Development of Technology for Detecting Damage by Earthquake from CCTV Camera Images Using AI

(Study period: FY2019)

OTE Masayuki, Senior Researcher,
SUGAWARA Kenji, Research Coordinator for Advanced Information Technology,
IKEDA Yuji (Ph. D.), Head, HOSOKAWA Takehiko, Guest Research Engineer,
Research Center for Infrastructure Management, Information Platform Division

key words: AI, CCTV camera, damage grasping, image processing

1. Introduction

The Ministry of Land, Infrastructure, Transport and Tourism ("MLIT") has installed more than 20,000 CCTV (Closed Circuit Television) cameras across the country in order to administer roads and rivers and has been also using them to collect information on the state of damage after occurrence of an earthquake.
In using those cameras, there is an issue, particularly in the event of a major earthquake, that it is a heavy burden for the personnel who address disasters to watch images of a lot of CCTV cameras set in wide area and quickly check whether any damage is on the screen.
In order to solve this issue, NILIM developed a technology ("difference detection") of detecting "variation" that may be damage as a difference from the images before and after earthquakes from FY2014 to FY2018.
Meanwhile, it is found from the study up to now that difference detection has an issue of falsely detecting rain, snow, etc. on the screen as "variation" that may be damage depending on season or weather (Fig. 1). Then, we studied in the current fiscal year a technology of reducing false detection of difference using AI (Artificial Intelligence).

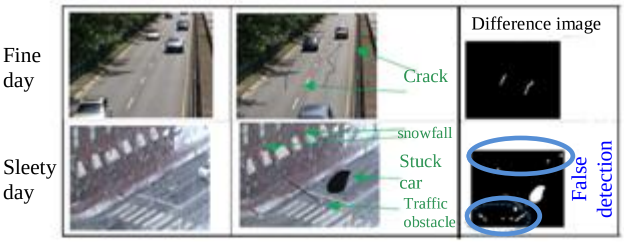

Fig. 1 Example of false detection caused by bad weather

2. Selection of AI algorithm

We studied about a method of conducting image processing with AI before conducting difference detection to reduce false detection in the subsequent difference detection. Since it is difficult to prepare a lot of damage images as teacher data, two algorithms were compared and examined in this study as AI algorithm; "Semantic Segmentation" and "Generative Adversarial Networks ("GAN"), both of which do not require damage images for teacher data.

"Semantic Segmentation" is used to create annotation images to detect "Which section does the input image indicate" using a model in which data created by adding information of such sections as "road" or "river" (hereinafter, "annotation") to images showing a normal state is machine-learned as teacher data.
Fig. 2 shows examples of image in a normal state and annotation image used for teacher data.

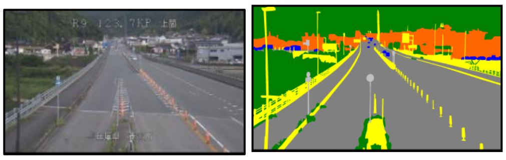

Fig. 2 CCTV camera image in a normal state (left) and annotation image (right)

"GAN" is an algorithm that converts an image similar to a horse, for example, into an image showing characteristics of horses when an image of a zebra is input to the model in which only horse images are machine-learned (Fig. 3). [1)] With application of this algorithm, it may be possible to create an image showing as if the state before suffering damage, where only the state of damage disappeared, leaving the season and weather in a normal state, when damage images are input to a model in which images of various seasons and weathers in a normal state are machine-learned. We attempted to detect only the state of damage by detecting differences between images thus created and input images.

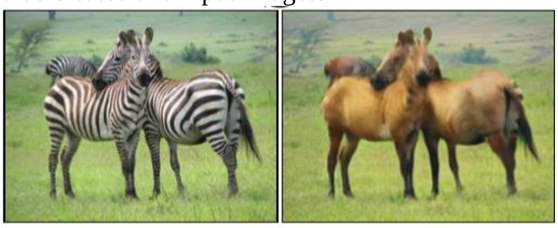

Fig. 3 Example of image generation by GAN Input image (left), Output image (right)

For Semantic Segmentation and GAN, we used a small volume of teacher data (110 images for Semantic Segmentation and 790 images for GAN). Semantic Segmentation showed a tendency of providing stable annotation images, while images created by GAN were unstable unless a model is created for each CCTV camera. Therefore, we selected Semantic Segmentation for this study and created models and evaluated accuracy.

3. Model building and accuracy evaluation results for Semantic Segmentation

We prepared 500 annotation images to be used as teacher data from still pictures of CCTV cameras for monitoring roads, rivers and wide-area. Of the 500 images, 350 images were used for learning, 75 images, for verification of learning, and 75 images, for testing. For the model that finished learning [3] with an existing data set [2], we built a model (the "Model") that finished deep learning using 350 annotation images for learning and evaluated accuracy of the Model. Fig. 4 shows examples of annotation images created by the Model and Fig. 5 shows the results of accuracy evaluation. Note that an evaluation index called "IoU" for object detection in image recognition was used for accuracy evaluation. IoU is expressed with the following formula (1).

$$\mathrm{IoU} = \frac{Intersection}{Union} \quad \text{-- (1)}$$

Intersection: A section where the correct answer image is overlapped with the generated image in the same area.
Union: A section occupied with an area of either the correct answer image or the generated image.

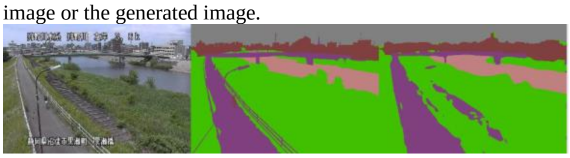

Fig. 4 CCTV camera image in a normal state (left), correct answer annotation image (middle), generated annotation image (right)

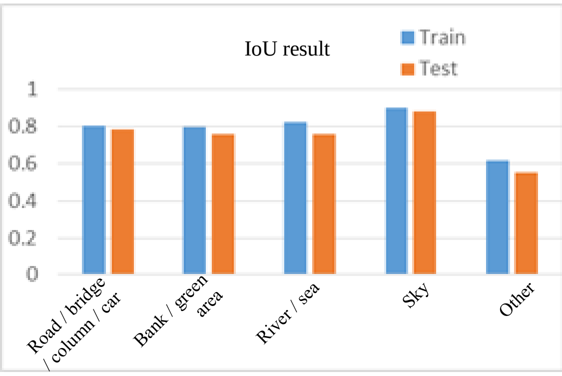

Fig. 5 Results of evaluation by IoU of the built model

Train: Evaluation result of the data for model building
Test: Evaluation result of the data for testing

Various IoUs released for Semantic Segmentation are between 0.8 and 0.9 at maximum. [4] Therefore, in a situation where the volume of learning data is limited to 350 images, results of the accuracy verification for the Model are considered generally reasonable.

4. Technology for reducing false detections using the Model

With the technology for reducing false detections using the Model, it is possible to identify the section where the detected vulnerable area is located because of the automatic generation of annotation images that recognize image sections as preliminary processing of difference detection. By examining the characteristics of damage images in the section, detection of damage points with less false detection is expected.

5. Conclusion

We are going to improve the technology for reducing false detection by verifying its effect using images of snowfall and rainfall and implement measures for supporting disaster response.

☞**See the following for details.**
1)https://arxiv.org/pdf/1703.10593.pdf
2)http://host.robots.ox.ac.uk/pascal/VOC/
3)https://towardsdatascience.com/deeplabv3-c5c749322ffa
4)https://paperswithcode.com/sota/semantic-segmentation-on-pascal-voc-2012