Road Damage Detection and Classification in Smartphone Captured Images Using Mask R-CNN

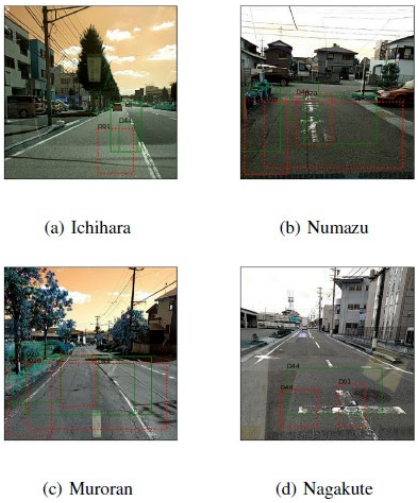
The experiment in this paper was done as a part of the Road Damage Detection and Classification Challenge, 2018 IEEE International Conference on Big Data Cup. The authors used instance detection based on convolutional neural network and classification methods to solve the problem, which is Mask R-CNN. Not only was Mask R-CNN fast, but also it gave them significant results. They used an NVIDIA GeForce 1080Ti and achieved a mean F1 score of 0.528 at an IoU of 50% to detect and classify different types of road damage, where the images were captured using smartphones.

The dataset used here contained 9053 annotated images of roads. The class label of the predicted bounding box was the same as the ground truth bounding box. IoU threshold was 50%. The evaluation was done by measuring the F1 score, where

$$F1 = 2 \times \frac{precision \times recall}{precision + recall}, \text{ here } precision = \frac{True\ Positive}{True\ Positive + False\ Positive},$$

$$\text{And } Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

The implementation of Mask R-CNN was done using the Feature Pyramid Network and ResNet101 as the backbone of the network, and weights from training on the MS-COCO dataset were used here. Images were resized to 512×512 pixels, and augmentation was performed by horizontally flipping the images.



(a) Ichihara (b) Numazu

(c) Muroran (d) Nagakute

Figure 1. Detection Results from trained Mask R-CNN Model

The problem of the paper was that bounding boxes detected the damaged regions, but no exact shape of the regions was identified. In other words, they didn't approach the instance segmentation method, which can precisely show the damaged regions. Bounding box annotation is faster in YOLO over Mask R-CNN.

References:

[1] Singh, J., & Shekhar, S. (2018). Road Damage Detection And Classification In Smartphone Captured Images Using Mask R-CNN. ArXiv, abs/1811.04535.