Road damage detection using morphological operations and RetinaNet model

Nilesh Maharjan, Nikesh Devkota, and Byung Wook Kim*
Changwon National University
20227087@gs.cwnu.ac.kr, *bwkim@changwon.ac.kr

Abstract

The recent technological developments in the field of image processing and computer vision have made it possible to easily monitor urban roads using various cameras, such as surveillance cameras, in-vehicle cameras, or smartphones. Moreover, road maintenance often requires significant amounts of manual labor, making it difficult to detect damage accurately. In this research project, we use RetinaNet model to analyze road images and effectively identify road damage using image contrast enhancement techniques, followed by several grayscale morphological operations during preprocessing stage of the image to recognize effective model with high accuracy using RDD2020 (Road Damage Detection 2020) dataset.

Ⅰ. Introduction

Road infrastructure is vital to economic growth and expansion and to the movement of goods and people. However, road surfaces deteriorate with time due to their location, age, amount of traffic, weather, engineering choices, and building materials. For this reason, several studies have been conducted to evaluate the impact of bad road conditions on economics and safety [1].

Researchers have used machine learning and deep learning-based algorithms to achieve outstanding results in road surface inspection, including pavement distress identification, crack detection, and pothole recognition. For instance, Zhang et al [2]. trained deep convolutional neural networks on manually annotated image patches taken with a smartphone for automatic pavement crack detection. Additionally, Anand et al. [3] demonstrated how to build a deep neural network structure using a GPU card and associated camera to recognize potholes and road cracks.

The purpose of the paper is to solve the issue of perceptibility of road cracks as objects in pavement distress while maintaining time performance and enhancing road damage detection performance. The images are preprocessed in numerous stages before being trained and tested in the RetinaNet model with pretrained backbone ResNet50. The images are scaled, followed by histogram equalization procedure, conversion to grayscale images, and utilized with various morphological processes to clearly visualize the pavement fractures and to further improve the detection performance of the neural network.

Ⅱ. Data preprocessing stage using histogram equalization and morphological operations

In this paper, we present a dataset from the Road Damage Detection Challenge. The dataset's annotations are available in PASCAL VOC format. A rectangular box is used to denote the cracked area, and each box has a crack class number attached. These crack classes include D00 (Longitudinal), D10 (Transverse), D20 (Alligator), and D40 (Discontinuous) (Potholes). Additionally, using stratify as the country of origin for each type of damage, we separated the training dataset into training (70%) and validation sets (30%).

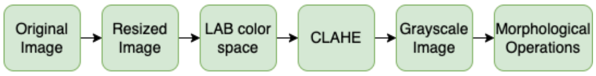

Fig. 1 Preprocessing stage

Fig. 2 Road damage image transformation during preprocessing stage

The preprocessing phase of the pavement crack detection model is shown in Figure 1. Original photos with different resolutions are reduced to 600 x 600 pixels before switching the color space from RGB (Red, Green, Blue) to LAB (L-channel, a-channel, b-channel). LAB is a color space that aims to increase saturation without changing brightness or hue in an image. The image is converted to LAB color space after which it is subjected to contrast-limited adaptive histogram equalization (CLAHE). In this process, the cumulative distribution function is computed after histogram clipping at a predetermined value; limits on contrast can be inferred from this process. Before morphological processes are applied to an acquired image, it is first transformed into a grayscale image. In grayscale morphology, the pixels' neighboring pixels are compared to that pixel to maintain the pixels with the largest (in the event of erosion) or smallest values (in the case of dilation). The pixel intensities of the road images are filtered or smoothed using these grayscale morphological functions to filter noise and fix uneven backgrounds. Moreover, these operations assist in altering bright areas at the expense of dark areas and vice versa. Fig. 2 shows the transformation of images during the preprocessing stage. In our experiment, two type of morphological operations are considered. A gray-level erosion function which reduces the brightness of pixels that are surrounded by neighbors with a lower intensity, whose neighborhood is defined by a structuring element. A gray- level closing function is utilized to eliminate dark patches that are isolated in bright areas and to smooth borders around the pavement cracks in the image.

Furthermore, in this study, a RetinaNet with a feature pyramid is employed to handle the trade-offs between pavement distress detection accuracy, real-time performance, and scalability. After evaluating our spread of bounding box sizes and ratios, we decide on the aspect ratios to these values [0.25, 0.535, 1.0, 1.868, 4.0] instead of the default [0.5, 1, 2] ratios.

Ⅲ. Results

In this section we describe the environment and assessment results of the experiment. Two Nvidia GeForce RTX 3080 GPUs were used in the experiment together with Tensorflow 2.6.0 and Keras 2.6.0, 7487 steps per epoch, and a batch size of 2 to train the model. As noted, we trained RetinaNet using pretrained backbone ResNet 50 model with and without performing morphological operations on the image data. Table I displays selected and further-trained RetinaNet models with various morphological operation conducted during the preprocessing stage. The best mAP (Mean Average Precision), and average reference time per image is shown in Table II. Furthermore, Fig 3 shows some successfully detected examples of pavement cracks.

Table 1: Mean Average Precision RetinaNet Models

Backbones with morphological operation	mAP	Avg Inference Time(s)
ResNet50 (None)	0.387825	0.156
ResNet50 (Erosion operation)	0.396325	0.1527
ResNet50 (Closing operation)	0.388	0.151

Table 2: Average Precision of four different classes

Class Name	Average Precision (Original Image)	Average Precision (Images after Closing)
D00	0.3798	0.3804
D10	0.2135	0.2157
D20	0.5850	0.6090
D40	0.3863	0.3802

Fig. 3 Successfully detected examples

IV. Conclusions

In this research, we provide a strategy based on contrast limited adaptive histogram equalization and grayscale morphological operations during preprocessing stage of the image to improve the performance of the model. The real-world dataset obtained from the smartphones during training of the models is preprocessed and tested on the RetinaNet model with ResNet50 backbone. Experiment results demonstrate the proposed approach in which morphological operations are applied performs better than the original image itself.

ACKNOWLEDGMENT

This research was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (NRF-2022R1A2B5B01001543).

REFERENCES

[1] D. Arya, H. Maeda, S. K. Ghosh, D. Toshniwal, and Y. Sekimoto, "RDD2020: An annotated image dataset for automatic road damage detection using deep learning," Data in brief, vol. 36, pp. 107133, 2021.

[2] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in 2016 IEEE International Conference on Image Processing (ICIP), 2016, pp. 3708-3712.

[3] S. Anand, S. Gupta, V. Darbari and S. Kohli, "Crack-pot: Autonomous Road crack and pothole detection," in 2018 Digital Image Computing: Techniques and Applications (DICTA), 2018, pp. 1-6.