and SGDM that is suitable for the purpose of our current study. However, it should be noted that there were some parameters, such as the mini-batch size, which were not thoroughly adjusted due to constraints in the experimental environment. Moving forward, we plan to try other object detection models that have not been considered and aim to improve the accuracy through increasing the dataset size and using other deep learning methods.

We also analyzed the location information of traffic mirrors using geographic data from OpenStreetMap (OSM) and image data from Google Street View (GSV) to generate a traffic mirror distribution map. We were able to detect and visualize approximately 60% of traffic mirrors from intersection images obtained from GSV. However, we identified the need to improve the traffic mirror detection model performance by increasing the amount of training data, and to acquire missing GSV images through other methods. For instance, we plan to improve the incomplete GSV image acquisition by obtaining images from multiple neighboring coordinates.

Our ultimate objective is to develop an advanced system that can effectively detect blind spots reflected in traffic mirrors by analyzing high-resolution street view data and geographical information, and provide timely alerts to pedestrians and vehicles, thereby improving safety in the vicinity of these blind spots.

Acknowledgment

This work was partially supported by JSPS KAKENHI Grant Numbers JP21K03962, JP21H01457, JP19K12240, JP22H03700, and the funding from Fukuoka University (Grant No.GW2310).

References

[1] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020.

[2] Fangrong Chang, Maosheng Li, Pengpeng Xu, Hanchu Zhou, Md Mazharul Haque, and Helai Huang. Injury severity of motorcycle riders involved in traffic crashes in hunan, china: a mixed ordered logit approach. *International journal of environmental research and public health*, 13(7):714, 2016.

[3] Dami Choi, Christopher J Shallue, Zachary Nado, Jaehoon Lee, Chris J Maddison, and George E Dahl. On empirical comparisons of optimizers for deep learning. *arXiv preprint arXiv:1910.05446*, 2019.

[4] Ali J Ghandour, Huda Hammoud, and Samar Al-Hajj. Analyzing factors associated with fatal road crashes: a machine learning approach. *International journal of environmental research and public health*, 17(11):4111, 2020.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[6] Yusuke Hino, Shintaro Ono, Noriaki Itagaki, and Yoshihiro Suda. Recognition of risky events reflected in road safety mirror considering ego vehicle's motion. *7th International Symposium on Future Active Safety Technology toward Zero Accidents (FAST-zero)*, 2021.9.

[7] Lin Hu, Xiqin Zhou, Xin Zhang, Fang Wang, Qiqi Li, and Wenguang Wu. A review on key challenges in intelligent vehicles: Safety and driver-oriented features. *IET Intelligent Transport Systems*, 15(9):1093–1105, 2021.

[8] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[9] Kazuhiro Kojima, Akihiko Sato, Fumihiro Taya, Yoshinari Kameda, and Yuichi Ohta. Naviview: visual assistance by virtual mirrors at blind intersection. In *Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005.*, pages 592–597. IEEE, 2005.

[10] Zhenning Li, Cong Chen, Yusheng Ci, Guohui Zhang, Qiong Wu, Cathy Liu, and Zhen Sean Qian. Examining driver injury severity in intersection-related crashes using cluster analysis and hierarchical bayesian models. *Accident Analysis & Prevention*, 120:139–151, 2018.

[11] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer, 2016.

[12] Shawky A Mohamed, Kishta Mohamed, and Husain A Al-Harthi. Investigating factors affecting the occurrence and severity of rear-end crashes. *Transportation research procedia*, 25:2098–2107, 2017.

[13] Midori Mori, Sadao Horino, Sou Kitajima, Masaru Ueyama, Takeshi Ebara, and Toru Itani. Ergonomics solution for crossing collisions based on field assessment of visual environment at urban intersections in japan. *Applied ergonomics*, 39(6):697–709, 2008.

[14] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 2015.

[15] Rusdi Rusli, Md Mazharul Haque, Mark King, and Wong Shaw Voon. Single-vehicle crashes along rural mountainous highways in malaysia: An application of random parameters negative binomial model. *Accident Analysis & Prevention*, 102:153–164, 2017.

[16] Rikiya Sato and Kazunori Onoguchi. Scene recognition for blind spot via road safety mirror and in-vehicle camera. In *2022 26th International Conference on Pattern Recognition (ICPR)*, pages 3714–3720. IEEE, 2022.

[17] Ilya Sutskever, James Martens, George Dahl, and Geoffrey Hinton. On the importance of initialization and momentum in deep learning. In *International conference on machine learning*, pages 1139–1147. PMLR, 2013.

[18] Chien-Yao Wang, Hong-Yuan Mark Liao, Yueh-Hua Wu, Ping-Yang Chen, Jun-Wei Hsieh, and I-Hau Yeh. Cspnet: A new backbone that can enhance learning capability of cnn. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 390–391, 2020.

[19] Hiroyuki Yoshida, Miyuki Nakano, and Midori Sugyaya. A study of accident prevention both of bicycle and pedestrian by approaching detection based on mobile gps. *SIG Technical Reports, Ubiquitous computing systems, 2015-UBI-46*, pages 1–8, 2015.

[20] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.