A COMPUTER VISION-BASED ROAD MARKING INSPECTION SYSTEM WITH STREET VIEW IMAGES

Junjie Wu*[1], Wen Liu[2], Yoshihisa Maruyama[2]
[1]Graduate School of Science and Engineering, Chiba University
Email: ayka5587@chiba-u.jp
[2]Graduate School of Engineering, Chiba University

KEY WORDS: Road Marking, Damage Inspection, Computer Vision, Deep Learning

ABSTRACT: Road marking as vital road infrastructure can convey abundant guidance and information to drivers and autonomous vehicles. However, it is inevitable that road marking will wear out over time and influence traffic safety. At the same time, the inspection and maintenance of road marking cost enormous human and economic resources. This study proposes a road marking damage inspection system using computer vision techniques with street view images captured by a regular digital camera mounted on a vehicle. The damage ratio of road marking is measured according to the undamaged part and region of road marking using semantic segmentation, inverse perspective mapping, and image thresholding approaches. Furthermore, road marking damage detectors using YOLOv8 algorithm are developed based on the damage ratio of road marking. The experimental results show that the proposed system successfully automates the inspection process for road markings.

1. INTRODUCTION

In recent years, there has been a rapid progress in Autonomous Vehicles (AVs). The Society of Automotive Engineers (SAE) defines six levels of driving automation ranging from level 0 (no automation) to level 5 (full automation) (Dimitris, 2019). Nowadays, although there are no SAE level 5 AV on roads, some systems can already support driver out of the wheel, reaching SAE level 3 (conditional automation) capabilities. With the development of AV technologies, several aspects should be enhanced to embrace the era of AVs, including traffic management, road infrastructure adaptation, revenue and budgeting, liability and insurance, police and emergency services, and social justice and equity (Felix, 2017). Among them, road infrastructure adaptation can be considered as one of the most urgent problems to be solved at present. On the one hand, roads and other road infrastructure (e.g., lane, marking, signal, etc.) require adjustments and maintenance to accommodate AVs. On the other hand, implementing smart road technologies that are compatible with AVs will enhance safety and efficiency on roads (Olatz, 2022).

The future road network will be a mix of human-driven vehicles and AVs, which makes road marking represents an important infrastructure element today and in future (Cha, 2017). Properly installed and well maintained road marking can convey abundant guidance and information to drivers and AVs. On the contrary, damaged road markings post great challenges to drivers and camera sensor based AVs, as traffic safety is dependent on the visibility of road markings (Chong, 2021). It is inevitable that road marking will wear out over time. Therefore, the inspection and maintenance of road marking are crucial. However, the previous practice was to manually inspect the degree of damage of road marking, which makes the maintenance work cost enormous human and economic resources (Kong, 2022). Considering this, an automated road marking damage inspection system using deep learning and computer vision techniques is proposed with street view images captured by a regular digital camera mounted on a vehicle in this study.

The remainder of the paper is organized as follows. Section 2 introduces the proposed method for road marking damage inspection system. Thereafter, the road marking damage detection dataset and the result of experiment are presented in Section 3. Finally, Section 4 concludes the paper.

2. PROPOSED METHOD

2.1 Road Marking Damage Assessment

In Figure 1, we show the overall flowchart of road marking damage assessment.

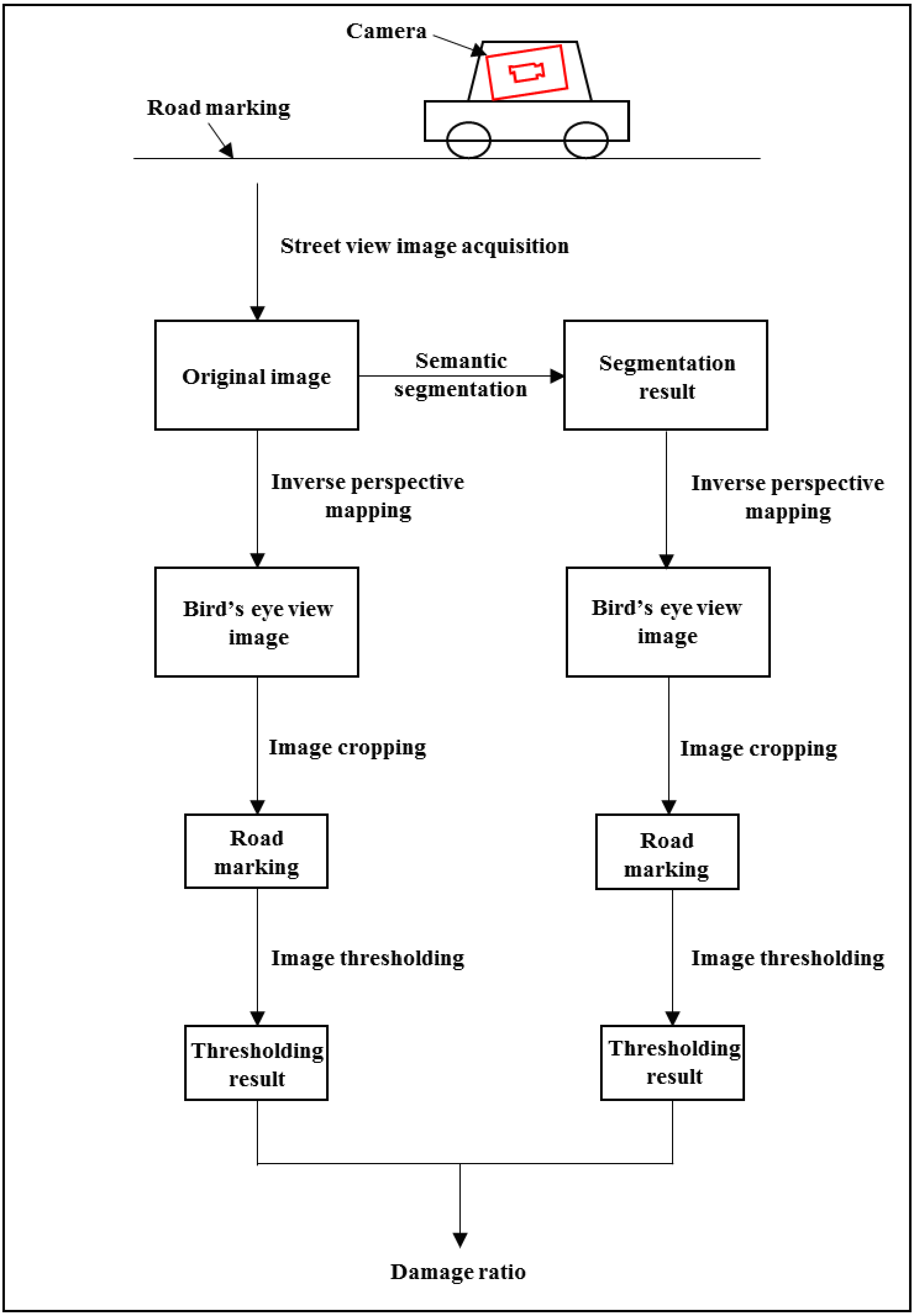


Figure 1. Overall flowchart of road marking damage assessment.

(1) Data Acquisition

For data acquisition, we use a front-view camera mounted on a vehicle shown in Figure 1. The camera used in this study can be a low-cost regular digital camera, this makes the proposed method cost-effective. Then, street view images are collected from three cities in Japan: Yokohama, Chofu, and Nogata, in November 2015, November 2015, and March 2017, respectively. Figure 2 shows the example of street view image captured.

Figure 2. Example of street view image.

(2) Semantic Segmentation

The second step is to perform semantic segmentation to the original street view image. In this study, we aim to measure the damage ratio of road marking according to the undamaged part and region of road marking. Therefore, semantic segmentation is used to extract integral region of road marking. On the other hand, the undamaged part can be extracted from original image using inverse perspective mapping (IPM), image cropping, and image thresholding. We use multiscale attention-based dilated convolutional neural network to tackle the original image. This neural network takes multiple-scale images that are resized from the original image as inputs to learn the attention weights of each scale, following merging the semantic predictions to obtain the final output. In addition, dilated convolution is adopted in the feature-extraction process to utilize a larger range of spatial-context information (Junjie, 2022). The segmentation result is promising, as the neural network successfully extract the region of road marking. Figure 3 shows the segmentation result of the example shown in Figure 2.

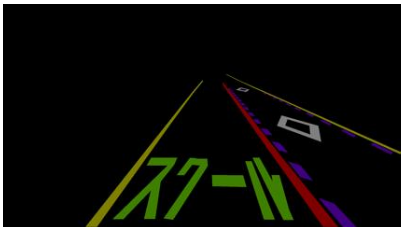


Figure 3. Example of segmentation result.

(3) Inverse Perspective Mapping

In the street view image captured, the pixels occupied by each road marking are different. As shown in Figure 2, the distant road marking of Approach to Pedestrian and Bicycle Crossing (prismatic road markings on the right side of road) is thinner than the nearby one. This is not conducive to the road marking damage assessment. Moreover, the street view image is horizontal, in which the perspective effect can negatively impact the road marking damage assessment (Kong, 2022). Therefore, the IPM is adopted to convert street view image to bird's eye view (BEV) image. Generally, IPM maps pixels of an image from horizontal view to vertical view through homography matrix, which is a transformation matrix (Massimo, 1998). In this study, we use Open Source Computer Vision Library (OpenCV) to conduct IPM on both original image and segmentation result. OpenCV provides more than 2500 optimized algorithms for computer vision task. We mainly use functions of $cv2.getPerspectiveTransform$ which takes as input the four pairs of corresponding points and outputs the transformation matrix, and $cv2.warpPerspective$ which applies the transformation matrix to the input image to get the BEV image. Figure 4 shows BEV images obtained through IPM.

Figure 4. Example of the obtained BEV images (left: original image; right: segmentation result).

(4) Image Cropping

In order to conduct damage assessment for road marking independently, each instance needs to be cropped from both original image and segmentation result. After cropping to preserve the proper range of road marking, we set the size of all of the road marking images to 80×78. Figure 5 shows the example of the obtained road marking image.

Figure 5. Example of the obtained road marking images (left: original image; right: segmentation result).

(5) Image Thresholding

Since the goal is to assess the damage ratio of road marking according to its area, we need to convert the areas of undamaged part and region to the numbers of pixels in an image to calculate. Image thresholding can perform basic segmentation in an image, and turn it into a binary image where pixels are either 0 or 1 (or 255) (Mehmet, 2004). Hence, we employ global thresholding to extract the undamaged part and region of road marking from cropped road marking images obtained from previous step. Figure 6 shows the thresholding results with OpenCV.

Figure 6. Example of the thresholding results (left: original image; right: segmentation result).

(6) Damage Ratio

The damage ratio of road marking can be measured as follows:

$$R = 1 - \frac{N_{ud}}{N_r} \tag{1}$$

Where N_{ud} is the number of pixels of the undamaged part of road marking, N_r is the number of pixels of the region of road marking, and R represents the damage ratio of road marking. Since IPM process in step (3) eliminates the perspective effect, every pixel on the road marking image obtained by previous step has the same area in the actual world. Moreover, the sizes of road marking images in previous step are set to be the same. Thus, the ratio of the area of undamaged part to the area of region of road marking, can be regarded as the ratio of the number of pixels of undamaged part to the number of pixels of region of road marking. For the example in Figure 6, the number of pixels of undamaged part is 2,573, the number of pixels of region is 3,841, that makes the damage ratio of 0.33.

2.2 Road Marking Damage Inspection System

The flowchart of road marking damage inspection system using object detection approach is shown in Figure 7. Inspired by Kong et al. (Kong, 2017), we divide the undamaged ratio $(1 - R)$ into four grades: $no\ damage$, $slight\ damage$, $moderate\ damage$, and $severe\ damage$. The undamaged ratio of 90% to 100% is regarded as $no\ damage$, and it means that there is no need for maintenance. The undamaged ratio of 75% to 90% is regarded as $slight\ damage$, and it means that there are minor defects that do not need immediate responses. The undamaged ratio of 50% to 75% is regarded as $moderate\ damage$, and it means that the defects need to be repaired. The undamaged ratio of 0 to 50% is regarded as $severe\ damage$, and it means that the defects require urgent responses.

Then, the results of road marking damage assessment are used as labelled data to train object detection models. The object detection model can detect and locate objects of interest in an image, that are damaged road markings with grade labels in this study. There are many outstanding object detection algorithms, i.e., Faster R-CNN (Ren, 2015), YOLO series (Joseph, 2015; Redmon, 2018), and SSD (Liu, 2016).After training and model validation, the weights obtained can automatically detect the damaged road marking with grade label from a street view image.

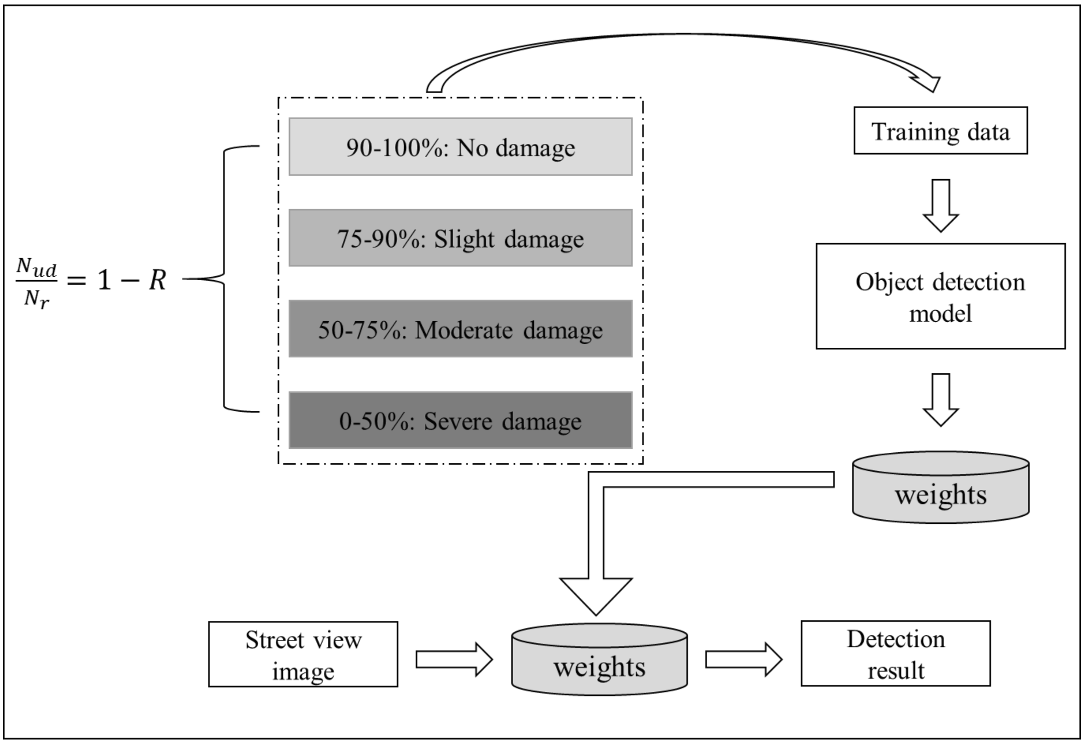


Figure 7. Flowchart of road marking damage inspection system.

3. RESULTS AND DISCUSSION

3.1 Road Marking Damage Detection Dataset

A road marking damage detection dataset is proposed in this study. In the road marking damage assessment, a total of 1036 street view images with 4374 road markings are processed. To be specific, the damage ratios of road markings are measured in the road marking damage assessment process mentioned in Section 2.1. Then, they are used as labelled data to create the dataset. An annotation tool called Roboflow is used to manually annotate the street view images. One street view image consists of 1,920 ×1,080 pixels. Road markings are annotated with bounding boxes, which is a commonly used annotation type in the field of object detection. The dataset is divided into the training and test sets at a ratio of approximately 8:2, which correspond to 810 and 226 images, respectively.

As mentioned earlier, the damaged road markings are divided into four grades that are $no\ damage$, $slight\ damage$, $moderate\ damage$, and $severe\ damage$ according to their damage ratios. On this basis, we divide road markings into line-type markings, arrow-type markings, block-type markings, and word & number-type markings according to their shapes. This makes the number of classes of the dataset 16. The shape of road marking, damage grade, class ID, and corresponding example are shown in Table 1.

3.2 Experimental Result

Detectors based on YOLOv8 architecture are built to tackle the proposed dataset. The YOLO series has achieved significant success in the realm of object detection since 2015. YOLOv8 is the latest version of YOLO series, which was released by Ultralytics on January 10, 2023. It is the most advanced and state-of-the-art model that offers high performance in terms of accuracy and speed. In YOLOv8, there are five different models that are: YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x. For instance, YOLOv8n represents YOLOv8 Nano, and it is the smallest and fastest model, while YOLOv8x represents YOLOv8 Extra Large, and it is the largest yet most accurate model.

The detectors are trained on a Linux Ubuntu 20.04 LTS with one NVIDIA GeForce RTX 3090 GPU with 24 GB video memory. For deep learning framework, we use PyTorch. After exploring several hyper-parameters, the weights of best fit are generated. The optimizer used is stochastic gradient descent (SGD) with momentum of 0.9. The learning rate and weights decay are set to be 0.001 and 0.0001, respectively. We train the model for 200 epochs with the batch size of 8. After training, the model is evaluated on the test set. We use mean average precision (mAP) as evaluation metrics (Paul, 2017). The results are shown in Table 2. Although the mAP is not very high, a reasonably accurate baseline is made in this study.

Table 1. The classification system of the proposed dataset.

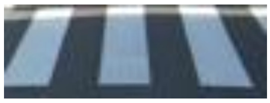
Shape of road marking	Damage grade	Class ID	Example
Line	No damage	LD0	
	Slight damage	LD1	
	Moderate damage	LD2	
	Severe damage	LD3	
Arrow	No damage	AD0	
	Slight damage	AD1	
	Moderate damage	AD2	
	Severe damage	AD3	
Block	No damage	BD0	
	Slight damage	BD1	
	Moderate damage	BD2	
	Severe damage	BD3	
Word & Number	No damage	WND0	
	Slight damage	WND1	
	Moderate damage	WND2	
	Severe damage	WND3	

Table 2. Detection performance using YOLOv8 models.

Model	mAP (%)
YOLOv8n	57.85
YOLOv8s	63.92
YOLOv8m	68.10
YOLOv8l	69.39
YOLOv8x	71.18

4. CONCLUSIONS

With the development of AVs, road infrastructure adaptation is becoming one of the most urgent problems to be solved. At the same time, road marking plays a vital role in the context of road infrastructure adaptation, as it can provide abundant guidance and information to the AVs and drivers. Considering this, this study is focused specifically on the damage inspection of road marking. We use Japanese street view images captured by a regular digital camera mounted on a vehicle as raw data. A method of road marking damage assessment is proposed based on semantic segmentation, IPM, and image thresholding. According to the damage ratios of road markings measured by this method, we design the road marking damage detection dataset. Furthermore, road marking damage detectors are developed using YOLOv8.

The road marking damage assessment process should be improved, particularly in terms of image cropping and thresholding. We conduct image cropping manually in this study, which should be replaced by automated process to eliminate human errors. For image thresholding, adaptive thresholding is more appropriate than global thresholding in this case, as the street view image has different lighting conditions in different areas. In addition, the prediction results of detectors are not very high due to the complexity of dataset and the lack of data. It is essential to continuously improve the dataset.

REFERENCES

Cha, Y.-J., Choi, W., Buyukozturk, O., 2017. Deep learning-based crack damage detection using convolutional neural networks, *Computer-Aided Civil and Infrastructure Engineering*, 32(5), pp. 361-78.

Chong, W., Shurong, L., Kai, W., Zijian, Z., Ying, W., 2021. Damage inspection for road markings based on images with hierarchical semantic segmentation strategy and dynamic homography estimation. *Automation in Construction*, Vol 131, 103876.

Dimitris, M., 2019. Long-term implications of automated vehicles: an introduction. *Transport Reviews*, Vol. 39, No. 1, pp. 1-8.

Felix, B., Kay, W.A., 2017. Literature review on surveys investigating the acceptance of automated vehicles. *Transportation* 44, pp. 1293-1306.

Joseph R., Ali F., 2018. YOLOv3: An Incremental Improvement. *arXiv*, 1804.02767.

Joseph R., Santosh D., Ross G., Ali F., 2015. You Only Look Once: Unified, Real-Time Object Detection. *Computer Vision and Pattern Recognition*, pp. 779-788.

Junjie, W., Wen, L., Yoshihisa, M., 2022. Automated Road-Marking Segmentation via a Multiscale Attention-Based Dilated Convolutional Neural Network Using the Road Marking Dataset. *Remote Sensing* 14, no. 18: 4508.

Kong, W., Teng, Z., Xin, M., Shuliang, Z., Min, C., Guonian L., 2022. Automatic Detection and Assessment of Pavement Marking Defects with Street View Imagery at the City Scale. *Remote Sensing* 14, no. 16: 4037.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, Cheng-Y., Berg, Alexander C., 2016. SSD: Single Shot MultiBox Detector. *European Conference on Computer Vision*.

Massimo, B., Alberto, B., Alessandra, F., 1998. Stereo inverse perspective mapping: theory and applications. *Image and Vision Computing*, Vol 16, Issue 8, pp. 585-590.

Mehmet, S., Bulent, S., 2004. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, Vol. 13, Issue 1.

Olatz, I., Nagore, I., Alfonso, B., Diego, B., 2022. Road Marking Damage Detection Based on Deep Learning for Infrastructure Evaluation in Emerging Autonomous Driving. *IEEE Transactions On Intelligent Transportation Systems*, Vol. 23, No. 11.

Paul, H., Vittorio, F., 2017. End-to-End Training of Object Class Detectors for Mean Average Precision. *Computer Vision – ACCV 2016*, Vol 10115.

Ren, S., He, K., Girshick, R., & Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems* 28.