Automating daily inspection for Expressways using anomaly detection model

Yuta Shirakawa[1], Satoshi Ito[1], Reiko Noda[1], Naoto Yoshitani[2], Masahide Wake[3], and Honoka Takano[3]

[1]*Media AI Laboratory, Corporate Research & Development Center, Toshiba Corporation*
yuta1.shirakawa@toshiba.co.jp
satoshi13.ito@toshiba.co.jp
reiko.noda@toshiba.co.jp

[2]*Maintenance Strategy Headquarters, Central Nippon Expressway Company Limited*
n.yoshitani.aa@c-nexco.co.jp

[3]*Social Infrastructure System Solutions Engineering Dept. ICT Solutions Div. Toshiba Digital Solutions Corporation*
masahide.wake@toshiba.co.jp
honoka1.takano@toshiba.co.jp

ABSTRACT

Because of the high speed at which vehicles travel on highways, even small irregularities on the road surface can lead to serious accidents. It is important to conduct daily visual inspections to detect these abnormalities at an early stage and to repair them quickly. We are considering replacing a part of visual inspection with automatic classification using image recognition. Automating inspections will make it possible to increase inspection frequency and expected to reduce the variation in quality due to the skill of inspectors. In this paper, we report on an AI-based inspection system and evaluation results using actual highway driving video captured by an in-vehicle camera.

1. INTRODUCTION

For the maintenance and preservation of healthy social infrastructure, it is important to evaluate its condition from time to time and to detect the occurrence of deformations at an early stage, using recognition technology of prognostic change. Expressways are one of the fundamental social infrastructures that are constantly subject to wear and tear caused by passing vehicles and daily inspections are conducted to maintain them.

However, as needs diversify and renewal work increases, the number of tasks other than inspections is also increasing. To respond to the increasing of tasks, Central Nippon Expressway Co., Ltd. is working on Innovative Conservation Management using Next-Generation Technology (i-MOVEMENT)" which uses ICT and robotics technologies to improve the efficiency of various tasks, including the daily inspections.

Figure 1 The sensing vehicle with multiple sensors. In this paper, we use an on-board camera in front of the vehicle to capture expressways surface.

In this paper, we report the results of applying anomaly detection methods to on-board camera images with the aim of replacing some of the visual inspections in daily operations. Normally, a large amount of training data is required to prepare such an anomaly detector, but it is difficult to collect many data from the product environment in advance. Therefore, we mitigate the cost of data collection in two ways: 1) we use weakly supervised learning, in which the labels required for training are simple, and 2) we also investigate performance improvement through domain adaptation which utilizes anomaly data from city streets, which have a different landscape domain from expressways.

This paper is organized as follows. In Section 2, we briefly review the daily inspections. In Section 3, we describe the anomaly detection method and domain adaptation settings, In Section 4, we report on the results of evaluation experiments conducted using on-board camera images. In Section 5 we conclude this paper.

2. DAILY INSPECTIONS IN EXPRESSWAYS

Expressways businesses provide road services to customers, but unlike railroad business, the passing vehicles are driven by the customers. Safety driving depends on their own skills, and traffic accidents are a constant occurrence. For this reason, expressway operator strives to detect road surface deformations that directly affect driving and repair them quickly to reduce traffic accidents as much as possible.

Maintenance inspections are performed daily to detect such deformations of the road surface. Currently, the inspections are conducted through visual inspections, sound inspections, and experiential inspections of the vehicle sway by skilled personnel.

Although the frequency of this daily inspection is defined according to the amount of traffic, which does not always allow for the detection of deformations earlier than the customer, and may result in damage to the customer, such as a flat tire caused by a pothole.

With the aim of automating the detection process and changing the frequency from daily basis to hourly basis, we are considering to install the sensing vehicle shown in Figure 1. The vehicle is equipped with cameras in four directions (front, rear, left, right), vibration sensors, and sound-collecting microphones, and has "AI image processing technology" that detects anomalies from the captured images.

3. AUTOMATIC INSPECTION

System Overview

In this section, we report on the applied anomaly detection method to realize the “AI image processing technology”. The framework of the application is shown in Figure 2. We are planning to capture the on-board camera video onto an in-vehicle laptop PC and perform anomaly analysis on the laptop PC. Power is supplied from car battery, and results are stored in the attached storage. This paper reports the results of analyzing the pre-recorded video on the server in the laboratory.

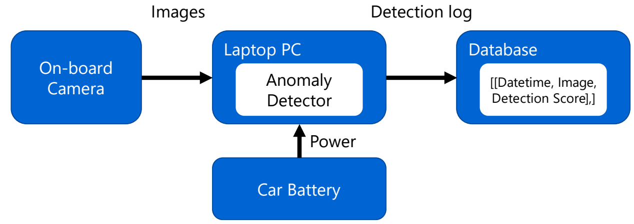


Figure 2 System Overview

Anomaly Detection

In this part, we describe our anomaly detection model applied to road inspections. Anomaly detection in road inspection is the task to detect the location of a pothole or other anomalies occurring on the road surface from a given image. In previous work, such anomaly detection was performed using object detection models (Maeda, Sekimoto, Seto, Kashiyama & Omata, 2018). Specifically, the object detection models are trained with the data and annotation label of the bounding boxes which surrounding the anomalies, and the trained model is applied to the given new image to estimate bounding boxes of anomalies.

However, the annotation cost of such bounding boxes is expensive for the following reasons. 1) It is difficult to collect enough data for training in the production environment. 2) It is difficult to determine the correct location of the anomaly and to annotate the accurate bounding boxes of the anomaly. 3) It is difficult to clearly determine the degree of deformation, and annotation varies among annotators.

To mitigate the annotation cost problem, we use an anomaly detection model based on weakly supervised learning. The method we use utilizes annotation of the presence or not-presence of anomalies. This type of annotation can be collected at lower cost than bounding boxes. Although the method does not use any location annotations, the trained model can estimate pixel-wise anomaly scores (Ito, 2022).

The overview of applied weakly supervised learning framework is shown in Figure 3. We first describe the inference process, followed by the training process and loss functions, and finally domain adaptation.

Inference phase: The anomaly score for each position in the image is calculated as a score map S using ConvNet, which consists of convolutional operations that extract local features of input image, and a subsequent sigmoid function shown in Eq. (1).

Training phase: First the model estimates pixel-level anomaly score map S for given image as in inference phase and converts the map into image-level anomaly score by taking max with GMP (global max pooling, Eq. (2)). The model parameters are optimized to minimize the BCE (binary cross-entropy, Eq. (3)) loss between the estimated image-level anomaly score p and the annotation label y by

Table 1 Evaluation dataset stats.

	Road is Damaged	Damage is Recorded	Damage is Observable	label	# of total samples	# train	# test
set1	✓	✓	✓	positive	113	84	29
set2	✓	✓		-			
set3				negative	2,329	1,746	583
set4	✓		✓	positive	291	218	73
RDD2020(train)	✓	✓	✓	positive	3,074	3,074	
RDD2020(train)				negative	17,967	17,967	

stochastic gradient descent. Making the maximum value of the score map close to the given label, the score for regions with visual features common to both images without and with anomalies approaches "0", while the score for regions with visual features found only in images with anomalies approaches "1".

$$S = Sigmoid\big(ConvNet(I)\big) = \frac{1}{1+e^{-ConvNet(I)}} \tag{1}$$

$$p = \mathrm{GMP}(S) = \max(S) \tag{2}$$

$$\mathrm{BCE}(p, y) = -y \log p - (1-y)\log(1-p) \tag{3}$$

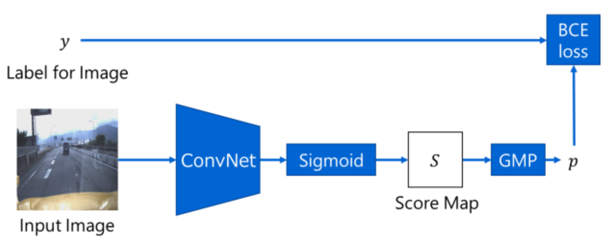


Figure 3 Anomaly Detection Model Overview

Domain Adaptation

Although the weakly supervised learning we applied trains a model without costly bounding box annotations of anomalies, it is still expensive to collect enough data in advance from the target domain, expressway.

Some previous studies on road surface anomaly detection published some datasets of anomalies and correspondent labels in city street domain. It would be good to leverage such pre-prepared data to improve an anomaly detection performance in our target domain.

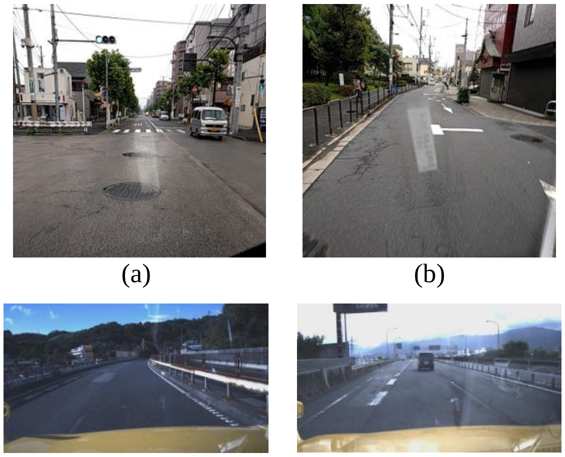
However, when applying a model trained only on such a different domain data to another target domain, it is affected by domain gaps and degrades test performance. Therefore, we confirm through experiments the effect of domain adaptation, which uses a small amount of data from the target domain in combination with a public dataset that is available. Specifically, we compared the detection performance before and after domain adaptation using a public dataset on city streets and data from an on-board camera that captured images of an Expressway.

4. EXPERIMENTS

4.1. CONDITIONS

To evaluate the effectiveness of the proposed method, we combine two type of datasets and evaluate the performance of anomaly detection. Specifically, we use public dataset of city street images as a source domain dataset and, dataset expressway image taken by on-board camera as a target domain dataset. We compared detection performance in the target domain by switching the data using for model training.

Target domain: Based on the videos captured by the on-board camera and the inspection records of holes and peeling, we prepare following subsets and use for the experiment. Set 1: Images that are in the inspection record and the anomaly can be visually confirmed in recorded image. Set 2: Some anomalies are recorded in the inspection records, but the correspondent anomalies cannot be visually confirmed in the image. Set 3: Images of the normal road surface collected from a time that is not in the inspection record. Set 4: Images on which skilled person visually confirms some anomalies other than potholes or repair marks but are not recorded in the inspection record. Of these image sets, we use 75% of sets 1, 3, and 4 as training data and the remaining 25% for evaluation data of target domain. **Source domain:** We use RDD2020 dataset (Arya, Maeda, Ghosh, Toshniwal & Sekimoto, 2021) as the source domain dataset to evaluate the effect of improvement through domain adaptation. Although RDD2020 has several anomaly labels, we use D40 (pothole) labels in this paper and only training subset are used in our experiments. Sample images from the source and target domains are shown in Figure 4. We summarize the data we use in this paper in Table 1.



(a) (b)

(c) (d)

Figure 4 (a) and (b) are image samples from Source Domain, and (c) and (d) are Target Domain

Evaluation metrics: We use AUC as the metric to compare the anomaly detection performance. The AUC is a metric for one class detection problem which has a value range of [0, 1]. The higher the detection performance, the closer to 1. Specifically, AUC are calculated by the area under the ROC curve, which shows the relationship between the false positive rate and the true positive rate as shown in Figure 5. It approaches 1 when there are few false positives and few false negative, approaches 0 when the model distinguish anomaly and normality oppositely, and is 0.5 when the model cannot distinguish between anomaly and normality (i.e., a chance rate).

Models: In this paper, we report results for the following models with different conditions for domain adaptation. (S) The anomaly detection model is trained only on the source domain. (S→T) Using a model trained in source domain as initial parameter and we conduct additional training in the target domain only. (S→S+T) After pre-training as in the model (S→T), we conduct additional training with data from both the source and target domains.

All models have the same model structure. In specifically to convert P2, P3, and P4 feature maps from ResNet50-FPN (Lin, Dollár, Girshick, He, Hariharan & Belongie, 2017) into score maps with convolution operation with 3×3 kernel, 1×1 kernel, and sigmoid function.

We use ImageNet pre-trained model to initialize parameter of (S). The input image size is also the same, resized to a fixed size of 512×512.

4.2. RESULTS

The results of the evaluation of detection performance in the target domain are shown in Table 2. Comparing the case where the model learned in the source domain was directly applied to the target domain with different domains (S) with the case where the model was domain adapted with a small amount of target domains (S→T), we can see that the improvement is from 61.25% to 80.50%. In addition, comparing domain adaptation with only the target domain (S→T) and domain adaptation with a mixture of source and target domains (S → S+T), we can confirm that the improvement is from 80.50% to 84.22%.

From above results we confirmed that: 1) the proposed method can detect anomalies; 2) collecting a small amount of training data in the target domain (only about 10% of the source domain) can improve the performance in the target domain and 3) continuous use of data from the source domain during domain adaptation can improve the performance.

The ROC curves corresponding to each condition in Table 2 are also shown in the Figure 5. We can see the same improvement as AUC shown in the table 2.

In Figure 6, we show the visualizing results of the anomaly scores estimated by the (S) and (S→S+T) models. From the results, we can see how domain adaptation improves false negatives and false positives. We also see that the anomaly score responds locally around the anomaly, even though the location of the anomaly, such as the bounding box, is not explicitly used for model training.

Table 2 Comparison of AUC

	use Source domain	use Target Domain	AUC
(S)	✓		61.25
(S→T)	(✓ Only for pretraining)	✓	80.50
(S→S+T)	✓	✓	84.22

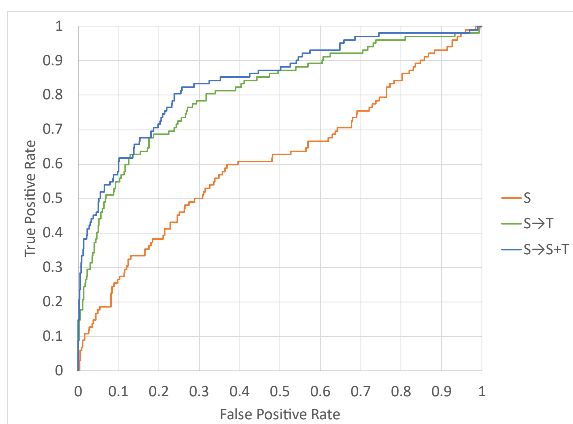


Figure 5 ROC Curve

5. CONCLUSION

With the aim of replacing visual inspections among daily operations, we report result of applying the anomaly detection model to actual in-vehicle camera images. From the results, we confirmed followings: 1) Using weakly supervised learning, we can train anomaly detection model which estimates pixel-wise anomaly scores from low collection cost training data without anomaly location labels; 2) Domain adaptation can improve anomaly detection performance if a small amount of training data in the target domain can be prepared.

REFERENCES

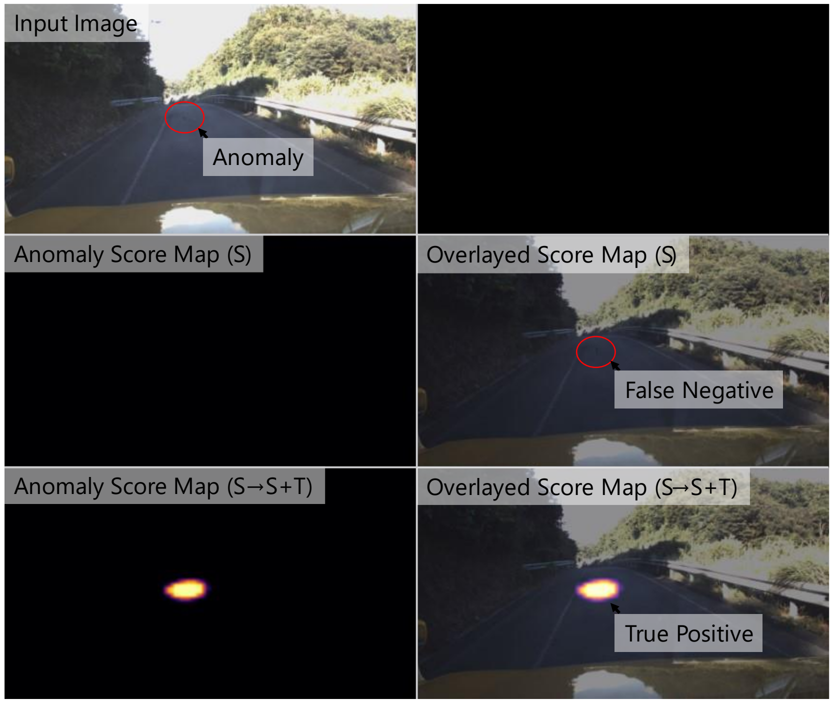
Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., Omata, H. (2018), Road Damage Detection Using Deep Neural Networks with Images Captured Through a Smartphone. Computer Aided Civil and Infrastructure Engineering, Vol. 33, Issue 12, pp 1127-1141, 2018. https://doi.org/10.1111/mice.12387

Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., & Sekimoto, Y. (2021), RDD2020: An annotated image dataset for automatic road damage detection using deep learning, Data in brief, Vol. 36, pp. 107133, 2021. doi:10.1016/j.dib.2021.107133

Lin, T., Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie S. (2017), Feature Pyramid Networks for Object Detection, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, https://doi.org/10.48550/arXiv.1612.03144

Ito, S. (2022), Multi-class road damage detection using multiple instance learning [Translated from Japanese.], *Vision Engineering Workshop 2022*

Yuta Shirakawa: He graduated from Graduate School of Information Sciences, Tohoku University in 2015 and joined to Research and Development Center, Toshiba. He worked on image recognition. He is a member of the Information Processing Society of Japan.



(a) Example of successful detection of anomaly in front of the vehicle.

(b) Example of false detection suppression to the repair trace on the left side.

Figure 6 Example of improvement by domain adaptation